Tissue Pattern Detection in Whole Slide Images Using *You-Only-Look-Once* Approach

by

Kimia Hemmatirad

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, 2022

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 paper, and report that I have authored or co-authored.

[1] Kimia Hemmatirad, Morteza Babaei, Mehdi Afshari, Danial Maleki, Mahjabin Sajadi, and H.R. Tizhoosh, **Quality Control of Whole Slide Images using the YOLO Concept**, Accepted to be presented at ICHI 2022, 10th IEEE International Conference on Healthcare Informatics.

[2] Areej Alsaafin, Kimia Hemmatirad, **Nucleus/Cell Segmentation in Digital Pathology: A Survey**, SYD 675 Course Project.

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 in writing course project [2]. Some parts of course project [2] are incorporated in Chapter 2.

Abstract

Analyzing digital pathology images is required for diagnostic conclusions by investigating tissue patterns and morphology. However, because of the large size of the whole slide images, manual evaluation can be time-consuming, expensive, and prone to inter- and intra-observer variability. Therefore, automated tissue structure detection and segmentation approaches have lately been intensively investigated in the field of digital pathology.

Generating a pixel level object annotation for histopathology images is expensive, time-consuming, and hard to achieve. In addition, in many applications precise object boundary is not needed. As a result, defection models with bounding box labels may be a smart solution. In this thesis, different techniques on tissue pattern detection, and segmentation in whole slide images have been explored. Specifically, YOLO-v4 (You-Only-Look-Once), a real-time object detector for microscopic images has been studied. YOLO uses a single neural network to predict several bounding boxes and class probabilities for the objects of interest. YOLO enhances detection performance by training on whole slide images. YOLO-v4 has been used in this thesis for two different applications to find specific tissue patterns in whole slide images. Comparisons with the segmentation techniques on the same dataset have been conducted as well.

The first application of tissue pattern recognition in this thesis is quality control of whole slide images. That includes detecting air bubble edges, tissue folds, which happens during glass slide preparation, and the presence of ink-markers on tissue glass slides, manually drawn by pathologists to highlight regions of interest on glass slides. In order to avoid rejecting a whole slide due to presence of artifacts and ink-markers, there are various approaches for detecting and eventually removing these artifacts. However traditional approaches to remove artifacts are mostly based on thresholding techniques combined with some mathematical morphology operations. In this thesis, YOLO-v4 has been employed for this purpose. The experiments showed 99.5% IOU calculation (intersection over union, also called Jaccard Index) for locating artifacts.

The second application of tissue pattern recognition in this thesis is glomeruli detection in kidney images. Glomeruli are groups of capillaries that help the body filter waste and extra fluids. In this application, YOLO-v4 has been trained to detect these patterns in kidney images. Multiple experiments have been designed and conducted based on different training data of two public datasets and also a private dataset from University of Michigan for fine-tuning the model, and tested on the private dataset from University of Michigan as an external validation on two different stained tissues, namely periodic acid–Schiff (PAS), and hematoxylin and eosin (H&E) stains. The results and the average specificity and sensitivity for all experiments along with the comparison of existing segmentation methods on the same datasets have been discussed in the result section.

Acknowledgements

First and foremost, I would like to express my deepest appreciation to my supervisors, Professor Hamid R. Tizhoosh, and Dr. Morteza Babaei who have been a great support and inspiration to me during my graduate studies.

I would also like to thank both of my committee members, Professor Michailovich, and Professor Ponnambalam for reviewing my thesis, and useful feedbacks.

Last but not least, I would like to thank my parents, my friends, Sahar and Rasoul, and my brother, Taha, for their encouragement, unconditional love and support, and for always believing in me.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

YOLO	You Only Look Once
AI	Artificial Intelligence
CNN	Convolutional Neural Network
DPM	Dimension-based Partitioning and Merging Clustering
SIFT	Scale-Invariant Feature Transform
HOG	Histogram of Oriented Gradients
WSI	Whole Slide Image
IOU	Intersection Over Union
PAS	Periodic Acid–Schiff
H&E	Hematoxylin and Eosin
CAD	Computer-aided diagnosis
DT	Distance transform
EDT	Euclidean DT
MO	Morphology operation
UE	Ultimate erosion
HIT	H-minima transform
HAT	H-maxima transform
TPR	True positive rate
LoG	Laplacian of Gaussian
MESR	Maximally stable extremal region
HT	Hough transform

Chapter 1

Introduction

For investigations of tissue morphology and, as a result, for making diagnostic conclusions, computational pathology approaches are fast and reliable compare to conventional microscopy-based workflows. On the other hand, any manual evaluation of tissue samples can be time-consuming, costly, and subject to both inter- and intra-observer variability because of the size and the complexity of the whole slide images (WSIs) [168]. Consequently, digital pathology researchers have recently focused their attention on automated ways to detect and segment tissue structures. Many quantitative studies, such as determining tissue types, rely on the accuracy of these tissue pattern segmentation, which is regarded as the foundation of automated image analysis. However, due to the complex tissue architectures in which tissue structures are frequently clustered into types, glands and organelles overlapping with each other, establishing precise segmentation is not a simple operation. This makes distinguishing these patterns from the background and mainly from each other a challenge. In addition, histopathological images typically contain noise and artifacts created during image acquisition and low contrast between foreground and background [168].

Segmentation models have been widely used in digital pathology to segment lesions, cells, and other regions of interest [71]. However, these models need pixel-level object annotations by the expert. Detailed labels (pixel-level) for histopathology images are expensive, time-consuming, and hard to achieve [81]. Moreover, in some of the applications in histopathology, only detecting the position of the specific tissue pattern without precisely outlining the borders may be required [168]. These techniques are called *tissue pattern detection*, and are usually faster compare to segmentation methods. The main advantage of detection models is that they construct a bounding box around the tissue of interest rather than pixel-level labelling, making its training much more convenient.

In this thesis, different techniques for detection and segmentation in digital pathology and the applications for each of these methods, accuracy, and performance on various datasets have been summarized.

Deep object detectors typically consist of two parts: a backbone trained on ImageNet and a head used to forecast object classes and bounding boxes. One-stage object detectors

1

and two-stage object detectors are the most common head types [18]. Regions with convolutional neural networks (R-CNN) [50] series, is a good example of a two-stage object detection category, and YOLO (you-only-look-once) is one of the examples for one-stage object detectors [18] which has been studied and explored on two different applications for detecting specific tissue patterns in this thesis.

YOLO is a simple concept with several advantages. To begin with, YOLO is very fast as it does not require a complicated pipeline for a regression problem. Furthermore, the mean average accuracy of YOLO is higher than that of comparable real-time systems. Therefore, the network can be considered as a real-time object detector. Second, with YOLO, context information about object classes is encoded as well as their appearance during training and testing, unlike sliding window and region proposal-based approaches [126]. A popular object recognition approach, Fast R-CNN [49], misidentifies background patches in an image as objects. In comparison to Fast R-CNN, YOLO creates half the number of background errors [126]. And third, YOLO learns to represent objects in a universally applicable way. YOLO surpasses the best detection algorithms like deformable parts models (DPM) and R-CNN when trained on natural images and evaluated on art. YOLO is less likely to fail when applied to new domains or unexpected inputs because of its high degree of generalizability [126].

In this thesis, YOLO-v4 has been employed in two different applications to find particular tissue patterns in WSIs. Comparisons on the same datasets, with segmentation approaches, will be performed.

1.1   Applications

1.1.1   Quality Control of Whole Slide Images

Computational pathology applies computer vision algorithms on WSIs. The digitization of tissue glass slides marks a significant change in the clinical diagnostic workflow [74]. One of the challenges in digital pathology is the presence of artifacts such as tissue fold, air bubbles, and ink-markers on archived cases. These artifacts may affect the focus points in digital scanners, and their presence may negatively affect the quality of the quality of tissue image and the subsequent diagnosis [74]. Manual review of WSIs requires experts, which is a laborious and time-consuming task. Traditional object detection systems often start by extracting features from input images using methods such as Haar tranform [114], scale-invariant feature transform (SIFT) [93], and histogram of oriented gradients (HOG) [35]. They often employ handcrafted morphology-based features. These approaches mostly fail to generalize and may not be precise enough for given applications. For example, HistoQC [74], a tool for WSI (whole slide image) acquisition quality control, can detect artifacts and output outliers using thresholding techniques mixed with mathematical morphological procedures. However, it may fail to deliver an accurate solution in many cases.

As the first application using detection models in this thesis, the YOLO-v4 model has been trained on 92 WSIs for training to detect air bubble edges, tissue folds, which can

happen during slide preparation, and ink-marked tissue glass slides, drawn by pathologists to highlight regions of interest on glass slides [74]. Based on the results that have been achieved on a dataset consisting of 15 WSIs, the method is not only fast but also highly accurate. The experiments showed 99.5% IOU (intersection over union, also called Jaccard Index) for locating artifacts. Figure 1.1 shows sample cases for tissue fold, air bubbles, and ink-markers on tissue in the training dataset.

Figure 1.1: Sample images in the dataset containing ink-markers, air bubbles, and tissue fold, along with the annotated bounding boxes.

1.1.2   Glomeruli Detection in Kidney Images

The histological evaluation of glomeruli is critical for identifying whether a kidney is transplantable [100]. The Karpinski score, which includes the ratio of sclerotic glomeruli to total glomeruli in a kidney segment, is critical for determining the necessity for a single or dual kidney transplant [5]. Immunopathology, clinical symptoms, etymology, and morphological abnormalities are all factors that go into classifying glomeruli disorders. To classify the glomeruli diseases, these anatomic structures need to be detected. Automated glomeruli identification frameworks for kidney biopsies conducted by pathologists can be quite helpful because manual examination of kidney samples is time-consuming and error-prone [5, 100]. It is crucial to identify glomeruli in the kidney, such that based on the features of this glomerulosclerosis, different diseases could be determined.

There are several segmentation methods to detect glomeruli in kidney images [21]. However, these methods require pixel-level annotation of the images. Labelling histopathological images at the pixel level is difficult, costly, and time-consuming [147]. In detection methods, only determining the location of a given tissue pattern, a glomerulus, is required without the need to precisely delineating its borders.

In the field of histopathology, the lack of image data, annotation, and labels has always been a problem [147]. Hence, it is very important to validate deep networks on their generalization capability. By training a network with public datasets, and then fine-tuning it with only limited number of data from the specific hospital or specific resource, we may be able to significantly improve the accuracy of the network on the validation set from the same resource.

In the second application, YOLO-v4 as a detection network has been trained to recognize all glumeruli in a given kidney image. Multiple experiments were designed and carried

out based on different training data from two public datasets to fine-tuning the model, and tested on the private dataset from the University of Michigan as an external validation on two differently stained tissues, namely periodic acid–Schiff (PAS) staining and hematoxylin and eosin (H&E) staining.

The first dataset is a collection of 31 SVS WSIs made available to the public. The annotation of the bounding boxes of these 31 WSIs has been performed by collaborating pathologists. This data is part of the WSI datasets generated in the European project AIDPATH [1]. The second dataset, has been used for HubMap competition [2]. TIFF files ranging in size from 500MB to 5GB make up the dataset containing 8 WSIs for the training, and 5 WSIS for testing. The segmentation annotation was provided for each of the WSIs in this competition. The generalization of the network has been tested by training on these two public datasets, followed by the external validation on the private dataset from University of Michigan. Another dataset that has been used for training and fine-tuning the models, is 7 PAS stained WSIs which is not a public dataset and has been collected from the University of Michigan annotated by expert pathologists. In Figure 1.2 three samples of the training dataset for the network has been shown. On the left, there is a sample from the first public dataset, the middle image is a sample from the second public dataset, and on the right there is an example of the private dataset from University of Michigan.
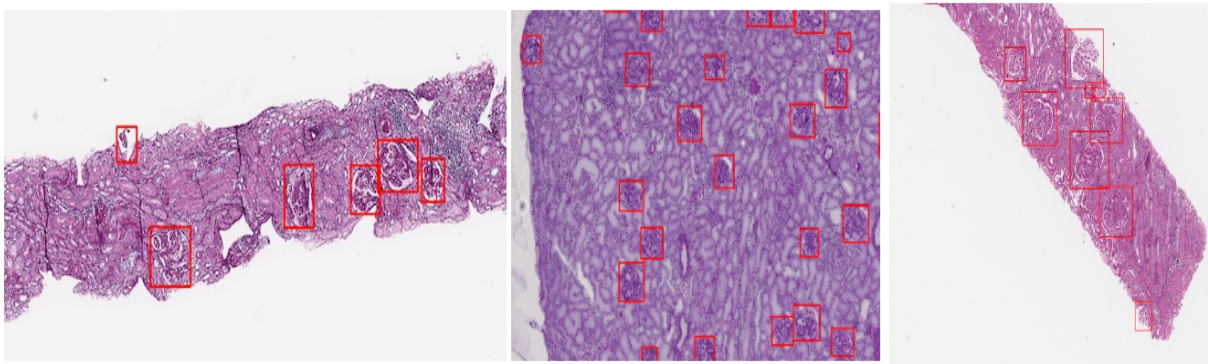
Figure 1.2: Three samples of the training dataset. On the left, there is a sample from the first public dataset, the middle is a sample from the second public dataset, and on the right there is an example from the private dataset from University of Michigan.

Three datasets, two public and one private, have been used to design and conduct 14 experiments. These experiments, have been trained on different combinations of the available datasets combing public datasets, and the private dataset, and all experiments have been validated on the private dataset from the University of Michigan with two staining, 20 PAS stained WSIs and 16 H&E stained WSIs. YOLO was serving as the detector network which will be described in details in the following sections.

In Figure 1.3, two samples of the private validation dataset along with the annotated bounding boxes have been shown. On the top is a sample tissue of H&E stained WSI,

---

[1] http://aidpath.eu/

[2] https://www.kaggle.com/c/hubmap-kidney-segmentation/overview

and on the bottom is a sample tissue of PAS stained WSI. The results, average specificity and sensitivity for all experiments, and comparison of existing segmentation methods on the same datasets have been discussed in the result section. In general, one could observe that the average specificity and sensitivity are higher on the PAS validation set because all of the images in the training dataset are PAS stained. Also, there is an improvement in average specificity and sensitivity while fine-tuning the network with only 7 PAS WSIs from the University of Michigan.
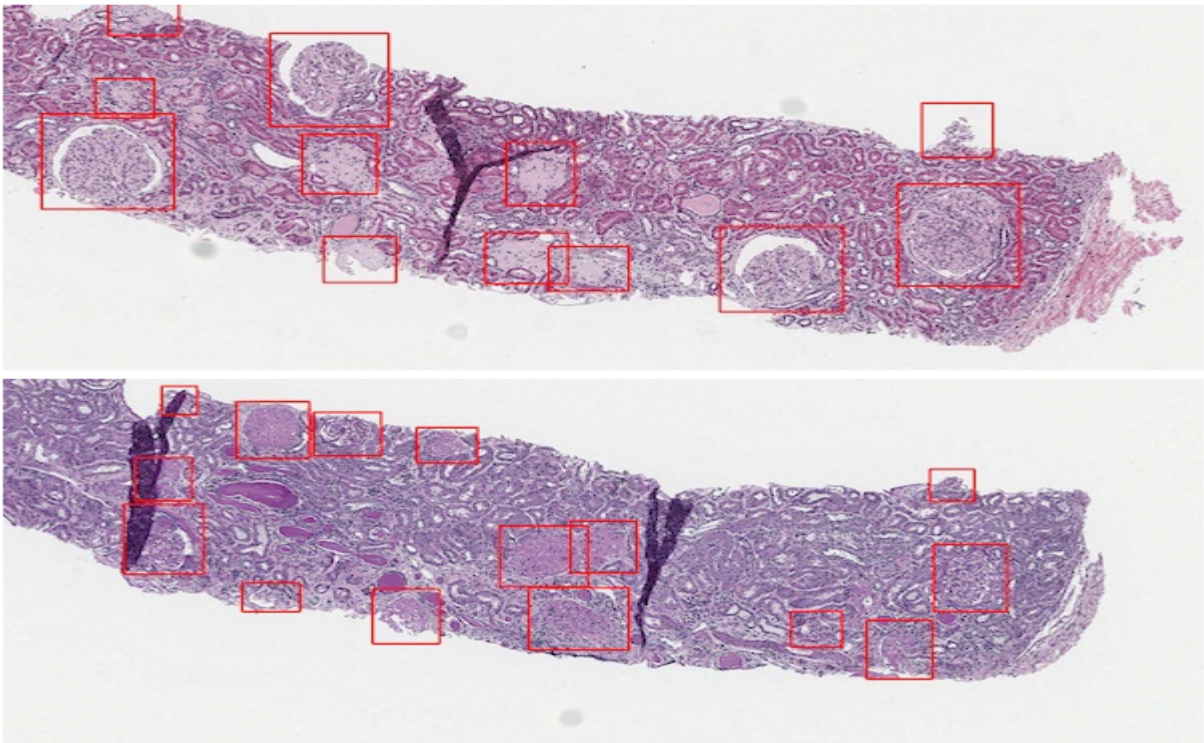
Figure 1.3: Two samples of the private validation dataset along with the annotated bounding boxes. On the top is a WSI with H&E-stained tissue, and the bottom shows a WSI with PAS-stained tissue.

1.2   Thesis Organization

The remainder of this thesis is structured as follows. Chapter 2 provides background information on different techniques for detection and segmentation in digital pathology and related research papers, along with the comparison between these different methods; Next, this chapter reviews the YOLO architecture network. After that, two different applications for detecting specific tissue patterns will be reviewed. The method, datasets attributes, and various designed experiments for the two different applications have been discussed in Chapter 3. Chapter 4 covers the results obtained for the experiments based on different performance metrics. Finally, the summary, conclusions and future directions are stated in Chapter 5.

Chapter 2

Literature Review

Digital pathology can help disease diagnoses by providing WSIs for image analysis to assis the diagnostic process [168]. Traditional microscopic diagnosis requires pathologists to physically examine glass slides under the microscope, a labor-intensive process that does not provide a digital workflow for assigning and sharing cases [12]. Machine learning advances have shown considerable potential in lowering pathologists' workload and assisting with diagnosis and treatment planning. Since pathologists mostly write the diagnostic report based on tissue morphology, tissue pattern identification and segmentation are crucial tasks in representing morphological information [57]. For example, the quality and accuracy of nuclei segmentation may be a key quantification in cancer diagnosis and grading [61, 57, 168]. In addition, for image classification tasks, one may extract features from a histology image based on the number of cells or other histology patterns [105, 168]. As a result, it is important to correctly detect and segment tissue patterns in pathology images.

In the literature, there are various reviews on automated pathology image analysis, generally called *computational pathology* [168]. Madabhushi and Lee [98] highlighted the advancement of computational image analysis approaches utilized in digital pathology for detection, segmentation, feature extraction, and tissue classification. Hamilton et al. [58] examined several image analysis applications, such as nuclear morphometry and tissue architecture analysis. Fuchs and Buhmann [43] defined a workflow of digital pathology as three main steps: data and ground-truth generation; image analysis including object detection, segmentation, and recognition; and medical statistics. Xing and Yang [168] reviewed automated nucleus/cell identification and segmentation algorithms using digital pathology and microscopy images.

In this section, the main categories of object detection and segmentation methods in digital pathology will be discussed based on [168], as well as the advantages, problems, and limitations of these methods. In general, **segmentation** refers to methods for identifying objects with specified boundaries, primarily at the pixel level, which can be useful for specifying characteristics or even grading illnesses in histopathological images. **Detection** approaches, on the other hand, are seen as acquiring the object position without precisely outlining its boundaries. Detailed (pixel-level) labeling for histopathological images is costly, time-consuming, and difficult. As detection does not need the object boundaries,

detection methods are easier to train compared to segmentation methods.

