# Multi-Magnification Search in Digital Pathology

by

Maral Rasoolijaberi

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Systems Design Engineering

Waterloo, Ontario, Canada, 2021

## Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

## Statement of Contributions

This thesis is partly based on the following papers that I have authored or co-authored.

[1] Maral Rasoolijaberi, Morteza Babaie, Abtin Riasatian, Sobhan Hemati, Parsa Ashrafi, Ricardo Gonzalez, and H.R. Tizhoosh, **Multi-Magnification Image Search in Digital Pathology**. Manuscript has been submitted for publication at IEEE Journal of Biomedical and Health Informatics. Submission date June 4, 2021.

[2] Abtin Riasatian, Maral Rasoolijaberi, Morteza Babaie, and H.R. Tizhoosh, **A Comparative Study of U-Net Topologies for Background Removal in Histopathology Images**, International Joint Conference on Neural Networks (IJCNN), 2020

[3] Abtin Riasatian, Danial Maleki, Shivam Kalra, Maral Rasoolijaberi, Morteza Babaie, Hamid Tizhoosh and others, **Fine-Tuning and Training of DenseNet for Histopathology Image Representation Using TCGA Diagnostic Slides**, Medical Image Analysis, 2021.

I am the first author of Paper [1] with a major contribution to the design, implementation, experimentation, evaluation, and writing. Paper [1] is incorporated in all chapters of this thesis. I have also contributed to the implementation, experimentation, and writing of Paper [2]. Some parts of paper [2] are incorporated in Chapter 2 and Chapter 4 of this thesis. My contribution to Paper [3] was to a lesser extend.

## Abstract

This research study investigates the effect of magnification on content-based image search in digital pathology archives and proposes to use multi-magnification image representation. Image search in large archives of digital pathology slides provides researchers and medical professionals with an opportunity to match records of current and past patients and learn from evidently diagnosed and treated cases. When working with microscopes, pathologists switch between different magnification levels while examining tissue specimens to find and evaluate various morphological features. Inspired by the conventional pathology workflow, this thesis investigates several magnification levels in digital pathology and their combinations to minimize the gap between AI-enabled image search methods and clinical settings. This thesis suggests two approaches for combining magnification levels and compares their performance. The first approach obtains a single-vector deep feature representation for a WSI, whereas the second approach works with a multi-vector deep feature representation. The proposed content-based searching framework does not rely on any pixel-level annotation and potentially applies to millions of unlabelled (raw) WSIs. This thesis proposes using binary masks generated by U-Net as the primary step of patch preparation to locating tissue regions in a WSI. As a part of this thesis, a multi-magnification dataset of histopathology patches is created by applying the proposed patch preparation method on more than 8,000 WSIs of TCGA repository.

The performance of both MMS methods is evaluated by investigating the top three most similar WSIs to a query WSI found by the search. The search is considered successful if two out of three matched cases have the same malignancy subtype as the query WSI. Experimental search results across tumors of several anatomical sites at different magnification levels, i.e., 20×, 10×, and 5× magnifications and their combinations, are reported in this thesis.

The experiments verify that cell-level information at the highest magnification is essential for searching for diagnostic purposes. In contrast, low-magnification information may improve this assessment depending on the tumor type. Both proposed search methods generally performed more accurately at 20× magnification or the combination of the 20× magnification with 10×, 5×, or both. The multi-magnification searching approach achieved up to 11% increase in F1-score for searching among some tumor types including the urinary tract and brain tumor subtypes compared to the single-magnification image search.

**Acknowledgements**

First and foremost, I would like to express my deepest appreciation to my supervisor, Professor Hamid R. Tizhoosh, for his enthusiasm for this research study, for his support, encouragement and patience throughout my degree. I am glad to be part of his research group, Kimia Lab, where so many researchers are free to explore creative ideas.

I would also like to thank both of my committee members, Professors Parsin Haji-Reza and Apurva Narayan, for donating their precious summertime to review my thesis and provide valuable feedback.

I gratefully acknowledge the assistance of co-authors Abtin, Morteza, and Parsa in writing papers and conducting the research.

Finally, I cannot forget to thank my family for their unconditional love and support, and for always believing in me.

## Dedication

This thesis is dedicated to passengers of flight ***PS752***. On January 8, 2020, this flight was shot down by missiles of the Islamic revolutionary guard. All 176 passengers, including my friends and colleagues, were killed in this tragedy.

# Table of Contents

# List of Figures

# List of Tables

# List of Abbreviations

**AI** Artificial Intelligence 1, 3, 5, 7, 8, 56

**BN** Batch Normalization 9, 10, 24

**CBIR** Content-Based Image Retrieval 2, 8, 9, 12, 14, 36

**CNN** Convolutional Neural Network 6, 9, 10, 12, 14–17, 22–24, 56

**DMMNs** Deep Multi-Magnification Networks 13

**MMS** Multi-Magnification Search iv, x, 4, 5, 16, 21, 22, 35–37, 39, 42, 51, 53, 55–58

**NRQ** Non-Rigid Quantization 19

**ReLU** Rectified Linear Units 9, 10, 24

**SIFT** Scale-Invariant Feature Transform 13, 14

**SNRQ** Sequential Non-Rigid Quantization 19, 56

**TCGA** The Cancer Genome Atlas iv, xii, 13, 15, 29, 34, 54, 57

**TM** Tri-Magnification ($5\times, 10\times, 20\times$ magnifications) 39, 42, 49, 51–54, 57

**WSI** Whole Slide Image iv, x–xii, 1–8, 12–24, 26, 29–31, 35, 36, 38–49, 51, 52, 54–58

# Chapter 1

# Introduction

Pathology is concerned with the microscopic study of tissue specimens and cells to investigate the nature of diseases. The analysis of tissue specimens by pathologists is the "gold standard" in assessing many diseases, including almost all types of cancer. Some practical aspects of pathology include differential diagnoses, biopsy interpretation, tumor grading, and tumor staging [40, 75]. For quite a long time, more than centuries, conventional light microscopy has been utilized as a well-established tool in pathology. With the progress in digital imaging techniques, the digital acquisition of pathology glass slides has gained considerable momentum in the recent decade. Worldwide, pathology laboratories and clinics are slowly beginning to trade in their light microscopes for digital whole-slide scanners, computers, and monitors to proceed towards digital pathology [42,60, 89]. Fig. 1.1 shows a sample WSI which is the digital version of a conventional pathology glass slide.

Digital pathology offers many advantages over conventional pathology, including more efficient workflows, more affordable and more reliable storage solutions for medical records, and a powerful venue for easier collaboration and telepathology. In addition to these advantages, digital pathology have provided an opportunity for the applications of AI and learning-based methods since large databases of WSIs are now available [42, 86]. The impressive performance of AI and deep learning methods in various tasks of digital pathology have provided a powerful argument for proponents of modern pathology to justify the advantages of going digital. AI-driven tools have extended the horizons of the pathologist's view beyond a microscopic glass slide and provided new insights into pathology data [42, 66, 86, 89].

The rest of the introduction chapter is organized as follows. Section 1.1 elaborates

on the motivation of this research study and thesis objectives. Section 1.2 declares thesis contributions. In the end, Section 1.3 provides the organization of the rest of the thesis.

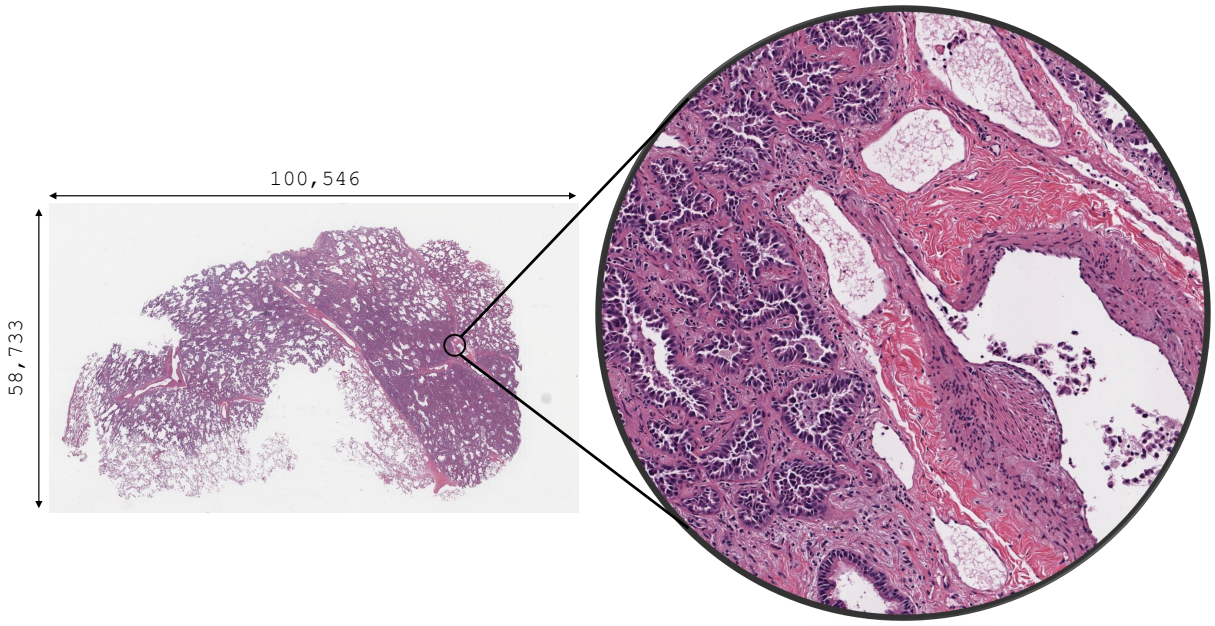



Figure 1.1: Illustration of a gigapixel WSI of lung adenocarcinoma with the size of $100,546 \times 58,733$ pixels.

## 1.1 Motivations

Image search and retrieval in digital pathology can assist pathologists and medical professionals in different diagnostic, research, and educational tasks. Experts users can search through thousands of digital tissue slides through databases, find the cases most similar to a query slide, compare current tissue samples with past patients, and recommend well-informed diagnoses and treatments. A multi-magnification image search engine allows pathologists to benefit from a content-based image search method that actively exploits different magnifications for better accuracy. Previous studies [41, 42, 52, 55, 56, 64] primarily defined content-based image search as a technology where the search input (query) is a digital image, not a textual description. A simple text-based search engine may use some keywords of a disease to find the most relevant cases. The advantage of CBIR is obtaining valuable visual information in large digital biopsy slides. In CBIR, the output

is determined based on the "content" of images, e.g., tissue morphology and cell nuclei distribution.

Pathologists usually use various lenses of a microscope in their detailed inspection of a tissue sample, switching between different magnification as needed. They usually start with low magnifications to identify regions of interest. Then, they look for diagnostically relevant regions to make preliminary diagnoses. Higher magnifications are often needed to confirm or rule out those diagnoses [12, 40, 92]. The tradition of utilizing the highest magnification power for many diagnosis tasks may be attributed to the fact that in several tumors, a high mitotic rate and atypical mitosis indicate malignancy [61].

As a logical expectation, an AI framework should emulate what expert pathologists do. Ultimately, the closer an AI model follows the pathologist's routines, the more realistic and reliable its results may become. Hence, this thesis investigated content-based search encompassing multiple magnification levels to parallel what pathologists do. In other words, since light microscopy in conventional pathology usually comes with interchangeable objective lenses, this thesis aims to develop the feature of adjustable digital magnification power for the proposed search engine to serve as a more powerful and more reliable AI-driven tool.

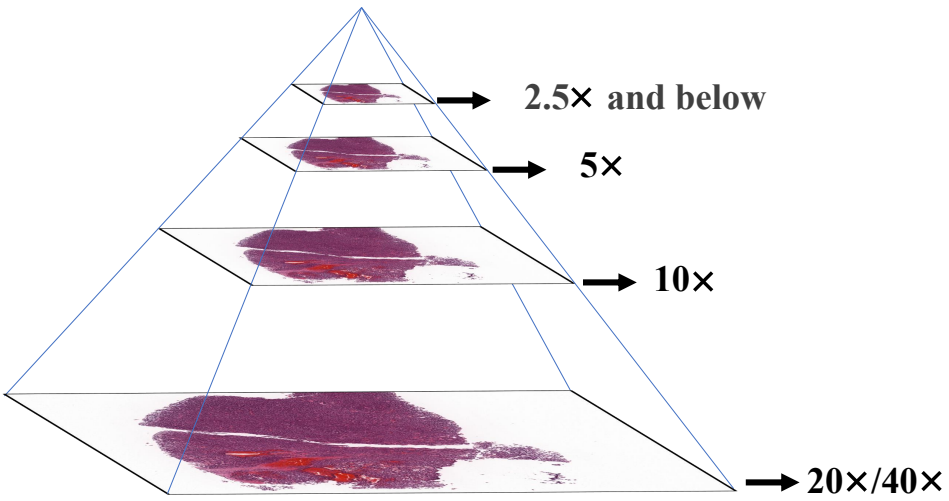



Figure 1.2: Pyramidal structures of a WSI in digital pathology. The base magnification of a WSI is usually 20× or 40×.

The multi-magnification data type of WSIs increased the motivation for this research study. WSIs are multi-resolution images. They usually consist of three to five different

magnification levels since a typical light microscope conventionally has around four objective lenses. However, the number of slide levels can vary depending on the scanning protocol, and device [86]. Fig. 1.2 displays the pyramidal data structure of a sample WSI. As it can be seen, the resolution of slide levels, i.e., the number of pixels representing a specimen, increases with the magnification level.

Even though distinct magnification levels are essential for various assessments, and despite the availability of several magnification levels, current methods in computational pathology mainly evaluate digital slides at a single magnification level, most often at 20×. Several morphological characteristics can be recognized at low magnifications, including the spatial distribution of normal and abnormal tissue components, tumor growth patterns, and heterogeneity. There are several examples of the low-magnification applications; first, identifying reactive epithelial alterations from dysplasia [76], secondly, estimating the extent of histological patterns for classification (e.g., in pure special type breast carcinomas [71]), and lastly, grading (e.g., in prostatic adenocarcinomas [62]). Hence, exploiting several magnification levels is valuable for different pathology tasks.

A histopathology search tool that can perform at various magnification levels is functionally valuable for several reasons. First, selecting the right magnification level for digital slide evaluation depends on the type of histopathology assessment. For instance, microinvasions are detectable at the highest magnification [23], whereas distinguishing some well-differentiated malignant tumors from benign tumors or non-tumoral lesions is challenging at high magnifications [16,57]. Also, each magnification level contains some pieces of histopathology information. At lower magnification levels, the contextual information of glands and tissues is exploited, e.g., architectural patterns, whereas, at higher magnification, cellular contents such as nuclei and cytoplasm are more distinctive.

## 1.2 Contributions

The main contributions of this thesis include:

i. Developing the feature of adjustable digital magnification power for an image search framework in digital pathology with proposing two methods to combine magnification information

ii. Investigating the effect of magnification on whole-slide image search by comparing the performance of two MMS methods at 5×, 10×, 20× magnifications and all of their possible combinations

iii. Creating a multi-magnification dataset of large-sized patches at three magnification levels by proposing and applying a patch preparation approach to a comprehensive and public database of unannotated WSIs

These main contributions aim to provide a multi-magnification image search and retrieval framework for digital pathology that can represent the content of WSIs at multiple magnifications and any combination of magnification levels. This multi-magnification image search tool is empowered with various AI methods and deep learning models to find and retrieve similar WSIs. This multi-magnification image search tool can provide new pathology data insights and assist pathologists in their diagnostic, research, and educational tasks.

## 1.3 Thesis Organization

The remainder of this thesis is structured as follows. Chapter 2 first provides background information and some core concepts used in this thesis. Next, this chapter reviews relevant research papers investigating multi-magnification for histopathology analysis in various tasks in digital pathology. Additionally, previous research studies that investigated the content-based image search in digital pathology are briefly reviewed. The proposed method for MMS is compiled in Chapter 3. Chapter 4 explains the preparation procedure for creating a multi-magnification dataset of patches for the MMS experiments. In Chapter 5, the experiments and evaluation measures to implement the MMS framework and quantitatively assess its performance are described. Chapter 6 first reports results and then analyzes those results and discusses findings from the investigation on the effect of magnification on search. Finally, conclusions and future directions are stated in Chapter 7.

# Chapter 2

# Background Literature

This chapter covers the core concepts used in this research thesis and reviews the relevant literature. At the outset, Section 2.1 provides brief background information on digital pathology and concisely reviews the impact of magnification on histopathology, from objective lenses on a conventional light microscope to magnification levels in a WSI. Section 2.2 briefly reviews content-based image search with an emphasis on image search in digital pathology. In addition, concise background information regarding image representation, search, and retrieval are provided. Section 2.3 overviews deep CNNs and presents background information about network architectures that have been utilized in this thesis for histopathology feature learning and tissue segmentation. In the end of this chapter, Section 2.4 reviews the related literature.

## 2.1 Magnification in Histopathology

At the outset of the clinical pathology routine, a part of the malignant (or suspicious) tissue is collected. After various stages of processing tissue specimens to chemically and physically stabilize them, thin-cut slices of tissue are stained for better visualization, then mounted on a glass slide. Next, the pathologist evaluates the biopsy glass slide at multiple magnification levels [58].

Conventionally, pathologists use a microscope set at various objectives (magnification levels) in their detailed inspection of the biopsy sample, zooming in and zooming out as needed. [58]. Pathologists must use various magnification levels and carefully analyze the glass slides containing tissue specimens to render pathology reports. The pathology

reports are used for many clinical decisions, including surgery decisions for tumor resection (removing an organ or tissue parts), and other treatment plans. Pathologists usually start with low magnifications to identify regions of interest. Then, they look for diagnostically relevant regions to make preliminary diagnoses. Higher magnifications are often needed to confirm or rule out those diagnoses [12, 40, 92].

Fig. 2.1 shows a pathology microscope that is used for the examination of biological tissue specimens. Objective lenses are essential parts of any light microscope. A notable feature of any objective lens, written with the largest font, is magnification. The type of lens design, tube length, and coverslip thickness are other features of an objective lens in optical microscopy [69]. Traditionally, microscopes used for clinical pathology usually have three to five objective powers.

The digitization of biopsy glass slides has recently become more prevalent in pathology practice. Telepathology, creating an efficient digital workflow, and new analytical tools for pathologists are the most evident benefits of moving towards digital pathology. Modern high-resolution scanners can quickly scan an entire glass slide, convert it into a gigapixel image, and produce a WSI in less than minutes [18]. The biopsy glass slide digitization has several other benefits for pathology laboratories and clinics as well. First, these medical centers must keep the records of their patients for at least ten years [58]. The digital version of glass slides, WSIs, take considerably less space for storage. Moreover, WSIs are more robust toward damaged, lost, or fading over time compared to biopsy glass slides. As another advantage, WSIs can be explored by image viewers, but biopsy glass slides should be put under a microscope. A typical conventional microscope has only two oculars and is not designed for multiple users (Fig. 2.1). Therefore, for applications such as education or consulting, using a WSI is easier and more practical than a microscope. Furthermore, despite the large size of scans (a typical WSI file is usually at least several hundred megabytes), recent storage and network sharing progress make it possible to share these files much faster than mailing glass samples for the purpose of consultations and acquiring second opinions [2]. Studies have shown that there is no noticeable difference between diagnoses rendered by pathologists using digitized images and diagnoses rendered using a microscope [8, 58]. Another significant benefit of digital pathology is that AI and computer vision methods can be applied on tissue scans to assist pathologists [66, 68, 89].

A whole-slide scanner digitizes the glass slide at several resolutions, and produces zoomable images in a pyramid data format. Fig. 1.2 shows pyramid data structure of a WSI. The pyramid structure of WSI allows smooth zooming to resemble a conventional microscope. Accordingly, the digital scan of the biopsy glass slide can be explored using image viewers such as a computer monitor in much the same way as standard microscopy [63].

Figure 2.1: Motic Binocular microscope that is commonly used for examination of biological tissue specimen in histopathology. The red arrow points at objectives of the microscope for setting the magnification level.

## 2.2 Content-based Image Search

Content-based image search is a type of search where the search input (query) is a digital image, not a textual description or any keywords. A simple text-based search engine may use some keywords of a disease to find the most relevant cases. The advantage of CBIR is exploiting the invaluable hidden information in the pixel values of images without demanding any external information. In the clinical pathology setting, there is no additional information regarding the new biopsy sample. Therefore, keyword-based searching for finding diagnosis-relevant cases is not an option because there is no diagnosis data concerning the new patient at the outset. However, CBIR can apply to raw pathology images without demanding any external information, and the output of the search is determined based on the content of images. Image search and retrieval has been one of the most popular and valuable AI-driven tools in digital pathology. An image search engine can find and retrieve similar WSIs and their associated metadata, such as annotations and pathology reports, based on the content similarity to the query WSI. Therefore, these search operations can assist pathologists in finding cases similar to the current patient, and they learn from those

already diagnosed and treated patients. CBIR methods rank the database in a decreasing order of similarity to find the best match for the query image [20, 42, 43].

