AN AI IMPLEMENTATION OF DIGITAL PATHOLOGY FOR THYROID CARCINOMA

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF
ENGINEERING AND NATURAL SCIENCES
OF ISTANBUL MEDIPOL UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE DEGREE OF
MASTER OF SCIENCE
IN
ELECTRICAL, ELECTRONICS ENGINEERING AND CYBER SYSTEMS

By
Gökhan ABBASOĞLU

May, 2024

AN AI IMPLEMENTATION OF DIGITAL PATHOLOGY FOR THYROID CARCINOMA

By Gökhan ABBASOĞLU

8 May 2024

We certify that we have read this dissertation and that in our opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Master of Science.

_____

Prof. Dr. Bahadır Kürşat Güntürk (Advisor)

_____

Prof. Dr. Hasan Fehmi Ateş

_____

Prof. Dr. Selim Akyokuş

Approved by the Graduate School of Engineering and Natural Sciences:

_____

Prof. Dr. Yasemin Yüksel Durmaz

Director of the Graduate School of Engineering and Natural Sciences

I hereby declare that all information in this document has been obtained and presented in accordance with the academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Signature :

Name, Surname: GÖKHAN ABBASOĞLU

ACKNOWLEDGEMENT

A significant part of this work took place during the COVID 19 pandemic period. Collecting data, processing and working on it, and meeting with people involved many difficulties. Some of the friends we worked with, including me, suffered serious illnesses and lost our relatives and loved ones.

First of all, I would like to thank my esteemed teacher Prof. Dr. Bahadır Kürşat Güntürk, who encouraged us, guided us and paved the way with solution suggestions at every opportunity throughout this process.

In addition, I would like to thank Ümraniye State Hospital Pathology Clinic Chief Prof. Dr. Itır Ebru Zemheri and her colleague Expert Dr. Cumhur Selçuk Topal for their great support and actual contributions to our work.

I would also like to thank the T.C. Sağlık Bakanlığı Anadolu Kuzey Kamu Hastaneleri Birliği (Ministry of Health Anatolia Northern Public Hospitals Association) and TÜBİTAK BİLGEM Management for signing a protocol that enabled a series of scientific studies, including this study.

I would also like to thank my colleagues at TÜBİTAK BİLGEM and MEDIPOL University for helping me with literature searches and data processing.

Gökhan ABBASOĞLU

May, 2024

CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

w : Angular Momentum
Ω : Resistance
λ : Wavelength

ABBREVIATIONS

AI	: Artificial Intelligence
ANN	: Artificial Neural Network
CNN	: Convolutional Neural Network
DP	: Digital Pathology
GUI	: Graphical User Interface
ML	: Machine Learning
WSI	: Whole Slide Image

TİROİD KANSERİ İÇİN BİR DİJİTAL PATOLOJİ YAPAY ZEKA UYGULAMASI

ÖZET

Gökhan ABBASOĞLU

Elektrik-Elektronik Mühendisliği, Yüksek Lisans

Tez Danışmanı: Prof. Dr. Bahadır Kürşat GÜNTÜRK

Mayıs, 2024

Dijital patoloji son yıllarda önem kazanan ve gelişmekte olan bir konudur. Patologların çoğu hala eski mikroskoplarını kullanmayı tercih etse de, yapay zekanın kullanıma sunulmasıyla birlikte dijital patoloji daha önce bahsedilmeyen birçok avantajı sunmaya adaydır. Bu çalışma, bu konuda daha sonra yapılacak çalışmalara da bir rehber teşkil edecek şekilde bir saha taramasının yanı sıra yapay zekanın patoloğun yardımcısı olarak konumlandırılabileceğini gösteren klinik bir çalışmayı ve ülke çapında uygulanabilecek bir dijital patoloji sistemi için yapılabilecek proje önerilerini içermektedir.

Hastalıkların ve özellikle kanserin teşhisinde patologların verdikleri kararlar teşhis tedavi sürecinde önemli bir rol oynamaktadır. Halbuki patologların koydukları teşhislerde %40'lara varan farklılaşmalar gözlemlenebilmektedir. Çoğu patolog ikinci bir görüş almak için hastadan alınan numunelerin slaytlarını başka bir patoloğa özel bir kurye ile göndermeyi tercih edebilmektedir. Dijital patoloji uygulamaları bu sürecin farklı noktalarında, süreci hızlandırmak, standardize etmek, tecrübe birikimi ve paylaşımını kolaylaştırmak, yeni araştırmalara alt yapı sunmak gibi önemli fonksiyonlar görebilir.

Patoloji insan vücudunun tümünü ilgi sahası içine aldığı için oldukça geniş bir uygulama pratiği vardır. Vücudun her bölgesinin farklı rahatsızlıkları için patolog görüşü istenebilmektedir. Bu yüzden incelenen bölgeye ve şüphelenilen hastalığa göre farklı yaklaşımlar takip edilmektedir. Bu sahanın tümünü ihata etmek çok uzun soluklu bir ekip çalışmasını gerektirir. Bu yüzden klinik çalışmamızı en yaygın kanser türlerinden biri olan tiroid kanseri üzerine yoğunlaştırdık. İlk olarak, taranan slaytlarda kanser şüphesi olabilecek bölgeleri bulup patoloğun önüne getirmenin çok önemli bir fayda sağlayacağını tespit ettik. Her ne kadar patologlar için malign bölgelerin %100'ünün teşhis edilmesi tatmin edici olarak kabul edilse de, pratikte %95'in üzerinde bir başarıma ulaşmanın iyi bir başlangıç olacağını kabul etmiştik. Yaptığımız çalışmalarda bu sahaya özel geliştirilmemiş açık kaynak kodlu mimarilerle bile %97 başarıma ulaşabildiğimizi gözlemledik.

Burada öncelikle dijital patalojinin önemli yararlarını ve kullanımında karşılaşılan zorlukları inceledik. Bu konuda daha önce yapılmış çalışmaları ve önerileri içeren geniş kapsamlı bir literatür taraması yaptık. Daha sonra teknolojiyi, patologların yerini almak yerine onlara yardımcı olacak şekilde kullanacak bir yol bulmaya çalıştık. Kullanıcının gerçek ihtiyaçlarını anlamak amacıyla doğrudan patologlarla (T.C. Sağlık Bakanlığı Ümraniye Devlet Hastanesi Patoloji Kliniği İstanbul) çalışarak, hasta isimlerinin

anonimleştirildiği hasta verilerini, teşhisleriyle birlikte, topladık. Patoloji slaytlarındaki malign (kötü huylu) bölgenin uzmana gösterilmesinin daha önemli bir hizmet olacağını gözlemledik. Model eğitimi için yeterli veri toplamak amacıyla tiroid histopatolojisi adı verilen tek bir alana yoğunlaştık. Toplanan verilerle, yapay zeka kullanarak malign bölgeyi belirlemek için segmentasyon yöntemleri ve sınıflandırma yöntemleri gibi çeşitli stratejiler denedik ve performanslarını karşılaştırdık. Sonunda sınıflandırmanın segmentasyondan daha iyi performans gösterdiği sonucuna vardık. Ancak daha iyi olabilecek bir strateji olarak, daha sonra yapılacak çalışmalar için birkaç farklı karma segmentasyon ve sınıflandırma stratejisi önerilebilir.

En popüler açık kaynak kodlu mimarilerden VGG_16, ResNet_18, Inception_v3, MobileNet_v3_small, EfficientNet_v2_s, SwinTransformer, ConvNext_tiny, SimpleCNN ile modeler oluşturup, eğitip, modellerin başarımını test ettik ve bu sonuçları değerlendirdik.

Ayrıca sonuç bölümünde, daha kaliteli ve hızlı bir sağlık hizmeti verebilmek için, hem yapılan saha taramalarından hem de çalışmalarımızın sonuçlarından hareketle ülke çapında uygulanabilecek 5 farklı proje önerisinde bulunduk.

Anahtar sözcükler: Dijital Patoloji, Yapay Zeka, Tiroid Histopatolojisi.

AN AI IMPLEMENTATION OF DIGITAL PATHOLOGY FOR THYROID CARCINOMA

ABSTRACT

Gökhan ABBASOĞLU

MSc in Electrical, Electronics Engineering and Cyber Systems

Advisor: Prof. Dr. Bahadır Kürşat GÜNTÜRK

May, 2024

Digital pathology is a developing issue that has gained importance in recent years. Although most pathologists still prefer to use their plain old microscopes, with the introduction of artificial intelligence, digital pathology is subject to offer many advantages previously unmentioned. This study includes a field survey that will serve as a guide for future studies on this subject, as well as a clinical study showing that artificial intelligence can be positioned as an assistant to the pathologist, and project suggestions for a digital pathology system that can be implemented nationwide.

In the diagnosis of diseases, especially in carcinoma, the decisions made by pathologists, play an important role in the diagnosis and treatment process. However, differences of up to 40% can be observed in the diagnoses made by pathologists. Many pathologists prefer to send slides of samples taken from the patient via a private courier to another pathologist to obtain a second opinion. Digital pathology applications can serve important functions at different points of this process, such as accelerating the process, standardizing it, facilitating the accumulation and sharing of experience, and providing infrastructure for new research.

Since pathology interested in the entire human body, it has a wide range of applications. A pathologist's opinion may be requested for different diseases of different parts of the body. Therefore, different approaches are followed depending on the body region examined and the suspected disease. Covering this entire area requires a very long-term team work. That's why we focused our clinical study on thyroid cancer, one of the most common types of cancer. First of all, we determined that finding areas that may be suspicious for cancer in the scanned slides and bringing them to the pathologist would be of great benefit. Although diagnosing 100% of malignant areas is considered satisfactory for pathologists, we accepted that achieving a success rate of over 95% would be a good start in practice. In our studies, we observed that we could achieve 97% performance even with open source architectures that were not developed specifically for this field.

Here, we first examined the important benefits of digital pathology and the difficulties encountered in its use. We conducted an extensive literature review, including previous studies and recommendations on this subject. We then tried to find a way to use the technology to assist pathologists rather than replace them. In order to understand the real needs of the user, we worked directly with pathologists (T.R. Ministry of Health Ümraniye State Hospital Pathology Clinic Istanbul) and collected patient data, where patient specific data such as names were anonymized, together with their diagnoses. We

observed that showing the malignant area on the pathology slides to the specialist would be a more valuable service. To collect sufficient data for model training, we concentrated on a single area called thyroid histopathology. With the collected data, we tried various strategies such as segmentation methods and classification methods to identify the malignant area using artificial intelligence and compared their performance. We finally concluded that classification outperforms segmentation. However, as a better strategy, several different mixed segmentation and classification strategies can be recommended for future studies.

We created and trained models with VGG_16, ResNet_18, Inception_v3, MobileNet_v3_small, EfficientNet_v2_s, SwinTransformer, ConvNext_tiny, SimpleCNN, which are among the most popular open source architectures, tested the performance of the models and evaluated these results.

In addition, in the conclusion section, we proposed 5 different projects that can be implemented throughout the country, based on both the field surveys and the results of our studies, in order to provide a higher quality and faster health service.

Keywords: Digital Pathology, Artificial Intelligence, Thyroid Histopathology.

CHAPTER 1

1. INTRODUCTION

Pathology is a branch of medicine that deals with examination of tissues, organs, bodily fluids, and sometimes autopsies to understand the nature and causes of diseases. Pathology plays a crucial role in modern medicine by providing insights into the mechanisms of diseases, aiding in the development of treatment strategies, and contributing to the overall understanding of human health.

Histopathology involves examining tissue samples under a microscope to identify abnormalities such as cancerous cells. Pathologists face several challenges in histopathology analysis, ranging from the complexity of interpreting tissue samples to handling large amounts of data. Some key challenges are listed below:

- **Variability in interpretation:** Histopathology analysis often involves subjective interpretation, as pathologists need to visually examine tissue samples under a microscope and identify subtle morphological features. Inter-observer variability, where different pathologists may reach different conclusions when examining the same sample, can be a significant challenge. This variability can affect diagnostic accuracy and patient outcomes.

- **Time-intensive process:** Histopathology analysis can be time-consuming, especially when pathologists need to review a large number of samples. This can lead to delays in diagnosis and treatment, particularly in cases where timely intervention is critical. Pathologists must balance thorough examination with efficiency to provide timely and accurate diagnoses.

- **Limited access to expertise:** In some regions, there may be a shortage of trained pathologists, particularly subspecialists with expertise in specific types of diseases or tissues. This shortage can result in delays in diagnosis and challenges in accessing specialized care for patients.

- **Complexity of data interpretation:** Histopathology data is complex and heterogeneous, consisting of intricate tissue structures, various cell types, and diverse disease presentations. Pathologists must integrate information from multiple sources, including clinical history, imaging studies, and laboratory tests, to make accurate diagnoses. This complexity increases the cognitive burden on pathologists and requires continuous education and training to stay updated on advancements in the field.
- **Handling large datasets:** With the advent of digital pathology and whole-slide imaging, pathologists now have access to vast amounts of digital data. However, analyzing and managing large datasets present challenges in terms of storage, processing, and interpretation. Pathologists need efficient tools and technologies to navigate through these datasets effectively while maintaining diagnostic accuracy.
- **Quality assurance and standardization:** Ensuring consistent quality in histopathology analysis is essential for accurate diagnosis and patient care. Quality assurance programs and standardization efforts aim to minimize variability among pathologists and improve the reliability of diagnoses. However, implementing and maintaining these standards across different institutions and regions can be challenging.

1.1. Digital Pathology

Digital pathology stands for the use of digital technology to analyze and interpret pathology information, particularly in the context of examining tissue samples. Instead of traditional microscopy, where pathologists examine glass slides under a microscope, digital pathology involves the capture, management, and medical interpretation of pathology images in a digital format. This field has seen significant advancements due to the integration of digital tools and technologies into pathology practices. Here are some key aspects of digital pathology:

Digital slide scanning: Tissue samples are converted into high-resolution digital images through the use of specialized slide-scanning equipment. These digital slides can then be viewed on a computer screen.

Telepathology: Digital pathology enables remote access to pathology slides, allowing pathologists to collaborate and consult with colleagues from different locations. This is particularly useful for sharing expertise, especially in areas with limited access to specialized pathology services.

Storage and archiving: Digital pathology allows for the efficient storage and retrieval of pathology images. This digital archiving makes it easier to manage large volumes of data and facilitates retrospective analysis.

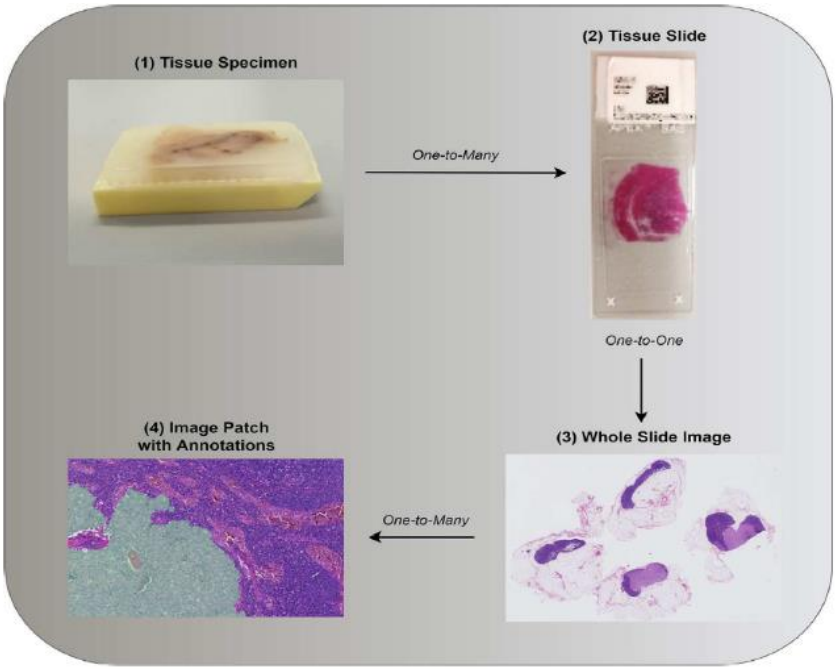
Image analysis and quantification: Digital pathology often involves the use of image analysis software, which can assist pathologists in quantifying various features of tissue samples. This can be valuable for research, as well as for providing more objective and reproducible diagnostic information.

Integration with electronic health records (EHR): Digital pathology can be integrated into electronic health record systems, allowing for seamless sharing of pathology results and reports with other healthcare providers.

Education and training: Digital pathology is also used for educational purposes, enabling the creation of digital teaching resources and virtual pathology labs for training medical professionals.

The adoption of digital pathology has the potential to enhance efficiency, collaboration, and accuracy in pathology practices. It offers benefits such as improved accessibility, faster turnaround times, and the integration of advanced technologies for image analysis and machine learning applications in pathology.

(a)



(b)

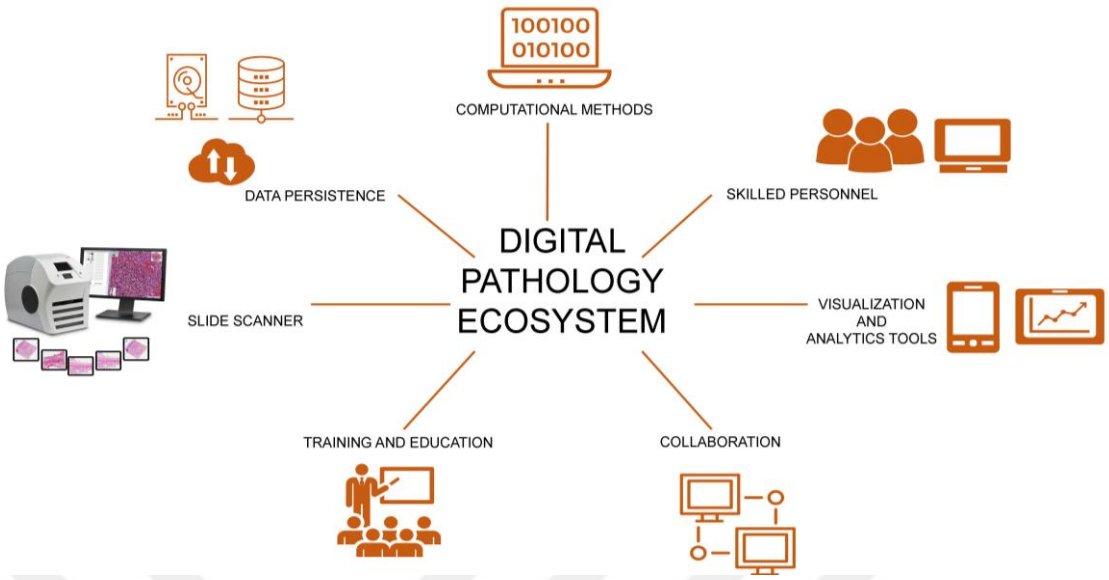


Figure 1.1: a) (a.1): The examination of tissue specimens is commonly explored as a potential indicator for patient diagnosis, prognosis, or other individual patient-related details. **(a. 2,3)** In both medical practice and scientific investigation, because of limited time, often only one tissue slide or its digital version is examined. Annotations associated with a single tissue sample may be provided, such as marking the presence of cancerous cells. **(a.4)** Given the immense gigapixel scale of Whole Slide Images (WSIs), additional image reduction is necessary for image analysis. Patches are typically extracted based on these annotations. This illustration was extracted from reference [1]. **b)** Depicts the representation of various elements in Digital Pathology.

After a patient has the tumor surgically removed, the tissue samples are subjected to a process of decalcification, stained using suitable dyes, and then placed onto glass slides. The effectiveness of a DP ecosystem relies on several essential internal components designed to execute the key steps outlined below:

1. Slide scanner
2. Storage
3. Tools for data analysis
4. Graphical User Interface (GUI) designed for user engagement and visual representation
5. Skilled personnel
6. Suite for collaboration
7. Support for training
8. Whole Slide Images (WSIs)

1.1.1. Slide scanner

Slide scanners are instruments with the ability to digitally capture high-magnification images of prepared histology glass slides, typically ranging from 40X to 100X. They transform these images into digital datasets known as whole slide images (WSI). Numerous commercially available slide scanners can achieve micron-level scanning, enabling the visualization of even cellular and sub-cellular structures. Over time, magnification capabilities have advanced from a 20X to a 100X objective, leading to substantial file sizes, often in the order of several gigabytes per slide. Some scanners provide options for highly compressing image data, resulting in smaller file sizes.

1.1.2. Storage

Digital whole slide images (WSIs) obtained through scanning must be stored for future use, and this can be achieved by archiving them on high-capacity encrypted hard disks or in a secure cloud repository. Various storage options offer a balance between security and flexibility. Encrypted hard disks ensure secure transportation of data, while cloud storage facilitates remote collaboration. Archiving data is essential for preserving slide information over time, facilitating future analysis and training endeavors.

1.1.3. Tools for data analysis

Processing digital files is essential to (1) extract tumor regions, (2) assess their grades, (3) recognize associated patterns, and other related tasks. Image analysis tools, whether they are freely available and open-source or proprietary, are utilized for these purposes. Some proprietary tools come with pre-established image analysis pipelines and suggest various built-in methods for processing. Certain DP (Digital Pathology) enterprises also offer the option of remotely analyzing images. Given the size of the images under examination, these tools are typically implemented on hardware with sufficient processing power and memory.

1.1.4. Graphical user interface (GUI) designed for user engagement and visual representation

The examined images must be sent back to the user to facilitate actions such as validation, data export, and viewing the analysis results. A specifically designed user interface (UI) captures this information and delivers essential insights to the user through either a

standalone application or a web interface. Effective visualization plays a pivotal role in offering a clear and concise summary of the results, enhancing the ease of understanding.

1.1.5. Skilled personnel

An essential element of the Digital Pathology (DP) ecosystem involves not only the infrastructure but also proficient analysts capable of translating the extensive knowledge of pathologists into a language understandable by machines. These experts must be familiar with diverse image-handling methods since Whole Slide Images (WSI) often have proprietary formats and require specialized tools for viewing. Expertise in handling huge size images, applying methods for retrieving tumor-associated information, conducting statistical examination, assessing the reliability of techniques, and effectively communicating these findings to team members is imperative.

1.1.6. Suite for collaboration

A discretionary feature involves facilitating collaboration among different participants within the group. Consequently, certain Digital Pathology (DP) infrastructures offer the capability for simultaneous collaboration. For instance, a pathologist and an image analyst can jointly review annotated data even if they are located in different places. As technological progress continues, this capability is anticipated to evolve into a crucial and standard feature in the future.

1.1.7. Support for training

Several universities have effectively implemented Digital Pathology (DP), bringing about a transformation in their educational systems. They have integrated student training in pathology topics by utilizing Whole Slide Images (WSI) archived during tumor studies. These archived samples not only serve as valuable resources for training professionals in tumor analysis but also enable students to shift from traditional learning with glass slides to collaborative digital platforms. Using suitable software tools, students can leverage the advantages of digital platforms. Instructors can adapt the curriculum to incorporate the use of digital WSIs, offering the same flexibility as a microscope, albeit in a virtual setting. These aspects highlight the potential of DP in overcoming the limitations associated with traditional glass slides in the context of education and training.

1.1.8. Whole slide images (WSIs)

The rise of digital imaging in pathology has been a key factor in the recent surge of virtual microscopy. Shifting from the analysis of glass slides to whole slide images (WSIs) has made analysis of digital images feasible. WSIs, intricate images produced by slide scanners through high-magnification scanning of glass slides, utilize advanced optic sensors for processing. There are two main approaches to image acquisition: tile-based approach and line-based approach. The former method generates small overlapping tiles from sensor output, while the second method scans linearly along a single axis, creating long strips of images that are later stitched together. This stitching results in a single large image with gigapixel resolution, significantly increasing its size. As WSIs in digital pathology are larger than those in other fields like radiology, they are typically stored in a lossless format like TIFF. To manage the increased image size, various compression levels are often applied, reducing the images to approximately 1/6th of their original size. This reduction facilitates image transportation, especially over networks. Given that the images are stitched at the highest magnification, they can be viewed at lower magnifications through image transformation, mimicking the functionality of a microscope and allowing observation of tissues at different objectives such as 5x, 10x, 20x, and 40x [27].

1.2. Challenges in Digital Pathology

The conversion of tissue glass slides into digital format presents both promising prospects and obstacles for computational imaging scientists. Although computational imaging holds potential for enhancing the quantitative understanding of diseases and advancing precision medicine, significant technical and computational hurdles must be addressed prior to computer-assisted image analysis in digital pathology can seamlessly integrate into routine clinical diagnostic workflows. Despite the opportunities presented by the advent of digital pathology and the commendable results achieved so far, evident challenges persist, hindering the automation and expeditious implementation of computer-assisted monitoring and diagnosis.

In the following subsections, several pressing obstacles, which required to be addressed, will be discussed in detail.

1.2.1. Absence of labelled data

Most AI and machine learning algorithms necessitate a large set of good quality, labeled and balanced images for training. To do this, a pathologist is required to manually label and annotate the region of interest (i.e., malignancy or disorders) in all whole slide images. There are three main factors which make the manual annotation and labelling of WSIs more challenging.

- Manual labelling of whole slide images is very time-consuming process because a typical glass-slide of size 20mm x 15mm produces a gigapixel image of size 80,000 x 60,000pixels.
- Also, manual annotations often cause a financial constraint to development of biomedical image processing applications. Reusing public ready-to-analyse data as training data in machine learning is another possibility, such as ImageNet [76] in natural images. In other words, although the use of crowdsourcing strategy may be more economical and faster, it can introduce more noise compared to applications developed for natural images.
- Problem in WSIs such as low resolution, blur, occlusion, overlapping, and ambiguity of features make the manual annotation and labeling step more challenging.

Table 1.1: Downloadable WSI database. This table was taken from [2].

Dataset or author's name	# slides or patches	Stain	Disease	Additional data
TCGA [301,302]	18,462	H&E	Cancer	Genome/transcriptome/epigenome
GTEx [303,304]	25,380	H&E	Normal	Transcriptome
TMAD [305,306]	3726	H&E/IHC		IHC score
TUPAC16 [307]	821 from TCGA	H&E	Breast cancer	Proliferation score for 500 WSIs, position for mitosis for 73 WSIs, ROI for 148 cases
Camelyon17 [50]	1000	H&E	Breast cancer (lymph node metastasis)	Mask for cancer region (in 500 WSIs with 5 WSIs per patient)
Köbel et al. [308,309]	80	H&E	Ovarian carcinoma	
KIMIA Path24 [310,311]	24	H&E/IHC and others	various tissue	

Table 1.2: Hand annotated histopathological images publicly available. This table was taken from [2].