2.1   Segmentation Methods in Digital Pathology

Segmentation refers to the methods that identify the objects with specific boundaries mainly in pixel level, which can be helpful for critical tasks such as grading diseases in histopathology images [168]. Veta et al. [157] categorized the cell/nuclei segmentation methods into two main categories: 1) some of these methods are designed to detect each nuclei before performing any segmentation technique, and 2) some other segmentation methods first segment parts/chunks of nuclei groups from the rest of the tissue, and then separate those into individual nuclei. Some of the conventional methods like Hough Transform [31], and center identifying methods to pinpoint the center of the nuclei with voting along the image gradient [120, 158], have been used in both mentioned different approaches for nuclei segmentation. Although there is a large number of different segmentation methods, they are mostly based on a few underlying algorithms [168]: intensity thresholding, morphology operation, watershed transform, deformable models, clustering, graph-based methods, and supervised classification. This discussion covers the most common nuclei/cell segmentation that have been commonly mentioned in the literature.

2.1.1   Image Thresholding

Thresholding methods could be the first and most straightforward method for tissue segmentation. Because of the difference between the intensity of cells and the background, there is a threshold that could separate the image into two segments [53]. This threshold could be a global threshold, or it could be a locally adaptive threshold. The global techniques determine the threshold based on the histogram of the entire image, whereas the local approach considers the neighbourhood attributes to derive an adaptive threshold value. To prevent losing important information during binarization, thresholding necessitates high image contrast. This makes thresholding-based approaches more efficient when used for clear region segmentation rather than delicate object extraction such as cell nucleus extraction [168].

The Otsu approach [112] is a classical approach to finding an optimal threshold by maximizing the variance between the statistical features of the image intensity histogram. Otsu thresholding has been utilized by many researchers to segment nuclei in histopathology images [86, 171, 163] . Before identifying cells, Suryani et al. [148] employed multi Otsu thresholding to detect blast cell nuclei in Acute Myeloid Leukemia (AML) images. Using several threshold settings, multi Otsu thresholding separates the image into several segments. In contrast to single Otsu thresholding, multi Otsu thresholding allows users to tailor the threshold value to the brightness level of each image. Multiple other applications use global thresholding, or local thresholding, which requires additional steps for managing pixel neighbourhoods [25, 80, 111, 82].

2.1.2 Mathematical Morphology Operations

Mathematical morphology uses the object shape to perform a variety of operations [138, 72]. Some segmentation methods use mathematical morphology theory to extract items with a preset form, such as a circular disk, square, or cross [145]. Such operations are separated into two main steps: topdown erosion and bottom-up dilation. Basic morphological processes such as erosion, dilation, opening, and closing are employed to produce further operations [168]. Erosion and dilation, for example, can be used to enlarge the borders of foreground pixels. Also, morphology operation can be used as a step in image preprocessing to enhance the image segmentation result [37, 72].

2.1.3 Watershed Transform

The wateshed method is one of the region accumulation segmentation methods [165]. It works by adding connected points, called catchment basin to the first selected base points (markers) in the image, until it forms various labelled regions [106]. It has been used for many previous nuclei/cell segmentation tasks [75, 1, 26, 109, 152, 8, 30]. The algorithm comes from the field of mathematical morphology, which is concerned with the topological representation of images [160, 134]. The regional minimum relates to the collection of pixels having the lowest regional elevation. The intensity minima are groups of connected pixels with grey levels that are strictly lower than their local neighbours. The watershed transform has several advantages, including its speed, straightforward, and intuitive implementation. However, it has drawbacks such as over-segmentation and noise sensitivity [55]. When the foreground and background can be clearly distinguished, segmentation-based watershed modification is efficient. Otherwise, due to the intensity differences in the foreground and background, it is prone to over-segmentation. Watershed is difficult to utilize to segment nuclei from the cluttered surroundings of histopathology pictures due to over-segmentation. There are some approaches to overcome this problem; for example, some researchers have used watershed in conjunction with deep neural networks [83, 40, 167, 108].

2.1.4 Active Contour Models (ACMs)

ACMs or deformable models, use gradient information to describe the contour of objects in an image by minimizing an energy function [36]. These models offer a good balance of efficiency and flexibility [36]. Geodesic and parametric snakes are the two primary implementations of deformable models for nucleus or cell segmentation, using implicit and explicit contour representations, respectively. They have been used in many existing nuclei/cell segmentation techniques [20, 32, 169]. Deformable models have suffered from convergence challenges when non-linearities and noise inhibit boundary recognition [168].

2.1.5 Clustering

Clustering is another technique used for nucleus and cell segmentation in microscopy images. It aims to partition an image into a number of disjoint groups (or clusters) by increasing inter-cluster similarity so that objects within the same cluster are more similar to each other than objects from different clusters [60]. There are several similarity (or dissimilarity) measures that may be used to allocate items to clusters. The most common unit of measurement is the Euclidean distance [168]. Clustering may not create the final segmented objects in nucleus/cell segmentation. Instead, it is mostly utilized to assist with object boundary extraction. Cluster centers, on the other hand, maybe utilized to identify and localize cell nuclei because each cluster represents one cell. Using the k-means technique [77], which is a fundamental iterative clustering algorithm, one can divide an image into k groups by associating each data point with the mean of the adjacent cluster. K-means has been utilized as a part of a variety of strategies for nucleus/cell segmentation that have been proposed in the literature. For example, Bhattacharjee et al. [16] presented a color segmentation approach based on a mix of k-means clustering and the watershed algorithm to differentiate contacting cell nuclei in histopathology images using a color segmentation method. According to [10], a similar method was used to separate nuclei from breast histopathology images before categorizing them into benign and malignant categories. In [56] the authors conducted nuclei detection by employing mathematical morphology and k-means clustering. Qu et al. [122] developed a unique technique based on k-means and the Voronoi diagram to solve the problem of insufficient nucleus annotation during the training phase of a neural network. The Fuzzy c-means (FCM) clustering method is another technique for segmentation that has been used in many previous research works [28, 85, 41, 151, 161]. Unlike *hard* clustering), which assigns each item to only one cluster, FCM is a *soft* clustering technique that enables an object to be linked with more than one cluster. FCM quantifies the strength of the relationship by measuring the membership degree of the association. When compared to the k-means method, FCM may provide an improved representation for objects that cannot be entirely allocated to a single cluster. As a consequence, soft clustering may be more resilient to ambiguity.

2.1.6 Convolutional Neural Networks (CNNs)

CNNs have been recently proposed for a variety of applications, including the identification and segmentation of nuclei/cells [154]. Compared to thresholding and other conventional approaches, deep CNNs are generally more successful. Nevertheless, their success is strongly dependent on the quantity and quality of the training data [168]. Deep learning approaches are capable of automatically recognizing, segmenting, and categorizing nucleus/cells for a variety of applications by learning characteristics in the input images. For segmenting nuclei, many works adopted the most prevalent neural network method, namely a CNN, for handling two-dimensional input [11, 38, 121, 119, 47]. **U-Net** [135] is a CNN which was proposed in 2015 for the first time. It has shown remarkable results in various segmentation challenges [23, 109]. U-Net has a U-shape architecture, which contains two symmetric components, called encoder and decoder [135, 124]. A CNN-like sub-network

called the encoder collects high-level information from the input image before sending it to the main network. There is a segmentation map created by the decoder sub-network (sometimes referred to as the expansion part). The U-Net architecture has been illustrated in Figure 2.1.
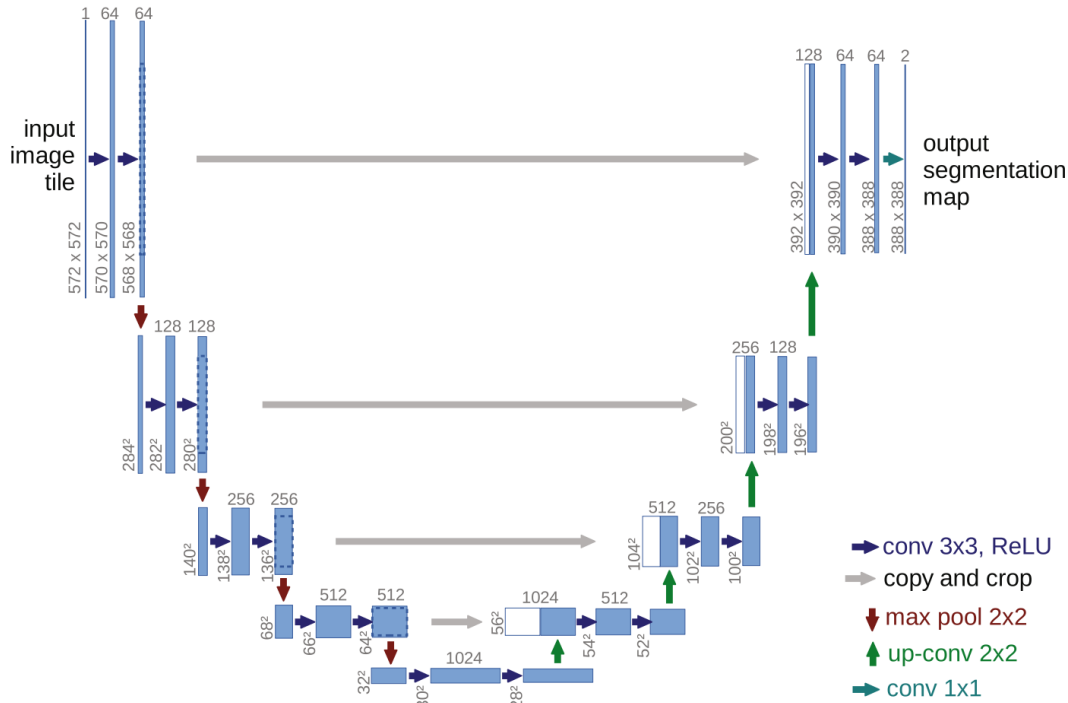
Figure 2.1: U-Net architecture. It has a U-shape architecture, containing two symmetric components, called encoder, and decoder. This image has been taken from [135].

2.2 Detection Methods in Digital Pathology

Detection approaches determine the object's position without defining its exact boundaries. There are generally markers or seeds found at the object centroids, one for each nucleus or cell. This section will cover nucleus/cell detection methods as getting the position of the item without precisely outlining its borders. An object's marker might be a single point or a limited area within the item. Clearly, marker identification can assist in the counting, tracking, and segmentation of nuclei and cells. Different types of histopathology images have a wide range of image characteristics that necessitate the use of different algorithms for nucleus/cell detection. However, the cell/nucleus detection methods can be categorized into conventional and learning-based approaches [168]. In conventional methods these groups will be discussed: distance transform, morphology operation, HIT/HAT, LoG filtering, MSER detection, HT, radial symmetry-based voting. In the learning-based detection, single and multi-stage deep neural networks in the literature will be reviewed.

2.2.1 Conventional Detection Algorithms in Digital Pathology

Distance Transform (DT)

DT approaches compute the exact Euclidean distance transform for each binary pixel and feature points to assign these pixels to the nearest feature point [104]. The edge pixels/voxels in a binary image are often used as feature points in nucleus/cell identification, whereas the Euclidean distance is used as the metric. Therefore, it is desirable that the local maxima in the created distance map match to the nuclei or cells' centroids. Over-segmentation may be avoided by using euclidean distance transform (EDT) in conjunction with watershed segmentation since the inverse distance map removes many unwanted local minima that would otherwise lead to a major oversegmentation problem [168]. Although this approach cannot detect overlapping cells, many applications are using this method for histopathological images [3, 170, 2, 59].

Mathematical Morphology

Morphological operations can provide filtering techniques for processing images using a certain *structuring element*, SE. Circular disk, square, cross, or other shapes can be to design suitable SEs. [145]. An object's geometrical and topological structure is examined to conduct image filtering. It is possible to perform fundamental morphological operations such as border extraction and filling holes using shift-invariant operators such as dilation, opening, and closing [145]. There are several ways to apply the binary mathematical morphology to grayscale images [165, 54]. These fundamental gray-level operators may be used to create certain often used image enhancement operators, such as top-hat and bottom-hat transformations, which make it easier to recover tiny objects from images [168, 145, 165]. In the field of marker identification, Ultimate Erosion (UE) is a well-known mathematical morphological procedure. Each linked component is subjected to a succession of erosion operators until the component is totally removed by a final erosion [39]. In this manner, items that come into contact with each other can be separated. It is possible that many markers will be produced for each item when using UE in noisy images. According to Park et al. [115], noise-resistant measurements of convexity may be used as a stopping condition for erosion in an enhanced UE operation. Starting with a binary image, it executes an erosion with a one-pixel-radius disk structure element for each non-convex linked component until the component is convex.

HIT/HAT Algorithms

HIT (H-minima transform) and HAT (H-maxima transform) algorithms are based on morphology operations for detecting cells in histopathological images [118, 123, 101, 162]; **LoG (Laplacian of Gaussian) filtering** can be very useful for identifying small blob objects, which usually correspond to the central regions of nuclei or cells in microscopy images [24, 91, 117];

Maximally Stable Extremal Regions (MSER)

The MSER detector can also be utilized for object detection [103]. An increasing intensity threshold is applied to the gray-level image. The result is a series of nesting extremal areas based on the intensity landscape level sets, with one region being deemed stable based on a local intensity minimum criteria. The MSER-based blob object identification technique typically consists of two fundamental phases in microscopy image analysis: Create a significant number of extremal areas, and then identify the ones that match to genuine nuclei or cells [168]. MSER nuclei detection in Pap smear microscopy images has been reported by using the eccentricity to assess and filter out those blobs that are less circular than actual nuclei [95]. Lu et al. [96], used the same approach to remove the undesirable extremal areas by combining the blob appearance information with the shape attributes.

Hough Transform (HT)

Because nuclei in pathological images frequently have circular or elliptical forms, the HT-based technique has been applied in a variety of applications [76, 116]; **Radial-Symmetry-Based Voting** [129, 130] is being used to detect the centroids of nuclei or cells in various applications [94, 156, 110];

2.2.2 Learning-based Detection Algorithms

This group of algorithms has been very useful in the field of pathology in dealing with histopathological images. As a machine learning approach, supervised learning attempts to infer a mapping function or model from training data. The identification of nuclei and cells is commonly presented as a classification issue at the pixel/voxel or superpixel/supervoxel level [168]. A specialized model is trained to convert fresh data instances into discrete labels [17]. It is possible to build a variety of classifiers that use different feature representations. This thesis will concentrate on deep neural networks (DNNs).

Deep Neural Networks (DNNs)

In recent years, neural nets have risen to the top of the rankings when it comes to high-quality object detection. Highlights of this literature are discussed in this section. Modern convolutional network detection had its start with the R-CNN article [50]. The R-CNN method was inspired by recent results in image classification [84] by taking the simple approach of cropping externally generated box suggestions from an input image and running a neural net classifier on these crops. Costs may rise due to the high number of crops required and the resulting duplication of calculations from overlapping crops. Cropping from an intermediate layer, Fast R-CNN [49] reduced the computational burden of feature extraction by running the entire image through a feature extractor only once [69]. Recent detectors are often made up of two parts: a backbone that has been pre-trained on ImageNet and a head used to forecast object classes and bounding boxes [18]. When using

GPU-based detectors, the backbone could be a commonly used model such as VGG [143], ResNet [63], ResNeXt [166], or DenseNet [68]. Also for CPU-based detectors, the backbone could be SqueezeNet [70], MobileNet [67, 137, 66, 150], or ShuffleNet [173, 97]. Head types can be categorized in two groups: One-stage object detectors like YOLO [126, 127, 128], Single Shot Detector (SSD) [92], and RetinaNet [90]; and two-stage object detectors like R-CNN [50] series, including fast R-CNN [49], faster R-CNN [131], R-FCN [34], and Libra R-CNN [113].

Faster R-CNN

Detection occurs in two steps when using Faster R-CNN [131]. Predictions of class-agnostic box proposals are made using features extracted at intermediate levels (e.g., "conv5") in the first step, known as the region proposal network (RPN). For each of these (usually 300) box suggestions, the feature extractor (e.g. "fc6" followed by "fc7") predicts a class and a class-specific box refinement, which are then fed into the feature extractor in the second step. The feature extractor is not used based on cropped image parts. The computation must be repeated for each region, therefore, the total duration depends on how many regions the RPN proposes. Faster R-CNN has had a significant impact in many applications since its release in 2015 [14, 140, 141, 172, 62].

R-FCN

While Faster R-CNN is several orders of magnitude faster than Fast R-CNN, Dai et al. [34] proposed the R-FCN (Region-based Fully Convolutional Networks). It is similar to Faster R-CNN but crops features from the last layer of features prior to prediction rather than cropping features from the layer where region proposals are predicted. For example, this method of reducing per-region calculation by shifting cropping to final layer reduces computing time. A position-sensitive cropping method is proposed instead of the more usual ROI pooling procedures in [49, 131] or the differentiable crop technique in [33] because Dai et al. claim that the object identification task requires localization representations that respect translation variance. Faster R-CNN can typically match the accuracy of the R-FCN model while operating at a lower cost thanks to the model's use of ResNet 101. R-FCN was recently used to segment instances in the latest TA-FCN model [89], which won the 2016 COCO instance segmentation competition. In Figure 2.2 the high-level chart of these three discussed object detection meta-architectures has been shown.

Single Shot Detector (SSD)

SSD [92] refers to designs that employ a single feed-forward convolutional network to predict classes and anchor offsets directly, without the need for a second stage classification operation per proposal. This method has been used in many applications [139, 126, 127].

| (a) SSD | (b) Faster R-CNN | (c) R-FCN |
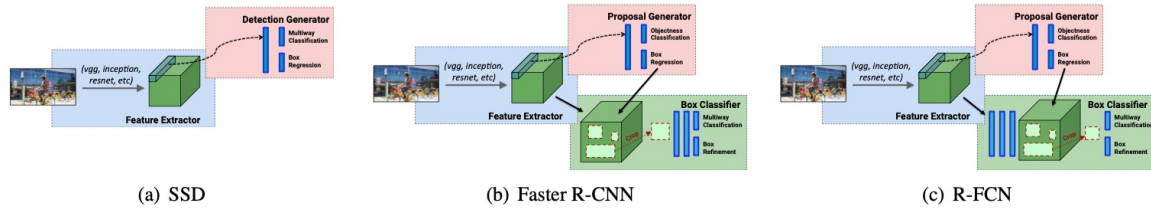
Figure 2.2: High-level chart of the three discussed object detection meta-architectures. This image has been taken from [69].

YOLO

YOLO is another single feed-forward convolutional network. Humans can tell instantaneously what objects are in an image, where they are, and how they interact with one other just glancing at the image. We are able to do difficult activities like driving because of the human visual system's speed and accuracy. In order to enable computers to drive automobiles without specialized sensors, assistive devices can provide real-time scene recognition to human users, and general-purpose responsive robotic systems can be unlocked with fast and precise object detection algorithms [126]. Detection technologies nowadays reuse classifiers to recognize certain shapes. In order to detect an object, these systems use a classifier for that object to analyze the picture at multiple sizes and locations. Sliding window systems like deformable parts models (DPMs) employ a classifier run at regularly spaced regions over the whole image [42].