In general, image search and retrieval require salient, discriminative, and representative features that are descriptive of the content of images. In image search, the similarity between image representations, not the raw pixel values, is employed to rank the images. In other words, each image on the database may be represented by a feature vector. Then, the similarity between those feature vectors can be measured to find the most similar matches. Basically, the similarity computation between feature vectors determines the performance of an image search method. Thus, features must be robust, discriminative, and efficient [20, 43].

Image representation by deep neural networks has become popular and commonly used for nearly a decade. Deep neural networks are composed of multiple processing layers to learn abstract-level representations of their input images. Before the popularity of deep networks, various hand-crafted feature descriptors have been investigated to represent the images. However, a clear shift from hand-crafted methods to deep CNN has been observed after the emerge of deep learning [20, 51].

## 2.3 Deep Convolutional Neural Networks

Deep neural networks can learn and discover high-level representations from the raw input data. Non-linear transformations at each layer of a deep neural network are applied to successively learn the abstract representation. In other words, after each layer of a deep neural network, a slightly more abstract level of representation is obtained by applying simple but non-linear transformations. As a result, very complex functions and high-level abstract representations can be learned with the composition of enough such transformations [51].

The most popular deep network architectures in computer vision have been CNNs in recent years. A typical CNN is composed of several convolutions, poolings, BNs, activation functions such as ReLU, and other nonlinear transformation layers. Deep CNNs are commonly trained by using the backpropagation learning algorithm. During the training phase of a deep CNN, the learnable parameters of its layers are changed to optimize a loss function. In this way, a CNN learns relevant patterns in thousands (even millions) of images.

### 2.3.1 DenseNet

DenseNet was proposed by Facebook AI Research in 2017 and its architecture was a breakthrough in CNNs [38]. DenseNet introduced dense connectivity between layers, intending to maximize information flow and reduce feature redundancy in deep CNNs. Therefore, the network could get deeper, and the content of images could be perceived more accurately and efficiently. In total, DenseNet-121 contains around 7 million learnable parameters.

Fig. 2.2 shows a single image that is passed through the DenseNet layers to obtain an abstract feature representation. DenseNet is composed of several (generally four) dense blocks. In each dense block, all layers are densely connected. Accordingly, there are $\frac{L(L+1)}{2}$ direct connections in the DenseNet architecture, where $L$ is the number of layers. Each layer of DenseNet applies transformations such as convolution, BN, ReLU, and pooling.

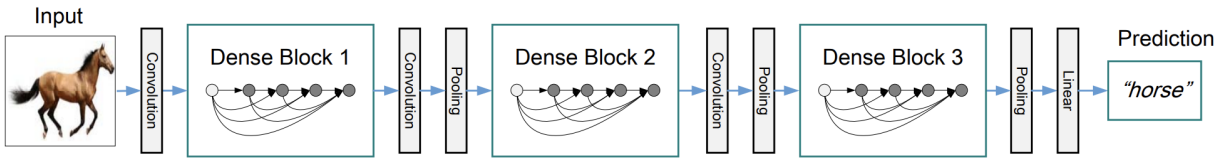



Figure 2.2: A deep DenseNet with three dense blocks. Image Source: [38]

DenseNet has been used as a feature extractor in various digital pathology tasks [34, 42, 43, 72]. For instance, authors in [72] fine-tuned and re-trained the DenseNet-121 architecture, initialized with ImageNet pre-trained weights [48], with histopathology patches.

### 2.3.2 U-Net

U-Net [74] is a convolutional neural network which firstly was proposed for the segmentation of neural structures in electron microscopic images in 2015. Since then, this network has shown impressive performance in several semantic segmentation tasks [15, 19, 65, 101].

Fig. 2.3 shows the U-Net architecture. The U-shape architecture of U-Net has two parts, called encoder and decoder. The first sub-network, known as the encoder, is very similar to a typical CNN and extracts high-level features to capture the image content. The decoder sub-network, also known as the expansion part, creates the desired segmentation map. The encoder and decoder parts are perfectly symmetric in the U-Net architecture. Also, there are shortcut connections between encoder and decoder layers in the network architecture of U-Net.

As the input image passes through the first sub-network, higher-level features are extracted. In the next sub-network, deep feature maps are combined with low-level feature maps from the encoder sub-network. The spatial resolution of feature maps are increased in the second sub-network to achieve an output mask with the same size as the input image. The connections between the encoder and decoder in U-Net architecture facilitate information propagation. In terms of connections in the U-Net architecture, feature maps from the encoder part are cropped and concatenated to feature maps in the decoder sub-network to retrieve local information. These connections enable the network to learn from a few samples.

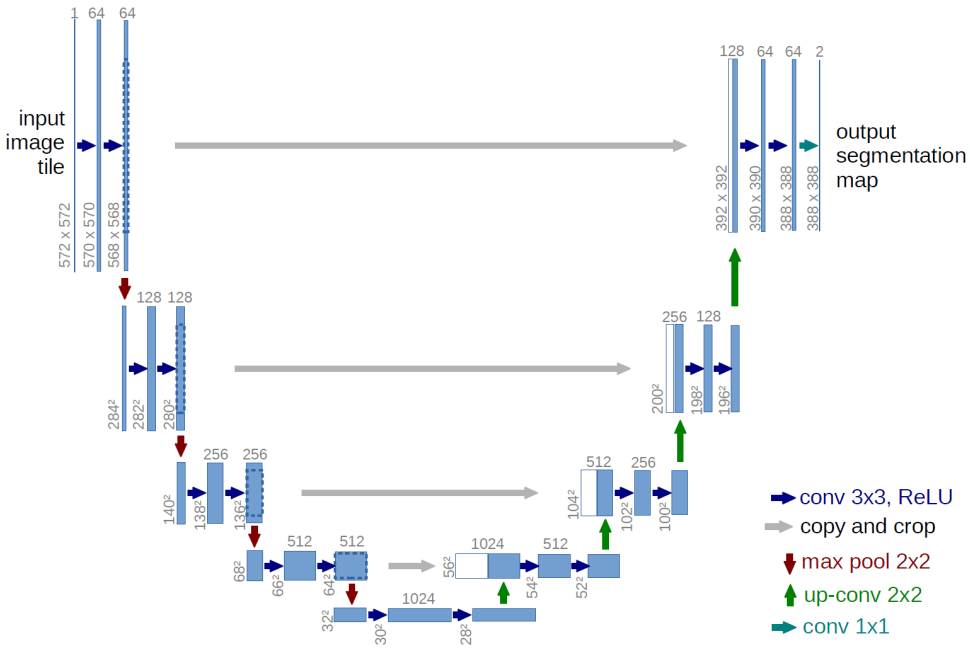



Figure 2.3: U-Net architecture [74]

## 2.4 Related Works

This section presents an overview of the literature related to multi-magnification or content-based image search in digital pathology. Section 2.4.1 reviews previous research studies in digital pathology that utilized multi-magnification in their investigations or investigated in various levels of magnification. Multi-magnification research studies showed promising

results and achievements in terms of accuracy and performance. In contrast to most of the multi-magnification research studies that utilized annotation provided by pathologists, this thesis did not use any manual expert annotations for neither training the proposed model nor testing. The last section of this chapter, Section 2.4.2, reviews previous research studies that targeted the task of content-based image search and CBIR in digital pathology.

### 2.4.1 Multi-Magnification Investigations

Early studies on utilizing multiple magnification levels for digital pathology, before the emerging of deep learning in the computer vision field, are, among others, based on wavelet approaches [97,98]. After the emergence of deep learning, there is an evident shift towards using learning-based approaches and CNNs for multi-magnification investigations in the recent literature of digital pathology.

Deep-Hipo [47] is among recent multi-magnification studies that combined high and low magnification information to locate cancerous regions. Having two receptive fields, Deep-Hipo takes two concentric patches at high and low magnification and concatenated their features to compute the cancer probability of the central pixel. The network architecture of Deep-Hipo consists of two branches, and the backbone of each branch is CAT-Net [93]. CAT-Net is texture-based CNN and consists of convolutional layers, max-poolings, and inception modules [87], followed by fully connected and softmax layers. In the Deep-Hipo paper, the authors used a private dataset of gastroscopic tumors with 94 WSIs. Pathologists manually annotated their dataset to determine cancer regions versus non-cancer regions, and allocated a binary label to the central pixel of each patch accordingly. Also, pathologists manually categorized their private stomach cancer dataset into four subtype classes, i.e., well-differentiated, moderately-differentiated, poorly-differentiated adenocarcinoma, and poorly cohesive carcinoma. In total, they obtained approximately 160k patches at each magnification level for training and validation. They reported comparative experiments with state-of-the-art CNNs in histopathology and computer vision. Their benchmark CNNs included NNet [80], VGG16 [82], DenseNet [38], Efficient-Net [88], Multi-Resolution Deep learning network (MRD-Net) [4] at $10 \times 20\times$ magnification, MRD-Net [4] at $5 \times 20\times$ magnification, Google Brain's Inception V3 (GB-INCV3) [87] , and also the single-magnification version of Deep-Hipo, which is CAT-Net [93]. Deep-Hipo could achieve higher accuracy results than benchmark CNNs, and outperformed in patch-based computing of the cancer probability. In terms of multi-magnification assessment, $20\times$ and $5\times$ outperformed other magnification configurations, i.e., $20 \times 10\times$ and $5 \times 10\times$ magnifications. As an external assessment, the performance of Deep-Hipo is evaluated on a subset

of the TCGA dataset consisted of Stomach Adenocarcinoma (TCGA-STAD) and Colon Adenocarcinoma (TCGA-COAD).

Much of the recent literature on histopathology image analysis based on multiple magnification levels pays particular attention to semantic segmentation tasks [28, 36, 90, 95]. HookNet is a multi-magnification model for histopathology tissue segmentation [95]. Having collected two manually annotated datasets of breast and lung tissues, Van Rijthoven et al. trained and evaluated their proposed HookNet model. The model architecture of HookNet included multiple branches to combine contextual information with fine details detectable only at the highest resolution. This network processed a pair of concentric patches at high and low magnification levels to create a segmented map of its input WSI. Results indicate that multi-magnification models are generally more accurate than single magnification models.

Similarly, authors in [36] adopted a multi-magnification approach and proposed DMMNs. DMMNs approached the semantic segmentation task in breast tissues via processing sets of patches from 20×, 10×, and 5× magnifications. The architecture of DMMNs is composed of various combinations of multiple encoders, decoders, and concatenation blocks. In the decoder part, feature maps of 5× and 10× are concatenated with 20× feature maps to recover local and contextual information. In other words, DMMNs enriched feature maps of the high magnification level with feature maps of lower magnification levels to acquire better spatial characteristics in the segmented image. Concerning the dataset, a pathologist partially annotated 38 breast WSIs and classified regions into six tissue subtypes to prepare a private dataset for the DMMNs study. The authors conclude that their proposed deep multi-magnification network outperforms single-magnification counterparts.

In another relevant research paper, two multi-magnification networks were presented to segment WSIs using a U-Net-based network architecture [28]. In this method, feature maps of lower resolutions were combined with feature maps of higher resolutions to create accurate binary masks. Their authors used WSIs of the CAMELYON16 dataset to train and test their proposed networks.

Authors in [103] proposed a WSI analysis framework for histopathology image classification and retrieval at multi-magnification levels to generate a probability map of malignancy. A fully annotated dataset with 145 containing epithelial breast tumors was employed in their experiment. As the preprocessing step, they utilized color deconvolution for stain normalization. Next, a superpixel segmentation approach was adopted for segmenting WSIs. They considered four magnifications (20×, 10×, 5×, and 2×) in the feature extraction and retrieval stages to capture contextual information. To find the regions, they applied the SIFT method. Next, they utilized a simple neural network with three fully

connected layers and a max-pooling layer to extract features from SIFT points. Finally, they binarized feature vectors intending to increase the efficiency of retrieval in large-scale databases. Their experiments showed that multi-magnification could improve the performance of histopathology classification of breast cancer images.

Many other papers in digital pathology have also hinted at combining context and detail information and investigation on multiple magnification levels, including cervical cell segmentation with a multi-scale CNN model [84], detecting regions of interests based on a superpixel algorithm [9], segmentation of the glands in the colon-rectal digital slides [25], cell classification using a CNN model with multi-scale input and multi-feature layers [104], urothelial carcinoma classification at multi-magnifications using a pre-trained network [100], and cancer subtype classification with a multi-instance CNN [30].

### 2.4.2 Image Search in Digital Pathology

There is a considerable literature on content-based image search and CBIR in digital pathology [5, 33, 34, 42, 43, 45, 59, 70, 72, 77, 85, 102, 103]. *SMILY* [33], allows a user to select a region of interest to obtain matches. A pre-trained network condenses an input image into a feature vector. The network architecture of SMILY is a deep ranking network that was pre-trained on the 5,000,000 natural images from 18,000 distinct classes. This network learned to extract discriminative features by computing and comparing the embeddings of input images. To evaluate the search performance in finding patches with the same histologic features, SMILY adopted a dataset manually annotated by pathologists. Top-5 scores have been reported for patch-based searches at 40×, 20×, 10×, and 5× magnification levels. In another experiment, pathologists compared the search results of SMILY to random search results. Their database contains about $10^9$ image patches and an embedding for each patch. SMILY is implemented as a web-based tool, and Google used 400 computers with 10 compute threads. Considering the rapid growth in histopathology images, this level of computation cost for future applications is not practical, and the higher level of efficiency is essential.

In another recent paper, [43] introduced a search engine, named *Yottixel*, for real-time whole slide image retrieval in the digital pathology. Yottixel is a combination of two words, first, "yotta", which is the largest decimal unit prefix in the metric system, denoting $10^{1024}$ and secondly "pixel". Thus, the term "Yottixel" alludes to the big-data nature of the WSI search in digital pathology. The authors used an unsupervised color-based clustering method to extract a set of images at 20× magnification from each WSI. They called the set of patches *Mosaic*, where around %5 of the tissue specimen was covered. Then, the Mosaic

was fed to pre-trained deep CNNs to extract deep features. In the next step, the feature vectors were barcoded, i.e., binarized, to create a *Bunch of Barcodes* for fast indexing of WSIs. The barcoding of gigapixel WSIs enables Yottixel to perform millions of searches in real-time. The performance of Yottixel evaluated on a private dataset of 300 WSIs and also a public dataset of 2020 WSIs (a portion of the TCGA archive).

KimiaNet [72] is another recent research study that proposed and reported image representation for search in digital pathology. In their research, the DenseNet topology was re-trained at several configurations. The training data were collected from a publicly available TCGA dataset without any pathologist annotation or manual delineation of regions of interest. A clustering-based approach was adopted to select histopathology images at 20× magnification based on a high-cellularity metric. Then, the type of malignancy associated with the WSI was used as the soft label of all extracted images from that WSI to fine-tune DenseNet. Around 240,000 histopathology images with the size of 1000 × 1000 pixels from more than 7,000 WSIs were selected for training the DenseNet at four different stages. Accordingly, this study proposed and reported the results of four feature extractors, i.e., re-training only the last dense block of the DenseNet-121 (fine-tuning 20 percent of network parameters) and continue to unfreeze the parameters of more dense blocks at different stages. After training the feature extractor and in the test phase, feature vectors were converted to binary codes using the *Min-Max* barcoding algorithm [43] to increase the efficiency of image search. KimiaNet was tested for image search on three public histopathology datasets for multi-organ WSI search. KimiaNet reported two types of image search: 1) image search across the entire dataset with the intention of finding the WSI with a similar tumor type to the query WSI among all WSIs, 2) Image search to find similar type of malignancy in an anatomical site, i.e., search for similar WSI with the similar tumor subtype among all WSIs of the same tumor site in their test dataset.

# Chapter 3

# Multi-Magnification Search (MMS)

This chapter explains the methods employed in the content-based WSI search framework to compare WSIs at multiple magnifications. The framework is based on representations provided by a deep convolutional neural network to find similar WSIs via feature matching.

The deep feature extraction procedure is elaborated in Section 3.1. Next, Section 3.2 and Section 3.3 explain two searching methods based on those deep features. Both searching methods can perform at a single magnification and also any combination of magnifications. The first MMS method, called single-vector, obtains a single multi-magnification feature vector of a WSI without demanding any pixel-level or regional annotation. The primary purpose of the feature vector representation for WSIs is to find the most similar WSIs to a query WSI based on the selected magnification level(s). The second MMS method, called multi-vector, performs on a multi-feature vector basis and compares patch feature vectors to find similar WSIs at each magnification, then takes a majority vote among magnifications.

## 3.1 Deep Feature Extraction

Pre-trained deep CNNs can learn to extract content-based features from their input image. An example is KimiaNet [72] which is a customized feature extractor for the histopathology images. The architecture of KimiaNet includes four dense blocks, with several convolutional and pooling layers and dense connections. KimiaNet has around seven million parameters in total. Model parameters of KimiaNet are adjusted to derive histological characteristics from input images.

### 3.1.1 Magnification-wise Fine-Tuning of KimiaNet

KimiaNet was originally trained at 20× magnification. Therefore, this thesis employs KimiaNet without any modification to extract tissue features from 20× magnification patches. To do so, all high-resolution patches are fed to KimiaNet, one by one, to generate a deep feature vector as a representation of each histopathology patch. Note that the term deep feature here indicates the output of the latest pooling layer in the KimiaNet model architecture.

The last DenseNet-121 block [38], i.e., 20 percent of KimiaNet layers, were re-trained with 10× and 5× magnification patches to adjust this feature extractor to those magnifications. In this fashion, the weights of the KimiaNet have been changed according to different magnification levels, attempting to imitate how a pathologist changes their microscope's objective lenses. As a result, different sets of parameters for the KimiaNet model architecture at three magnification levels were obtained. The model architecture and configurations of feature extractors are the same for all magnification levels. Details of fine-tuning procedure is presented in chapter 5, Section 5.3.1, on page 37.

### 3.1.2 Representing Patches with KimiaNet Deep Features

After training, in the feature extraction phase, the input of KimiaNet is still a histopathology patch. The output is a feature vector with a size of $1024 \times 1$ representing the input patch. As mentioned before, dimension reduction is not the primary purpose of utilizing pre-trained deep CNNs. Feature vectors generated by KimiaNet contain histological characteristics and distinct features of the input patches as the network has been trained with histopathology patches and their corresponding primary diagnostic labels.

Note that most deep learning approaches in digital pathology, including KimiaNet, are applied on patches rather than the entire WSI due to the extremely large image size of WSIs [72,86]. See Chapter 4 for more details related to patching. This thesis proposed two methods for the WSI representation and search using KimiaNet patch features. Section 3.2 and Section 3.3 explain single-vector and multi-vector methods to represent an entire WSI based on its patch features. WSI representations can be compared to find and retrieve the most similar WSIs.

## 3.2 Single-Vector WSI Representation: Median Aggregation

In the context of WSI feature representation, the goal is to obtain a single feature space so that WSIs with similar histological features are close to each other. Up to this point, patches, i.e., square images acquired from tissue regions of a WSI at a specific magnification level, have been fed to KimiaNet to generate deep feature vectors. In other words, each deep feature vector (the output of KimiaNet as the feature extractor) represents a patch, not the entire WSI. The median aggregation approach first creates a feature space for each magnification level. To do so, all patch feature vectors of a WSI are aggregated by taking the median value with respect to feature positions to create a single vector representing of the WSI at a specific magnification.

Next, the aggregated vectors of different magnifications are concatenated to create a multi-magnification vector representation. Fig. 3.1 shows this procedure.

The Euclidean distance between median feature vectors in the dataset can be computed to compare WSIs. In the single magnification searching, each WSI is represented by a single-vector. Therefore, the pair-wise Euclidean distance can be calculated between single-vector features to find matches. In the double-magnification searching, each WSI is still represented by a single vector that is the concatenation of two single-magnification vectors. Therefore, the euclidean distance is calculated between those two concatenated vectors for double-magnification searching in a similar way. If all three magnification levels are considered, the concatenation of all magnification generates a single-vector Tri-magnification representation for each WSI. Again, the pair-wise Euclidean distances are calculated accordingly for searching at multiple magnifications.

## 3.3 Multi-Vector WSI Representation: Median-of-Mins

One may also extract many patches and attempt to represent a WSI with all feature vectors at the same time. The *median-of-mins* approach performs patch-to-patch, patch-to-slide, and slide-to-slide comparisons [43]. The procedure is repeated for all magnification levels in an independent manner.

