INTEGRATING HYPERSPECTRAL IMAGING AND MICROSCOPY FOR
HEPATOCELLULAR CARCINOMA DETECTION FROM H&E STAINED
HISTOPATHOLOGY IMAGES

A THESIS SUBMITTED TO

THE GRADUATE SCHOOL OF INFORMATICS OF

THE MIDDLE EAST TECHNICAL UNIVERSITY

BY

UMUT ÇİNAR

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

IN

THE DEPARTMENT OF INFORMATION SYSTEMS

JUNE 2023

INTEGRATING HYPERSPECTRAL IMAGING AND MICROSCOPY FOR HEPATOCELLULAR CARCINOMA DETECTION FROM H&E STAINED HISTOPATHOLOGY IMAGES

Submitted by **UMUT ÇİNAR** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy in Information Systems Department, Middle East Technical University** by,

Date: _19.06.2023_

I hereby declare that all information in this document has been obtained and presented in accordance with 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.

Name, Last name :   Umut Çinar

Signature   :   _____

To the valiant cancer patients and their steadfast families,
who inspire me with their strength. May this work aid in
advancing cancer research and improving treatments.

ABSTRACT

INTEGRATING HYPERSPECTRAL IMAGING AND MICROSCOPY FOR HEPATOCELLULAR CARCINOMA DETECTION FROM H&E STAINED HISTOPATHOLOGY IMAGES

Çinar, Umut

Ph.D, Department of Information Systems

Supervisor: Prof. Dr. Yasemin Yardımcı Çetin

June 2023, 110 pages

In this study, we introduce a new method for classifying Hepatocellular Carcinoma (HCC) using a hyperspectral imaging system (HSI) integrated with a light microscope. We developed a custom imaging system that captures 270 bands of hyperspectral images from healthy and cancerous tissue samples with HCC diagnosis taken from a liver microarray slide. To build an accurate classification model, we utilized Convolutional Neural Networks (CNNs) with 3D convolutions (3D-CNN). These convolutions incorporate both spectral and spatial features within the hyperspectral cube to train a robust classifier. By leveraging 3D convolutions, we can collect distinctive features automatically during CNN training without requiring manual feature engineering on hyperspectral data. Our proposed method is compact and can be applied effectively in medical HSI applications. Additionally, we addressed the class imbalance problem in the dataset by utilizing the focal loss function as the CNN cost function. This function emphasizes hard examples to learn and prevents overfitting caused by the lack of inter-class balancing. Our empirical results indicate that hyperspectral data outperforms RGB data in liver cancer tissue classification, and increased spectral resolution leads to higher classification accuracy. Furthermore, we found that spectral and spatial features are both critical for training an accurate classifier for cancer tissue classification.

Keywords: hyperspectral imaging, hyperspectral microscopy, human carcinoma detection, convolutional neural networks, 3D convolutions

ÖZ

H&E BOYALI HİSTOPATOLOJİ GÖRÜNTÜLERİNDEN HEPATOSELLÜLER KARSİNOM TESPİTİ İÇİN HİPERSPEKTRAL GÖRÜNTÜLEME VE MİKROSKOP ENTEGRASYONU

Çinar, Umut

Doktora, Bilişim Sistemleri Bölümü

Tez Yöneticisi: Prof. Dr. Yasemin Yardımcı Çetin

Haziran 2023, 110 sayfa

Bu çalışmada, ışık mikroskobu ile entegre bir hiper-spektral görüntüleme sistemi (HSI) kullanarak Hepatosellüler Karsinom (HCC) sınıflandırması için yeni bir yöntem sunuyoruz. HCC tanısı alan sağlıklı ve kanserli doku örneklerinden oluşan karaciğer mikro-dizi slaytlarından 270 bant hiper-spektral görüntü elde etmek için özel bir görüntüleme sistemi geliştirdik. Doğru bir sınıflandırma modeli oluşturmak için, 3 boyutlu evrişimlerle (3B-CNN) Evrişimsel Sinir Ağları (CNN) kullandık. Bu evrişimler, hiper-spektral küpte hem spektral hem de mekânsal özellikleri içerir ve sağlam bir sınıflandırıcı eğitmek için kullanılır. 3B evrişimler kullanarak, CNN eğitimi sırasında hiper-spektral verilere manuel özellik mühendisliği yapmaya gerek kalmadan otomatik olarak ayırt edici özellikler toplayabiliriz. Önerilen yöntemimiz kompakt olup tıbbi HSI uygulamalarında etkili bir şekilde uygulanabilir. Ayrıca, veri kümesindeki sınıf dengesizliği sorununu, CNN maliyet fonksiyonu olarak odak kayıp fonksiyonunu kullanarak ele aldık. Bu fonksiyon, zor örneklerin öğrenilmesini vurgular ve sınıflar arası dengesizliğin eksikliği nedeniyle meydana gelen aşırı uyumu önler. Deneysel sonuçlarımız, karaciğer kanseri doku sınıflandırmasında hiper-spektral verilerin RGB verilerinden daha iyi performans gösterdiğini ve artan spektral boyutun daha yüksek sınıflandırma doğruluğuna yol açtığını göstermektedir. Ayrıca, kanser doku sınıflandırması için doğru bir sınıflandırıcı eğitmekte spektral ve mekânsal özelliklerin her ikisinin de kritik olduğunu bulduk.

Anahtar Sözcükler: hiperspektral görüntüleme, hiperspektral mikroskopi, insan karsinomu tespiti, evrişimsel sinir ağları, 3B evrişimler

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

HSI	Hyperspectral imaging
UV	Ultraviolet
SWIR	Shortwave infrared
H&E	Hematoxylin and Eosin
HCC	Hepatocellular carcinoma
HBV	Hepatitis B virus
HCV	Hepatitis C virus
3D-CNN	3D Convolutional Neural Networks
VNIR	Visible Near Infrared
AOTF	Acousto-Optic Tunable Filter
PCA	Principal Component Analysis
CAD	Computer-aided diagnosis
CT	Computed tomography
MRI	Magnetic resonance imaging
CNN	Convolutional neural network
SVM	Support vector machines
ILD	Interstitial lung diseases
HRCT	High-resolution computed tomography
LR	Logistic regression
ANN	Artificial neural network
AUC	Area Under the ROC Curve
GAN	Generative Adversarial Network
WSI	Whole-slide image
FCN	Fully convolutional network
Bi-LSTM	Bidirectional long short-term memory
GuSA	Guided soft attention network
MCIL	Multiple clustered instance learning
MIL	Multiple instance learning
RNN	Recurrent neural network
ROI	Region-of-interest
NIR	Near-infrared
DLP	Digital light processing
AMD	Age-related-macular-degeneration
LS-SVM	Least squares support vector machine
SLIC	Simple linear iterative clustering
RBC	Red blood cell
SAM	Spectral angle mapper
WBC	White blood cells
ASPP	Atrous Spatial Pyramid Pooling
ResNet	Residual Network

ReLU	The Rectified Linear Unit
CE	Cross entropy
BCE	Balanced cross entropy
FL	Focal Loss
NA	Numerical aperture
EPE	Expected prediction error
EMC	Effective model complexity
MCC	Matthews Correlation Coefficient
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative

CHAPTER 1

INTRODUCTION

Hyperspectral imaging (HSI) is an advanced remote sensing technology that has garnered significant attention lately due to its ability to provide detailed spectral information across a wide range of wavelengths. This non-invasive technique captures along the electromagnetic spectrum, providing a continuous spectral curve for each pixel in an image [1]. By exploiting the unique spectral signatures of various materials, hyperspectral imaging has found applications in numerous fields, including agriculture, environmental monitoring, mineral exploration, and defense, among others [2]. As seen in Figure 1, unlike multispectral imaging, which records data at a limited number of spectral bands, hyperspectral imaging acquires data at hundreds or thousands of neighboring spectral bands, enabling the extraction of precise information about the composition and properties of the imaged objects. This high-dimensional data, known as hyperspectral data cubes, consist of two spatial dimensions and one spectral dimension. Hyperspectral data cubes enable the identification, classification, and quantification of materials, even those that are not distinguishable in the visible spectrum or conventional imaging devices. The process of acquiring hyperspectral images involves the use of specialized sensors, which capture the reflected, emitted, or transmitted radiation from the target scene. These sensors can employ various scanning mechanisms, such as whisk-broom, push-broom, or global imaging techniques, depending on the utilized imaging device [3]. The acquired hyperspectral data is then subjected to various preprocessing steps, including radiometric and atmospheric corrections, to ensure the accurate representation of the ground reflectance. That is, hyperspectral imaging is a powerful remote sensing tool that provides spectral information for a wide range of applications. Its ability to capture high-resolution, continuous spectral data enables the detection and characterization of materials that would be otherwise indiscernible.

The potential applications of hyperspectral imaging (HSI) in scientific research are broad and varied. For instance, in geology, HSI can reveal mineralogical compositions of soils or rocks that cannot be discerned by the naked eye, aiding in resource exploration and management [4]. In agriculture, hyperspectral data can provide information about crop health, soil quality, and pest infestation, thereby guiding precision farming practices for improved yield and sustainability [5]. Additionally, in the field of environmental science, HSI can track changes in ecosystems over time, assess water quality, and monitor pollution levels [6]. The wealth of information provided by HSI, coupled with advancements in data processing techniques and machine learning algorithms, offers the potential to uncover hidden patterns and insights that could further our understanding of various phenomena. Therefore, the continued development and refinement of HSI technology and its integration with other technological advancements could pave the way for significant breakthroughs across multiple fields.

1

ELECTROMAGNETIC SPECTRUM
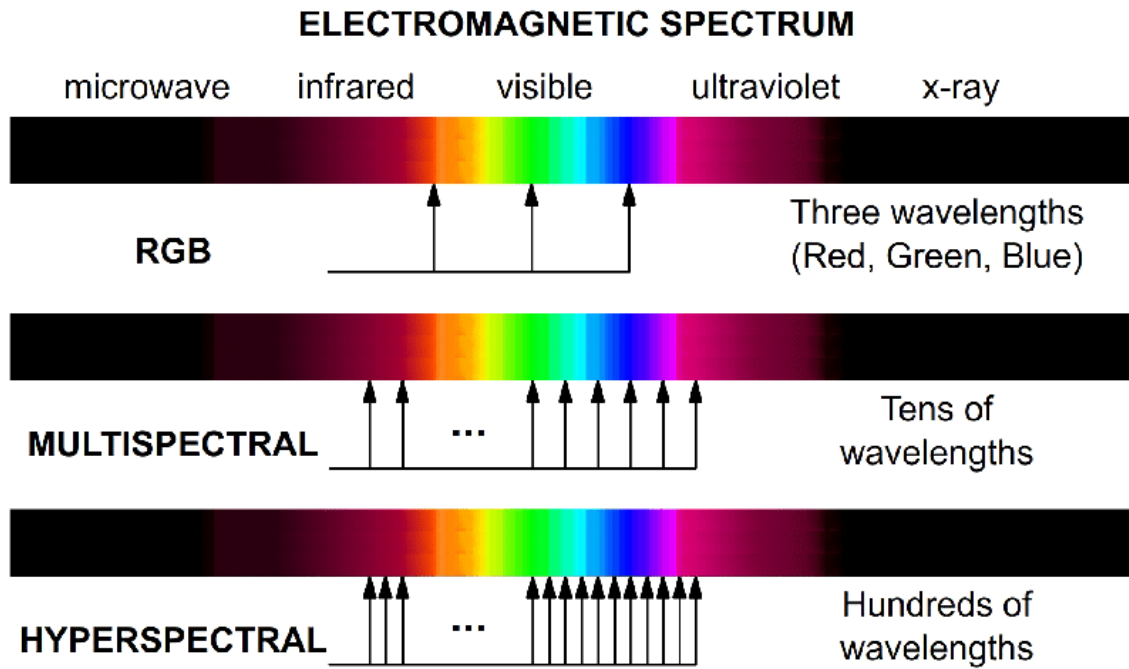
Figure 1 RGB, Multispectral and Hyperspectral coverage on electromagnetic spectrum [7]

On the other hand, histopathology is a major discipline within the field of pathology, focusing on the microscopic examination of tissue samples to diagnose diseases and evaluate treatment outcomes. It involves the study of morphological and cellular changes in tissues, as well as the identification of specific disease markers, to provide accurate, detailed information on the underlying pathological processes [8]. Histopathology plays a critical role in various aspects of medical practice, such as clinical diagnostics, prognostication, and therapy decision-making, particularly in oncology, where it aids in tumor classification, staging, and grading. The conventional workflow in histopathology involves several steps, including tissue fixation, embedding, sectioning, staining, and finally, microscopic examination by a pathologist [9]. Among the most common staining techniques is Hematoxylin and Eosin (H&E) staining, which highlights the basic structure and composition of the tissue by differentiating between the cell nuclei and cytoplasm [10]. Additionally, immunohistochemistry and in situ hybridization techniques are employed to detect specific antigens or nucleic acid sequences within the tissue, providing vital information on the molecular characteristics of the disease [11]. In fact, the application of HSI in histopathology holds great potential for the discovery of novel biomarkers by identifying spectral signatures associated with specific pathological states [12]. This can ultimately lead to a better understanding of disease pathogenesis and facilitate the development of targeted therapeutic interventions. Despite the invaluable insights gained from histopathological analysis, the process is labor-intensive, time-

consuming, and prone to intra- and inter-observer variability [13]. Moreover, the evaluation of large-scale, high-resolution histopathological images demands considerable expertise and experience, leading to potential diagnostic discrepancies and delays in patient care. To address these challenges, recent advancements in digital pathology, computational and imaging techniques such as HSI have paved the way for the development of automated histopathological analysis systems, which leverage image processing, machine learning, and deep learning algorithms to boost diagnostic accuracy, efficiency, and reproducibility [14]. One of the key strengths of HSI is its ability to capture spatial and spectral information from a sample, generating a continuous spectrum for each pixel in the image. This detailed chemical and structural information enables HSI to reveal subtle differences in tissue compositions and structures that may be indistinguishable using conventional histopathology techniques. As a result, HSI may produce more precise identification and differentiation of various tissue types, leading to enhanced diagnostic accuracy and reduced inter-observer variability. Integrating hyperspectral imaging into histopathology tasks can offer significant research value and has the potential to open new horizons for diagnostic practices and broaden the understanding of disease processes.

1.1. Motivation

In 2020, an estimated 960,000 new liver cancer cases were reported, resulting in approximately 830,000 deaths attributed to this disease [15]. Liver cancer is the sixth most prevalent type of cancer and the third leading cause of cancer-related deaths worldwide. Hepatocellular carcinoma (HCC) accounts for the majority of liver cancer cases, with an incidence rate of 80%. As depicted in Figure 2, the HCC tumors develop in the liver, which is located in the upper right part of the abdomen, beneath the diaphragm and above the stomach. The liver is one of the largest organs in the human body and plays a vital role in processing nutrients, metabolizing drugs, and detoxifying harmful substances. HCC, as a primary liver cancer, has its origins in the cells of the liver itself. It arises due to a series of genetic mutations that occur within the liver cells, leading to uncontrolled cellular growth and the formation of a tumor. Given the crucial functions of the liver, such as nutrient processing, drug metabolism, and toxin detoxification, the onset of HCC can significantly disrupt these processes and lead to severe health consequences. The high prevalence and mortality rate of HCC underscores the importance of ongoing research into effective diagnostic and treatment strategies for this aggressive form of cancer.
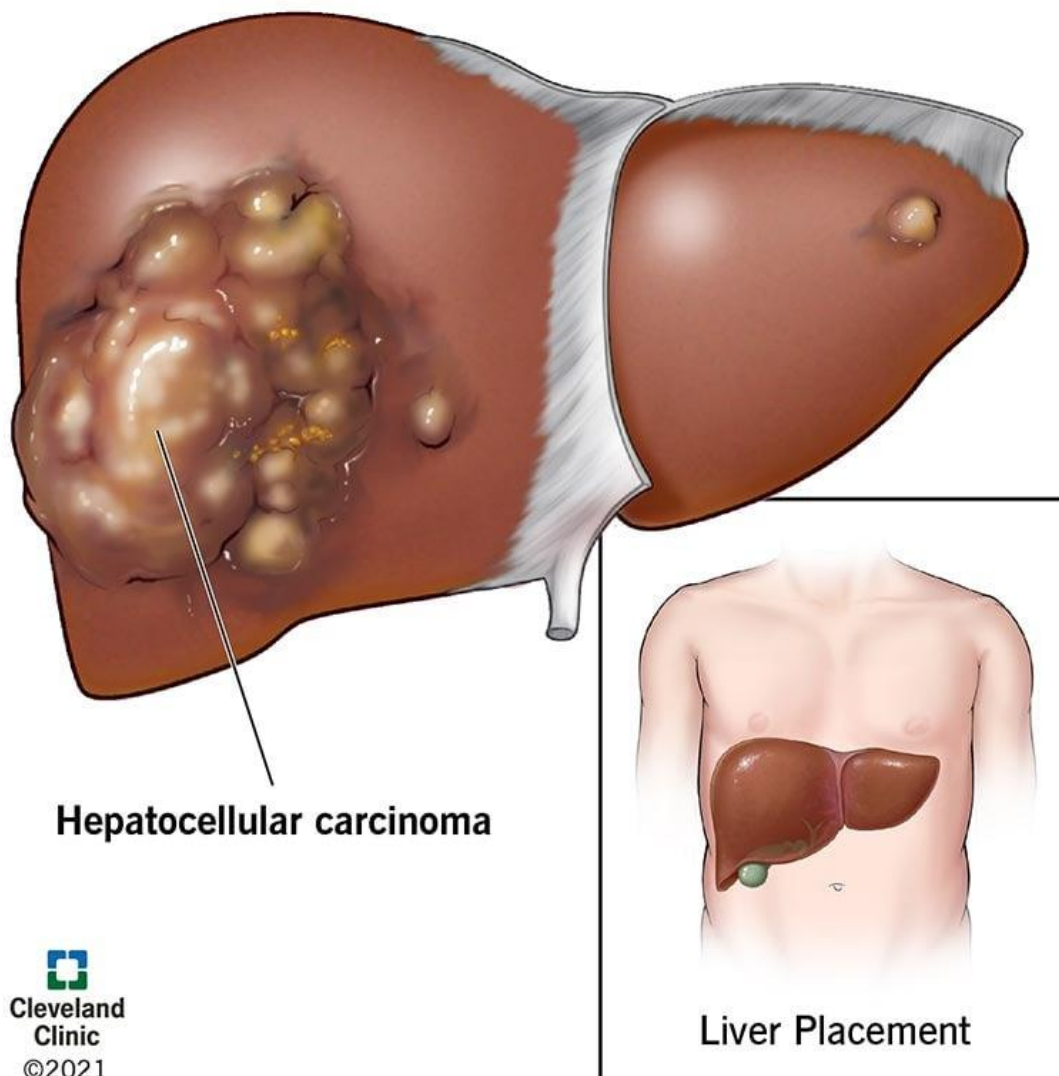
Figure 2 Hepatocellular Carcinoma (HCC) [16]

In terms of appearance, HCC tumors can vary widely [17]. Their visual characteristics can be influenced by factors such as the stage of the cancer, the overall health of the liver, and the underlying cause of the liver disease. In general, HCC tumors often appear as solid masses within the liver tissue. They can be single or multiple, and their sizes can vary greatly. Early-stage tumors may be small, often less than 3 centimeters in diameter, and can be hard to distinguish from surrounding healthy liver tissue. As the cancer progresses, the tumors can grow larger and may have an irregular or lobulated shape. Their color often differs from the surrounding liver tissue, usually being paler due to the different cell structure. Late stage HCC can also lead to a varied appearance in the liver as the tumors may cause the liver to enlarge or change shape. However, it's important to note that these characteristics are typically not visible to the naked eye. The presence and appearance of