Region proposal techniques like R-CNN generate possible bounding boxes in an image and then run a classifier on these suggested boxes. In post-processing, the bounding boxes are refined, duplicate detections are eliminated, and the boxes' scores are recalculated depending on other objects in the scene [51]. Because each component must be learned individually, these detection pipelines are slow and difficult to optimize. YOLO, using a single convolutional network, may predict several bounding boxes and class probabilities at the same time. YOLO is trained on entire images and directly improves the detection performance. Compared to more traditional approaches, this unified model provides some advantages in object detection. In this study, YOLO has been studied on two different applications: **WSI Quality control**, and **glomeruli detection** in kidney images.

2.3 WSI Quality Control

Using a microscope, a trained pathologist may examine tissue samples and make a diagnosis of a wide range of diseases. The majority of malignancies are diagnosed through a biopsy [146]. The biopsy technique begins with the removal of a tiny portion of the questionable tissue. A microscopic portion of tissue is mounted on a glass slide following tissue preparation, which may include drying, embedding in paraffin, cutting, staining and other preparation steps [144]. Artifacts, such as tissue folding, air bubbles, and tearing can be introduced into the tissue preparation process at any point [74]. Under the light

microscope, pathologists can examine the glass slide in order to generate a diagnostic report. Computed-Assisted Diagnosis (CAD) can boost accuracy and produce more accurate findings by reducing pathologists' workload, which in turn reduces observation mistakes. This examination may be time-consuming as well as error-prone due to human factors [74].

Artifacts, such as folded tissue, may have a detrimental impact on the diagnosis regardless of whether we employ digitization or microscopy. Tissue folds can arise during the processing of the tissue while cutting tiny block slices [9]. There are two examples of folded tissue depicted in the Figure 2.3. Other aberrations, such as blur, might affect computerized algorithms when they are implemented using digital technology [9].
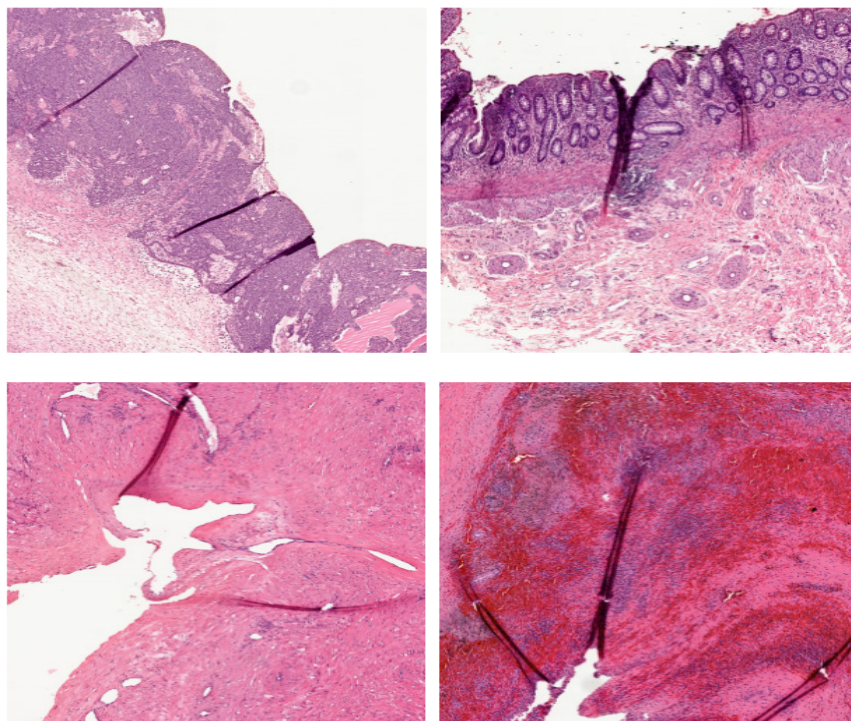
Figure 2.3: Sample tissue folds which can happen in the laboratory during glass slide preparation.

To protect the stained glass slides from damage, the slides are usually fixed on the tissue sections using mounting media [125]. During floating of the tissue cuts in the water bath or covering, artifacts such as air bubbles may emerge negatively affecting the inspecting of important tissue morphology features [44]. Figure 2.4 shows two samples of tissue slides containing air bubbles.

Computational pathology requires the digitization of the tissue glass slides to support diagnosis with computers [43]. Once the glass slides are digitized, marker signs on glass slides, which pathologists may use to highlight some diagnostic hints, may appear in the digital images as well. These colored strips can occlude tissue and mislead the interpretation of the tissue information [99]. Figure 2.5 shows two samples of ink-marker images.
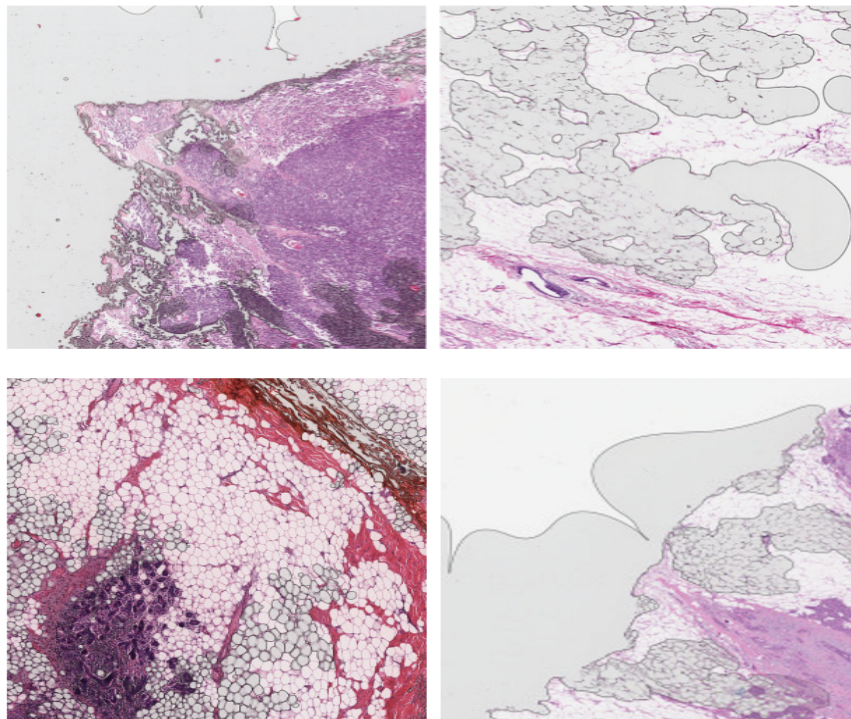
Figure 2.4: Sample tissue slides containing air bubble which can happen in the laboratory during slide preparation.

Instead of rejecting the entire slide because of artifacts and ink-markers, which is not an efficient approach, there are some techniques for the detection and removal of these artifacts [99]. Traditional object detection systems usually commence by feature extraction from the input images using methods like Haar [114], SIFT [93], and HOG [35]. They generally use handcrafted features based on morphology, thresholding of pixels, and voxel clustering [107] [73]. These methods may be effective, but the segmentation may not be sufficiently accurate. Moreover, the number of diagnostically relevant but missed lesions is high [35]. HistoQC [74], a tool for quality control of WSI acquisition, can discover artifacts and outputs outliers. However, the method to remove artifacts is mostly based on thresholding techniques combined with mathematical morphology operations. Two pathologists were asked to label the mask images produced by the tool as either acceptable (if the overlap between pathologists' visual artifact-free mask and tools' artifact-free mask is over 85%) or not acceptable (if the this condition is not satisfied). Works have reported 95.5% average accuracy for such approaches but with limited data [74]. Bautista et al. [13] proposed a method using color shifting to magnify the color difference and detect tissue fold. These methods have inevitable drawbacks and are dependent on the manual setting of thresholds and other parameters. Therefore, it is an urgent need to utilize more automated methods for segmentation.

The other segmentation method widely employed is specialized deep networks that enable complex data distribution modelling. They have also been applied in several image analysis, and detection problems in the medical domain [136]. CycleGan was employed by
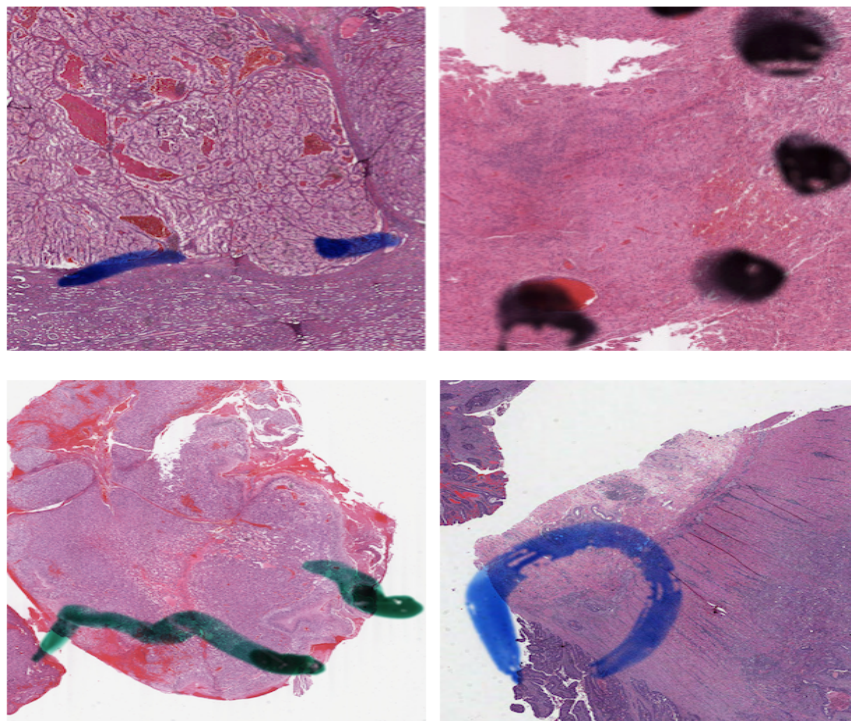
Figure 2.5: Sample ink-markers on tissue which occur when pathologists highlight regions of interest on glass slides. This may happen for consulting other pathologists, or for educational or research purposes.

Venkatesh et al. [155] for the reconstruction of marker removed WSI and was utilized by Zhu et al. [175] for the restoration of the original non-marked image tiles. Besides, Ali et al. [4] separated marked tiles with a binary classification and used the marked tiles as inputs for a YOLO3 network. Unlike the YOLO object detection, which is very fast and can even be used as a preprocessing step for more specialized networks, these segmentation networks and deep feature extraction are rather slow. Ramachandran et al. [48] proposed a method for lung nodule detection and localization using the YOLO-based deep learning, which results in low false positive rates with high sensitivity.

2.4   Glomeruli Detection in Kidney Images

A significant stage in determining whether a kidney is transplantable is the histological examination of renal samples by experienced pathologists [5, 100]. The histopathology evaluation of the number of globally sclerosed glomeruli in relation to the overall number of glomeruli is an essential examination for accepting or rejecting a donor's kidneys [5]. In Figures 2.6, and 2.7 samples of glomeruli objects in kidney images have been shown.

Waste and excess fluids are expelled from the human body by glomeruli, which are clusters of capillaries responsible for expulsion. It is possible to group glomerular disor-
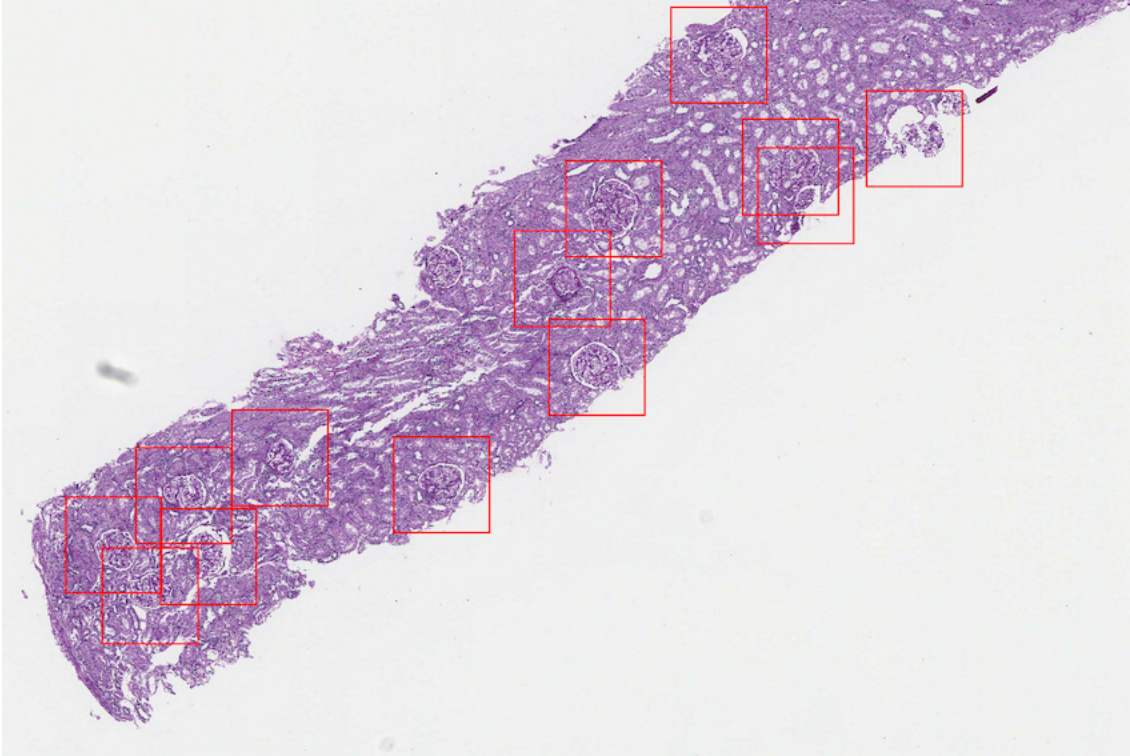
Figure 2.6: Glomeruli Samples.
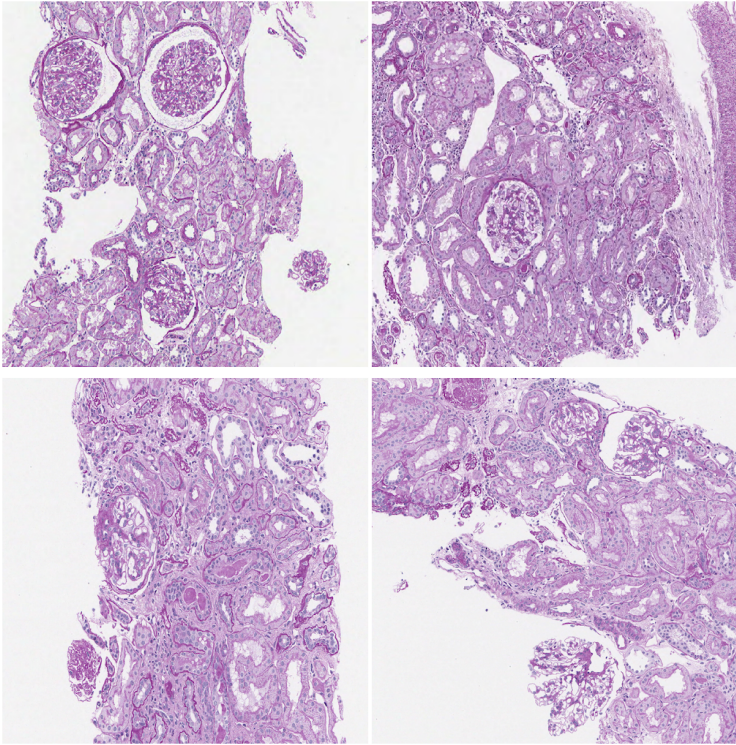
Figure 2.7: Glomeruli Samples.

ders according on their clinical symptoms, etymology, immunopathology, or morphological changes [5, 100]. A condition known as "glomerulosclerosis" is the result of the kidney lesion changing its morphology; this sclerosis can impact the kidney in many ways, depending on whether or not it is global or partial [21]. The number of glomeruli detected in each kidney biopsy should be counted in daily practice. Per kidney biopsy, about 20 to 30 cuts are made [21]. Additionally, glomeruli that are completely sclerosed must be noted (the entire glomerulus). Detection of localized sclerosis will provide further information regarding the patient's condition. Each pathology report should include this information because the number of glomeruli assessed must be representative enough to determine a diagnosis [159]. On the other hand, if the sample has numerous sclerosed glomeruli, this may suggest that the patient has chronic kidney disease with dead glomeruli. As a result, the patient may not be suited for some medications, which will help to define adequate treatment [159]. This information is also entered into the national register for glomerulonephritis by the [21]. Time-consuming and tiresome, the count of glomeruli is a painstaking process. Because of this, image processing methods that can identify and categorize the glomerulus are needed.

With the emergence of deep learning networks, various options for computer vision tasks such as glomeruli object identification, semantic segmentation, and instance segmentation became available [21]. For instance, some works provide a detailed assessment of object identification and instance segmentation algorithms [174]. Some other give a complete review for semantic segmentation [46]. Several recent research in digital pathology have used deep neural networks for glomeruli detection, and segmentation [88, 79, 27, 6, 45, 100, 78, 142, 153, 21].

For Glomeruli detection, YOLO has been applied on kidney images for the first time in this thesis and compared the existing segmentation method with U-Net on the same validation dataset. There are two different tissue staining images in the validation dataset.

2.4.1 Tissue Staining

Staining is used to emphasize essential characteristics of the tissue, as well as improve contrast. Hematoxylin is a typical stain dye used in this technique that gives the nuclei a bluish hue, whereas eosin (another stain dye used in histology) gives the cell's nucleus a pinkish tint [7].

2.4.2 Periodic Acid-Schiff (PAS)

A staining technique called periodic-Acid Schiff (PAS) is used in histochemistry and histology to show that carbohydrates and carbohydrate compounds like polysaccharides, mucin, glycogen, and fungal cell wall components are found in the cells. It has been used to look for glycogen in places like the skeletal muscles, the liver, and the heart muscles. PAS staining works with both formalin-fixed, paraffin-embedded, and frozen tissue sections [164].

2.4.3 Hematoxylin and Eosin (H&E) Staining

There are two types of histological stains that come together to make H&E: Hematoxylin and Eosin. The hematoxylin stains cell nuclei purple, and eosin stains the extracellular matrix as well as the cytoplasm pink. Other anatomic tissue structures take on different shades and hues of these two colors [29]. There are two parts of a cell that are called the nucleus and the cytoplasm. Pathologists can easily tell them apart, and the overall patterns of coloration from the stain show the general layout and distribution of cells and give an overall image of a tissue morphology [164].

Chapter 3

Method, Data Preparation & Experiments

3.1 Method

Predicting one or more object locations, determining their classes, and drawing a bounding box around the object is the definition of an object detection task. In many existing detection systems, multiple classifiers are applied to an image at many locations and scales to calculate the high-scoring regions of the image for detecting a region of interest. In this thesis, the YOLO approach (You Only Look Once) [126] has been trained for both applications on detecting tissue patterns in WSIs, one is artifacts and manual ink-markers detection, and the second one is glomeruli detection in kidney images.

One of the essential advantages of YOLO over classifier-based systems is the speed of this model. YOLO is faster than R-CNN more than 1000 times [50], and 100 times faster than Fast R-CNN [49]. Predictions based on YOLO are with a single network evaluation, while R-CNN requires many network evaluations for a single image. More importantly, as an object detector, YOLO does not require a detailed pixel-level annotation; labels for YOLO are just bounding boxes around the target objects.

3.1.1 Network Architecture of YOLO

By combining separate components of other object detection networks, like the ones using a sliding window, or region-based techniques, YOLO can predict all image objects for all the classes based on the information from the whole image only by looking at the image once [126]. In other words, the network models the entire image at once along with all of its individual objects. End-to-end training and real-time speeds are made possible by the YOLO architecture while high average accuracy is maintained [126].