### 3.3.1 Binarization of Feature Vectors (Feature Barcoding)

Working with many feature vectors when operating on large WSI archives can be prohibitively time-consuming. The calculation of Hamming distance between binary codes is much faster than calculating Euclidean distance between full-precision feature vectors. In large-scale search problems, reducing the demand for memory and computational resources is another advantage of binary feature vectors. Hence, hashing feature vectors would be required in case of multi-vector WSI representation.

Instead of the *Min-Max* barcoding algorithm as proposed before [43], this research study uses the SNRQ, an unsupervised hashing approach, for barcoding patch feature vectors. The NRQ method is an extended version of the iterative quantization (ITQ) method [27]. NRQ applies dimensionality reduction and utilizes non-rigid transformations along with rigid transformations such as rotation to further reduce the quantization error caused by hashing. Employing non-rigid transformations allow a more robust quantization since the degree of freedom in non-rigid transformations is higher. Sequential NRQ (SNRQ) is an efficient implementation of NRQ based on sequential updates.

### 3.3.2 WSI to WSI Distance Calculation

Let assume the test dataset contain $n_p$ patches from $n_s$ WSIs. First, pair-wise Hamming distances between barcoded feature vectors of all patches are calculated, generating a *patch-to-patch-distance* matrix with dimension of $n_p \times n_p$. Note that distances between patches of the same WSI are set to infinity for convenient search purposes. In the next step, the minimum distance from each patch to all patches of each WSI is calculated, generating the *patch-to-slide-distance* matrix of dimension $n_p \times n_s$. Finally, pair-wise distances between WSIs are calculated by finding the median of patch-to-slide distances, generating a *slide-to-slide-distance* matrix of dimension $n_s \times n_s$.

This *median-of-mins* matching process repeated for all magnification levels. Figure 3.2 illustrates the three independent search processes for a query WSI at 5×, 10×, and 20× magnifications. In terms of multi-magnification search, all retrieved WSIs are considered at selected magnification levels for later evaluation.

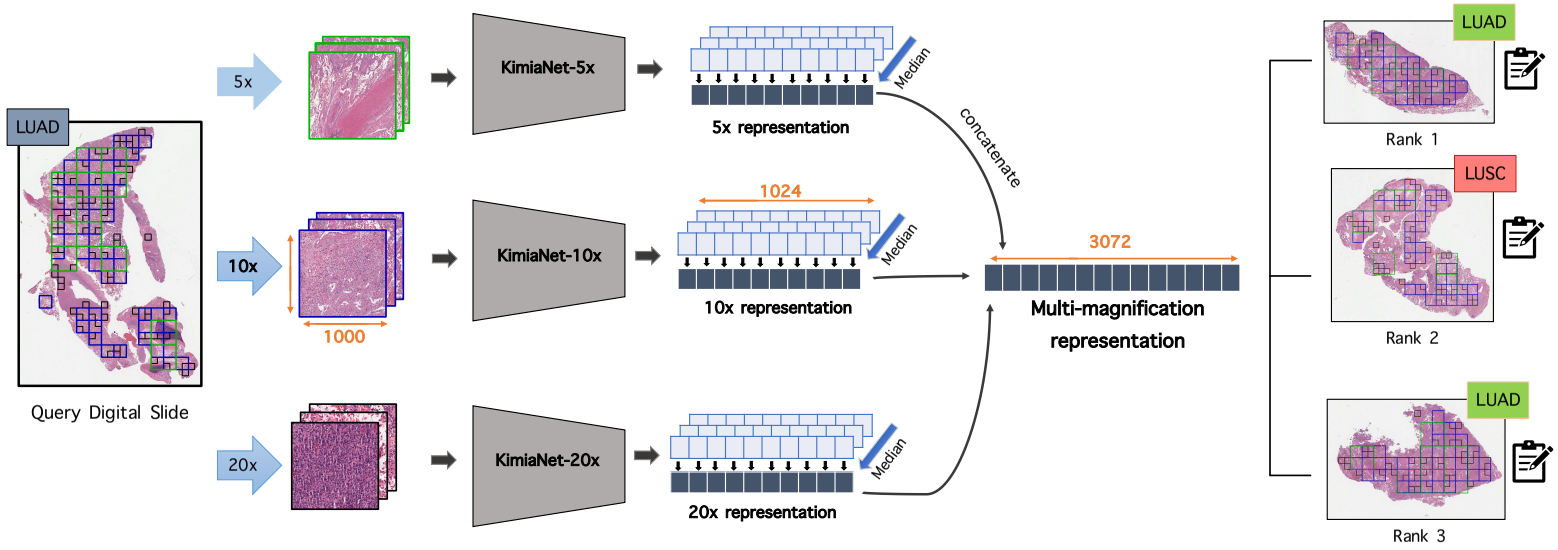



Figure 3.1: Searching framework for the single-vector WSI representation method using the median aggregation approach (Section 3.2). The pen-and-pad icon stands for metadata such as pathology reports containing primary diagnosis.

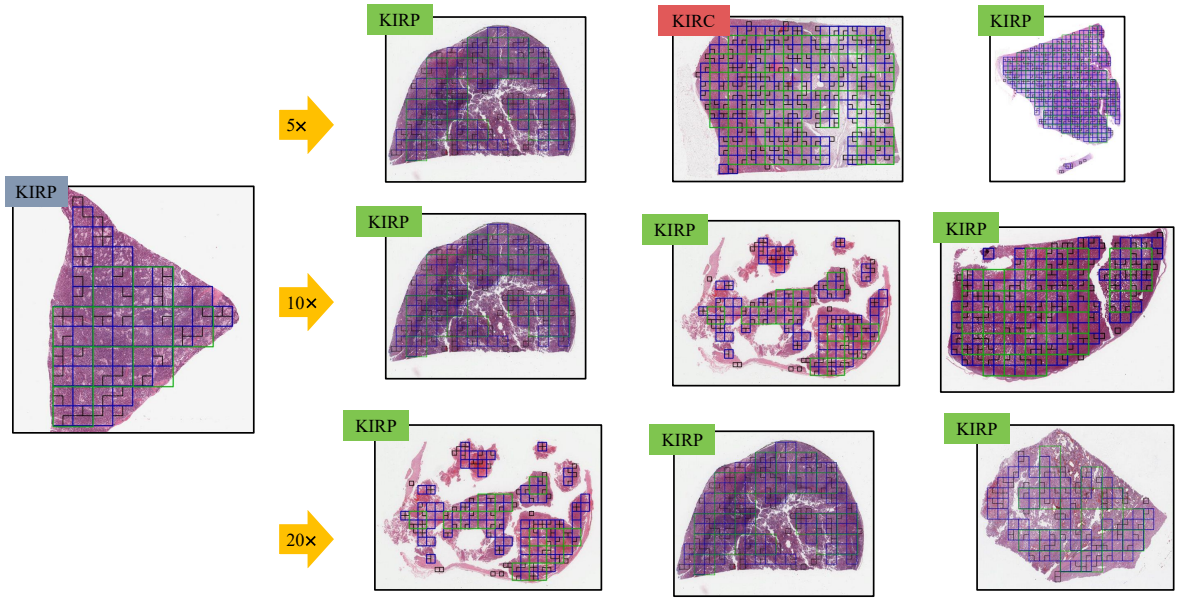



Figure 3.2: Image search among urinary tract WSIs using the multi-vector method. The search has been performed three times based on extracted patches at 5×, 10×, and 20× magnification. The search at a single magnification is successful if at least two of the three search results show correct primary diagnosis. The multi-magnification search (MMS) is successful if the majority of the search results at selected magnifications indicates the same disease type as the query WSI.

# Chapter 4

# Data Preparation, Methods and Experiments

Gigapixel WSIs, with a common image size of 100,000×100,000 pixels for large tissue samples, are too large to be processed by typical CNNs directly. To address this challenge, a patch-based approach is usually adopted [5, 26, 43, 72, 86]. This research study also adopts a patching approach to overcome the computational challenges of utilizing a deep CNN. The term "patch" here refers to a rather small sub-image of a WSI with manageable dimensions.

In the first step of multi-magnification image search (MMS), patches with a manageable image size should be extracted from tissue regions of a WSI at various magnification levels. This chapter first addresses the challenge of locating the tissue specimen regions in the gigapixel WSI for patch extraction. The task of tissue localization can be addressed by hand-operated annotation by pathologists. In many research studies [28, 36, 47, 93, 95, 103], expert pathologists annotated regions of interest in a small private dataset, and patches were extracted from those regions accordingly. It must be pointed out that only expert pathologists with a high level of domain knowledge can annotate WSIs. Therefore, the hand-operated annotation of WSIs is not a feasible task for a large-scale dataset. Even for a small dataset, the delineation of gigapixel WSIs is tedious, time-consuming, and subject to errors and variability. Accordingly, this thesis did not use any pathologist's pixel annotations or manual delineations. The following sections explain the proposed method for creating a multi-magnification dataset of histopathology patches without demanding any pixel-level annotation by pathologists.

## 4.1 Tissue Segmentation

Although for a human operator the discrimination between tissue and background is obvious, the identification of tissue regions in WSIs could be challenging for computers mainly due to the existence of color variations and artifacts. This thesis uses binary masks to identify tissue regions and avoid patch extraction from blank areas or artifacts. In a binary mask, pixels belonging to tissue regions have a value of one, whereas pixels belonging to background regions have zero value.

To overcome the challenge of background removal in histopathology images, the performance of two methods for generating binary masks were evaluated. The First method, otsu thresholding [67], is hand-crafted and is one of the best traditional methods for foreground and background segmentation. The second method for generating a binary mask is utilizing pre-trained U-Net, which can be considered as a deep learning approach.

### 4.1.1 Otsu Thresholding

Otsu binarization method is a well-known algorithms to classify pixels into foreground and background. The otsu algorithm is a robust iterative thresholding method that has been widely used to compute the optimal threshold. This hand-crafted image thresholding method performs the best for images with only two distinct pixel values. Since the image histogram of an image with only two distinct values would only consist of two peaks, a decent threshold would be in the middle of those two values. In this way, the Otsu's algorithm determines an optimal global threshold value, and for every pixel the same threshold value is applied to generate the corresponding binary mask.

The otsu thresholding is not cable of identification of artifacts such as marker traces and extra stain. Identification of those non-tissue areas as tissue regions leads to patch extraction from non-informative background, and can confuse any CNNs. Moreover, some tissue types such as alveolar tissue types, fatty tissues, and tissues with poor staining are more challenging to detect with the otsu thresholding method.

### 4.1.2 U-Net

Experiments on U-Net architecture with different network backbones have been performed to select the best topology for the tissue segmentation task. After comparing a wide range of backbone networks, the U-Net architecture with MobileNet [37] backbone was utilized

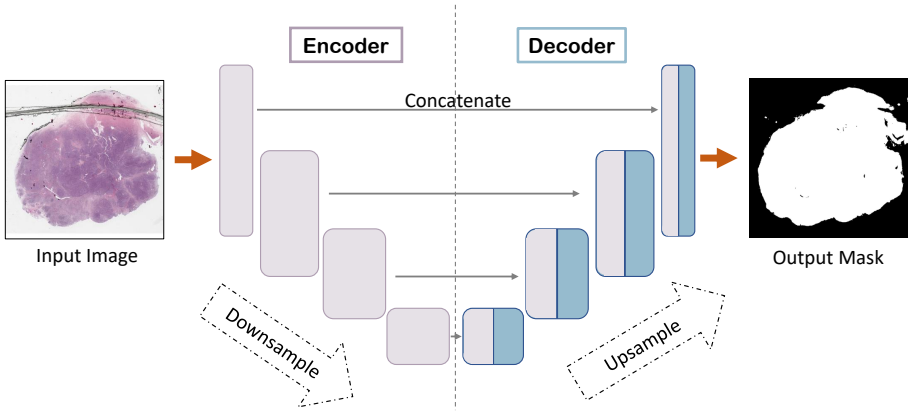
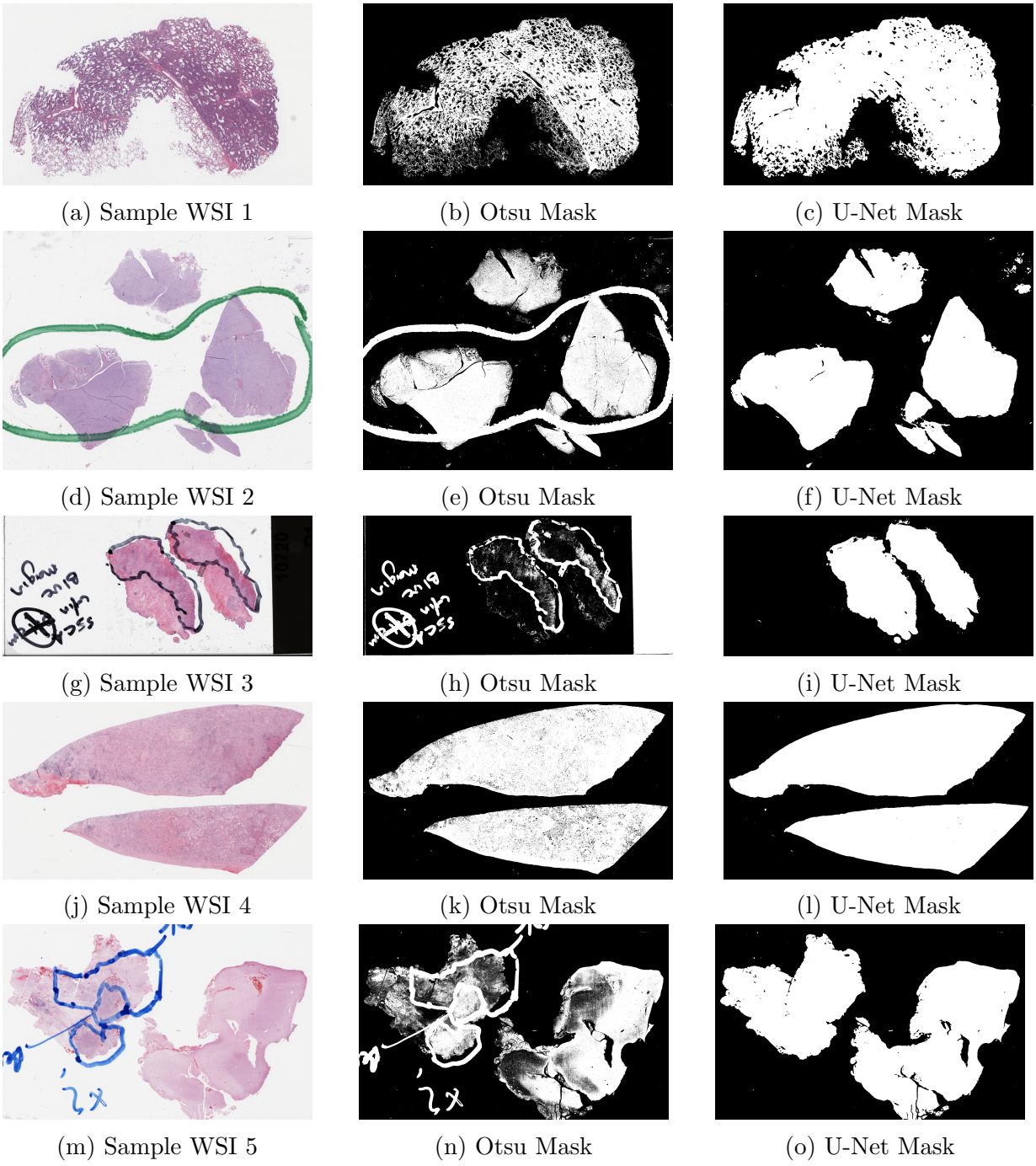



Figure 4.1: U-Net for tissue segmentation. Each block shows a feature map.

for the tissue segmentation. U-Net with MobileNet backbone has been trained to segment input thumbnail WSI, i.e., WSI at one of its low magnifications, into tissue and non-tissue regions. Non-tissue regions include the blank background and artifacts such as bubbles, tissue folds, extra stains, broken glass, debris, and marker traces.

Fig. 4.1 shows the U-Net proposed approach for generating binary masks. The input image is a WSI at low magnification. As the input image passes through the first sub-network, higher-level features are extracted. In the next sub-network, deep feature maps are combined with low-level feature maps from the encoder sub-network. The spatial resolution of feature maps are increased in the second sub-network to achieve an output mask with the same size as the input image. The connections between the encoder and decoder in U-Net architecture facilitate information propagation. In terms of connections in the U-Net architecture, feature maps from the encoder part are cropped and concatenated to feature maps in the decoder sub-network to retrieve local information.

The encoder part of the U-Net is a feature extractor, and it can be a pre-trained CNN, e.g., MobileNet [37] and EfficientNet [88] without the last fully connected layers. The decoder part of the U-Net with MobileNet backbone consists of 5 decoder blocks. The structure of each decoder block is made up of one 2d-upsampling layer and two repetitions of 2d-convolution, BN, and ReLU activation. Four skip connections connect layers from the encoder part, usually the output of ReLU activation at a certain layer of each encoder block, to the last four decoder blocks, after the up-sampling layer. A 2d-convolution layer with Sigmoid activation is the last layer of the network. The Segmentation_Models library [101] with various backbones for U-Net has been used for the experiments on tissue segmentation with U-Net.



(a) Sample WSI 1 (b) Otsu Mask (c) U-Net Mask

(d) Sample WSI 2 (e) Otsu Mask (f) U-Net Mask

(g) Sample WSI 3 (h) Otsu Mask (i) U-Net Mask

(j) Sample WSI 4 (k) Otsu Mask (l) U-Net Mask

(m) Sample WSI 5 (n) Otsu Mask (o) U-Net Mask

Figure 4.2: Comparison between Otsu and U-Net methods for generating binary masks.

## 4.2 Patch Extraction Using a Binary Mask

After generating a binary mask, the binary mask image is divided into grids. If a pixel has the value of one in a binary mask, that pixel belongs to tissue regions. If a pixel has zero value in a binary mask, that pixel does not belong to tissue regions. Therefore, all the pixels in that grid area are summed up to find out whether that grid area, i.e., patch, contains mainly tissue regions or not. Next, qualified patches, i.e., patches containing tissue regions more than a threshold, are selected. In this procedure, the tissue percentage threshold parameter is identical for all WSIs to make the process as automated as possible. The procedure is repeated for different magnification levels to generate a unique dataset of patches at each magnification.

The steps of the patch extraction procedure without demanding on any pathologist's annotation is outlined in algorithm 1. In the literature [43], a small percentage, usually a range between 5 percent to 20 percent of patches of tissue regions, is randomly selected by preserving the spatial diversity at the high magnification due to a large and unmanageable number of patches. The empirical experiments have shown than sampling tissue regions with a ratio around 20 percent is representative enough [43]. The parameter *SamplingRate* in algorithm 1 determines this percentage. This sampling step, i.e., step 15 in Algorithm 1, It only is applied to high-magnification where there is redundancy, and the number of patches is too large and unmanageable. Normal Distribution was used to sample various tissue types all over the input WSI.

Fig. 4.3 show the procedure of path extraction. To make the procedure faster and more efficient, binary masks were generated at the lowest magnification level, usually $1\times$ or $2.5\times$ magnification. Next, the window size for gridding the binary mask was scaled from the target magnification to the mask magnification to find tissue regions in that binary mask. For example, a $1000 \times 1000$ patch at $20\times$ magnification (a target magnification for patching) is equivalent to $125 \times 125$ area at $2.5\times$ magnification (the mask magnification).