Dataset or paper	Image size (px)	# images	Stain	Disease	Additional data	Potential usage
KIMIA960 [312,313]	308 × 168	960	H&E/IHC	various tissue		Disease classification
Bio-segmentation [314,315]	896 × 768, 768 × 512	58	H&E	Breast cancer		Disease classification
Bioimaging challenge 2015 [316,125]	2040 × 1536	269	H&E	Breast cancer		Disease classification
GlaS [317,318]	574–775 × 430–522	165	H&E	Colorectal cancer	Mask for gland area	Gland segmentation
BreakHis [319,46]	700 × 460	7909	H&E	Breast cancer		Disease classification
Jakob Nikolas et al. [313,320]	1000 × 1000	100	IHC	Colorectal cancer	Blood vessel count	Blood vessel detection
MITOS-ATYPIA-14 [321]	1539 × 1376, 1663 × 1485	4240	H&E	Breast cancer	Coordinates of mitosis with a confidence degree/six criteria to evaluate nuclear atypia	Mitosis detection, nuclear atypia classification
Kumar et al. [322,323]	1000 × 1000	30	H&E	Various cancer	Coordinates of annotated nuclear boundaries	Nuclear segmentation
MITOS 2012 [26,324]	2084 × 2084, 2252 × 2250	100	H&E	Breast cancer	Coordinates of mitosis	Mitosis detection
Janowczyk et al. [325,326]	1388 × 1040	374	H&E	Lymphoma	None	Disease classification
Janowczyk et al. [325,326]	2000 × 2000	311	H&E	Breast cancer	Coordinates of mitosis	Mitosis detection
Janowczyk et al. [325,326]	100 × 100	100	H&E	Breast cancer	Coordinates of lymphocyte	Lymphocyte detection
Janowczyk et al. [325,326]	1000 × 1000	42	H&E	Breast cancer	Mask for epithelium	Epithelium segmentation
Janowczyk et al. [325,326]	2000 × 2000	143	H&E	Breast cancer	Mask for nuclei	Nuclear segmentation
Janowczyk et al. [325,326]	775 × 522	85	H&E	Colorectal cancer	Mask for gland area	Gland segmentation
Janowczyk et al. [325,326]	50 × 50	277,524	H&E	Breast cancer	None	Tumor detection
Gertych et al.[327]	1200 × 1200	210	H&E	Prostate cancer	Mask for gland area	Gland segmentation
Ma et al.[328]	1040 × 1392	81	IHC	Breast cancer		TIL analysis
Linder et al. [329,330]	93–2372 × 94–2373	1377	IHC	Colorectal cancer	Mask for epithelium and stroma	Segmentation of epithelium and stroma
Xu et al. [331]	Various size	717	H&E	Colon cancer		
Xu et al. [331]	1280 × 800	300	H&E	Colon cancer	Mask for colon cancer	Segmentation

Table 1.3: A widely used breast histopathology image dataset that is publicly accessible. The "Detail" column in the fourth column indicates the number of classes. This table was sourced from reference [3].

Datasets	Year	Staining	Detail	Magnification	Dataset size	Website
ICPR 2012	2012	H&E	\	40×	50 images corresponding to 50 high-power fields in 5 different biopsy slides	Closed
IDC	2014	H&E	\	40×	277,524 patches are from 162 IDC breast cancer histopathological slides (198,738 IDC negative, 78,786 IDC positive)	[45]
BreaKHis	2015	H&E	4	40×, 100×, 200×,400×	7,909 histopathology images	[46]
Bioimaging 2015 breast histology classification challenge	2015	H&E	4	200×	249 images for training, 20 image for testing and an extended testset of 16 images	[47]
TUPAC 2016	2016	H&E	\	40×	500 for training and 321 for testing breast cancer histopathology WSIs	[48]
Camelyon 2016	2016	H&E	2	40×,10×,1×	400 WSIs of lymph node	[49]
Camelyon 2017	2017	H&E	\	40×	200 WSIs of lymph node	[50]
BACH	2018	H&E	4	\	Part A: 400 microscopy images Part B: 30 whole-slide images	[51]

Within the digital pathology domain, there exists a limited selection of publicly accessible datasets that include labeled and manually annotated images, outlined in **Table 1.1**, **Table 1.2**, and **Table 1.3**. These datasets prove beneficial when the analysis purpose aligns with the conditions of the slides (e.g., stain) and attributes of the images (e.g., magnification level and resolution). It's worth noting that these datasets are tailored to specific diseases or cell types, and as a result, numerous tasks may not find coverage within these datasets.

Numerous studies have endeavored to address this issue, and the majority of approaches can be categorized into one of the following groups: 1) enhancing the efficiency of labeled data augmentation, 2) incorporating weak labels or unlabeled information, or 3) employing models and parameters (such as fine-tuning or transfer learning) from other tasks.

1.2.2. Tremendous variability in different cancer types

Several fundamental tissue types exist, including muscle, epithelial, connective, and nervous tissues. However, when histopathology or Whole Slide (WS) images are viewed by computers as a 2D array of pixels, the number of patterns derived from these tissues becomes almost limitless. Furthermore, textural variations in tissue

types that constitute an organ add another layer of complexity. This extensive variability poses a significant challenge for image analysis algorithms in recognizing tissues. Consequently, achieving optimal performance with state-of-the-art deep networks necessitates obtaining numerous training cases for each variation. Obtaining such datasets, especially with labeled data, can be a challenge.

1.2.3. The non-boolean characteristic of diagnostic reports

As a medical document, a pathology report provides details about a diagnosis, including conditions like carcinoma. A sample of the suspicious tissue undergoes examination in a laboratory for the investigate the investigation of the disease, where a pathologist studies it with a microscope. Additional tests may be requested for further insights. The pathology report encompasses these findings, the diagnosis, information on the extent of cancer spread, and other relevant details. Pathologists use this report to determine the most effective treatment. **Section 1.2.4** will elaborate on the specific content of a standard digital pathology report.

1.2.4. Content of a pathology report

A typical report consists of four main parts including identifying information, gross description, microscopic description and grade. Below, a detail description is given about content of pathology report.

Gross description: The pathologist provides a description of the tissue sample without resorting to a microscope. This description may encompass details such as size, shape, color, weight, and what does it look like. Cancer measurements are commonly expressed in centimeters. It's essential to note that size is just one aspect of the overall assessment, as larger tumors might not necessarily indicate faster growth compared to smaller ones.

Microscopic examination: involves the pathologist cutting the tissue into thin sections, placing them on slides, applying dye for staining, and thoroughly examining them under a microscope. During this process, the pathologist observes the characteristics of cancer cells, their comparison to normal cells, and determines whether they have invaded surrounding tissues.

Grade: The pathologist assesses the cancer cells in comparison to healthy cells, utilizing specific scales tailored for different types of cancers. Tumor grade serves as an indicator of the likelihood of growth and spread. Generally, the grades can be interpreted as follows:

- Grade 1: Well-differentiated or low grade - The cells exhibit slight variations from regular cells, and they are not growing rapidly.
- Grade 2: Moderately differentiated or moderate grade - The cells deviate from normal cells in appearance and are growing at a moderate pace.
- Grade 3: Poorly differentiated or high grade - The cells markedly differ from normal cells, and they are growing or spreading rapidly.

Most of the published research papers in literature, about classification problems in digital pathology, largely deal with binary classification problem such as “benign” or “malignant”.

However, as described above, diagnosis in pathology has a complex nature. Hence, binary decision system may not be directly used in practice. In addition to problem of binary decision, in literature, most of the available datasets only contains data labelled at the case-level. However, in order to automatically return a comprehensive report, datasets which contain labels at cell-level should be datasets which contain labels at cell-level should be provided.

1.2.5. Scanner variability across laboratories

Various whole slide scanners, produced by different vendors, are presently accessible in the market, capable of both brightfield and fluorescence imaging. Each scanner employs distinct compression methods, sizes, illumination, objectives, and resolutions. Additionally, they generate images in proprietary file formats unique to each scanner. The absence of a standardized image format has the potential to impede the timely creation of extensive datasets.

Furthermore, employing various scanners from different brands results in the acquisition of diverse scanned Whole Slide Images (WSIs). While developing a universal framework immune to scanner variability is a challenging task, an automated system that generates a pathology report based on a provided WSI should demonstrate resilience to these variations.

1.2.6. Artifacts and color variability

Since the data is gathered from various origins, there is an absence of standard in staining procedures. Different artifacts may emerge at the different stages of the entire sample preparation process and also during imaging. These artifacts encompass factors like ischemia times, fixation times, microtome-related issues, variations in staining reagents, and imaging-related irregularities such as uneven lighting, variations in focus, image fragmentation, and the presence of fluorescence residues and bleed-through. **Figure 1.2** provides illustrations of such artifacts.

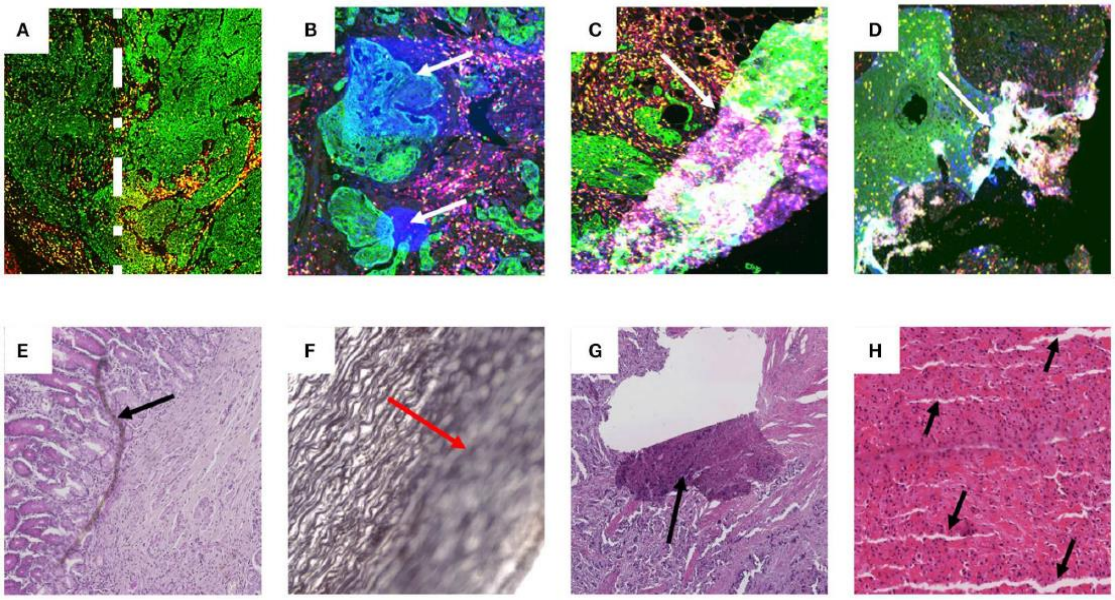


Figure 1.2: Instances of artifacts are evident in both fluorescence and brightfield images. Images **(A–D)** serve as illustrations of multiplex immunofluorescence (IF) images featuring various artifacts. These images were acquired from slides stained with Pan-cytokeratin (green), DAPI (blue), CD3 (yellow), and CD8 (red). **(A)** Demonstrates a higher concentration of Pan-cytokeratin on the right side compared to the left, delineated by a dashed white line. **(B)** White arrows point to areas with artificially increased intensity in the DAPI channel during imaging. **(C,D)** White arrows draw attention to tears and folds in the tissue, resulting in blurred images and fluorescence anomalies. Images **(E–H)** display instances of artifacts in brightfield images labeled with H&E **(E,G,H)** or Verhoeff's elastic stain **(F)**. **(E)** A black arrow highlights a foreign object beneath the coverslip. **(F)** A red arrow indicates an area out of focus. **(G)** A black arrow points to a tear in the tissue. **(H)** Black arrows identify cutting artifacts. All images were captured using a 20× objective on a Zeiss Axioscan.z1 microscope. This figure was sourced from reference [1].

As the human brain undergoes training, it develops the ability to disregard artifacts and variations in staining, focusing on the visual information crucial for precise diagnosis. To achieve a similar proficiency in deep learning models, two primary approaches are typically employed. The first method explicitly eliminates artifacts,

employing techniques such as image filters, and standardizes color variability. Conversely, the second approach adopts a more indirect strategy by supplementing the data with synthetically generated data that encompasses a diverse range of artifacts and staining variations. This incorporation ensures that learning about these aspects becomes an inherent aspect of the training process.

1.2.7. The substantial size of whole slide images (WSI)

A Whole Slide Image (WSI) is obtained by scanning a glass slide at an extremely high resolution, typically around 0.25 micrometers/pixel (equivalent to 40X magnification under a microscope). Pathologists generally examine high-resolution gigapixel images, commonly known as WSIs, with sizes often exceeding 80,000 by 60,000 pixels. However, deep neural networks typically run on much smaller image resolutions, often not exceeding 600 by 600 pixels. To address this issue, for AI algorithms, a patching strategy is employed, dividing the image into numerous smaller tiles, as well as for traditional computer vision methods. Even with this patching strategy, down sampling may be necessary to input these images into a deep network (refer to **Figure 1.3**).

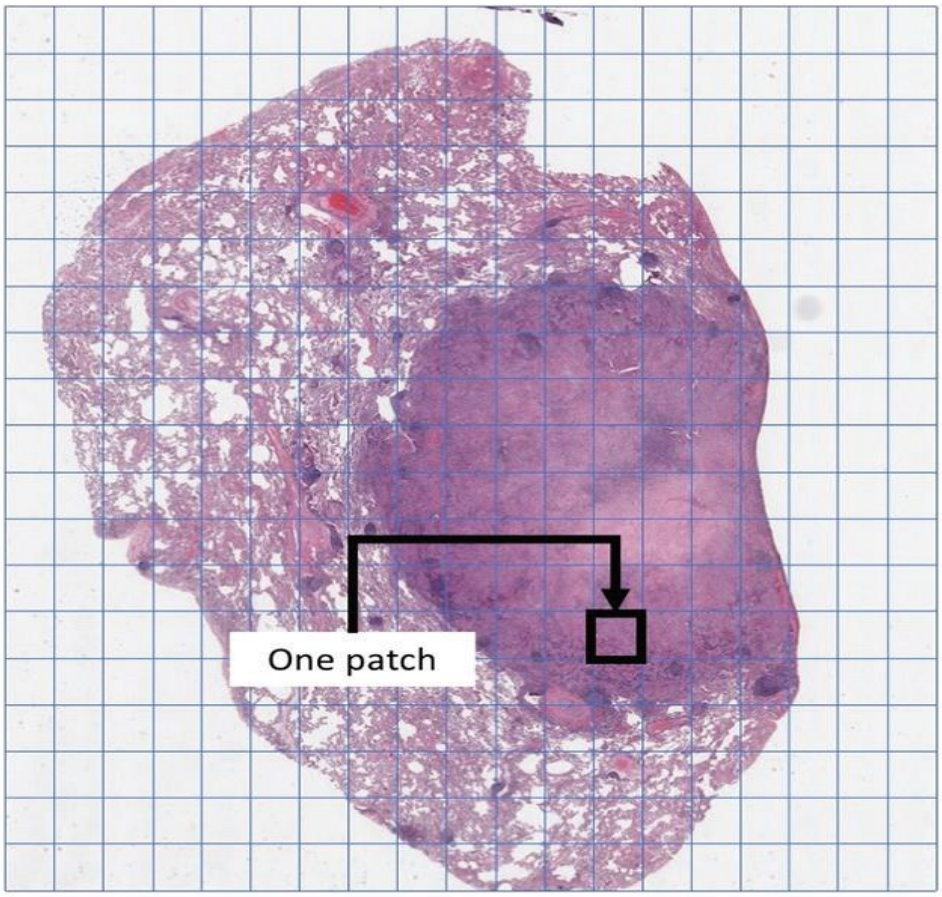


Figure 1.3: Typically, patching is employed to represent extensive scans. For example, each patch might consist of an image measuring 1000 pixels × 1000 pixels at a magnification of ×20. This illustration is sourced from reference [4].

Reducing the resolution of these patches may lead to the omission of essential information and details present in the scanned Whole Slide Image (WSI). In simpler terms, employing deep neural networks with larger input sizes necessitates more extensive network architectures and a considerably larger number of neurons. Training such networks becomes more challenging and possibly unfeasible, particularly when dealing with limited training datasets.

1.2.8. Absence of clarity and understandability in deep convolutional neural networks (CNNs)

Deep Convolutional Neural Networks (CNNs) are employed as cutting-edge approaches in computer vision challenges to extract clinically relevant features and information. While these networks have achieved notable success in tasks like object and scene recognition, it becomes challenging to elucidate the specific rationale behind a network's decision when handling Whole Slide Images (WSIs). In

essence, the intricate calculations involving millions of multiplications and additions within a deep Artificial Neural Network (ANN) contribute to the network's output. However, a trained CNN lacks a detailed explanation to comprehend the reasoning behind its decisions and ensure the extraction of clinically meaningful features.

In the pathology community, a typical pathology report does not only contain whether the glass slide includes cancer cells or not, but also it consists of four main parts including identifying information, Gross description, Microscopic description, and Grade as explained in **Section 1.2.4.**

The result produced by a trained Convolutional Neural Network (CNN) is typically deemed insufficient in the medical field for substantiating the underlying justifications behind a particular decision. For a trustworthy diagnosis, an automated system should furnish a comprehensive, easily understandable, and interpretable report.

1.2.9. Various levels of magnification

Tissues that contains significant information, typically involves cells, and various tissues exhibit distinct cellular features. High-power field microscopic images effectively capture details about cell shapes, while lower-power field images are better at revealing structural information, such as glandular structures composed of multiple cells (refer to **Figure 1.4**). Cancerous tissues display abnormalities at both cellular and structural levels, emphasizing the importance of examining images at various magnifications, each providing essential information. Pathologists identify diseases by examining various details, spanning from cellular to tissue levels, accomplished by adjusting the magnification of microscopes. In machine learning, studies often use images at different magnifications [57]–[59]. Handling images at their original resolution is hard to handle, so they are frequently resized to correspond to various magnifications for analysis. The most informative magnification for diagnosis is a matter of debate [14],[39],[109]. However, there are cases where using both high and low magnification images together enhances accuracy, potentially influenced by the specific diseases, tissues, and machine learning algorithms involved.

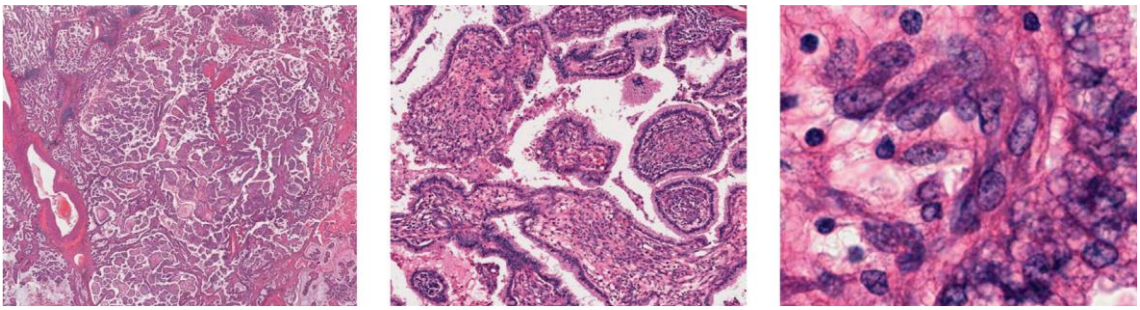


Figure 1.4: Various levels of magnification for a histopathological image are depicted. The images on the right provide a magnified perspective of the area outlined by the red box in the left images. The image on the far left clearly shows a papillary structure, whereas the image on the far right distinctly depicts the nuclei of individual cells. This illustration was taken from reference [2].

1.2.10. Whole slide image (WSI) as an unordered, texture-like image

Pathological images differ from images of cats and dogs in nature, as they exhibit a repetitive pattern of minimal components, typically cells. Consequently, they are more akin to textures than distinct objects. Convolutional Neural Networks (CNNs) achieve a certain degree of shift invariance through pooling operations. Moreover, traditional Convolutional Neural Networks (CNNs) have the ability to learn patterns resembling textures through data augmentation, which involves shifting tissue images with a small stride. However, various techniques, including the gray-level co-occurrence matrix [110], local binary pattern [111], Gabor filter bank, and recently developed deep texture representations using a CNN [64], [112], extensively exploit texture structures. Deep texture representations are computed using the correlation matrix of feature maps in a CNN layer. Transforming CNN features into texture representations helps achieve invariance concerning cell position while capitalizing on well-established CNN representations. Another benefit of deep texture representation is its suitability for processing large image sizes, such as those found in Whole Slide Images (WSI), as there are no limitations on input image size. Despite the potential ambiguity in distinguishing texture from non-texture, it is clear that a single cell or structure does not constitute a texture. The optimal method, therefore, depends on the specific object being analyzed.

1.2.11. Limitations in storage technology

One of the most common problem of digital pathalogy is the limitations in storage technology. Generally, for a session of one patient, there must be 3-5 slides to be

examined. A pathalogist looks at these slides with microscobes at different magnification levels from 4X to 40X. During this examination a pathalogist try to find out; (i) whether the samples has enough quality for diagnosis, (ii) whether there exist abnormal patterns both in tissue clusters and also in tissue shapes (especially in nucleus), (iii) whether abnormal patterns (if exist) shows malignancy, (iv) and finally the type of the malignancy [2]. To do the same thing with digital pathology each slide must be scanned with a resolution at least X40. This means we need 1 GB-10GB storage area for each session of each patient. A common pathology clinic in Türkiye have 50–100 patients each day. So, we need 50 GB-1TB storage area for each day. There are several efforts to diminish the size of the images through compression or by eliminating unnecessary parts from images, but these efforts will provide us a few weeks instead of a few days before to run out physical limits of a common computer system. A widely adopted resolution to this issue involves utilizing cloud technology for storage as an alternative to local storage systems. This idea also addresses another problem which is called "second opinion" between pathologists. In some cases, pathologists need a second opinion for his/her diagnosis because the tissue patterns may not be so clear to make a certain decision. At this time the slides have to be sent to other pathologist by courier via certain protocol. This process sometimes takes several weeks. But if the slide images are stored on the cloud and can be accessed by another pathologist, the consultation time reduces to minutes or hours. This is a great idea but must be handled carefully because of patient rights. The control of access to patient data must be given by patient himself/herself at this time. Another important point for this solution is: it, of course, requires internet connection with high upload speeds. Today most of the great hospitals has high speed Internet connections in great cities, but in peripherals, the facilities may not be enough for cloud computing. A batch process may be designed for this type of hospitals. Recent data may be stored in local storages while the whole data transferred to the cloud during off time.

CHAPTER 2

2. THEORETICAL PART

In pathology, Machine Learning (ML) is applied for the diagnosis of different carcinoma forms, including colon, cervical, breast, gastric and lung carcinoma. Its utilization encompasses tasks such as distinguishing between benign and malignant conditions, grading diseases, analyzing staining patterns, and screening for early tumor detection.

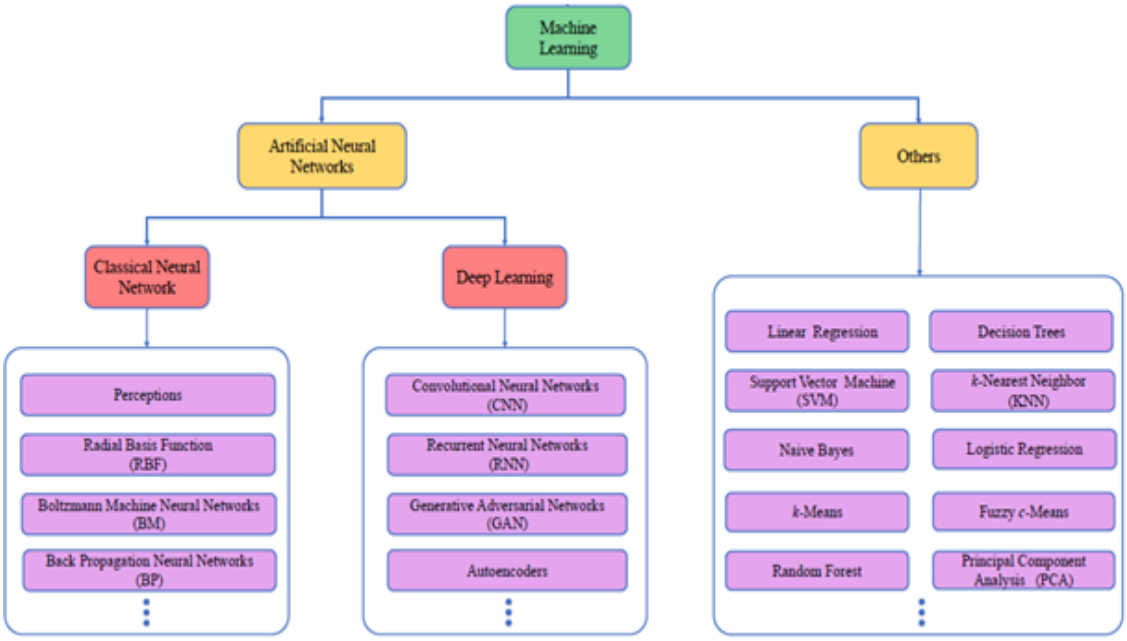


Figure 2.1: The configuration of Artificial Neural Network (ANN) technology within the AI knowledge system is depicted in this illustration. The figure was taken from reference [3].

It is essential to highlight that Artificial Neural Networks (ANNs), within the realm of machine learning, hold significance in diagnosis of pathology. ANN methods, encompassing both classical and deep neural networks, represent complex mathematical models that mimic the structure and functionality of the human

neural network. In recent times, several studies have been performed with the objective of segmenting, extracting features, and classifying Whole Slide Images (WSIs) for pathological diagnosis.

In the following subsections, image analysis and AI methods will be analyzed in detail for different cancer fields including thyroid, and breast.

2.1. Analysis of Images and Application of AI in Thyroid Cancer

Thyroid carcinoma stands out as one of the frequently diagnosed types of cancer worldwide. To facilitate analysis, the cases were categorized into five diagnostic groups, considering cytologic and surgical diagnoses: PTC, A/S PTC, A/S benign, NIFTP, and benign, as outlined in **Table 2.1**.

Table 2.1: Diagnoses before and after surgery for cases within each diagnostic category. This table was taken from reference [5].

Category	FNA Cytopathology	n	Surgical Histopathology
PTC	Malignant PTC (Bethesda VI)	9	PTC, classical variant
A/S PTC	Atypia of undetermined significance (Bethesda III)	3	PTC, classical variant
	Suspicious for malignancy (Bethesda V)	2	
NIFTP	Benign (Bethesda II)	2	NIFTP
	Atypia of undetermined significance (Bethesda III)	2	
	Suspicious for follicular neoplasm (Bethesda IV)	1	
	Malignant (Bethesda VI)	1	
A/S Benign	Atypia of undetermined significance (Bethesda III)	3	Benign nodule/follicular adenoma
		1	Multinodular goiter
		1	Multinodular goiter; lymphocytic thyroiditis
Benign	Benign, negative for malignant cells (Bethesda II)	9	No surgical follow-up
		1	Multinodular goiter

Abbreviations: A/S, atypical/suspicious; FNA, fine-needle aspiration; NIFTP, noninvasive follicular thyroid neoplasm with papillary-like features; PTC, papillary thyroid carcinoma.