4

HCC tumors are usually determined through imaging studies such as ultrasound, computed tomography (CT), and magnetic resonance imaging (MRI), and confirmed through biopsy, where a small sample of the tumor is examined under a microscope. Figure 3 highlights the abnormal cell development and complex morphology of this disease. In part A, the figure provides a microscopic view of dysplastic foci in which small cell changes, a possible precursor to cancer, can be observed, stained using hematoxylin and eosin. Part B of the figure shows a liver with cirrhosis, a scarring of the liver often caused by long-term liver damage, containing a single, nodular HCC, indicated by an arrow. In Part C, the figure presents a liver with cirrhosis, which hosts multiple nodules of HCC, further emphasizing the multifaceted nature of this disease. Lastly, part D exhibits a case of multicentric HCC, indicating the presence of multiple, independent tumors within the liver. This visual overview presents the wide spectrum of HCC presentations, illustrating the disease's complex development and progression within the liver.
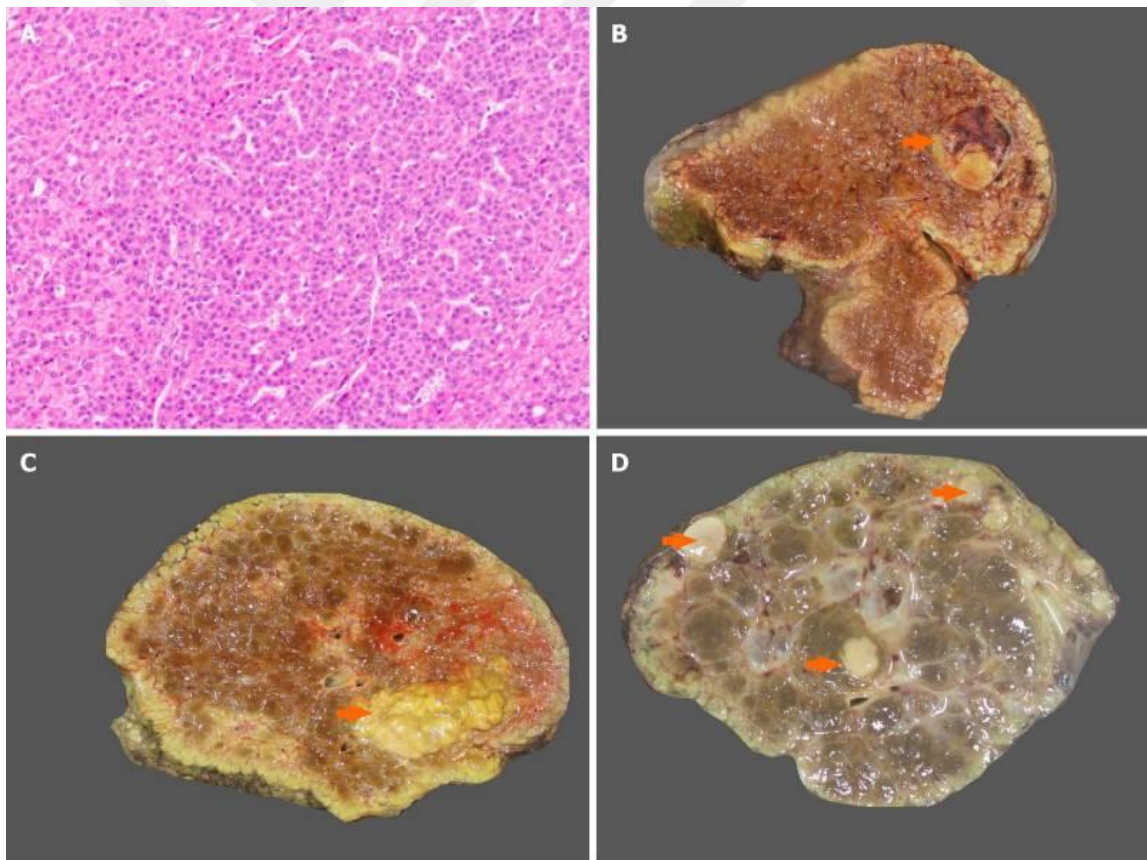
Figure 3 Changes in cell structure and the overall appearance of HCC. (A) Abnormal cell development demonstrating a small cell transformation (stained with H&E); (B) Nodular HCC located within a liver affected by cirrhosis (indicated by an arrow); (C) HCC comprising multiple nodules in a cirrhotic liver (marked by an arrow); D: HCC with multiple origins or centers (pointed by an arrow) [18].

HCC develops as a result of a multi-step process of malignant transformation of hepatocytes, the primary cell type in the liver. This complex process is influenced by a variety of risk factors and involves multiple stages of liver damage and cellular changes. Initially, as depicted in Figure 4, the hepatocytes may experience injury and inflammation due to factors such as chronic viral hepatitis (hepatitis B and C), exposure to toxins such as aflatoxin, alcoholic liver disease, or non-alcoholic steatohepatitis [19]. This chronic inflammation can lead to an increased rate of cell turnover and, eventually, the development of fibrosis or scar tissue. If the inflammation and injury persist, this fibrosis can progress to cirrhosis, a condition where the normal liver architecture is replaced by scar tissue and regenerative nodules. This altered environment can lead to genomic instability, providing an opportunity for the acquisition of additional genetic mutations.
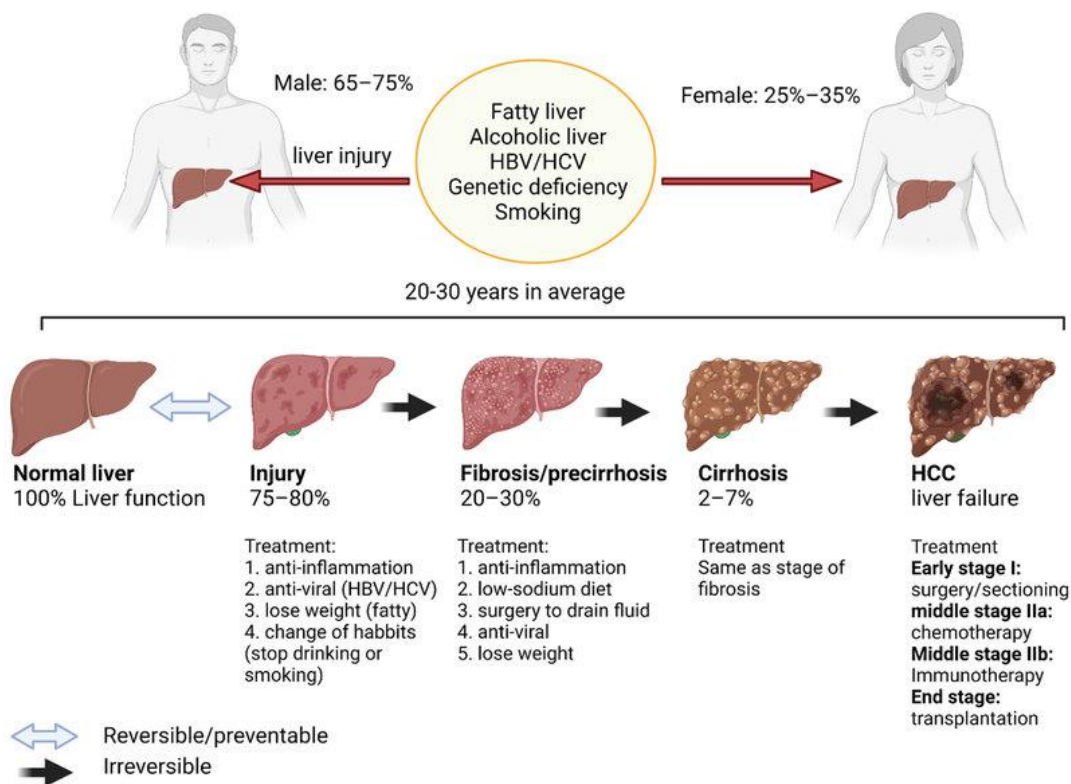
Figure 4 The transition from normal liver to HCC [20]

At the cellular level, these genetic mutations may alter the normal regulation of cell division and death, leading to uncontrolled cell proliferation. These mutations may affect key oncogenes and tumor suppressor genes, which are crucial for controlling cell growth and division. Over time, these mutated cells may undergo further changes, such as alterations in their size, shape, and organization, a process known as dysplasia. A key step in the development of HCC is the activation of telomerase, an enzyme that allows cancer cells to divide indefinitely. The mutated hepatocytes may also develop the ability to invade surrounding tissues and spread to other parts of the body, a process known as metastasis. As a result of these cumulative changes, dysplastic hepatocytes can eventually transform

into fully malignant HCC cells. These cells may form a tumor, or multiple tumors, within the liver, leading to the clinical disease known as HCC. It is important to note that the exact sequence and nature of these mutations can vary between individuals and may be influenced by a variety of factors such as the underlying cause of liver disease, individual genetics, and lifestyle factors.

Normal cells and cancer cells, as visualized in Figure 5 and Figure 6, exhibit several notable morphological differences. This disparity results from the aberrant growth and survival signals characteristic of cancer, which allow these cells to evade the body's usual mechanisms of cellular control and homeostasis [21]. Healthy hepatocytes have a distinct, uniform appearance under the microscope. These cells are generally large and polygonal, with abundant cytoplasm and a round, centrally located nucleus. The liver tissue architecture is orderly, with hepatocytes arranged in cords or plates separated by blood-filled sinusoidal spaces [22]. In contrast, HCC cells exhibit significant cellular and architectural irregularity. Firstly, their size can vary considerably, with some cells appearing much larger or smaller than typical hepatocytes. HCC cells can also exhibit changes in shape, including a less defined, more irregular cellular outline [23]. The nucleus of cancer cells often appears larger and more irregular, a feature known as pleomorphism. The nuclear to cytoplasmic ratio may be increased, and multiple nuclei may be present in some cells. Additionally, the chromatin within the nucleus may be coarser, and there may be prominent nucleoli, which are regions within the nucleus where ribosome synthesis occurs [24]. Architecturally, HCC tissue lacks the regular cord-like arrangement of normal liver tissue. Instead, it may exhibit a trabecular, pseudo-glandular, or solid pattern. Increased cell density, unregulated angiogenesis, and the presence of necrotic regions are also standard features of HCC. These morphological changes are a direct result of the genetic alterations that drive cancer development, reflecting changes in genes regulating cell growth, division, and survival. They provide valuable diagnostic and prognostic information and are often used to grade the severity of the disease [25].

The abnormal morphology of HCC cells and the disorganized tissue architecture they form are not mere consequences of cancer, but they actively contribute to the disease's progression and spread. For instance, the increased cell density seen in HCC tumors can promote a hypoxic environment, which can induce further genetic mutations and enable the survival of more aggressive cancer cells [26]. Unregulated angiogenesis, the process by which tumors create new blood vessels, ensures that rapidly proliferating cancer cells receive necessary nutrients, thereby facilitating tumor growth and offering a potential pathway for metastasis, or spread to other body parts [27]. Necrotic regions, areas of cell death caused by the harsh tumor environment, can stimulate inflammation and further promote tumor growth [28]. The morphological characteristics of HCC cells can also impact treatment response. For example, changes in cell size and shape can alter how cells interact with therapeutics, and nuclear changes can affect the cell's response to treatments that target DNA or disrupt cell division. As such, understanding these morphological differences between normal and cancer cells is not only crucial for diagnosing and grading HCC but also for developing and optimizing treatment strategies.
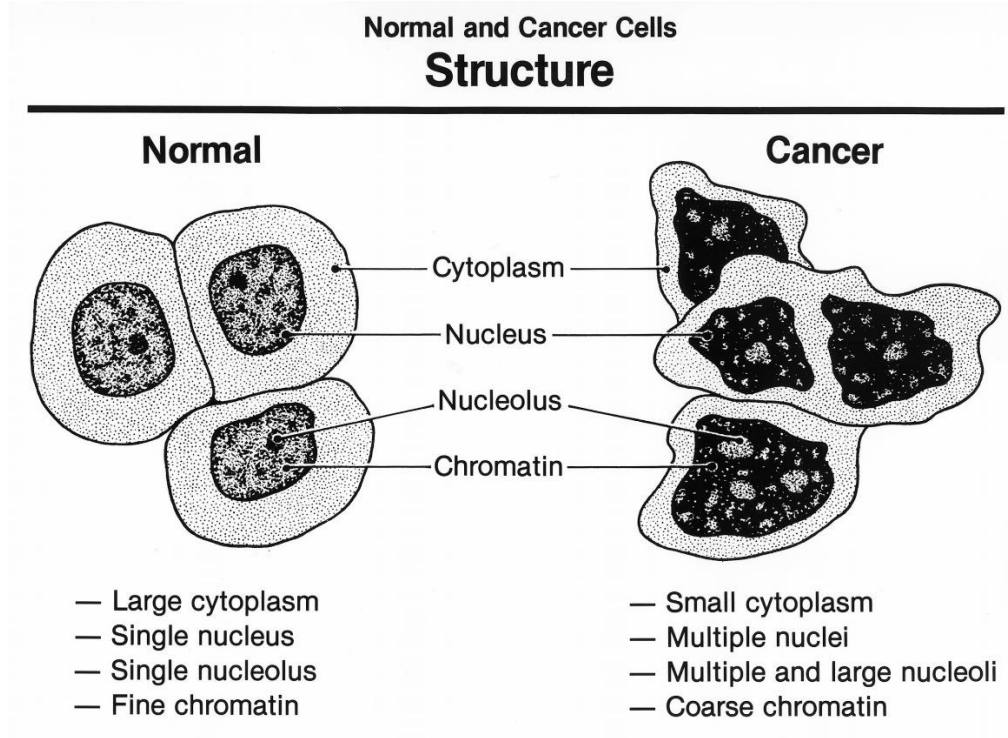
Figure 5 Characteristic structures of normal and cancer cells [29].
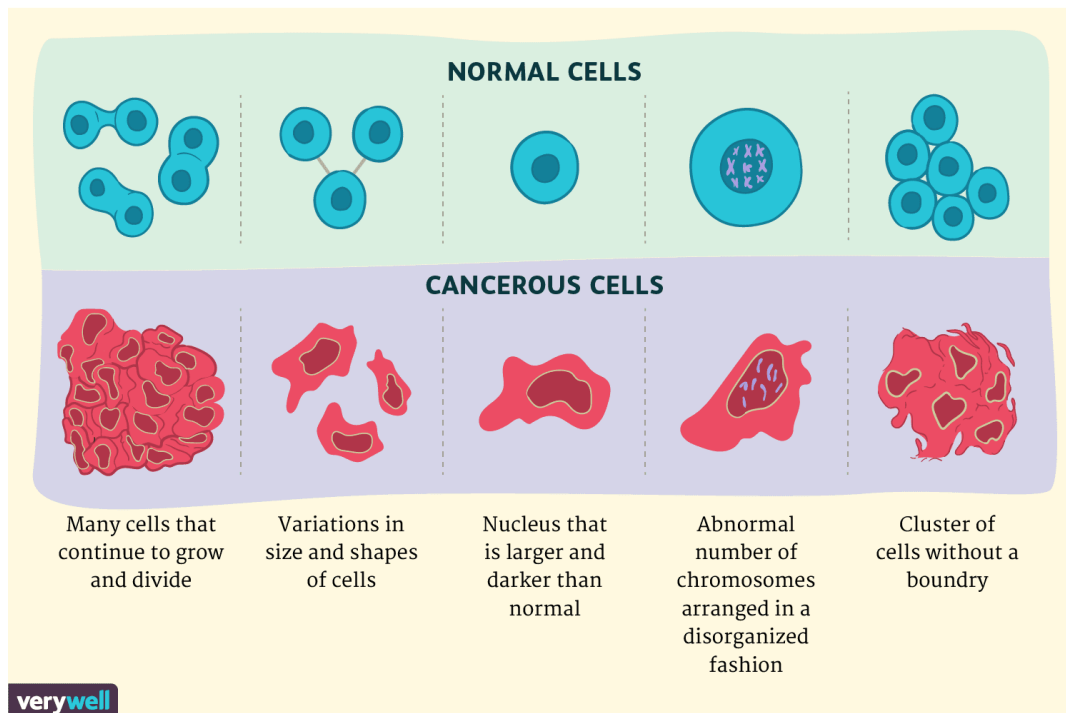
Figure 6 Differences between normal and cancer cells [30]

From the clinical perspective, ultrasonography plays a vital role in the early detection and monitoring of tumor nodules. As a non-invasive imaging technique, it allows for the real-time visualization of the liver and other internal organs, enabling the identification of abnormal growths or changes in organ structure. Regular ultrasonographic surveillance is particularly crucial for individuals at high risk of developing HCC such as those with chronic hepatitis or cirrhosis. In patients with sufficient liver function and small, solitary tumors, resection is typically the first-line treatment option. This surgical procedure involves the removal of the tumor along with a margin of healthy tissue, with the aim of eliminating all cancerous cells. The success of liver resection as a treatment largely depends on the tumor's size, location, and number, as well as the patient's overall health and liver function [31]. In cases where the disease is more advanced or the patient's health status does not allow for resection, other treatment strategies may be pursued. Liver transplantation is an option for patients with early-stage HCC who also have severe underlying liver disease. It involves replacing the diseased liver with a healthy one from a donor [32]. Trans-arterial chemoembolization (TACE) is another option, particularly for patients with larger or multiple tumors that cannot be removed surgically. TACE involves the delivery of chemotherapy drugs directly to the tumor through the blood vessels feeding it, followed by the embolization (blockage) of these vessels to starve the tumor of its blood supply [33-34]. Finally, molecular-targeted therapies have emerged as a promising treatment approach, especially for advanced HCC. These therapies work at the molecular level, interfering with specific signaling pathways that cancer cells rely on for growth and survival. One example is tyrosine kinase inhibitors, which block the activity of enzymes known as tyrosine kinases, involved in several cellular processes, including cell growth and division [35]. Treatment for HCC is usually personalized, based on factors such as the stage of the disease, the patient's overall health and liver function, and the genetic makeup of the tumor. In some cases, a combination of treatment approaches may be used to achieve the best possible outcome [36].

On the other hand, a biopsy is a procedure that involves the removal of a small amount of tissue for examination under a microscope. In the case of HCC, a liver biopsy can be a critical component of the diagnostic process, providing definitive evidence of cancer [37]. During a liver biopsy, a thin needle is inserted through the skin and into the liver to collect a small tissue sample. This sample is then examined by a pathologist, a doctor who specializes in interpreting laboratory tests and evaluating cells, tissues, and organs to diagnose disease. The biopsy sample can provide vital information about the nature of the liver lesions identified on imaging studies. It can confirm whether a lesion is indeed cancerous, differentiate HCC from other types of liver cancer, and detect the presence of other liver diseases such as cirrhosis. Moreover, it can provide information about the grade of the tumor, which is a measure of how much cancer cells resemble normal cells and gives an indication of how quickly the tumor is likely to grow and spread [38]. In terms of treatment, the results of a liver biopsy can guide the choice of therapeutic strategies. For example, the presence of certain molecular markers in the tumor tissue can suggest that a patient might respond to specific targeted therapies [39]. However, it's worth noting that biopsy is an invasive procedure and carries risks, including bleeding and infection. Therefore, it's typically reserved for cases where imaging studies alone are insufficient for

a definitive diagnosis. In fact, pathology tools are instrumental in diagnosing HCC, assessing disease grading, determining the risk of recurrence post-surgery, and formulating effective treatment strategies, including drug therapies. The development of cancer entails complex interactions of genetic mutations within cells, culminating in the emergence of distinguishable cancerous tissues. Expert pathologists can recognize these tissues using biopsy procedures, which involve extracting and examining tissue samples from the affected area. Nevertheless, these tissue examination processes can be laborious, time-consuming, and prone to interobserver variability, underscoring the need for more efficient and accurate diagnostic methods [40]. Hyperspectral imaging, when integrated with microscopy, has the potential to significantly enhance diagnostic accuracy for HCC detection from histopathology images. By providing a broad range of spectral information, hyperspectral imaging allows for better differentiation of cancerous tissues from healthy ones. It can effectively address the challenges faced by traditional pathology methods and contributing to improved decision-making in cancer diagnosis.