**Algorithm 1:** Patch Extraction

**Input :** WSI Path
**Output:** Set of Patches

$m_x \leftarrow$ (Set the Magnification Level)
*PatchSize* $\leftarrow$ (Set the Patch Size)
*Threshold* $\leftarrow$ (Set a Threshold for Binary Mask)
*SamplingRate* $\leftarrow$ (Set the Sampling Ratio)
**procedure**
    wsi $\leftarrow$ Open WSI using OpenSlide Library
    $m_0 \leftarrow$ Objective-Power of the Highest Magnification (20× or 40× )
    $m_{target} \leftarrow$ ( Determine the Slide Level based on $m_0$ and $m_x$)
    thumbnail $\leftarrow$ Get Thumbnail at Lowest Magnification Level
    mask $\leftarrow$ Feed thumbnail to Pre-trained U-Net
    GridSize $\leftarrow$ Set Girding size the based on mask size, *PatchSize*, and $m_{target}$
    Grids $\leftarrow$ Grid the Mask to GridSize
    **Loop:** Calculate the Tissue Percentage for Each Grid ;
    Mask-Coordinates $\leftarrow$ Select Grids with Tissue Percentage above the Threshold
    Mask-Coordinates $\leftarrow$ Sample Mask-Coordinates (Only at high-magnification)
    Patch-Coordinates $\leftarrow$ Scale Mask-Coordinates to $m_x$
    Patches $\leftarrow$ Read the WSI Regions at Patch-Coordinates with size of *PatchSize*
**end procedure**

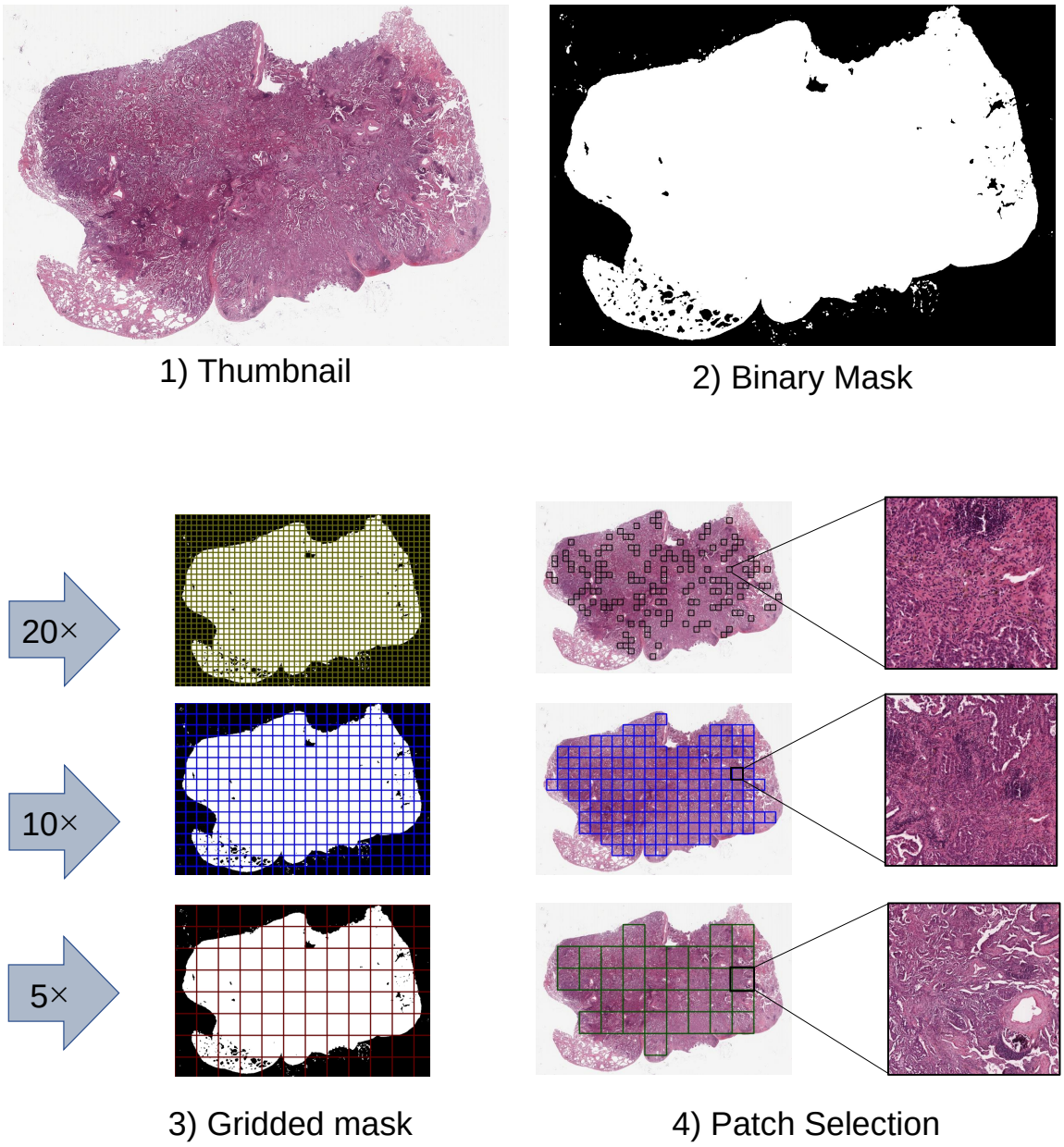



Figure 4.3: The pipeline of pre-processing, patch extraction procedure at 5×, 10×, and 20× magnification levels. A binary mask is utilized for the tissue specimen detection.

## 4.3 Creating a Reusable Multi-Magnification Dataset

As a part of this research study, this thesis aims to create a multi-magnification dataset of histopathology patches for various deep learning-based cancer research purposes rather than a dataset only limited to this study. The gigapixel WSIs are too large to be used for many deep learning approaches. Hence, a multi-magnification dataset of large-size patches is created from extracting tissue regions of more than 8,600 WSIs. This multi-magnification multi-organ datasets consist of 1000 × 1000 patches, with three sub-datasets at 5×, 10×, and 20× magnifications.

### 4.3.1 WSI Repository

The Cancer Genome Atlas, in short TCGA, is an open-access and rich resource of digital pathology data for worldwide cancer research [1]. This repository was used as the source of the multi-magnification dataset. TCGA is one the most comprehensive publicly available dataset of hematoxylin and eosin (H&E)-stained whole slides in digital pathology. In the literature [18], diagnostic slides of the TCGA data repository are categorized to 13 anatomical sites and 33 human tumor types. The tumor types of the TCGA dataset are presented in Table 4.1. Note that there is neither regional nor patch-level pathologist's annotation for WSIs used in this research study. However, WSIs in the TCGA dataset are associated with pathology information such as morphology, primary diagnosis, tissue or organ of origin, patient age at the time of diagnosis, tumor stage, medical center, gender, race, and ethnicity. This information applies to the entire WSI.

### 4.3.2 Data Collection and Preprocessing

At the outset, formalin-fixed paraffin-embedded human pathology samples of the TCGA dataset were collected for this research, i.e., only diagnostic slides were considered. Frozen sections, generally showing somewhat lower image quality, were eliminated to maintain a more consistent evaluation. Many of the remaining slides did not include 20× or higher magnification levels. Therefore, this research study only included permanent section diagnostic slides with 20× or higher levels of magnification. Next, these slides were categorized based on morphology codes, primary sites, and diagnosis, again omitting the groups with lower than 20 slides due to search purposes in the test phase. This grouping allowed having a minimum of two WSIs from each class in the test dataset. Finally, 8,611 permanent H&E stained digital slides were recorded from 13 anatomic sites (categorization

based on [18]). The anatomical site of a digital slide indicates the organ or body system biopsied. Anatomical sites of this dataset consists of Brain, Endocrine, Gastrointestinal tract, Gynecological, Haematopoietic, Liver/Pancreaticobiliary, Melanocytic malignancies, Prostate/Testis, Pulmonary, Urinary tract, Breast, Head and Neck, Soft Tissue, and Thymus. Next, the dataset was split with a ratio of roughly 80, 10, and 10 percent with respect to the anatomical sites into the train, test, and validation sub-datasets, respectively. Therefore, each train, test, and validation dataset includes 7,126, 744, and 741 H&E WSIs, respectively.

Next, concerning each WSI, regions containing tissue specimens were segmented to remove background region. As mentioned in Section 4.1, the term background here refers to non-informative pixels in a WSI, including white-coloured backgrounds of a glass slide and artifacts such as ink marker and extra stain. This thesis utilized a pre-trained U-Net model with the MobileNet backbone to generate binary masks at the lowest magnification level (see Section 4.1.2 for more details on U-Net for tissue segmentation). The binary mask facilitates identifying the background and foreground of a WSI. After recognizing tissue regions (foreground), high-, medium-, and low-resolution patches with a fixed size of $1000 \times 1000$ pixels were extracted. Fig. 4.5 shows a WSI and its extracted patches at three magnification levels. Some samples of patches at various magnification levels are presented in Fig. 4.6. Patches were extracted in an automated way and without any pathologist supervision or delineation of regions of interest. Fig. 4.4 shows three sample WSIs and the location of their extracted patches.

A total of 242,202 histopathology $20\times$ magnification patches were selected from those 7,126 H&E digital slides and used for re-training the DenseNet to find the model parameters for KimiaNet at $20\times$ magnification [72]. Concerning the $10\times$ magnification level, 190,257 patches were collected to fine-tune KimiaNet. The last DenseNet-121 block, i.e., 20 percent of KimiaNet layers, were re-trained with those $10\times$ magnification patches. In a similar way, the previous Dense-block was fine-tuned again with 167,746 patches acquired at $5\times$ magnification. All versions of KimiaNet were trained to classify patches from 30 primary diagnoses.

In the search phase and concerning the test dataset, 20% of tissue regions of 744 WSIs were sampled at $20\times$ magnification, resulting in a collection of 91,287 patches. The patch selection with the sample ratio of 20 percent generates a manageable collection of histopathology patches at $20\times$ magnification. Other studies in the literature followed similar sampling settings with a sample ratio of 5 to 20 percent to collect data from gigapixel WSIs [43]. As the patching algorithm tries to distribute the location of patches over the entire specimen areas, patches are selected from diverse tissue regions. Moreover, this sampling helps to reduce data redundancy in the patch dataset since many patches, especially

patches extracted from some tissue types such as fatty tissues, can be considered the same piece of data. According to empirical evidence, sampling from well-distributed locations at 20× magnification with a sample ratio of 20 percent is sufficiently representative. Since WSI image sizes at 10× and 5× magnifications are smaller than 20× magnification, even the sample ratio of 100% at 10× and 5× magnifications creates a manageable patch collection. Therefore, every patch that mainly contains tissue texture (and not much background) was extracted at 10× and 5× magnifications. The number of overall collated patches in for the test dataset is 97,389 and 20,397 concerning the 10× and 5× magnification, respectively.

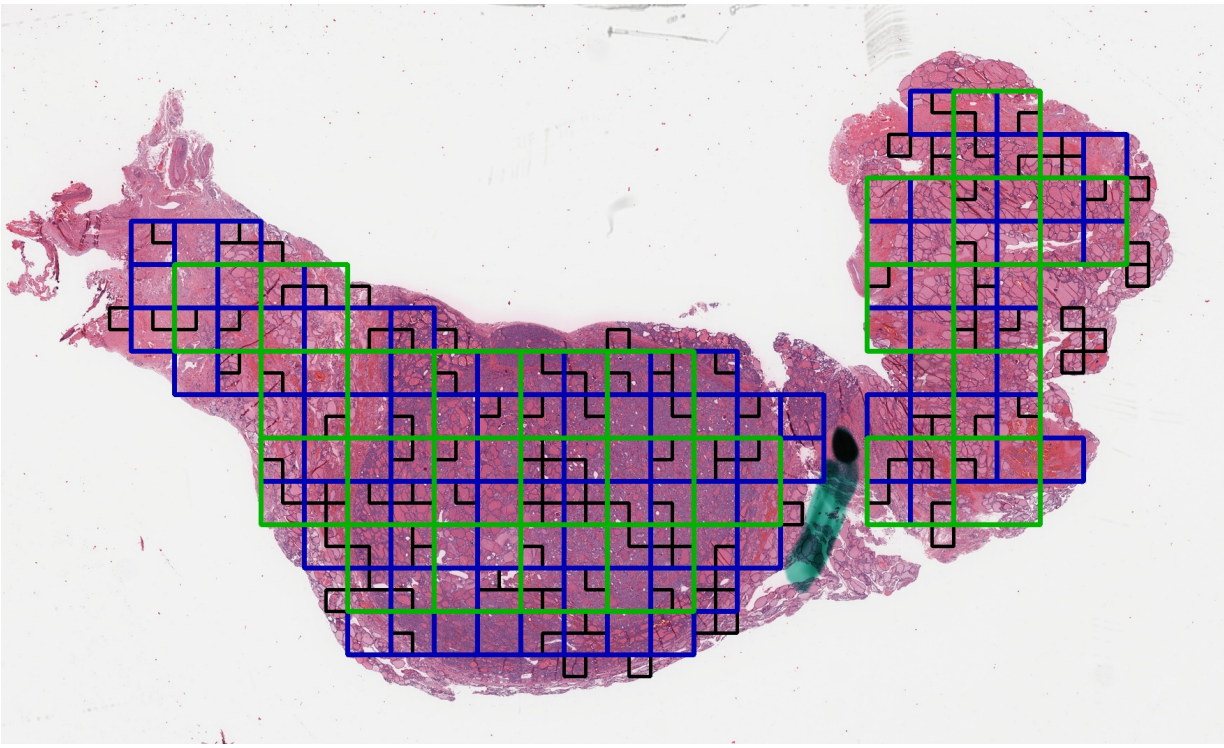
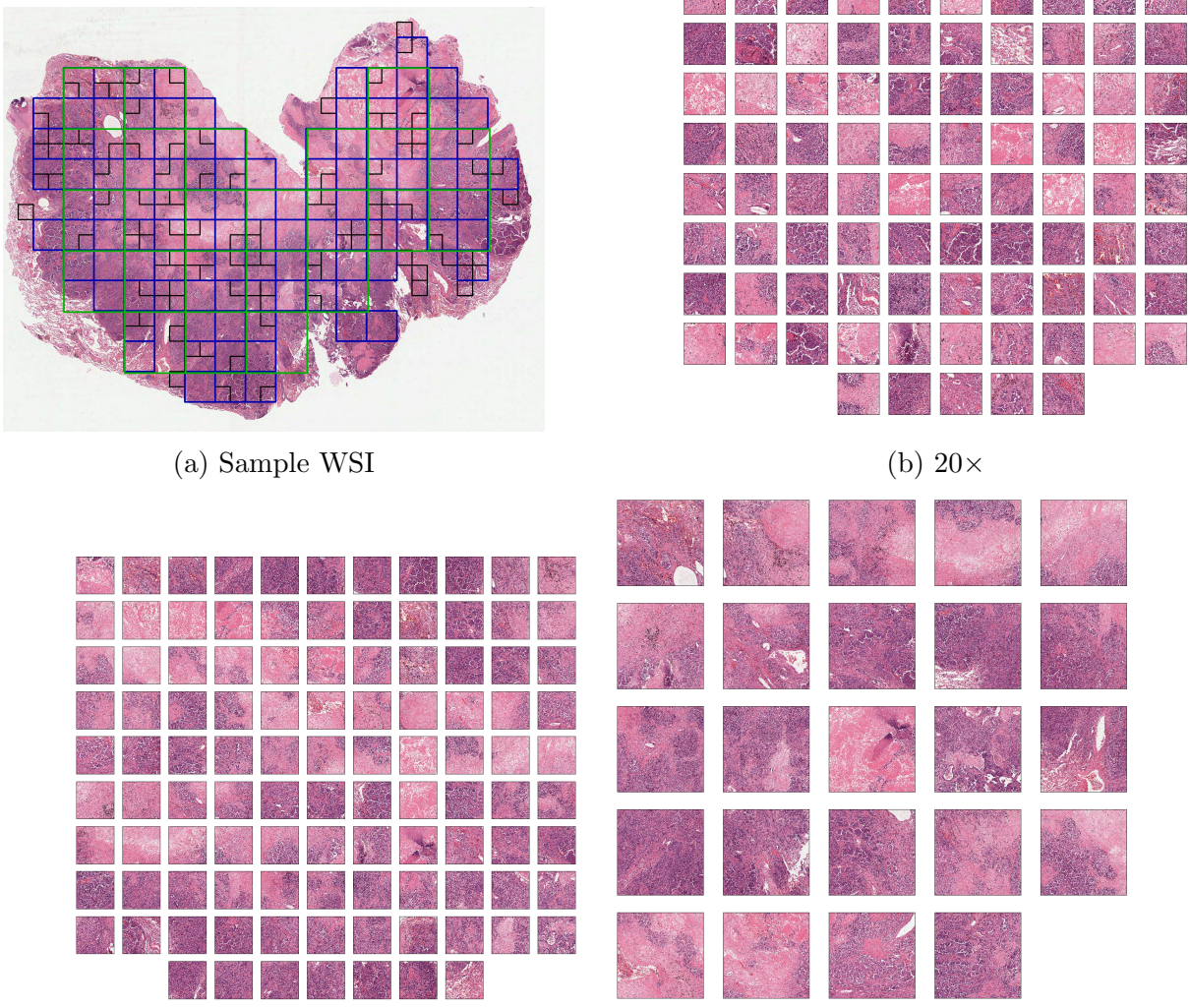


Figure 4.4: Illustration of multi-magnification patch selection in a gigapixel WSI. Green, blue, and black bounding boxes indicate patches at 5×, 10×, and 20× magnification, respectively. All patches are of size 1000 × 1000 pixels.



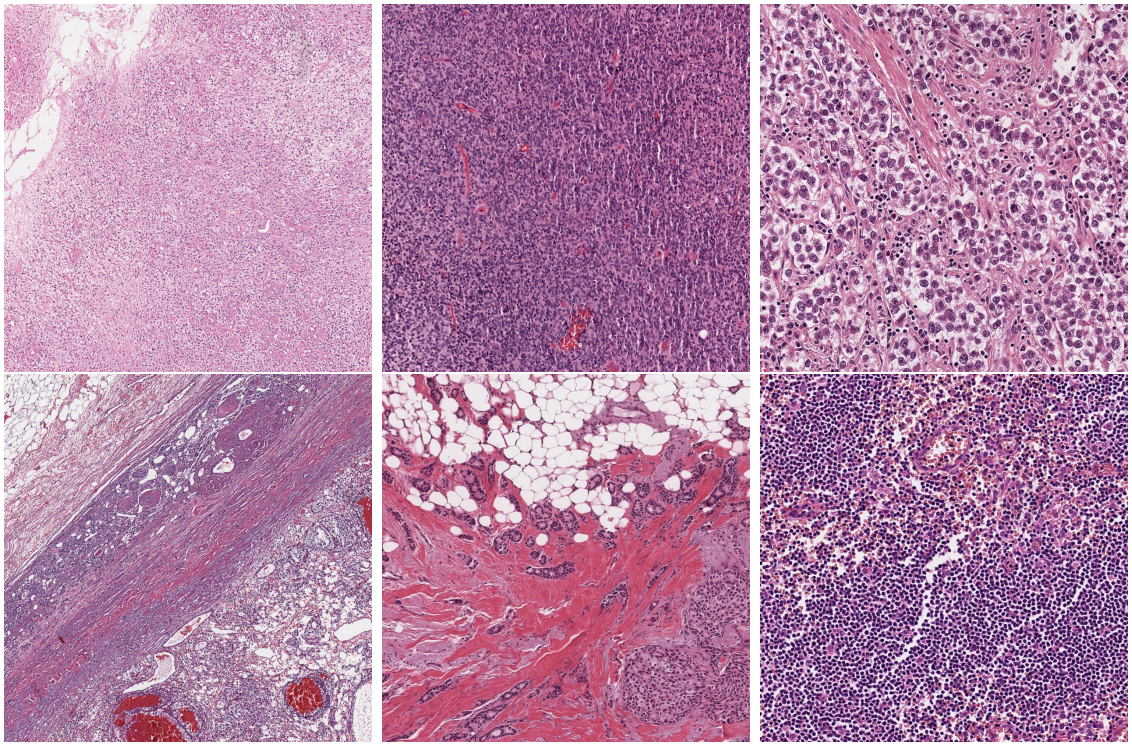
(a) Sample WSI

(b) 20×

(c) 10×

(d) 5×

Figure 4.5: Patch collections of a WSI at 5×, 10×, and 20× magnifications. All patches at all magnification levels have the same size. 20% of tissue regions at 20× is sampled to make a manageable collection and avoid redundancy in dataset.



(a) 5× (b) 10× (c) 20×

Figure 4.6: Sample patches of 5×, 10×, and 20× magnification datasets.

Table 4.1: TCGA cancer types. TCGA Codes are sorted in alphabetical order.

| TCGA Code | Primary Diagnosis |
|---|---|
| ACC | Adrenocortical Carcinoma |
| BLCA | Bladder Urothelial Carcinoma |
| BRCA | Breast Invasive Carcinoma |
| CESC | Cervical squamous cell carcinoma Endocervical adenocarcinoma |
| CHOL | Cholangiocarcinoma |
| COAD | Colon Adenocarcinoma |
| DLBC | Lymphoid Neoplasm Diffuse Large B-cell Lymphoma |
| ESCA | Esophageal Carcinoma |
| GBM | Glioblastoma Multiforme |
| HNSC | Head and Neck Squamous Cell Carcinoma |
| KICH | Kidney Chromophobe |
| KIRC | Kidney Renal Clear Cell Carcinoma |
| KIRP | Kidney Renal Papillary Cell Carcinoma |
| LGG | Brain Lower Grade Glioma |
| LIHC | Liver Hepatocellular Carcinoma |
| LUAD | Lung Adenocarcinoma |
| LUSC | Lung Squamous Cell Carcinoma |
| MESO | Mesothelioma |
| OV | Ovarian Serous Cystadenocarcinoma |
| PAAD | Pancreatic Adenocarcinoma |
| PCPG | Pheochromocytoma and Paraganglioma |
| PRAD | Prostate Adenocarcinoma |
| READ | Rectum Adenocarcinoma |
| SARC | Sarcoma |
| SKCM | Skin Cutaneous Melanoma |
| STAD | Stomach Adenocarcinoma |
| TGCT | Testicular Germ Cell Tumors |
| THCA | Thyroid Carcinoma |
| THYM | Thymoma |
| UCEC | Uterine Corpus Endometrial Carcinoma |
| UCS | Uterine Carcinosarcoma |
| UVM | Uveal Melanoma |

# Chapter 5

# Experiments on MMS

This chapter explains the performed experiments to implement and evaluate the proposed MMS framework. Section 5.1 explains procedures of MMS performance assessment. Section 5.2 describes the evaluation metrics. Section 5.3 provides experimental setup details.