An $S \times S$ grid is generated on any given image. A grid cell is responsible for identifying an object whose center lies within that grid cell. Boxes and confidence ratings are predicted

for each grid square. If the model is certain that the box contains an object, it will give it a high confidence score [126]. This confidence score is calculated based on

$$Pr(Object) = IOU_{pred}^{truth} \qquad (3.1)$$

The confidence score should be 0 if there is no predicted object present in a cell. If there is at least one predicted object in that cell, for the confidence score to be accurate, it must be equal to the intersection over union (IOU) between the predicted box and the ground truth. The probabilities of each C conditional class, $Pr(Class_i|Object)$, are also predicted in each grid cell. The location of object's grid cell determines these probabilities. No matter how many boxes B there are in a grid cell, the network can only forecast one set of class probabilities. For the evaluation of the network, the network computes the class-specific confidence scores for each box based on

$$Pr(Class_i|Object) * Pr(Object) * IOU_{pred}^{truth} = Pr(Class_i) * IOU_{pred}^{truth} \qquad (3.2)$$

Both the likelihood that a certain class will be found in the box and how well the predicted box will fit the item are represented by these scores [126]. In Figure 3.1, the whole procedure has been visually summarized.
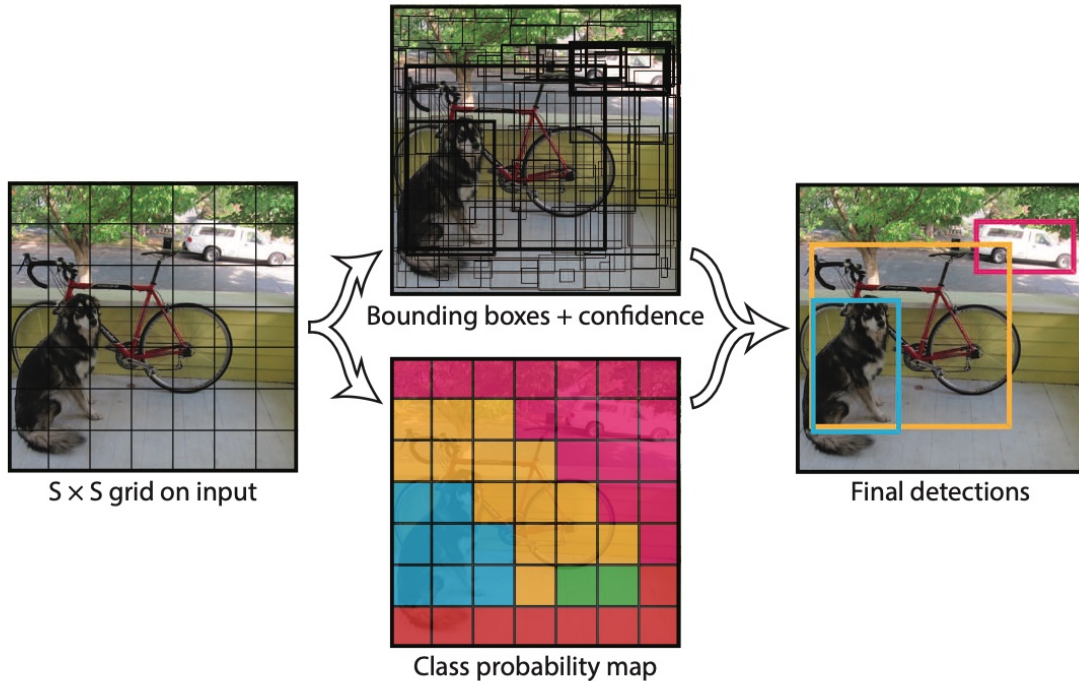
Figure 3.1: For each grid cell, YOLO forecasts the B bounding boxes, their confidence levels, and their C class probabilities. This image has been taken from [126].

YOLO was inspired by GoogleNet model for image classification [149]. It has 24 convolutional layers, followed by two fully connected layers make up the detecting network. The full network has been shown in Figure 3.2.
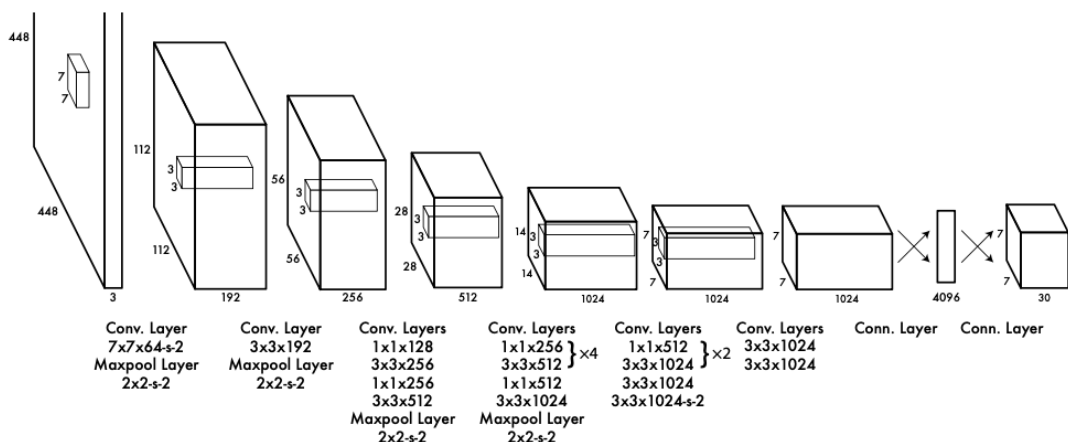
Figure 3.2: YOLO Architecture: 24 convolutional layers and two fully linked layers make up the detecting network. This image has been taken from [126].

3.1.2 YOLO-v4

The YOLO-v4 [18] model has been trained for both applications in this thesis. The codes of all steps involved are available on GitHub [1], and in this experiment, the codes have been modified in a way to train a custom dataset. The implementation of a new architecture in the backbone in YOLO-V4 compared to YOLO-V3 has made an essential improvement in the mAP (mean Average Precision) and the number of FPS (Frame per Second) by 10% and 12%, respectively, when trained and tested on COCO dataset[2]. In Figure 3.3 the comparison between YOLO-V4 and other object detector methods has been shown. The new architecture in the backbone is a deep neural network composed mainly of convolution layers, and the main objective is to extract features. The backbone selection is a key step and can improve object detection performance; mostly pre-trained neural networks are used to train the backbone [18].

3.2 WSI Quality Control

Instead of rejecting the whole slide because of artifacts and ink-markers, which is not very efficient, there are some ways to find and remove these artifacts. Experiments and dataset for detecting these artifacts, and WSI quality control using YOLO-v4 has been discussed in this section.

---

[1]https://github.com/AlexeyAB/darknet
[2]https://cocodataset.org/
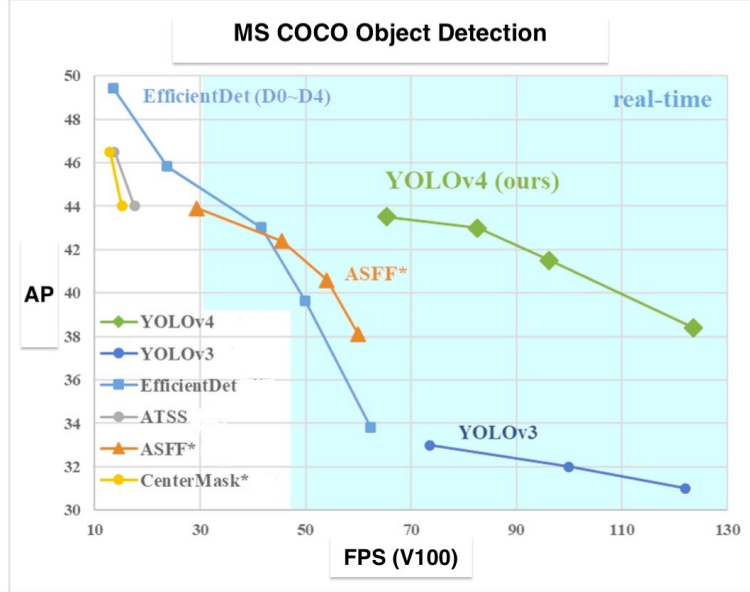
Figure 3.3: YOLOv4 vs. other object detectors. In comparison to EfficientDet, YOLOv4 is twice as fast, improves YOLOv3's average precision (AP) and Frames per Second (FPS) by 10% and 12%, respectively. This image has been taken from [18].

3.2.1 Dataset

For the quality control experiments, training and validation sets of WSIs with folded tissue, ink-marked tissue WSIs, and tissue slides with air bubbles have been created containing 92 WSIs for training (40 of folded tissues, 40 of ink-marked tissues, and 12 tissue slides with air bubbles) and 15 WSIs for validation (6 of folded tissues, 6 of ink-marked tissues, and 3 of tissue slides with air bubbles). Images at 2.5x magnification (roughly equivalent to 2000×3000 pixels) have been labeled. For folded tissue dataset, 46 WSIs from [9], for the ink-marked tissue dataset, 46 WSIs from [99], and for the tissue slides with air bubbles, 15 WSIs from *HistoQC* repository [3] have been used.

In Figures 3.4, 3.5, and 3.6 3 WSIs in the training dataset, containing ink-marker, tissue-fold, and air bubbles with annotated bounding boxes have been shown. For using these images for the YOLO model, first, labelling target objects (tissue fold and ink-marker on tissue) in each image using the YOLO format 3.2.1 is required. For labelling the images, using the above networks' output images and segmented areas, the coordinates of the bounding boxes around each detected object have been calculated.

A *txt* file for the annotation coordinates, with the same name for each image *jpg* file has been created:

$$\text{object-class}, x_C, y_C, \text{width}, \text{height} \tag{3.3}$$

where

---

[3]http://histoqcrepo.com/

- *object-class* is integer object number from 0 to $(classes - 1)$

- x_C, y_C, *width, height* are float values relative to width and height of image, it can be equal from (0.0 to 1.0]

- attention: x_C, y_C are center of bounding box rectangle
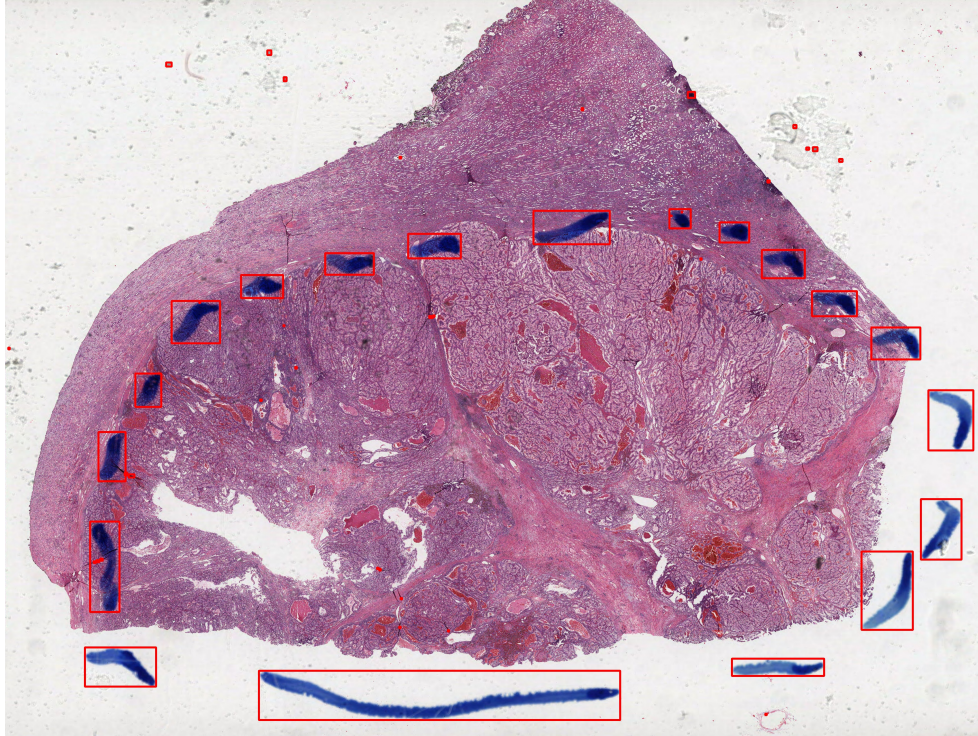
Figure 3.4: Sample WSI containing ink-markers.

3.2.2   Experiments

The dataset in this study contains 92 WSIs for training and 15 WSIs for validation. The network aims to detect three different class objects, namely tissue folds, air bubble edges in tissue, and ink-marker on a tissue. Pre-trained weights [4] have been used to fine-tune the network using on the training data. The configuration of the network has been shown in Table 3.1. In this table,

- **Batch** stands for how many images are used in the forward pass to compute a gradient and update the weights via back-propagation,

- **subdivisions** stands for the number of blocks in which the batch is subdivided,

---
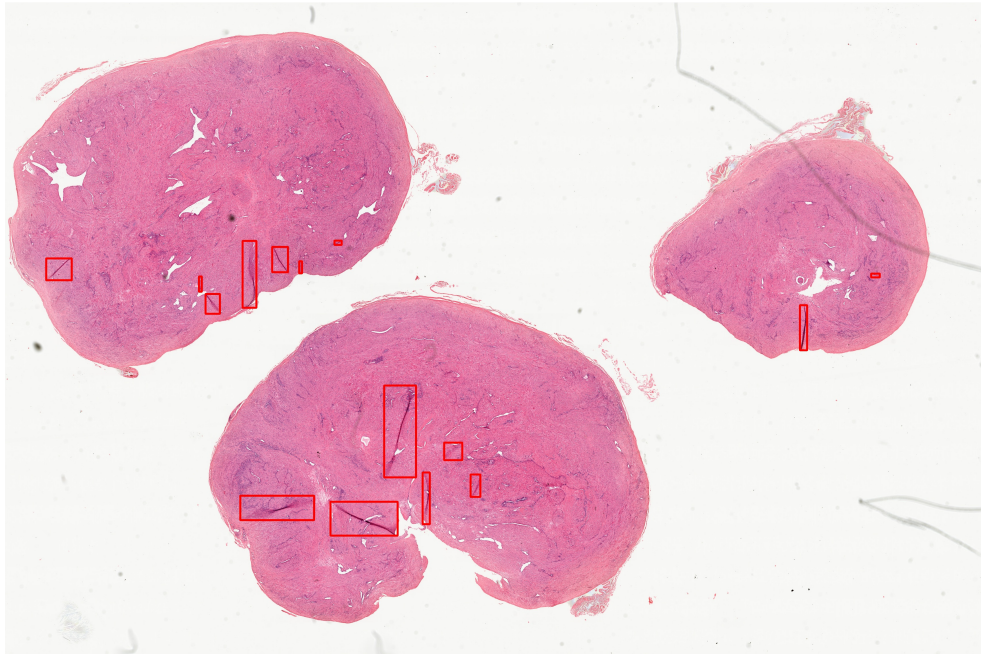
Figure 3.5: Sample WSI containing tissue folds.
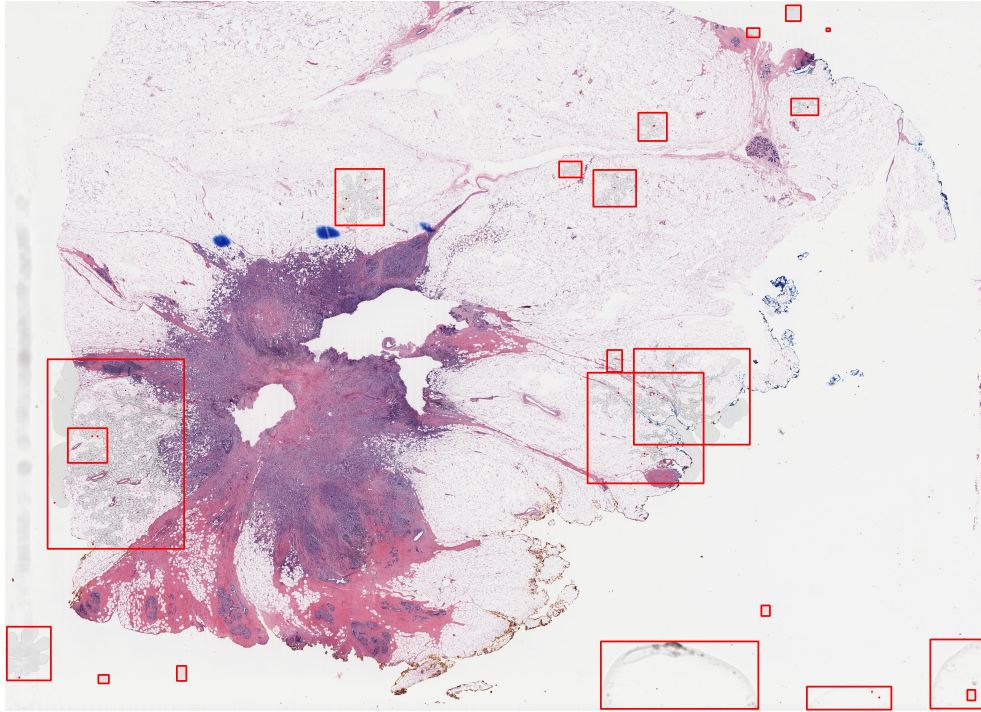
Figure 3.6: Sample WSI containing air bubbles.

Table 3.1: The configuration of the network WSI for quality control

learning rate	batch	subdivisions	policy	steps	scales	max batches	filters	activation
0.001	64	16	steps	4800, 5400	0.1,0.1	6000	24	linear

- **policy** mean using the steps and scales parameters bellow to adjust the learning rate during training,

- **steps** means adjust the learning rate after 3200 and 3600 batches,

- **scales** means re-scale the current learning rate by the corresponding factor once the number of steps is reached,

- **max batches** is the maximum number of iterations,

- **filters** stands for how many convolutional kernels there are in a layer, and

- **activation** defines the activation function.

For training YOLO, usually around 2000 iterations are sufficient for each class (object) [18]. The network was trained for 6000 iterations and the common object detection evaluation metrics for training were monitored to obtain the best weights in chapter 4 4.

3.3 Glomeruli Detection in Kidney Images

Glomeruli are clusters of capillaries responsible for expulsion, remove waste and excess fluids from the human body [21]. The histological examination of glomeruli is crucial in determining whether or not a kidney may be transplanted [5]. Manual examination of kidney samples is time-consuming and error-prone. Therefore, automated glomeruli identification frameworks for kidney biopsies can be beneficial [5]. This topic is one of the applications studied in this thesis.

3.3.1 Dataset

Kimia Lab and the pathology department of the University of Michigan are collaborating on a project for developing a computational kidney disease diagnosis model. As a part of this project, Kimia Lab has received a glomeruli dataset with bounding box annotations created by nephrologists. To expand the training data two different public datasets, plus a private dataset from the University of Michigan have been used in this study.