Thyroid pathology presents substantial potential for employing image analysis and AI algorithms, with a wealth of existing literature exploring the use of these computational tools to scanned thyroid pathology images. In recent years, AI algorithms have been specifically designed for thyroid pathology applications. Thyroid specimens are commonly utilized to deliver a diagnosis known for its safety, accuracy, and cost-effectiveness. Despite the reliability of this approach, diagnosing some of the thyroid lesions can be difficult for pathologists due to limited and non-discriminative pathological features in collected samples. Automated algorithms addressing thyroid lesions and quantifying cellular parameters, such as staining intensity and nuclear properties (e.g., nuclear area and elongation factor of nuclei), are poised to play a crucial role in providing decision support for pathologists, particularly in cases with indeterminate findings.

Certainly, thyroid pathology holds significant potential for the utilization of automated algorithms. The growing occurrence of thyroid nodules, alongside the emergence of new conditions such as noninvasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP), presents a significant challenge for pathologists in diagnosing these lesions. Furthermore, the rate of uncertain diagnoses from fine-needle aspiration (FNA) remains notably high. Factors contributing to the increased prevalence of thyroid nodules and cancer include expanded ultrasound screenings and related variables. Since papillary thyroid carcinoma (PTC) typically displays less aggressive behavior with lower rates of recurrence and metastasis compared to other cancers, accurately distinguishing PTC and its variations becomes essential for precise patient management.

In the existing literature, the majority of studies focus on automated image analysis for extracting and classifying pathological features, specifically assessing stain intensity and evaluating nuclear features to distinguish between different types of thyroid cancers. Across these studies, a consistent finding is a strong correlation between automated AI assessment systems and manual assessments. Particularly in cases deemed indeterminate, automated assessment systems are often considered more reliable, faster, and easier to execute.

The majority of researches examined in this study focus on automated image analysis for the objective quantification of histological features, especially assessing immunohistochemical stain intensity and obtaining nuclear features. The primary aim is often to quantify distinctions among different types of thyroid carcinoma and non-neoplastic lesions. Across all these studies, consistently high correlation coefficients were observed in comparison to manual assessments.

2.2. Analysis of Images and The Application of AI in Breast Cancer

Breast cancer is the most commonly diagnosed cancer and a leading cause of cancer-related deaths among women. According to the World Health Organization (WHO), approximately 2.1 million women worldwide are diagnosed with breast cancer each year. In 2018, around 627,000 women lost their lives to the disease, accounting for about 15% of all cancer-related deaths among women. Breast tissue is classified into four types: normal, benign, in-situ carcinoma, and invasive carcinoma.

2.2.1. Applying classical artificial neural networks (ANNs) in breast histopathological image analysis (BHIA)

This section presents a summary of research in Breast Histopathological Image Analysis (BHIA) using classical Artificial Neural Network (ANN) methods. Subsequently, we conduct an analysis and summary of the chapter. The relevant studies are classified into tasks related to classification and segmentation depending on their objectives. The contributions, methodologies, and outcomes of each paper are subsequently outlined.

Table 2.2: Histopathology and traditional Artificial Neural Networks (ANNs) for the analysis of breast cancer images. The ANNs mentioned include Multi-Layer Perceptron (MLP), Probabilistic Neural Networks (PNN), Multi-layer Neural Network (MNN), with metrics such as Accuracy (Acc) and Sensitivity (Sn) considered. The "Detail" column in the second column indicates the number of classes and segmentation regions. This table was taken from reference [2]

Aim	Detail	Year	Reference	Team	Data Information	ANN type	Evaluation
Classification	3	2006	[35]	S. Petushi, et al.	24 slide images, H&E staining	Neural network	Acc = 90%
	2	2010	[36]	A. Osareh, et al.	Dataset 1: 692 specimens of fine needle aspirates of breast lumps, Dataset 2: 295 microarrays	PNN	Acc = 98.8%
	4	2011	[37]	S. Singh, et al.	1080 images, H&E staining, (1080 for training, 387 for validation, 387 for test)	Feed forward back propagation neural network	Acc = 95%
	3	2011	[38]	Y. Zhang, et al.	361 images, H&E staining, (760×570)	MLP	Acc = 95.22%
		2013	[39]				
		2013	[40]				
	3	2013	[41]	C. Loukas, et al.	65 regions of interests H&E staining, (20 grade I, 20 grade II, 25 grade III)	PNN	Acc = 87%
	2	2017	[42]	K. Shukla, et al.	70 images, H&E staining, (35 non-cancerous and 35 cancerous)	MLP	Acc = 80%, Sn = 82.9%, AUC = 89.2%
Segmentation	Nuclei	2013	[43]	M. Kowal, et al.	500 cytological samples, H&E staining	Competitive neural network	Acc = 98.7%
	Nuclei	2013	[44]	A. Mouelhi, et al.	24 microscopic images IHC staining, (2048×1360)	MLP	Acc = 95.5%

2.2.1.1. Tasks involving classification

In reference [35], the assessment of two suggested texture features is carried out utilizing third-party software (LNKnet package), which incorporates a neural

network classifier. The study incorporates 536 samples for training the classifier and 526 samples for testing, leading to an attained accuracy of 90%.

In reference [36], a variety of classifiers, including Support Vector Machine (SVM), k-Nearest Neighbor (KNN), and Probabilistic Neural Networks (PNN), are utilized along with techniques for ranking, selecting, and extracting features to discern between benign and malignant breast tumors. The SVM classifier achieves the highest overall accuracies in breast cancer diagnosis, with rates of 98.80% on dataset 1 (692 specimens) and 96.33% on dataset 2 (295 microarrays). Meanwhile, PNN achieves overall accuracies of 97.23% and 93.39% on dataset 1 and dataset 2, respectively.

In reference [37], a three-layer forward/back Artificial Neural Network (ANN) classifier is employed to categorize four types of H&E stained breast histopathology images using eight features. The study encompasses 1808 training samples, 387 validation samples, and 387 test samples, achieving an overall accuracy of approximately 95%.

In references [38]–[40], a scheme for automatic breast cancer classification is proposed based on histopathological images. Features like edge, texture, and intensity are extracted, and individual ANN classifiers are designed for each feature. An ensemble learning method called "random subspace ensemble" is then applied to select and combine these classifiers, resulting in a classification accuracy of 95.22% on a public image dataset.

In reference [41], breast cancer histopathology images captured at low magnification (10x) are classified into three malignancy grades. Initially, thirty texture features are extracted, and feature selection techniques are utilized to pinpoint more efficient information. A Probabilistic Neural Network (PNN) classifier is then developed using the chosen features. The experiment, which involves 65 images, attains an overall accuracy of approximately 87%.

In reference [42], morphological characteristics are utilized for distinguishing between cancerous and non-cancerous cells in histopathological images. The study, conducted on a dataset comprising 70 histopathological images, utilizes a multi-layer perceptron based on a feed-forward artificial neural network model. It

achieves an accuracy of 80%, sensitivity of 82.9%, and an Area Under the Curve (AUC) of 89.2%, respectively.

2.2.1.2. Tasks involving segmentation

In [43], regions of breast cancer in microscopic images of needle biopsies are segmented through the utilization of a competitive neural network, which serves as a clustering-based method. The process involves extracting 21 shape, texture, and topological features. The network is subsequently utilized to group images into distinct regions using these features. The study, performed on a dataset comprising over 500 images, yields an overall accuracy of about 98.7%.

In [44], an outlined supervised segmentation method is introduced for the identification of breast cancer nuclei, employing a multilayer neural network and a color active contour model. The approach is assessed on 24 images, attaining an average accuracy of 95.5%. The process flow is depicted in **Figure 2.2**.

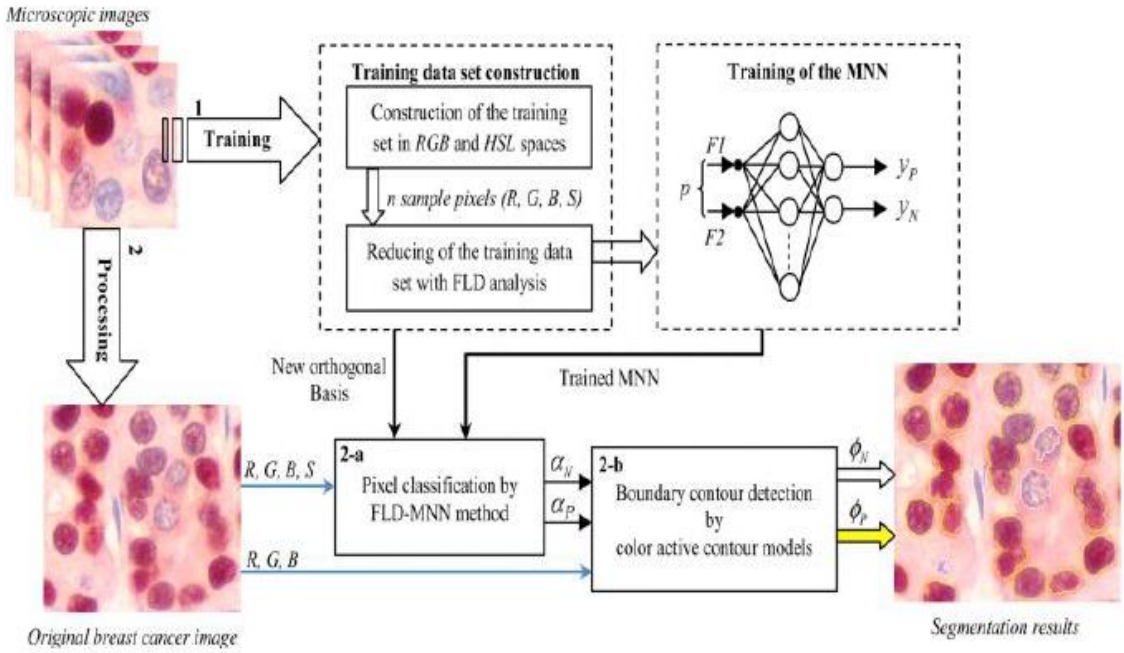


Figure 2.2: The diagram outlines the segmentation approach suggested for the detection of carcinoma nuclei in [44]. The yellow boundary signifies the delineation of the positively identified nucleus, whereas the white boundary represents the delineation of the negatively identified nucleus. This figure aligns with **Fig.2** in the original paper.

Table 2.3: Breast cancer image analysis using histopathology and conventional Artificial Neural Networks (ANNs) such as Multi-Layer Perceptron (MLP), Probabilistic Neural Networks (PNN), Multi-layer Neural Network (MNN).

Performance metrics include Accuracy (Acc) and Sensitivity (Sn). The "Detail" column in the second column provides information on the number of classes and segmentation regions. This table was taken from reference [2]

Aim	Detail	Year	Reference	Team	Data Information	ANN type	Evaluation
Classification	3	2006	[35]	S. Petushi, et al.	24 slide images, H&E staining	Neural network	Acc = 90%
	2	2010	[36]	A. Osareh, et al.	Dataset 1: 692 specimens of fine needle aspirates of breast lumps, Dataset 2: 295 microarrays	PNN	Acc = 98.8%
	4	2011	[37]	S. Singh, et al.	1080 images, H&E staining, (1080 for training, 387 for validation, 387 for test)	Feed forward back propagation neural network	Acc = 95%
	3	2011	[38]	Y. Zhang, et al.	361 images, H&E staining, (760×570)	MLP	Acc = 95.22%
		2013	[39]				
		2013	[40]				
	3	2013	[41]	C. Loukas, et al.	65 regions of interests H&E staining, (20 grade I, 20 grade II, 25 grade III)	PNN	Acc = 87%
	2	2017	[42]	K. Shukla, et al.	70 images, H&E staining, (35 non-cancerous and 35 cancerous)	MLP	Acc = 80%, Sn = 82.9%, AUC = 89.2%
Segmentation	Nuclei	2013	[43]	M. Kowal, et al.	500 cytological samples, H&E staining	Competitive neural network	Acc = 98.7%
	Nuclei	2013	[44]	A. Mouelhi, et al.	24 microscopic images IHC staining, (2048×1360)	MLP	Acc = 95.5%

2.2.1.3. Summary

Based on the above review, it is evident that the Artificial Neural Networks (ANNs) employed in Breast Histopathological Image Analysis (BHIA) around 2012 were predominantly classical neural networks. While classical neural networks exhibit remarkable performance across various domains, they do have certain limitations, including a tendency to overfit, slow training speeds, and the reliance on parameter settings based on experience. During that period, the computational speed of computers was relatively low, and there was a shortage of sufficient data for training computer systems. Consequently, extracting effective ANN features from raw data was challenging. As a result, classical neural networks in BHIA primarily functioned as classifiers. Concerning the selection of features, numerous research endeavors have utilized texture features and morphological features for segmentation and classification. **Table 2.3** presents an overview of various teams' efforts utilizing classical neural networks in the examination of histopathological images of breast cancer.

2.2.2. Utilizing deep neural networks for BHIA

In the examination of breast histopathology images, several publicly accessible datasets are commonly employed in digital pathology applications. This section categorizes related research based on the utilized datasets, followed by a chronological summary of each paper's motivation, contribution, methods, and results.

2.2.2.1. Camelyon tasks

The "Camelyon Grand Challenge" represents an initiative aimed at assessing computational systems designed for the automated identification of metastatic breast cancer within Whole Slide Images (WSIs) of sentinel lymph node biopsies.

Related works of camelyon 2016

In [90], a deep convolutional neural network (DCNN) is constructed for this purpose, achieving an Area Under the Curve (AUC) of 97%. [91] introduces a deep learning method based on GoogLeNet, using 270 images for training and 130 for testing, ultimately achieving an AUC of 92.5%. Under the same experimental conditions, [92] proposes a recurrent visual attention model with three primary components, resulting in a 96% AUC.

In [93], a rapid and dense screening framework named ScanNet is presented for metastatic breast cancer detection from Whole Slide Images (WSIs). ScanNet, implemented using the VGG-16 network, yields a Free Response Operating Characteristic (FROC) of 0.8533 and an AUC of 98.75%.

[94] introduces Multiple Magnification Feature Embedding (MMFE), a transfer learning approach for breast cancer detection in digital pathology images without requiring network training. MMFE simulates the diagnostic process of a medical professional by initially observing a low-resolution image to identify suspicious areas and then switching to a high-resolution image for confirmation. Experimental results demonstrate significant improvements in model training and prediction speed without compromising performance.

[95] summarizes the Camelyon 2016, revealing that out of 32 submitted algorithms, 25 are based on deep learning methods, and the top 19 performing algorithms all employ deep convolutional neural network (DCNN) approaches.

Related works of camelyon 2017

In the Camelyon 2017 challenge [96], a deep learning framework is introduced for the detection of four types of breast cancer from Whole Slide Images (WSIs), designed to accommodate limited computational resources. The architecture employs two Convolutional Neural Networks (CNNs) arranged in a cascade, with subsequent local maxima extraction and Support Vector Machine (SVM) classification of the identified local maxima regions. The experimental setup involves using 300 images for training, 200 for validation, and 500 for testing, ultimately achieving an accuracy of 92%.

2.2.2.2. The tasks in the "TCUG16" challenge

"In the 2012 International Conference on Pattern Recognition (ICPR), a competition known as the "mitotic figure recognition contest" took place, featuring a dataset comprising 50 High Power Fields (HPF) from 5 different slides scanned by three types of equipment at 40x magnification. Each HPF measured 512x512 pixels, containing a total of 326 mitotic cells on images from both scanners and 322 mitotic cells on the multispectral microscope [115].

In [116] and [117], a combination of manually crafted color, texture, and shape features were utilized alongside machine learning features extracted by a multi-layer CNN. This method achieved F1-scores of up to 65.9% on color scanners and 58.9% on multispectral scanners. Similarly, in [118], handcrafted features and DCNN features were employed in an ensemble learning process, yielding an F1-score of 73.5%.

For mitosis detection in breast histology images, [119] introduced a deep max-pooling CNN trained to classify each pixel in the image into a labeled region. With 26 images for training, 9 for validation, and 15 for testing, the method achieved an F1-score of 78.2%. Additionally, a similar approach in [120] resulted in an F1-score of 61.1%.

In [121], a novel deep cascade convolutional neural network (CasCNN) was developed for mitosis detection, comprising two parts. Initially, a full CNN served as a rough retrieval model to identify and locate mitotic candidates with high sensitivity. Subsequently, a fine recognition model based on cross-domain knowledge transfer singled out mitoses from the rough model. Using both the ICPR12 and ICPR14 datasets, the method achieved a precision of 80.4%, recall of 72.2%, and an F1-score of 78.8% on ICPR12, and precision of 46%, recall of 50.7%, and an F1-score of 48.2% on ICPR14.

In the summary of the ICPR 2012 contest provided in [115], 17 teams submitted results, with the IDSIA team achieving the best performance. The IDSIA team trained a CNN on ground truth mitosis from the training dataset, using it to calculate a map of mitosis probabilities on the entire image, resulting in a recall of 70%, accuracy of 89%, and an F-measure of 78%."

2.2.2.3. Tasks related to "Invasive Ductal Carcinoma (IDC)"

"In [132], a novel Convolutional Neural Network (CNN) model designed for the identification of Invasive Ductal Carcinoma (IDC) cells in histopathological slides is introduced. Derived from the Inception architecture, this model incorporates a multi-level batch normalization module between each convolutional step. The experiment involves the utilization of 94,543 patches for training, 31,514 for validation, and 151,465 for testing. Ultimately, the model achieves a balanced accuracy of 89% and an F1-score of 90%.

2.2.2.4. Summary

The adoption of deep artificial neural networks (ANNs) in the realm of Breast Histopathological Image Analysis (BHIA) has experienced a notable surge since 2012, with CNN-based methodologies emerging as predominant in this trend. Several factors contribute to the prevalence of CNNs in this domain:

i. **High-performance GPU computing**: The availability of high-performance Graphics Processing Units (GPUs) enables the training of networks with more layers, allowing for the development of deeper and more intricate architectures.

ii. **Increased availability of datasets:** Various institutions have released datasets containing breast histopathological images, addressing the challenge posed by limited labeled public datasets. The growing abundance of training data helps counteract the risk of overfitting.

iii. **Automated feature learning:** Deep learning, particularly CNNs, excels in autonomously learning features from data. This stands in contrast to traditional image classification methods reliant on manually designed features, eliminating the complexity and limitations associated with artificial feature extraction.

iv. **Widespread application of CNNs:** CNNs have demonstrated success across diverse domains, including natural language processing, object recognition, and image classification. This broad applicability establishes a robust foundation for their effective utilization in analyzing histopathological images of breast cancer.

The endeavors of various research teams in employing deep neural networks for the analysis of breast histopathological images are summarized in an overview of the examined studies on deep neural network approaches for Biomedical Image Analysis (BHIA) tasks in reference [3]

2.2.3. Evaluation of methods involving deep ANNs

In recent years, transfer learning strategies have become more prevalent in the classification of breast histopathological images using deep artificial neural networks (ANNs). The articles referenced in this analysis include [57], [58], [65], [66], [79], [83], [86], [87], [101], [107], [109], [113], [122], [126], [127], [130]. Transfer learning involves applying knowledge acquired from one task to address another [158]. **Figure 2.3** illustrates two main approaches to implementing transfer learning: (1) Fine-tuning the parameters in the pretraining network based on the required tasks (e.g., [58], [66], [101], [122], [126], [130]). (2) Utilizing a pre-trained network as a feature extractor and subsequently using these features to train a new classifier (e.g., [57], [65], [79], [83], [86], [87], [107], [109], [113], [127], [138]). VGG16 [159], VGG19, and ResNet50 [160] are popular pre-trained CNN models in transfer learning due to their more comprehensive architectures [65]. The key reasons for their popularity are as follows: First, the intrinsic complexity and

diversity of breast tissue pathological images make labeling challenging, and the cost of expert labeling is high, resulting in limited publicly available labeled datasets. Transfer learning effectively addresses the issue of small datasets [161]. Second, in the classification of breast histopathology images, most pre-trained models are derived from the ImageNet Large Scale Visual Recognition Challenge [162], demonstrating stable performance on specific tasks and safe applicability in breast cancer classification tasks. Finally, the transfer learning process contributes to enhancing accuracy or reducing training time [163], making it a crucial factor in its widespread adoption.

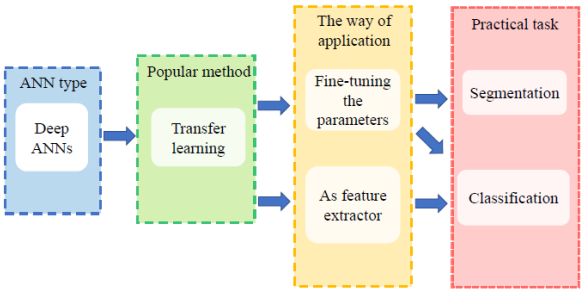


Figure 2.3: The commonly employed techniques in deep Artificial Neural Networks (ANN) for tasks related to Biomedical Image Analysis (BHIA).

2.2.4. Reviewing the approaches excelling in each task

In various review tasks, notable methods have been introduced. For instance, in the BreakHis dataset task, the most promising outcomes are reported in [81]. Here, a compact SEResNet model is devised, incorporating a combination of residual modules and Squeeze-and-Excitation blocks, effectively minimizing training parameters. Additionally, a novel learning rate scheduler called Gaussian error scheduler is proposed, achieving excellent performance without intricate fine-tuning of the learning rate. Moving to the Camelyon 2016 dataset task, optimal results are presented in [93]. To identify metastatic breast cancer in WSIs, the ScanNet framework is introduced, implemented based on the VGG-16 network with modifications to the last three fully connected layers, resulting in faster performance on tumor localization tasks that even surpasses human performance on WSI classification tasks. In the ICPR 2012 dataset task, superior outcomes are attained in [121], introducing the CasCNN, a novel deep cascaded convolutional neural network designed for mitosis detection. The CasCNN stands out for its ability

to significantly reduce detection time while maintaining satisfactory accuracy. Lastly, in the Bioimaging 2015 Breast Histology Classification Challenge dataset task, [129] presents the most effective method. This approach utilizes the Inception Recurrent Residual Convolutional Neural Network (IRRCNN) model, a hybrid DCNN architecture improved upon inception, residual networks, and RCNN architectures. The IRRCNN demonstrates superior recognition performance using equal or fewer network parameters compared to its counterparts.

2.3. Deep Learning for Histopathology

Deep learning has shown promising results in automating various tasks within histopathology analysis. Deep learning models can learn to identify patterns and features within the images that may be indicative of certain diseases or conditions. Traditional machine learning methods often relied on manually crafted features, which could be time-consuming and less effective at capturing complex patterns. Deep learning models, on the other hand, can learn hierarchical representations of features directly from the data, potentially leading to more accurate results.

2.3.1. Possible applications of deep learning in histopathology

Deep learning can be in various ways in histopathology. For example, a deep learning model could be trained to classify tissue samples as either cancerous or non-cancerous based on the features extracted from the images. These models can learn from large datasets of labeled examples, allowing them to generalize to new, unseen cases.

Another possible application is predictive modeling. For instance, researchers may use deep learning models to predict patient outcomes based on histopathology images. By analyzing patterns in the data, these models can provide valuable insights into prognosis and treatment planning.

Deep learning models can assist pathologists in their workflow by automating repetitive tasks, such as screening slides for abnormalities or prioritizing cases for further review. This can help reduce the burden on pathologists and improve efficiency in diagnostic workflows.

Another active research field of deep learning in histopathology is the discovery of novel biomarkers or disease subtypes through the analysis of large-scale datasets. Researchers

can leverage deep learning techniques to uncover hidden patterns in histopathology data that may not be apparent through traditional analysis methods.

Existing applications in the literature can be categorized according to the objectives as follows.

- **Diagnosis:** Deep learning models can automatically detect and classify various tissue structures, cell types, and pathological features, aiding pathologists in identifying diseases such as cancer. Moreover, deep learning algorithms can learn complex patterns and relationships within histopathology images, potentially uncovering subtle abnormalities that may be overlooked by human observers. Through continuous training on large annotated datasets, these models can improve their performance and generalize well to diverse tissue types and disease presentations.
- **Tumor subtyping:** Deep learning techniques have shown significant promise in tumor subtyping within histopathology analysis. Models can distinguish subtle morphological variations and molecular characteristics associated with specific tumor subtypes, providing valuable insights for personalized treatment strategies. Additionally, deep learning-based approaches can streamline the subtyping process, reducing the burden on pathologists and potentially accelerating diagnosis and treatment decisions.
- **Tumor grading:** Traditionally, tumor grading involves subjective assessment by pathologists to evaluate the aggressiveness and differentiation level of cancer cells. Models can accurately classify cancerous cells based on their morphological characteristics, such as nuclear size, shape, and staining intensity. By leveraging large datasets of annotated histopathology images, deep learning algorithms can learn to differentiate between different tumor grades with high sensitivity and specificity, providing valuable insights into disease prognosis and treatment planning.
- **Tumor staging:** Conventionally, tumor staging involves manual examination by pathologists to determine the size of the tumor, extent of local invasion, and presence of metastasis. However, deep learning models can automate this process by analyzing digitized histopathology images and extracting relevant features indicative of tumor stage. These models can learn to recognize subtle morphological changes and spatial relationships between tumor cells, enabling

precise staging predictions. By training on large annotated datasets, deep learning algorithms can generalize across diverse cancer types and anatomical locations, providing robust staging assessments.

- **Evaluation of pathological features:** It may be critical to look into some specific features, such as mitosis and budding, in diagnosis and treatment decisions. It is time consuming to manually detect and quantify such features. Deep learning models can efficiently extract features from histopathology images and quantify them according to predefined rules, enabling automatic assessment of pathological characteristics.
- **Evaluation of biomarkers:** Deep learning is revolutionizing the evaluation of biomarkers in histopathology, offering a highly sophisticated approach to deciphering complex molecular signatures within tissue samples. Traditionally, assessing biomarkers involves labor-intensive and subjective interpretation by pathologists. Deep learning models can automatically analyze digitized histopathology images and detect subtle variations in biomarker expression levels with remarkable precision. By integrating deep learning-based approaches with immunohistochemical staining and image analysis techniques, researchers can streamline biomarker evaluation, enhance reproducibility, and uncover novel biomarker associations.
- **Evaluation of genetic changes:** Traditional methods for assessing genetic changes, such as fluorescence in situ hybridization (FISH) or polymerase chain reaction (PCR), are often time-consuming and labor-intensive. Deep learning is rapidly transforming the evaluation of genetic changes in histopathology, offering a powerful approach to deciphering complex molecular alterations within tissue samples. Deep learning models can learn from vast datasets to recognize patterns indicative of genetic changes, including mutations, gene amplifications, and chromosomal rearrangements. By integrating deep learning-based approaches with genomic profiling techniques, researchers can streamline the evaluation of genetic changes, accelerate the identification of driver mutations, and unravel the molecular landscape of cancer subtypes.
- **Prognosis prediction:** Traditionally, prognosis prediction relied on manual assessment by pathologists based on histological features and clinical data. Deep learning models can learn from large datasets of annotated images and clinical

outcomes to identify subtle biomarkers and histological characteristics associated with disease progression or patient survival. By integrating deep learning-based prognostic models with clinical data, such as patient demographics and treatment history, researchers can develop more comprehensive and personalized prognostic tools.