## 5.1 Assessment of the Search Performance

This section explains the procedures to evaluate the performance of the proposed MMS framework. Generally, an image search method can show the results after retrieval and leave the evaluation to the expert user. However, for the purpose of benchmarking, the performance of the search must be somehow quantified. The input of MMS engine is a WSI as the query image, and its output is a set of most similar whole-slide images in the test dataset matched to that query WSI based on the selected magnification. The proposed MMS is applied to digital slides from several anatomical sites, at three magnification levels and their combinations.

Various techniques can be used to evaluate the quality of retrieved WSIs from different points of view. The accuracy of content-based image search methods can be measured if the search is treated like classification. In this regard, this thesis investigated the top-three most similar WSIs to a query WSI found by the search. Next, the majority vote among those three most similar WSIs to the query slide were examined to predict the diagnosis. One may examine the top-$n$ search results for $n > 3$ if larger test datasets are available. Finally, the prediction via majority vote was compared to the actual diagnosis

of the query WSI. In other words, this experiment considers a search successful if two out of three matched cases have the same subtype of malignancy to the query slide. Taking the majority vote is much more rigorous than the top-$n$ accuracy in computer vision that assumes correctness if at least one of the search results is correct.

The anatomical site (organ) of any digital slide is given since the organ of biopsy is known before any assessment. As the primary site is a pivotal knowledge about a WSI, the search space for each query WSI was limited to WSIs with the same anatomical organ. Hence, this experiment would be interested in identifying primary diagnoses for each primary site. In other words, each query WSI was compared to all other WSIs of the same primary site, with the intention of finding cases with the same subtype of malignancy. Then, the top WSIs with minimum distances to that query glswsi were selected. The term "*distance*" here depends on the image search method. Concerning the single-vector approach, the distance metric is the Euclidean distances between WSI embeddings. Concerning the multi-vector method, the distance metric is the median of minimum Hamming distances between hashed feature vectors.

This search experiment requires two text labels for each WSI to evaluate the quality of retrieved WSIs based on the diagnosis similarity. The first label is the primary anatomical site which is the organ biopsied. The second label is the primary diagnosis which indicates the type of malignancy. It must be pointed out that the proposed MMS does not rely on any text information to find matches, and the labels have been used only for evaluation purposes. The following explains how the accuracy of the search was estimated.

This experiment used the "*leave-one-patient-out*" approach, i.e., excluding one WSI, finding similar cases to this WSI in the search space, and repeating this process for all WSIs in the test dataset. The search is successful if two out of three matched cases have the same subtype of malignancy to the query slide. In other words, This experiment considered all WSIs one by one as the search query, and retrieved the three most similar WSI to the query WSI in the test dataset. Note that among the 13 anatomical sites, the classification (majority voting) experiments concern 9 anatomical sites with more than one subtype.

## 5.2 Search Evaluation Measures

*Precision*, *Recall* and *F1-score* are the standard evaluation metrics to assess the performance of image search and CBIR methods [20, 43, 72]. *Precision* indicates the fraction of retrieved documents that are relevant to the query, and is defined as the number of relevant images retrieved by a search divided by the total number of images retrieved by

that search. *Recall* indicates the fraction of the relevant documents that are successfully retrieved, and is defined as the number of relevant images retrieved by a search divided by the total number of relevant images existing in the dataset. *F1-score* is the harmonic mean of the precision and recall. The formula for calculating *F1-score* is shown in Eq. (5.1).

$$\text{F1-score} = \frac{2 \times \text{ Precision} \times \text{ Recall}}{\text{Precision} + \text{Recall}} \tag{5.1}$$

## 5.3 Experimental Setups

Details of conducted experiments for implementation and evaluation of the proposed MMS methods are provided in the following sub-sections. First, Section 5.3.1 describes the fine-tuning setting of the deep feature extractor, i.e., KimiaNet. Next, the implementation procedure of single-vector method and multi-vector method is overviewed in Section 5.3.2 and Section 5.3.3, respectively.

### 5.3.1 Fine-Tuning KimiaNet for $5\times$ and $10\times$ Magnifications

In two independent experiments, the last dense block of KimiaNet was re-trained with $10\times$ and $5\times$ patches. The fine-tuning procedure was implemented using the Tensorflow Keras framework. At the outset, the KimiaNet model was initialized with KimiaNet-IV pre-trained weights [72]. The input of the network was batches of $1000 \times 1000$ patches at $10\times$ or $5\times$ magnification, with the batch size of 128, and their TCGA-codes as corresponding labels ( Table 4.1 on page 34). One Hot encoding was applied on class labels to fine-tune the KimiaNet as a classifier. Also, data augmentation was performed in the original image patches by 90,180 and 270 degrees rotation. In terms of the experimental setup of training, cross-entropy loss function and Adam optimizer [44] were utilized. Two Tesla V100 GPUs with 32GB memory for each GPU were used for fine-tuning the KimiaNet.

After training, the network has learned to convert the input histopathology to a high-level representation. To extract features and represent histopathology patches, we used the last average pooling layer of the fine-tuned network. Accordingly, the size of each feature vector is 1024.

### 5.3.2 Experimental Setups for Single-vector Method

Seven `Python` Dictionaries were generated for experiments on single-vector method since we have evaluated this method on a magnification-basis, i.e., at 5×,10×,20×, 5 × 10×, 5 × 20×, 10 × 20×, and 5 × 10 × 20× magnifications. In each dictionary, a `key` is the name of a WSI, and its corresponding `value` is a single vector representation of that WSI at that magnification(s). In the retrieval phase, the test dataset must be ranked based on the euclidean distances (similarity) to the query WSI to retrieve the most similar matches. `Scipy` library[1] was used for calculating euclidean distances between WSI representations.

#### T-SNE Experiment

We used t-SNE [94] to visualize single-vector representations. Single-vector representations are high-dimensional feature vectors, with a dimension of 1024, 2048, or 3096, depending on the magnification settings. `Sklearn` [2] library were used to convert high-dimensional embeddings to low-dimensional embeddings for visualization. Considering the very high number of dimensions, we used PCA for the dimensionality reduction to 64 before applying the t-SNE method. The parameters of t-sne, i.e., perplexity, learning rate, and the number of iterations, were set empirically to 15, 800, and 2000, respectively.

### 5.3.3 Experimental Setups for Multi-Vector Method

After the deep feature extraction from KimiaNet, the first step is barcoding the feature vectors to speed up the distance calculation between binary codes. Parameters in the SNRQ method for barcoding were set empirically. The algorithm was run for 70 iterations, with the control quantization power of 3.

Next, a Python class with three functions was implemented for patch-to-patch, patch-to-WSI, and WSI-to-WSI distance calculations, respectively. Several `Numpy` and `Scipy` built-in function were used for implementing the median-of-mins algorithm in the multi-vector method.

In the WSI-to-WSI matrix (`2D- Numpy array`), `row[i]` shows the distance between $i^{th}$ WSI to all other WSIs in the test dataset. Accordingly, in the search phase, the most similar WSIs to the $i^{th}$ WSI are associated with indices of the minimum values, i.e., minimum distances, along the `row[i]`.

[1] https://docs.scipy.org/doc/scipy/reference/generated/scipy.spatial.distance.cdist.html
[2] https://scikit-learn.org/stable/modules/generated/sklearn.manifold.TSNE.html

# Chapter 6

# Results and Analysis

This chapter reports and analyzes the search results of the MMS framework. At the outset, Section 6.1 discusses the discriminative power of the single-vector method using the t-SNE visualization. The second section, Section 6.2, reports search performances at different mixtures of magnification levels for several anatomical sites. Section 6.3 analyzes the search results with respect to the anatomical sites. Finally, Section 6.4 discusses the most important findings of the MMS.

## 6.1 T-SNE Visualization of Single-Vector Method

Fig. 6.1 and Fig. 6.2 visualize the TM and 20× magnification representations of 744 WSIs from 13 anatomical sites on the test dataset, respectively. WSIs with the same anatomical site are shown by the same color in this visualization. WSIs with different tumor types of the same anatomical are shown by same color and different markers. The presented t-SNE is associated with the WSI representation via single-vector method. The multi-vector WSI representation method is based on the patch-to-patch distances and does not offer any single vector representation for a WSI. Thus, there is no t-SNE visualization associated with the multi-vector method.

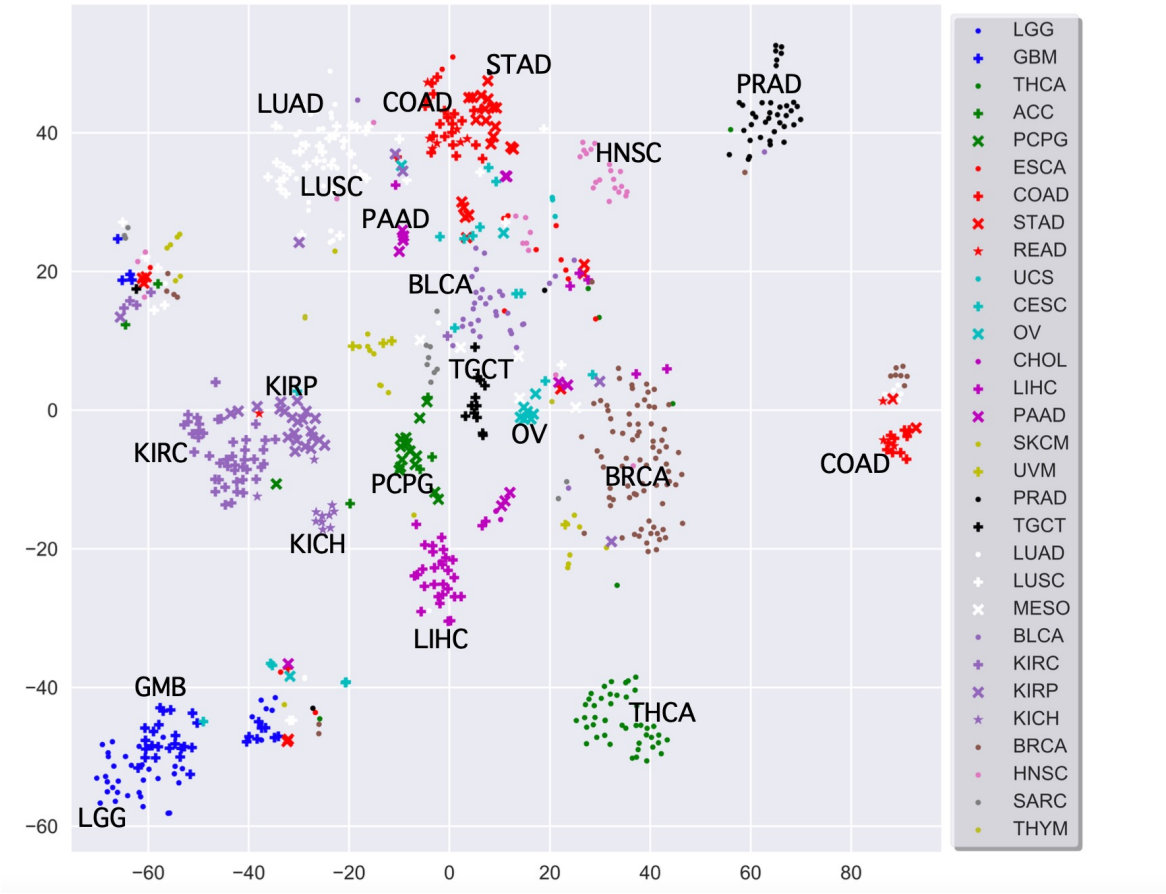



Figure 6.1: T-SNE visualization of TriMagnification embeddings (Single-Vector method). Each point of this t-SNE picture displays a feature vector associated with one WSI in the test dataset. Each color is associated with one anatomical site. Tumor types of the same anatomical are indicated by different markers, e.g., plus, x, point, and star.

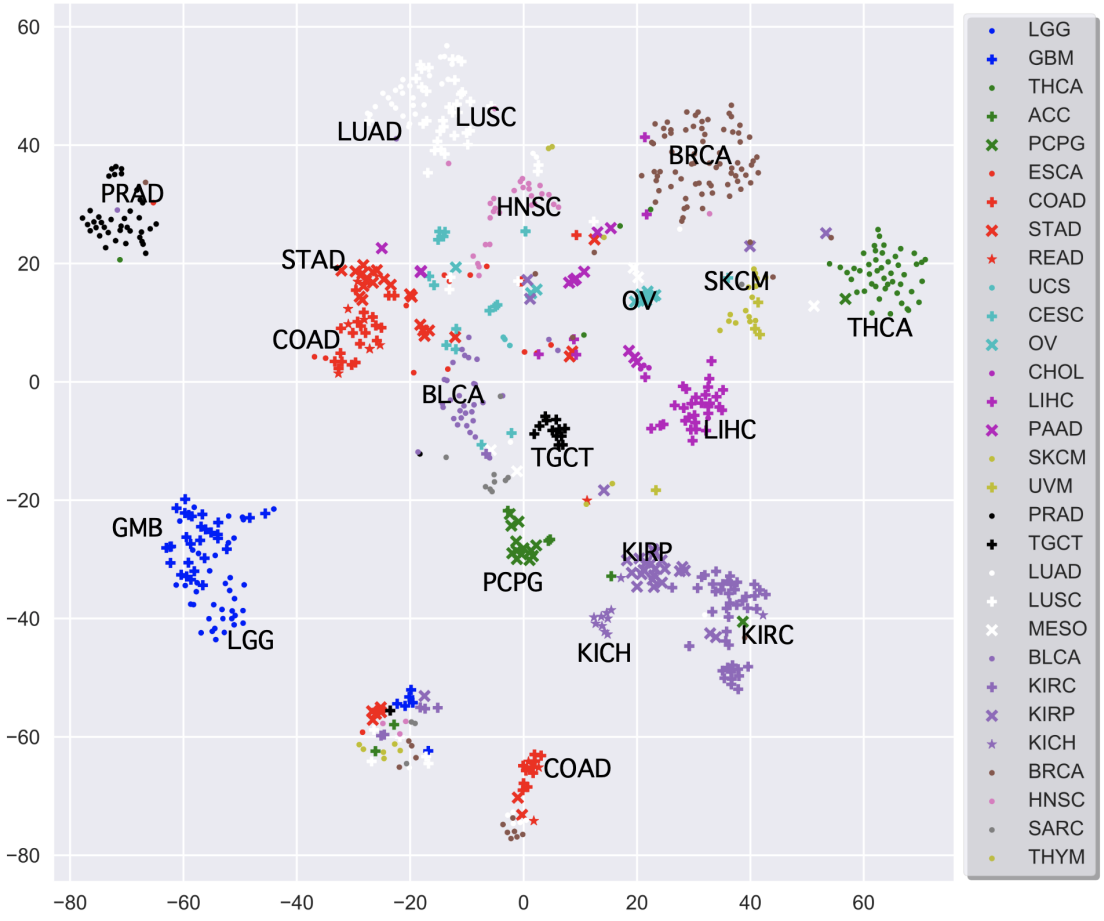



Figure 6.2: T-SNE visualization of 20× Magnification embeddings (Single-Vector method). Each point of this t-SNE picture displays a feature vector associated with one WSI in the test dataset. Each color is associated with one anatomical site. Tumor types of the same anatomical are indicated by different markers, e.g., plus, x, point, and star.

The t-SNE visualizations show clear class discrimination of WSI features. It can be seen from the data in the t-SNE visualization that tissue types and subtypes are generally well clustered. Although the single-vector search method is not aware of the anatomical sites (feature extractors trained with primary diagnosis labels), tumors derived from the same anatomical site are generally formed in close clusters. The most apparent results are brain, lung, and kidney with 2, 2, and 3 distinct sub-clusters associated with tumor subtypes

(primary diagnoses), respectively. Concerning kidney tumour sub-types, there are three distinct and adjacent clusters displaying each subtype in both Fig. 6.2 and Fig. 6.2. These clusters are associated with Kidney Renal Papillary Cell Carcinoma (KIRP), Kidney Renal Clear Cell Carcinoma (KIRC), and Kidney Chromophobe (KICH). However, the WSI feature vectors associated with the anatomical site of the Bladder, i.e., Bladder Urothelial Carcinoma (BLCA), has not created a single cluster, not in proximity to kidney tumors. This may be attributed to the fact that although [18] considers KIRP, KIRC, KICH, and BLCA subtypes of urinary tract tumors, many pathology books consider bladder and kidney two distant tumor types [61].

A white-colour cluster of lung tumors, i.e., LUAD and LUSC, is also evident in both Fig. 6.2 and Fig. 6.2. Two main clusters associated with gastrointestinal tract tumors (red colour) were recognized as well. Concerning prostate/testis tumors, two disjoint clusters (green colour) were observed in both t-SNE visualizations at 20× magnifciation and TM. Testis tumors (THCA) are perfectly clustered with considerable margin from other tumor types. When it comes to pancreas/liver tumors (black colour), there are two distant clusters associated with pancreas (PAAD) and liver (LIHC). Some consider the two tumor types for pancreas and liver since they have different pathology features [61].

## 6.2 Search Accuracy Report of MMS

This section reports the results of several experiments performed to evaluate the performance of two MMS methods. These experimental results investigate the impact of magnification on diagnosis-wise image search.