Public Dataset 1   The first public dataset consists of 31 WSIs in SVS format. With the size range between 21651 × 10498 pixels and 49799 × 32359 pixels acquired at 20x to preserve image quality and information while requiring significantly less computational time than images taken at other magnifications [22]. A glomerulus may lose structural information due to the lower resolution and poor image quality. It is also important to note that employing magnifications such as 40x would increase the model size, slowing down the training process [21]. This data is part of the WSI datasets generated within the *European project AIDPATH* [5]. A biopsy needle with an outside diameter of between 100 nm and 300 nm was used to obtain tissue samples. Once the paraffin blocks were ready, the tissue portions were cut into 4 m pieces and coloured with through PAS staining [22]. It is common to employ PAS stain to color polysaccharides found in kidney tissue and to highlight glomerular basement membranes because of its effectiveness [133]. These images contain different types of glomeruli labeled by Bueno et al. approach [21]. This dataset has two parts, DATASET_A, which contains the raw 31 WSIs, and DATASET_B, which is 2340 glomeruli images, 1170 normal glomeruli and 1170 sclerosed glomeruli. Because of the lack of exact coordinates of the extracted glomeruli, the exact coordinates of the glomeruli bounding boxes were extracted by a pathologist at Kimia Lab [6]. An annotated WSI sample of the first public dataset has been shown in Figure 3.7.

Public Dataset 2   This dataset has been used for *HubMap competition* [7]. TIFF files ranging in size from 500MB to 5GB make up the dataset containing 8 images for the training and 5 images for the test. RLE-coded and uncoded (JSON) annotations are included in the training set. The annotations identify glomeruli that have been divided into sections. Also, anatomical structural segmentations are included in both the training and public test sets. The bounding boxes of these anatomical structures for using these annotations for the YOLO object detector have been created based on manual contours. Figure 3.8 is an example of the procedure to generate a bounding box from manual delineation. This bounding box is found by calculating the upper left most and lower right-most coordinates in the delineation. An annotated WSI of the second dataset has been shown in Figure 3.9.

University of Michigan Data   This private dataset has been collected from the University of Michigan and annotated by expert pathologists. The training dataset consists of 7 Periodic acid–Schiff (PAS) stained WSIs for fine-tuning the models that have been trained on the mentioned public datasets. Annotated WSI sample of this dataset has been shown in Figure 3.10. Beside these 7 PAS stained WSIs for training, 20 PAS stained WSIs, and 16 H&E stained WSIs have been used for validation. The images show that the University of Michigan data has the same type as the first public dataset, namely needle biopsy images, unlike the second public dataset, which are surgical excisions. This difference would affect the results obtained by each public dataset.

---

[5]http://aidpath.eu/
[6]Dr. Ricardo Gonzalez, an anatomic pathologist, visiting Kimia Lab from May 2021 till April 2022.
[7]https://www.kaggle.com/c/hubmap-kidney-segmentation/overview
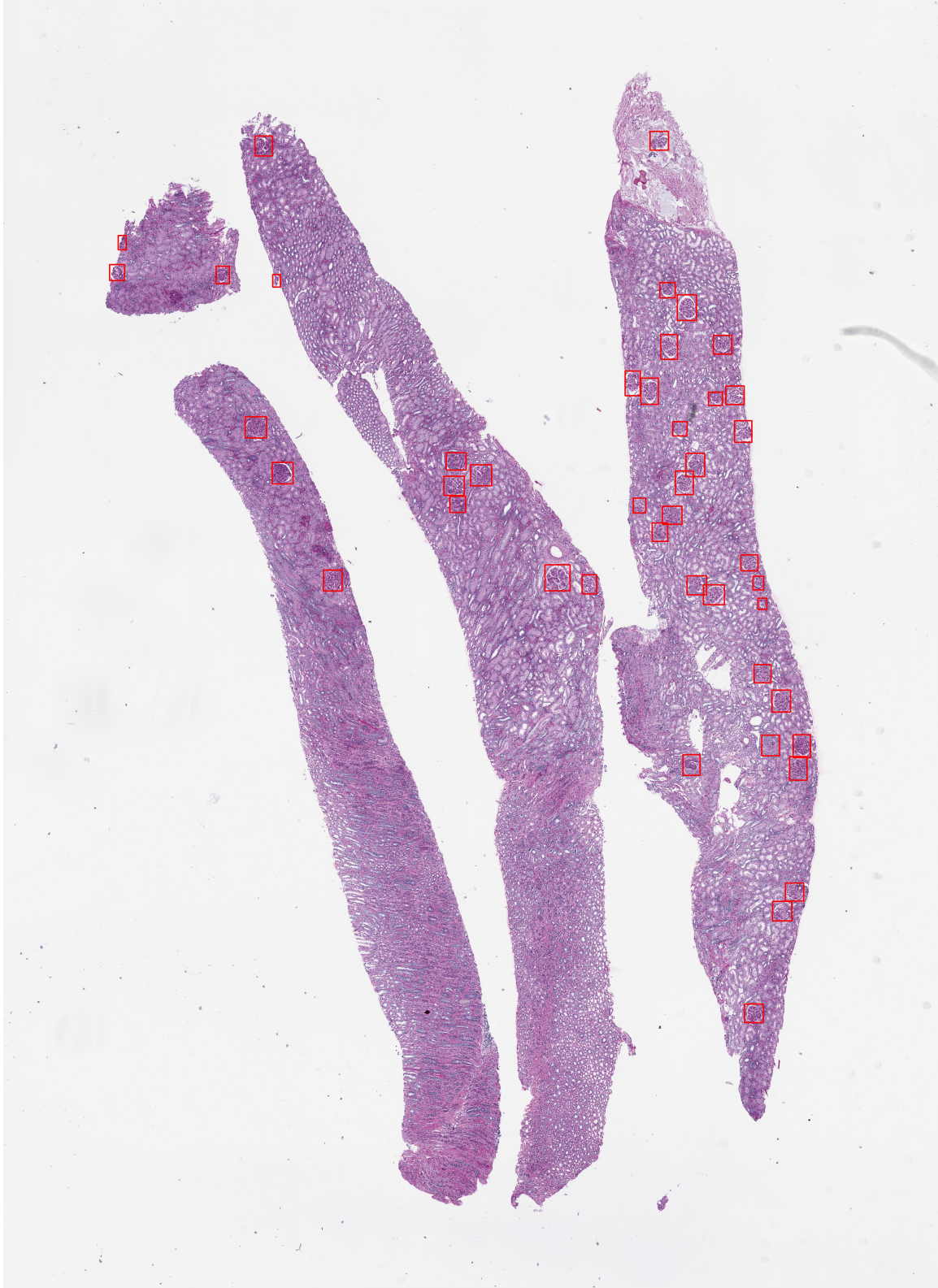
Figure 3.7: Annotated WSI sample from the public dataset 1.
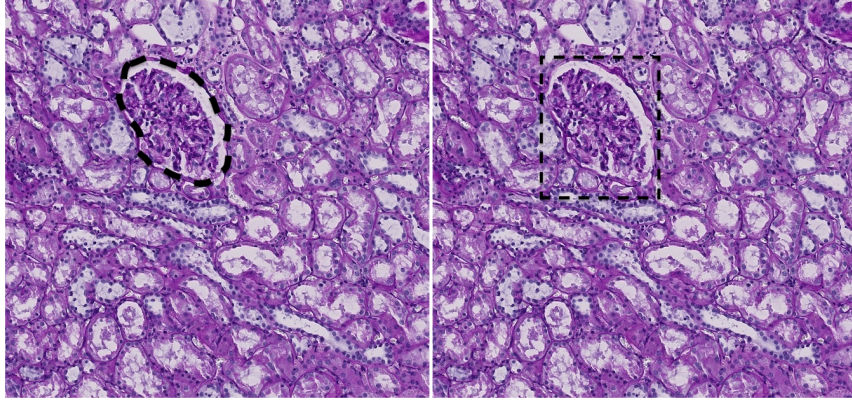
Figure 3.8: Extracted bounding boxes (right) from manual delineations of a glomerulus (left).

3.3.2   Experiments

A total of 7 different combination of datasets (using two public datasets, and a private dataset) selected for the training of YOLO object detector, resulting in a total of 7 different models. The 7 training datasets have been evaluated on two different validation datasets with different stains from the University of Michigan: One contains 20 PAS stained images and the other one contains 16 H&E stained images. All experiments, along with the explanation of the training and the validation dataset, have been reported in Table 3.3. And the configurations of the network for all 7 different training datasets have been described in Table 3.2. In this table,

- **Batch** stands for how many images are used in the forward pass to compute a gradient and update the weights via back-propagation,

- **subdivisions** stands for the number of blocks in which the batch is subdivided,

- **policy** means using the steps and scales parameters bellow to adjust the learning rate during training,

- **steps** means adjust the learning rate after 3200 and 3600 batches,

- **scales** means re-scale the current learning rate by the corresponding factor once the number of steps is reached,

- **max batches** is the maximum number of iterations,

- **filters** stands for how many convolutional kernels there are in a layer, and

- **activation** defines the activation function.

Many studies have been performed to identify glomeruli functional tissue units in human kidneys. Recently, there was a competition, **Hacking the Kidney**, launching in Kaggle

Table 3.2: The configuration of the network for all training datasets for glomeruli detection

learning rate	batch	subdivisions	policy	steps	scales	max batches	filters	activation
0.001	40	16	steps	4800, 5400	0.1,0.1	6000	18	linear

Table 3.3: All 7 training sets along with the test experiments using public daatsets 1 and, the proviate dataset from University of Michigan (UMICH).

Experiment	Training Dataset	Test Dataset
1	31 WSIs from public dataset 1	20 PAS WSIs from UMICH dataset
		16 H&E WSIs from UMICH dataset
2	31 WSIs from public dataset 1, fine-tuned with 7 PAS WSIs from UMICH dataset	20 PAS WSIs from UMICH dataset
		16 H&E WSIs from UMICH dataset
3	8 WSIs from public dataset 2	20 PAS WSIs from UMICH dataset
		16 H&E WSIs from UMICH dataset
4	8 WSIs from public dataset 2, fine-tuned with 7 PAS WSIs from UMICH dataset	20 PAS WSIs from UMICH dataset
		16 H&E WSIs from UMICH dataset
5	31 WSIs from public dataset 1, and 8 WSIs from public dataset 2	20 PAS WSIs from UMICH dataset
		16 H&E WSIs from UMICH dataset
6	31 WSIs from public dataset 1, and 8 WSIs from public dataset 2, fine-tuned with 7 PAS WSIs from UMICH dataset	20 PAS WSIs from UMICH dataset
		16 H&E WSIs from UMICH dataset
7	7 PAS stained WSIs from UMICH dataset	20 PAS WSIs from UMICH dataset
		16 H&E WSIs from UMICH dataset
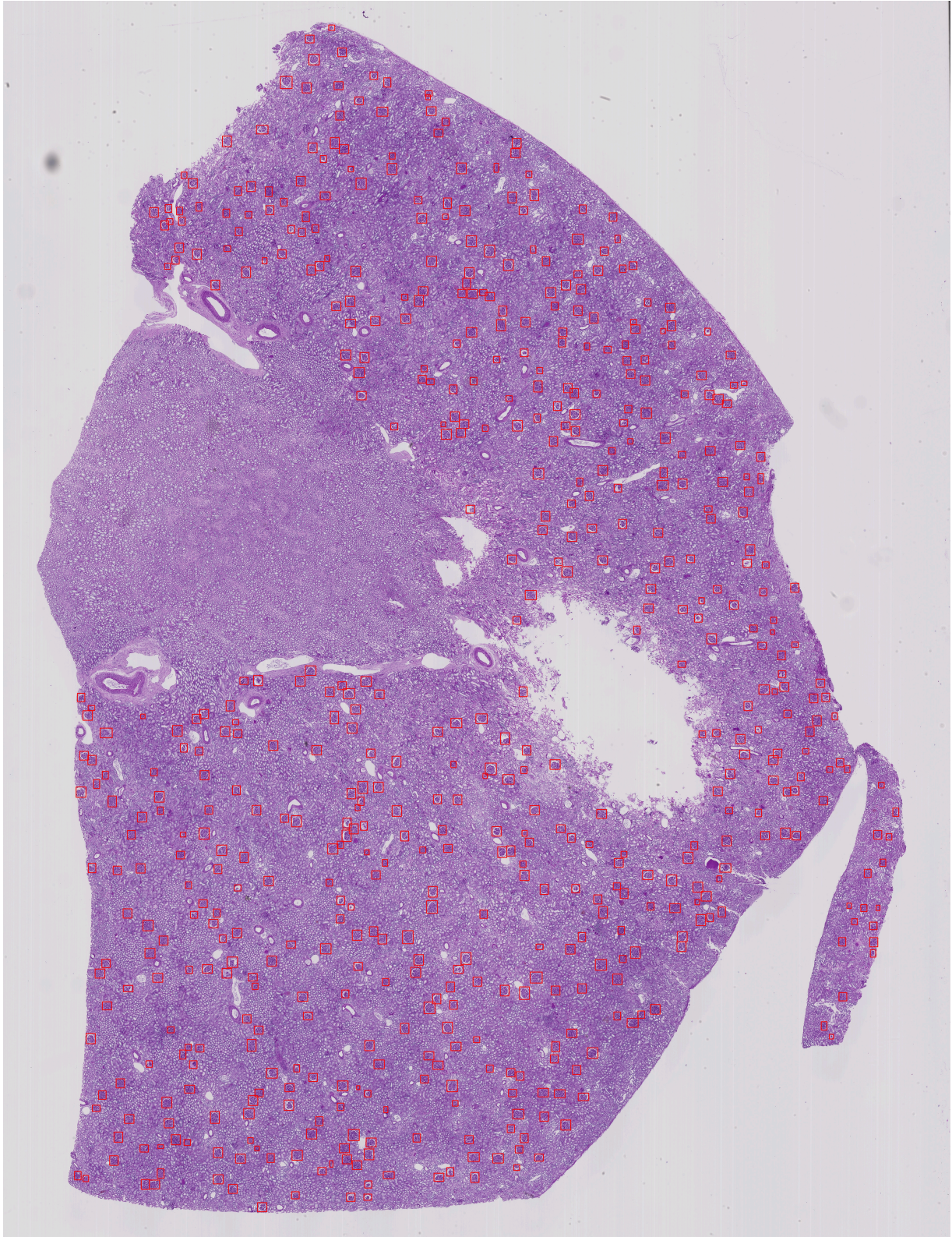
Figure 3.9: Annotated WSI sample from the public dataset 2.

Figure 3.10: Annotated WSI sample from the University of Michigan's private dataset.

to segment glomeruli in kidney images [8]. The dataset provided for the competition was public dataset 2, discussed in the dataset section. TIFF files, ranging in size from 500MB to 5GB, make up the dataset containing eight images for the training and five images for the test. RLE-coded and uncoded (JSON) annotations are included in the training and validation sets. The authors of a study [52] compare the five winning algorithms between more than a thousand teams that participated in the above competition. They assess the accuracy and performance of the five top algorithms, and the codes are available online [9].

To compare a segmentation model with the detection model in this thesis, the first team's algorithm has been chosen as the benchmark. The accuracy on the same validation dataset e.i. 20 PAS stained images and 16 H&E images from the University of Michigan has been calculated based on the explanation for the winning proposal [10]. They have used a single U-Net SeResNext101 architecture with Convolutional Block Attention Module (CBAM), hypercolumns, and deep supervision. Their network read 1024×1024 pixel patches and then downsample them to 320×320 patches. SGD is the optimizer for their model, trained using binary cross-entropy. Training is perfromed for 20 epochs, with a learning rate of 10^{-4} to 10^{-6} and a batch size of 8 images. Their final weights trained on the whole training dataset have been used to validate and compare their network on the University of Michigan dataset, which contains 20 PAS stained images and 16 H&E stained images. The results are provided in Chapter 4. Note that this is not possible to fine-tune the mentioned segmentation model with the external validation set (University of Michigan WSIs) as the external WSIs do not contain the pixel-level annotation. For comparing the segmentation model with YOLO, the segmentation area is enclosed with the smallest possible rectangle (the upper left most and lower right-most coordinates) and use these rectangles as the segmentation model output. Figure 3.8 depicts the process visually.

---

[8]https://www.kaggle.com/competitions/hubmap-kidney-segmentation
[9]https://github.com/cns-iu/ccf-research-kaggle-2021/
[10]https://www.kaggle.com/c/hubmap-kidney-segmentation/discussion/238198

Chapter 4

Results & Analysis

4.1   WSI Quality Control

One of the common evaluation metrics for detection and segmentation is the IoU measure (intersection over union, also called Jaccard Index [15]) which can determine the overlap between two regions divided by their union. The validation dataset 3.2.1 pre-defined an IoU threshold (i.e., 0.5) in classifying whether the prediction is a true positive or a false positive.

Another evaluation metric for object detection models like R-CNN [49] and YOLO [18] the is mean [132]. The mAP compares the ground-truth bounding box to the detected box and returns a score. The higher the score, the more accurate is the detection [64].

The validation dataset contains 15 WSIs that include 199 unique objects in total including tissue fold, air bubbles in tissue, or ink-marker on tissue. The purpose of the network is to detect these 199 objects. To better understand the training procedure and obtain the best weights for the network, the evaluation stage has been done with weights after every 1,000 iterations. The comparison between the results based on the above evaluation metrics on the validation dataset is reported in Tables 4.1 & 4.2. By having True Positives as TP, False Positives as FP, False Negatives as FN, and True Negative as TN, The formula of the evaluation metrics are as follows [19]:

$$\text{Precision} \;=\; \frac{TP}{TP + FP} \tag{4.1}$$

$$\text{Recall} \;=\; \frac{TP}{TP + FN} \tag{4.2}$$

$$\text{F1 Score} \;=\; 2 \times \frac{Precision \times Recall}{Precision + Recall} = \frac{TP}{TP + \frac{1}{2}(FP + FN)} \tag{4.3}$$

Some of the results for ink-marker detection have been shown in Figures 4.1 and 4.2. The left picture in Figures 4.1 and 4.2 demonstrates the bounding boxes of the actual ink-marker areas, and on the right, the predicted detection areas has been shown. The

Table 4.1: Comparison of the results based on **common evaluation metrics** and the number of iteration for training on the validation dataset

Number of Iterations	TP	FP	FN	Precision	Recall	F1 Score
1000	110	394	89	0.22	0.55	0.31
2000	149	133	50	0.53	0.75	0.62
3000	141	89	58	0.61	0.71	0.66
4000	142	34	57	0.81	0.71	0.76
5000	192	32	7	0.86	0.96	0.91
6000	199	0	0	1.00	1.00	1.00

Table 4.2: Comparison of the results based on **object detection evaluation metrics** and the number of iterations for training on the validation dataset for tissue fold (TF), ink-marker (IM) and air bubble (AB)

Number of Iterations	Average IoU	AP for TF	AP for IM	AP for AB	mAP
1000	15.83%	25.63%	85.59%	0.85%	37.36%
2000	39.17%	59.38%	91.17%	25.33%	58.63%
3000	45.95%	63.73%	87.02%	1.58%	50.78%
4000	61.60%	77.93%	93.93%	43.04%	71.64%
5000	72.78%	96.92%	98.56%	86.46%	93.98%
6000	99.50%	100%	100%	100%	100%

results that have been obtained, are more accurate compared to HistoQC [74]. The results that have been reported in HistoQC [74] are based on two pathologists' reviews. Each pathologist has reviewed 250 samples separately. Overall, HistoQC and the experts agreed on 94% (235 of 250) of cases for expert 1 and 97 percent (242 of 250) of cases for expert 2.

Some results of air bubble edge detection are demonstrated in Figure 4.3. The left picture demonstrates the bounding boxes of the actual air bubble boundary areas, and on the right, the predicted detection areas have been shown. A sample for tissue fold detection is illustrated in Figure 4.4.