Based on these application types (objectives), a list of deep learning applications for different cancers types is provided in **Table 2.4.** The list is not complete as the literature is growing fast. We can also direct the readers to some recent survey papers by Ahmed et al. (2022) [212] and Springenberg et al. (2023) [277].

Table 2.4: A list of deep learning based approaches for different cancer types.

Cancer type	**Application type**	**References**
Breast cancer	Diagnosis	(Araujo et al., 2017 [215]; Bejnordi, Zuidhof et al., 2017 [219]; Bejnordi et al., 2018 [215]; Mercan et al., 2018 [256]; Maleki et al., 2023 [256])
	Tumor subtyping	(Jiang et al., 2019 [244])
	Tumor grading	(Wan et al., 2017 [291])
	Tumor staging	(Cruz-Roa et al., 2017 [227]; Cruz-Roa et al., 2018 [228]; Bejnordi, Veta et al., 2017 [219], Liu et al., 2019 [254]; Steiner et al., 2018 [278])
	Evaluation of pathological features	(Veta et al., 2015 [287]; Saha et al., 2017 [287]; Veta et al., 2019 [286]; Turkki et al., 2016 [284])
	Evaluation of biomarkers	(Vandenberghe et al., 2017 [285]; Mandair et al., 2023 [257])
	Evaluation of genetic changes	(Mondol et al., 2023 [260])

Cervical cancer	Diagnosis	(Zhang et al., 2017 [295])
	Tumor subtyping	(Wu et al., 2018 [294])
	Tumor staging	(Liu et al., 2023 [])
Colorectal cancer	Diagnosis	(Kainz et al., 2017; Awan et al., 2017 [245])
	Tumor subtyping	(Korbar et al., 2017 [248])
	Evaluation of pathological features	(Weis et al., 2018 [293])
	Evaluation of genetic changes	(Kather et al., 2019 [246])
Gastric cancer	Diagnosis	(Wang et al., 2019 [290])
	Evaluation of genetic changes	(Sharma et al., 2017 [272])
Glioma	Tumor grading	(Zhuge et al., 2020 [299])
	Prognosis prediction	(Mobadersany et al., 2018 [259])
Lung cancer	Tumor subtyping	(Teramoto et al., 2017 [282]; Coudray et al., 2018 [226]; Gertych et al.i 2019 [235]; Wei et al., 2019 [292]; Aprupe et al., 2019 [214])
	Evaluation of biomarkers	(Sha et al., 2019 [271]; Wang et al., 2018 [288])
	Prognosis prediction	(Zhang et al., 2024 [296])
Prostate cancer	Tumor grading	(Arvaniti et al., 2018 [216])
	Evaluation of genetic changes	(Schaumberg et al., 2016 [269])
Thyroid cancer	Diagnosis	(Guan et al., 2019 [237])
	Tumor subtyping	(Wang et al., 2019 [289])

CHAPTER 3

3. EXPERIMENTAL PART

The thyroid, a small gland shaped like a butterfly located in the neck just in front of the windpipe, produces hormones that influence heart rate and body temperature [193]. Tumors related to branchial-thymic tissues can occur in or around the thyroid [194]. Thyroid follicular cells, also known as thyrocytes, constitute the primary cell type in the thyroid gland and are responsible for producing and secreting thyroid hormones, namely thyroxine (T4) and triiodothyronine (T3) [195]. Pathologists use the term "Benign Thyrocyte nodule" to describe a group of non-cancerous conditions in the thyroid gland. Typically, this diagnosis follows a procedure known as fine-needle aspiration or FNA. Symptoms involved in this condition are enlarged thyroid gland or some small flat groups called micro follicles around the thyroid gland [196]. The histiocyte is another type of tissue macrophage and their job is to clear out Neutrophils once they've reached the end of their lifespan [197]. Neutrophils, a subtype of white blood cells, are responsible for eliminating and breaking down bacteria and fungi. Lymphocytes, also white blood cells and integral to the immune system, come in two primary types: B cells and T cells. B cells generate antibodies employed to combat invading bacteria, viruses, and toxins [198]. Related to these three tissue types, the thyroid glands can develop three major types of cancerous conditions [199]. The diagnosis of these conditions can be done with some chemical laboratory tests (which are time-consuming, slow, and require expert staff) as well as with help of microscopic image analysis. In this paper, we have proposed a microscopic image analysis method for the diagnosis of three cancer types, benign thyrocyte (non-cancer), histiocyte (cancer type 1), and lymphocyte (cancer type 2), with help of convolutional neural networks.

A convolutional neural network (CNN) is a deep learning-based classification model that does image recognition and required no manual feature extractions from the images. CNN

development in deep learning is boosted to analyze immense data for development and further improvement of the computation power [200]. LeNet is a pattern recognition-based CNN architecture that was used by the banks so they can read the cheques with high performance [201]. ‘‘ImageNET Challenge’’ was held in which different teams were given the task to classify the natural images from thousands of categories [202].

Automatic feature extraction is a healthy change in artificial intelligence that overcome the flaws of manual extraction of features from images [203]. In short, deep learning has become a popular approach in computer vision for the categorization of text, images, and sound [12],[13]. Medical imaging uses deep learning as a tool for rapid diagnoses of different disorders, for example, common cancer [206], neurological abnormalities [207], kidney and liver infections [208], and musculoskeletal injuries [207]. CNN is applied in disease recognition utilizing medical image datasets sourced from INbreast and BreakHis for breast cancer assessment [209], the Danish Lung Cancer Screening Trial for lung nodule detection [210], and the Brain Tumor Segmentation Challenge (BraTS) for identifying brain cancer [211]. In this study, we have used locally generated thyroid histopathologic microscopic images and classified two types of thyroid cancers in the non-cancer class.

3.1. Method Discussion

3.1.1. Components of deep learning based analysis

3.1.1.1. Model architectures

For analyzing histopathology images; classification, segmentation and detection model architectures are typically used. In classification applications, deep learning models are used to classify histopathology images into different tumor types or disease states. In segmentation applications, tissue regions are segmented from background and specific structures, like nuclei or tumor boundaries, are identified. In detection applications abnormalities such as tumors or metastasis within tissue samples are detected and identified.

Classification architectures: Deep learning architectures for image classification have evolved significantly over the years, with several key models demonstrating remarkable performance across various datasets and tasks. One of the pioneering architectures is the

Convolutional Neural Network (CNN), which has become the cornerstone of image classification tasks due to its ability to automatically learn hierarchical features from raw pixel data. LeNet-5, proposed by LeCun et al. (1998) [250], was one of the earliest CNN architectures designed for handwritten digit recognition. AlexNet, introduced by Krizhevsky et al. (2012) [249], significantly advanced CNNs' capabilities by employing deeper architectures and ReLU activation functions, achieving breakthrough results in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) competition. VGGNet, proposed by Simonyan and Zisserman (2014) [275], further deepened networks and introduced smaller filter sizes, demonstrating improved performance on image classification tasks. GoogLeNet (Inception), introduced by Szegedy et al. (2014) [166], introduced the inception module, enabling efficient use of computational resources by employing multiple filter sizes within a single layer. ResNet (Residual Network), proposed by He et al. (2015), introduced skip connections to alleviate the vanishing gradient problem, enabling training of extremely deep networks with over a hundred layers. These architectures laid the foundation for subsequent advancements, including MobileNet (Howard et al., 2017 [242]), DenseNet (Huang et al., 2017) [243], EfficientNet (Tan et al., 2019) [281] and more, each offering unique features and optimizations to further improve image classification performance.

Recently, transformer-based architectures that have been developed for image classification tasks. The Vision Transformer (ViT) architecture, introduced by Dosovitskiy et al. (2020) [230], applies the transformer architecture, originally proposed for natural language processing tasks, to the field of computer vision. Unlike traditional convolutional neural networks (CNNs), which process images using convolutional layers, ViT divides an input image into fixed-size patches and flattens them into sequences of tokens. These token sequences are then fed into a transformer encoder, which consists of multiple layers of self-attention mechanisms and feed-forward neural networks. The self-attention mechanism allows the model to capture global dependencies between patches, while the feed-forward networks enable non-linear transformations of the token representations. ViT has shown impressive performance on various image classification benchmarks, demonstrating the efficacy of transformers in computer vision tasks.

The Swin Transformer, proposed by Liu et al. (2021) [255], further advances the transformer architecture for image classification by introducing hierarchical representations and multi-scale windows. Unlike ViT, which processes images at a fixed resolution, Swin Transformer adopts a hierarchical design that gradually aggregates information across multiple levels of spatial resolutions. This hierarchical processing enables the model to capture both local and global contextual information efficiently. Additionally, Swin Transformer introduces the concept of shifted windows, where each transformer block operates on non-overlapping windows of the input feature map. By leveraging shifted windows, Swin Transformer reduces computational complexity and memory requirements while maintaining strong performance on image classification tasks. Overall, Swin Transformer represents a significant advancement in transformer-based architectures for computer vision, offering state-of-the-art results on various benchmarks.

Segmentation architectures: Deep learning architectures for image segmentation aim to accurately delineate objects or regions of interest within images. Several notable architectures have been developed for this purpose, each offering unique features and capabilities.

U-Net, introduced by Ronneberger et al. (2015) [266], is a widely used architecture for biomedical image segmentation. It consists of a contracting path, which captures context and downsamples the input image, and an expansive path, which enables precise localization and upsamples the feature map to the original resolution. U-Net employs skip connections between corresponding layers in the contracting and expansive paths to preserve spatial information and facilitate precise segmentation.

DeepLab, proposed by Chen et al. (2014) [224], is a family of architectures designed for semantic image segmentation. It utilizes atrous (dilated) convolutions to increase the receptive field of convolutional layers without increasing the number of parameters. DeepLab also incorporates spatial pyramid pooling modules to capture multi-scale contextual information and refine segmentation results.

Mask R-CNN, introduced by He et al. (2017) [238], extends the Faster R-CNN architecture for object detection to perform instance segmentation. It predicts both

bounding boxes and segmentation masks for each object in an image, enabling precise delineation of object boundaries. Mask R-CNN achieves this by adding a parallel branch to the Faster R-CNN network, which generates pixel-wise segmentation masks.

UNet++, proposed by Zhou et al. (2018) [298], enhances the U-Net architecture by incorporating nested and dense skip pathways. It introduces dense convolutional blocks within skip connections to capture more fine-grained features and improve segmentation accuracy. UNet++ achieves superior performance compared to U-Net by enabling better feature reuse and propagation across different scales.

DeepLabv3+, an extension of the DeepLab architecture, introduced by Chen et al. (2018) [224], incorporates encoder-decoder modules with atrous spatial pyramid pooling and bilinear upsampling. It employs a depthwise separable convolutional backbone for efficient feature extraction and context aggregation, enabling high-resolution semantic segmentation with improved accuracy.

Detection architectures: Deep learning architectures for object detection in images aim to identify and localize objects within images, typically by predicting bounding boxes and class labels for each object instance. Several prominent architectures have been developed for this task:

Faster R-CNN, introduced by Ren et al. (2015) [265], is a seminal architecture for object detection that consists of two main components: a region proposal network (RPN) for generating candidate object bounding boxes and a region-based convolutional neural network (CNN) for classifying and refining these proposals. Faster R-CNN achieves state-of-the-art performance by jointly optimizing the region proposal and object detection tasks.

YOLO (You Only Look Once), proposed by Redmon et al. (2016) [263], is an efficient object detection architecture that frames object detection as a regression problem, directly predicting bounding boxes and class probabilities from a single pass through the network. YOLO divides the input image into a grid and predicts bounding boxes and class probabilities for each grid cell, achieving real-time performance on object detection tasks.

SSD (Single Shot MultiBox Detector), introduced by Liu et al. (2016) [253], is another efficient object detection architecture that combines the advantages of region proposal methods (like Faster R-CNN) and single-shot detection methods (like YOLO). SSD

predicts object bounding boxes and class labels at multiple scales using feature maps from different convolutional layers, enabling accurate detection of objects of varying sizes.

RetinaNet, proposed by Lin et al. (2017) [251], addresses the problem of class imbalance in object detection by introducing a novel focal loss function that downweights the contribution of easy examples during training. This allows RetinaNet to effectively handle datasets with a large number of background examples, leading to improved performance on object detection tasks.

EfficientDet, introduced by Tan et al. (2020) [280], extends the EfficientNet architecture for object detection by incorporating a scalable and efficient compound scaling method. EfficientDet achieves state-of-the-art performance by balancing model efficiency and accuracy across different model scales, enabling effective object detection on resource-constrained devices.

3.1.1.2. Loss functions

Loss functions play a crucial role in training deep learning models for image classification tasks by quantifying the difference between predicted and ground truth labels (Goodfellow et al., 2016 [236]; Lin et al., 2017 [251]). Choosing an appropriate loss function depends on the specific characteristics of the classification task, such as the number of classes, class imbalance, and desired model behavior. Several loss functions are commonly used for this purpose:

- **Cross-entropy loss (Log Loss):** Cross-entropy loss is widely used for multi-class classification tasks. It measures the dissimilarity between predicted class probabilities and the ground truth one-hot encoded labels. Cross-entropy loss encourages the model to assign high probabilities to the correct classes and low probabilities to incorrect classes.
- **Binary cross-entropy loss:** Binary cross-entropy loss is used for binary classification tasks. It computes the cross-entropy loss for each class independently and averages them. Binary cross-entropy loss is suitable when the classification task involves distinguishing between two classes.
- **Categorical hinge loss:** Categorical hinge loss is an alternative to cross-entropy loss, particularly useful for multi-class classification with SVM-like behavior. It penalizes incorrect predictions based on the margin between predicted class scores and the ground truth scores.

- **Weighted cross-entropy loss:** Weighted cross-entropy loss assigns different weights to classes based on their importance or frequency in the dataset. It helps address class imbalance issues by penalizing misclassifications of minority classes more heavily.
- **Focal loss:** Focal loss is designed to address class imbalance by down-weighting the loss for well-classified examples. It introduces a modulating factor that reduces the loss contribution of easy examples, focusing training on hard examples and improving performance on imbalanced datasets.

Loss functions for image segmentation tasks aim to measure the discrepancy between the predicted segmentation masks and the ground truth masks. The loss functions used for classification problems, such as cross-entropy and focal loss, can also be adapted for segmentation by treating each pixel as a binary classification problem (foreground vs. background). In addition, there are loss functions that are more suitable for segmentation problems (Ronneberger et al., 2015 [267]):

- **Dice loss:** Dice loss is based on the Dice coefficient, which measures the overlap between the predicted and ground truth masks. It's computed as twice the intersection area divided by the sum of areas of the predicted and ground truth masks. Dice loss encourages accurate segmentation by penalizing deviations from the ground truth mask.
- **Jaccard loss (Intersection over Union Loss):** Jaccard loss, also known as the intersection over union (IoU) loss, measures the similarity between the predicted and ground truth masks. It's computed as one minus the IoU, encouraging the model to maximize the overlap between segmented regions.

Object detection tasks typically have two loss components: localization loss and classification loss (Ren et al., 2015 [265]; Redmon et al. [264], 2017; Lin et al., 2017 [251]). Localization loss measures the discrepancy between predicted bounding box coordinates and ground truth bounding box coordinates. Common localization loss functions include L1 Loss and L2 Loss (Euclidean Loss). Classification loss quantifies the difference between predicted class probabilities and ground truth class labels. Common classification loss functions include cross-entopy loss and focal loss.

3.1.1.3. Optimization methods

Parameter optimization methods aim to find the optimal set of model parameters that minimize a predefined loss function, thereby improving the model's performance on a given task. The optimization algorithms differ in their convergence properties, computational efficiency, and robustness to various optimization challenges. The choice of optimization algorithm depends on factors such as the characteristics of the dataset, the architecture of the neural network, and the computational resources available. Several optimization algorithms are commonly used in deep learning:

- Stochastic Gradient Descent (SGD) is a foundational optimization algorithm used in deep learning. It updates model parameters iteratively based on the gradient of the loss function with respect to each parameter. SGD works by taking small steps in the direction of the negative gradient to minimize the loss function.
- Mini-batch Gradient Descent is an extension of SGD that updates model parameters using mini-batches of training data rather than the entire dataset at once (Hinton et al., 2020 [240]). Mini-batch gradient descent reduces computational burden and improves convergence speed by leveraging parallelism and vectorization.
- Adagrad (Adaptive Gradient Algorithm) is an adaptive learning rate optimization algorithm that scales the learning rate for each parameter based on the historical gradients accumulated during training (Duchi et al., 2011 [231]). Adagrad performs well in settings with sparse data but may suffer from diminishing learning rates over time.
- RMSprop (Root Mean Square Propagation) is a variant of Adagrad that addresses its diminishing learning rate issue by using an exponentially decaying average of squared gradients to adjust the learning rate (Tieleman & Hinton, 2012 [283]). RMSprop is effective for non-stationary problems and converges faster than vanilla SGD.
- Adam (Adaptive Moment Estimation) is a popular optimization algorithm that combines elements of momentum and RMSprop (Kingma & Ba, 2014 [274]). It maintains exponentially decaying averages of past gradients and squared gradients to adaptively adjust the learning rate for each parameter. Adam is known for its robustness and efficiency in training deep neural networks.

3.1.1.4. Data augmentation

Data augmentation techniques are crucial for enhancing the generalization capability of deep learning models in image classification, segmentation and detection tasks (Shorten & Khoshgoftaar, 2019 [273]; Cubuk et al., 2019 [229]). Some common techniques include rotation, horizontal and vertical flipping, scaling, translation, brightness and contrast adjustment, noise addition and color shifting (modifying hue, saturation and intensity). These data augmentation techniques help improve the robustness, generalization, and performance of deep learning models especially when the dataset size is limited.

3.1.1.5. Performance evaluation metrics

Evaluation metrics for image classification: Evaluation metrics for image classification assess the performance of models in correctly categorizing images into predefined classes (Sokolova & Lapalme, 2009 [276]; Boyd & Eng, 2020 [222]). Some common metrics include:

- **Accuracy:** Accuracy measures the proportion of correctly classified images out of the total number of images. It's a straightforward metric but might not be sufficient for imbalanced datasets.
- **Precision:** Precision measures the ratio of true positive predictions to the total number of positive predictions. It quantifies the model's ability to correctly classify positive instanc es and avoid false positives.
- **Recall (Sensitivity):** Recall measures the ratio of true positive predictions to the total number of actual positive instances in the dataset. It quantifies the model's ability to capture all positive instances, minimizing false negatives.
- **F1 Score:** The F1 score is the harmonic mean of precision and recall, providing a balance between the two metrics. It's useful when there's an imbalance between classes or when both precision and recall are crucial.
- **ROC curve and AUC:** Receiver Operating Characteristic (ROC) curve plots the true positive rate (sensitivity) against the false positive rate (1 - specificity) at various threshold values. Area Under the ROC Curve (AUC) summarizes the ROC curve, providing a single value that represents the model's overall performance across all thresholds.

- **Confusion matrix:** A confusion matrix summarizes the actual versus predicted class labels, providing insights into the model's performance across different classes. From the confusion matrix, various metrics such as precision, recall, and F1 score can be derived.

Evaluation metrics for segmentation: Evaluation metrics for image segmentation assess the performance of models in accurately delineating objects or regions of interest within images (García-García et al., 2017; Chavhan et al., 2020 [332]). Some common metrics include:

- **Intersection over union (IoU):** IoU measures the overlap between the predicted segmentation mask and the ground truth mask, calculated as the intersection area divided by the union area. It ranges from 0 (no overlap) to 1 (perfect overlap) and is a widely used metric for image segmentation evaluation.
- **Dice coefficient:** The Dice coefficient is another measure of overlap between the predicted and ground truth masks, calculated as twice the intersection area divided by the sum of areas of the predicted and ground truth masks. Like IoU, it ranges from 0 to 1, with higher values indicating better segmentation accuracy.
- **Pixel accuracy:** Pixel accuracy measures the proportion of correctly classified pixels in the segmentation mask out of the total number of pixels. It provides a pixel-level assessment of segmentation accuracy but may not be informative for imbalanced datasets.
- **Mean intersection over union (mIoU):** mIoU computes the average IoU across all classes or regions of interest in the dataset. It provides a comprehensive measure of segmentation performance, considering both class-wise and overall segmentation accuracy.
- **F1 score:** F1 score can also be adapted for segmentation tasks, considering each pixel as a binary classification problem (foreground vs. background). It computes the harmonic mean of precision and recall at the pixel level, providing a balanced measure of segmentation accuracy.
- **Precision-recall curve (PR Curve):** Similar to classification tasks, precision-recall curves can be generated for segmentation tasks by varying segmentation thresholds. Area Under the PR Curve (AUC-PR) summarizes the overall segmentation performance across different thresholds.

Evaluation metrics for detection: Evaluation metrics for object detection in images assess the performance of models in accurately detecting and localizing objects within images (Everingham et al., 2010; Lin et al., 2014). The evaluation metrics for segmentation can also be used for detection, where segments are replaced with bounding boxes. Some common metrics include:

- **Intersection over union (IoU):** IoU measures the overlap between the predicted bounding boxes and the ground truth bounding boxes. It's calculated as the intersection area divided by the union area of the two bounding boxes. IoU values range from 0 (no overlap) to 1 (perfect overlap), with higher values indicating better detection accuracy.
- **Average precision (AP):** AP computes the average precision-recall curve for object detection by varying the detection threshold. It measures the trade-off between precision and recall at different confidence thresholds for detecting objects.
- **Mean average precision (mAP):** mAP calculates the average AP across multiple object classes or categories. It provides a comprehensive measure of detection performance, considering both class-wise and overall detection accuracy.
- **Precision-recall curve (PR Curve):** Similar to AP, precision-recall curves plot the precision against recall at different confidence thresholds for object detection. Area Under the PR Curve (AUC-PR) summarizes the overall detection performance across different thresholds.
- **F1 score:** F1 score is the harmonic mean of precision and recall, providing a balanced measure of detection accuracy. It's particularly useful when there's an imbalance between positive and negative samples in the dataset.
- **False positive rate (FPR):** FPR measures the ratio of falsely detected objects (false positives) to the total number of actual negative instances in the dataset. It quantifies the model's ability to avoid false alarms during object detection.
- **Receiver operating characteristic (ROC) curve and area under the curve (AUC-ROC):** ROC curve plots the true positive rate (sensitivity) against the false positive rate (1-specificity) at various confidence thresholds. AUC-ROC summarizes the overall detection performance across all thresholds.

3.1.2. Interpretability and explainability

Interpreting deep learning models' decisions in histopathology analysis poses significant challenges due to the inherent complexity of these models and the lack of transparency in their decision-making process. Deep learning models, particularly convolutional neural networks (CNNs), are capable of learning intricate patterns and features from large-scale histopathology image datasets, enabling accurate disease diagnosis and prognosis prediction. However, understanding how these models arrive at their decisions and explaining their predictions to clinicians and pathologists is essential for building trust and facilitating their adoption in clinical practice.

One major challenge in interpreting deep learning models' decisions in histopathology analysis is the black-box nature of these models. CNNs operate by transforming input images through multiple layers of non-linear transformations, making it challenging to discern which features contribute most to their predictions. Additionally, CNNs can learn complex, abstract representations that may not align with human-understandable concepts, further complicating interpretation.

To address these challenges, various techniques for interpretability and explainability have been proposed:

Saliency maps: Saliency maps highlight the most discriminative regions of input images that contribute to the model's prediction. Techniques like Gradient-weighted Class Activation Mapping (Grad-CAM) visualize the gradient of the predicted class score with respect to the input image pixels, providing insights into which image regions the model focuses on (Selvaraju et al., 2017).

Attention mechanisms: Attention mechanisms, inspired by human visual attention, enable models to selectively focus on relevant image regions. Techniques such as self-attention mechanisms highlight important features and regions within images, aiding in understanding the model's decision-making process (Zhou et al., 2016).

Layer-wise relevance propagation (LRP): LRP attributes the model's prediction to individual input features, propagating relevance scores backward through the network layers. By visualizing which input features contribute most to the model's decision, LRP provides insights into the model's internal representations (Bach et al., 2015).

Feature visualization: Feature visualization techniques generate synthetic input images that maximally activate specific neurons or feature maps in the network. By visualizing

the patterns that excite individual neurons, these techniques offer insights into the types of features learned by the model (Simonyan et al., 2013).

Model distillation: Model distillation involves training a smaller, more interpretable model to mimic the predictions of a larger, complex model. The smaller model is easier to interpret while retaining most of the larger model's predictive performance, making it more suitable for deployment in clinical settings (Hinton et al., 2015).

These techniques aim to enhance the interpretability and explainability of deep learning models in histopathology analysis, enabling clinicians and pathologists to understand and trust the models' predictions. However, further research is needed to develop more interpretable deep learning architectures and techniques tailored specifically for histopathology applications.

Among these techniques, Grad-CAM is widely adopted. It is used for visualizing the areas of an image that a convolutional neural network (CNN) focuses on when making predictions. It provides insight into which regions of the input image contribute the most to the model's decision.

The key idea behind Grad-CAM is to use the gradients of the predicted class score with respect to the feature maps of the last convolutional layer of the CNN. These gradients are used to compute the importance of each feature map for the predicted class, effectively highlighting the regions of the input image that activate the most relevant features.

The Grad-CAM technique works as follows:

- **Forward Pass:** The input image is fed forward through the CNN until the last convolutional layer. Feature maps are generated at each convolutional layer, capturing different levels of abstraction in the input image.
- **Compute Gradient:** Gradients of the predicted class score with respect to the feature maps of the last convolutional layer are computed using backpropagation. These gradients represent how much each feature map contributes to the predicted class score.
- **Global Average Pooling:** The gradients are then globally averaged across the spatial dimensions of each feature map. This step ensures that the importance scores are calculated based on the entire feature map rather than individual spatial locations.

- **Weighted Combination:** The importance scores are multiplied by the corresponding feature maps to generate the class activation map. This map highlights the regions of the feature maps that are most relevant for predicting the target class.
- **Heatmap Visualization:** Finally, the class activation map is upsampled to the original image size and overlaid onto the input image as a heatmap. The heatmap visually indicates the regions of the input image that the model focuses on when making predictions for the target class.