Detailed results obtained from the magnification analysis of proposed single-vector searching methods among 744 test digital slides are reported in Tables 6.1 to 6.3. Search results of the proposed multi-vector searching methods is presented in Tables 6.4 to 6.6. The performance of both MMS methods are evaluated and reported in terms of precision, recall, and F1-score. There are seven columns in each table, corresponding with seven experiments at 5×, 10×, 20×, 5×10×, 5×20×, 10×20×, and all three magnifications (TM), i.e., 5 × 10 × 20×. These experiments were conducted independently on nine anatomical sites with more than one tumour subtype. Note that according to categorization proposed by [18], 9 anatomical sites among 13 anatomical sites in the dataset have more than one tumor subtypes. In each experiment, three most similar WSIs to a query WSI were retrieved. The search was evaluated as correct if the majority of retrieved WSIs have the same type of malignancy (primary diagnosis) as the query WSI. This process was repeated

Table 6.1: Percentage of precision search results for the Single-Vector method at three levels of magnification and their combinations. Search considered successful if the majority of retrieved WSIs from an anatomical site has the same tumor type as the query WSI in the leave-one-out procedure.

| Site | Subtype | #samples | 20x | 10x | 5x | 5x10x | 5x20x | 10x20x | TM |
|---|---|---|---|---|---|---|---|---|---|
| Brain | GBM | 35 | 72 | 70 | 72 | 69 | 72 | 72 | 80 |
| Brain | LGG | 39 | 96 | 92 | 85 | 83 | 96 | 96 | 100 |
| Endocrine | ACC | 6 | 100 | 0 | 0 | 0 | 100 | 100 | 0 |
| Endocrine | PCPG | 15 | 74 | 90 | 75 | 69 | 78 | 78 | 72 |
| Endocrine | THCA | 51 | 98 | 82 | 88 | 89 | 96 | 96 | 96 |
| Gastro | COAD | 32 | 61 | 61 | 60 | 53 | 61 | 61 | 56 |
| Gastro | ESCA | 14 | 88 | 75 | 83 | 86 | 88 | 88 | 86 |
| Gastro | READ | 12 | 38 | 27 | 18 | 18 | 38 | 36 | 31 |
| Gastro | STAD | 30 | 77 | 65 | 71 | 69 | 79 | 79 | 81 |
| Gynaeco | CESC | 17 | 85 | 65 | 75 | 71 | 81 | 85 | 81 |
| Gynaeco | OV | 10 | 100 | 75 | 80 | 73 | 100 | 100 | 100 |
| Gynaeco | UCS | 3 | 100 | 0 | 0 | 0 | 100 | 100 | 100 |
| Liver/panc | CHOL | 4 | 43 | 33 | 25 | 0 | 43 | 50 | 50 |
| Liver/panc | LIHC | 35 | 94 | 94 | 100 | 100 | 97 | 94 | 97 |
| Liver/panc | PAAD | 12 | 100 | 59 | 64 | 69 | 100 | 100 | 91 |
| Melanocytic | SKCM | 24 | 86 | 86 | 88 | 89 | 86 | 86 | 86 |
| Melanocytic | UVM | 4 | 0 | 0 | 100 | 100 | 0 | 0 | 0 |
| Prostate/testis | PRAD | 40 | 98 | 97 | 91 | 91 | 98 | 98 | 98 |
| Prostate/testis | TGCT | 13 | 100 | 86 | 100 | 100 | 100 | 100 | 100 |
| Pulmonary | LUAD | 38 | 74 | 62 | 58 | 52 | 74 | 74 | 74 |
| Pulmonary | LUSC | 43 | 85 | 68 | 62 | 55 | 85 | 85 | 80 |
| Pulmonary | MESO | 5 | 75 | 50 | 0 | 0 | 75 | 75 | 67 |
| Urinary tract | BLCA | 34 | 92 | 83 | 91 | 88 | 92 | 92 | 97 |
| Urinary tract | KICH | 11 | 100 | 70 | 83 | 83 | 100 | 100 | 90 |
| Urinary tract | KIRC | 50 | 91 | 85 | 75 | 74 | 91 | 91 | 94 |
| Urinary tract | KIRP | 28 | 88 | 79 | 83 | 72 | 88 | 88 | 96 |

Table 6.2: Percentage of recall search results for the Single-Vector method at three magnification and their combinations. Search considered successful if the majority of retrieved WSIs from an anatomical site has the same tumor type as the query WSI in the leave-one-out procedure.

| Site | Subtype | #samples | 20x | 10x | 5x | 5x10x | 5x20x | 10x20x | TM |
|---|---|---|---|---|---|---|---|---|---|
| Brain | GBM | 35 | 97 | 94 | 85 | 85 | 97 | 97 | 100 |
| Brain | LGG | 39 | 66 | 63 | 72 | 66 | 66 | 66 | 76 |
| Endocrine | ACC | 6 | 33 | 0 | 0 | 0 | 17 | 17 | 0 |
| Endocrine | PCPG | 15 | 93 | 60 | 60 | 60 | 93 | 93 | 87 |
| Endocrine | THCA | 51 | 98 | 100 | 100 | 100 | 100 | 100 | 100 |
| Gastro | COAD | 32 | 69 | 72 | 81 | 72 | 72 | 72 | 72 |
| Gastro | ESCA | 14 | 50 | 46 | 42 | 46 | 50 | 50 | 43 |
| Gastro | READ | 12 | 42 | 33 | 17 | 17 | 42 | 42 | 33 |
| Gastro | STAD | 30 | 80 | 57 | 61 | 60 | 77 | 73 | 73 |
| Gynaeco | CESC | 17 | 100 | 87 | 86 | 80 | 100 | 100 | 100 |
| Gynaeco | OV | 10 | 80 | 60 | 89 | 80 | 70 | 80 | 80 |
| Gynaeco | UCS | 3 | 67 | 0 | 0 | 0 | 67 | 67 | 33 |
| Liver/panc | CHOL | 4 | 75 | 25 | 25 | 0 | 75 | 75 | 75 |
| Liver/panc | LIHC | 35 | 97 | 83 | 97 | 94 | 97 | 97 | 94 |
| Liver/panc | PAAD | 12 | 67 | 83 | 70 | 92 | 75 | 75 | 83 |
| Melanocytic | SKCM | 24 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| Melanocytic | UVM | 4 | 0 | 0 | 25 | 25 | 0 | 0 | 0 |
| Prostate/testis | PRAD | 40 | 100 | 95 | 100 | 100 | 100 | 100 | 100 |
| Prostate/testis | TGCT | 13 | 92 | 92 | 69 | 69 | 92 | 92 | 92 |
| Pulmonary | LUAD | 38 | 82 | 66 | 54 | 45 | 82 | 82 | 76 |
| Pulmonary | LUSC | 43 | 79 | 70 | 72 | 67 | 79 | 79 | 81 |
| Pulmonary | MESO | 5 | 60 | 20 | 0 | 0 | 60 | 60 | 40 |
| Urinary tract | BLCA | 34 | 97 | 100 | 88 | 88 | 97 | 97 | 100 |
| Urinary tract | KICH | 11 | 82 | 64 | 50 | 45 | 82 | 82 | 82 |
| Urinary tract | KIRC | 50 | 96 | 90 | 98 | 96 | 96 | 96 | 98 |
| Urinary tract | KIRP | 28 | 79 | 54 | 54 | 46 | 79 | 79 | 89 |

Table 6.3: Percentage of F1-score search results for Single-Vector method at three magnification and their combinations. Search considered successful if the majority of retrieved WSIs from an anatomical site has the same tumor type as the query WSI in the leave-one-out procedure.

| Site | Subtype | #samples | 20x | 10x | 5x | 5x10x | 5x20x | 10x20x | TM |
|---|---|---|---|---|---|---|---|---|---|
| Brain | GBM | 35 | 83 | 80 | 78 | 76 | 83 | 83 | 89 |
| Brain | LGG | 39 | 78 | 75 | 78 | 74 | 78 | 78 | 87 |
| Endocrine | ACC | 6 | 50 | 0 | 0 | 0 | 29 | 29 | 0 |
| Endocrine | PCPG | 15 | 82 | 72 | 67 | 64 | 85 | 85 | 79 |
| Endocrine | THCA | 51 | 98 | 90 | 93 | 94 | 98 | 98 | 98 |
| Gastro | COAD | 32 | 65 | 66 | 68 | 61 | 66 | 66 | 63 |
| Gastro | ESCA | 14 | 64 | 57 | 56 | 60 | 64 | 64 | 57 |
| Gastro | READ | 12 | 40 | 30 | 17 | 17 | 40 | 38 | 32 |
| Gastro | STAD | 30 | 79 | 61 | 65 | 64 | 78 | 76 | 77 |
| Gynaeco | CESC | 17 | 92 | 74 | 80 | 75 | 89 | 92 | 89 |
| Gynaeco | OV | 10 | 89 | 67 | 84 | 76 | 82 | 89 | 89 |
| Gynaeco | UCS | 3 | 80 | 0 | 0 | 0 | 80 | 80 | 50 |
| Liver/panc | CHOL | 4 | 55 | 29 | 25 | 0 | 55 | 60 | 60 |
| Liver/panc | LIHC | 35 | 96 | 88 | 99 | 97 | 97 | 96 | 96 |
| Liver/panc | PAAD | 12 | 80 | 69 | 67 | 79 | 86 | 86 | 87 |
| Melanocytic | SKCM | 24 | 92 | 92 | 94 | 94 | 92 | 92 | 92 |
| Melanocytic | UVM | 4 | 0 | 0 | 40 | 40 | 0 | 0 | 0 |
| Prostate/testis | PRAD | 40 | 99 | 96 | 95 | 95 | 99 | 99 | 99 |
| Prostate/testis | TGCT | 13 | 96 | 89 | 82 | 82 | 96 | 96 | 96 |
| Pulmonary | LUAD | 38 | 78 | 64 | 56 | 48 | 78 | 78 | 75 |
| Pulmonary | LUSC | 43 | 82 | 69 | 67 | 60 | 82 | 82 | 80 |
| Pulmonary | MESO | 5 | 67 | 29 | 0 | 0 | 67 | 67 | 50 |
| Urinary tract | BLCA | 34 | 94 | 91 | 90 | 88 | 94 | 94 | 99 |
| Urinary tract | KICH | 11 | 90 | 67 | 62 | 59 | 90 | 90 | 86 |
| Urinary tract | KIRC | 50 | 93 | 87 | 85 | 83 | 93 | 93 | 96 |
| Urinary tract | KIRP | 28 | 83 | 64 | 65 | 57 | 83 | 83 | 93 |

Table 6.4: Percentage of Precision search results for the Multi-Vector method at three magnification and their combinations. Search considered successful if the majority of retrieved WSIs from an anatomical site has the same tumor type as the query WSI in the leave-one-out procedure.

| Site | Subtype | #samples | 20x | 10x | 5x | 5x10x | 5x20x | 10x20x | TM |
|---|---|---|---|---|---|---|---|---|---|
| Brain | GBM | 35 | 79 | 74 | 59 | 72 | 79 | 79 | 79 |
| Brain | LGG | 39 | 88 | 85 | 66 | 78 | 88 | 86 | 86 |
| Endocrine | ACC | 6 | 0 | 0 | 25 | 0 | 0 | 0 | 0 |
| Endocrine | PCPG | 15 | 78 | 75 | 71 | 75 | 75 | 79 | 78 |
| Endocrine | THCA | 51 | 94 | 91 | 92 | 93 | 93 | 96 | 94 |
| Gastro | COAD | 32 | 71 | 60 | 67 | 67 | 71 | 67 | 66 |
| Gastro | ESCA | 14 | 78 | 67 | 71 | 83 | 75 | 75 | 75 |
| Gastro | READ | 12 | 33 | 20 | 44 | 38 | 50 | 43 | 43 |
| Gastro | STAD | 30 | 69 | 64 | 54 | 62 | 69 | 68 | 66 |
| Gynaeco | CESC | 17 | 89 | 75 | 81 | 78 | 84 | 84 | 81 |
| Gynaeco | OV | 10 | 100 | 100 | 88 | 88 | 86 | 86 | 100 |
| Gynaeco | UCS | 3 | 60 | 100 | 100 | 100 | 75 | 75 | 67 |
| Liver/panc | CHOL | 4 | 20 | 40 | 17 | 25 | 17 | 29 | 25 |
| Liver/panc | LIHC | 35 | 89 | 91 | 91 | 91 | 94 | 91 | 91 |
| Liver/panc | PAAD | 12 | 78 | 69 | 50 | 67 | 82 | 80 | 67 |
| Melanocytic | SKCM | 24 | 89 | 86 | 88 | 88 | 89 | 86 | 86 |
| Melanocytic | UVM | 4 | 100 | 0 | 33 | 50 | 100 | 0 | 0 |
| Prostate/testis | PRAD | 40 | 100 | 97 | 95 | 98 | 98 | 100 | 98 |
| Prostate/testis | TGCT | 13 | 93 | 86 | 85 | 92 | 92 | 93 | 92 |
| Pulmonary | LUAD | 38 | 71 | 76 | 69 | 76 | 70 | 77 | 76 |
| Pulmonary | LUSC | 43 | 74 | 71 | 62 | 62 | 67 | 77 | 67 |
| Pulmonary | MESO | 5 | 80 | 0 | 0 | 0 | 50 | 75 | 50 |
| Urinary tract | BLCA | 34 | 92 | 92 | 83 | 92 | 92 | 97 | 97 |
| Urinary tract | KICH | 11 | 91 | 83 | 57 | 80 | 91 | 100 | 100 |
| Urinary tract | KIRC | 50 | 91 | 92 | 80 | 87 | 91 | 92 | 86 |
| Urinary tract | KIRP | 28 | 91 | 91 | 78 | 91 | 95 | 96 | 100 |

Table 6.5: Percentage of recall search results for the Multi-Vector method at three magnification and their combinations. Search considered successful if the majority of retrieved WSIs from an anatomical site has the same tumor type as the query WSI in the leave-one-out procedure.

| Site | Subtype | #samples | 20x | 10x | 5x | 5x10x | 5x20x | 10x20x | TM |
|---|---|---|---|---|---|---|---|---|---|
| Brain | GBM | 35 | 89 | 85 | 59 | 76 | 89 | 86 | 86 |
| Brain | LGG | 39 | 79 | 74 | 66 | 74 | 79 | 79 | 79 |
| Endocrine | ACC | 6 | 0 | 0 | 17 | 0 | 0 | 0 | 0 |
| Endocrine | PCPG | 15 | 93 | 80 | 67 | 80 | 80 | 100 | 93 |
| Endocrine | THCA | 51 | 98 | 100 | 98 | 98 | 100 | 100 | 100 |
| Gastro | COAD | 32 | 75 | 66 | 84 | 81 | 84 | 81 | 78 |
| Gastro | ESCA | 14 | 50 | 46 | 42 | 38 | 43 | 43 | 43 |
| Gastro | READ | 12 | 25 | 25 | 33 | 25 | 25 | 25 | 25 |
| Gastro | STAD | 30 | 83 | 60 | 54 | 70 | 83 | 77 | 77 |
| Gynaeco | CESC | 17 | 94 | 100 | 93 | 93 | 94 | 94 | 100 |
| Gynaeco | OV | 10 | 70 | 60 | 78 | 70 | 60 | 60 | 60 |
| Gynaeco | UCS | 3 | 100 | 67 | 67 | 67 | 100 | 100 | 67 |
| Liver/panc | CHOL | 4 | 25 | 50 | 25 | 25 | 25 | 50 | 25 |
| Liver/panc | LIHC | 35 | 94 | 86 | 86 | 91 | 91 | 89 | 91 |
| Liver/panc | PAAD | 12 | 58 | 75 | 50 | 67 | 75 | 67 | 67 |
| Melanocytic | SKCM | 24 | 100 | 100 | 91 | 96 | 100 | 100 | 100 |
| Melanocytic | UVM | 4 | 25 | 0 | 25 | 25 | 25 | 0 | 0 |
| Prostate/testis | PRAD | 40 | 98 | 95 | 95 | 98 | 98 | 98 | 98 |
| Prostate/testis | TGCT | 13 | 100 | 92 | 85 | 92 | 92 | 100 | 92 |
| Pulmonary | LUAD | 38 | 66 | 68 | 51 | 50 | 61 | 71 | 58 |
| Pulmonary | LUSC | 43 | 79 | 84 | 82 | 86 | 79 | 84 | 86 |
| Pulmonary | MESO | 5 | 80 | 0 | 0 | 0 | 20 | 60 | 20 |
| Urinary tract | BLCA | 34 | 97 | 100 | 100 | 100 | 100 | 100 | 100 |
| Urinary tract | KICH | 11 | 91 | 91 | 40 | 73 | 91 | 91 | 91 |
| Urinary tract | KIRC | 50 | 96 | 94 | 90 | 94 | 96 | 96 | 98 |
| Urinary tract | KIRP | 28 | 75 | 75 | 50 | 71 | 75 | 89 | 75 |

Table 6.6: Percentage of F1-score search results for the Multi-Vector search at three magnification and their combinations. Search considered successful if the majority of retrieved WSIs from an anatomical site has the same tumor type as the query WSI in the leave-one-out procedure.

| Site | Subtype | #samples | 20x | 10x | 5x | 5x10x | 5x20x | 10x20x | TM |
|---|---|---|---|---|---|---|---|---|---|
| Brain | GBM | 35 | 84 | 79 | 59 | 74 | 84 | 82 | 82 |
| Brain | LGG | 39 | 83 | 79 | 66 | 76 | 83 | 82 | 82 |
| Endocrine | ACC | 6 | 0 | 0 | 20 | 0 | 0 | 0 | 0 |
| Endocrine | PCPG | 15 | 85 | 77 | 69 | 77 | 77 | 88 | 85 |
| Endocrine | THCA | 51 | 96 | 95 | 95 | 95 | 96 | 98 | 97 |
| Gastro | COAD | 32 | 73 | 63 | 74 | 73 | 77 | 73 | 71 |
| Gastro | ESCA | 14 | 61 | 55 | 53 | 53 | 55 | 55 | 55 |
| Gastro | READ | 12 | 29 | 22 | 38 | 30 | 33 | 32 | 32 |
| Gastro | STAD | 30 | 76 | 62 | 54 | 66 | 76 | 72 | 71 |
| Gynaeco | CESC | 17 | 91 | 86 | 87 | 85 | 89 | 89 | 89 |
| Gynaeco | OV | 10 | 82 | 75 | 82 | 78 | 71 | 71 | 75 |
| Gynaeco | UCS | 3 | 75 | 80 | 80 | 80 | 86 | 86 | 67 |
| Liver/panc | CHOL | 4 | 22 | 44 | 20 | 25 | 20 | 36 | 25 |
| Liver/panc | LIHC | 35 | 92 | 88 | 88 | 91 | 93 | 90 | 91 |
| Liver/panc | PAAD | 12 | 67 | 72 | 50 | 67 | 78 | 73 | 67 |
| Melanocytic | SKCM | 24 | 94 | 92 | 89 | 92 | 94 | 92 | 92 |
| Melanocytic | UVM | 4 | 40 | 0 | 29 | 33 | 40 | 0 | 0 |
| Prostate/testis | PRAD | 40 | 99 | 96 | 95 | 98 | 98 | 99 | 98 |
| Prostate/testis | TGCT | 13 | 96 | 89 | 85 | 92 | 92 | 96 | 92 |
| Pulmonary | LUAD | 38 | 68 | 72 | 59 | 60 | 65 | 74 | 66 |
| Pulmonary | LUSC | 43 | 76 | 77 | 71 | 72 | 72 | 80 | 76 |
| Pulmonary | MESO | 5 | 80 | 0 | 0 | 0 | 29 | 67 | 29 |
| Urinary tract | BLCA | 34 | 94 | 96 | 91 | 96 | 96 | 99 | 99 |
| Urinary tract | KICH | 11 | 91 | 87 | 47 | 76 | 91 | 95 | 95 |
| Urinary tract | KIRC | 50 | 93 | 93 | 85 | 90 | 93 | 94 | 92 |
| Urinary tract | KIRP | 28 | 82 | 82 | 61 | 80 | 84 | 93 | 86 |

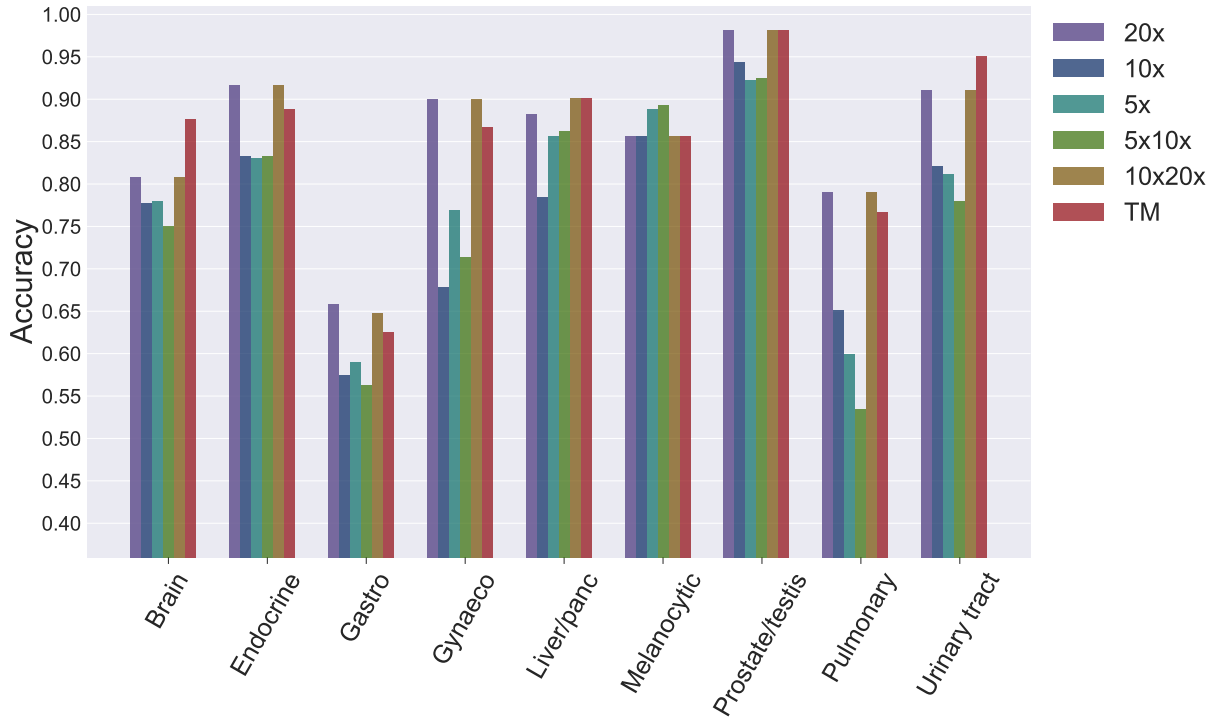



Figure 6.3: Accuracy histogram of Single-Vector (*Median Aggregation*) with respect to the magnification level and anatomical site. Vertical axis shows fraction of correct subtype classifications. For each anatomical site, the total image search accuracy at different magnification levels is visualized.

for all WSIs in the test dataset through "leave-one-out" approach, and the accuracy of the experiment was calculated accordingly.