The motivation behind this study is to explore and develop the potential of hyperspectral imaging in conjunction with microscopy for the diagnosis of HCC. Our ultimate goal is to enhance the diagnostic accuracy and reduce the burden on pathologists. By leveraging the richness of spectral information available through hyperspectral imaging, this research aims to overcome the limitations associated with conventional histopathology, such as interobserver variability and time-consuming procedures. Moreover, the study aims to develop a new approach that may aid in the automatic detection of HCC, potentially enhancing treatment outcomes and reducing mortality rates. Additionally, this study will contribute to the broader understanding of the implications of HSI on histopathology research, which is essential for the development of new cancer detection methods with other tissue samples such as brain, breast, lung, colon, skin and cervix. By tackling the major challenges in liver cancer diagnosis, this PhD thesis seeks to make an important contribution to the field of cancer research and positively affect the lives of patients impacted by HCC.

1.2. The Purpose of the Study

The primary objective of this study is to develop a comprehensive and robust decision support tool for pathologists by utilizing deep learning methodologies and integrating hyperspectral imaging with microscopy devices. Hyperspectral imaging has demonstrated significant potential for the analysis of histopathological images, as it provides data within a broader interval of the electromagnetic spectrum. It enables the extraction of more detailed information about the composition and characteristics of the sample [41]. To achieve this integration, we aim to construct a biological tissue image capture system in our laboratory. This system combines a hyperspectral camera and a light microscope with a 3D-printed motorized stepper, allowing precise control over the sample positioning and facilitating the acquisition of high-quality hyperspectral images.

Another purpose of this research is to investigate the applicability of deep learning methods on hyperspectral images captured from microscopy tissue samples. Deep learning has emerged as a powerful tool for analyzing complex data and extracting valuable information from large datasets [42]. In particular, deep learning algorithms, such as convolutional neural networks (CNN), have demonstrated remarkable success in the analysis of medical images and improved the diagnostic accuracy in various fields of pathology [43]. By exploring the potential of machine learning algorithms on hyperspectral histopathology images, we aim to uncover useful insights and improve the accuracy of histopathological assessments, ultimately contributing to better clinical outcomes for patients.

This study also aims to evaluate the classification performance of HCC using the proposed integrated hyperspectral imaging and machine learning methods. HCC is a common form of liver cancer, and accurate diagnosis and classification are essential for the development of effective treatment strategies. Traditional histopathological evaluation is often subjective, with considerable inter-observer variability, and may not capture the full complexity of the disease [44]. Additionally, as suggested by Trichromatic [45] and Opponent Process Theories [46], human color vision is limited by the color sensing capabilities of three types of cone cells in the human eye. Therefore, by leveraging the broad spectral information provided by hyperspectral imaging and the advanced classification competence of deep learning algorithms, an improvement in the classification performance of HCC is aimed.

1.3. Contribution of the Thesis

In this research, we propose a new HCC tumor detection framework, which exploits the synergistic potential of hyperspectral imaging and microscopy in conjunction with 3D Convolutional Neural Networks (3D-CNN). To facilitate our approach, we have developed a custom, in-house microscopy-based biological tissue image-capturing system. This system seamlessly integrates a push broom visible-near-infrared (VNIR) hyperspectral camera with a light microscopy device. Hence, it enables the collection of an extensive range of spectral data from liver tissue samples, covering wavelengths from 400 nm to 800 nm. The acquired images from each sample are segmented into smaller patches as inputs to the data processing pipeline. These patches are subsequently fed into a custom 3D convolution-based CNN learner, which generates a strong cancer tissue prediction model. This model incorporates both spectral and spatial features during the training phase of the classification model, thereby enhancing its performance. The 3D convolution operation is involved in extracting local spectral features within the hyperspectral cube, leading to a more comprehensive understanding of the tissue samples. In addressing the issue of class imbalance in our dataset similar to most of medical studies, we have employed the focal loss function as the CNN cost function [47]. Utilizing a fine-tuned focal loss function in our model not only mitigates the class imbalance problem but also considerably improves the overall classification performance of the model. Our empirical results demonstrate the superiority of 3D convolutions in terms of classification

accuracy when compared to 2D convolutions applied to the same dataset. Furthermore, our experiments highlight the improved performance of hyperspectral data over its RGB counterpart in the context of tumor tissue classification.

Different from existing literature, our study employs a broader range of spectral bands for tissue classification, maximizing the potential benefits of hyperspectral imaging. Although Acousto-Optic Tunable Filter (AOTF) is utilized in tissue classification tasks, some studies have raised concerns about its reliability for radiometric measurements due to the lack of homogeneity in diffraction efficiency [48-50]. To address these concerns, we have opted for a hyperspectral VNIR camera as the primary imaging equipment, ensuring more reliable and accurate data acquisition. Furthermore, we introduce a comprehensive and robust classification framework based on CNN and 3D convolutions. In contrast to other studies, we provide a clear comparison of the classification performance achieved using hyperspectral data, Principal Component Analysis (PCA) of hyperspectral data, and RGB datasets. This comparative analysis offers valuable insights into the effectiveness of various data types in the context of HCC tumor detection, potentially guiding future research in the field.

The contributions of our study can be summarized as follows;

1. We have successfully constructed a biological tissue image capture system in our laboratory by integrating a hyperspectral camera and a light microscope with a 3D-printed motorized stepper. This integration allows for the acquisition of high-resolution hyperspectral images, which can be further analyzed to improve histopathological assessments and diagnostic accuracy.

2. Our research demonstrates that hyperspectral data significantly enhances classification performance when compared to traditional RGB data. While RGB data can represent the spatial features of tumor tissues in fine detail, hyperspectral imaging captures both spatial and spectral features of tumor tissues, which in turn leverages the deep neural network classification accuracy.

3. The proposed method takes advantage of the hyperspectral cube by utilizing 3D convolutional neural networks. 3D kernels enable the extraction of voxel information in a compact manner, allowing the learner to collect both spectral and spatial features via a single convolution operation. This approach results in a more efficient and accurate feature extraction process.

4. Our method does not require manual feature engineering in the pre-processing or post-processing stages of the classification pipeline. By employing 3D convolutional neural networks, we achieve better generalization performance with a simpler network topology, streamlining the overall image analysis process.

5. In this study, we address the class imbalance problem, a common challenge in many medical image analysis studies. We have implemented the focal loss

12

function within our classification model, which compensates for class imbalance by using a focusing parameter in the cross-entropy function. This approach enhances the learner's sensitivity to misclassified samples and improves model generalization without causing overfitting.

1.4. Thesis Outline

The remainder of this dissertation is structured as follows:

- Chapter 2 presents an extensive literature review, focusing on computer-aided diagnosis, the application of hyperspectral imaging in the medical domain, and the integration of microscopy with hyperspectral instruments. We also elaborate theoretical aspects of CNNs and their essential components in this chapter. This chapter sets the foundation for understanding the current state of research in these areas and identifying potential research opportunities.

- In Chapter 3, an extensive description of our methodology is provided, containing the process of data acquisition and the steps involved in the implementation of deep learning techniques. This chapter elucidates the rationale behind our approach, detailing the various stages of the proposed framework and the underlying techniques employed.

- Chapter 4 presents the results of our experimental evaluations, comparing the performance of various sets of parameters and configurations used by the learners. The chapter showcases the effectiveness of our proposed framework, highlighting the improvements achieved in comparison to alternative experiment configurations and dataset options.

- In Chapter 5, a thorough discussion of our findings is presented, accompanied by a concise conclusion of our study. This chapter also includes the limitations of our research and offers recommendations for future research directions, aimed at advancing the field of hyperspectral imaging and computer-aided diagnosis in the medical domain.

CHAPTER 2

BACKGROUND

2.1. Computer-aided Diagnosis in Medicine

The field of computer-aided diagnosis (CAD) in medicine has shown notable advancements in recent years, driven by the rapid progress in computational techniques, artificial intelligence, and medical imaging technologies [51]. As an essential component of modern medical practice, CAD systems have been applied to a variety of imaging techniques, including computed tomography (CT), magnetic resonance imaging (MRI), and mammography. Variety of CAD applications are shown to enhance the diagnostic process and improve patient outcomes [52]. In this sub-section, we aim to provide a comprehensive understanding of CAD in medicine, discussing its evolution, key publications, techniques, applications, and challenges in the field.

CAD systems have become increasingly vital in enhancing diagnostic accuracy and efficiency across various medical domains. As healthcare professionals face growing workloads and the need for precise diagnosis, CAD systems offer valuable assistance through automated detection and classification of diseases, significantly reducing the time and effort required for manual interpretation of medical images [53]. The integration of advanced machine learning and deep learning techniques has further improved the performance of CAD systems, enabling them to improve precision in identifying and diagnosing a wide range of medical conditions, such as cancer, cardiovascular diseases, and neurological disorders [54]. The application of CAD systems in medicine not only improves diagnostic accuracy but also plays a crucial role in advancing precision medicine through the incorporation of artificial intelligence, machine learning, and radiomics.

From phonocardiography signal processing field, a technique is used to analyze heart sounds and murmurs to diagnose various cardiac conditions [55]. The proposed methodology involves the application of conventional signal processing techniques, such as wavelet transform, adaptive filtering, empirical mode decomposition, and time-frequency analysis, to extract and analyze essential features from the phonocardiogram signals. These methods allow for better representation and understanding of the heart sounds, improving the accuracy and reliability of the diagnostic process. On the other hand, the study [56] proposes an automated detection system for arrhythmias using different intervals of tachycardia ECG segments with CNNs. The authors preprocess the ECG signals to remove noise and artifacts and then extract various morphological features from the ECG segments. A CNN is used for feature learning and classification of the ECG segments into normal or arrhythmic classes. The proposed methodology demonstrates high performance in detecting different types of arrhythmias, offering a potential tool for real-time arrhythmia monitoring and diagnosis.

15

The study [57] presents a computer-aided pattern classification system for dermoscopy images, which aids in the diagnosis of skin lesions. The authors employ an image processing pipeline containing RGB to Lab color space conversion, morphological pre-processing, feature extraction with discrete wavelet transform and local binary patterns. A pattern classification system, incorporating support vector machines (SVM) is then utilized to classify the skin lesions into different categories. Another study [58] introduces a CNN model for lung pattern classification in interstitial lung diseases (ILDs). The authors use high-resolution computed tomography (HRCT) images from different hospitals and scanners as input to train the CNN model, which learns to recognize different ILD patterns. The model topology consisted of five convolutional layers with 2x2 kernels and one average pooling layer following each convolution layer. The system outperforms traditional machine learning methods with its 85.5% accuracy and other feature-based classification techniques, showcasing the potential of deep learning in medical image analysis and ILD diagnosis. This approach allows for more accurate, efficient, and automated classification of ILD patterns. In another study [59], the authors compare logistic regression (LR) and artificial neural network (ANN) models in estimating breast cancer risk. The authors utilize a dataset containing various demographic, medical, and clinical factors, which are used as input features for both the LR and ANN models. The performance of the two models is compared in terms of predictive accuracy, calibration, and discrimination. The results indicate that the ANN model outperforms the LR model in risk estimation, suggesting that ANN models can be more effective in predicting breast cancer risk. Another example of a CAD prediction model has been devised with a machine learning-based framework for long-term Type 2 diabetes risk prediction utilizing features derived from the English Longitudinal Study of Ageing Database [60]. Their methodology involved multiple steps, including data preprocessing to handle missing values and normalize data, feature selection using the Correlation-based Feature Selection technique to identify the most informative variables, and risk prediction using several machine learning algorithms, such as Naïve Bayes Decision Trees, Logistic Regression and ANN. Additionally, the authors employed a majority voting ensemble method to combine the predictions of the individual classifiers, improving the overall accuracy and robustness of the model. The ensemble approach has led to provide Area Under the ROC Curve (AUC) of 0.885 classification performance. The study demonstrated the strength of machine learning techniques in predicting Type 2 diabetes risk with high accuracy.

From the field of automatic mammography inspection, the paper [61] presents a prospective study on 12,860 patients at a community breast center to investigate the effectiveness of CAD methods in screening mammography. The study compared the performance of radiologists in detecting breast cancer with and without the assistance of a CAD system. The CAD system used in the study was based on the detection of microcalcifications and masses in mammograms using image processing techniques, such as Gaussian filtering, thresholding, and morphological operations. The study evaluated the sensitivity, specificity, and recall rates of radiologists with and without CAD. The results indicated a significant increase in cancer detection rates when using CAD, particularly in detecting small and early-stage cancers. The study indicated the potential of CAD systems in improving mammography inspection performance. However, in

another study [62], the authors examined the influence of CAD on screening mammography performance in a large-scale, multicenter study. The authors analyzed data from over 400,000 mammograms performed by expert radiologists to evaluate the sensitivity, specificity, positive predictive values, and recall rates before and after the implementation of CAD. The CAD system used in the study was designed to detect microcalcifications and masses in mammograms using image processing techniques and pattern recognition algorithms. The study revealed that the use of CAD did not significantly improve the overall diagnostic accuracy of mammography but increased the recall rate, suggesting the need for further optimization of CAD algorithms to reduce false-positive detections and improve clinical utility. In a more recent study, Lehman et al. [63] investigated the diagnostic accuracy of digital screening mammography with and without CAD using data from the Breast Cancer Surveillance Consortium. The study included mammograms from over 320,000 women and compared the cancer detection rates, recall rates, and positive predictive values between the two approaches. The CAD system employed in the study was based on the detection of suspicious patterns in mammograms with image processing. The results again demonstrated no significant improvement in diagnostic accuracy with the addition of CAD.

In the study [64], the authors proposed a Generative Adversarial Network (GAN)-based synthetic medical image augmentation method to improve the performance of CNN in liver lesion classification. The authors used GANs to generate synthetic liver lesion images based on real CT scan images, with the generator network learning to create realistic images by competing against a discriminator network that distinguishes between real and synthetic images. The study demonstrated that the inclusion of these synthetic images in the training process led to improved classification performance by 7% on real medical images, highlighting the potential of GAN-based augmentation in medical image analysis. Another study [65] presents a new an end-to-end lung cancer screening system using a three-dimensional deep learning model on low-dose chest CT scans. The proposed model, a 3D CNN, was designed to automatically detect lung nodules and classify them as malignant or benign. The authors used a large dataset of over 42,000 CT scans to train the model, with data augmentation techniques to address the class imbalance issue. In addition, they employed a sliding window approach and a two-stage classification process, which involved a candidate generation stage to identify potential nodules and a false positive reduction stage to refine the results. The study demonstrated that the proposed model detected lung cancer nodules with high sensitivity and specificity, outperforming radiologists in certain cases, suggesting the potential of deep learning in lung cancer screening and diagnostics. CNN based models are also utilized in the study [66] for the automatic recognition of the thyroid gland and tissues of the neck in ultrasound images. The framework consists of R-CNN based deep learning approach operating on 2D thyroid ultrasound data. The proposed method was experimented on a dataset of 17 patients, and the model has performed better than state-of-the-art. Another CNN model is developed in the study [67] for diagnosing meniscus tears using MRI. The authors presented two different models; one model to detect meniscal tears and another model for determining tear type. The model topology was based on AlexNet, and the study demonstrated the efficacy of the proposed CNN model in diagnosing meniscus tears with AUC score of

0.924. El-Dahshan et al. [68] provided a survey on computer-aided diagnosis of human brain tumors using MRI and proposed a new hybrid algorithm utilizing the feedback pulse-coupled neural networks, the discrete wavelet transform and principle component analysis. A feed-forward back-propagation neural network-based learner is employed to train a model by using 101 images of dataset. The classification performance was 99%, demonstrating the efficiency of the proposed hybrid method for brain tumor detection. In the study [69], a hybrid device is presented for thyroid nodule diagnosis. The proposed system integrates diffuse optics and ultrasound technologies to provide a non-invasive and radiation-free approach for detecting and characterizing thyroid nodules. The diffuse optics technique involves the use of near-infrared light to measure tissue absorption and scattering properties, while the ultrasound component provides high-resolution images of the thyroid gland and surrounding tissues. The combination of these two methods enables the system to obtain complementary information, enhancing the diagnostic accuracy and reducing unnecessary biopsies in thyroid nodule evaluation.

2.2. Computer-aided Diagnosis in Histopathology

Histopathology, a vital aspect of medical diagnostics, involves the microscopic examination of tissue samples to identify and analyze diseases, particularly cancer. Accurate and timely histopathological analysis is essential for determining the appropriate course of treatment, predicting patient outcomes, and monitoring disease progression [70]. However, manual analysis of histological slides can be labor-intensive, time-consuming, and prone to intra and inter-observer variability, which may lead to inconsistent and suboptimal diagnoses [71]. This highlights the importance of developing and incorporating advanced computational methods, such as computer-aided diagnosis systems, to augment and support the work of pathologists and ultimately improve patient care. The field of computer-aided diagnosis in histopathology has gained significant attention in recent years, as researchers strive to develop solid algorithms that can assist pathologists in the precise and efficient interpretation of histopathological images [72]. Machine learning and deep learning techniques have become the cornerstone of these developments, offering improved diagnostic accuracy, reduced subjectivity, and increased efficiency in the analysis of histological slides [73]. This sub-section aims to provide a comprehensive overview of the progress made in computer-aided diagnosis in histopathology, focusing on key methodologies, challenges, and future directions.

The study [74] proposes computer-aided diagnosis system for histopathological images of the endometrium, utilizing a combination of a VGG-16 based CNN and attention mechanisms. The methodology primarily consisted of two components: convolution layers generating high-level features from the input images, and a spatial and channel-wise attention mechanism that adaptively emphasized the most relevant regions and channels in the feature maps. The attention mechanisms enabled the model to focus on the most discriminative parts of the images, improving classification performance. The authors have compared the performance of the proposed CNN topology with the well-known topologies such as AlexNet, VGG-16, Inception-V3 and ResNet-50. The proposed

18

method has outperformed the existing models and demonstrated 76.91% of classification accuracy with 4 categories of endometrial tissue types. In another study [75], the authors introduced a cervical histopathology dataset for computer-aided diagnosis of precancerous lesions, comprising 1,839 whole-slide images (WSIs) of cervical tissues with different grades of cervical intraepithelial neoplasia (CIN). The authors also presented a two-step segmentation and classification framework that used a deep learning-based encoder-decoder architecture for segmentation of the epithelium, and a CNN for classification of the segmented regions into different CIN grades. The proposed method demonstrated high accuracy in segmenting the epithelium and classifying the CIN grades, paving the way for reliable and automated diagnosis of precancerous cervical lesions.

The researchers in [76] developed a computer-aided diagnosis system for efficient breast cancer detection from histopathological images by combining a fully convolutional network (FCN) and a bidirectional long short-term memory (Bi-LSTM) model. The FCN was employed to extract spatial features from the input images, while the Bi-LSTM was utilized to capture the contextual information between different regions. The proposed hybrid model effectively combined the advantages of both FCN and Bi-LSTM, resulting in improved performance for breast cancer detection in comparison to other state-of-the-art methods. Similarly, the authors in [77] presented a guided soft attention network (GuSA) for classification of breast cancer histopathology images. The GuSA integrated a soft attention mechanism into a CNN, allowing the model to focus on the most relevant regions in the input images for accurate classification. Additionally, the neural network training can be interpreted with the help of the activation maps generated by GuSA. Thus, the method is also useful to make neural network decisions transparent as the activations can be monitored by the human experts. Another CNN implementation is presented in the study [78]. The authors proposed a large-scale tissue histopathology image analysis framework that employed deep convolutional activation features for classification, segmentation, and visualization. The authors have divided the input images into 336x336 image patches and utilized AlexNet architecture with transfer learning to train a strong classifier. At the final step of classification pipeline, an SVM model is employed for classifying the features generated by CNN. The proposed model has tested on both brain and colon tissue samples stained with H&E. The authors have concluded that the automatically generated features by CNN are more discriminative than the manually extracted features when feeding to the SVM classifier. The framework demonstrated promising results; 97.5% accuracy for classification and 84% accuracy for segmentation tasks.