4.2 Glomeruli Detection in Kidney Images

Immunopathology, clinical symptoms, etymology, and morphological abnormalities are all factors that go into classifying glomeruli disorders [102]. To classify the glomeruli diseases, these objects need to be detected first. Therefore, the average sensitivity and specificity of the detection matters.
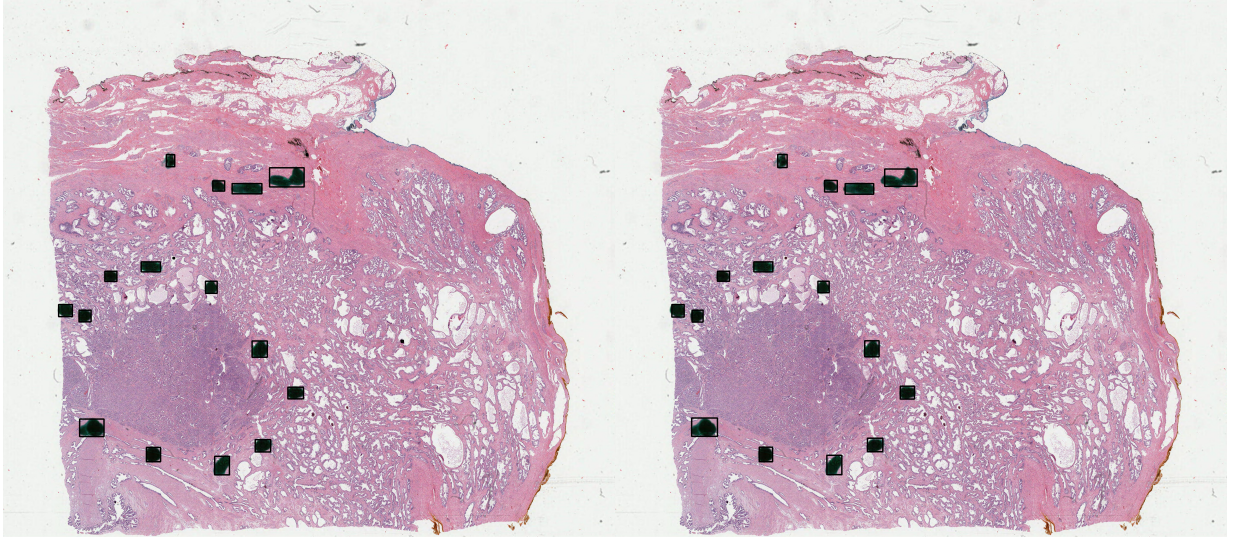
Figure 4.1: Sample ink-marked tissue detection from the dataset
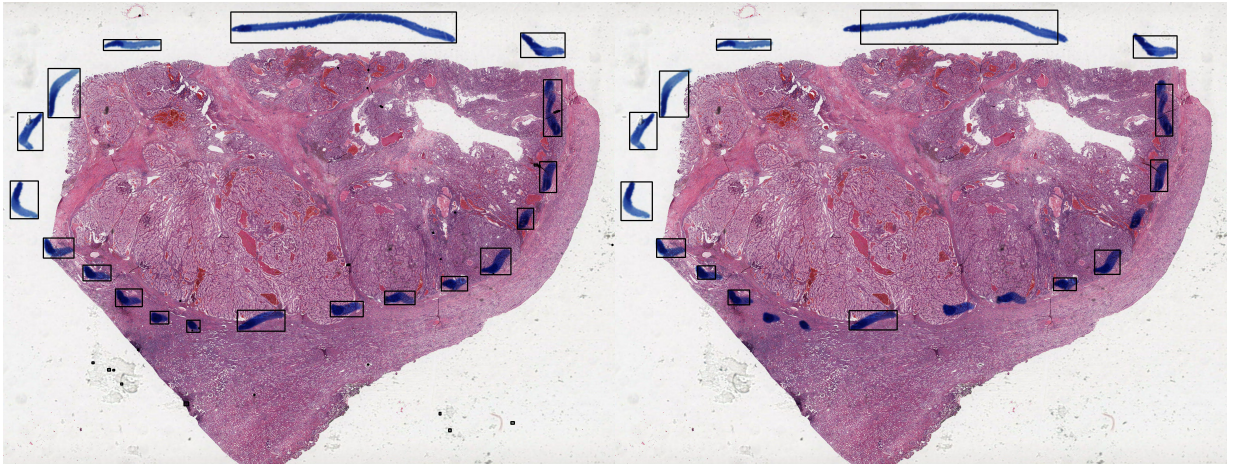
Figure 4.2: Sample ink-marked tissue detection from the dataset

By having True Positives as TP, False Positives as FP, False Negatives as FN, and True Negative as TN, The formula of sensitivity and specificity metrics [87] is as follows:

$$\text{Sensitivity} \;\; = \;\; \frac{TP}{TP + FN} \tag{4.4}$$

$$\text{Specificity} \;\; = \;\; \frac{TN}{TN + FP} \tag{4.5}$$

For computing true positives, false positives, false negatives, and true negatives, the IoU measure has been used (intersection over union) to determine the overlap between two boundaries divided by their union. Our dataset pre-defined an IoU threshold (i.e., 0.5) in classifying whether the prediction is a true positive or a false positive. Also, false negative would be those glomeruli objects that any predicted bounding boxes have not covered. Moreover, the true negatives were calculated based on the area of the whole slide tissue

37
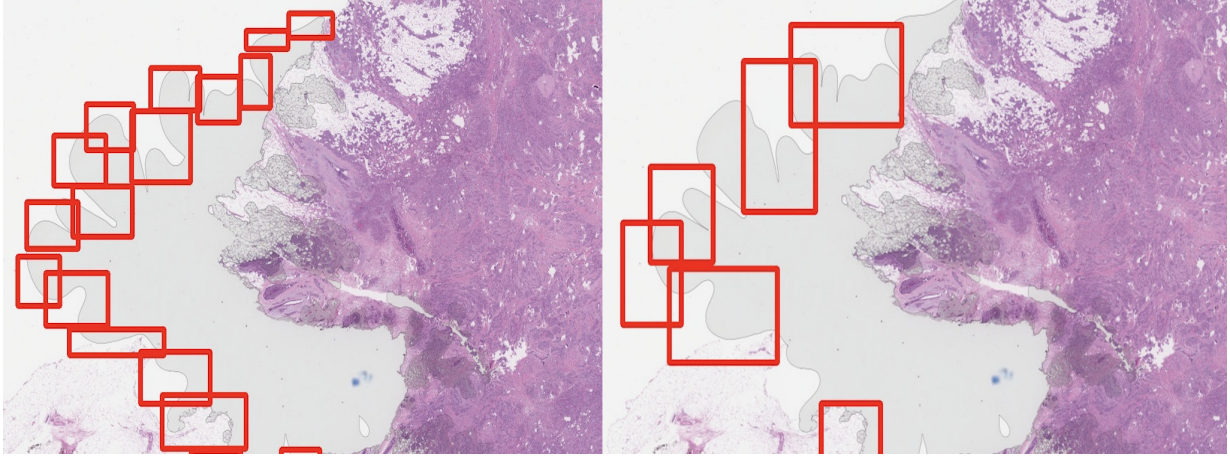
Figure 4.3: Sample air bubble tissue detection from the dataset
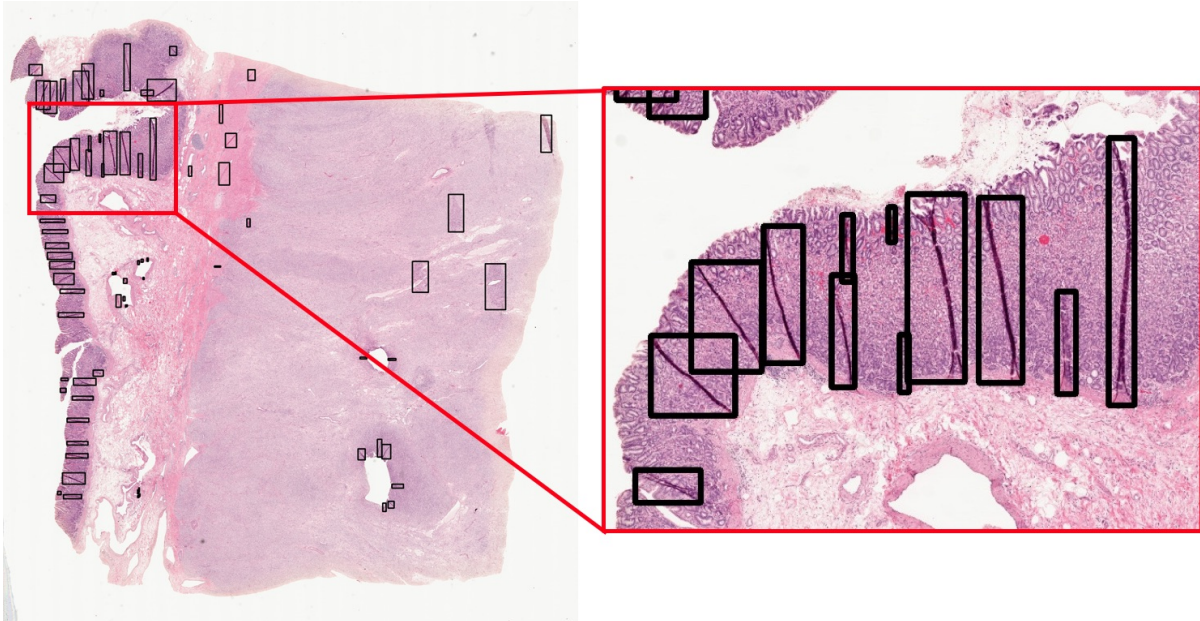
Figure 4.4: Sample tissue fold detection from the dataset

minus those predicted areas that were not containing any glomeruli.

As mentioned in the last chapter 3, 7 training datasets with two public and one private datasets have been created and validated on two datasets with different stains from the University of Michigan. In this chapter, the average sensitivity and specificity have been calculated for all these experiments for all images, along with the comparison with the existing segmentation method.

4.2.1 PAS Validation Set

Two public datasets and a private dataset from the University of Michigan, all PAS stained, were used to train YOLO and validated on 20 PAS stained images. Different experiments were designed and evaluated on these images. Average sensitivity and specificity values for each experiment can be seen in Table 4.3 along with the comparison with the segmentation method explained in Chapter 3 (used for Hacking the Kidney Competition).

The ROC (receiver operator characteristics) curves [65] for all the experiments on these 20 PAS stained images have been shown in Figure 4.5. As it has been reported in Table 4.3, the segmentation results have a high average specificity with lower sensitivity which means the network has low number of false positives. However, it can only predict half of the true negative glomeruli objects. Furthermore, using the external validation set (University of Michigan WSIs) to fine-tune this segmentation model is not feasible since the external WSIs do not contain pixel-level annotation.

Among the YOLO experiments, one experiment was done with a training set containing only 7 PAS stained images from the University of Michigan, with average sensitivity and specificity equal to 85%, and 80%, respectively, which can show the network is performing well on the validation from the same resource with limited training data. Another three experiments have been performed only on public datasets. They have been evaluated on an external validation dataset which is the data from the University of Michigan.

By examining the network on an external dataset, the generalization of the network can be assessed. Also, it is evident that after fine-tuning the network with only 7 PAS stained images from the University of Michigan on the same dataset, the average sensitivity has a considerable improvement. For example, the average sensitivity and specificity changed from 45%, and 98% to 74%, and 94% respectively. The results may significantly change if there is more data of the same resource as the validation dataset for fine-tuning the network.

Another important point would be the difference between the results of experiments trained on the first public dataset and the second one. It has been shown that by combining both datasets, the accuracy could drop off compared to only training on the first public dataset, and the reason may be related to the difference between the images from the second public dataset and the images from the University of Michigan. The images from the first public dataset and images from the University of Michigan are needle biopsy images. In contrast, the second public dataset consists of excision tissue samples. The phrase "needle biopsy" refers to a procedure in which a specific needle is inserted into a suspicious region of the skin in order to collect cells. During a "surgical biopsy", a surgeon creates an incision in your skin in order to reach the suspicious cells. As shown in Figures 3.9 and 3.10 number of glomeruli and the size of the glomeruli compared to the whole image are one of the differences between needle biopsy and surgical biopsy.

Table 4.3: Average sensitivity and average specificity were reported for seven different experiments designed with two public datasets and a private dataset from the University of Michigan (UMICH). All the PAS stained images and evaluated on 20 PAS stained images, along with the comparison with a segmentation method using U-NET.

Dataset	Average Sensitivity	Average Specificity
31 WSIs from public dataset 1	82%	95%
31 WSIs from public dataset 1, fine-tuned with 7 PAS stained WSIs from UMICH dataset	85%	89%
8 WSIs from public dataset 2	45%	98%
8 WSIs from public dataset 2, fine-tuned with 7 PAS stained WSIs from UMICH datase	74%	94%
31 WSIs from public dataset 1, and 8 WSIs from public dataset 2	75%	95%
31 WSIs from public dataset 1, and 8 WSIs from public dataset 2, fine-tuned with 7 PAS stained WSIs from UMICH dataset	83%	96%
7 PAS stained WSIs from UMICH dataset	85%	80%
Segmentation Method (HubMap Competition)	48%	99%

4.2.2 H&E Validation Set

A total of 16 H&E stained images from the University of Michigan have been used as a validation dataset for all training datasets described in the previous section. Comparison between the average sensitivity and average specificity for all seven experiments using YOLO, with two public datasets, as well as a private dataset from the University of Michigan and the segmentation method explained in Chapter 3 that was used for Hacking the Kidney Competition are provided in Table 4.4. ROC curves for all experiments on these 16 H&E stained images have been shown in Figure 4.6. There is a considerable difference between the validation results on PAS stained images and H&E stained images. This substantial difference is explainable because of the difference in tissue staining of training and validation datasets.

Same as in the Table 4.3, because of the high average specificity and low sensitivity shown in Table 4.4, the network's segmentation results practically never show false positives. However, only half of the ground truth negative glomeruli objects can be predicted by this method.

As well as the Table 4.3, the results have been improved by fine-tuning the training dataset with only seven images from the University of Michigan. The difference between
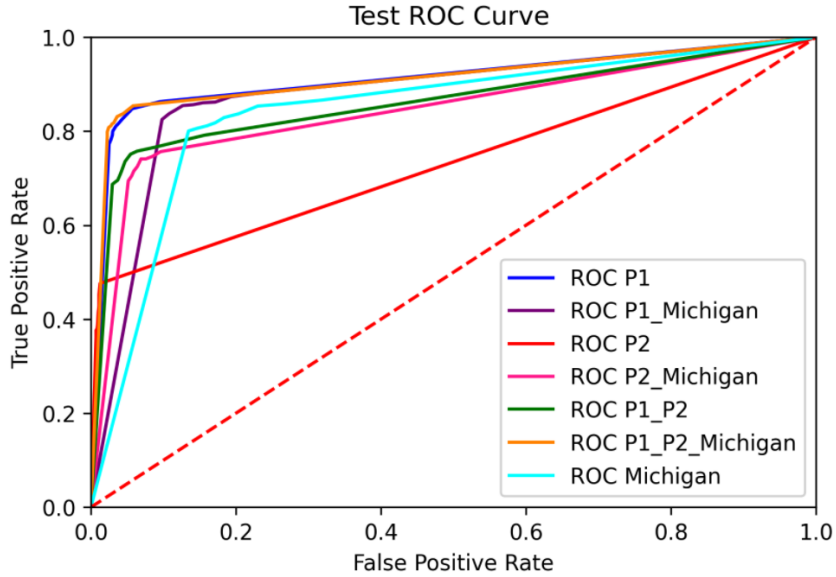
Figure 4.5: ROC curve for 20 PAS stained images, and the comparison for all designed experiments using YOLO. $P1$ indicates the first public dataset, $P2$ indicates the second public dataset, and $Michigan$ is the data from University of Michigan for fine-tuning the models.

the outcomes of experiments trained on the first public dataset and the second is still significant. Because of the differences in images between the second dataset which are surgical biopsy images and those from the University of Michigan that are needle biopsy images, it has been demonstrated that by combining both datasets, accuracy can drop.

Table 4.4: Average sensitivity, and average specificity reported for different seven experiments designed with two public datasets and a private dataset from the University of Michigan (UMICH), all PAS stained and evaluated on 16 H&E stained images, along with the comparison with a segmentation method using U-NET

| Dataset	Average Sensitivity	Average Specificity
31 WSIs from public dataset 1	51%	95%
31 WSIs from public dataset 1, fine-tuned with 7 PAS stained WSIs from UMICH dataset	67%	89%
8 WSIs from public dataset 2	30%	85%
8 WSIs from public dataset 2, fine-tuned with 7 PAS stained WSIs from UMICH datase	59%	90%
31 WSIs from public dataset 1, and 8 WSIs from public dataset 2	58%	94%
31 WSIs from public dataset 1, and 8 WSIs from public dataset 2, fine-tuned with 7 PAS stained WSIs from UMICH dataset	70%	96%
7 PAS stained WSIs from UMICH dataset	70%	86%
Segmentation Method (HubMap Competition)	47%	99%
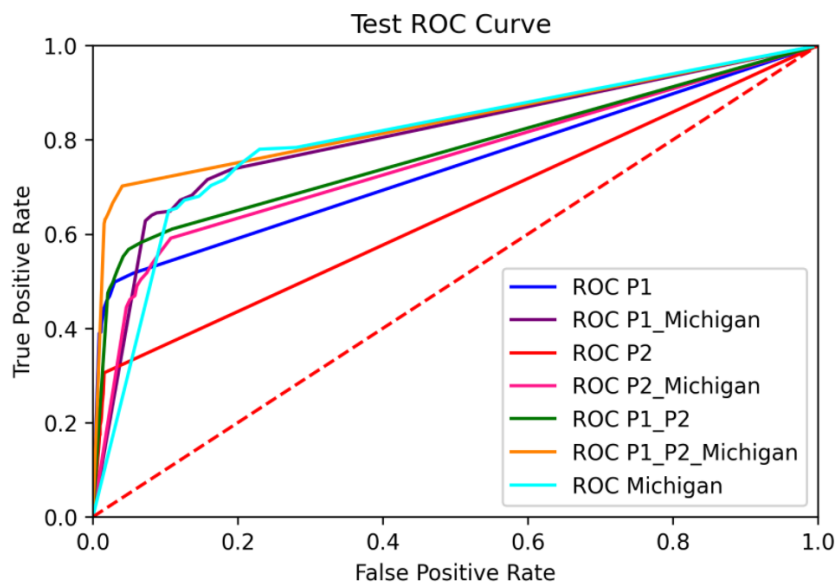
Figure 4.6: ROC curve for 16 H&E stained images, and the comparison for all the designed experiments using YOLO. $P1$ indicates the first public dataset, $P2$ indicates the second public dataset, and $Michigan$ is the data from University of Michigan for fine-tuning the models

Chapter 5

Summary & Conclusions

There have been several technological advances across health care and digital pathology in recent years. Automated segmentation and pixel analysis of digital pathology images may identify diagnostic patterns and visual cues, leading to more reliable and consistent diagnostic categorization. In this thesis, different detection and segmentation methods have been investigated for various applications. Specifically, the YOLO-v4 framework (You-Only-Look-Once), a real-time object detector for microscopic images, has been studied for two different applications: Quality control of WSIs, and glomeruli object detection in kidney images.

Quality control via artifacts and ink-marker detection in histopathology slides could save time in clinical practice by avoiding re-scanning glass slides. Tissue glass slides with these artifacts may affect the diagnosis and will be mostly rejected if detected by the laboratory staff. In this thesis, a procedure based on the YOLO-V4 deep model was proposed to detect folded and ink-marked tissue and the presence of air bubbles in tissue in large scan regions. The YOLO-V4 model has been trained on a in-house curated dataset with high accuracy. The approach is very fast, detecting the target objects in less than 0.5 seconds per image. For future works, a larger dataset with a broader representation of more artifacts can boost the generalization of quality control in digital pathology.