Since the search is interpreted as a classification task to measure the search performance numerically, the classification accuracy can be also calculated accordingly. Figs. 6.3 and 6.4 show how accurate the search methods are in the tumor subtype classification of each anatomical site. Fig. 6.3 shows the overall classification accuracy achieved at each magnification using the *single-vector* method. For example, this histogram shows that classifying brain WSIs into GBM and LGG are more than 85 percent accurate using the single-vector method at TM. Note that classification means majority voting between retrieved WSIs. Fig. 6.4 presents the overall classification accuracy of *multi-vector* method at each magnification for searching among different anatomical sites in the test dataset.

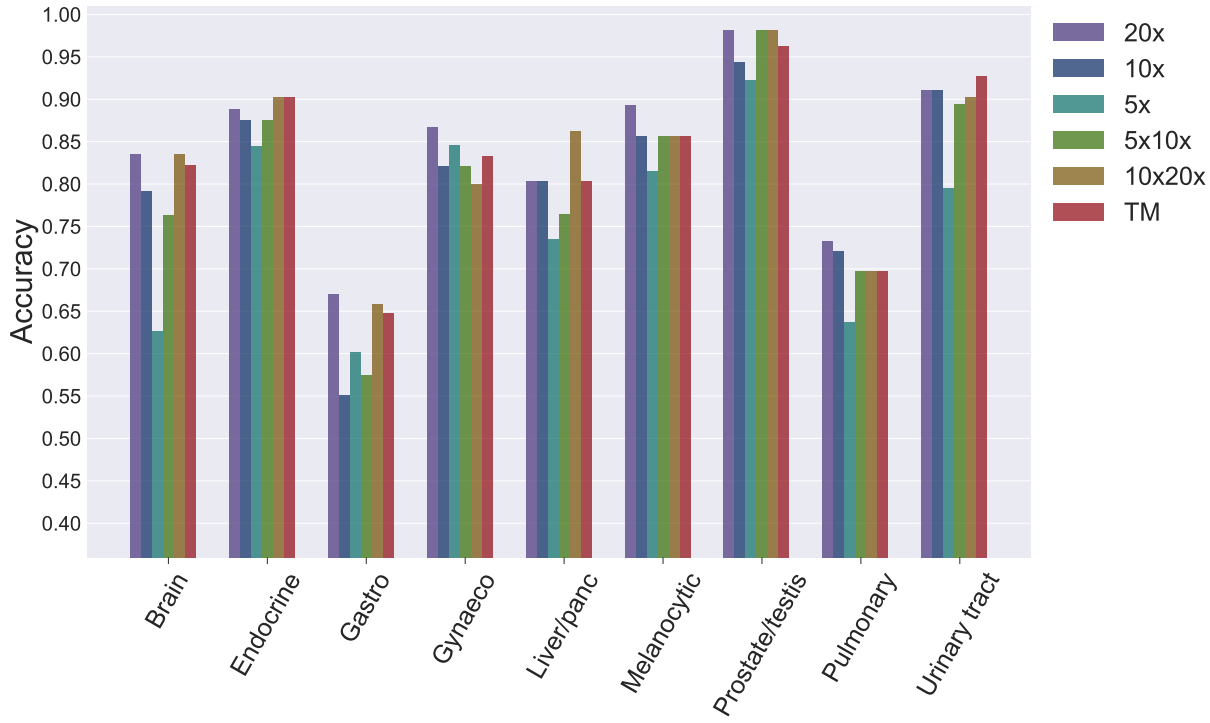



Figure 6.4: Accuracy histogram of Multi-Vector method (*median-of-mins*) with respect to the magnification level and anatomical site. Vertical axis shows fraction of correct subtype classifications. For each anatomical site, the total image search accuracy at different magnification levels is visualized.

## 6.3 Analysis of Results

This section analyzes the search performance. The aim here is to suggest the best search method, and above that, to suggest the most effective magnification level (or combination of levels) relevant to each tumor type.

Since F1-score is the harmonic mean of the precision and recall, this measure can reflect both specificity and sensitivity of the search method. Accordingly, F1-score was selected for the comparison of search performance in the following analysis of results. Note that F1-score has been the most common way to evaluate the search performance in the literature [20, 42, 43, 72]. Accordingly, this section analyzes the search mainly based on the F1-score results presented in Table 6.3 and Table 6.6.

### Brain

Using TM instead of 20× magnification improved the single-vector method performance, increasing the F1-score for Glioblastoma Multiforme (GBM) and Low-grade Gliomas (LGG) from 0.83 and 0.78 to 0.89 and 0.87, respectively. Results indicate that MMS using the single-vector algorithm at TM may be the most appropriate configuration for searching among brain tumors.

### Endocrine

The class imbalance challenge in this dataset is noticeable. Adrenocortical carcinoma (ACC), pheochromocytoma and paraganglioma (PCPG), and thyroid carcinoma (THCA) have 6, 15, and 51 samples, respectively. This disparity in the number of subtype samples markedly affects the search performance. There are not enough samples of ACC tumors, causing confusion in finding two or more similar cases and resulting in zero F1-scores in the evaluation using some magnifications/methods. Overall, results indicate that utilizing 10× magnification along with 20× magnification has been beneficial for the multi-vector method, improving the F1-score accuracy of PCPG and THCA by three and two percent, respectively, compared to 20× magnification. The results support using the single-vector method at 20× magnification or at the $10 \times 20\times$ magnifications.

### Gastrointestinal Tract

The single-vector method at 20× magnification achieved the most accurate results concerning the image search among gastrointestinal tract tumors. However, the multi-vector method at $5 \times 20\times$ magnification surprisingly showed 12 percent F1-score improvement compared to the previously suggested configuration for searching for colon adenocarcinoma (COAD) cases.

### Gynecological Tumors

The limited number of samples for uterine carcinosarcoma tumors (UCS) negatively affected the performance of the single-vector method at 10× and 5× magnifications and their combination. Interestingly, the multi-vector method showed satisfactory performance across all magnifications. However, more samples are required for a more reliable assessment. Overall, the single-vector method at 20× magnification is the best MMS configuration for diagnosis-related searches among gynecological WSIs.

### Liver, Pancreaticobiliary

The single-vector method at TM achieved the most accurate results in overall. Interestingly, single-vector method 5× achieved the F1-score of 0.99 concerning liver hepatocellular carcinoma (LIHC).

### Melanocytic Malignancies

Since there are only four samples of uveal melanoma (UVM) tumors, only three samples were left in the *leave-one-out* approach. Therefore, finding two or more UVM matches within three retrieved samples can be considered slightly unlikely. Overall, the single-vector method at 5×, 5 × 10×, and the multi-vector method at 20× and 5 × 20× magnifications achieved the most accurate results. Since the computational complexity using the single-vector method at 5× magnification is less than other configurations, representing WSIs with melanocytic malignancies using the single-vector method at 5× magnification is preferred.

### Prostate

Image search at 20× achieved accurate results with 0.99 and 0.96 F1-score for prostate adenocarcinoma (PRAD) and testicular germ cell tumors (TGCT), respectively. These results are regardless of the searching method, i.e., both single-vector and multi-vector methods achieved the same F1-score results. Combining lower magnification levels information with 20× magnification information using the single-vector method also achieved similar results as for the 20× magnification.

### Pulmonary

Top search results concerning lung tumors, i.e., lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC), were achieved using the single-vector method at 20× magnification, and also the combination of 20× magnification with one lower magnification, i.e., 10 × 20× and 5 × 20×. Concerning mesothelioma (MESO) diseases, the multi-vector method showed an unanticipated improvement in the F1-score compared to the single-vector method at 20× magnification, improving from 67 to 80 percent.

#### Urinary Tract

The most apparent advantage of utilizing multi-magnification can be seen in searching among urinary tract tumor cases. Both MMS methods achieved higher-accuracy results for most urinary tract tumor cases at TM compared to single magnification counterparts. For instance, 10 and 4 percent F1-accuracy improvement in the single-vector method and the multi-vector method were observed for KIRP, respectively. These results indicate that employing low and medium magnification along with high magnification can improve the assessment of urinary tract tumors.

## 6.4 Findings and Discussion

This section first discusses the most important findings from the magnification-wise analysis of results. Next, Section 6.4.2 compares two proposed MMS methods. Finally, a discussion about the search dataset is provided in Section 6.4.3.

### 6.4.1 High Magnification for Diagnosis-based Evaluations

The results confirmed that the high magnification information played an essential role in diagnosis-based evaluations. The searching at 20× magnification or its combination with other magnifications outperformed searching at other magnification levels. Image search at 20× magnification, the combination of $10 \times 20\times$ magnifications, or TM reported the highest accuracy results in most anatomical sites. The general trend of successful searches is also visible among searching at the mentioned magnifications regardless of the searching methods. As can be seen in Fig. 6.3 and Fig. 6.4, both search methods, i.e., single-vector and multi-vector, achieved more than 80 percent classification accuracy at the mentioned magnifications in 7 out of 9 anatomical sites. The 20× magnification (both searching methods) acquired above 80 percent F1-score accuracy in the classification of 13 tumor subtypes. This value for $10 \times 20\times$ and TM is 12 and 13, respectively. Both MMS methods could achieve an F1-score above 65 percent at 20× and $10 \times 20\times$ in more than 20 out of 26 primary diagnosis types. It must be pointed out that the highest magnification is usually used by pathologists for confirming a diagnosis [12, 40, 61, 92].

### 6.4.2 Single-vector vs. Multi-vector Comparison

Using the single-vector method for the brain, liver/pancreas, and urinary tract tumor cases, TM achieved the highest overall classification accuracy (Fig. 6.3), and the highest F1-score in most subtypes (Table 6.3) in comparison with all other magnification settings. Results show that exploiting single-vector method at TM for representing WSIs can improve diagnostic assessments. Accordingly, if one of the anatomical sites mentioned earlier is a diagnostic consideration, evaluation based on multiple magnification levels using the single-vector method is highly recommended. Concerning the multi-vector method, 20× magnification has achieved slightly higher accuracy compared to the TM in many cases. This method at TriMagnification (TM) votes between retrieved WSIs of single-magnifications. As a potential solution to improve the accuracy at TM, a weighted majority voting approach may emphasize the 20× magnification. The suggested approach may be more successful with trainable weights. In comparison between methods at TM, the single-vector method tends to be more successful in combining all magnifications than the multi-vector method. That means the single-vector WSI representation performed more accurately than the multi-vector at TM in finding cases with the same tumor subtypes in most anatomical sites. The reported search accuracy associated with the brain, gynecological, liver/pancreaticobiliary, prostate/testis, and urinary tract tumors supports this finding. Another advantages of the single-vector method over the multi-vector method are less computational complexity, more efficient storage, and faster indexing and searching.

### 6.4.3 A Discussion on the Dataset

The results of this thesis study provide further support for the hypothesis that there is a direct relationship between the number of WSIs in the test dataset and the search accuracy [42]. ACC, USC, CHOL, UVM, and MESO are tumor subtypes with less than six samples in the test dataset. Due to the limited number of samples, the search framework was unsuccessful in finding two or more similar WSIs at some magnifications. Besides, top accuracy results are associated with frequent tumor types with a relatively high number of samples, i.e., BLCA, PRAD, and THCA, with 34, 40, and 51 samples, respectively. Many of the "primary diagnosis" labels obtained from the TCGA referred to groups of tumors with different morphologies (e.g., LGG, THCA, ESCA, CESC, OV, CHOL, SKCM, PRAD, and TGCT). It is possible that multiple magnifications improved the assessment of urinary tract tumors because only specific types of tumors were included in this anatomical site (i.e., BLCA, KICH, KIRC, and KIRP). However, this improvement was not seen for the specific types of tumors included in "Pulmonary".

This thesis performed multi-magnification content-based image searches on a diverse and multi-organ WSI repository. Results indicated that the search performance varies depending on anatomical site and disease type. In other words, MMS performed very accurately in some anatomical sites, such as prostate/testis, while it performed less accurately in some other tumor types, e.g., tumors of the gastrointestinal tract. This difference might be attributed to the fact that the histology and architectural complexity of tumors varies depending on the tumor types and diseases [61]. To put it simply, the diagnosis of some tumor types is more challenging than others, even for expert pathologists. This difference in search performance has also been observed in previous research studies [42, 43, 72].

# Chapter 7

# Summary and Conclusions

Multi-Magnification image search in digital pathology is a useful AI-enabled tool that can assist pathologists in many diagnostic, research and educational tasks and provide new insights into pathology data. Inspired by conventional pathology that analyzes glass slides at different magnification levels, this thesis proposed the function of adjustable magnification for a search engine and suggested using multi-magnification for content-based image search and retrieval in digital pathology. Based on the tumor type and the histology assessment, the magnification level can be adjusted in the proposed MMS framework to search more efficiently and accurately.

This thesis investigated three magnification levels and their combinations to search among unannotated (unlabelled) WSIs. MMS first represents the content of WSIs at different magnification levels using deep features and then measures the similarity between WSIs according to those representations. MMS works based on the deep feature vectors and uses KimiaNet, a deep CNN with DenseNet-121 network architecture, for extracting high-level histopathology information from patches. In this thesis, the last dense block of the KimiaNet was fine-tuned with 5× and 10× patches to learn features at low and medium magnifications more accurately. MMS is empowered with two independent search methods to measure similarities between WSIs. The first MMS method, single-vector representation, combines patch feature vectors by taking their median values to represent a WSI at a given magnification level. Subsequently, single vectors of all magnifications are concatenated to construct a single-magnification vector to represent each WSI. The second MMS method, multi-vector representation, measures the similarity between WSIs based on the median-of-minimum distance calculation of all patch feature vectors at any magnification. The deep feature vectors associated with patches are first converted to binary codes, i.e., barcodes, using the SNRQ hashing method to accelerate searching.

This thesis created a multi-magnification dataset of histopathology patches not only for MMS investigations but also for other deep learning-based studies in digital pathology. Since the manual annotation of gigapixel WSIs is not practical for large-scale datasets, this thesis proposed and used an automated way for patch extraction. Binary masks, generated by U-Net with MobileNet backbone, were used to detect tissue regions and remove non-informative background and artifacts regions from a WSI. In the next step of patch preparation, binary masks were split at various scales to find tissue patches at the selected magnification level. This approach was applied on a subset of the TCGA dataset with 13 anatomical sites (30 tumor subtypes) to create a multi-magnification dataset of large-size patches. The dataset was used for the magnification-oriented feature learning and evaluations in this thesis.

The experiments on MMS verified the capability of both MMS methods in finding WSIs with the same malignancy among WSIs of the same affected anatomical site. When search is evaluated as a classification task, MMS could achieve more than 80 percent accuracy in most of the anatomical sites. Significant accuracy improvements for diagnostic tasks were achieved concerning kidney tumors, e.g., F1-score of 0.93 for kidney renal papillary cell carcinoma (KIRP) using the TM search compared to 0.83 F1-score using the single magnification (20×) counterpart. The results confirmed that enriching high magnification search information with low magnification is a promising way to increase search accuracy in some cases. Also, the results showed that the highest magnification, containing the cell-level information and detailed morphological features, is an essential resolution for diagnostic tasks (the most accurate results were achieved using the 20× magnification or a combination of 20× magnification with other magnifications). These results are supported by the fact that pathologists often use high magnifications to confirm or rule out a diagnosis.

With tumor-oriented analysis of search results, this thesis proposed the best configuration of the MMS framework, in terms of method and magnification, to search for similar WSIs among WSIs of each anatomical site. These experimental results and analysis indicated that depending on the tumor type, the image search should use different magnifications and combinations of magnifications for more accurate and reliable search and retrieval. Therefore, magnification level selection is a valuable function of a digital pathology search platform.

## 7.1 Potential Areas for Future Research

Future work can include applying the proposed MMS methods in additional archives of WSIs in digital pathology. MMS framework does not require any manual annotations by

the pathologist and can easily be applied on raw databases of WSIs upon their availability. Without demanding any external information or annotation, the MMS can be verified on datasets of WSIs of various organs, and the search outputs can be evaluated by expert pathologist or based on corresponding metadata.

The data preparation procedure can be extended and improved to select patches more wisely. The binary mask can be extended to a multi-class mask by segmentation of various tissue types. The extended version of U-Net for tissue segmentation can potentially boost the MMS performance, and more generally, the accuracy of histopathology feature learning and representation.

The future direction of MMS might be driven by employing improved deep learning models for feature learning. Self-supervised and unsupervised learning approaches potentially can provided new insights into pathology data.

MMS has shown promising performance in diagnosis-oriented retrieval of WSIs. The performance of MMS can be further evaluated based on other morphological characteristics. Moreover, future research might investigate MMS for pan-cancer analysis to consider morphological similarities among tumors of different organs.

# References


[1] The cancer genoum atlas (tcga) dataset. https://portal.gdc.cancer.gov/. Accessed: November 2019.

[2] Shaimaa Al-Janabi, André Huisman, and Paul J Van Diest. Digital pathology: current status and future perspectives. *Histopathology*, 61(1):1–9, 2012.

[3] Raja S Alomari, Ron Allen, Bikash Sabata, and Vipin Chaudhary. Localization of tissues in high-resolution digital anatomic pathology images. In *Medical Imaging 2009: Computer-Aided Diagnosis*, volume 7260, page 726016. International Society for Optics and Photonics, 2009.

[4] Najah Alsubaie, Muhammad Shaban, David Snead, Ali Khurram, and Nasir Rajpoot. A multi-resolution deep learning framework for lung adenocarcinoma growth pattern classification. In *Annual Conference on Medical Image Understanding and Analysis*, pages 3–11. Springer, 2018.

[5] Morteza Babaie, Shivam Kalra, Aditya Sriram, Christopher Mitcheltree, Shujin Zhu, Amin Khatami, Shahryar Rahnamayan, and Hamid R Tizhoosh. Classification and retrieval of digital pathology scans: A new dataset. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 8–16, 2017.

[6] Morteza Babaie, Shivam Kalra, Aditya Sriram, Christopher Mitcheltree, Shujin Zhu, Amin Khatami, Shahryar Rahnamayan, and Hamid R. Tizhoosh. Classification and retrieval of digital pathology scans: A new dataset. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, July 2017.

[7] Péter Bándi, Rob van de Loo, Milad Intezar, Daan Geijs, Francesco Ciompi, Bram van Ginneken, Jeroen van der Laak, and Geert Litjens. Comparison of different methods for tissue segmentation in histopathological whole-slide images. In *2017*

*IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, pages 591–595. IEEE, 2017.

[8] Thomas W Bauer, Lynn Schoenfield, Renee J Slaw, Lisa Yerian, Zhiyuan Sun, and Walter H Henricks. Validation of whole slide imaging for primary diagnosis in surgical pathology. *Archives of pathology & laboratory medicine*, 137(4):518–524, 2013.

[9] Babak Ehteshami Bejnordi, Geert Litjens, Meyke Hermsen, Nico Karssemeijer, and Jeroen AWM van der Laak. A multi-scale superpixel classification approach to the detection of regions of interest in whole slide histopathology images. In *Medical Imaging 2015: Digital Pathology*, volume 9420, page 94200H. International Society for Optics and Photonics, 2015.

[10] AD Belsare and MM Mushrif. Histopathological image analysis using image processing techniques: An overview. *Signal & Image Processing*, 3(4):23, 2012.

[11] Aïcha BenTaieb and Ghassan Hamarneh. Predicting cancer with a recurrent visual attention model for histopathology images. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 129–137. Springer, 2018.

[12] Tad T Brunye, Patricia A Carney, Kimberly H Allison, Linda G Shapiro, Donald L Weaver, and Joann G Elmore. Eye movements as an index of pathologist visual expertise: a pilot study. *PloS one*, 9(8):e103447, 2014.

[13] Daniel Bug, Friedrich Feuerhake, and Dorit Merhof. Foreground extraction for histopathological whole slide imaging. In *Bildverarbeitung für die Medizin 2015*, pages 419–424. Springer, 2015.

[14] Wouter Bulten, Péter Bándi, Jeffrey Hoven, Rob van de Loo, Johannes Lotz, Nick Weiss, Jeroen van der Laak, Bram van Ginneken, Christina Hulsbergen-van de Kaa, and Geert Litjens. Epithelium segmentation using deep learning in h&e-stained prostate specimens with immunohistochemistry as reference standard. *Scientific reports*, 9(1):1–10, 2019.

[15] Wouter Bulten, Christina A Hulsbergen-van de Kaa, Jeroen van der Laak, Geert JS Litjens, et al. Automated segmentation of epithelial tissue in prostatectomy slides using deep learning. In *Medical Imaging 2018: Digital Pathology*, volume 10581, page 105810S. International Society for Optics and Photonics, 2018.