A weakly supervised approach for histopathology cancer image segmentation and classification is proposed in [79]. The authors employed a multiple clustered instance learning (MCIL) approach in order to streamline image-level or patch-level classification and segmentation tasks. The researchers effectively identify cancerous regions in whole slide images without requiring strong annotations, such as detailed pixel-level ground truth. The proposed method combines the benefits of multiple-instance learning and clustered learning. A similar approach proposed in the study [80]. The researchers presented a weakly supervised deep learning method for the classification and localization

of colorectal cancer in histopathology images. The authors proposed a three-stage learning framework that first utilized a CNN-based classifier with only global image-level labels to perform classification on the lower resolution image. In the second stage, a patch-wise classification model was trained using another CNN model, doing cancer vs non-cancer binary classification. Finally, the probabilities of all patch images are combined onto the whole spatial plane and the heatmap visualization depicting cancerous regions is constructed. The proposed method posed 94.6% accuracy on classification and 92% accuracy on colorectal cancer localization tasks.

The study [81] introduced a comprehensive multiple instance learning (MIL) approach with center embeddings for histopathology classification. The proposed method utilized a deep learning-based MIL framework, where bag feature learning is improved by center loss. This approach also enabled the authors to decrease ambiguity in instance labels. The center loss was found out effective to reduce intra-class distances. Thus, more instance level predictions are achieved, effectively reducing false positives. The proposed method is tested on two distinct WSI datasets and the classification accuracy has reached to 92.31%. An automated brain and breast tissue classification system using deep learning techniques is presented in [82]. As the main classifier, a modified version of InceptionV3 CNN is employed. The proposed method is tested on two datasets; H&E stained brain tissue dataset with 3 classes and H&E stained breast tissue dataset with 2 classes. In order to prevent overfitting on the small datasets, the authors have implemented a dropout layer in CNN. The authors have reported that applying transfer learning form one tissue dataset (brain) to other tissue dataset (breast) boosts the classification accuracy significantly. The study [83] proposed deep learning models for the histopathological classification of gastric and colonic epithelial tumors. The authors have devised two distinct models by using CNNs and recurrent neural networks (RNNs) with H&E stained stomach and colon tissue samples. Firstly, a CNN model with a slimmer version of inception-V3 network is trained. Then, an RNN model is trained by using the features generated by the previous CNN network. The feature vectors from CNN are revealed by removing the last fully-connected layer from the inception-V3 network. There was no clear difference of accuracy between CNN and RNN based classifications, and the models were able to reach up to 0.99 AUC scores.

The researchers in [84] explored the performance of CNNs for histopathology image classification by comparing the performance of training a network from scratch versus utilizing pre-trained networks. They employed VGG16 and Inception-V3 networks to perform transfer learning on a dataset of histopathology images. The authors highlighted the benefits of transfer learning, demonstrating that pre-trained networks significantly outperformed those trained from scratch. They also found that the fine-tuning of Inception-V3 network further enhanced classification performance whereas the fine-tuning of VGG16 had no significant effect on the accuracy. Moreover, in the study [85], transfer learning was employed for automated classification of histopathology images. The author utilized a dataset of grayscale and color histopathological images from 24 whole-slide tissue images, using pre-trained CNN architectures including ResNet50 and DenseNet-161. The results demonstrated that the ResNet50 model achieved the highest

classification accuracy among the tested architectures. The study revealed the effectiveness of transfer learning in histopathology image classification tasks and emphasized the importance of selecting the appropriate pre-trained CNN model.

The authors in [86] presented a deep learning-based method for classifying H&E stained breast tissue images from biopsies. Utilizing the Inception-v3 CNN and images from the BACH challenge, the method classifies the images into four categories: normal tissue, benign lesion, in situ carcinoma, and invasive carcinoma. By extracting patches based on nuclear density and discarding non-epithelial regions, the researchers effectively focus on the most informative regions of the images, as visual signs of tumors are most apparent in epithelium. The classification of each patch with high nuclear density leads to the determination of the entire image's class through majority voting. This approach achieved an average accuracy of 85% across the four classes and 93% for non-cancer (normal or benign) versus malignant (in situ or invasive carcinoma). In the study [87], the authors investigated the use of deep learning for classification in lung cancer histopathology images with three classes Adenocarcinoma (LUAD) and squamous cell carcinoma (LUSC) and normal samples. The researchers implemented InceptionV3 network, and the whole-slide input images are divided into smaller patches as the pre-processing step. The model's performance was comparable to experienced pathologists, with an average AUC of 0.97, and was validated on independent datasets of frozen tissues, formalin-fixed paraffin-embedded tissues, and biopsies. Additionally, the network was trained to predict the ten most commonly mutated genes in LUAD, and successfully predicted six of them with AUCs ranging from 0.733 to 0.856. The model was further tested on various cohorts to identify its limitations, and its performance was evaluated on region-of-interest (ROI) selections made by a pathologist. In the paper [88], the researchers developed a method for the classification of colorectal tissues using ensemble deep learning methods with two neural network architectures; VGG16 and CapsNet. The researchers utilized adaptive gamma correction as a part of color enhancement stage and test data augmentation to improve performance and reduce uncertainty in predictions. The network hyperparameters were determined by applying random search. In the context of ensemble learning, the authors deployed 5 flows composed of different combinations of learners such as independent, snapshot ensemble and soft voting ensemble learning options. They achieved an accuracy of 93-98% for all tissue classes, including tumor, normal, and stroma/others.

2.3. Hyperspectral Imaging in Medical Applications

HSI imaging technique combines the principles of spectroscopy and imaging to obtain spatial and spectral information from a sample, generating a three-dimensional data cube that reveals a comprehensive view of the examined area [89]. The use of hyperspectral imaging in medical applications offers a non-invasive and non-ionizing approach for the visualization, diagnosis, and monitoring of various diseases and conditions. Recent developments in hyperspectral imaging systems and data processing techniques have expanded the range of potential applications in medicine, including in dermatology,

ophthalmology, oncology, and surgical guidance, among others [90]. Various studies have demonstrated the potential of HSI in detecting and characterizing tissue abnormalities, differentiating between healthy and diseased tissue, and monitoring treatment response [91]. As a result, hyperspectral imaging has emerged as a promising tool for improving patient care and reducing healthcare costs. The integration of machine learning and artificial intelligence techniques with hyperspectral imaging has further propelled the technology's potential in medical applications, paving the way for more efficient and accurate diagnostic and prognostic tools [92]. Potential benefits of hyperspectral imaging for medical applications include enhanced diagnostic accuracy, reduced need for invasive procedures, and improved patient outcomes. By providing detailed information on the biochemical composition of tissues, HSI can help healthcare professionals differentiate between healthy and diseased tissue, enabling more accurate diagnoses and targeted treatment plans. Furthermore, the non-invasive nature of HSI minimizes patient discomfort and reduces the risk of complications associated with invasive diagnostic procedures [93]. In addition, HSI can facilitate real-time intraoperative guidance for surgeons, improving surgical precision and reducing the likelihood of postoperative complications [94]. Overall, the use of hyperspectral imaging in medical applications has the potential to improve disease detection, treatment planning, and patient care.

In the study presented in [95], HSI was applied to forensic medicine for the first time, aiming to demonstrate the potential of this technology in postmortem examinations. HSI offers potential benefits to forensic investigations by providing more detailed and accurate information about samples. The study highlights how HSI methods can enhance the analysis of physical evidence, improve the identification of trace materials, and aid in the detection of concealed or altered evidence. In the paper, the authors provide a comparative summary of existing HSI systems for applying to forensic medicine domain. Then, the authors in paper [96] developed a hyperspectral imaging system specifically designed for in vivo optical diagnostics. The system utilized an Acousto-optic Tunable Filter spectrometer coupled with a CCD camera for acquiring hyperspectral data in the visible spectral range (573-695 nm). The authors have worked on mice and chicken organ samples during the experimentations. The spectral signatures of various samples have been extracted and illustrated in order to display the potential of HSI on capturing wide range of spectral data. The researchers argued that HSI coupled with fiber-optic equipment could provide a high potential for surgery-assist applications like endoscopy in real-time. The researchers in [97] investigated the use of medical hyperspectral imaging to facilitate residual tumor identification during surgery. The authors employed a programmable liquid crystal tunable filter with a spectral range of 450-700 nm, and artificial breast cancer rat model was adapted for the experimentation. The imaging system was used to collect hyperspectral cubes from various tissue types including connective tissues, oxygenated tissues, muscle, tumor and blood. The manual classifications were verified by using histopathological analysis of the samples. The researchers provided the comparison of spectral signatures from distinct tissue types. The proposed application could be used for guidance to surgeons during cancer resection procedures and potentially reducing the risk of tumor recurrence. By using a larger wavelength interval, Wehner et al. [98] presented a near-infrared (NIR) digital light processing (DLP) hyperspectral imaging system

22

specifically designed for medical applications. The system used a DLP-based spectral engine capable of operating in the NIR range (760-1600 nm), offering improved penetration depth in biological tissues. The proposed system can be used to measure the amount of water in the sample, thus, with the help of this technology, various applications are possible in surgery guidance field.

In the work presented in [99], tongue tumor detection was explored using medical hyperspectral images. The authors employed an AOTF imaging system with a spectral range of 600-1000 nm. The methodology involved the development of an automatic tongue tumor detection algorithm, which included preprocessing and a classification with dictionary based sparse representations. The proposed method does a pixel-wise classification, and the comparative results with SVM and Relevance Vector Machine RVM indicated the superiority of the method. The algorithm effectively differentiated between healthy tongue tissue and tumors with the accuracy of 96.5%. Another study in [100], the authors explored the potential of HSI as a new modality for Age-related-macular-degeneration (AMD) disease. Without applying extra biochemicals to the patients, the hyperspectral system can detect stage of degeneration and oxidative status of proteins from retinal ganglion cells. Since the known proteins from eyes has pre-defined absorbance within the spectral range of 400-700 nm. The methodology involved acquiring hyperspectral data from various ocular tissues, including the retina, iris, and cornea, and analyzing the data to differentiate between healthy and pathological conditions. HSI was shown for the early detection of eye diseases, such as AMD.

The work elaborated in papers [101-102] investigated the use of hyperspectral imaging and spectral-spatial classification for prostate cancer detection. The authors used a hyperspectral imaging system operating in the visible and near-infrared spectral range (500-950 nm). Least squares support vector machines (LS-SVMs) are employed for classifying hyperspectral data captured from a rat tumor model with in vivo imaging. The proposed tumor detection algorithm posed 92.8% sensitivity and 96.9% specificity. A method for gastric cancer detection using hyperspectral imaging is introduced in [103]. The authors used an HSI system with a spectral range of 400-800 nm and 72 spectral bands images. The methodology involved acquiring hyperspectral data from gastric mucosa during endoscopic procedures and developing an algorithm to detect gastric cancer based on the extracted spectral features. The proposed algorithm was based on a supervised method to identify ideal threshold parameters for tumor versus healthy discrimination and it has reached 85.6% classification accuracy. Gerstner et al. [104] investigated the application of hyperspectral imaging for the examination of mucosal surfaces in patients. The authors employed an HSI system with a spectral range of 590-680 nm, and the sample images were collected by microlaryngoscopy. The captured hyperspectral cubes were analyzed with the help of PCA in order to determine spectral profiles of the mucosal surfaces of humans. The study demonstrated the potential of HSI for in vivo evaluation of mucosal surfaces and early detection of pathological conditions, such as lesions and inflammation.

Calin et al. [105] conducted a preliminary study on the characterization of burns using hyperspectral imaging. The authors employed an HSI system with a spectral range of 400-1000 nm and 346 spectral bands. The methodology involved acquiring hyperspectral data from human skin with various degrees of burns, followed by the extraction of spectral signatures and the use of linear spectral unmixing to classify the burn severity. The endmembers are detected using PCA, then, spectral unmixing is performed to obtain abundance maps. The authors reports that the HSI data analysis is correlated with the clinical examination, and HSI is a promising tool for burn wound grading. Similarly, Wahabzada et al. [106] explored the use of hyperspectral imaging for monitoring wound healing in a 3D wound model. Unlike the traditional techniques in wound healing research, hyperspectral imaging is invasive and it does not require tissue biopsy. The methodology involved acquiring hyperspectral data from in vitro 3D wound models at various stages of healing and applying unsupervised clustering algorithms, such as k-means and extreme hierarchical clustering. The study showed that HSI could effectively monitor the wound healing process, providing valuable insights into tissue regeneration and the effectiveness of various treatment strategies.

The researchers in [107] developed a super-pixel-based spectral classification method for the detection of head and neck cancer using hyperspectral imaging. The methodology involved acquiring hyperspectral data from head and neck tissue samples from 11 female mice. The collected data was firstly decomposed into first two principle components by using PCA. Then the principle components were segmented into superpixels with the help of simple linear iterative clustering (SLIC) method, which is a modified version of k-means clustering. Finally, the super-pixel regions were classified using SVM, and the tumor regions were identified. The super-pixel-based method aims to reduce computational complexity and improve the classification performance by considering the spatial context of the data. The super pixel approach in classification has resulted in 93% sensitivity and 0.85 specificity, offering a robust segmentation model. A similar super-pixel approach is also reported in the study [108]. The hyperspectral imaging for monitoring neoplastic progression in a mouse model of oral carcinogenesis was investigated. The authors collected hyperspectral data from in vivo mouse models with various stages of oral cancer and using various classifiers to differentiate between healthy and cancerous tissues. The study experimented six different classifiers on the data including random forest, linear discriminate analysis, SVM, naïve bayes, k-nearest neightnors and decision trees, and SVM approach provided the best classification accuracy by 88.9%.

Li et al. [109] proposed an automation method for red blood cell (RBC) counting using microscopic hyperspectral imaging technology. The authors employed a hyperspectral microscope covering the spectral range of 550-1000 nm. The methodology involved acquiring hyperspectral data from blood smears, performing image preprocessing, and employing morphological analysis to segment and count RBCs. The image processing pipeline included PCA for dimensionality reduction, a combination of k-means and spectral angle mapper (SAM) algorithms for segmentation tasks. Similarly, Huang et al. [110] developed a blood cell classification method based on hyperspectral imaging using

modulated Gabor features and CNN. Different from conventional CNN, the kernels were combined with multi-scale and orientation Gabor filters by calculating dot product of them. The classification performance of the proposed method compared with well-known methods such as SVM, VGG16 and CNN without Gabor filters. The proposed method outperformed all alternatives by reaching 97.65% classification accuracy. Furthermore, in the study [111], the researchers proposed a 3D attention network for the classification of white blood cells (WBC) from microscopy hyperspectral images. The authors employed a hyperspectral microscope covering the spectral range of 450-780 nm with a spatial resolution of 0.6 µm and 51 channels images. 3D convolutional neural networks-based model was adapted in the study in order to benefit from both spectral and spatial features at the same convolution operation. 3D convolutions improved the classification performance significantly by reaching 96.72% accuracy while traditional methods provided accuracy around 90%. The study highlights the value of 3D-CNN, combined with microscopic hyperspectral imaging as a promising tool for WBC classification. On the other hand, Verebes et al. presented hyperspectral enhanced dark field microscopy as a method for imaging blood cells [112]. The study integrated dark field microscopy with HSI to acquire high-resolution images of red and white blood cells, revealing their morphological and spectral features. The authors demonstrated that this technique could be used to differentiate between different types of blood cells and identify cellular abnormalities, suggesting its potential utility in the field of clinical diagnostics and hematology research.

By using brain tissue for research, Fabelo et al. [113] presented an in-vivo hyperspectral human brain image database for brain cancer detection, the first of its kind. The study aimed to develop a surgery guidance tool that can be used in real-time. The authors employed a hyperspectral camera system with a spectral range of 400-1000 nm. In total, the publicly published dataset contained 36 hyperspectral images collected from 22 distinct patients together with their corresponding labels. The study provided a valuable resource for researchers working on hyperspectral imaging-based brain cancer detection methods. The authors have also provided the classification results with the available machine learning and deep learning methods. For tumor detection task from tongue tissues, the researchers in [114] proposed a dual stream network in hyperspectral images. The dual stream network combined spatial and spectral information for improved tumor detection and classification. The first stream had the responsibility of classifying pixel signatures, and the second stream was classifying the first three principle components in order to embed the textural information into the final model. Although the dataset was very small with 7 patients, the resultant classifier was able to achieve 79.8% accuracy. In another research [115], liver cancer grade differentiation is investigated using multiphoton microscopy images. The study adapts the VGG-16 CNN topology to train a classifier on a dataset containing three liver cancer disease grades, yielding over 90% differentiation accuracy and demonstrating the efficacy of deep learning approaches with multiphoton fluorescence imagery.

2.4. Integration of Hyperspectral Imaging and Microscopy in Histopathology

The integration of HSI and microscopy in histopathology has emerged as a promising approach for fostering disease diagnosis, particularly in the field of cancer detection. Histopathological examination of tissue samples is an essential tool for accurate cancer diagnosis, prognosis, and treatment planning [116]. The traditional method of histopathological examination involves manual inspection of stained tissue sections under a microscope, which is time-consuming, labor-intensive, and subjective to inter- and intra-observer variability [117]. The development of advanced imaging techniques such as hyperspectral imaging, which combines spatial and spectral information, has the potential to improve histopathology literature by providing a more objective and accurate assessment of tissue samples. Hyperspectral imaging is a handy imaging technique that acquires a continuous spectrum of light reflected or transmitted from a sample at each spatial location [118]. This provides rich spatial and spectral information, allowing for a more detailed analysis of the tissue samples. In recent years, the integration of HSI with microscopy has gained attention as a powerful tool for histopathology, offering enhanced diagnostic capabilities and the ability to visualize subtle differences in tissue composition, which may be missed in conventional histopathological examination [119].

Several studies have demonstrated the potential of hyperspectral microscopy in histopathology for various applications, such as cancer detection, tissue classification, and biomarker identification [120]. By combining the high spatial resolution of microscopy with the spectral information provided by HSI, researchers can gain a deeper understanding of the tissue samples and identify disease-specific spectral signatures. This has led to the development of more accurate and objective diagnostic algorithms that can potentially improve patient outcomes and streamline the histopathological examination process. The advantages of combining HSI and microscopy are numerous. One of the primary benefits is the enhanced diagnostic accuracy, as the rich spatial and spectral information provided by the integrated system allows for better differentiation between healthy and diseased tissue, and even between different types of diseases. This can result in a more accurate and timely diagnosis, ultimately improving patient care. Additionally, the integration of HSI and microscopy can lead to a more objective and standardized histopathological examination process, reducing inter- and intra-observer variability and minimizing the potential for human error [121]. For example, the researchers developed a hyperspectral imaging microscopy technique for identifying and quantifying fluorescently-labeled cells in highly auto-fluorescent tissue samples [122]. The authors utilized an AOTF hyperspectral imaging system to acquire images of tissue samples with multiple fluorescent labels. They then applied spectral flat-field correction and linear unmixing algorithms to separate the signals from different fluorophores, as well as the autofluorescence background. This approach allowed for the accurate identification and quantification of fluorescently-labeled components such as nuclei, green fluorescent protein, autofluorescence and background. The authors reported that hyperspectral data improves the detection accuracy when the tissues are highly auto-fluorescent.

One study [123] employs Atrous Spatial Pyramid Pooling (ASPP) blocks to extract multi-scale texture features from H&E stained HCC histopathology images utilizing RGB data. By incorporating ASPP blocks after each max-pool layer, a multi-scale sample space is generated, enabling the effective use of texture features in images by deep neural networks. This approach achieved a four-category classification accuracy of 90.93% for HCC images. The study [124] provides an early example of integrating hyperspectral imaging with a light microscope, proposing a reference hardware system to capture hyperspectral cubes from microscopy tissue slides. The system combines an imaging spectrograph with an epi-fluorescence microscope and utilizes various wavelength light sources for sample illumination. A motorized mover is employed to move the sample slide during data capture, with a CCD camera capturing the reflected light and custom software developed for visualizing and storing hyperspectral data. Another study [125] proposes a hyperspectral image analysis-based approach for identifying ALK-positive and ALK-negative tumors. Sixty-channel hyperspectral data from lung cancer tissues is captured using an AOTF-based hyperspectral imaging system. An SVM segmentation algorithm is applied to classify lung tissue images into cell nucleus, cytoplasm, and blank area. Segmentation accuracy is calculated using manual ground truth data provided by a lung cancer expert and evaluated for each class to determine a treatment prescription targeting ALK-positive and ALK-negative tumor variation. In a related study [126], an AOTF-powered hyperspectral imaging system is used to gather hyperspectral data from 30 channels of bile duct tissue samples. Deep CNNs, specifically Inception-V3 [127] and Restnet50 [128], are employed to develop a prediction model. A spectral interval convolution method is proposed to adapt hyperspectral data for deep learning architectures. CNN experiments are conducted using image patches as input, and a random forest-based approach combines image patch predictions from the same scene to provide scene-level predictions. The authors report tumor detection accuracies of 0.93 with hyperspectral data and 0.92 with RGB data.