Glomeruli detection, as the first step of classifying the glomeruli diseases following by diagnosing different kidney diseases, is essential and critical in digital pathology. Because of the large number of these objects in the kidney, glomeruli detection could help pathologists save considerable time by computerized quantification. This thesis trained YOLO-v4 with seven different training datasets consisting of two public datasets and a private dataset from the University of Michigan. Moreover, the networks were evaluated on 20 PAS stained images and 16 H&E stained images from the University of Michigan. By training YOLO-v4 on the first public dataset, and fine-tuning by only 7 PAS stained images from the University of Michigan, experiments achieved 85% average sensitivity and 89% average specificity while validating the network on 20 PAS stained images from the University of Michigan, which was the best result out of different training datasets. For evaluating the network on H&E stained images, 70% average sensitivity and 96% average specificity were obtained while training on both public datasets, followed by fine-tuning on the 7 PAS

stained images. Also, final weights of a segmentation method based on U-Net have been used to evaluate the results on the same validation datasets. The model could achieve high specificity and lower sensitivity, making this method rather unreliable compared to YOLO with higher sensitivity. Moreover, obtaining pixel-level WSI annotations for the network is time-consuming. This makes the whole procedure for fine-tuning the model with limited data harder than detection methods like YOLO, which only requires a bounding box around the target objects.

References

[1] Mahmoud Abdolhoseini, Murielle G Kluge, Frederick R Walker, and Sarah J Johnson. Segmentation of heavily clustered nuclei from histopathological images. *Scientific reports*, 9(1):1–13, 2019.

[2] Umesh Adiga, Brian L Bell, Larissa Ponomareva, Debbie Taylor, Roland Saldanha, Sandra Nelson, and Thomas J Lamkin. Mapping infected cell phenotype. *IEEE transactions on biomedical engineering*, 59(8):2362–2371, 2012.

[3] Umesh Adiga, Ravi Malladi, Rodrigo Fernandez-Gonzalez, and C Ortiz de Solorzano. High-throughput analysis of multispectral images of breast cancer tissue. *IEEE Transactions on Image Processing*, 15(8):2259–2268, 2006.

[4] Sharib Ali, Nasullah Khalid Alham, Clare Verrill, and Jens Rittscher. Ink removal from histopathology whole slide images by combining classification, detection and image generation models. In *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, pages 928–932. IEEE, 2019.

[5] Nicola Altini, Giacomo Donato Cascarano, Antonio Brunetti, Irio De Feudis, Domenico Buongiorno, Michele Rossini, Francesco Pesce, Loreto Gesualdo, and Vitoantonio Bevilacqua. A deep learning instance segmentation approach for global glomerulosclerosis assessment in donor kidney biopsies. *Electronics*, 9(11):1768, 2020.

[6] Nicola Altini, Giacomo Donato Cascarano, Antonio Brunetti, Francescomaria Marino, Maria Teresa Rocchetti, Silvia Matino, Umberto Venere, Michele Rossini, Francesco Pesce, Loreto Gesualdo, et al. Semantic segmentation framework for glomeruli detection and classification in kidney histological sections. *Electronics*, 9(3):503, 2020.

[7] Hani A Alturkistani, Faris M Tashkandi, and Zuhair M Mohammedsaleh. Histological stains: a literature review and case study. *Global journal of health science*, 8(3):72, 2016.

[8] T Ahmad Aris, AS Abdul Nasir, and WA Mustafa. Analysis of distance transforms for watershed segmentation on chronic leukaemia images. *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, 10(1-16):51–56, 2018.

[9] Morteza Babaie and Hamid R Tizhoosh. Deep features for tissue-fold detection in histopathology images. In *European Congress on Digital Pathology*, pages 125–132. Springer, 2019.

[10] Qanita Bani Baker, Sajda Banat, Eman Eaydat, Mohammad Alsmirat, et al. Automated detection of benign and malignant in breast histopathology images. In *2018 IEEE/ACS 15th International Conference on Computer Systems and Applications (AICCSA)*, pages 1–5. IEEE, 2018.

[11] Partha Pratim Banik, Rappy Saha, and Ki-Doo Kim. An automatic nucleus segmentation and cnn model based classification method of white blood cell. *Expert Systems with Applications*, 149:113211, 2020.

[12] Laura Barisoni, Kyle J Lafata, Stephen M Hewitt, Anant Madabhushi, and Ulysses GJ Balis. Digital pathology and computational image analysis in nephropathology. *Nature Reviews Nephrology*, 16(11):669–685, 2020.

[13] P. Bautista and Y. Yagi. Detection of tissue folds in whole slide images. *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 3669–3672, 2009.

[14] Sean Bell, C Lawrence Zitnick, Kavita Bala, and Ross Girshick. Inside-outside net: Detecting objects in context with skip pooling and recurrent neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2874–2883, 2016.

[15] Maxim Berman, Amal Rannen Triki, and Matthew B Blaschko. The lovász-softmax loss: A tractable surrogate for the optimization of the intersection-over-union measure in neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4413–4421, 2018.

[16] Subrata Bhattacharjee, Hyeon-Gyun Park, Cho-Hee Kim, Deekshitha Prakash, Nuwan Madusanka, Jae-Hong So, Nam-Hoon Cho, and Heung-Kook Choi. Quantitative analysis of benign and malignant tumors in histopathology: Predicting prostate cancer grading using svm. *Applied Sciences*, 9(15):2969, 2019.

[17] Christopher M Bishop and Nasser M Nasrabadi. *Pattern recognition and machine learning*, volume 4. Springer, 2006.

[18] Alexey Bochkovskiy, Chien-Yao Wang, and Hong-Yuan Mark Liao. Yolov4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*, 2020.

[19] Michael Buckland and Fredric Gey. The relationship between recall and precision. *Journal of the American society for information science*, 45(1):12–19, 1994.

[20] Elzbieta Budginaitė, Mindaugas Morkūnas, Arvydas Laurinavičius, and Povilas Treigys. Deep learning model for cell nuclei segmentation and lymphocyte identification in whole slide histology images. *Informatica*, 32(1):23–40, 2021.

[21] Gloria Bueno, M Milagro Fernandez-Carrobles, Lucia Gonzalez-Lopez, and Oscar Deniz. Glomerulosclerosis identification in whole slide images using semantic segmentation. *Computer methods and programs in biomedicine*, 184:105273, 2020.

[22] Gloria Bueno, Lucia Gonzalez-Lopez, Marcial Garcia-Rojo, Arvydas Laurinavicius, and Oscar Deniz. Data for glomeruli characterization in histopathological images. *Data in brief*, 29:105314, 2020.

[23] 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.

[24] Jiyun Byun, Mark R Verardo, Baris Sumengen, Geoffrey P Lewis, BS Manjunath, and Steven K Fisher. Automated tool for the detection of cell nuclei in digital microscopic images: application to retinal images. *Mol Vis*, 12(105-07):949–60, 2006.

[25] Cristina Callau, Marylène Lejeune, Anna Korzynska, Marcial García, Gloria Bueno, Ramon Bosch, Joaquín Jaén, Guifré Orero, Teresa Salvadó, and Carlos López. Evaluation of cytokeratin-19 in breast cancer tissue samples: a comparison of automatic and manual evaluations of scanned tissue microarray cylinders. *BioMedical Engineering OnLine*, 14(2):1–11, 2015.

[26] Maoyong Cao, Shuang Wang, Lu Wei, Laxmisha Rai, Dong Li, Hui Yu, and Dan Shao. Segmentation of immunohistochemical image of lung neuroendocrine tumor based on double layer watershed. *Multimedia Tools and Applications*, 78(7):9193–9215, 2019.

[27] Giacomo Donato Cascarano, Francesco Saverio Debitonto, Ruggero Lemma, Antonio Brunetti, Domenico Buongiorno, Irio De Feudis, Andrea Guerriero, Michele Rossini, Francesco Pesce, Loreto Gesualdo, et al. An innovative neural network framework for glomerulus classification based on morphological and texture features evaluated in histological images of kidney biopsy. In *International Conference on Intelligent Computing*, pages 727–738. Springer, 2019.

[28] Mustafa Çetin, Zümray Dokur, and Tamer Ölmez. Fuzzy local information c-means algorithm for histopathological image segmentation. In *2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT)*, pages 1–6. IEEE, 2019.

[29] John KC Chan. The wonderful colors of the hematoxylin–eosin stain in diagnostic surgical pathology. *International journal of surgical pathology*, 22(1):12–32, 2014.

[30] Germán Corredor, Xiangxue Wang, Cheng Lu, Vamsidhar Velcheti, Eduardo Romero, and Anant Madabhushi. A watershed and feature-based approach for automated detection of lymphocytes on lung cancer images. In *Medical Imaging 2018: Digital Pathology*, volume 10581, page 105810R. International Society for Optics and Photonics, 2018.

[31] Eric Cosatto, Matt Miller, Hans Peter Graf, and John S Meyer. Grading nuclear pleomorphism on histological micrographs. In *2008 19th International Conference on Pattern Recognition*, pages 1–4. IEEE, 2008.

[32] Yuxin Cui, Guiying Zhang, Zhonghao Liu, Zheng Xiong, and Jianjun Hu. A deep learning algorithm for one-step contour aware nuclei segmentation of histopathology images. *Medical & biological engineering & computing*, 57(9):2027–2043, 2019.

[33] Jifeng Dai, Kaiming He, and Jian Sun. Instance-aware semantic segmentation via multi-task network cascades. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3150–3158, 2016.

[34] Jifeng Dai, Yi Li, Kaiming He, and Jian Sun. R-fcn: Object detection via region-based fully convolutional networks. *Advances in neural information processing systems*, 29, 2016.

[35] Navneet Dalal and Bill Triggs. Histograms of oriented gradients for human detection. In *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, volume 1, pages 886–893. Ieee, 2005.

[36] Ricard Delgado-Gonzalo, Virginie Uhlmann, Daniel Schmitter, and Michael Unser. Snakes on a plane: A perfect snap for bioimage analysis. *IEEE Signal Processing Magazine*, 32(1):41–48, 2014.

[37] Leyza Baldo Dorini, Rodrigo Minetto, and Neucimar Jeronimo Leite. Semiautomatic white blood cell segmentation based on multiscale analysis. *IEEE journal of biomedical and health informatics*, 17(1):250–256, 2012.

[38] Dalí FD dos Santos, Thaína AA Tosta, Adriano B Silva, Paulo R de Faria, Bruno AN Travençolo, and Marcelo Z do Nascimento. Automated nuclei segmentation on dysplastic oral tissues using cnn. In *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*, pages 45–50. IEEE, 2020.

[39] Edward R Dougherty. *Digital image processing methods*. CRC Press, 2020.

[40] Dennis Eschweiler, Thiago V Spina, Rohan C Choudhury, Elliot Meyerowitz, Alexandre Cunha, and Johannes Stegmaier. Cnn-based preprocessing to optimize watershed-based cell segmentation in 3d confocal microscopy images. In *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, pages 223–227. IEEE, 2019.

[41] Andi Rezky Fatmasari and Ingrid Nurtanio. A colour space based detection for cervical cancer using fuzzy c-means clustering. In *Proceedings of the 6th International Conference on Bioinformatics and Biomedical Science*, pages 137–141, 2017.

[42] Pedro F. Felzenszwalb, Ross B. Girshick, David McAllester, and Deva Ramanan. Object detection with discriminatively trained part-based models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(9):1627–1645, 2010.

[43] Thomas J Fuchs and Joachim M Buhmann. Computational pathology: challenges and promises for tissue analysis. *Computerized Medical Imaging and Graphics*, 35(7-8):515–530, 2011.

[44] Nidhal Tawfeeq Furat, Nada AS Alwan, and Basim M Khashman. Optimization of digital histopathology image quality. *IAES International Journal of Artificial Intelligence*, 7(2):71, 2018.

[45] Jaime Gallego, Anibal Pedraza, Samuel Lopez, Georg Steiner, Lucia Gonzalez, Arvydas Laurinavicius, and Gloria Bueno. Glomerulus classification and detection based on convolutional neural networks. *Journal of Imaging*, 4(1):20, 2018.

[46] Alberto Garcia-Garcia, Sergio Orts-Escolano, Sergiu Oprea, Victor Villena-Martinez, and Jose Garcia-Rodriguez. A review on deep learning techniques applied to semantic segmentation. *arXiv preprint arXiv:1704.06857*, 2017.

[47] Srishti Gautam, Arnav Bhavsar, Anil K Sao, and KK Harinarayan. Cnn based segmentation of nuclei in pap-smear images with selective pre-processing. In *Medical Imaging 2018: Digital Pathology*, volume 10581, page 105810X. International Society for Optics and Photonics, 2018.

[48] Jose George, Shibon Skaria, VV Varun, et al. Using yolo based deep learning network for real time detection and localization of lung nodules from low dose ct scans. In *Medical Imaging 2018: Computer-Aided Diagnosis*, volume 10575, page 105751I. International Society for Optics and Photonics, 2018.

[49] Ross Girshick. Fast r-cnn. In *Proceedings of the IEEE international conference on computer vision*, pages 1440–1448, 2015.

[50] Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 580–587, 2014.

[51] Ross B. Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pages 580–587, 2014.

[52] Leah L Godwin, Yingnan Ju, Naveksha Sood, Yashvardhan Jain, Ellen M Quardokus, Andreas Bueckle, Teri Longacre, Aaron Horning, Yiing Lin, Edward D Esplin, et al. Robust and generalizable segmentation of human functional tissue units. *bioRxiv*, 2021.

[53] Rafael C Gonzalez and Richard E Woods. Digital image processing, prentice hall. *Upper Saddle River, NJ*, 2008.

[54] Rafael C Gonzalez, Richard E Woods, and Barry R Masters. Digital image processing, 2009.

[55] Vicente Grau, AUJ Mewes, M Alcaniz, Ron Kikinis, and Simon K Warfield. Improved watershed transform for medical image segmentation using prior information. *IEEE transactions on medical imaging*, 23(4):447–458, 2004.

[56] Peng Guo, Koyel Banerjee, R Joe Stanley, Rodney Long, Sameer Antani, George Thoma, Rosemary Zuna, Shelliane R Frazier, Randy H Moss, and William V Stoecker. Nuclei-based features for uterine cervical cancer histology image analysis with fusion-based classification. *IEEE journal of biomedical and health informatics*, 20(6):1595–1607, 2015.

[57] Metin N Gurcan, Laura E Boucheron, Ali Can, Anant Madabhushi, Nasir M Rajpoot, and Bulent Yener. Histopathological image analysis: A review. *IEEE reviews in biomedical engineering*, 2:147–171, 2009.

[58] Peter W Hamilton, Peter Bankhead, Yinhai Wang, Ryan Hutchinson, Declan Kieran, Darragh G McArt, Jacqueline James, and Manuel Salto-Tellez. Digital pathology and image analysis in tissue biomarker research. *Methods*, 70(1):59–73, 2014.

[59] Nathalie Harder, Megan Bodnar, Roland Eils, David L Spector, and Karl Rohr. 3d segmentation and quantification of mouse embryonic stem cells in fluorescence microscopy images. In *2011 IEEE international symposium on biomedical imaging: From nano to macro*, pages 216–219. IEEE, 2011.

[60] Trevor Hastie, Robert Tibshirani, Jerome H Friedman, and Jerome H Friedman. *The elements of statistical learning: data mining, inference, and prediction*, volume 2. Springer, 2009.

[61] Tomohiro Hayakawa, VB Prasath, Hiroharu Kawanaka, Bruce J Aronow, and Shinji Tsuruoka. Computational nuclei segmentation methods in digital pathology: a survey. *Archives of Computational Methods in Engineering*, 28(1):1–13, 2021.

[62] 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.

[63] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Identity mappings in deep residual networks. In *European conference on computer vision*, pages 630–645. Springer, 2016.

[64] Paul Henderson and Vittorio Ferrari. End-to-end training of object class detectors for mean average precision. In *Asian conference on computer vision*, pages 198–213. Springer, 2016.

[65] Zhe Hui Hoo, Jane Candlish, and Dawn Teare. What is an roc curve?, 2017.

[66] Andrew Howard, Mark Sandler, Grace Chu, Liang-Chieh Chen, Bo Chen, Mingxing Tan, Weijun Wang, Yukun Zhu, Ruoming Pang, Vijay Vasudevan, et al. Searching for mobilenetv3. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 1314–1324, 2019.

[67] 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.

[68] 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.

[69] Jonathan Huang, Vivek Rathod, Chen Sun, Menglong Zhu, Anoop Korattikara, Alireza Fathi, Ian Fischer, Zbigniew Wojna, Yang Song, Sergio Guadarrama, et al. Speed/accuracy trade-offs for modern convolutional object detectors. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7310–7311, 2017.

[70] Forrest N Iandola, Song Han, Matthew W Moskewicz, Khalid Ashraf, William J Dally, and Kurt Keutzer. Squeezenet: Alexnet-level accuracy with 50x fewer parameters and¡ 0.5 mb model size. *arXiv preprint arXiv:1602.07360*, 2016.

[71] Humayun Irshad, Antoine Veillard, Ludovic Roux, and Daniel Racoceanu. Methods for nuclei detection, segmentation, and classification in digital histopathology: a review—current status and future potential. *IEEE reviews in biomedical engineering*, 7:97–114, 2013.

[72] Paul T. Jackway and Mohamed Deriche. Scale-space properties of the multiscale morphological dilation-erosion. *IEEE transactions on pattern analysis and machine intelligence*, 18(1):38–51, 1996.

[73] Colin Jacobs, Eva M Van Rikxoort, Thorsten Twellmann, Ernst Th Scholten, Pim A De Jong, Jan-Martin Kuhnigk, Matthijs Oudkerk, Harry J De Koning, Mathias Prokop, Cornelia Schaefer-Prokop, et al. Automatic detection of subsolid pulmonary nodules in thoracic computed tomography images. *Medical image analysis*, 18(2):374–384, 2014.

[74] Andrew Janowczyk, Ren Zuo, Hannah Gilmore, Michael Feldman, and Anant Madabhushi. Histoqc: an open-source quality control tool for digital pathology slides. *JCO clinical cancer informatics*, 3:1–7, 2019.

[75] Dongyao Jia, Chuanwang Zhang, Nengkai Wu, Zhigang Guo, and Hairui Ge. Multilayer segmentation framework for cell nuclei using improved gvf snake model, watershed, and ellipse fitting. *Biomedical Signal Processing and Control*, 67:102516, 2021.

[76] Zeng Jiankui, He Zishu, Mathini Sellathurai, and Liu Hongming. Modified hough transform for searching radar detection. *IEEE geoscience and remote sensing letters*, 5(4):683–686, 2008.

[77] Xin Jin and Jiawei Han. *K-Means Clustering*, pages 563–564. Springer US, Boston, MA, 2010.

[78] Tsuyoshi Kato, Raissa Relator, Hayliang Ngouv, Yoshihiro Hirohashi, Osamu Takaki, Tetsuhiro Kakimoto, and Kinya Okada. Segmental hog: new descriptor for glomerulus detection in kidney microscopy image. *Bmc Bioinformatics*, 16(1):1–16, 2015.

[79] Yoshimasa Kawazoe, Kiminori Shimamoto, Ryohei Yamaguchi, Yukako Shintani-Domoto, Hiroshi Uozaki, Masashi Fukayama, and Kazuhiko Ohe. Faster r-cnn-based glomerular detection in multistained human whole slide images. *Journal of Imaging*, 4(7):91, 2018.