Grad-CAM provides interpretable visualizations of CNN predictions, enabling users to understand which parts of the input image contribute most to the model's decision-making process. It has applications in various domains, including medical imaging, object detection, and natural language processing, where interpretability is crucial for model deployment and decision-making. There several other methods built upon the Grad-CAM technique, including Grad-CAM++ (Chattopadhay et al., 2018 [223]), Ablation-CAM (Desai et al., 2020 [262]), XGrad-CAM (Fu et al., 2020 [233]) and Eigen-CAM (Bany et al., 2020 [261])

3.2. Methodology and Experimental Results

As an experimental part of thesis, we investigate classification of thyroid histopathology images on an original dataset, obtained from Ümraniye Devlet Hastanesi Patoloji Kliniği, using various deep learning architectures. Our objective is to evaluate whether it is possible to reliably distinguish benign and malignant tumors from thyroid histopathology images. We tested the well-known and widely used models, including VGG16 (Simonyan and Zisserman, 2014 [275]), ResNet (He , et al,. 2015 []), Inception (Szegedy, et al,. 2015 [279]), MobileNet (Howard et al., 2017 [242]), EfficientNet (Tan, Le, 2020 [280]), Swin Transformer (Liu, et al., 2021 [255]) and ConvNeXt (Liu, et al., 2022 [300]). For these models, we used the pre-trained weights, modified the last layer to fit to our task, and fine-tuned the model parameters with our dataset. In addition to transfer learning approach, we built a custom CNN model, which has much less number of parameters; and trained it from scratch. Finally, we used explainability techniques to see if help to identify regions of interest, which can guide pathologists.

3.2.1. Dataset

The dataset used in this study was generated in Ümraniye Devlet Hastanesi Pathology Clinique as Thyroid Histopathology data. The images taken from an Olympus DP20 Microscopic Camera used by the Pathology Clinique where the samples are usual pap stained pathology lamels of the patients. Names and other private information of the patients have been hidden by the doctors before the images were taken, in order to protect patient privacy. The dataset contains images of three classes, which are benign thyrocyte, histiocyte, and lymphocyte. The benign class contains 100 microscopic images, the histiocyte class has 17 microscopic images whereas the thyrocyte class has 20 images of resolution 1600 x 1200. To feed the balanced data to the network, we kept 20 images from benign thyrocytes and lymphocytes while we used all 17 images from the histiocyte.

So, the dataset organized as 100 benign and 148 malignant images, each one is of size 1600-by-1200 pixels. **Figure 3.1** shows two examples from the dataset.

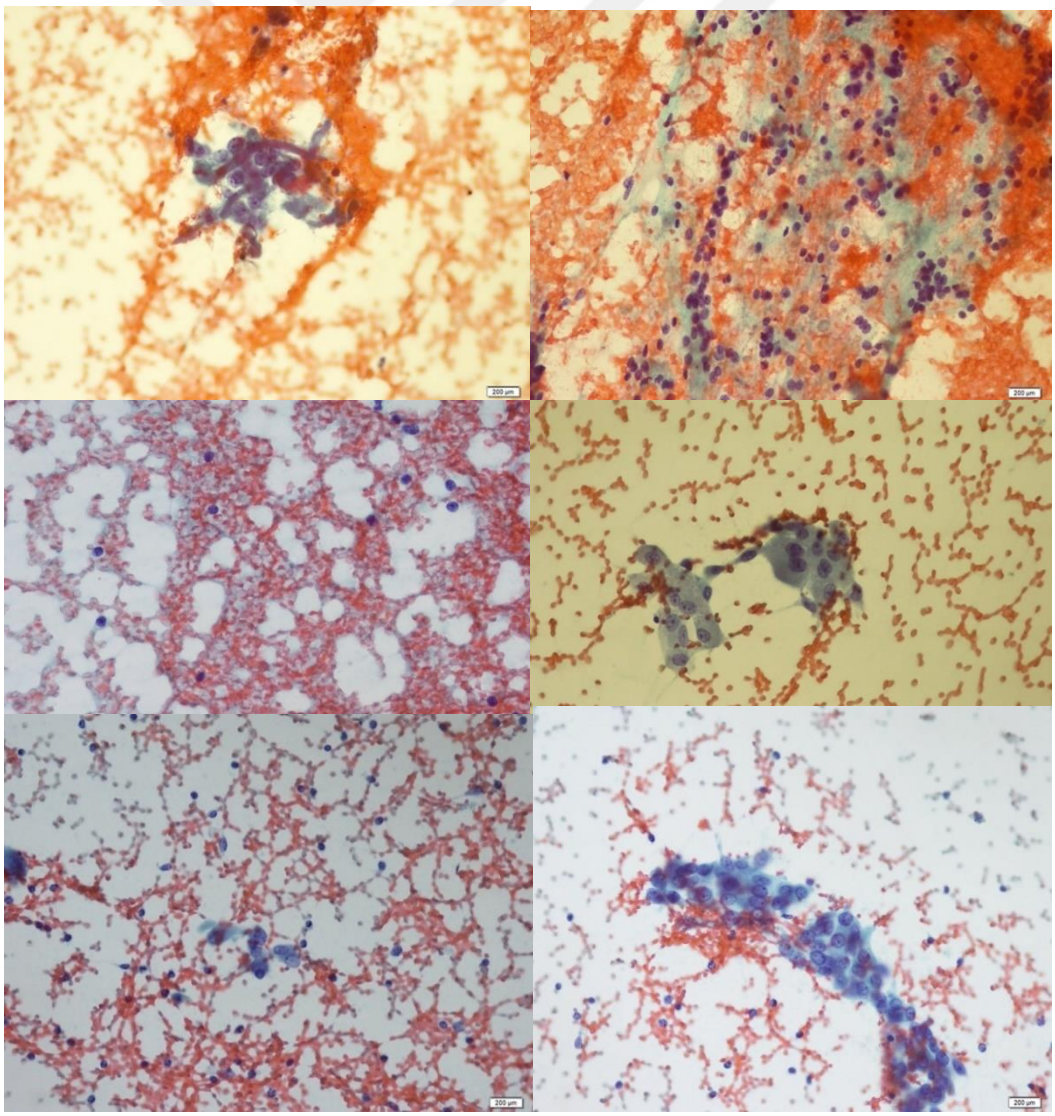


Figure 3.1: Examples of benign and non-benign (malignant) histopathology images. (Top row) Ben ign. (Middle left) Non-benign (lymphocytosis). (Middle right) Non-benign (histiocytosis). (Bottom row) Non-benign (atypical).

3.2.2. Model training

The dataset is split into training and testing parts with 80-to-20 percent ratio. There are 198 images for training and 50 images for testing. To compensate for the low number images, we utilized data augmentation techniques. The training stage consists of the following data augmentation steps:

- Random horizontal flip, which randomly flips the image horizontally with a 50% chance.
- Random vertical flip, which randomly flips the image vertically with a 50% chance.
- Random rotation, which randomly rotates the image by up to 20 degrees in either direction.
- Random crop, randomly crops a 512-by-512 patch from the original image. Padding with reflection is used if the image is smaller than 512-by-512.

The pre-trained models are modified by replacing the final layer to match the class size of 2. Mini-batch stochastic gradient descent technique is used with a batch size of 16, learning rate of 0.01, and cross-entropy loss as the loss function. Train and test loss values and accuracies are stored throughout the iterations. A batch of input samples is given **Figure 3.2.**

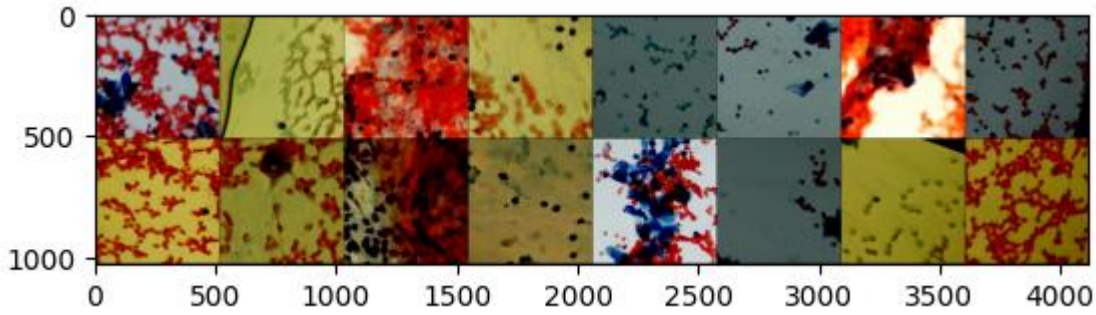


Figure 3.2: A batch of input samples during the training process.

The custom model, which we designed and called SimpleCNN, is essentially a CNN architecture. As shown in **Figure 3.3**, it has convolutional, batch normalization and pooling layers, placed in sequence. Final layers are average pooling and fully connected

layers. The convolutional layers extract the features in the images. Batch normalization layers help better handle variations in image colors; and max pooling layers provide positional invariance by handling translational variations.

```
SimpleCNN(
  (conv1): Conv2d(3, 32, kernel_size=(3, 3), stride=(1, 1), padding=(2, 2))
  (conv2): Conv2d(32, 64, kernel_size=(3, 3), stride=(1, 1), padding=(2, 2))
  (conv3): Conv2d(64, 128, kernel_size=(3, 3), stride=(1, 1), padding=(2, 2))
  (conv4): Conv2d(128, 256, kernel_size=(3, 3), stride=(1, 1), padding=(2, 2))
  (conv5): Conv2d(256, 512, kernel_size=(3, 3), stride=(1, 1), padding=(2, 2))
  (bn1): BatchNorm2d(32, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (bn2): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (bn3): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (bn4): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (bn5): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (pool): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
  (avg): AvgPool2d(kernel_size=8, stride=8, padding=0)
  (fc): Linear(in_features=2048, out_features=2, bias=True)
)

==========================================================================================
Layer (type:depth-idx)                   Output Shape              Param #
==========================================================================================
SimpleCNN                                [1, 2]                    --
├─Conv2d: 1-1                            [1, 32, 514, 514]         896
├─BatchNorm2d: 1-2                       [1, 32, 514, 514]         64
├─MaxPool2d: 1-3                         [1, 32, 257, 257]         --
├─Conv2d: 1-4                            [1, 64, 259, 259]         18,496
├─BatchNorm2d: 1-5                       [1, 64, 259, 259]         128
├─MaxPool2d: 1-6                         [1, 64, 129, 129]         --
├─Conv2d: 1-7                            [1, 128, 131, 131]        73,856
├─BatchNorm2d: 1-8                       [1, 128, 131, 131]        256
├─MaxPool2d: 1-9                         [1, 128, 65, 65]          --
├─Conv2d: 1-10                           [1, 256, 67, 67]          295,168
├─BatchNorm2d: 1-11                      [1, 256, 67, 67]          512
├─MaxPool2d: 1-12                        [1, 256, 33, 33]          --
├─Conv2d: 1-13                           [1, 512, 35, 35]          1,180,160
├─BatchNorm2d: 1-14                      [1, 512, 35, 35]          1,024
├─MaxPool2d: 1-15                        [1, 512, 17, 17]          --
├─AvgPool2d: 1-16                        [1, 512, 2, 2]            --
├─Linear: 1-17                           [1, 2]                    4,098
==========================================================================================
Total params: 1,574,658
Trainable params: 1,574,658
Non-trainable params: 0
...
Input size (MB): 3.15
Forward/backward pass size (MB): 267.53
Params size (MB): 6.30
Estimated Total Size (MB): 276.97
==========================================================================================
```

Figure 3.3: Custom CNN model, which is called SimpleCNN.

As the models are trained; train loss, train accuracy, test loss and test accuracy are recorded and tracked to analyze the convergence and to determine if there is over- or under-fitting situations. A set of typical convergence curves are shown in **Figure 3.4**.

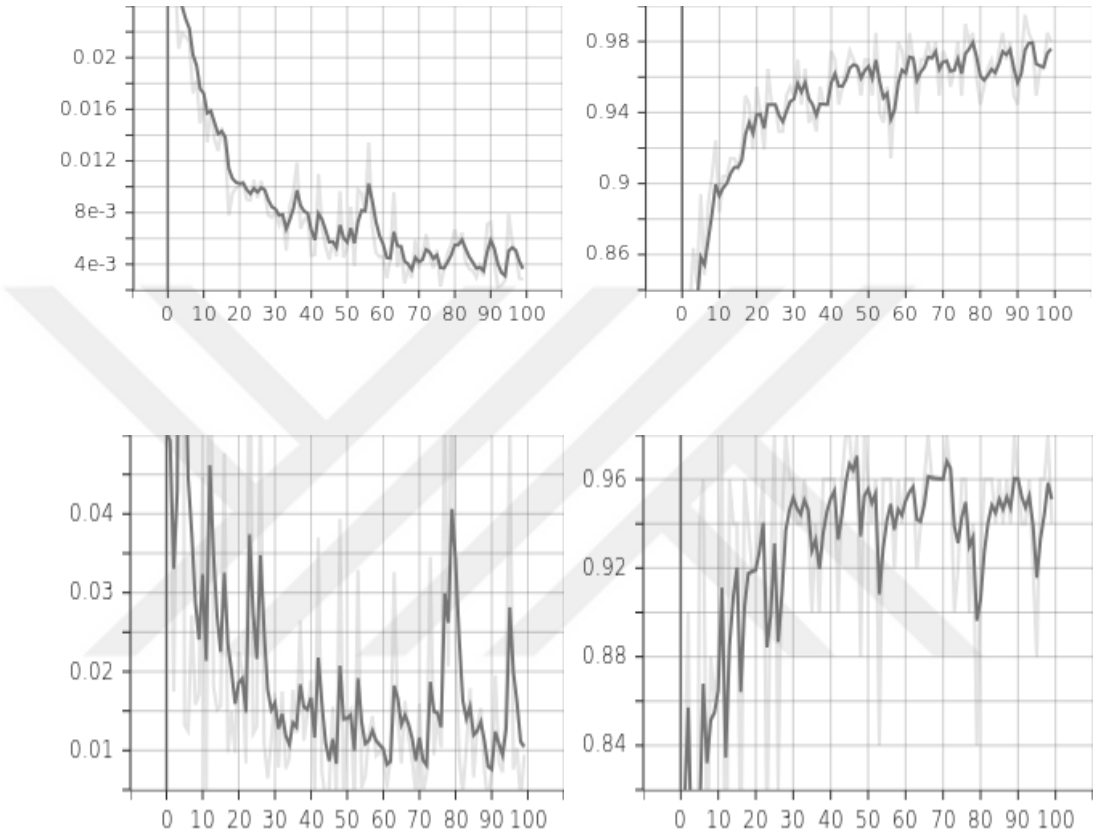


Figure 3.4: Typical convergence curves throughout the iterations (epochs). (Top left) Training loss. (Top right) Train accuracy. (Bottom left) Test loss. (Bottom right) Test accuracy. (Note: The curves shown are for the Swin-Transformer training process.)

Once the training is completed, the confusion matrices for each model extracted. In **Figure 3.5**, confusion matrices for MobileNet and SimpleCNN architectures are shown. In the figure, label "0" indicates benign cases, label "1" indicates malignant cases.

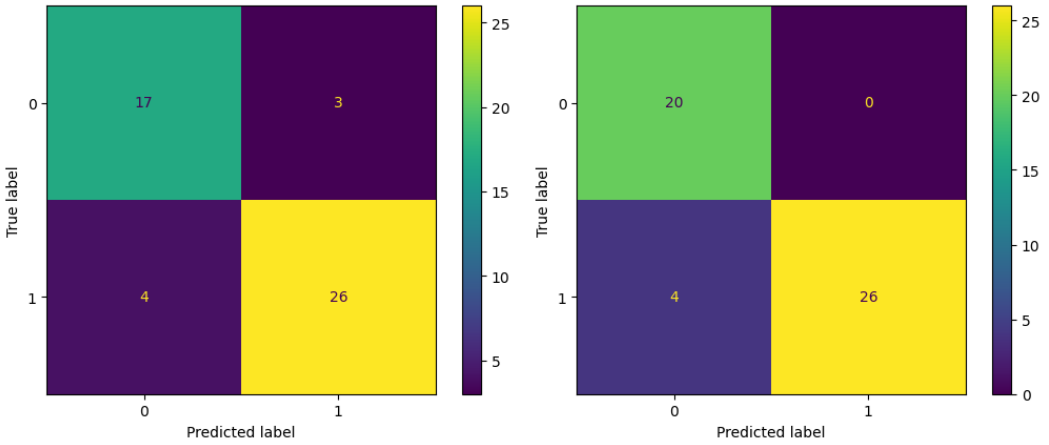


Figure 3.5: Confusion matrices for (left) MobileNet and (right) SimpleCNN architectures. Label “0” is benign; label “1” is malignant.

CHAPTER 4

4. RESULTS AND DISCUSSION

This project aims to develop a deep learning-based approach to automatically detect thyroid cancer using thyroid histopathology imaging. In the first stage, our basic approach is to position AI as the pathologist's assistant. Because there exist 200 patches for one slide and at least 10 slides per patient. The pathologist should examine all of these 2.000 patches in a very short time such as 15 minutes and should never pass a malign case, since the entire treatment procedure will be based on the condition of malignancy.

In **Table 4.1**, performances (accuracy, precision, recall and F1 scores) of various architectures are shown.

As it can be seen on the table, all the architectures have produced satisfactory results. One of the main reasons for this success can be considered as the data collection strategy. Rather than using publicly available datasets or slide scanner outputs, we chose to work with pathologists to create our own dataset. Each image in the dataset was selected by the pathologists as a significant case example. And the source of the images was a microscope giving digital outputs. As a result, the quality of the training set has increased considerably. For further studies these models can be tested against other publicly available datasets and outputs of slide scanners.

Considering the criterion priority, since we aimed to put AI as an assisting mechanism for pathologists at the beginning, it is more important for us to catch all the positives than to decrease false positives. So, for our primary purposes the high recall and accuracy values is more important than the others.

Also, the speed performance of the model may be another important criteria, since the number of images to be examined for one patient is very high. In order to achieve the speed requirements, the number of parameters must not be so high.

Table 4.1: Performance of different architectures on the test set.

	Number of parameters	Accuracy	Precision	Recall	F1
VGG_16	134,268,738	0.94	1.00	0.90	0.95
ResNet_18	11,689,512	0.92	0.96	0.90	0.93
Inception_v3	27,161,264	0.94	0.94	0.97	0.95
MobileNet_v3_small	2,542,856	0.86	0.90	0.87	0.88
EfficientNet_v2_s	21,458,488	0.92	1.00	0.87	0.93
SwinTransformer	28,288,354	0.96	0.96	0.97	0.96
ConvNext_tiny	28,589,128	0.94	0.94	0.97	0.95
SimpleCNN	1,574,658	0.92	1.00	0.87	0.93

Considering the goals of the project, we can say that the first priority is to capture all true positives. Since recall, also known as sensitivity or true positive rate, measures the ability of a model to correctly identify all relevant instances (true positives) in the dataset, recall performance of the models is more significant for this study.

Three of the models, “Inception_v3”, “SwinTransformer” and “ConvNext_tiny”, have 0,97 as the highest recall value. The number of parameters are nearly the same so we can assume their time performance are similar. And these three models can be used as a suitable starting point for further studies.

Finally, in **Figure 3.4** and **Figure 3.5**, we provide regions of interest (heatmaps) that are effective in the decision for benign and non-benign cases, respectively. We provide results for various techniques, including Grad-CAM, Grad-CAM++, XGrad-CAM and Eigen-CAM).

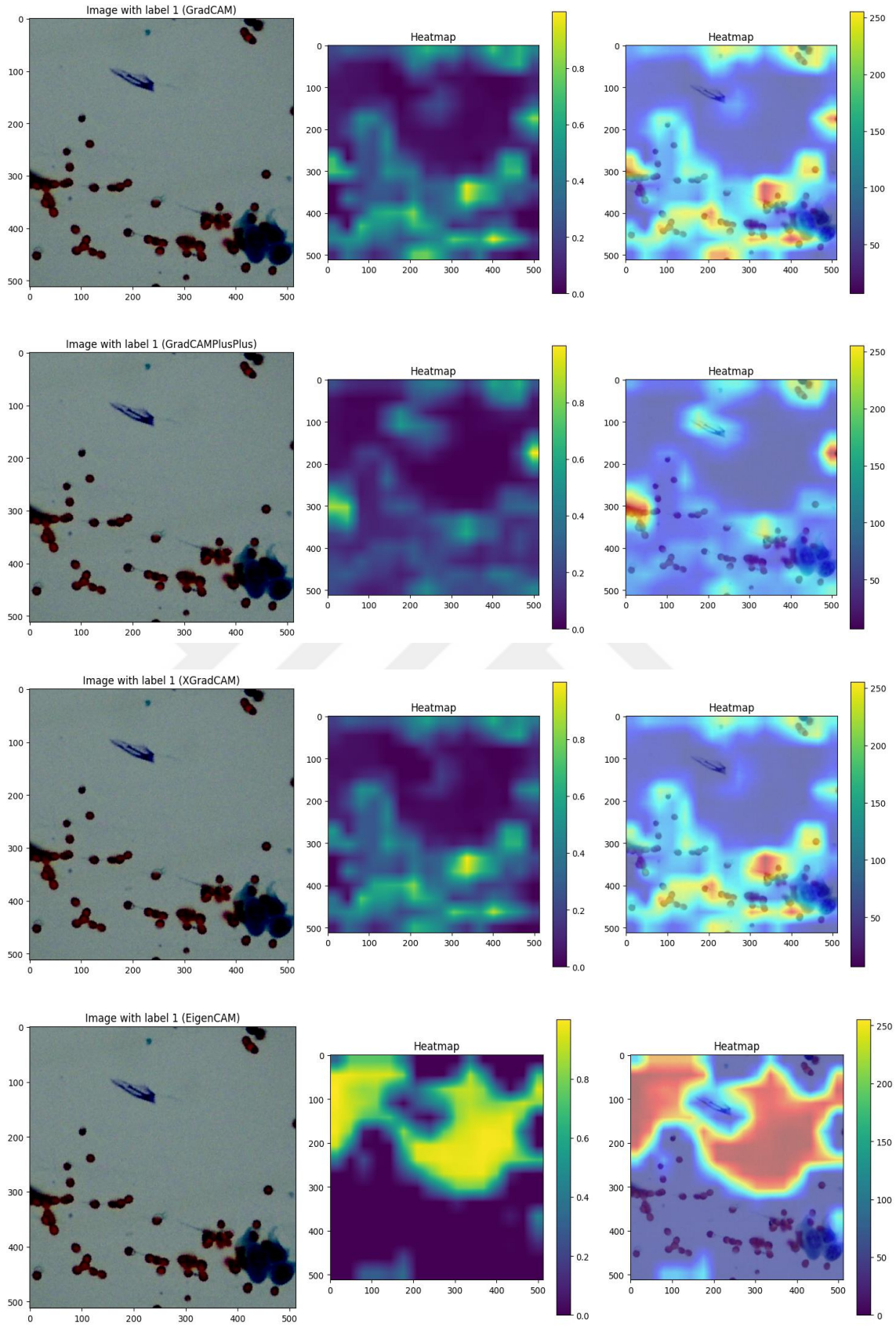


Figure 4.1: Regions of interest (heatmaps) that are effective in the decision for a benign case.

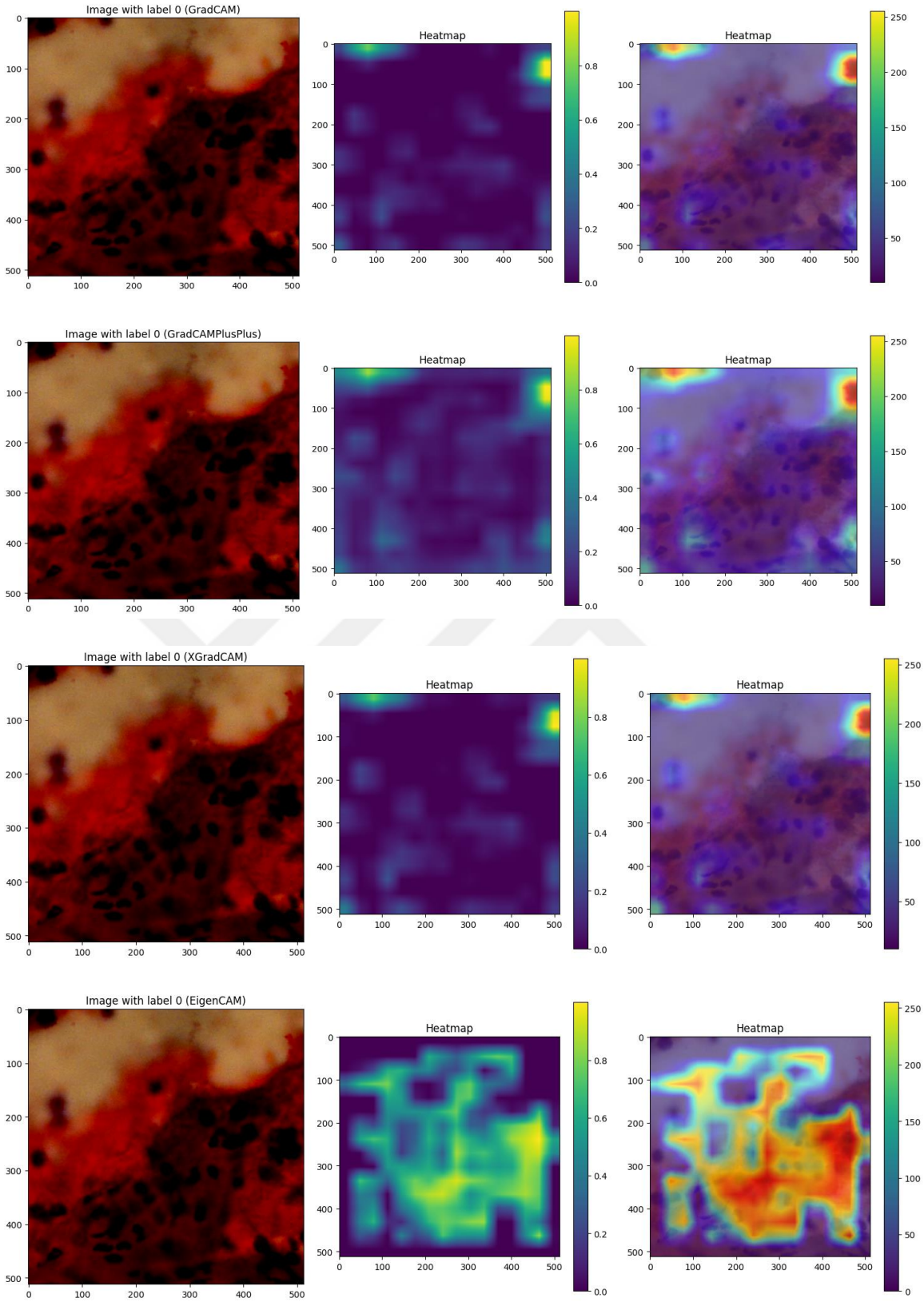


Figure 4.2: Regions of interest (heatmaps) that are effective in the decision for a malignant case.