2.5. Convolutional Neural Networks

CNNs have emerged as a powerful and versatile tool for various computer vision and image processing tasks, such as object recognition, segmentation, and classification [129]. The effectiveness of CNNs can be attributed to their ability to exploit the spatial structure and local patterns within an image, owing to their unique architectural characteristics, such as local connectivity, shared weights, and pooling layers [130]. While the foundational work of LeCun et al. [131] pioneered the today's CNN framework, significant advancements in computational resources and algorithmic refinements have led to the development of deeper and more intricate architectures. The advancements in CNN models have resulted in improved performance on a wide range of applications [132]. This introduction sub-section aims to provide a comprehensive overview of the principles and evolution of CNNs, emphasizing their mathematical foundations, architectural design choices, and practical applications in the context of this thesis.

CNNs were inspired by biological processes, specifically the organization and functioning of the mammalian visual cortex. In the human visual system, neurons are arranged in a hierarchical manner, with the early stages of visual processing, such as the primary visual cortex (V1), being sensitive to simple features like edges and orientations, while the higher stages (e.g., V2, V4, and IT) respond to more complex and abstract features, such as shapes, textures, and objects [133]. This hierarchical feature extraction process is mirrored in convolutional networks through their successive layers of convolutions and nonlinear activations. Additionally, the concept of local receptive fields in the visual cortex, where neurons respond to stimuli only within a specific region of the visual field [134], has been incorporated into convolutional networks in the form of local connectivity. Furthermore, the weight-sharing mechanism employed in convolutional networks, wherein neurons within a convolutional layer share the same set of weights and biases, reflects the observation that certain neurons in the visual cortex exhibit similar response properties [135]. That is, convolutional networks are a simplified and abstracted representation of the complex processes underlying biological perception. The theory of CNNs has been heavily influenced by these biological insights and led to their high performance in various computer vision and image processing tasks. In fact, CNNs exhibit a notable advantage over other image classification algorithms due to their minimal reliance on pre-processing techniques. Traditional image classification methods often require extensive pre-processing, including feature engineering, normalization, and dimensionality reduction, to achieve optimal performance. However, CNNs possess the built-in ability to learn optimal filters through automated training processes. This eliminates the need for manual feature engineering and allows the network to adaptively capture useful patterns and structures present in the input data. As a result, CNNs can efficiently discover and exploit hierarchical features that emerge from successive layers of convolutions and nonlinear activations.

Over the years, several popular CNN architectures have been developed, each demonstrating state-of-the-art performance in various computer vision tasks. One of the earliest successful CNN architectures, LeNet-5, was developed by Yann LeCun and his team in the late 1990s. Primarily designed for handwritten digit recognition, it utilized a simple architecture with alternating convolutional and pooling layers, followed by fully connected layers [136]. In 2012, the introduction of AlexNet marked a significant improvement for the field. With deeper layers, ReLU activations, and dropout for regularization, it achieved a notable performance improvement over traditional computer vision methods [137]. On the other hand, VGGNet was proposed in 2014. Known for its deep architecture with up to 19 layers, it demonstrated the effectiveness of using small 3x3 convolutional filters and multiple layers for better feature extraction [138]. Following this, researchers at Google introduced the Inception architecture, also known as GoogLeNet, in 2014. This novel architecture employed inception modules, which allowed for more efficient use of computational resources by incorporating parallel convolutional layers with varying filter sizes in a single module [139]. Subsequently, in 2015, the Residual Network (ResNet) was developed by Kaiming He and his team at Microsoft Research. By introducing the concept of residual connections to tackle the problem of vanishing gradients and training difficulties in very deep networks, they enabled the

training of networks with over a hundred layers [140]. In 2016, DenseNet was introduced, connecting each layer to every other layer in a feed-forward fashion. This design promoted feature reuse and reduced the number of parameters [141]. A year later, Google researchers developed MobileNet, designed specifically for resource-constrained environments such as mobile devices. It employed depth-wise separable convolutions to reduce the number of parameters and computational cost while maintaining competitive performance [142]. Finally, in 2019, EfficientNet was proposed, introducing a systematic approach for scaling up CNNs called compound scaling. By adjusting the depth, width, and resolution of the network simultaneously, it achieved state-of-the-art performance with fewer parameters and lower computational cost compared to other models [143]. These popular CNN architectures have served as a foundation for many other variations and have been widely used in various computer vision tasks, such as image classification, object detection, and semantic segmentation.

As seen in Figure 7, a CNN model is composed of an input layer, hidden layers, and an output layer, each serving a distinct role in the processing and classification of input data. The input layer is responsible for receiving and processing raw data, such as images or time-series signals, and passing it to the subsequent hidden layers. Hidden layers, which consist of multiple convolutional, pooling, and fully connected layers, perform the primary computational tasks within the network. Convolutional layers are designed to learn hierarchical features from the input data through the application of filters or kernels. Pooling layers serve to down-sample and introduce invariance to minor changes in the data. Fully connected layers, often found near the end of the network, integrate the high-level features learned in previous layers and map them to the final output space. Finally, the output layer, which usually consists of a SoftMax activation function, generates the network's predictions by assigning probabilities to each class or label. At the next sub-sections, we go through the building blocks of CNN models.
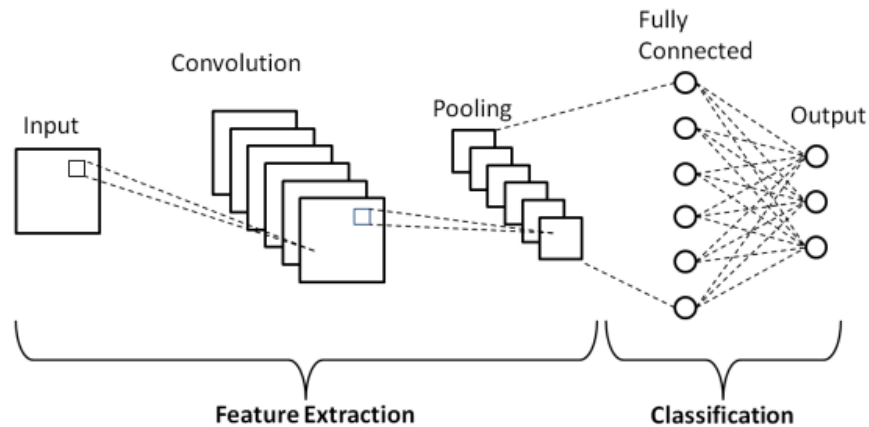
Figure 7: A sample CNN topology sketch.

2.1.1. Convolutional Layers

In the convolutional layer, a collection of filters with learnable parameters is employed to analyze the input data. These filters have dimensions smaller than the input volume and are designed to create activation maps consisting of neurons. The process involves sliding the filter across the input's width and height, calculating dot products between the filter and input at every spatial location. The resulting output volume is formed by stacking these activation maps along the depth axis. As the filters have smaller dimensions than the input, neurons in the activation maps are connected only to a limited local region of the input volume, resulting in a small receptive field for each neuron, equivalent to the filter size. The convolutional layer's local connectivity enables the network to learn filters that optimally respond to specific input regions, thus, it leverages the input's spatial local correlation. Moreover, since the activation map is produced by convolving the filter with the input, filter parameters are shared across all local positions, reducing the number of parameters. As a result, it leads to efficient expression, learning, and improved generalization capabilities.

A. 2D Convolutions

2D convolution takes place at the convolutional layers to derive features from local regions on feature maps from the preceding layer. Following this, an additive bias is incorporated, and the output is processed through a sigmoid function. In a formal manner, 2D convolution operator is defined in Equation 1 as follows;

$$v_{ij}^{xy} = f\left(b_{ij} + \sum_{m} \sum_{p=0}^{P_i-1} \sum_{q=0}^{Q_i-1} w_{ijm}^{pq} v_{(i-1)m}^{(x+p)(y+q)} \right) \tag{1}$$

where v_{ij}^{xy} refers to the value at position (x, y) in the jth feature map of the ith layer. The index value m denotes the input feature maps from the $(i-1)^{th}$ layer that are connected to the current feature map, P_i and Q_i represent the height and width of the kernel, respectively. The kernel value at position pq for the mth feature map in the previous layer is denoted by w_{ijm}^{pq}, and b represents the bias. The activation function is represented by $f(.)$. By stacking multiple layers of convolution and pooling layers together, custom CNN topologies can be built. The kernel weights and biases are learnt during the supervised training stage.

B. 3D Convolutions

In this study, we utilized the 3D convolution operation proposed in a previous study [144].

$$v_{ij}^{xyz} = f\left(b_{ij} + \sum_{m}\sum_{p=0}^{P_i-1}\sum_{q=0}^{Q_i-1}\sum_{r=0}^{R_i-1} w_{ijm}^{pqr} v_{(i-1)m}^{(x+p)(y+q)(z+r)}\right) \tag{2}$$

The mathematical formulation of 3D kernels can be expressed using Equation 2, where v_{ij}^{xyz} refers to the value at position (x, y, z) in the jth feature map of the ith layer. The index value m denotes the input feature maps from the $(i-1)^{th}$ layer that are connected to the current feature map, and P_i, Q_i, and R_i represent the height, width, and depth of the kernel, respectively. The kernel value at position pqr for the mth feature map in the previous layer is denoted by w_{ijm}^{pqr}, and b represents the bias. The activation function is represented by $f(.)$.

The key distinction between traditional 2D-CNN and 3D-CNN lies in the convolution operation performed on the convolution layers. In 2D-CNN classifiers, the kernel slides along two dimensions (x and y) on the data, while in 3D-CNN classifiers, the kernel slides along three dimensions (x, y, and z) on the data. The use of 3D-shaped kernels in convolutions enables the description of features in both spatial and spectral directions. This allows the final classification model to embed spectral information, in addition to spatial features such as texture and shape attributes.

2.1.2. Pooling Layers

Pooling layers are used to reduce the spatial dimensions of the feature maps, thereby reducing the computational complexity of the network. Pooling layers are typically placed between successive convolutional layers in a CNN architecture. There are two common types of pooling operations: max pooling and average pooling. Max pooling calculates the maximum value within a specified region of the input feature map, whereas average pooling computes the average value within the same region. The regions are determined by the pooling layer's window size and stride, which define the dimensions of the pooling operation and the step size by which the window is moved across the input feature map, respectively.

Max pooling operation can be formulated seen in Equation 3, as follows;

$$y_{kij} = max_{(p,q)\in R_{ij}}(x_{kpq}) \tag{3}$$

where y_{kij} is the output value related to k^{th} feature map, x_{kpq} is the input value at coordinate (p, q), which is located inside pooling window R_{ij} representing local neighborhood around the coordinates (i, j). Similarly, average pooling can be defined as seen in Equation 4.

$$y_{kij} = \frac{1}{|R_{ij}|} \sum_{(p,q)\in R_{ij}} x_{kpq} \tag{4}$$

where the term, $|R_{ij}|$, represents the size of the pooling region R_{ij}.

2.1.3. Fully Connected Layers

Fully connected layers, also referred to as dense layers, act as the final phase of a CNN. While the convolutional and pooling layers are employed for extracting features and reducing spatial dimensions, the fully connected layers function as the network's final stage. These layers merge the high-level features obtained from preceding layers and execute classification or regression tasks. Within a fully connected layer, every neuron is connected to all neurons in the layer before it, enabling the integration of information from all input neurons. Figure 8 illustrates the connections between input and output neurons in case of fully connected layers. This extensive connectivity allows the layer to learn intricate patterns and relationships among the high-level features derived from the convolutional and pooling layers. In the majority of CNN architectures, one or more fully connected layers are positioned after the last convolutional or pooling layer, succeeded by an output layer that aligns with the number of classes or target values specific to the problem domain.
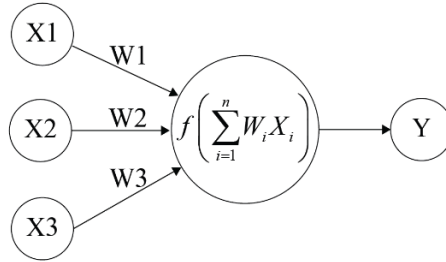
Figure 8: Sketch showing fully connected neuron (Y) interacting with input neurons (X_i).

Formally, the fully connected layers can be formulated as depicted in Equation 5.

$$y_j(x) = f\left(\sum_{i=1}^{n} w_i x_i + b_{j0}\right) \tag{5}$$

where y_i is the output value for the given input neurons, x. i is the index of input neuron, j is the index of the current fully connected neuron.

2.1.4. Activation Functions

Activation functions serve as crucial elements in the structure and operation of Convolutional Neural Networks (CNNs), along with other neural network variants. Their main objective is to incorporate non-linearity within the network, enabling it to learn and

represent intricate, non-linear associations between input features and corresponding outputs. These functions are utilized on the outputs of neurons across different layers in the network, encompassing convolutional, pooling, and fully connected layers. The introduction of non-linearity through activation functions allows the network to capture complex patterns and adapt to a wide range of problem domains. Without non-linear activation functions, the network would essentially become a linear function, limiting its capacity to model complex data and relationships. By applying activation functions to the outputs of neurons, the network gains the ability to learn hierarchical representations and generalize well to unseen data.

Among the commonly employed activation functions, the sigmoid function, as formulated at Equation 6, is a continuous and differentiable mapping from input values to the range of 0 to 1, which has been widely used in early neural network models [145]. Nonetheless, the vanishing gradient problem, characterized by diminishing gradient magnitudes during backpropagation, can impede learning in deep networks utilizing sigmoid activation [146]. The hyperbolic tangent (tanh) function, as formulated at Equation 7, which maps input values to a range between -1 and 1, shares similar characteristics with the sigmoid function but also encounters the vanishing gradient issue [147]. The Rectified Linear Unit (ReLU) activation, as formulated at Equation 8, has gained prominence due to its computational efficiency, simplicity, and effectiveness in mitigating the vanishing gradient problem [148]. In multi-class classification problems, the softmax function, as formulated at Equation 9, is typically employed in the output layer, normalizing output values into a probability distribution [149].

Sigmoid
$$f(x) = \frac{1}{1 + e^{-x}} \tag{6}$$

Tanh
$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{7}$$

ReLU
$$f(x) = \max(0, x) \tag{8}$$

SoftMax
$$f(x_i) = \frac{e^{x_i}}{\sum_{j=1}^{K} e^{xj}} \tag{9}$$

2.1.5. Cost Functions

Cost functions, often referred to as loss functions or objective functions, are essential components in the optimization of CNNs and other machine learning algorithms [130]. They quantify the discrepancy between the predicted outputs and the actual target values, serving as an indicator of the model's performance. The primary goal during training is to minimize the loss function by updating the model's parameters, thereby enhancing the network's capacity to generalize and make precise predictions for unseen data [150]. In CNNs, the choice of loss functions depends on the specific task and problem domain. For

classification tasks, cross-entropy loss is among the most frequently employed loss functions [151]. Cross-entropy loss, as formulated in Equation 10, sometimes referred to as log loss, evaluates the divergence between the predicted probability distribution and the true distribution of class labels. This loss function is particularly well-suited for multi-class classification problems because it promotes high probabilities for the correct class labels and low probabilities for incorrect labels. The utilization of cross-entropy loss in CNNs encourages the model to accurately differentiate between various classes while maintaining a high degree of confidence in its predictions. Nevertheless, the commonly used cross entropy (CE) function is not as effective in handling imbalanced datasets.

$$CE(p_t) = -\log(p_t) \tag{10}$$

where p_t is given by

$$p_t = \begin{cases} p & y = 1 \\ 1 - p & y = -1 \end{cases} \tag{11}$$

where $y \in \{-1,1\}$ is the ground truth class and $p \in [0,1]$ is the classifier's output probability value for the class $y = 1$.

When the class imbalance is extreme, the cross-entropy-based models can assign too much weight to the well-classified examples, which can cause the minority class to be neglected. To address this issue, the balanced cross-entropy (BCE) function, as shown in Equation 12, is often used to handle class imbalance in traditional CE functions.

$$CE(p_t) = -a_t \log(p_t) \tag{12}$$

where a_t is a weighting factor hyperparameter and defined as;

$$a_t = \begin{cases} a & if\ y = 1 \\ 1 - a & otherwise \end{cases} \tag{13}$$

where $a \in [0,1]$. The imbalanced nature of our dataset can cause the loss contribution of well-classified examples to dominate the minority class in cross-entropy-based models, particularly in cases of extreme class imbalance. Although the balanced cross entropy (BCE) function can help address this issue, it cannot balance the loss between easy and hard examples. Fortunately, focal loss, as formulated in Equation 14, is derived from the cross-entropy loss function by introducing a modulating factor $-a(1 - p_t)^\gamma$.

$$FL(p_t) = -a(1 - p_t)^\gamma \log(p_t) \tag{14}$$

where $\gamma \geq 0$ is the focusing parameter. The parameter a serves the purpose of emphasizing the training on the minority class instead of treating all classes equally

important. On the other hand, the parameter γ is known as the focusing parameter that prioritizes the samples resulting in large errors, also known as hard examples [47].

2.1.6. Batch Normalization

Batch normalization is a technique introduced in 2015 to address the internal covariate shift problem and improve the training of CNNs and other deep learning models [152]. Internal covariate shift refers to the change in the distribution of inputs to a layer during training, resulting from the continuous updating of weights and biases in preceding layers. This shift can lead to slower convergence, as the learning process must constantly adapt to the changing input distributions, and may also contribute to overfitting. Batch normalization aims to reduce the internal covariate shift by normalizing the inputs to each layer, ensuring that they have a consistent mean and variance during the training process. As shown in Algoritm-1, the technique operates by normalizing the input activations for each mini-batch, scaling and shifting them to have a mean of zero and a variance of one. After normalization, learnable parameters (gamma and beta) are introduced, allowing the network to adjust the normalized values to the desired mean and variance, which is crucial for the model to learn more complex relationships.

Input: Values of x over a mini-batch: $\beta = \{x_{1.....m}\}$ Parameters to be learned: β, γ Fixed parameter: $\in$ is used to avoid division by zero error.	
Output: $\{y_i = BN_{\beta,\gamma}(x_i)\}$	
$\mu_\beta \leftarrow \frac{1}{m}\sum_{i=1}^{m} x_i$	//mini-batch mean
$\sigma_\beta^2 \leftarrow \frac{1}{m}\sum_{i=1}^{m}(x_i - \mu_\beta)^2$	//mini-batch variance
$\widehat{x_i} \leftarrow \frac{x_i - \mu_\beta}{\sqrt{\sigma_\beta^2 + \in}}$	//normalize
$y_i \leftarrow \gamma\widehat{x_i} + \beta$	//scale and shift

Algorithm-1: Batch Normalization Transform.

Incorporating batch normalization into CNNs offers several benefits. First, it accelerates the training process by allowing the use of higher learning rates without the risk of divergence or instability. This is because the normalization process reduces the dependence of gradients on the parameter scale, making the network more robust to large parameter updates [153]. Second, batch normalization acts as a form of regularization. This effect is achieved as the normalization process introduces noise into the model due to its mini-batch processing algorithm. This makes the model less sensitive to small noise

in input data [154]. Finally, batch normalization has been shown to improve generalization performance, leading to state-of-the-art results on various computer vision tasks [155].