[16] Alcides Chaux and Antonio L Cubilla. Diagnostic problems in precancerous lesions and invasive carcinomas of the penis. In *Seminars in diagnostic pathology*, volume 29, pages 72–82. Elsevier, 2012.

[17] Pingjun Chen and Lin Yang. tissueloc: Whole slide digital pathology image tissue localization. *J. Open Source Software*, 4(33):1148, 2019.

[18] Lee Ad Cooper, Elizabeth G Demicco, Joel H Saltz, Reid T Powell, Arvind Rao, and Alexander J Lazar. Pancancer insights from the cancer genome atlas: the pathologist's perspective. *The Journal of pathology*, 244(5):512–524, 2018.

[19] Hao Dong, Guang Yang, Fangde Liu, Yuanhan Mo, and Yike Guo. Automatic brain tumor detection and segmentation using u-net based fully convolutional networks. In *annual conference on medical image understanding and analysis*, pages 506–517. Springer, 2017.

[20] Shiv Ram Dubey. A decade survey of content based image retrieval using deep learning. *IEEE Transactions on Circuits and Systems for Video Technology*, 2021.

[21] Hamed Erfankhah, Mehran Yazdi, Morteza Babaie, and Hamid R Tizhoosh. Heterogeneity-aware local binary patterns for retrieval of histopathology images. *IEEE Access*, 7:18354–18367, 2019.

[22] Anna Fabijańska. Segmentation of corneal endothelium images using a u-net-based convolutional neural network. *Artificial intelligence in medicine*, 88:1–13, 2018.

[23] PL Fitzgibbons, JL Connolly, S Bose, College of American Pathologists, et al. Protocol for the examination of resection specimens from patients with invasive carcinoma of the breast. *College of American Pathologists. Version: Breast Invasive Resection*, 4(0.0), 2020.

[24] Shereen Fouad, David Randell, Antony Galton, Hisham Mehanna, and Gabriel Landini. Unsupervised morphological segmentation of tissue compartments in histopathological images. *PloS one*, 12(11):e0188717, 2017.

[25] Yi Gao, William Liu, Shipra Arjun, Liangjia Zhu, Vadim Ratner, Tahsin Kurc, Joel Saltz, and Allen Tannenbaum. Multi-scale learning based segmentation of glands in digital colonrectal pathology images. In *Medical Imaging 2016: Digital Pathology*, volume 9791, page 97910M. International Society for Optics and Photonics, 2016.

[26] Baris Gecer, Selim Aksoy, Ezgi Mercan, Linda G Shapiro, Donald L Weaver, and Joann G Elmore. Detection and classification of cancer in whole slide breast histopathology images using deep convolutional networks. *Pattern recognition*, 84:345–356, 2018.

[27] Yunchao Gong, Svetlana Lazebnik, Albert Gordo, and Florent Perronnin. Iterative quantization: A procrustean approach to learning binary codes for large-scale image retrieval. *IEEE transactions on pattern analysis and machine intelligence*, 35(12):2916–2929, 2012.

[28] Feng Gu, Nikolay Burlutskiy, Mats Andersson, and Lena Kajland Wilén. Multi-resolution networks for semantic segmentation in whole slide images. In *Computational Pathology and Ophthalmic Medical Image Analysis*, pages 11–18. Springer, 2018.

[29] David A Gutman, Mohammed Khalilia, Sanghoon Lee, Michael Nalisnik, Zach Mullen, Jonathan Beezley, Deepak R Chittajallu, David Manthey, and Lee AD Cooper. The digital slide archive: A software platform for management, integration, and analysis of histology for cancer research. *Cancer research*, 77(21):e75–e78, 2017.

[30] Noriaki Hashimoto, Daisuke Fukushima, Ryoichi Koga, Yusuke Takagi, Kaho Ko, Kei Kohno, Masato Nakaguro, Shigeo Nakamura, Hidekata Hontani, and Ichiro Takeuchi. 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*, pages 3852–3861, 2020.

[31] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[32] Lei He, L Rodney Long, Sameer Antani, and George R Thoma. Histology image analysis for carcinoma detection and grading. *Computer methods and programs in biomedicine*, 107(3):538–556, 2012.

[33] Narayan Hegde, Jason D Hipp, Yun Liu, Michael Emmert-Buck, Emily Reif, Daniel Smilkov, Michael Terry, Carrie J Cai, Mahul B Amin, Craig H Mermel, et al. Similar image search for histopathology: Smily. *NPJ digital medicine*, 2(1):1–9, 2019.

[34] Sobhan Hemati, Shivam Kalra, Cameron Meaney, Morteza Babaie, Ali Ghodsi, and HR Tizhoosh. Cnn and deep sets for end-to-end whole slide image representation learning. *Proceedings of Machine Learning Research–Under Review*, 1:11, 2021.

[35] Sobhan Hemati, Mohammad Hadi Mehdizavareh, Shojaeddin Chenouri, and Hamid R Tizhoosh. A non-alternating graph hashing algorithm for large scale image search. *arXiv preprint arXiv:2012.13138*, 2020.

[36] David Joon Ho, Dig VK Yarlagadda, Timothy M D'Alfonso, Matthew G Hanna, Anne Grabenstetter, Peter Ntiamoah, Edi Brogi, Lee K Tan, and Thomas J Fuchs. Deep multi-magnification networks for multi-class breast cancer image segmentation. *Computerized Medical Imaging and Graphics*, 88:101866, 2021.

[37] Andrew G Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*, 2017.

[38] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger. Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4700–4708, 2017.

[39] R Inbaraj and G Ravi. A survey on recent trends in content based image retrieval system. *Journal of Critical Reviews*, 7(11):10–31838, 2020.

[40] Thomas Jaarsma, Halszka Jarodzka, Marius Nap, Jeroen JG van Merriënboer, and Henny PA Boshuizen. Expertise in clinical pathology: Combining the visual and cognitive perspective. *Advances in Health Sciences Education*, 20(4):1089–1106, 2015.

[41] Mayank Jain and Divakar Singh. A survey on cbir on the basis of different feature descriptor. *Journal of Advances in Mathematics and Computer Science*, pages 1–13, 2016.

[42] Shivam Kalra, Hamid R Tizhoosh, Sultaan Shah, Charles Choi, Savvas Damaskinos, Amir Safarpoor, Sobhan Shafiei, Morteza Babaie, Phedias Diamandis, Clinton JV Campbell, et al. Pan-cancer diagnostic consensus through searching archival histopathology images using artificial intelligence. *NPJ digital medicine*, 3(1):1–15, 2020.

[43] Shivam Kalra, HR Tizhoosh, Charles Choi, Sultaan Shah, Phedias Diamandis, Clinton JV Campbell, and Liron Pantanowitz. Yottixel–an image search engine for large

archives of histopathology whole slide images. *Medical Image Analysis*, 65:101757, 2020.

[44] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[45] Daisuke Komura, Keisuke Fukuta, Ken Tominaga, Akihiro Kawabe, Hirotomo Koda, Ryohei Suzuki, Hiroki Konishi, Toshikazu Umezaki, Tatsuya Harada, and Shumpei Ishikawa. Luigi: Large-scale histopathological image retrieval system using deep texture representations. *biorxiv*, page 345785, 2018.

[46] Daisuke Komura and Shumpei Ishikawa. Machine learning methods for histopathological image analysis. *Computational and structural biotechnology journal*, 16:34–42, 2018.

[47] Sai Chandra Kosaraju, Jie Hao, Hyun Min Koh, and Mingon Kang. Deep-hipo: Multi-scale receptive field deep learning for histopathological image analysis. *Methods*, 179:3–13, 2020.

[48] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25:1097–1105, 2012.

[49] Meghana Dinesh Kumar, Morteza Babaie, and Hamid R Tizhoosh. Deep barcodes for fast retrieval of histopathology scans. In *2018 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE, 2018.

[50] Meghana Dinesh Kumar, Morteza Babaie, Shujin Zhu, Shivam Kalra, and Hamid R Tizhoosh. A comparative study of cnn, bovw and lbp for classification of histopathological images. In *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 1–7. IEEE.

[51] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *nature*, 521(7553):436–444, 2015.

[52] Thomas M Lehmann, Marc Oliver Güld, Christian Thies, Benedikt Fischer, Klaus Spitzer, Daniel Keysers, Hermann Ney, Michael Kohnen, Henning Schubert, and Berthold B Wein. Content-based image retrieval in medical applications. *Methods of information in medicine*, 43(04):354–361, 2004.

[53] Ying Liu, Dengsheng Zhang, Guojun Lu, and Wei-Ying Ma. A survey of content-based image retrieval with high-level semantics. *Pattern recognition*, 40(1):262–282, 2007.

[54] Yun Liu, Krishna Gadepalli, Mohammad Norouzi, George E Dahl, Timo Kohlberger, Aleksey Boyko, Subhashini Venugopalan, Aleksei Timofeev, Philip Q Nelson, Greg S Corrado, et al. Detecting cancer metastases on gigapixel pathology images. *arXiv preprint arXiv:1703.02442*, 2017.

[55] L Rodney Long, Sameer Antani, Thomas M Deserno, and George R Thoma. Content-based image retrieval in medicine: retrospective assessment, state of the art, and future directions. *International Journal of Healthcare Information Systems and Informatics (IJHISI)*, 4(1):1–16, 2009.

[56] Dimitrios Markonis, Markus Holzer, Sebastian Dungs, Alejandro Vargas, Georg Langs, Sascha Kriewel, and Henning Müller. A survey on visual information search behavior and requirements of radiologists. *Methods of information in Medicine*, 51(06):539–548, 2012.

[57] Aarti Mathur, Matthew T Olson, and Martha A Zeiger. Follicular lesions of the thyroid. *Surgical Clinics*, 94(3):499–513, 2014.

[58] Michael T McCann, John A Ozolek, Carlos A Castro, Bahram Parvin, and Jelena Kovacevic. Automated histology analysis: Opportunities for signal processing. *IEEE Signal Processing Magazine*, 32(1):78–87, 2014.

[59] Neville Mehta, Alomari Raja'S, and Vipin Chaudhary. Content based sub-image retrieval system for high resolution pathology images using salient interest points. In *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 3719–3722. IEEE, 2009.

[60] Mrinal Mohit. *Automated histopathological analyses at scale*. PhD thesis, Massachusetts Institute of Technology, 2017.

[61] Diana Weedman Molavi. *The practice of surgical pathology: a beginner's guide to the diagnostic process*. Springer, 2017.

[62] Rodolfo Montironi, Liang Cheng, Alessia Cimadamore, Roberta Mazzucchelli, Marina Scarpelli, Matteo Santoni, Francesco Massari, and Antonio Lopez-Beltran. Narrative review of prostate cancer grading systems: will the gleason scores be replaced by the grade groups? *Translational Andrology and Urology*, 10(3):1530, 2021.

[63] Sanjay Mukhopadhyay, Michael D Feldman, Esther Abels, Raheela Ashfaq, Senda Beltaifa, Nicolas G Cacciabeve, Helen P Cathro, Liang Cheng, Kumarasen Cooper, Glenn E Dickey, et al. Whole slide imaging versus microscopy for primary diagnosis in surgical pathology: a multicenter blinded randomized noninferiority study of 1992 cases (pivotal study). *The American journal of surgical pathology*, 42(1):39, 2018.

[64] Henning Müller, Nicolas Michoux, David Bandon, and Antoine Geissbuhler. A review of content-based image retrieval systems in medical applications—clinical benefits and future directions. *International journal of medical informatics*, 73(1):1–23, 2004.

[65] Peter Naylor, Marick Laé, Fabien Reyal, and Thomas Walter. Segmentation of nuclei in histopathology images by deep regression of the distance map. *IEEE transactions on medical imaging*, 38(2):448–459, 2018.

[66] Muhammad Khalid Khan Niazi, Anil V Parwani, and Metin N Gurcan. Digital pathology and artificial intelligence. *The lancet oncology*, 20(5):e253–e261, 2019.

[67] Takayuki Otsu and Masatoshi Yoshida. Role of initiator-transfer agent-terminator (iniferter) in radical polymerizations: Polymer design by organic disulfides as iniferters. *Die Makromolekulare Chemie, Rapid Communications*, 3(2):127–132, 1982.

[68] Liron Pantanowitz, Paul N Valenstein, Andrew J Evans, Keith J Kaplan, John D Pfeifer, David C Wilbur, Laura C Collins, and Terence J Colgan. Review of the current state of whole slide imaging in pathology. *Journal of pathology informatics*, 2, 2011.

[69] David W Piston. Choosing objective lenses: the importance of numerical aperture and magnification in digital optical microscopy. *The Biological Bulletin*, 195(1):1–4, 1998.

[70] Xin Qi, Daihou Wang, Ivan Rodero, Javier Diaz-Montes, Rebekah H Gensure, Fuyong Xing, Hua Zhong, Lauri Goodell, Manish Parashar, David J Foran, et al. Content-based histopathology image retrieval using cometcloud. *BMC bioinformatics*, 15(1):1–17, 2014.

[71] EA Rakha, K Allison, IO Ellis, R Horii, S Masuda, F Penault-Llorca, et al. Invasive breast carcinoma: general overview. *Breast tumours WHO classification of tumours. 5th Ed. Lyon: World Health Organization*, pages 82–101, 2019.

[72] Abtin Riasatian, Morteza Babaie, Danial Maleki, Shivam Kalra, Mojtaba Valipour, Sobhan Hemati, Manit Zaveri, Amir Safarpoor, Sobhan Shafiei, Mehdi Afshari, et al.

Fine-tuning and training of densenet for histopathology image representation using tcga diagnostic slides. *Medical Image Analysis*, page 102032, 2021.

[73] Abtin Riasatian, Maral Rasoolijaberi, Morteza Babaei, and Hamid R Tizhoosh. A comparative study of u-net topologies for background removal in histopathology images. In *2020 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE, 2020.

[74] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[75] Raphael Rubin, David S Strayer, Emanuel Rubin, et al. *Rubin's pathology: clinicopathologic foundations of medicine*. Lippincott Williams & Wilkins, 2008.

[76] Nikhil A Sangle, Shari L Taylor, Mary J Emond, Michelle Depot, Bergein F Overholt, and Mary P Bronner. Overdiagnosis of high-grade dysplasia in barrett's esophagus: a multicenter, international study. *Modern Pathology*, 28(6):758–765, 2015.

[77] Roger Schaer, Sebastian Otálora, Oscar Jimenez-del Toro, Manfredo Atzori, and Henning Müller. Deep learning-based retrieval system for gigapixel histopathology cases and the open access literature. *Journal of pathology informatics*, 10, 2019.

[78] Tiffany L Sellaro, Robert Filkins, Chelsea Hoffman, Jeffrey L Fine, Jon Ho, Anil V Parwani, Liron Pantanowitz, and Michael Montalto. Relationship between magnification and resolution in digital pathology systems. *Journal of pathology informatics*, 4, 2013.

[79] Taimoor Shakeel Sheikh, Yonghee Lee, and Migyung Cho. Histopathological classification of breast cancer images using a multi-scale input and multi-feature network. *Cancers*, 12(8):2031, 2020.

[80] Fahime Sheikhzadeh, Rabab K Ward, Dirk van Niekerk, and Martial Guillaud. Automatic labeling of molecular biomarkers of immunohistochemistry images using fully convolutional networks. *PLoS One*, 13(1):e0190783, 2018.

[81] Milad Sikaroudi, Benyamin Ghojogh, Amir Safarpoor, Fakhri Karray, Mark Crowley, and Hamid R Tizhoosh. Offline versus online triplet mining based on extreme distances of histopathology patches. *arXiv preprint arXiv:2007.02200*, 2020.

[82] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[83] David RJ Snead, Yee-Wah Tsang, Aisha Meskiri, Peter K Kimani, Richard Crossman, Nasir M Rajpoot, Elaine Blessing, Klaus Chen, Kishore Gopalakrishnan, Paul Matthews, et al. Validation of digital pathology imaging for primary histopathological diagnosis. *Histopathology*, 68(7):1063–1072, 2016.

[84] Youyi Song, Ling Zhang, Siping Chen, Dong Ni, Baiying Lei, and Tianfu Wang. Accurate segmentation of cervical cytoplasm and nuclei based on multiscale convolutional network and graph partitioning. *IEEE Transactions on Biomedical Engineering*, 62(10):2421–2433, 2015.

[85] Akshay Sridhar, Scott Doyle, and Anant Madabhushi. Content-based image retrieval of digitized histopathology in boosted spectrally embedded spaces. *Journal of pathology informatics*, 6, 2015.

[86] Chetan L Srinidhi, Ozan Ciga, and Anne L Martel. Deep neural network models for computational histopathology: A survey. *Medical Image Analysis*, page 101813, 2020.

[87] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1–9, 2015.

[88] Mingxing Tan and Quoc V Le. Efficientnet: Rethinking model scaling for convolutional neural networks. *arXiv preprint arXiv:1905.11946*, 2019.

[89] Hamid Reza Tizhoosh and Liron Pantanowitz. Artificial intelligence and digital pathology: challenges and opportunities. *Journal of pathology informatics*, 9, 2018.

[90] Hiroki Tokunaga, Yuki Teramoto, Akihiko Yoshizawa, and Ryoma Bise. Adaptive weighting multi-field-of-view cnn for semantic segmentation in pathology. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 12597–12606, 2019.

[91] Katarzyna Tomczak, Patrycja Czerwińska, and Maciej Wiznerowicz. The cancer genome atlas (tcga): an immeasurable source of knowledge. *Contemporary oncology*, 19(1A):A68, 2015.

[92] Darren Treanor, Chee Hooi Lim, Derek Magee, Andy Bulpitt, and Phil Quirke. Tracking with virtual slides: a tool to study diagnostic error in histopathology. *Histopathology*, 55(1):37–45, 2009.

[93] Nelson Zange Tsaku, Sai Chandra Kosaraju, Tasmia Aqila, Mohammad Masum, Dae Hyun Song, Ananda M. Mondal, Hyun Min Koh, and Mingon Kang. Texture-based deep learning for effective histopathological cancer image classification. In *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 973–977, 2019.

[94] Laurens Van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of machine learning research*, 9(11), 2008.

[95] Mart van Rijthoven, Maschenka Balkenhol, Karina Siliņa, Jeroen van der Laak, and Francesco Ciompi. Hooknet: Multi-resolution convolutional neural networks for semantic segmentation in histopathology whole-slide images. *Medical Image Analysis*, 68:101890, 2021.

[96] Ji Wan, Dayong Wang, Steven Chu Hong Hoi, Pengcheng Wu, Jianke Zhu, Yongdong Zhang, and Jintao Li. Deep learning for content-based image retrieval: A comprehensive study. In *Proceedings of the 22nd ACM international conference on Multimedia*, pages 157–166, 2014.

[97] James Z Wang. Pathfinder: multiresolution region-based searching of pathology images using irm. In *Proceedings of the AMIA Symposium*, page 883. American Medical Informatics Association, 2000.

[98] James Z Wang, John Nguyen, Kin-koi Lo, Christine Law, and Donald Regula. Multiresolution browsing of pathology images using wavelets. In *Proceedings of the AMIA Symposium*, page 430. American Medical Informatics Association, 1999.

[99] John N Weinstein, Eric A Collisson, Gordon B Mills, Kenna R Mills Shaw, Brad A Ozenberger, Kyle Ellrott, Ilya Shmulevich, Chris Sander, and Joshua M Stuart. The cancer genome atlas pan-cancer analysis project. *Nature genetics*, 45(10):1113–1120, 2013.

[100] Rune Wetteland, Kjersti Engan, Trygve Eftestøl, Vebjørn Kvikstad, and Emilius AM Janssen. Multiscale deep neural networks for multiclass tissue classification of histological whole-slide images. *arXiv preprint arXiv:1909.01178*, 2019.

[101] Pavel Yakubovskiy. Segmentation models. https://github.com/qubvel/segmentation_models, 2019.

[102] Lei Zheng, Arthur W Wetzel, John Gilbertson, and Michael J Becich. Design and analysis of a content-based pathology image retrieval system. *IEEE transactions on information technology in biomedicine*, 7(4):249–255, 2003.

[103] Yushan Zheng, Zhiguo Jiang, Haopeng Zhang, Fengying Xie, Yibing Ma, Huaqiang Shi, and Yu Zhao. Histopathological whole slide image analysis using context-based cbir. *IEEE transactions on medical imaging*, 37(7):1641–1652, 2018.

[104] Konstantinos Zormpas-Petridis, Henrik Failmezger, Ioannis Roxanis, Matthew Blackledge, Yann Jamin, and Yinyin Yuan. Capturing global spatial context for accurate cell classification in skin cancer histology. In *Computational Pathology and Ophthalmic Medical Image Analysis*, pages 52–60. Springer, 2018.