When we investigate the results, we see that including Grad-CAM, Grad-CAM++ and XGrad-CAM provide comparable results; the heatmaps focus on the nuclei regions. This is convenient with the behavior of the pathologists since they first look at the nuclei regions and sizes.

In clinical applications it is also important to direct the pathologist to the problem area. Since the size of the image is very large, it needs several times to zoom in and out to recognize malignancy in a slide. So, such a practice can be a serious time saving feature.

When we compare the heatmaps for the malign cases, especially Eigen CAM produced a very meaningful result for pathologists. The malign region where the pathologists have to examine is given exactly in the heatmap. This model can be used as a starting point for further studies.

CHAPTER 5

5. CONCLUSION AND FUTURE WORK

In this project, we have classified the three types of thyroid histopathology images based on cancer and non-cancer classes. We have a locally generated microscopic imaging dataset for three classes which are benign thyrocyte (non-cancer), histiocyte (cancer type 1), and lymphocyte (cancer type 2). We proposed a 15 layers convolutional neural network architecture which was originally derived from the darknet 19. The initial images underwent cropping using the sliding window cropping technique and were then inputted into the CNN. Various training methods were experimented with, and the highest accuracy of 99.4% was achieved for the 3-class classification using a random image split approach. Additionally, an accuracy of 97.4% was attained with a random subject split method. Our suggested CNN model offers superior accuracy and is computationally less intricate, making it well-suited for efficient implementation in edge computing.

For future work, a better approach may be a combination of segmentation and classification methods for diagnosis. Although classification methods perform better, it is vulnerable to artifacts caused by different scanners or different pap stains. So, to obtain patterns significant to pathologists, several segmentation methods may be used to detect cell boundaries and vein boundaries etc.. After this process then classification may be used to train the network which will be more accurate. Nevertheless, this type of work needs a larger data set taken from different types of slide scanners in order to prove efficiency.

5.1. Clinical Adoption and Regulatory Challenges

The adoption of deep learning-based histopathology tools in clinical settings is steadily growing, with increasing recognition of the potential benefits they offer for improving diagnostic accuracy, efficiency, and patient outcomes. However, several challenges

remain, particularly concerning regulatory approval and integration into clinical workflows.

5.1.1. Current status of clinical adoption

Research and development: Many deep learning-based histopathology tools are still in the research and development phase, undergoing validation and refinement before clinical deployment. These tools cover various applications, including image classification, segmentation, tumor detection, grading, and prognostication.

Pilot studies and clinical trials: Pilot studies and clinical trials are being conducted to evaluate the performance and clinical utility of deep learning-based histopathology tools. These studies assess factors such as diagnostic accuracy, interobserver variability, impact on workflow efficiency, and patient outcomes.

Integration into clinical practice: Some healthcare institutions have started integrating deep learning-based histopathology tools into their clinical workflows on a limited scale. These integrations often involve collaboration between pathologists, radiologists, data scientists, and software engineers to develop tailored solutions that meet the specific needs and requirements of the healthcare institution.

Challenges and barriers: Despite the potential benefits, the adoption of deep learning-based histopathology tools faces several challenges, including regulatory hurdles, data privacy concerns, technical limitations, workflow integration issues, and the need for extensive validation and clinical validation.

5.1.2. Regulatory challenges and FDA approval

FDA approval process: In the United States, deep learning-based histopathology tools are subject to regulatory oversight by the Food and Drug Administration (FDA). Depending on the intended use and risk classification, these tools may require FDA clearance or approval before clinical use. The FDA approval process typically involves preclinical testing, clinical validation studies, and submission of regulatory documents demonstrating safety, effectiveness, and performance.

Validation requirements: Validating deep learning-based histopathology tools for clinical use involves demonstrating their accuracy, reliability, reproducibility, and clinical utility. Validation studies must adhere to regulatory guidelines and standards, such as

those outlined by the FDA, Clinical Laboratory Improvement Amendments (CLIA), and College of American Pathologists (CAP).

Data requirements: Deep learning algorithms require large, high-quality datasets for training, validation, and testing. Ensuring the availability of annotated histopathology images with ground truth labels can be challenging, particularly for rare diseases or specialized applications.

Interpretability and explainability: Regulatory agencies emphasize the importance of interpretability and explainability in AI-based medical devices. Deep learning models must provide transparent and understandable outputs, allowing clinicians to interpret the results and make informed decisions.

Post-market surveillance: After FDA approval, deep learning-based histopathology tools are subject to post-market surveillance to monitor their safety and performance in real-world clinical settings. Manufacturers are required to report adverse events, monitor user feedback, and conduct periodic reviews to ensure ongoing compliance with regulatory requirements.

In summary, while deep learning-based histopathology tools hold great promise for improving diagnostic accuracy and patient care, their clinical adoption is contingent upon overcoming regulatory challenges, ensuring validation and compliance with regulatory requirements, and addressing technical and workflow integration issues. Collaboration between stakeholders, including clinicians, researchers, regulators, and industry partners, is essential to navigate these challenges and facilitate the responsible and effective integration of AI systems into clinical practice.

5.2. Future Directions and Challenges

In the rapidly evolving field of deep learning-based histopathology analysis, several emerging trends and future research directions are shaping the landscape:

Multi-modal integration: Integrating histopathology data with other modalities, such as genomics, radiology, and clinical data, holds immense potential for improving disease diagnosis, prognosis, and treatment planning. Combining information from multiple sources can provide a more comprehensive understanding of disease mechanisms and heterogeneity, enabling personalized medicine approaches. Future research will focus on

developing deep learning frameworks capable of effectively integrating and analyzing multi-modal data to extract clinically relevant insights.

Federated learning for privacy-preserving analysis: Federated learning is a promising approach for conducting collaborative deep learning across multiple institutions or organizations while preserving data privacy. In histopathology analysis, federated learning enables the aggregation of data from diverse sources without sharing sensitive patient information, thus addressing privacy concerns and regulatory constraints. Future research will explore federated learning techniques tailored for histopathology data and develop robust frameworks for secure and efficient model training across distributed datasets.

Model generalization and robustness: Improving the generalization and robustness of deep learning models in histopathology analysis is a critical research area. Deep learning models trained on large datasets may exhibit overfitting or lack robustness to variations in image quality, staining protocols, or tissue characteristics. Future research will focus on developing techniques for data augmentation, domain adaptation, and model regularization to enhance model generalization and robustness across diverse histopathology datasets. Additionally, research efforts will explore the use of explainable AI techniques to understand model behavior and identify potential sources of bias or errors.

Semi-supervised and self-supervised learning: Semi-supervised and self-supervised learning methods leverage unlabeled data to improve model performance in histopathology analysis. By leveraging abundant unlabeled histopathology images, these techniques can enhance model training and address data scarcity issues. Future research will investigate novel semi-supervised and self-supervised learning approaches tailored for histopathology data, enabling more efficient and scalable model development.

Interpretable and explainable AI: As deep learning models become increasingly complex, there is a growing emphasis on developing interpretable and explainable AI techniques for histopathology analysis. Interpretable models facilitate trust, transparency, and clinical acceptance by providing insights into model predictions and decision-making processes. Future research will focus on developing interpretable deep learning architectures, visualization techniques, and post-hoc explanations tailored for histopathology applications.

Overall, these emerging trends and future research directions hold the potential to advance the field of deep learning-based histopathology analysis, enabling more accurate, efficient, and clinically relevant solutions for disease diagnosis, prognosis, and treatment. Collaboration between researchers, clinicians, industry partners, and regulatory bodies will be essential to drive innovation and translate research findings into clinical practice.

5.3. Further Studies for Türkiye

In the field of pathology, a series of projects can be developed nationwide, in order to provide a much higher level of medical service with the opportunities offered by technology.

5.3.1. An economic slide scanner project

Slide scanners are usually highly expensive devices. This is mostly due to high resolution and precision request of the operation itself. There are several new numerical techniques in digital imaging such as quantitative phase imaging or digital holographic microscopy, which allows to obtain high resolution and precision by using ordinary cameras. Some of these techniques may be used to develop an economic slide scanner. This makes widespread use of slide scanners possible.

The project should also include an application software both for controlling device and analyzing scanned images. It must assist pathologists in examining and quantifying various features of tissue samples. By the introduction of several image processing and AI techniques, preselection of malign areas will be possible. Examining a few images instead of 200 images per slide will be very valuable for pathologists. Furthermore, the possibility of overlook will be decreased.

5.3.2. Digital pathology cloud project

In **section 1.2.11**, it's noted that**,** because WSI images are large and often need to be shared for second opinions, storing them securely on a shared cloud platform would be highly beneficial. This storage must be under the control of Ministry of Health and integrated with E-nabiz (a digital healthcare platform in Türkiye governed by Ministry of Health). The cloud service must be safe and reliable. TUBITAK BILGEM's "Safir Depo" may be a suitable starting point.

The project should also incorporate services and user interfaces needed for uploading, sharing, and examining the images. Just as other E-nabiz data, images must be also accessible from mobile devices in order to ease the examination for second opinion. It should be possible to upload images from different slide scanners and also from microscopes which provides digital images. The device information must be included in the image data for further processing.

Such a project, even without AI advancements, is subject to solve most of the problems of pathology practitioners and provide a more valuable healthcare service for patients.

5.3.3. Digital pathology AI project

Once the Digital Pathology Cloud Project was implemented, it will provide necessary background for the use of artificial intelligence in digital pathology.

Firstly, in order to make data anonymous (i.e. do not contain patient specific data and violate patients' rights) a study for preparation of these datasets must be done under the control of Ministry of Health. With the inclusion of related diagnose and prognose data, many studies in the field of AI, may be done to support diagnose and prognose processes. But at least, at the very beginning, having a continuously improving AI model for the Economic Slide Scanner Project discussed in **section 5.1.1** would be highly beneficial for identifying images containing malignant tissues. Such a model will be updated from the cloud by the application software of developed slide scanner. The model will be sharpened by the new training data as the pathologists upload new images.

Here, several AI and image processing tools may be offered as a service. TUBITAK BILGEM's Artificial Intelligence tool set called "Safir Zeka" may be a suitable starting point for these services.

5.3.4. Digital pathology training and challenge portal

The Digital Pathology Cloud Project and Digital Pathology AI Project have the potential to introduce numerous new services for students and researchers.

By the use of existing images and diagnostic data, some training programs and exams can be prepared for Medical students.

There can be also certain challenges like Camelyon in order to provide motivation on AI researchers and developers.

All of these services will be offered via this portal. Also, with the inclusion of related Universities and Institutions, the results of these studies can be used for further enhancements in diagnose and prognose processes.

5.3.5. Use of multispectral analysis in digital pathology

A novel field of exploration involves the development of an improved scanner capable of conducting multispectral analysis by illuminating slides with varied wavelengths during scanning. This device may help two different research areas.

Firstly, the images at different wavelengths can be used to enhance the resolution and accuracy, by using some numerical techniques.

Secondly, with the introduction of Machine Learning and suitable stochastic analysis techniques, some meaningful relations between malignancy and the illumination wavelength can be found. This will contribute to transforming pathology practices, reducing analysis time for practitioners, and enhancing diagnostic accuracy.

BIBLIOGRAPHY

[1] Dimitriou, Neofytos, O. Arandjelović, and P. D. Caie. "Deep learning for whole slide image analysis: An overview." Frontiers in Medicine 6 (2019).

[2] Komura, Daisuke, and S. Ishikawa. "Machine learning methods for histopathological image analysis." Computational and structural biotechnology journal 16 (2018): 34-42.

[3] Zhou, Xiaomin, et al. "A Comprehensive Review for Breast Histopathology Image Analysis Using Classical and Deep Neural Networks." IEEE Access 8 (2020): 90931-90956.

[4] Tizhoosh, H. Reza, and L. Pantanowitz. "Artificial intelligence and digital pathology: challenges and opportunities." Journal of pathology informatics 9 (2018).

[5] Chain, et al. "Digital image-assisted quantitative nuclear analysis improves diagnostic accuracy of thyroid fine-needle aspiration cytology." Cancer cytopathology 127.8 (2019): 501-513.

[6] F. Bray, et al. Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: A Cancer Journal for Clinicians, 68(6):394–424, 2018.

[7] World Health Organization, (2020). [Onlion]. Available: www.who.int/ cancer/prevention/diagnosis-screening/breast-cancer/en/.

[8] G. Murtaza, et al. Deep learning-based breast cancer classification through medical imaging modalities: state of the art and research challenges. Artificial Intelligence Review, pages 1–66, 2019.

[9] Janowczyk, Andrew, and A. Madabhushi. "Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases." Journal of pathology informatics 7 (2016).

[10] A. Basavanhally, et al. Multi-field-of-view strategy for image-based outcome prediction of multi-parametric estrogen receptor-positive breast cancer histopathology: Comparison to oncotype DX. J Pathol Inform 2011; 2: S1.

[11] C. Genestie, et al. Comparison of the prognostic value of scarff-bloom-richardson and nottingham histological grades in a series of 825 cases of breast cancer: Major importance of the mitotic count as a component of both grading systems. Anticancer Res 1998; 18: 571-6.

[12] P.A. Humphrey. Gleason grading and prognostic factors in carcinoma of the prostate. Mod Pathol 2004; 17: 292-306.

[13] H. Irshad, A. Veillard, L. Roux, D. Racoceanu. Methods for nuclei detection, segmentation, and classification in digital histopathology: A review-current status and future potential. IEEE Rev Biomed Eng 2014; 7: 97-114.

[14] A.C. Ruifrok, D.A. Johnston. Quantification of histochemical staining by color deconvolution. Anal Quant Cytol Histol 2001; 23: 291-9.

[15] A.H. Beck, et al. Systematic analysis of breast cancer morphology uncovers stromal features associated with survival. Sci Transl Med 2011; 3: 108ra113.

[16] A. Basavanhally, et al. Incorporating domain knowledge for tubule detection in breast histopathology using o'callaghan neighborhoods. In: SPIE Medical Imaging. (Computer-Aided Diagnosis). Vol. 7963. SPIE; 2011. p. 796310.

[17] K. Sirinukunwattana, D. Snead, N. Rajpoot. A Random Polygons Model of Glandular Structures in Colon Histology Images. In: Biomedical Imaging (ISBI), 2015 IEEE 12th. International Symposium on; April, 2015. p. 1526-9.

[18] A. Cruz-Roa, et al. Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks. In: SPIE Medical Imaging. Vol. 9041. ;2014. p. 904103-904103-15.

[19] H. Fatakdawala, et al. Expectation-maximization-driven geodesic active contour with overlap resolution (EMaGACOR): Application to lymphocyte segmentation on breast cancer histopathology. IEEE Trans Biomed Eng 2010; 57: 1676-89.

[20] C. Arteta, V. Lempitsky, J.A. Noble, A. Zisserman. Learning to detect cells using non-overlapping extremal regions. In: Ayache N, editor. International Conference on Medical Image Computing and Computer Assisted Intervention. (Lecture Notes in Computer Science). MICCAI, Springer; 2012. p. 348-56.

[21] D.C. Ciresan, A. Giusti, L.M. Gambardella, J. Schmidhuber. Mitosis detection in breast cancer histology images with deep neural networks. Med ImageComput Comput Assist Interv 2013; 16(Pt 2): 411-8.

[22] H. Chang, L. Loss, B. Parvin. Nuclear Segmentation in H and E Sections via Multi-reference Graph-cut (mrgc). International Symposium Biomedical Imaging; 2012.

[23] H. Wang, et al. Mitosis detection in breast cancer pathology images by combining handcrafted and convolutional neural network features. J Med Imaging (Bellingham) 2014; 1: 034003.

[24] N. Orlov, et al. WND-CHARM: Multi-purpose image classification using compound image transforms. Pattern Recognit Lett 2008; 29: 1684-93.

[25] Veta M, et al. Assessment of algorithms for mitosis detection in breast cancer histopathology images. Med Image Anal 2015; 20: 237-48.

[26] L. Roux, et al. Mitosis detection in breast cancer histological images An ICPR 2012 contest. J Pathol Inform 2013; 4: 8.

[27] Arunachalam, H. Babu. Computational Methods for Histopathological Whole Slide Image Analyaaasis of Osteosarcoma. Diss. 2018.

[28] C. Senaras, et al. Creating synthetic digital slides using conditional generative adversarial networks: application to ki67 staining. In Medical Imaging 2018: Digital Pathology, volume 10581, page 1058103. International Society for Optics and Photonics, 2018.

[29] P. Moeskops and M. Veta. Domain-adversarial neural networks to address the appearance variability of histopathology images. In Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: Third

International Workshop, DLMIA 2017, and 7th International Workshop, ML-CDS 2017, Held in Conjunction with MICCAI 2017, Quebec City, QC, Canada, September 14, Proceedings, volume 10553, page 83. Springer, 2017.

[30] O. Sertel, U.V. Catalyurek, H. Shimada, M.N. Gurcan (2009a) Computer-aided prognosis of neuroblastoma: Detection of mitosis and karyorrhexis cells in digitized histological images. In: Annual international conference of the IEEE engineering in medicine and biology society, pp 1433–1436.

[31] H. Kong, M. Gurcan, K. Belkacem-Boussaid (2011a) Partitioning histopathological images: an integrated framework for supervised color-texture segmentation and cell splitting. IEEE Trans Med Imaging 30(9):1661–1677

[32] A. Tabesh, et al. (2007) Multifeature prostate cancer diagnosis and gleason grading of histological images. IEEE Trans Med Imaging 26(10):1366–1378

[33] Y. Wang, et al. (2009) Assisted diagnosis of cervical intraepithelial neoplasia (cin). IEEE J Sel Top Signal Process 3(1): 112–121

[34] R. Awan, et al. (2017). Glandular morphometrics for objective grading of colorectal adenocarcinoma histology images. Scientific reports, 7(1), 16852.

[35] S. Petushi, P. Garcia, M. Haber, and et al. Large-scale Computations on Histology Images Reveal Grade-differentiating Parameters for Breast Cancer. BMC Medical Imaging, 6(14): 1–11, 2006.

[36] A. Osareh and B. Shadgar. Machine learning techniques to diagnose breast cancer. In 2010 5th International Symposium on Health Informatics and Bioinformatics, pages 114–120. IEEE, 2010.

[37] S. Singh, P. Gupta, and M. Sharma. Breast Cancer Detection and Classification of Histopathological Images. International Journal of Engineering Science and Technology, 3(5): 4228–4332, 2011.

[38] Y. Zhang, B. Zhang and W. Lu. Breast Cancer Classification from Histological Images with Multiple Features and Random Subspace Classifier Ensemble. In Proc. of AIP 1371(1), pages 19–28, 2011.

[39] Y. Zhang, B. Zhang, F. Coenen, and W. Lu. Breast Cancer Diagnosis from Biopsy Images with Highly Reliable Random Subspace Classifier Ensembles. Machine Vision and Applications, 24(7): 1405–1420, 2013.

[40] Y. Zhang, B. Zhang, and W. Lu. Breast Cancer Histological Image Classification with Multiple Features and Random Subspace Classifier Ensemble. In T. D. Pham and L. C. Jain, editors, Knowledge-based Systems in Biomedicine, SCI 450, pages 27–42. Springer, Germany, 2013.

[41] C. Loukas, S. Kostopoulos, A. Tanoglidi, and et al. Breast Cancer Characterization based on Image Classification of Tissue Sections Visualized under Low Magnification. Computational and Mathematical Methods in Medicine, 2013: 1–8, 2013.

[42] K. Shukla, A. Tiwari, S. Sharma, and et al. Classification of histopathological images of breast cancerous and noncancerous cells based on morphological features. Biomedical and Pharmacology Journal, 10(1): 353– 366, 2017.

[43] M. Kowal, P. Filipczuk, A. Obuchowicz, and J. Korbicz. Computer-aided Diagnosis of Breast Cancer Based on Fine Needle Biopsy Microscopic Images. Computers in Biology and Medicine, 43(10): 1563–1572, 2013.

[44] A. Mouelhi, M. Sayadi, and F. Fnaiech. A Supervised Segmentation Scheme Based on Multilayer Neural Network and Color Active Contour Model for Breast Cancer Nuclei Detection. In Proc. of ICEESA, pages 1–6, 2013.

[45] IDC: Invasive Ductal Carcinoma, (2014). [Onlion]. Available: http://www.andrewjanowczyk.com/use-case-6-invasive-ductal-carcinoma-idc-segmentation/.

[46] BreakHis: Breast Cancer Histopathological Database BreakHis, (2015). [Onlion]. Available: http://web.inf.ufpr.br/vri/databases/breast-cancer-histopathological-database-breakhis/.

[47] Bioimaging 2015 Breast Histology Classification Challenge, (2015). [Onlion]. Available: https://rdm.inesctec.pt/dataset/nis-2017-003.

[48] TUPAC: The Tumor Proliferation Assessment Challenge 2016, (2016). [Onlion]. Available: http://tupac.tue-image.nl/.

[49] Camelyon 2016: Camelyon Grand Challenge 2016, (2016). [Onlion]. Available: https://camelyon16.grand-challenge.org/Data/.

[50] Camelyon 2017: Camelyon Grand Challenge 2017, (2017). [Onlion]. Available: https://camelyon17.grand-challenge.org/.

[51] BACH: The Grand Challenge on BreAst Cancer Histology images, (2018). [Onlion]. Available: https://iciar2018-challenge. grand-challenge.org/.

[52] F. Spanhol, L. Oliveira, C. Petitjean, and L. Heutte. A Dataset for Breast Cancer Histopathological Image Classification. IEEE Transactions on Biomedical Engineering, 63(7): 1455–1462, 2015.

[53] N. Bayramoglu, J. Kannala, and J. Heikkilä. Deep learning for magnification independent breast cancer histopathology image classification. In Proc. of ICPR, pages 2440–2445. IEEE, 2016.

[54] F. Spanhol, L. Oliveira, C. Petitjean, and L. Heutte. Breast Cancer Histopathological Image Classification Using Convolutional Neural Networks. In Proc. of IJCNN, page Online, 2016.

[55] F. Spanhol, et al. Deep Features for Breast Cancer Histopathological Image Classification. In Proc. of SMC, pages 1868–1873, 2017.

[56] F. Spanhol. Automatic Breast Cancer Classification from Histopathological Images: A Hybrid Approach. PhD Thesis: Federal University of Parana, Brazil, 2018.

[57] Y. Song, J. Zou, H. Chang, and W. Cai. Adapting Fisher Vectors for Histopathology Image Classification. In Proc. of ISBI 2017, pages 600–603, 2017.

[58] W. Zhi, et al. Using Transfer Learning with Convolutional Neural Networks to Diagnose Breast Cancer from Histopathological Images. In Proc. of ICONIP 2017, pages 669–676, 2017.

[59] E. Nejad, L. Affendey, R. Latip, and I. Ishak. Classification of Histopathology Images of Breast into Benign and Malignant using a Single-layer Convolutional Neural Network. In Proc. of ICISPC 2017, pages 50–53, 2017.

[60] Q. Li and W. Li. Using Deep Learning for Breast Cancer Diagnosis. Technical Report: Chinese University of Hong Kong, China, 2017. Z. Han, B. Wei, Y. Zheng, and et al. Breast Cancer Multi-classification from Histopathological Images with Structured Deep Learning Model. Scientific Reports, 7(4172): 1–10, 2017.

[61] Z. Han, et al. Breast Cancer Multi-classification from Histopathological Images with Structured Deep Learning Model. Scientific Reports, 7(4172): 1–10, 2017.

[62] K. Das, et al. Classifying histopathology wholeslides using fusion of decisions from deep convolutional network on a collection of random multi-views at multi-magnification. In Proc. of ISBI 2017, pages 1024–1027. IEEE, 2017.

[63] B. Wei, Z. Han, X. He, and Y. Yin. Deep learning model-based breast cancer histopathological image classification. In Proc. of ICCCBDA, pages 348–353. IEEE, 2017.

[64] N. Motlagh, et al. Breast cancer histopathological image classification: A deep learning approach. bioRxiv, page 242818, 2018.

[65] R. Mehra, et al. Breast cancer histology images classification: Training from scratch or transfer learning? ICT Express, 4(4): 247–254, 2018.

[66] M. Nawaz, A. Sewissy, and T. Soliman. Automated Classification of Breast Cancer Histology Images Using Deep Learning Based Convolutional Neural Networks. International Journal of Computer Science and Network Security, 18(4): 152–160, 2018.

[67] A. Nahid, A. Mikaelian, and Y. Kong. Histopathological Breast-image Classification with Restricted Boltzmann Machine Along with Backpropagation. Biomedical Research, 29(10): 2068–2077, 2018.

[68] A. Nahid, M. Mehrabi, and Y. Kong. Histopathological Breast Cancer Image Classification by Deep Neural Network Techniques Guided by Local Clustering. BioMed Research International, 2018: 1–20, 2018.

[69] A. Nahid and Y. Kong. Histopathological Breast-image Classification Using Local and Frequency Domains by Convolutional Neural Network. Information, 9(19): 1–26, 2018.

[70] B. Du, et al. Breast Cancer Histopathological Image Classification via Deep Active Learning and Confidence Boosting. In Proc. of ICANN 2018, pages 109–116, 2018.

[71] G. Lee, M. Bajger, and K. Clark. Deep Learning and Color Variability in Breast Cancer Histopathological Images: A Preliminary Study. In Proc. of SPIE 10718, page Online, 2018.

[72] M. Nawaz, A. Sewissy, and T. Soliman. Multi-class Breast Cancer Classification using Deep Learning Convolutional Neural Network. International Journal of Advanced Computer Science and Applications, 9(6): 316–332, 2018.

[73] Z. Gandomkar, P. Brennan, and C. Mello-Thoms. A Framework for Distinguishing Benign from Malignant Breast Histopathological Images Using Deep Residual Networks. In Proc. of SPIE 10718, page Online, 2018.

[74] D. Bardou, K. Zhang, and S. Ahmad. Classification of Breast Cancer Based on Histology Images Using Convolutional Neural Networks. IEEE Access, 6:24680–24693, 2018.

[75] K. Das, et al. Multiple instance learning of deep convolutional neural networks for breast histopathology whole slide classification. In Proc. of ISBI 2018, pages 578–581. IEEE, 2018.

[76] R. Mehra, et al. Automatic Magnification Independent Classification of Breast Cancer Tissue in Histological Images Using Deep Convolutional Neural Network. In International Conference on Advanced Informatics for Computing Research, pages 772–781. Springer, 2018.

[77] S. Cascianelli, et al. Dimensionality reduction strategies for cnn-based classification of histopathological images. In International Conference on Intelligent Interactive Multimedia Systems and Services, pages 21–30. Springer, 2018.

[78] B. Xu, et al. Look, Investigate, and Classify: A Deep Hybrid Attention Method for Breast Cancer Classification. arXiv preprint arXiv: 1902.10946, 2019.