2.1.7. *Dropout*

Dropout is a regularization technique widely used in CNNs to prevent overfitting and improve generalization capabilities. It was first introduced by Srivastava et al. in their paper [156]. As depicted in Figure 9, the main idea behind dropout is to randomly deactivate, or "drop," a proportion of neurons in the network during training, which forces the model to learn redundant representations of the input data. In CNNs, dropout is typically applied after fully connected layers or after pooling layers, depending on the architecture. During the training phase, dropout randomly selects a predefined fraction of neurons and sets their activations to zero. This process effectively simulates the presence of multiple subnetworks that are trained in parallel, each sharing a portion of the overall network's weights. As a result, this ensemble of networks can yield more robust predictions, making the model less reliant on individual neurons and less prone to overfitting [157].

(a)   **(b)**
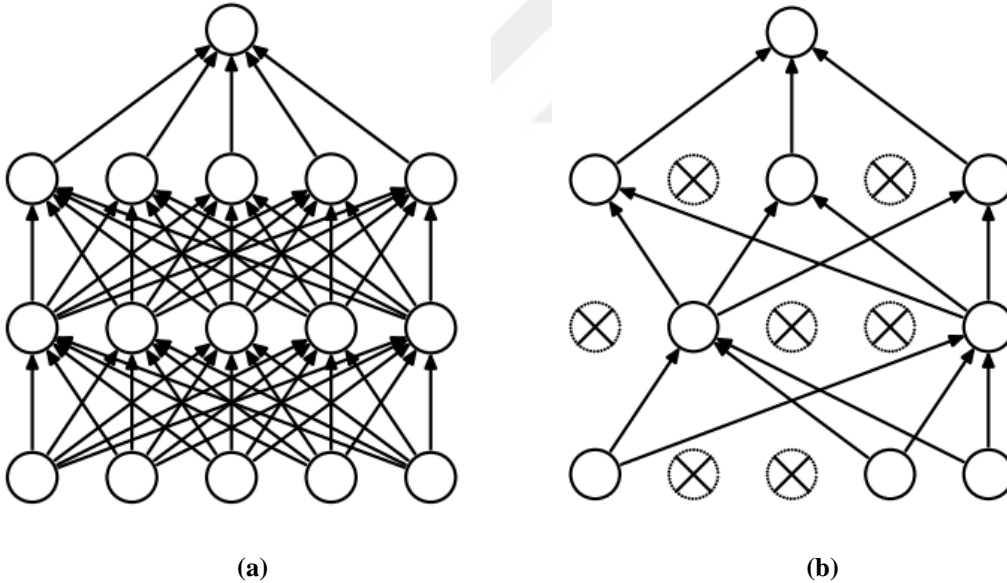
Figure 9: (a) Standard Neural Net (b) the network after applying dropout

METHODOLOGY

3.1. Data Acquisition

In this research, we developed a system for capturing hyperspectral microscopy images in our laboratory by integrating a Headwall A-series VNIR model push-broom hyperspectral camera and a Euromex Oxion light microscope. An illustration of our system is available at Figure 10. A real-world photo of our data acquisition system is presented in Figure 11.
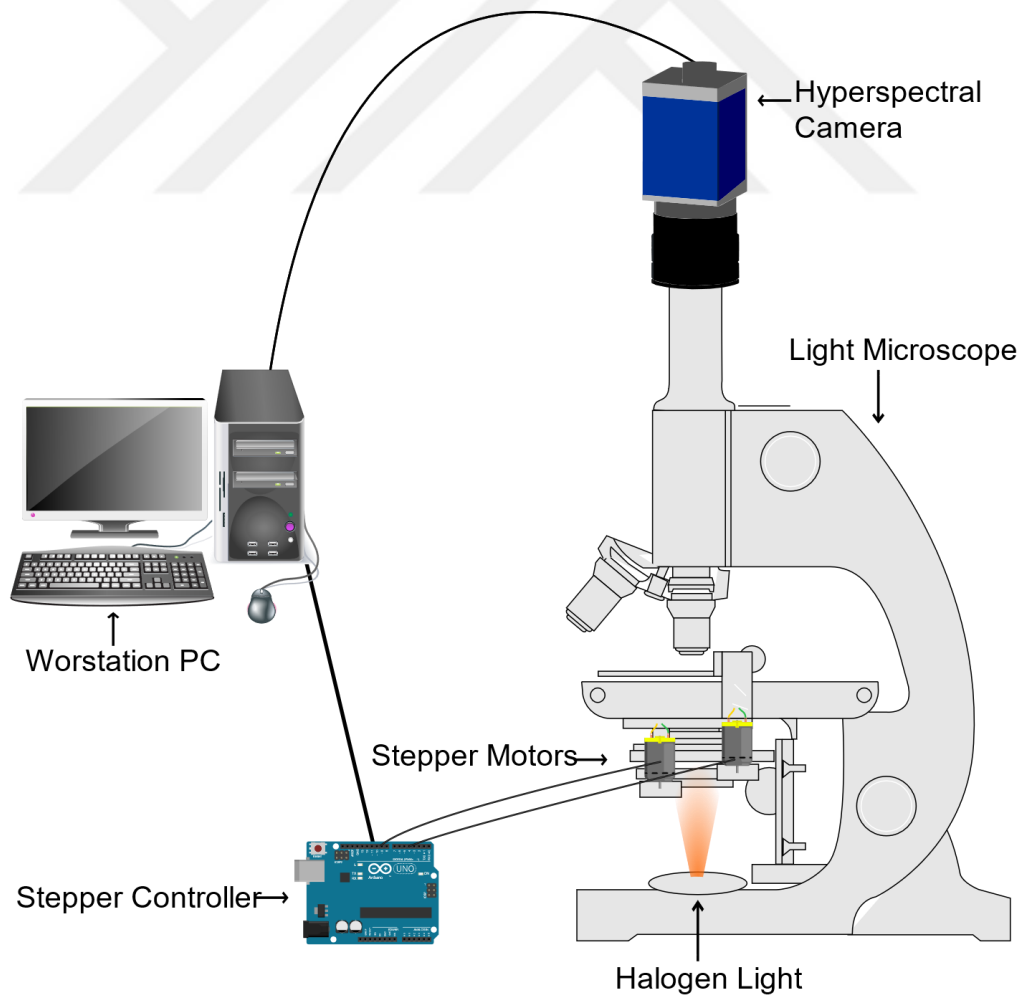
Figure 10: Illustration of our data acquisition system.
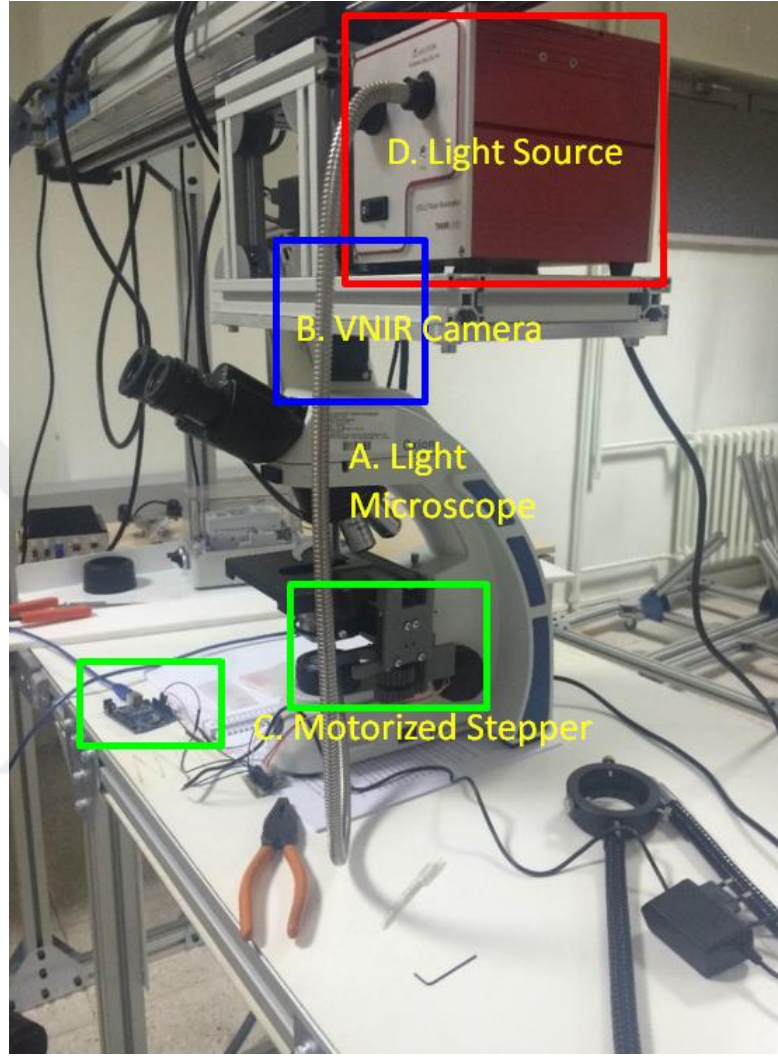
Figure 11: Data Acquisition system. (A) Light microscope; (B) VNIR Camera; (C) Motorized Stepper; (D) Light Source.

Light microscopes, also known as optical microscopes, function based on the principle of manipulating light to magnify and resolve minute objects utilizing a system of lenses [158]. Basic components of a light microscope are illustrated in Figure 12. The image formation process in a light microscope can be studied in two primary categories; magnification and resolution. Magnification involves enlarging the specimen's appearance, which is achieved by the combination of the objective lens, situated near the specimen, and the eyepiece (ocular lens), through which the user observes the magnified image [159]. The total magnification is the product of the magnification factors of the objective lens and the eyepiece. Resolution, on the other hand, is the capacity of the microscope to distinguish two closely spaced points as separate entities [160]. In light microscopes, the resolution is determined by the wavelength (λ) of the light employed and the numerical aperture (NA) of the objective lens. The numerical aperture is a number that depends on the refractive index (n) of the medium between the specimen and the objective

lens, and the half-angle of the maximum cone of light (α) that can enter the lens. Mathematically, numerical aperture can be formatted as Equation 15.

$$NA = n \sin \theta$$

(15)

Whereas the limit of resolution (d) can be calculated using the Abbe's diffraction limit formula as Equation 16.

$$d = \frac{\lambda}{2NA}$$

(16)

The performance and functionality of a light microscope are contingent upon various components working in harmony. These include the eyepiece for image observation, objective lenses with different magnification powers (e.g., 4x, 10x, 40x, 100x), the stage for holding the specimen, and the light source for illumination (e.g., LED, halogen lamp, or mirror). The condenser, situated below the stage, collects and concentrates light onto the specimen, while the diaphragm controls the light intensity by adjusting its aperture. The coarse and fine focus knobs enable users to adjust the specimen's focus by moving the stage or the objective lens assembly vertically.

(A)

(B)
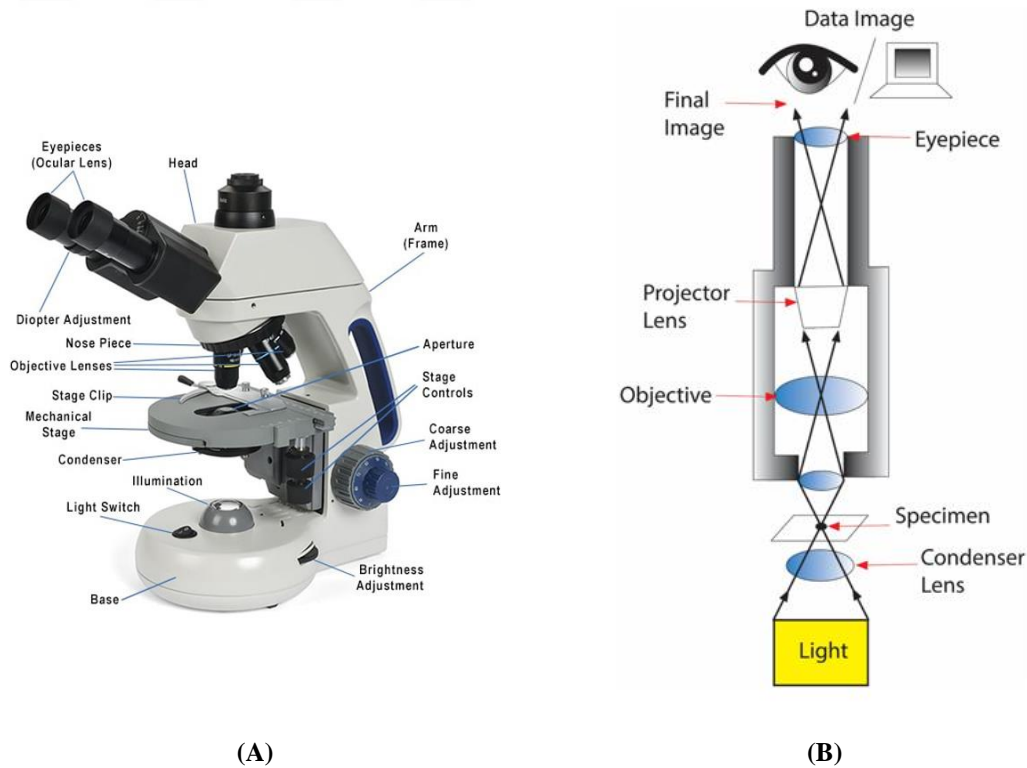
Figure 12: Components of a light microscope. (A) Microscopy device with components (B) Internal optical structure of a light microscope

Our primary aim of integrating hyperspectral camera with light microscope is to establish a robust and versatile imaging system that benefit from the advantages of both technologies. By fusing the high-resolution imaging capabilities of light microscopes with the comprehensive spectral data procured from hyperspectral cameras, we can attain an in-depth understanding of the specimens. High-resolution imaging facilitated by light microscopes enables the visualization of intricate details and structures in specimens, which is crucial for precise histopathology tasks. Nonetheless, conventional microscopy techniques frequently exhibit limitations in delivering extensive chemical or material information pertaining to the sample. This is the context in which hyperspectral cameras become functional. Hyperspectral cameras capture images over an extensive range of wavelengths, yielding a continuous spectrum for each pixel within the image. This spectral information empowers our classification models to identify materials, chemical compounds.

We configured the objective lens of the light microscope to view samples with 40× magnification, and the hyperspectral camera can capture 408 spectral bands between 400 nm to 1000 nm. To achieve optimum image quality, we calibrated and verified our imaging setup by using a microscope stage calibration slide. The spatial resolution of our imaging system is calculated as 0.55 microns, and we measured a liver cell nucleus around 12–18 pixels in diameter and 6.6 to 9.9 microns, which correlates with clinical measurements of human liver cell size [161].

Along with hyperspectral images, the camera simultaneously captures RGB images of the same scene. To obtain data with the proper geometry from our hyperspectral camera, we designed a motorized moving table hardware solution that gradually moves tissue samples while the camera is in capture mode. We controlled the motor speed using a small Arduino device optimized to capture tissue sample images with the highest resolution along the track direction. The tissue samples are illuminated from the bottom by a 3200K 150 W halogen light source, and all images are captured in a dark room without other light sources. For radiometric calibration, we collected white references from an empty glass slide illuminated by the halogen lamp and dark references by blinding the camera sensor with its lens cap. An example of captured data from healthy and unhealthy classes and corresponding tissue components, including cell and background, is presented in Figure 13. The spectra sketches in Figure 13 are obtained by averaging the selected regions' area from the sample image captured with 40× lens magnification. We observed that normal and tumor cell samples transmit different spectral signatures for their respective components.
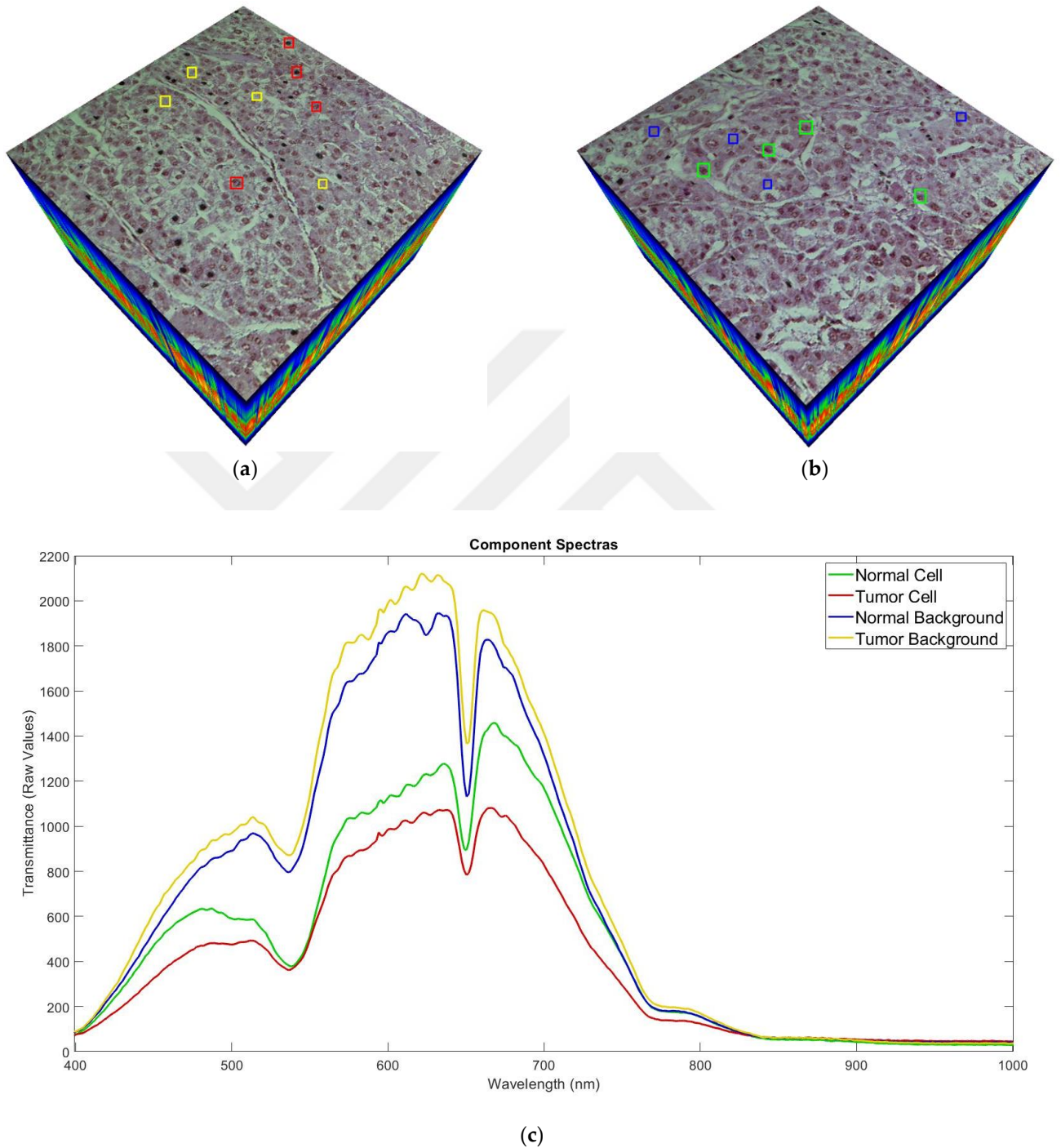
Figure 13: Hyperspectral tissue samples dataset. (a) Tumor (hepatocellular carcinoma, HCC) tissue sample, tumor cells (red), tumor background tissue (yellow); (b) Normal (Healthy) tissue sample, normal cells (green), normal background tissue (blue); (c) Spectra comparison plotting of the given components.

Upon examining the tissue spectra depicted in Figure 13, it becomes evident that there are two distinct dips around the 540 nm and 650 nm wavelengths. A well-known characteristic dip of eosin in H&E staining is situated around the 540 nm wavelength [162-163]. This

particular finding is consistent with previous research studies that have investigated hyperspectral data of liver tissue samples [164-165].

3.2. Classification

Hyperspectral imaging has great potential for classification tasks by fusing spectral and spatial data into machine learning models. However, these models may overfit the training data due to the high dimensionality of the dataset, particularly in small datasets. This issue is especially relevant for complex classifiers like CNN, which may learn random noise instead of meaningful relationships between classes [166]. Additionally, manual feature engineering can limit a model's generalization capability by restricting the feature space. In contrast, deep learning models can automatically find useful features and identify nonlinear relationships between features. Therefore, our study aims to develop a fully automatic classification model with high generalization capability to address the HCC detection problem.