[80] Stephen J Keenan, James Diamond, W Glenn McCluggage, Hoshang Bharucha, Deborah Thompson, Peter H Bartels, and Peter W Hamilton. An automated machine vision system for the histological grading of cervical intraepithelial neoplasia (cin). *The Journal of pathology*, 192(3):351–362, 2000.

[81] Mahendra Khened, Avinash Kori, Haran Rajkumar, Ganapathy Krishnamurthi, and Balaji Srinivasan. A generalized deep learning framework for whole-slide image segmentation and analysis. *Scientific reports*, 11(1):1–14, 2021.

[82] Jun Kong, Fusheng Wang, George Teodoro, Yanhui Liang, Yangyang Zhu, Carol Tucker-Burden, and Daniel J Brat. Automated cell segmentation with 3d fluorescence microscopy images. In *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, pages 1212–1215. IEEE, 2015.

[83] Marek Kowal, Michał Żejmo, Marcin Skobel, Józef Korbicz, and Roman Monczak. Cell nuclei segmentation in cytological images using convolutional neural network and seeded watershed algorithm. *Journal of Digital Imaging*, 33(1):231–242, 2020.

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

[85] Rajesh Kumar, Subodh Srivastava, and Rajeev Srivastava. A fourth order pde based fuzzy c-means approach for segmentation of microscopic biopsy images in presence of poisson noise for cancer detection. *Computer methods and programs in biomedicine*, 146:59–68, 2017.

[86] Shyam Lal, Russel Desouza, M Maneesh, Anirudh Kanfade, Aman Kumar, Gokul Perayil, Kumar Alabhya, Amit Kumar Chanchal, and Jyoti Kini. A robust method for nuclei segmentation of h&e stained histopathology images. In *2020 7th International Conference on Signal Processing and Integrated Networks (SPIN)*, pages 453–458. IEEE, 2020.

[87] Abdul Ghaaliq Lalkhen and Anthony McCluskey. Clinical tests: sensitivity and specificity. *Continuing education in anaesthesia critical care & pain*, 8(6):221–223, 2008.

[88] David Ledbetter, Long Ho, and Kevin V Lemley. Prediction of kidney function from biopsy images using convolutional neural networks. *arXiv preprint arXiv:1702.01816*, 2017.

[89] Yi Li, Haozhi Qi, Jifeng Dai, Xiangyang Ji, and Yichen Wei. Fully convolutional instance-aware semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2359–2367, 2017.

[90] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision*, pages 2980–2988, 2017.

[91] Tony Lindeberg. Feature detection with automatic scale selection. *International journal of computer vision*, 30(2):79–116, 1998.

[92] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C Berg. Ssd: Single shot multibox detector. In *European conference on computer vision*, pages 21–37. Springer, 2016.

[93] David G Lowe. Object recognition from local scale-invariant features. In *Proceedings of the seventh IEEE international conference on computer vision*, volume 2, pages 1150–1157. Ieee, 1999.

[94] Gareth Loy and Alexander Zelinsky. Fast radial symmetry for detecting points of interest. *IEEE Transactions on pattern analysis and machine intelligence*, 25(8):959–973, 2003.

[95] Zhi Lu, Gustavo Carneiro, and Andrew P Bradley. Automated nucleus and cytoplasm segmentation of overlapping cervical cells. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 452–460. Springer, 2013.

[96] Zhi Lu, Gustavo Carneiro, and Andrew P Bradley. An improved joint optimization of multiple level set functions for the segmentation of overlapping cervical cells. *IEEE Transactions on Image Processing*, 24(4):1261–1272, 2015.

[97] Ningning Ma, Xiangyu Zhang, Hai-Tao Zheng, and Jian Sun. Shufflenet v2: Practical guidelines for efficient cnn architecture design. In *Proceedings of the European conference on computer vision (ECCV)*, pages 116–131, 2018.

[98] Anant Madabhushi and George Lee. Image analysis and machine learning in digital pathology: Challenges and opportunities. *Medical image analysis*, 33:170–175, 2016.

[99] Danial Maleki, Mehdi Afshari, Morteza Babaie, and Hamid R Tizhoosh. Ink marker segmentation in histopathology images using deep learning. In *International Symposium on Visual Computing*, pages 359–368. Springer, 2020.

[100] Jon N Marsh, Matthew K Matlock, Satoru Kudose, Ta-Chiang Liu, Thaddeus S Stappenbeck, Joseph P Gaut, and S Joshua Swamidass. Deep learning global glomerulosclerosis in transplant kidney frozen sections. *IEEE transactions on medical imaging*, 37(12):2718–2728, 2018.

[101] Hela Masmoudi, Stephen M Hewitt, Nicholas Petrick, Kyle J Myers, and Marios A Gavrielides. Automated quantitative assessment of her-2/neu immunohistochemical expression in breast cancer. *IEEE transactions on medical imaging*, 28(6):916–925, 2009.

[102] Antonio Mastrangelo, Jessica Serafinelli, Marisa Giani, and Giovanni Montini. Clinical and pathophysiological insights into immunological mediated glomerular diseases in childhood. *Frontiers in Pediatrics*, 8:205, 2020.

[103] Jiri Matas, Ondrej Chum, Martin Urban, and Tomás Pajdla. Robust wide-baseline stereo from maximally stable extremal regions. *Image and vision computing*, 22(10):761–767, 2004.

[104] Calvin R Maurer, Rensheng Qi, and Vijay Raghavan. A linear time algorithm for computing exact euclidean distance transforms of binary images in arbitrary dimensions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(2):265–270, 2003.

[105] 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, 2015.

[106] Erik Meijering. Cell segmentation: 50 years down the road [life sciences]. *IEEE signal processing magazine*, 29(5):140–145, 2012.

[107] Keelin Murphy, Bram van Ginneken, Arnold MR Schilham, BJ De Hoop, Hester A Gietema, and Mathias Prokop. A large-scale evaluation of automatic pulmonary nodule detection in chest ct using local image features and k-nearest-neighbour classification. *Medical image analysis*, 13(5):757–770, 2009.

[108] Peter Naylor, Marick Laé, Fabien Reyal, and Thomas Walter. Nuclei segmentation in histopathology images using deep neural networks. In *2017 IEEE 14th international symposium on biomedical imaging (ISBI 2017)*, pages 933–936. IEEE, 2017.

[109] 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.

[110] Jie Ni, Maneesh K Singh, and Claus Bahlmann. Fast radial symmetry detection under affine transformations. In *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pages 932–939. IEEE, 2012.

[111] C Ortiz de Solorzano, E Garcia Rodriguez, Arthur Jones, Dan Pinkel, Joe W Gray, Damir Sudar, and Stephen J Lockett. Segmentation of confocal microscope images of cell nuclei in thick tissue sections. *Journal of Microscopy*, 193(3):212–226, 1999.

[112] Nobuyuki Otsu. A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics*, 9(1):62–66, 1979.

[113] Jiangmiao Pang, Kai Chen, Jianping Shi, Huajun Feng, Wanli Ouyang, and Dahua Lin. Libra r-cnn: Towards balanced learning for object detection. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 821–830, 2019.

[114] Constantine P Papageorgiou, Michael Oren, and Tomaso Poggio. A general framework for object detection. In *Sixth International Conference on Computer Vision (IEEE Cat. No. 98CH36271)*, pages 555–562. IEEE, 1998.

[115] Chiwoo Park, Jianhua Z Huang, Jim X Ji, and Yu Ding. Segmentation, inference and classification of partially overlapping nanoparticles. *IEEE transactions on pattern analysis and machine intelligence*, 35(3):1–1, 2012.

[116] Jong Wook Park and Sang Uk Lee. Recovery of corrupted image data based on the nurbs interpolation. *IEEE Transactions on Circuits and Systems for Video Technology*, 9(7):1003–1008, 1999.

[117] Huiming Peng, Xiaobo Zhou, Fuhai Li, Xiaofeng Xia, and Stephen TC Wong. Integrating multi-scale blob/curvilinear detector techniques and multi-level sets for automated segmentation of stem cell images. In *2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 1362–1365. IEEE, 2009.

[118] Marina E Plissiti and Christophoros Nikou. Overlapping cell nuclei segmentation using a spatially adaptive active physical model. *IEEE Transactions on Image Processing*, 21(11):4568–4580, 2012.

[119] Justin Tyler Pontalba, Thomas Gwynne-Timothy, Ephraim David, Kiran Jakate, Dimitrios Androutsos, and April Khademi. Assessing the impact of color normalization in convolutional neural network-based nuclei segmentation frameworks. *Frontiers in bioengineering and biotechnology*, page 300, 2019.

[120] Xin Qi, Fuyong Xing, David J Foran, and Lin Yang. Robust segmentation of overlapping cells in histopathology specimens using parallel seed detection and repulsive level set. *IEEE Transactions on Biomedical Engineering*, 59(3):754–765, 2011.

[121] Aiping Qu, Zhiming Cheng, Xiaofeng He, and Yue Li. An integration convolutional neural network for nuclei instance segmentation. In *2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 1104–1109. IEEE, 2020.

[122] Hui Qu, Pengxiang Wu, Qiaoying Huang, Jingru Yi, Gregory M Riedlinger, Subhajyoti De, and Dimitris N Metaxas. Weakly supervised deep nuclei segmentation

using points annotation in histopathology images. In *International Conference on Medical Imaging with Deep Learning*, pages 390–400. PMLR, 2019.

[123] Francesco Raimondo, Marios A Gavrielides, Georgia Karayannopoulou, Kleoniki Lyroudia, Ioannis Pitas, and Ioannis Kostopoulos. Automated evaluation of her-2/neu status in breast tissue from fluorescent in situ hybridization images. *IEEE Transactions on Image Processing*, 14(9):1288–1299, 2005.

[124] Maral Rasoolijaberi. Multi-magnification search in digital pathology. Master's thesis, University of Waterloo, 2021.

[125] Shamala Ravikumar, R Surekha, Rooban Thavarajah, et al. Mounting media: An overview. *Journal of Dr. NTR University of Health Sciences*, 3(5):1, 2014.

[126] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 779–788, 2016.

[127] Joseph Redmon and Ali Farhadi. Yolo9000: better, faster, stronger. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7263–7271, 2017.

[128] Joseph Redmon and Ali Farhadi. Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767*, 2018.

[129] Daniel Reisfeld, Haim Wolfson, and Yehezkel Yeshurun. Context-free attentional operators: The generalized symmetry transform. *International Journal of Computer Vision*, 14(2):119–130, 1995.

[130] Daniel Reisfeld and Yehezkel Yeshurun. Preprocessing of face images: Detection of features and pose normalization. *Computer vision and image understanding*, 71(3):413–430, 1998.

[131] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems*, 28, 2015.

[132] Stephen Robertson. A new interpretation of average precision. In *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*, pages 689–690, 2008.

[133] Roscoe R Robinson. *Renal disease in children: Clinical evaluation and diagnosis*. Springer Science & Business Media, 2012.

[134] Jos BTM Roerdink and Arnold Meijster. The watershed transform: Definitions, algorithms and parallelization strategies. *Fundamenta informaticae*, 41(1, 2):187–228, 2000.

[135] 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.

[136] Holger R Roth, Le Lu, Amal Farag, Hoo-Chang Shin, Jiamin Liu, Evrim B Turkbey, and Ronald M Summers. Deeporgan: Multi-level deep convolutional networks for automated pancreas segmentation. In *International conference on medical image computing and computer-assisted intervention*, pages 556–564. Springer, 2015.

[137] Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, and Liang-Chieh Chen. Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4510–4520, 2018.

[138] Oliver Schmitt and Maria Hasse. Morphological multiscale decomposition of connected regions with emphasis on cell clusters. *Computer Vision and Image Understanding*, 113(2):188–201, 2009.

[139] Pierre Sermanet, David Eigen, Xiang Zhang, Michaël Mathieu, Rob Fergus, and Yann LeCun. Overfeat: Integrated recognition, localization and detection using convolutional networks. *arXiv preprint arXiv:1312.6229*, 2013.

[140] Abhinav Shrivastava and Abhinav Gupta. Contextual priming and feedback for faster r-cnn. In *European conference on computer vision*, pages 330–348. Springer, 2016.

[141] Abhinav Shrivastava, Abhinav Gupta, and Ross Girshick. Training region-based object detectors with online hard example mining. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 761–769, 2016.

[142] Olivier Simon, Rabi Yacoub, Sanjay Jain, John E Tomaszewski, and Pinaki Sarder. Multi-radial lbp features as a tool for rapid glomerular detection and assessment in whole slide histopathology images. *Scientific reports*, 8(1):1–11, 2018.

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

[144] Mohamed Slaoui and Laurence Fiette. Histopathology procedures: from tissue sampling to histopathological evaluation. In *Drug safety evaluation*, pages 69–82. Springer, 2011.

[145] Pierre Soille. *Morphological image analysis: principles and applications*. Springer Science & Business Media, 2013.

[146] Christopher S Stewart, Bradley C Leibovich, Amy L Weaver, and Michael M Lieber. Prostate cancer diagnosis using a saturation needle biopsy technique after previous negative sextant biopsies. *The Journal of urology*, 166(1):86–92, 2001.

[147] PJ Sudharshan, Caroline Petitjean, Fabio Spanhol, Luiz Eduardo Oliveira, Laurent Heutte, and Paul Honeine. Multiple instance learning for histopathological breast cancer image classification. *Expert Systems with Applications*, 117:103–111, 2019.

[148] E Suryani, EI Asmari, and B Harjito. Image segmentation of acute myeloid leukemia using multi otsu thresholding. In *Journal of Physics: Conference Series*, volume 1803, page 012016. IOP Publishing, 2021.

[149] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott E Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. corr abs/1409.4842 (2014). *arXiv preprint arXiv:1409.4842*, 2014.

[150] Mingxing Tan, Bo Chen, Ruoming Pang, Vijay Vasudevan, Mark Sandler, Andrew Howard, and Quoc V Le. Mnasnet: Platform-aware neural architecture search for mobile. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2820–2828, 2019.

[151] Jing Rui Tang, Nor Ashidi Mat Isa, and Ewe Seng Ch'ng. A fuzzy-c-means-clustering approach: Quantifying chromatin pattern of non-neoplastic cervical squamous cells. *PloS one*, 10(11):e0142830, 2015.

[152] Afaf Tareef, Yang Song, Heng Huang, Dagan Feng, Mei Chen, Yue Wang, and Weidong Cai. Multi-pass fast watershed for accurate segmentation of overlapping cervical cells. *IEEE transactions on medical imaging*, 37(9):2044–2059, 2018.

[153] Maja Temerinac-Ott, Germain Forestier, Jessica Schmitz, Meyke Hermsen, JH Bräsen, Friedrich Feuerhake, and Cédric Wemmert. Detection of glomeruli in renal pathology by mutual comparison of multiple staining modalities. In *Proceedings of the 10th International Symposium on Image and Signal Processing and Analysis*, pages 19–24. IEEE, 2017.

[154] Andrea Vedaldi and Karel Lenc. Matconvnet: Convolutional neural networks for matlab. In *Proceedings of the 23rd ACM international conference on Multimedia*, pages 689–692, 2015.

[155] Bairavi Venkatesh, Tosha Shaht, Antong Chen, and Soheil Ghafurian. Restoration of marker occluded hematoxylin and eosin stained whole slide histology images using generative adversarial networks. In *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, pages 591–595. IEEE, 2020.

[156] Mitko Veta, A Huisman, Max A Viergever, Paul J van Diest, and Josien PW Pluim. Marker-controlled watershed segmentation of nuclei in h&e stained breast cancer biopsy images. In *2011 IEEE international symposium on biomedical imaging: from nano to macro*, pages 618–621. IEEE, 2011.

[157] Mitko Veta, Josien PW Pluim, Paul J Van Diest, and Max A Viergever. Breast cancer histopathology image analysis: A review. *IEEE transactions on biomedical engineering*, 61(5):1400–1411, 2014.

[158] Mitko Veta, Paul J Van Diest, Robert Kornegoor, André Huisman, Max A Viergever, and Josien PW Pluim. Automatic nuclei segmentation in h&e stained breast cancer histopathology images. *PloS one*, 8(7):e70221, 2013.

[159] Neil J Vickers. Animal communication: when i'm calling you, will you answer too? *Current biology*, 27(14):R713–R715, 2017.

[160] Luc Vincent and Pierre Soille. Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 13(06):583–598, 1991.

[161] P Viswanathan. Fuzzy c means detection of leukemia based on morphological contour segmentation. *Procedia Computer Science*, 58:84–90, 2015.

[162] Carolina Wählby, I-M Sintorn, Fredrik Erlandsson, Gunilla Borgefors, and Ewert Bengtsson. Combining intensity, edge and shape information for 2d and 3d segmentation of cell nuclei in tissue sections. *Journal of microscopy*, 215(1):67–76, 2004.

[163] Khin Yadanar Win and Somsak Choomchuay. Automated segmentation of cell nuclei in cytology pleural fluid images using otsu thresholding. In *2017 International Conference on Digital Arts, Media and Technology (ICDAMT)*, pages 14–18. IEEE, 2017.

[164] D Wittekind. Traditional staining for routine diagnostic pathology including the role of tannic acid. 1. value and limitations of the hematoxylin-eosin stain. *Biotechnic & histochemistry*, 78(5):261–270, 2003.

[165] Qiang Wu, Fatima Merchant, and Kenneth Castleman. *Microscope image processing*. Elsevier, 2010.

[166] Saining Xie, Ross Girshick, Piotr Dollár, Zhuowen Tu, and Kaiming He. Aggregated residual transformations for deep neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1492–1500, 2017.

[167] Xinpeng Xie, Yuexiang Li, Menglu Zhang, and Linlin Shen. Robust segmentation of nucleus in histopathology images via mask r-cnn. In *International MICCAI Brainlesion Workshop*, pages 428–436. Springer, 2018.

[168] Fuyong Xing and Lin Yang. Robust nucleus/cell detection and segmentation in digital pathology and microscopy images: a comprehensive review. *IEEE reviews in biomedical engineering*, 9:234–263, 2016.

[169] Jun Xu, Lei Gong, Guanhao Wang, Cheng Lu, Hannah Gilmore, Shaoting Zhang, and Anant Madabhushi. Convolutional neural network initialized active contour model with adaptive ellipse fitting for nuclear segmentation on breast histopathological images. *Journal of Medical Imaging*, 6(1):017501, 2019.

[170] Pingkun Yan, Xiaobo Zhou, Mubarak Shah, and Stephen TC Wong. Automatic segmentation of high-throughput rnai fluorescent cellular images. *IEEE transactions on information technology in biomedicine*, 12(1):109–117, 2008.

[171] Faliu Yi, Junzhou Huang, Lin Yang, Yang Xie, and Guanghua Xiao. Automatic extraction of cell nuclei from h&e-stained histopathological images. *Journal of Medical Imaging*, 4(2):027502, 2017.

[172] Sergey Zagoruyko, Adam Lerer, Tsung-Yi Lin, Pedro O Pinheiro, Sam Gross, Soumith Chintala, and Piotr Dollár. A multipath network for object detection. *arXiv preprint arXiv:1604.02135*, 2016.

[173] Xiangyu Zhang, Xinyu Zhou, Mengxiao Lin, and Jian Sun. Shufflenet: An extremely efficient convolutional neural network for mobile devices. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 6848–6856, 2018.

[174] Zhong-Qiu Zhao, Peng Zheng, Shou-tao Xu, and Xindong Wu. Object detection with deep learning: A review. *IEEE transactions on neural networks and learning systems*, 30(11):3212–3232, 2019.

[175] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision*, pages 2223–2232, 2017.