[79] M. Bhuiyan, et al. Transfer Learning and Supervised Classifier Based Prediction Model for Breast Cancer. In Big Data Analytics for Intelligent Healthcare Management, pages 59–86. Elsevier, 2019.

[80] J. Xie, R. Liu, J. Luttrell IV, and C. Zhang. Deep learning-based analysis of histopathological images of breast cancer. Frontiers in Genetics, 10, 2019.

[81] Y. Jiang, L. Chen, H. Zhang, and X. Xiao. Breast cancer histopathological image classification using convolutional neural networks with small SEResNet module. PLOS ONE, 14(3), 2019.

[82] J. Hu, L. Shen, and G. Sun. Squeeze-and-excitation networks. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 7132–7141, 2018.

[83] M. Thuy and V. Hoang. Fusing of deep learning, transfer learning and gan for breast cancer histopathological image classification. In International Conference on Computer Science, Applied Mathematics and Applications, pages 255–266. Springer, 2019.

[84] T. Karras, S. Laine, and T. Aila. A style-based generator architecture for generative adversarial networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 4401–4410, 2019.

[85] P. Isola, et al. Image-to-image translation with conditional adversarial networks. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 1125–1134, 2017.

[86] J. Matos, et al. Double transfer learning for breast cancer histopathologic image classification. arXiv preprint arXiv:1904.07834, 2019.

[87] S. Saxena, S. Shukla and M. Gyanchandani. Pre-trained convolutional neural networks as feature extractors for diagnosis of breast cancer using histopathology. International Journal of Imaging Systems and Technology, 2020.

[88] M. Gour, S. Jain and T. Sunil Kumar. Residual learning based cnn for breast cancer histopathological image classification. International Journal of Imaging Systems and Technology, 2020.

[89] L. Breiman. Random forests. Machine Learning, 45(1):5–32, 2001.

[90] Y. Liu, et al. Detecting Cancer Metastases on Gigapixel Pathology Images. arXiv: Camelyon Grand Challenge 2016, 2017.

[91] D. Wang, et al. Deep Learning for Identifying Metastatic Breast Cancer. arXiv: Camelyon Grand Challenge 2016, 2016.

[92] A. BenTaieb and G. Hamarneh. Predicting Cancer with a Recurrent Visual Attention Model for Histopathology Images. In Proc. of MICCAI 2018, pages 129–137, 2018.

[93] H. Lin, H. Chen, Q. Dou, and et al. Scannet: A fast and dense scanning framework for metastastic breast cancer detection from wholeslide image. In Proc. of WACV, pages 539–546. IEEE, 2018.

[94] H. Pang, W. Lin, C. Wang, and C. Zhao. Using Transfer Learning to Detect Breast Cancer without Network Training. In Proc. of CCIS, pages 381–385. IEEE, 2018.

[95] B. Bejnordi, and et al. Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases inWomen With Breast Cancer. JAMA, 318(22): 2199–2210, 2017.

[96] L. Chervony and S. Polak. Fast Classification of Whole Slide Histopathology Images for Breast Cancer Detection. Camelyon Grand Challenge 2017, 2017.

[97] A. Golatkar, D. Anand, and A. Sethi. Classification of Breast Cancer Histology Using Deep Learning. arXiv: Breast Cancer Histology Challenge 2018, 2018.

[98] K. Nazeri, A. Aminpour, and M. Ebrahimi. Two-stage Convolutional Neural Network for Breast Cancer Histology Image Classification. arXiv: Breast Cancer Histology Challenge 2018, 2018.

[99] K. Kiambe. Breast Histopathological Image Feature Extraction with Convolutional Neural Networks for Classification. ICSES Transactions on Image Processing and Pattern Recognition, 4(2): 4–12, 2018.

[100] N. Ranjan, P. Machingal, S. Jammalmadka, and et al. Hierarchical Approach for Breast Cancer Histopathology Images Classification. In Proc. of MIDL 2018, pages 1–7, 2018.

[101] S. Vesal, N. Ravikumar, A. Davari, S. Ellmann, and A. Maier. Classification of Breast Cancer Histology Images Using Transfer Learning. In ter Haar B. Romeny, A. Campilho, F. Karray, editor, International Conference Image Analysis and Recognition, pages 812–819. Springer, Springer: Cham, 2018.

[102] C. Ferreira, et al. Classification of Breast Cancer Histology Images Through Transfer Learning Using a Pre-trained Inception Resnet V2. In ter Haar B.

Romeny, A. Campilho, F. Karray, editor, International Conference Image Analysis and Recognition, pages 763–770. Springer, Springer: Cham, 2018.

[103] Q. Vu, et al. Micro and Macro Breast Histology Image Analysis by Partial Network Re-use. In International Conference Image Analysis and Recognition, pages 895–902. Springer, 2018.

[104] M. Kohl, C. Walz, F. Ludwig, S. Braunewell, and M. Baust. Assessment of breast cancer histology using densely connected convolutional networks. In International Conference Image Analysis and Recognition, pages 903–913. Springer, 2018.

[105] Y. Wang, L. Sun, K. Ma, and J. Fang. Breast cancer microscope image classification based on cnn with image deformation. In International Conference Image Analysis and Recognition, pages 845–852. Springer, 2018.

[106] R. Awan, et al. Context-aware learning using transferable features for classification of breast cancer histology images. In International Conference Image Analysis and Recognition, pages 788–795. Springer, 2018.

[107] H. Cao, S. Bernard, L. Heutte, and R. Sabourin. Improve the performance of transfer learning without fine-tuning using dissimilarity-based multiview learning for breast cancer histology images. In International conference image analysis and recognition, pages 779–787. Springer, 2018.

[108] Y. Vang, Z. Chen, and X. Xie. Deep learning framework for multi-class breast cancer histology image classification. In International Conference Image Analysis and Recognition, pages 914–922. Springer, 2018.

[109] R. Yan, F. Ren, Z. Wang, and et al. Breast cancer histopathological image classification using a hybrid deep neural network. Methods, 2019.

[110] M. Schuster and K. Paliwal. Bidirectional recurrent neural networks. IEEE Transactions on Signal Processing, 45(11): 2673–2681, 1997.

[111] Dataset proposed by Yan, (2019). [Onlion]. Available: http://ear.ict.ac.cn/?page_id=1616.

[112] K. Roy, D. Banik, D. Bhattacharjee, and M. Nasipuri. Patch-based system for Classification of Breast Histology images using deep learning. Computerized Medical Imaging and Graphics, 71:90–103, 2019.

[113] S. Kassani, et al. Breast cancer diagnosis with transfer learning and global pooling. arXiv preprint arXiv: 1909.11839, 2019.

[114] T. Kausar, M. Wang, M. Idrees, and Y. Lu. Hwdcnn. Multi-class recognition in breast histopathology with haar wavelet decomposed image-based convolution neural network. Biocybernetics and Biomedical Engineering, 39(4):967–982, 2019.

[115] L. Roux, et al. Mitosis detection in breast cancer histological images an icpr 2012 contest. Journal of Pathology Informatics, 4, 2013.

[116] C. Malon, et al. Mitotic figure recognition: Agreement among pathologists and computerized detector. Analytical Cellular Pathology, 35(2): 97–100, 2012.

[117] C. Malon and E. Cosatto. Classification of Mitotic Figures with Convolutional Neural Networks and Seeded Blob Features. Journal of Pathology Informatics, 4(8): Online, 2013.

[118] H. Wang, A. Cruz-Roa, A. Basavahally, and et al. Cascaded Ensemble of Convolutional Neural Networks and Handcrafted Features for Mitosis Detection. In Proc. of SPIE 9041, page Online, 2014.

[119] D. Ciresan, A. Giusti, L. Gambardella, and J. Schmidhuber. Mitosis Detection in Breast Cancer Histology Images with Deep Neural Networks. In Proc. of MICCAI 2013, pages 411–418, 2013.

[120] M. Veta. Breast Cancer Histopathology Image Analysis. PhD Thesis in Utrecht University, Netherlands, 2014.

[121] H. Chen, Q. Dou, X. Wang, and et al. Mitosis detection in breast cancer histology images via deep cascaded networks. In Thirtieth AAAI Conference on Artificial Intelligence, 2016.

[122] N. Wahab, A. Khan, and Y. Lee. Transfer learning based deep CNN for segmentation and detection of mitoses in breast cancer histopathological images. Microscopy, 68: 216â˘A¸ S–233, 2019.

[123] C. Li, et al. Weakly supervised mitosis detection in breast histopathology images using concentric loss. Medical Image Analysis, 53: 165â˘A¸ S–178, 2019.

[124] M. Veta, et al. Predicting breast tumor proliferation from whole-slide images: the tupac16 challenge. Medical Image Analysis, 54:111–121, 2019.

[125] T. Araujo, et al. Classification of Breast Cancer Histology Images Using Convolutional Neural Networks. PLOS ONE, 12(6): 1–14, 2017.

[126] A. Mahbod, I. Ellinger, R. Ecker, and et al. Breast Cancer Histological Image Classification Using Fine-tuned Deep Network Fusion. In Proc. of ICIAR 2018, pages 754–762, 2018.

[127] A. Rakhlin, A. Shvets, V. Iglovikov, and A. Kalinin. Deep Convolutional Neural Networks for Breast Cancer Histology Image Analysis. In Proc. of ICIAR 2018, pages 737–744, 2018.

[128] Y. Li, J. Wu, and Q. Wu. Classification of Breast Cancer Histology Images Using Multi-Size and Discriminative Patches Based on Deep Learning. IEEE Acess, 7: 21400–21408, 2019.

[129] M. Alom, et al. Breast Cancer Classification from Histopathological Images with Inception Recurrent Residual Convolutional Neural Network. Journal of Digital Imaging, pages 1–13, 2019.

[130] H. Ahmad, S. Ghuffar, and K. Khurshid. Classification of Breast Cancer Histology Images Using Transfer Learning. In Proc. of IBCAST, pages 328–332, 2019.

[131] A. Cruz-Roa, A. Basavanhally, F. González, and et al. Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks. In Medical Imaging 2014: Digital Pathology, volume 9041, page 904103. International Society for Optics and Photonics, 2014.

[132] F. Romero, A. Tang, and S. Kadoury. Multi-Level Batch Normalization In Deep Networks For Invasive Ductal Carcinoma Cell Discrimination In Histopathology Images. arXiv preprint arXiv:1901.03684, 2019.

[133] J. Wu, al. Histopathological Image Classification Using Random Binary Hashing based PCANet and Bilinear Classifier. In Proc. of EUSIPCO, pages 2050–2054, 2016.

[134] B. Bejnordi, et al. Deep learning-based assessment of tumor-associated stroma for diagnosing breast cancer in histopathology images. In Proc. of ISBI 2017, pages 929–932. IEEE, 2017.

[135] Z. Wang, et al. Classification of Breast Cancer Histopathological Images Using Convolutional Neural Networks with Hierarchical Loss and Global Pooling. In Proc. of ICIAR 2018, pages 745–753, 2018.

[136] B. Gecer, et al. Detection and Classification of Cancer in Whole Slide Breast Histopathology Images Using Deep Convolutional Networks. Pattern Recognition, 84: 345–356, 2018.

[137] H. Couture, et al. Image analysis with deep learning to predict breast cancer grade, ER status, histologic subtype and intrinsic subtype. NPJ Breast Cancer, 4(1): 30, 2018.

[138] S. Khan, et al. A Novel Deep Learning based Framework for the Detection and Classification of Breast Cancer Using Transfer Learning. Pattern Recognition Letters, 2019.

[139] T. Qaiser, et al. Her 2-challenge contest: a detailed assessment of automated her 2 scoring algorithms in whole slide images of breast cancer tissues. Histopathology, 72(2): 227–238, 2018.

[140] B. Pang, Y. Zhang, Q. Chen, and et al. Cell Nucleus Segmentation in Color Histopathological Imagery Using Convolutional Networks. In Proc. of CCPR, pages 1–5, 2010.

[141] H. Su, et al. Region segmentation in histopathological breast cancer images using deep convolutional neural network. In Proc. of ISBI, pages 55–58. IEEE, 2015.

[142] J. Xu, et al. A deep convolutional neural network for segmenting and classifying epithelial and stromal regions in histopathological images. Neurocomputing, 191: 214–223, 2016.

[143] F. Xing, Y. Xie and L. Yang. An automatic learning-based framework for robust nucleus segmentation. IEEE Transactions on Medical Imaging, 35(2): 550–566, 2016.

[144] X. Pan, et al. Accurate segmentation of nuclei in pathological images via sparse reconstruction and deep convolutional networks. Neurocomputing, 229: 88–99, 2017.

[145] P. Naylor, M. Laé, F. Reyal and T. Walter. Nuclei segmentation in histopathology images using deep neural networks. In Proc. of ISBI 2017, pages 933–936. IEEE, 2017.

[146] P. Naylor. Dataset, (2017). Available: http://cbio.mines-paristech.fr/~pnaylor/BNS.zip.

[147] Y. Cui, G. Zhang, Z. Liu, Z. Xiong, and J. Hu. A Deep Learning Algorithm for One-step Contour Aware Nuclei Segmentation of Histopathological Images. arXiv preprint arXiv: 1803.02786, 2018.

[148] N. Kumar, et al. A dataset and a technique for generalized nuclear segmentation for computational pathology. IEEE Transactions on Medical Imaging, 36(7): 1550–1560, 2017.

[149] S. Mejbri, et al. Deep Analysis of CNN Settings for New Cancer whole-slide Histological Images Segmentation: The Case of Small Training Sets. In 6th International conference on BioImaging (BIOIMAGING 2019), pages 120–128, 2019.

[150] B. Priego-Torres, D. Sanchez-Morillo, M. Fernandez-Granero and M. Garcia-Rojo. Automatic segmentation of whole-slide h&e stained breast histopathology images using a deep convolutional neural network architecture. Expert Systems with Applications, page 113387, 2020.

[151] J. Xu, et al. Stacked Sparse Autoencoder (SSAE) for Nuclei Detection on Breast Cancer Histopathology Images. IEEE Transactions on Medical Imaging, 35(1): 119–130, 2016.

[152] G. Litjens, et al. Deep Learning as a Tool for Increased Accuracy and Efficiency of Histopathology Diagnosis. Scientific Reports, 6(26286): 1–11, 2016.

[153] A. Cruz-Roa, H. Gilmore, A. Basavanhally, and et al. Highthroughput adaptive sampling for whole-slide histopathology image analysis (HASHI) via convolutional neuralnetworks: Application to invasive breast cancer detection. PLOS ONE, 13(5), 2018.

[154] M. Saha, C. Chakraborty, and D. Racoceanu. Efficient deep learning model for mitosis detection using breast histopathology images. Computerized Medical Imaging and Graphics, 64: 29–40, 2018.

[155] Z. Zainudin, S. Shamsuddin, and S. Hasan. Deep Layer CNN Architecture for Breast Cancer Histopathology Image Detection. In International Conference on Advanced Machine Learning Technologies and Applications, pages 43–51. Springer, 2019.

[156] K. Beevi, M. Nair, and G. Bindu. Automatic mitosis detection in breast histopathology images using convolutional neural network based deep transfer learning. Biocybernetics and Biomedical Engineering, 39(1): 214–223, 2019.

[157] M. Gardner and S. Dorling. Artificial neural networks (the multilayer perceptron) A review of applications in the atmospheric sciences. Atmospheric Environment, 32(14-15): 2627–2636, 1998.

[158] R. Ribani and M. Marengoni. A survey of transfer learning for convolutional neural networks. In Proc. of SIBGRAPI-T, pages 47–57. IEEE, 2019.

[159] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv: 1409.1556, 2014.

[160] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 770–778, 2016.

[161] O. Hadad, et al. Classification of breast lesions using cross-modal deep learning. In ISBI 2017, pages 109–112.IEEE, 2017.

[162] O. Russakovsky, et al. Imagenet large scale visual recognition challenge. International Journal of Computer Vision, 115(3): 211–252, 2015.

[163] D. Sarkar, R. Bali, and T. Ghosh. Hands-On Transfer Learning with Python: Implement advanced deep learning and neural network models using TensorFlow and Keras. Packt Publishing Ltd, 2018.

[164] M. Alom, et al. Improved inception-residual convolutional neural network for object recognition. Neural Computing and Applications, pages 1–15, 2018.

[165] M. Alom, M. Hasan, C. Yakopcic, and T. Taha. Inception recurrent convolutional neural network for object recognition. arXiv preprint arXiv: 1704.07709, 2017.

[166] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi. Inception-v4, inception-resnet and the impact of residual connections on learning. In Thirty-first AAAI conference on artificial intelligence, 2017.

[167] M. Liang and X. Hu. Recurrent convolutional neural network for object recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 3367–3375, 2015.

[168] Y. Chen, et al. Deep transfer learning for histopathological diagnosis of cervical cancer using convolutional neural networks with visualization schemes. Journal of Medical Imaging and Health Informatics, 10(2): 391–400, 2020.

[169] S. Sornapudi, et al. Deep learning nuclei detection in digitized histology images by superpixels. Journal of Pathology Informatics, 9, 2018.

[170] F. Sheikhzadeh, R. Ward, D. van Niekerk, and M. Guillaud. Automatic labeling of molecular biomarkers of immunohistochemistry images using fully convolutional networks. PLOS ONE, 13(1), 2018.

[171] M. Wu, et al. Automatic classification of cervical cancer from cytological images by using convolutional neural network. Bioscience Reports, 38(6), 2018.

[172] S. Gautam, et al. Considerations for a pap smear image analysis system with cnn features. arXiv preprint arXiv: 1806.09025, 2018.

[173] Y. Song, et al. Accurate cervical cell segmentation from overlapping clumps in pap smear images. IEEE Transactions on Medical Imaging, 36(1): 288–300, 2016.

[174] S. Padi, et al. Comparison of artificial intelligence-based approaches to cell function prediction. Informatics in Medicine Unlocked, 18: 100270, 2020.

[175] D. Kusumoto, et al. Automated deep learning-based system to identify endothelial cells derived from induced pluripotent stem cells. Stem Cell Reports, 10(6): 1687–1695, 2018.

[176] E. Ito, T. Sato, D. Sano, E. Utagawa, and T. Kato. Virus particle detection by convolutional neural network in transmission electron microscopy images. Food and Environmental Virology, 10(2): 201–208, 2018.

[177] D. Matuszewski and I. Sintorn. Minimal annotation training for segmentation of microscopy images. In ISBI 2018, pages 387–390. IEEE, 2018.

[178] S. Kosov, K. Shirahama, C. Li, and M. Grzegorzek. Environmental microorganism classification using conditional random fields and deep convolutional neural networks. Pattern Recognition, 77: 248–261, 2018.

[179] S. Javadi and S. Mirroshandel. A novel deep learning method for automatic assessment of human sperm images. Computers in Biology and Medicine, 109: 182–194, 2019.

[180] J. Riordon, C. McCallum, and D. Sinton. Deep learning for the classification of human sperm. Computers in Biology and Medicine, 111: 103342, 2019.

[181] A. Houston, et al. Medical data mining on the internet: Research on a cancer information system. Artificial Intelligence Review, 13(5-6): 437–466, 1999.

[182] T. Markiewicz, et al. Miap–web-based platform for the computer analysis of microscopic images to support the pathological diagnosis. Biocybernetics and Biomedical Engineering, 36(4): 597–609, 2016.

[183] N. Alqahtani, et al. Deep learning convolutional neural networks to predict porous media properties. In SPE Asia Pacific oil and gas conference and exhibition. Society of Petroleum Engineers, 2018.

[184] S. Karimpouli and P. Tahmasebi. Segmentation of digital rock images using deep convolutional autoencoder networks. Computers & Geosciences, 126: 142–150, 2019.

[185] F. Dong, et al. (2014) Computational pathology to discriminate benign from malignant intraductal proliferations of the breast. PLoS One 9(12): e114885. doi:10.1371/journal.pone.0114885

[186] A. Tabesh, et al. (2007) Multifeature prostate cancer diagnosis and gleason grading of histological images. IEEE Trans Med Imaging 26(10): 1366–1378

[187] Y. Song, et al. (2015), Accurate segmentation of cervical cytoplasm and nuclei based on multi-scale convolutional network and graph partitioning. IEEE transactions on biomedical engineering PP(99): 1–1

[188] A. Chekkoury, et al. (2012) Automated malignancy detection in breast histopathological images. In: Medical Imaging 2012: Computer-Aided Diagnosis, San Diego, California, vol 8315:831505

[189] E. Reinhard, M. Adhikhmin, B. Gooch, P. Shirley, (2001), Color transfer between images. IEEE Comput Gr Appl 21(5): 34–41

[190] S. Kothari, et al. (2011), Automatic batch-invariant color segmentation of histological cancer images. In: IEEE International symposium on biomedical imaging: from nano to macro, IEEE, pp 657–660

[191] A.M. Khan, N. Rajpoot, D. Treanor, D. Magee. (2014) Non-linear mapping approach to stain normalisation in digital histopathology images using image-

specific colour deconvolutionieee transactions on biomedical engineering. IEEE Trans Biomed Eng 61(6): 1729–1738

[192] X. Li, K.N. Plataniotis, (2015), A complete color normalization approach to histopathology images using color cues computed from saturation-weighted statistics. IEEE Trans Biomed Eng 62(7): 1862–1873

[193] Fagman, Henrik, and M. Nilsson. "Morphogenesis of the thyroid gland." Molecular and cellular endocrinology 323.1 (2010): 35-54.

[194] Nixon, J. Iain, et al. "Metastasis to the thyroid gland: a critical review." Annals of surgical oncology 24.6 (2017): 1533-1539.

[195] Dralle, Henning, et al. "Follicular cell-derived thyroid cancer." Nature reviews Disease primers 1.1 (2015): 1-18.

[196] Roger, P. Pierre, et al. "Signal transduction in the human thyrocyte and its perversion in thyroid tumors." Molecular and cellular endocrinology 321.1 (2010): 3-19.

[197] Tousseyn, Thomas, and De Wolf-Peeters. "T cell/histiocyte-rich large B-cell lymphoma: an update on its biology and classification." Virchows Archiv 459.6 (2011): 557-563.

[198] Jara, L. Evelyn, et al. "Modulating the function of the immune system by thyroid hormones and thyrotropin." Immunology letters 184 (2017): 76-83.

[199] Xing, Mingzhao. "Molecular pathogenesis and mechanisms of thyroid cancer." Nature Reviews Cancer 13.3 (2013): 184-199.

[200] J.P. Cohen, P. Morrison, and L.J.U.h.g.c.i.c.-c.-d. Dao, Covid-19 image data collection. arXiv 2003.11597, 2020. 2020.

[201] Y. LeCun, Y. Bengio, and G.J.n. Hinton, Deep learning. 2015. 521(7553): p. 436-444

[202] J. Deng, et al. Imagenet: A large-scale hierarchical image database. in 2009 IEEE conference on computer vision and pattern recognition. 2009. IEEE.

[203] D. Shen, G. Wu, and H.-I.J.A.r.o.b.e. Suk, Deep learning in medical image analysis. 2017. 19: p. 221-248.

[204] P. Semberecki, and H. Maciejewski. Deep learning methods for subject text classification of articles. in 2017 Federated Conference on Computer Science and Information Systems (FedCSIS). 2017. IEEE.

[205] M. Tschannen, et al. Heart sound classification using deep structured features. in 2016 Computing in Cardiology Conference (CinC). 2016. IEEE.

[206] F.A. Jaffer, and R.J.J. Weissleder. Molecular imaging in the clinical arena. 2005. 293(7): p. 855-862.

[207] B. Rim, et al. Deep Learning in Physiological Signal Data: A Survey. 2020. 20(4): p. 969.

[208] A. Rehman, F.G.J.M.T. Khan, and Applications, A deep learning based review on abdominal images. 2020: p. 1-32.

[209] L. Li, et al. Multi-task deep learning for fine-grained classification and grading in breast cancer histopathological images. 2020. 79(21): p. 14509-14528.

[210] S. Hamidian, et al. 3D convolutional neural network for automatic detection of lung nodules in chest CT. in Medical Imaging 2017: Computer-Aided Diagnosis. 2017. International Society for Optics and Photonics.

[211] S. Iqbal, et al. Brain tumor segmentation in multi-spectral MRI using convolutional neural networks (CNN). 2018. 81(4): p. 419-427.

[212] Redmon, Joseph, and A. Farhadi. "YOLO9000: better, faster, stronger." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.

[213] A.A. Ahmed, M. Abouzid, & E. Kaczmarek. (2022), Deep Learning Approaches in Histopathology. Cancers 2022, 14, 5264.

[214] L. Aprupe, G. Litjens, T.J. Brinker, J. van der Laak, & N. Grabe (2019), Robust and accurate quantification of biomarkers of immune cells in lung cancer micro-environment using deep convolutional neural networks. PeerJ, 7, e6335.

[215] T. Araujo, et al. (2017), Classification of Breast Cancer Histology Images Using Convolutional Neural Networks. PLoS ONE, 12, e0177544.

[216] E. Arvaniti, et al. (2018), Automated Gleason grading of prostate cancer tissue microarrays via deep learning. Scientific reports, 8(1), 12054.

[217] S. Bach, et al. (2015), On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation. PloS ONE, 10(7), e0130140.

[218] B.E. Bejnordi, et al. (2018), Using Deep Convolutional Neural Networks to Identify and Classify Tumor-Associated Stroma in Diagnostic Breast Biopsies. Mod. Pathol. 31, 1502–1512.

[219] Bejnordi, et al. & CAMELYON16 Consortium, (2017), Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. Jama, 318(22), 2199-2210.

[220] B.E. Bejnordi, et al. Context-Aware Stacked Convolutional Neural Networks for Classification of Breast Carcinomas in Whole-Slide Histopathology

[221] Images. J. Med. Imaging 2017, 4, 044504.

[222] K. Boyd, K. Eng, (2020), Evaluation metrics for classification and regression models. Radcliffe Department of Medicine, University of Oxford.

[223] A. Chattopadhay, A. Sarkar, P. Howlader, & V.N. Balasubramanian (2018, March), Grad-cam++: Generalized gradient-based visual explanations for deep convolutional networks. In 2018 IEEE winter conference on applications of computer vision (WACV) (pp. 839-847). IEEE.

[224] L.C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, A.L. Yuille, (2018), DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. IEEE transactions on pattern analysis and machine intelligence, 40(4), 834-848.

[225] L.C. Chen, Y. Zhu, G. Papandreou, F. Schroff, H. Adam, (2018), Encoder-decoder with atrous separable convolution for semantic image segmentation. In Proceedings of the European Conference on Computer Vision (ECCV) (pp. 801-818).

[226] N. Coudray, et al. (2018), Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning. Nature medicine, 24(10), 1559-1567.

[227] A. Cruz-Roa, et al. (2017), Accurate and reproducible invasive breast cancer detection in whole-slide images: A Deep Learning approach for quantifying tumor extent. Scientific reports, 7(1), 46450.