In machine learning and statistics, the bias-variance tradeoff is a fundamental concept, elucidating the intricate relationship between a model's complexity and its ability to generalize to unseen data [167]. Grasping this tradeoff is vital for selecting suitable models, avoiding overfitting or underfitting issues, and achieving optimal model performance. Bias and variance are two components of a classifier's expected prediction error that help to explain the tradeoff between fitting the training data (low bias) and generalizing to new data (low variance). Bias quantifies the systematic deviation introduced by a model's assumptions. In other words, bias represents the discrepancy between the model's expected predictions and the true underlying relationship within the data. Whereas variance implies a model's sensitivity to minor fluctuations in the training data, illustrating the inconsistency of the model's predictions across disparate training sets.

The bias-variance tradeoff arises from the observation that models with low bias typically exhibit high variance, and vice versa [168]. Consequently, there exists an implicit tradeoff between a model's capacity to fit the training data (low bias) and its capacity to generalize to unseen data (low variance). Finding a balance between these two conflicting objectives is crucial for optimizing model performance. In the case of high bias and low variance, trained models make strong assumptions about the data, potentially oversimplifying the underlying relationship. This can lead to underfitting, where the model fails to capture the true structure within the data. High-bias models generally exhibit low variance, as their predictions are relatively consistent across diverse training sets. However, in the case of low bias and high variance, trained models possess greater complexity, enabling them to fit the training data more closely. This can result in overfitting, wherein the model captures the noise in the training data instead of the true underlying relationship. Overfit models often have high variance, as their predictions can be highly sensitive to minor noise in the training data. The performance of a model can be measured using its expected prediction error (EPE), which can be decomposed into three components: bias, variance, and

irreducible error. The irreducible error represents the noise inherent in the data, which cannot be reduced by any model.

The decomposition of EPE is inferred in the study [169]. Let the training set contains points, $x_1, x_2, \ldots, x_n$, and label values y_i assigned to each x_i

$$y = f(x) + \varepsilon \tag{17}$$

Where the noise, ε, has zero mean and variance σ^2

We seek for a function $\hat{f}(x; D)$, approximating the true function $f(x)$ on the training dataset $D = \{(x_1, y_1) \ldots, (x_n, y_n)\}$. Assuming the loss is measured with mean squared error, the objective is to minimize the term, $(y - \hat{f}(x; D))^2$. The expected error for our classification, EPE, can be formulated as follows:

$$E_{D,\varepsilon}\left[(y - \hat{f}(x; D))^2\right] = (Bias_D[\hat{f}(x; D)])^2 + Var_D[\hat{f}(x; D)] + \sigma^2 \tag{18}$$

where

$$Bias_D[\hat{f}(x; D)] = E_D[\hat{f}(x; D)] - E[y(x)] \tag{19}$$

and

$$Var_D[\hat{f}(x; D)] = E_D[(E_D[\hat{f}(x; D)] - \hat{f}(x; D))^2] \tag{20}$$

Informally, the expected prediction error (EPE) can be written as:

$$EPE = Bias^2 + Variance + Irreducible\ Error \tag{21}$$

Achieving both low bias and low variance in classification models is an ideal target for machine learning applications. These models aim to accurately capture the underlying relationships within the data while still maintaining their ability to generalize to unseen data. Given the significance of the bias-variance tradeoff, multiple classification models with varying levels of complexity and flexibility needs to be experimented. By comparing the performance of these models, the optimal balance between bias and variance for our specific dataset and problem domain can be found.

Contrary to the traditional understanding of the bias-variance tradeoff, recent studies have shown that overparameterized deep learning models can achieve better generalization performance than simpler models with low variance [170]. This phenomenon, known as double descent, challenges the conventional rule of thumb that more complex models are prone to overfitting. It suggests that the overparameterized neural networks can outperform simpler models in most scenarios.
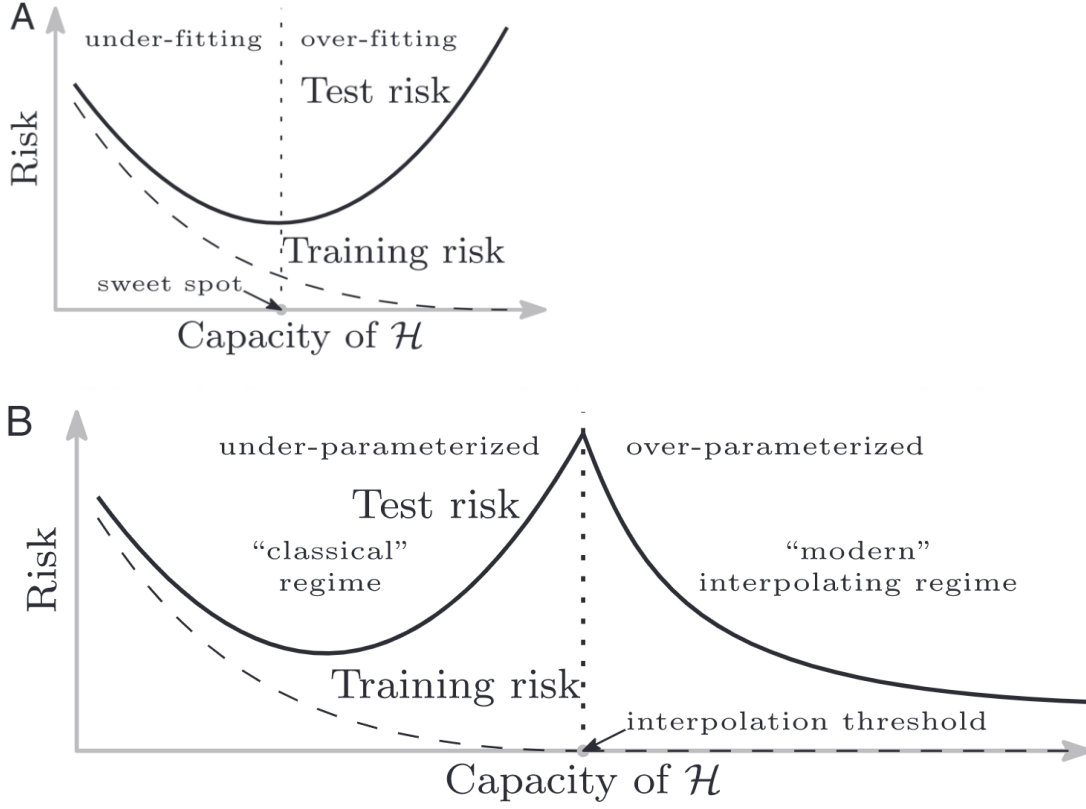
Figure 14: Curves depicting training risk (dotted line) and test risk (continuous line). (A) The traditional U-shaped risk curve resulting from the bias-variance tradeoff. (B) The double descent risk curve, containing the U-shaped risk and the observed behavior from employing high-capacity function classes.

As illustrated in Figure 14, the double descent curve is characterized by three distinct regions: the underparameterized, the interpolating, and the overparameterized regions. In the underparameterized region, models exhibit high bias and low variance, leading to underfitting. As model complexity increases, the interpolation threshold is reached, and the models begin to fit the training data perfectly. This point marks the transition to the interpolating region, where models exhibit low bias but increasing variance, which may result in overfitting. Unexpectedly, as model complexity continues to increase, entering the overparameterized region, the test error starts to decrease again. In this region, models have many parameters that allow them to capture complex relationships within the data. The decrease in test error suggests that these overparameterized models can generalize better than simpler models, contrary to the traditional understanding of the bias-variance tradeoff. A more recent study in [171] elaborates the implications of double descent issue on CNN models. According to the experimental results, effective model complexity (EMC) can be formulated as follows;

$$\text{EMC}_{D,\epsilon}(\mathcal{T}) := \max\left\{ n \mid E_{S \sim D^n}\big[Error_S\big(\mathcal{T}(S)\big)\big] \leq \epsilon \right\} \tag{22}$$

where $Error_S(M)$ is the mean error of model M for the given train samples S, $\mathcal{T}$ represents the network training procedure applied on the distribution D, ϵ is a parameter greater than zero. The theorem also states that there is a threshold value called interpolation threshold such that $\text{EMC}_{D,\epsilon}(\mathcal{T}) = n$, where n is the size of dataset.

The double descent phenomenon remains an active research topic, and its underlying mechanisms are not yet fully understood. Nonetheless, the lottery ticket hypothesis [172] seems as one possible explanation for the functioning of this phenomenon. The lottery ticket hypothesis suggests that large neural networks contain small, sparse subnetworks that, when trained in isolation, can achieve similar performance to the original network with a fraction of the parameters. These small subnetworks are referred to as "winning tickets." The hypothesis argues that the success of large neural networks can be attributed to the presence of these winning tickets. Connecting the double descent phenomenon to the lottery ticket hypothesis, we can infer that the improved generalization performance observed in overparameterized models might be due to the existence of these winning tickets. In other words, the larger models contain subnetworks that can effectively learn the underlying patterns in the data without overfitting, which contributes to their better generalization capabilities.

The double descent phenomenon has significant implications for model selection and training in modern machine learning. First, it suggests that the overparameterized models offer better generalization performance. However, it is essential to carefully monitor the model's performance on validation dataset to ensure that it does not overfit. Second, the double descent phenomenon highlights the importance of properly regularizing overparameterized models. Regularization techniques, such as dropout, can help to control the model's capacity. Then, regularization can mitigate the risk of overfitting while still allowing the model to capture complex relationships within the data [173].

In this study, we utilized a CNN-based learner with 3D convolutions to fully utilize automatic feature learning in deep learning. 3D-CNN models have been extensively employed in various fields such as 3D object recognition [174], video action recognition [175], and medical image recognition [176]. 3D-CNN learners effectively utilize the spatial-spectral data and demonstrate high generalization performance for hyperspectral data. Therefore, we utilized a 3D-CNN learner to extract both spectral signature information encoded within a 3D hyperspectral cube and textural information available on the spatial plane. The activation function utilized in this study is the non-saturating ReLU function, which was suggested in a previous study [177]. ReLU introduces non-linearity into the neural network, which enables the learning of complex and non-linear relationships between inputs and outputs. This non-linearity is crucial for the expressive power and generalization capabilities of deep learning models. Furthermore, ReLU promotes sparse activation of neurons by outputting zero for negative input values. This sparsity in activation fosters more effective internal representations, diminishing the likelihood of overfitting and enhancing model interpretability. ReLU also addresses the vanishing gradient problem, which can cut off the learning process when gradients become excessively small. By having a derivative that is either zero or one, depending on

the input value, ReLU accelerates the convergence of gradient-based optimization algorithms. Due to its simplicity and non-linear properties, ReLU is particularly well-suited for deep neural networks with multiple layers, thus contributing to the scalability of deep learning models.

We have adapted the 3D-CNN topology suggested in the study [144] for our classification task. The details of our 3D-CNN topology are presented in Table 1 and illustrated in Figure 15. The network incorporates max-pooling layers between consecutive convolution layers to reduce the model's complexity and the number of parameters [178]. Additionally, a batch normalization layer follows each max-pooling layer to minimize internal covariate shift, which accelerates the training process by normalizing the mean to zero and standard deviation to one. Therefore, the optimizer algorithm can use a higher learning rate. Instead of a conventional fully connected layer, a global average pooling layer is used to generate feature maps into a 2D structure before feeding to the final dense layer. As highlighted in [179], this layer is immune to overfitting since it has no parameters to optimize, and it is insensitive to spatial translations in the input since it involves spatial averaging. Thus, the proposed model can address overfitting related to texture features in the training set and minimize the noise impact caused by the stepper motor's small vibrations.

Table 1: 3D-CNN Topology with parameters.

Layer	Parameters	Output Size
Input		$100 \times 100 \times 270$
Convolution3D	Kernel Size: $3 \times 3 \times 3$ Number of filters: 4 Activation: ReLU Padding: same	$100 \times 100 \times 270 \times 4$
Max-pooling3D	Pool Size: $3 \times 3 \times 3$ Strides: $2 \times 2 \times 2$ Padding: same	$50 \times 50 \times 135 \times 4$
BatchNormalization		$50 \times 50 \times 135 \times 4$
Convolution3D	Kernel Size: $3 \times 3 \times 3$ Number of filters: 8 Activation: ReLU Padding: same	$50 \times 50 \times 135 \times 8$
Max-pooling3D	Pool Size: $3 \times 3 \times 3$ Strides: $2 \times 2 \times 2$ Padding: same	$25 \times 25 \times 68 \times 8$
BatchNormalization		$25 \times 25 \times 68 \times 8$
Convolution3D	Kernel Size: $3 \times 3 \times 3$ Number of filters: 16 Activation: ReLU Padding: same	$25 \times 25 \times 68 \times 16$
Max-pooling3D	Pool Size: $3 \times 3 \times 3$ Strides: $2 \times 2 \times 2$ Padding: same	$13 \times 13 \times 34 \times 16$
BatchNormalization		$13 \times 13 \times 34 \times 16$
Convolution3D	Kernel Size: $3 \times 3 \times 3$ Number of filters: 32 Activation: ReLU Padding: same	$13 \times 13 \times 34 \times 32$
Max-pooling3D	Pool Size: $3 \times 3 \times 3$ Strides: $2 \times 2 \times 2$ Padding: same	$7 \times 7 \times 17 \times 32$
BatchNormalization		$7 \times 7 \times 17 \times 32$
GlobalAveragePooling3D		32
Dense	Units: 512 Activation: ReLU	512
Dropout	Drop rate: 0.1	512
Dense (Classification)	Units: 1	1
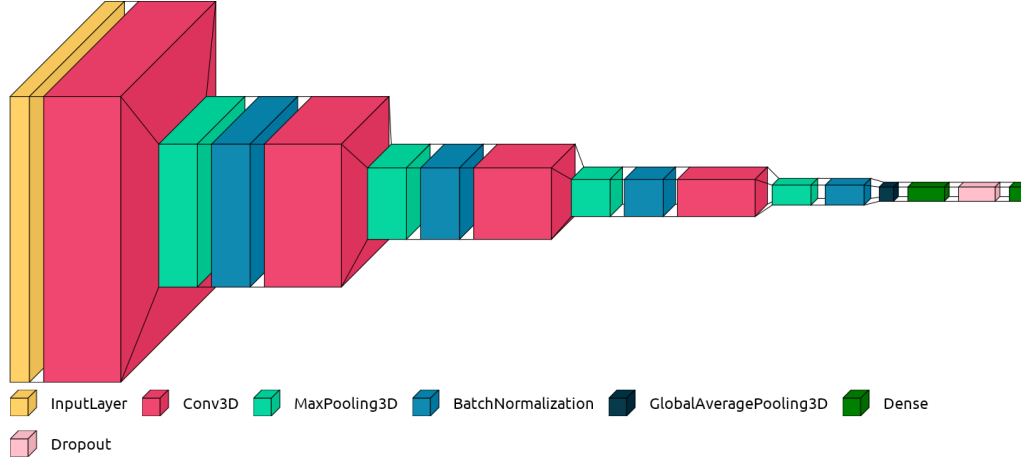
Figure 15: 3D-CNN topology sketch.

To train our CNN model, we followed the recommended practice of using a $3 \times 3 \times 3$ kernel size [154]. We have started with 4 filters at the first convolution layer, then doubled the number at each successive convolution layer, up to 32 filters. Doubling the number of filters in this way helps to capture increasingly complex and abstract features in the hierarchical structure of the network as it progresses through the layers [127]. In the early layers of a CNN, the filters are responsible for detecting simple and low-level features such as edges and textures. As the network progresses through the layers, the receptive field of the neurons increases, and they become capable of detecting more complex and high-level patterns in the input data. By increasing the number of filters, the model is given the capacity to learn a greater variety of higher-level features. Additionally, doubling the number of kernels helps to balance the reduction in spatial dimensions caused by pooling layers. When pooling layers reduce the spatial dimensions of the feature maps, the number of feature maps is increased to preserve the representational power of the network.

We utilized the Adam optimizer for training, which has been shown to be effective in deep learning models [180], with default parameters ($\beta_1 = 0.9$ and $\beta_2 = 0.999$) and a learning rate of 0.001. During training, we used a batch size of 128 and trained the models for 100 epochs, while also applying a dropout rate of 10% to regulate the training. As with many medical studies, our dataset is imbalanced due to a small number of healthy samples compared to tumor samples [181]. This imbalance can lead to CNN classifiers being biased towards the majority class and generating false positives, which is not ideal for medical applications. To address this issue, we have used the focal loss function, which is specifically designed to handle class imbalance in datasets. In our dataset, the presence of a significant class imbalance means that easy positives may dominate the training process, resulting in excessive focus on these examples. To address this issue, we have employed the focal loss function, which can down-weight the loss contribution of easy examples and increase the loss contribution of hard examples.

CHAPTER 4

EXPERIMENTS

4.1.Dataset

The liver tissue array used in this study was obtained from Biomax LV962 (TissueArray.Com LLC, Derwood, MD, USA), which is a commercially available H&E-stained liver tissue slide. The tissue microarray included a range of healthy and unhealthy cases, including 3 normal liver tissues, 1 cancer adjacent liver tissue, 1 metastatic adenocarcinoma, 1 cavernous hemangioma, 4 liver cirrhosis, 3 cholangiocarcinoma, and 32 hepatocellular carcinoma samples. For each case, two tissue samples were included in our dataset, consisting of normal (healthy) and hepatocellular carcinoma (unhealthy) classes. Our dataset comprises 6 healthy and 54 unhealthy tissue samples, as shown in Table 2.

Table 2: Dataset statistics with the class distribution.

	Healthy	**Unhealthy**	**Total**
Training Samples	2	18	20
Training Cases	1	9	10
Validation Samples	2	18	20
Validation Cases	1	9	10
Testing Samples	2	18	20
Testing Cases	1	9	10
Total Samples	6	54	60
Total Cases	3	27	30

The dataset was evenly partitioned into three distinct subsets: training, validation, and testing sets. Each subset includes different patient samples, and there is no overlap between them.

The dataset consists of sample images that have a resolution of 1000×2000 pixels and were captured using a $40\times$ microscopy lens magnification. As shown in Figure 16, to facilitate visualization, an RGB representation was generated from the hyperspectral cube by synthesizing red, green, and blue bands using three normal distributions with a standard deviation of 25 and mean values of 630, 540, and 480, respectively. That is, the corresponding normal distributions for each red, green and blue bands are used as weights to multiply with the spectral dimension. By this way, single band values are obtained from the hyperspectral data. The sample images were divided into smaller patches of size $S\times S$ pixels, where S is a parameter. Some patches contained blank areas without any tissue samples. To ensure a reliable dataset, image patches with more than 50% blank area were automatically removed. Eliminating blank area patches helps to capture actual variance in the dataset, and approximately 5% of the patches are discarded with this method.

(**a**)   (**b**)
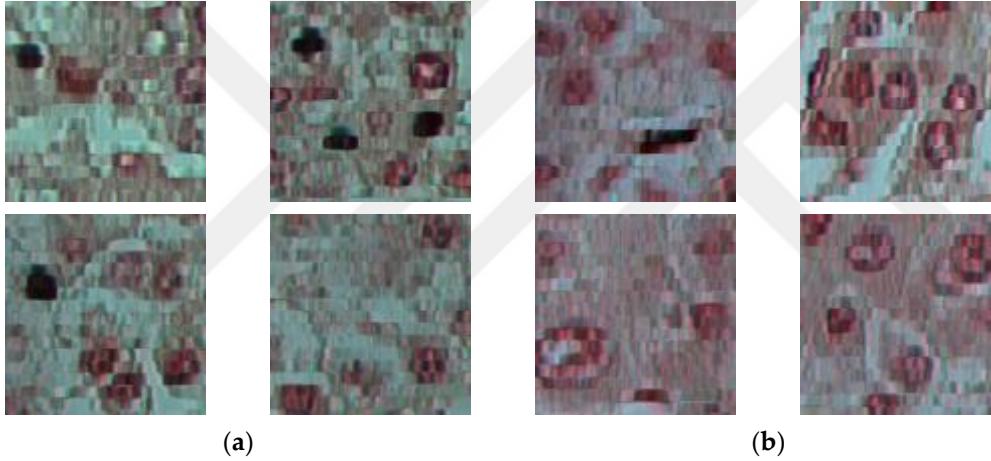
Figure 16: Sample patch images taken with 40X magnification, and the image size is 100×100 pixels. (a) Tumor sample patches; (b) Healthy sample patches.