[228] A. Cruz-Roa, et al. (2018), High-throughput adaptive sampling for whole-slide histopathology image analysis (HASHI) via convolutional neural networks: Application to invasive breast cancer detection. PloS one, 13(5), e0196828.

[229] E.D. Cubuk, B. Zoph, D. Mane, V. Vasudevan, Q.V. Le, (2019), AutoAugment: Learning augmentation policies from data. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 113-123).

[230] A. Dosovitskiy, et al. (2020), An image is worth 16x16 words: Transformers for image recognition at scale. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 3156-3164).

[231] J. Duchi, E. Hazan, Y. Singer, (2011), Adaptive subgradient methods for online learning and stochastic optimization. Journal of machine learning research, 12(7).

[232] M. Everingham, L. Van Gool, C.K. Williams, J. Winn, A. Zisserman, (2010), The pascal visual object classes (VOC) challenge. International journal of computer vision, 88(2), 303-338.

[233] R. Fuet, et al. (2020), Axiom-based grad-cam: Towards accurate visualization and explanation of cnns. arXiv preprint arXiv: 2008.02312.

[234] A. García-García, et al. (2017), A review on deep learning techniques applied to semantic segmentation. arXiv preprint arXiv: 1704.06857.

[235] A. Gertych, et al, (2019), Convolutional neural networks can accurately distinguish four histologic growth patterns of lung adenocarcinoma in digital slides. Scientific reports, 9(1), 1483.

[236] I. Goodfellow, Y. Bengio, A. Courville, (2016), Deep learning. MIT press.

[237] Q. Guan, et al. (2019), Deep convolutional neural network VGG-16 model for differential diagnosing of papillary thyroid carcinomas in cytological images: a pilot study. Journal of Cancer, 10(20), 4876.

[238] K. He, G. Gkioxari, P. Dollár, R. Girshick, (2017). Mask R-CNN. In Proceedings of the IEEE international conference on computer vision (pp. 2961-2969).

[239] K. He, X. Zhang, S. Ren, J. Sun, (2016), Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

[240] G. Hinton, N. Srivastava, K. Swersky, (2020), Lecture 6a: Overview of Mini-Batch Gradient Descent.

(https://www.cs.toronto.edu/~tijmen/csc321/slides/lecture_slides_lec6.pdf) (Last accessed: March 28, 2024)

[241] G. Hinton, O. Vinyals, J. Dean, (2015), Distilling the knowledge in a neural network. arXiv preprint arXiv:1503.02531.

[242] A.G. Howard, et al. (2017), MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. arXiv preprint arXiv:1704.04861.

[243] G. Huang, Z. Liu, L. Van Der Maaten, K.Q. Weinberger, (2017), Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4700-4708).

[244] Y. Jiang, L. Chen, H. Zhang, X. Xiao, (2019), Breast cancer histopathological image classification using convolutional neural networks with small SE-ResNet module. PloS one, 14(3), e0214587.

[245] P. Kainz, M. Pfeiffer, M. Urschler, (2017), Segmentation and classification of colon glands with deep convolutional neural networks and total variation regularization. PeerJ, 5, e3874.

[246] J. N. Kather, et al. (2019), Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. Nature medicine, 25(7), 1054-1056.

[247] D.P. Kingma, J. Ba, (2014), Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

[248] B. Korbar, et al. (2017), Deep learning for classification of colorectal polyps on whole-slide images. Journal of pathology informatics, 8(1), 30.

[249] A. Krizhevsky, I. Sutskever, G.E. Hinton, (2012), ImageNet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105).

[250] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, (1998), Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278-2324.

[251] T.Y. Lin, P. Goyal, R. Girshick, K. He, P. Dollár, (2017), Focal loss for dense object detection. In Proceedings of the IEEE international conference on computer vision (pp. 2980-2988).

[252] T.Y. Lin, et al. (2014), Microsoft coco: Common objects in context. In European conference on computer vision (pp. 740-755).

[253] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, (2016), SSD: Single shot multibox detector. In European conference on computer vision (pp. 21-37).

[254] Y. Liu, et al. (2019). Artificial intelligence–based breast cancer nodal metastasis detection: insights into the black box for pathologists. Archives of pathology & laboratory medicine, 143(7), 859-868.

[255] Liu, Z., et al. (2021). Swin transformer: Hierarchical vision transformer using shifted windows. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 13194-13203).

[256] A. Maleki, M. Raahemi, H. Nasiri, (2023). Breast cancer diagnosis from histopathology images using deep neural network and XGBoost. Biomedical Signal Processing and Control, 86, 105152.

[257] D. Mandair, J.S. Reis-Filho, A. Ashworth, (2023), Biological insights and novel biomarker discovery through deep learning approaches in breast cancer histopathology. NPJ breast cancer, 9(1), 21.

[258] C. Mercan, et al. (2018) Multi-Instance Multi-Label Learning for Multi-Class Classification of Whole Slide Breast Histopathology Images. IEEE Trans. Med. Imaging 2018, 37, 316–325.

[259] P. Mobadersany, et al. (2018), Predicting cancer outcomes from histology and genomics using convolutional networks. Proceedings of the National Academy of Sciences, 115(13), E2970-E2979.

[260] R.K. Mondol, et al. (2023), hist2rna: an efficient deep learning architecture to predict gene expression from breast cancer histopathology images. Cancers, 15(9), 2569.

[261] M.B. Muhammad, M. Yeasin, (2020, July), Eigen-cam: Class activation map using principal components. In 2020 international joint conference on neural networks (IJCNN) (pp. 1-7). IEEE.

[262] H.G. Ramaswamy, (2020), Ablation-cam: Visual explanations for deep convolutional network via gradient-free localization. In proceedings of the IEEE/CVF winter conference on applications of computer vision (pp. 983-991).

[263] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, (2016), You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788).

[264] J. Redmon, A. Farhadi, (2017). YOLO9000: better, faster, stronger. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 7263-7271).

[265] S. Ren, K. He, R. Girshick, J. Sun, (2015), Faster R-CNN: Towards real-time object detection with region proposal networks. In Advances in neural information processing systems (pp. 91-99).

[266] O. Ronneberger, P. Fischer, T. Brox, (2015). U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention (pp. 234-241).

[267] O. Ronneberger, P. Fischer, T. Brox, (2015). U-net: Convolutional networks for biomedical image segmentation. In Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18 (pp. 234-241). Springer International Publishing.

[268] M. Saha, C. Chakraborty, I. Arun, R. Ahmed, S. Chatterjee, (2017), An advanced deep learning approach for Ki-67 stained hotspot detection and proliferation rate scoring for prognostic evaluation of breast cancer. Scientific reports, 7(1), 3213.

[269] A.J. Schaumberg, M.A. Rubin, T.J. Fuchs, (2016), H&E-stained whole slide image deep learning predicts SPOP mutation state in prostate cancer. BioRxiv, 064279.

[270] R.R. Selvaraju, et al. (2017), Grad-CAM: Visual explanations from deep networks via gradient-based localization. In Proceedings of the IEEE International Conference on Computer Vision (pp. 618-626).

[271] L. Sha, et al. (2019). Multi-field-of-view deep learning model predicts nonsmall cell lung cancer programmed death-ligand 1 status from whole-slide hematoxylin and eosin images. Journal of pathology informatics, 10(1), 24.

[272] H. Sharma, N. Zerbe, I. Klempert, O. Hellwich, P. Hufnagl, (2017). Deep convolutional neural networks for automatic classification of gastric carcinoma using whole slide images in digital histopathology. Computerized Medical Imaging and Graphics, 61, 2-13.

[273] C. Shorten, T.M. Khoshgoftaar, (2019), A survey on image data augmentation for deep learning. Journal of Big Data, 6(1), 1-48.

[274] K. Simonyan, A. Vedaldi, A. Zisserman, (2013), Deep inside convolutional networks: Visualising image classification models and saliency maps. arXiv preprint arXiv: 1312.6034.

[275] K. Simonyan, A. Zisserman, (2014), Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

[276] M. Sokolova, G. Lapalme, (2009), A systematic analysis of performance measures for classification tasks. Information Processing & Management, 45(4), 427-437.

[277] M. Springenberg, et al. (2023), From modern CNNs to vision transformers: Assessing the performance, robustness, and classification strategies of deep learning models in histopathology. Medical Image Analysis, 87, 102809.

[278] D.F. Steiner, et al. (2018), Impact of deep learning assistance on the histopathologic review of lymph nodes for metastatic breast cancer. The American journal of surgical pathology, 42(12), 1636-1646.

[279] C. Szegedy, et al. (2015), Going deeper with convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1-9).

[280] M. Tan, B. Chen, R. Pang, V. Vasudevan, Q.V. Le, (2020), EfficientDet: Scalable and efficient object detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 10781-10790).

[281] M. Tan, Q.V. Le, (2019), EfficientNet: Rethinking model scaling for convolutional neural networks. In Proceedings of the 36th International Conference on Machine Learning (Vol. 97, pp. 6105-6114).

[282] A. Teramoto, T. Tsukamoto, Y. Kiriyama, H. Fujita, (2017), Automated classification of lung cancer types from cytological images using deep convolutional neural networks. BioMed research international, 2017.

[283] T. Tieleman, (2012), Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude. COURSERA: Neural networks for machine learning, 4(2), 26.

[284] R. Turkki, N. Linder, P.E. Kovanen, T. Pellinen, J. Lundin, (2016), Antibody-supervised deep learning for quantification of tumor-infiltrating immune cells in hematoxylin and eosin stained breast cancer samples. Journal of pathology informatics, 7(1), 38.

[285] M.E. Vandenberghe, et al. (2017), Relevance of deep learning to facilitate the diagnosis of HER2 status in breast cancer. Scientific reports, 7(1), 45938.

[286] M. Veta, et al. (2019), Predicting breast tumor proliferation from whole-slide images: the TUPAC16 challenge. Medical image analysis, 54, 111-121.

[287] M. Veta, et al. (2015). Assessment of algorithms for mitosis detection in breast cancer histopathology images. Medical image analysis, 20(1), 237-248.

[288] S. Wang, et al. (2018). Comprehensive analysis of lung cancer pathology images to discover tumor shape and boundary features that predict survival outcome. Scientific reports, 8(1), 10393.

[289] Y. Wang, et al. (2019), Using deep convolutional neural networks for multi-classification of thyroid tumor by histopathology: a large-scale pilot study. Annals of translational medicine, 7(18).

[290] S. Wang, et al. (2019), RMDL: Recalibrated multi-instance deep learning for whole slide gastric image classification. Medical image analysis, 58, 101549.

[291] T. Wan, J. Cao, J. Chen, Z. Qin, (2017), Automated grading of breast cancer histopathology using cascaded ensemble with combination of multi-level image features. Neurocomputing, 229, 34-44.

[292] J.W. Wei, et al. (2019), Pathologist-level classification of histologic patterns on resected lung adenocarcinoma slides with deep neural networks. Scientific reports, 9(1), 3358.

[293] C.A. Weis, et al. (2018). Automatic evaluation of tumor budding in immunohistochemically stained colorectal carcinomas and correlation to clinical outcome. Diagnostic pathology, 13, 1-12.

[294] M. Wu, C. Yan, H. Liu, Q. Liu, Y. Yin, (2018), Automatic classification of cervical cancer from cytological images by using convolutional neural network. Bioscience reports, 38(6), BSR20181769.

[295] L. Zhang, et al. (2017), DeepPap: deep convolutional networks for cervical cell classification. IEEE journal of biomedical and health informatics, 21(6), 1633-1643.

[296] Y. Zhang, et al. (2024), Histopathology images-based deep learning prediction of prognosis and therapeutic response in small cell lung cancer. NPJ digital medicine, 7(1), 15.

[297] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, A. Torralba, (2016). Learning deep features for discriminative localization. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 2921-2929).

[298] Z. Zhou, M.M.R. Siddiquee, N. Tajbakhsh, J. Liang, (2018), UNet++: A nested U-Net architecture for medical image segmentation. In Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support (pp. 3-11).

[299] Y. Zhuge, et al. (2020), Automated glioma grading on conventional MRI images using deep convolutional neural networks. Medical physics, 47(7), 3044-3053.

[300] L. Zhuang, et al. (2022), A ConvNet for the 2020s. https://openaccess.thecvf.com/content/CVPR2022/papers/Liu_A_ConvNet_for_the_2020s_CVPR_2022_paper.pdf.

[301] J.N. Weinstein et al. The cancer genome atlas pan-cancer analysis project. Nat Genet 2013; 45:1113–20.

[302] Genomic data commons data portal (legacy archive). https://portal.gdc.cancer.gov/ legacy-archive/, Accessed date: 30 August 2017.

[303] The genotype-tissue expression (GTEx) project. Nat Genet 2013;45:580–5. https:// doi.org/10.1038/ng.2653.

[304] Biospecimen Research Database. https://brd.nci.nih.gov/brd/image-search/ searchhome, Accessed date: 30 August 2017.

[305] R.J. Marinelli, et al. The Stanford tissue microarray database. Nucleic Acids Res 2008; 36: D871–7. https:// doi.org/10.1093/nar/gkm861.

[306] TMAD main menu. https://tma.im/cgi-bin/home.pl, Accessed date: 29 November 2017.

[307] MICCAI grand challenge: tumor proliferation assessment challenge (TUPAC16). MICCAI Gd Chall Tumor Prolif Assess Chall TUPAC16. http://tupac.tue-image.nl/, Accessed date: 29 November 2017.

[308] A. BenTaieb, H. Li-Chang, D. Huntsman, G. Hamarneh, A structured latent model for ovarian carcinoma subtyping from histopathology slides. Med Image Anal 2017; 39: 194–205. https://doi.org/10.1016/j.media.2017.04.008.

[309] Ovarian carcinomas histopathology dataset. http://ensc-mica-www02.ensc.sfu.ca/ download/, Accessed date: 29 November 2017.

[310] M. Babaie, et al. Classification and retrieval of digital pathology scans: a new dataset. ArXiv170507522 Cs; 2017.

[311] KimiaPath24: dataset for retrieval and classification in digital pathology. KIMIA Lab; 2017.

[312] M.D. Kumar, M. Babaie, S. Zhu, S. Kalra, H.R. Tizhoosh. A comparative study of CNN, BoVW and LBP for classification of histopathological images. ArXiv171001249 Cs; 2017.

[313] KIMIA Lab: Image Data and Source Code. http://kimia.uwaterloo.ca/kimia_lab_data_Path960.html, Accessed date: 29 November 2017.

[314] E.D. Gelasca, J. Byun, B. Obara, B.S. Manjunath. Evaluation and benchmark for biological image segmentation. 2008 15th IEEE Int. Conf. Image Process; 2008. p. 1816–9. https://doi.org/10.1109/ICIP.2008.4712130.

[315] Bio-segmentation | center for bio-image informatics | UC Santa Barbara. http://bioimage.ucsb.edu/research/bio-segmentation, Accessed date: 29 November 2017.

[316] Bioimaging Challenge 2015 Breast Histology Dataset - CKAN. https://rdm.inesctec. pt/dataset/nis-2017-003, Accessed date: 1 December 2017.

[317] K. Sirinukunwattana, J.P.W. et al. Gland segmentation in colon histology images: the glas challenge contest. Med Image Anal 2017; 35: 489–502. https://doi.org/10.1016/j.media.2016.08.008.

[318] BIALab@Warwick: GlaS Challenge Contest. https://warwick.ac.uk/fac/sci/dcs/research/tia/glascontest/, Accessed date: 29 November 2017.

[319] F.A. Spanhol, L.S. Oliveira, C. Petitjean, L. Heutte, Breast cancer histopathological image classification using convolutional neural networks. 2016 Int. Jt. Conf. Neural Netw. IJCNN; 2016. p. 2560–7. https://doi.org/10.1109/IJCNN.2016.7727519.

[320] J.N. Kather, et al. Continuous representation of tumor microvessel density and detection of angiogenic hotspots in histological whole-slide images. Oncotarget 2015; 6: 19163–76. https://doi.org/ 10.18632/oncotarget.4383.

[321] MITOS-ATYPIA-14 - Dataset. https://mitos-atypia-14.grand-challenge.org/dataset/, Accessed date: 29 November 2017.

[322] N. Kumar, et al. A Dataset and a technique for generalized nuclear segmentation for computational pathology. IEEE Trans Med Imaging 2017; 36: 1550–60. https://doi.org/10.1109/TMI.2017. 2677499.

[323] Nucleisegmentation. Nucleisegmentation. http://nucleisegmentationbenchmark.weebly.com, Accessed date: 29 November 2017.

[324] Mitosis detection in breast cancer histological images. http://ludo17.free.fr/mitos_2012/index.html, Accessed date: 29 November 2017.

[325] A. Janowczyk, A. Madabhushi, Deep learning for digital pathology image analysis: a comprehensive tutorial with selected use cases. J Pathol Inform 2016; 7: 29. https:// doi.org/10.4103/2153-3539.186902.

[326] A. Janowczyk. Andrew Janowczyk - Tidbits from along the way. http://www.andrewjanowczyk.com, Accessed date: 29 November 2017.

[327] A. Gertych, et al, Machine learning approaches to analyze histological images of tissues from radical prostatectomies. Comput Med Imaging Graph 2015; 46: 197–208. https://doi.org/10. 1016/j.compmedimag.2015.08.002.

[328] Z. Ma, et al. Data integration from pathology slides for quantitative imaging of multiple cell types within the tumor immune cell infiltrate. Diagn Pathol 2017; 12. https://doi.org/10.1186/ s13000-017-0658-8.

[329] N. Linder, et al. Identification of tumor epithelium and stroma in tissue microarrays using texture analysis. Diagn Pathol 2012; 7: 22. https://doi.org/10.1186/1746-1596-7-22.

[330] egfr colon stroma classification. http://fimm.webmicroscope.net/supplements/epistroma, Accessed date: 29 November 2017

[331] Y. Xu, et al. Large-scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features. BMC Bioinformatics 2017; 18: 281. https://doi.org/10. 1186/s12859-017-1685-x.

[332] G. B. Chavhan, Z. Alsabban, P. S. Babyn, S. Punwani, (2020), A review of common evaluation metrics for delineation of structures in medical imaging. European Journal of Radiology, 126, 108940.

CURRICULUM VITAE

Photo

Name Surname : Gökhan ABBASOĞLU

Place and Date of Birth :

E-Mail :

EDUCATION:

B.Sc. : 1992, METU, Engineering F., Electrics and Electronics Eng.

MBA : 1995, İstanbul Üniversitesi İşletme İktisadı Enstitüsü, İşletmecilik İhtisas Programı

PROFESSIONAL EXPERIENCE AND REWARDS:

2018 – : TÜBİTAK BİLGEM (Business Development Specialist, Head of Analytic Devices and Systems Department)

- Business development activities for the products and services of BİLGEM Artificial Intelligence Institute and Block Chain Applications Department.
- Cooperation and Technology Transfer opportunities with the private sector.

2008 –2018 : Bilgi Sanatı Teknoloji Hizm. Ltd. (Owner)

- Application and Consultation of Building Electric and Electronic Systems.
- Fire Alarm Systems, Camera Systems, Access Control Systems, CO Alarm Systems.
- Smart Home Systems, Building Automation Systems.
- Wireless data transfer, wireless network solutions in wide area.
- Digital Signage solutions
- Consultation of ERP Systems in construction industry (TOKİ, Emlak Konut, Akkom).

2005 - 2008 : Belbim A.Ş. - İstanbul (Manager of Business and Project Development, Project Leader)

- Technical deputy of General Manager.
- Project Leadership of Akyolbil Project which has 13.500.000 USD budget. This project includes fleet management of 5.000 İETT busses and 800 passenger information units at bus stops. The project requires high skills in several

technologies including GPS, GPRS, Wi Fi, Embedded System Development, Mobile Software development, Database, Server Systems, SOA, Web Development etc.

- Formation of project team.
- Preparation of specifications for hardware, software and firmware.
- Preparation of contracts for subsidiaries.
- Management of both software development group and firmware development group.
- Project Leadership of İsfalt Vehicle Tracking System.
- Project Leadership of İSKİ Vehicle Tracking System.
- Project Leadership of Personel Attandance Control System with smartcard.
- Management of most important product of Belbim which is called "Akbil". Akbil has 500.000.000 USD revenue per year and 25.000.000 USD profit for Belbim.
- Management of both software development group and firmware development group.
- Business development activities for marketing.
- Management of Customer Call Center of Akbil.
- Management of several Research and development Projects.
- Management of an agile project called "Electronic Ticket Integration Project". This project was a challenging one which was actually evaluated to take three months but was finished in three weeks with a successful project management execution. At the end of project our team is awarded by General Secretary of Metropolitan Municipality and General Manager.
- Reorganization studies for Research and Development Teams.
- Organization of a new team and automation system for Bug Tracking System of Akbil.
- Organization of a new team for stochastic analysis of electronic ticketing system. By the use of this system we developed a new quality control system for Akbil which takes its philosophy from six sigma.
- Field Research for new smartcard system and vehicle tracking system in Monaco, Scotland and England.
- Management and controlling of several projects in Projects Group Directorates.
- Consulting Sales department in many sales activities.

- Management of Vehicle Tracking Project for IETT. Established a software development team for this project. Developed specifications for subcontractors and contributed to subcontractor selection process. Prepared reports for candidate firms.
- Contributed to Taxi Call Center Project in analysis phase.
- Contributed to IBB Call Center Project in analysis phase.
- Established a version control and issue tracking system for Projects Group.
- Analysis and Design of One Time Barcode Ticket Project, Implementation of business logic and persistency layer by using Java, Hibernate, Tomcat in a 3 tier architecture.
- Contribute to ERP Implementation Project in İstanbul Municipality
- Corporations. Participated in preparation of ERP Selection Guide Handbook, process analysis and improvement of several departments, preparation of ERP Requirements Lists and interviews with all candidate firms including SAP, Oracle, Microsoft, IFS.
- Developed an XML based solution for sales department to follow sales activities using MS Infopath and Share Point Portal Server.

1997 - 2005 : MedArt Ltd.- İstanbul (General Manager, joint owner(%33))

- Project Leadership of Erdoğanlar Alüminyum A.Ş. ERP Project.
- Conceptual and Architectural Analysis and Design of Marketing
- Management Module including Customer Relationship Management and
- Sales activities management in Erdoğanlar ERP. A new technology Software AG's Tamino XML Server is used as persistency mechanism instead of a relational database. This was a web project in a three tier architecture where JSP lies in both client side and middle tier.
- Conceptual and Architectural Analysis and Design of Manufacturing Planning Module. Implementation of Business Logic layer by using Java, Hibernate and Sybase ASA in a three tier architecture.
- Analysis Design and Implementation of Manufacturing and Inventory Control Project using Power Builder and Sybase ASA in a client Server environment.

- Development of MedArt Hospital Information System Software Package by using Power Builder and Microsoft SQL Server. Implementation of this package in several hospitals.
- Development of MedArt Policlinic Information System Software Package by using Power Builder and Sybase SQL Anywhere. This product is implemented in several clinics.
- Development of Laboratory and Radiology Information System Software Package by using Power Builder and Sybase SQL Anywhere.
- Development of Alpsan A.Ş. Plastic Profile Manufacturing Controlling System Power Builder and Microsoft SQL Server.
- Development of MedArt the World of Doctor - office management package by using Power Builder and Sybase SQL Anywhere.
- Development of İ.Ü. Microbiology Department Laboratory Record System. The system includes barcodes, electronic workflow, electronic signature, Patient Record System, exper systems as a whole and no paper is used until the patient takes his report from a KIOSK. The success of the system is mentioned in some newspapers and TV's.
- Implemented several Hardware and Network Projects for customers.
- Contributed marketing and sales activities of all these products.

2003 - 2004 : Petran Bilgisayar - İstanbul (Software Development Manager, joint owner(%25))

- Project Leadership Petran B-B and B-C project. An online buying and selling site has developed.
- Management and planning of Java Training Program for Software Development Department.
- Planning and management of migration of Software Infrastructure to Java platform.
- Organized training programs for Technical Service personnel.
- Implemented several Hardware and Network Projects for customers.
- Contributed marketing and sales activities of software, hardware and service.

1998 - 2003 : Erdemsoft Training Center - İstanbul (Part Time Trainer)

- MCSE Training programs. Over 85% of trainees is succeeded.

1998 - 2001 : Bebim PEC Training Center - İstanbul (Part Time Training Coordinator)

- Prepared and gave presentations for Beşiktaş branch. After three months of starting income of the branch exceeded $100.000. This amount was three times of preceding three months.
- Prepared education programs for MCSE, English Prep, MS Office, Compti A, A+ certifications.
- Performed MCSE, Compti A and A+ trainings.
- Training Coordinator for Bakırköy Branch.
- Gave MS Office trainings.
- Contribute Bebim to take Microsoft Technical Training Center Certificate.

1996 - 1997 :Black Sea Coast Security Forces - Samsun (Miltary)

- Developed an inventory record system, a cryptology system and
- Implemented TURPAK packet switching network.

1994 - 1996 : Belbim A.Ş. – İstanbul (Project Manager)

- Developed an Inventory control system over VAX, VMS operating system and Ingres Database by using Ingres 4 GL. This was the first successful project of Belbim A.Ş. in 1994.
- Conceptual analysis of Accounting Project.
- Implemented Several Network and Hardware projects.
- Requirement analysis for several departments of İstanbul Municipality.

1993 - 1994 : MedTek A.Ş. – İstanbul (Product Technical Support)

1992 - 1993 : Doğa Madencilik A.Ş. – İstanbul (Technical Manager)

- Foundation of a granite treating factory including four cutting unit and one polishing unit. This factory was the third one in Turkey at that time, later it became the first one in the Europe.

1991 - 1992 : TUBITAK MODISA-5 Project – Ankara (Technical Consultant)

Prepared electro mechanic project for an Industrial Robot arm. Developed a simulation program in C, where robot arm learns moving patterns with a three layer neural network.

AN AI IMPLEMENTATION OF DIGITAL PATHOLOGY FOR THYROID CARCINOMA

ORİJİNALLİK RAPORU

%12	%10	%11	%1
BENZERLİK ENDEKSİ	İNTERNET KAYNAKLARI	YAYINLAR	ÖĞRENCİ ÖDEVLERİ

BİRİNCİL KAYNAKLAR

1	arxiv.org İnternet Kaynağı	%4
2	Daisuke Komura, Shumpei Ishikawa. "Machine Learning Methods for Histopathological Image Analysis", Computational and Structural Biotechnology Journal, 2018 Yayın	%2
3	www.researchgate.net İnternet Kaynağı	%2
4	www.arxiv-vanity.com İnternet Kaynağı	%1
5	Ilaria Girolami, Stefano Marletta, Liron Pantanowitz, Evelin Torresani et al. "Impact of Image Analysis and Artificial Intelligence in Thyroid Pathology, with particular reference to cytological aspects.", Cytopathology, 2020 Yayın	<%1
6	link.springer.com İnternet Kaynağı	<%1