The final number of patches with varying patch size and their class distributions are given in Table 3. Our hyperspectral imaging system is capable of outputting 408 bands between 400 and 1000 nm. However, upon manual inspection of the samples, we observed that the bands above 800 nm had a low signal-to-noise ratio. Thus, we only used the first 270 bands between 400 and 800 nm to reduce computational cost and prevent the classifier from being presented with flawed information.

Table 3: Dataset statistics for the patch level data

	Healthy	**Unhealthy**	**Total**
Patch Size, S = 50			
Training Samples	1536	13824	15360
Validation Samples	1536	13824	15360
Testing Samples	1536	13824	15360
Class Total	**4608**	**41472**	
Patch Size, S = 100			
Training Samples	384	3456	3840
Validation Samples	384	3456	3840
Testing Samples	384	3456	3840
Class Total	**1152**	**10368**	
Patch Size, S = 150			
Training Samples	148	1333	1481
Validation Samples	148	1333	1481
Testing Samples	148	1333	1481
Class Total	**444**	**3999**	
Patch Size, S = 200			
Training Samples	95	885	950
Validation Samples	95	885	950
Testing Samples	95	885	950
Class Total	**285**	**2655**	

4.2. Hardware and Software Configuration

The AI server used in our study is equipped with eight NVIDIA V100 Tensor Core 32GB GPUs that feature 5,120 Tensor Cores, providing up to 1 petaflop of AI computing performance. Additionally, the server is powered by dual Intel Xeon E5-2620 v3 CPUs and has 128 GB of DDR4 memory. Our setup allows us to train eight different models simultaneously. For deep learning programming, we utilized Python 3.8, Keras 2.3.1 with Tensorflow 2.0, CUDA for GPU acceleration, and Ubuntu 18.04 as the primary operating system.

4.3. Evaluation Metrics

To assess the performance of our classifier, we have utilized various metrics, including accuracy, precision, recall, F1 score, and Matthews Correlation Coefficient (MCC). These metrics are formulated in Equations 23 to 27 and are commonly used in medical studies [182] to evaluate classifiers that address class imbalance problems. The MCC metric is particularly suitable for our study and is calculated using Equation 27. The MCC metric ranges between -1 and 1, with 1 indicating perfect predictions, 0 representing random predictions, and -1 implying complete disagreement between the classifier's prediction and the observation. To compute these metrics, we use True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) output metrics from the classifier.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{23}$$

$$Precision = \frac{TP}{TP + FP} \tag{24}$$

$$Recall = \frac{TP}{TP + FN} \tag{25}$$

$$F1Score = 2.\frac{precision.recall}{precision + recall} \tag{26}$$

$$MCC = \frac{TP.TN - FP.FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{27}$$

4.4. Quantitative Results and Discussion

In order to assess the effectiveness of our proposed method, we trained multiple CNN classifiers using various configurations, but with the same base topology outlined in Table 1. At the first three experiments, we identified the optimal network topology for HCC classification task. That is, through experimentation, we determined depth of network,

number of filters to employ and input image size parameters. Afterwards, by using the optimized CNN topology, we investigated the implications of cost function, spectral data resolution and type of convolution operator aspects. As mentioned in [47], the hyperparameters γ and α in the focal loss function are specific to the dataset, and must be adjusted accordingly. Therefore, to ensure a fair evaluation, we optimized these hyperparameters for the dataset related configurations. Additionally, we compared the classification performance of various spectral resolutions, including hyperspectral, sampled hyperspectral, PCA of hyperspectral, and RGB. Subsequently, we explored the implications of convolution operation by comparing the results of 2D and 3D convolution-based CNNs. Lastly, we conducted an experiment by rotating our dataset splits between training, validation, and testing subsets to prove the generalization capability of our models for the data.

4.4.1. *Experiment-1: CNN Number of Convolution Layers*

For the first experiment setup, we have conducted experiments to evaluate the impact of varying the number of convolution layers in our base CNN topology given in Table 1. The primary goal was to understand how the depth of the network influences its performance. For the experiments, we have used varying number of convolution layers followed by max-pooling and batch normalization layers. The results of classification are presented in Table 4. The results demonstrate that a CNN with four convolution layers achieves the best overall performance, with an accuracy of 0.97, precision of 0.999, recall of 0.968, F1-score of 0.984, and MCC of 0.860. This suggests that a four-layer CNN can effectively capture the hierarchical features available in the images. On the other hand, a CNN with only two convolution layers produces significantly lower performance, with an accuracy of 0.745. The lower performance can be attributed to the shallow architecture's limited capacity to learn complex features, leading to reduced discriminative power and weaker classification results. As the number of convolution layers increases to three, there is an obvious improvement in performance, with an accuracy of 0.902. This highlights the benefits of adding more layers to capture higher-level features and improve the model's generalization capabilities. As the depth of the CNN is further increased to five and six layers, the performance remains relatively high, with accuracies of 0.962 and 0.965, respectively. While the performance does not surpass that of the four-layer CNN, it demonstrates that increasing the depth beyond four layers does not lead to a substantial degradation in performance. This suggests that deeper CNNs can still maintain strong classification performance, although the gains may be marginal compared to the optimal four-layer CNN.

Table 4: Classification results for varying topology depth configurations. (focusing parameter $\gamma = 2$, weighting factor $\alpha = 0.50$, 3D convolutions, patch size S=100)

Number of Convolution Layers	Accuracy	Precision	Recall	F1-Score	MCC
2	0.745	0.936	0.77	0.83	0.201
3	0.902	0.975	0.914	0.937	0.583
4	**0.97**	**0.999**	**0.968**	**0.984**	**0.86**
5	0.962	0.991	0.966	0.976	0.814
6	0.965	0.995	0.967	0.98	0.834

4.4.2. *Experiment-2: CNN Number of Filters*

In the second experiment setup, we have investigated the impact of varying the number of filters used in each layer of CNN. The primary objective was to understand how the number of filters in each layer influences the performance of the network. For this purpose, we have started with the number of filters, n, then doubled the number at each successive convolution layer, up to $8n$. The results, in Table 5, indicate that the optimal number of filters for the CNN in this study is the configuration 4-8-16-32, which achieves the highest performance across all evaluation metrics: accuracy (0.97), precision (0.999), recall (0.968), F1-score (0.984), and MCC (0.86). This suggests that the 4-8-16-32 configuration effectively captures the hierarchical features present in the images, balancing the complexity of the model and its capacity to generalize to unseen data. When comparing the performance of CNNs with different kernel configurations, we can observe that the lowest performance is achieved by the 1-2-4-8 configuration, with an accuracy of 0.703. This can be attributed to the limited capacity of the network to learn complex features due to the insufficient number of filters, which results in weaker classification performance. As the number of filters increases, such as in the 2-4-8-16 and 8-16-32-64 configurations, the performance improves significantly, with accuracies of 0.859 and 0.966, respectively. This improvement highlights the benefits of using more filters to capture higher-level features. The configuration with the highest number of filters, 16-32-64-128, results in a performance close to the optimal configuration, with an accuracy of 0.969. This suggests that further increasing the number of filters does not necessarily lead to a significant improvement in performance.

Table 5: Classification results for varying number of filters for each convolution layer. (focusing parameter γ = 2, weighting factor α = 0.50, 3D convolutions, patch size S=100).

Number of Filters	Accuracy	Precision	Recall	F1-Score	MCC
1-2-4-8	0.703	0.935	0.72	0.803	0.174
2-4-8-16	0.859	0.973	0.868	0.912	0.491
4-8-16-32	**0.97**	**0.999**	**0.968**	**0.984**	**0.86**
8-16-32-64	0.966	0.996	0.966	0.981	0.838
16-32-64-128	0.969	0.997	0.968	0.983	0.852

4.4.3. Experiment-3: CNN Topology Input Resolution

The third experiment focused on analyzing the effect of patch size (S) on classification performance. This parameter is crucial for the classification approach since it governs the variation in textural features on a single patch image. These features comprise different components such as cell nucleus, cytoplasm, and blank areas in the tissue sample. It is important to note that the cropped patches' size should not be too small to overlook critical textural features. Conversely, the classifier may concentrate solely on dense areas when the parameter is too large. We analyzed the performance of the model with different patch sizes, ranging from 50x50 to 200x200, as the results are shown in Table 6. The results reveal that the 100x100 patch size achieved the highest performance across all evaluation metrics: accuracy (0.970), precision (0.999), recall (0.968), F1-score (0.984), and MCC (0.860). This suggests that a patch size of 100x100 provides an optimal balance between capturing spatial information and minimizing computational complexity, leading to better generalization capabilities. The 50x50 patch size exhibited slightly lower performance, with an accuracy of 0.929 and an MCC of 0.722. While this configuration may require less computational resources, the smaller patch size might limit the model's ability to learn more complex and discriminative features, thus affecting its overall performance. When the patch size was increased to 150x150, the accuracy and MCC dropped slightly to 0.961 and 0.774, respectively. This indicates that a larger patch size may lead to increased model complexity without providing significant improvements in performance. Lastly, the 200x200 patch size demonstrated the lowest performance among the tested configurations, with an accuracy of 0.924 and an MCC of 0.615. The increase in patch size resulted in a trade-off between the model's capacity to capture spatial information and its ability to generalize to new data, ultimately leading to diminished performance. Consequently, we set the patch size parameter to 100 for the remaining experiments.

Table 6: Classification results for varying patch size parameter. (focusing parameter $\gamma = 2$, weighting factor $\alpha = 0.50$, 3D convolutions).

Patch Size (S)	Accuracy	Precision	Recall	F1-Score	MCC
50×50	0.929	0.996	0.924	0.961	0.722
100×100	0.970	0.999	0.968	0.984	0.860
150×150	0.961	0.973	0.983	0.967	0.774
200×200	0.924	0.968	0.947	0.945	0.615

4.4.4. Experiment-4: Optimal Cost Function

During the fourth phase of our experiments, we identified the optimal hyperparameters γ and α for the HSI dataset. We compared the performance of the balanced cross-entropy function with that of the focal loss function. The focal loss function employs a weighting factor α that differentiates between loss values for minority (healthy) and majority (tumor) classes, balancing the influence of negative and positive examples on the loss. The focusing parameter γ reduces the loss contribution from well-classified, easy examples while maintaining a high loss contribution for hard examples. This parameter adjusts the level of focus on hard examples during training. The optimal values for α and γ depend on the severity of imbalance and the distribution of hard and easy examples in the dataset. As reported in the paper [47] that introduced the focal loss function, modifying the focusing parameter γ has a greater impact than modifying the weighting factor α. As shown in Table 7, empirical results show that the parameter values $\gamma = 2$ and $\alpha = 0.5$ produce the best classification performance for our HSI dataset, despite the extreme class imbalance. The same α value of 0.5 is chosen for both positive and negative classes, as the focusing parameter γ down-weights easy positives and requires less focus on negatives. Therefore, the model focuses on hard examples rather than intentionally emphasizing the minority class. We conclude that γ is the critical factor in the loss function, while α should be optimized for each γ value. In the balanced cross-entropy configuration, the classifier produces a low precision due to a relatively high false positive rate. However, when α is set to 0.25 in balanced cross-entropy, there is an obvious improvement in precision due to the reduction in false positives. These results highlight the importance of α in the behavior of the cost function in balanced cross-entropy form. Nevertheless, focal loss with different combinations of γ and α values generally outperforms CE and BCE. As the γ value increases, the performance tends to improve, with the best results obtained for $\gamma=2.0$. The choice of α also impacts performance, with $\alpha=0.50$ appearing to provide the optimal balance between emphasizing hard-to-classify examples and avoiding overfitting.

Table 7: Classification results with HSI dataset for varying cost functions and respective parameter sets. (patch size parameter S = 100, 3D convolutions).

Loss Function	Accuracy	Precision	Recall	F1-Score	MCC
CE	0.891	0.921	0.961	0.906	0.277
BCE ($\gamma = 0, \alpha = 0.25$)	0.901	0.938	0.953	0.919	0.412
BCE ($\gamma = 0, \alpha = 0.75$)	0.902	0.916	0.981	0.909	0.278
FL ($\gamma = 1.5, \alpha = 0.25$)	0.918	0.982	0.926	0.949	0.646
FL ($\gamma = 1.5, \alpha = 0.50$)	0.922	0.978	0.935	0.949	0.643
FL ($\gamma = 1.5, \alpha = 0.75$)	0.930	0.969	0.952	0.949	0.638
FL ($\gamma = 2.0, \alpha = 0.25$)	0.960	0.993	0.962	0.976	0.811
FL ($\gamma = 2.0, \alpha = 0.50$)	**0.970**	**0.999**	**0.968**	**0.984**	**0.860**
FL ($\gamma = 2.0, \alpha = 0.75$)	0.955	0.983	0.966	0.969	0.766
FL ($\gamma = 2.5, \alpha = 0.25$)	0.948	0.998	0.943	0.972	0.782
FL ($\gamma = 2.5, \alpha = 0.50$)	0.958	0.999	0.955	0.978	0.818
FL ($\gamma = 2.5, \alpha = 0.75$)	0.953	0.986	0.962	0.969	0.768

4.4.5. Experiment-5: Spectral Resolution

We conducted a further experiment to assess the classification performance of different spectral resolutions. We compared various versions of our dataset, including a hyperspectral dataset with 270 bands (HSI), as well as sampled hyperspectral datasets with 90 (HSI-90), 30 (HSI-30), and 10 (HSI-10) bands. We also included two PCA-based versions of the hyperspectral dataset with nine (PCA-9) and three (PCA-3) principal components, as well as RGB versions of our dataset. To generate the HSI-90, HSI-30, and HSI-10 datasets, we sampled individual bands from the HSI dataset with a constant frequency. Additionally, we utilized the PCA method to reduce the dimensionality of the HSI dataset. Specifically, we applied the PCA algorithm described in [183], which is an incremental technique designed for calculating the PCA of large datasets. As Figure 17 suggests, we selected the first nine principal components using a variance threshold value of 0.1%. Additionally, we have generated a PCA dataset using the first three principal components to compare its performance with the RGB dataset. The total variance of the PCA-3 dataset was 93.46%, whereas the PCA-9 dataset had a total variance of 98.60%. Alongside hyperspectral datasets, we have also incorporated RGB data captured by the

hyperspectral camera, which consists of three individual bands taken from the 630, 540, and 480 nm wavelengths. We trained a 3D-CNN model using RGB images with the same topology as that of the hyperspectral datasets. The hyperparameters of the focal loss function of 3D-CNN for the datasets were fine-tuned empirically, and the results are presented in Tables 7-12. The experimentation results indicate that HSI outperforms RGB with the highest accuracy and MCC score. Moreover, the sampled HSI datasets demonstrate that an increased number of hyperspectral bands result in better classification performance. Specifically, we observed that the classification accuracy is directly proportional to the number of bands included in the dataset. Among the sampled hyperspectral datasets, the PCA-9 dataset achieved the second-best classification accuracy since it retains most of the variance from the original HSI dataset. The PCA-3 dataset exhibited lower accuracy than the PCA-9 dataset but still outperformed the RGB dataset in terms of accuracy and MCC score.
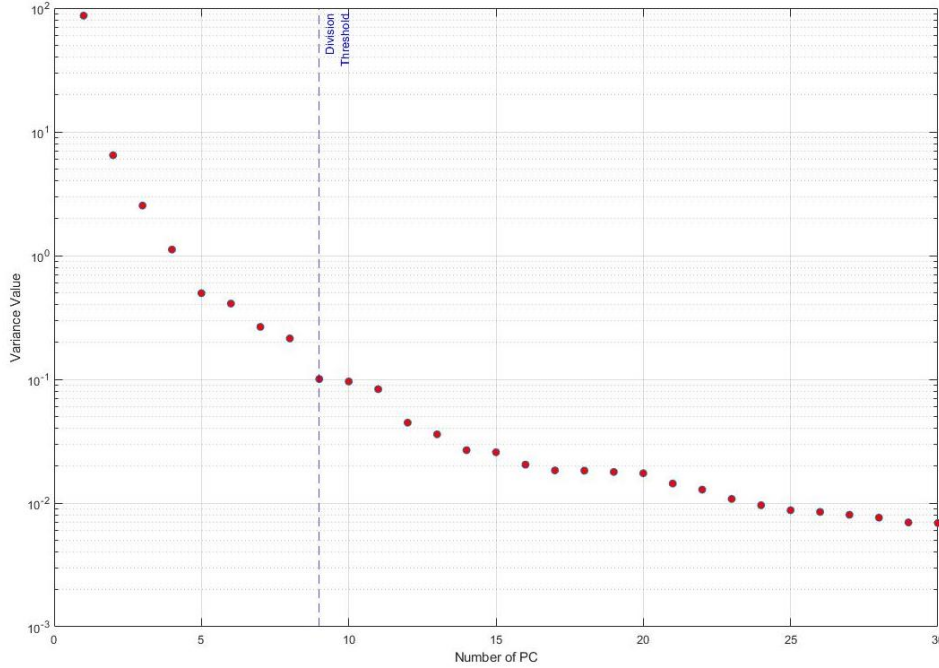
Figure 17: The plot showing principle component ranking and their variance values. The y-axis is log-scaled for better visualization.

Table 8: Classification results with HSI-90 dataset for varying cost functions and respective parameter sets. (patch size parameter S = 100, 3D convolutions).

Loss Function	Accuracy	Precision	Recall	F1-Score	MCC
FL ($\gamma = 1.5, \alpha = 0.25$)	0.885	0.973	0.897	0.927	0.539
FL ($\gamma = 1.5, \alpha = 0.50$)	0.892	0.975	0.903	0.932	0.559
FL ($\gamma = 1.5, \alpha = 0.75$)	0.897	0.971	0.913	0.933	0.554
FL ($\gamma = 2.0, \alpha = 0.25$)	0.919	0.980	0.929	0.949	0.645
FL ($\gamma = 2.0, \alpha = 0.50$)	**0.950**	**0.990**	**0.955**	**0.970**	**0.767**
FL ($\gamma = 2.0, \alpha = 0.75$)	0.924	0.986	0.929	0.954	0.679
FL ($\gamma = 2.5, \alpha = 0.25$)	0.914	0.977	0.926	0.944	0.619
FL ($\gamma = 2.5, \alpha = 0.50$)	0.885	0.977	0.894	0.929	0.552
FL ($\gamma = 2.5, \alpha = 0.75$)	0.857	0.970	0.868	0.91	0.473

Table 9: Classification results with HSI-30 dataset for varying cost functions and respective parameter sets. (patch size parameter S = 100, 3D convolutions).

Loss Function	Accuracy	Precision	Recall	F1-Score	MCC
FL ($\gamma = 1.5, \alpha = 0.25$)	0.877	0.970	0.890	0.921	0.510
FL ($\gamma = 1.5, \alpha = 0.50$)	0.885	0.973	0.897	0.927	0.537
FL ($\gamma = 1.5, \alpha = 0.75$)	0.885	0.971	0.90	0.926	0.53
FL ($\gamma = 2.0, \alpha = 0.25$)	0.889	0.974	0.900	0.930	0.55
FL ($\gamma = 2.0, \alpha = 0.50$)	**0.931**	**0.985**	**0.937**	**0.957**	**0.692**
FL ($\gamma = 2.0, \alpha = 0.75$)	0.909	0.975	0.923	0.941	0.598
FL ($\gamma = 2.5, \alpha = 0.25$)	0.883	0.971	0.897	0.925	0.524
FL ($\gamma = 2.5, \alpha = 0.50$)	0.861	0.973	0.870	0.914	0.495
FL ($\gamma = 2.5, \alpha = 0.75$)	0.857	0.97	0.868	0.91	0.473

Table 10: Classification results with HSI-10 dataset for varying cost functions and respective parameter sets. (patch size parameter S = 100, 3D convolutions)

Loss Function	Accuracy	Precision	Recall	F1-Score	MCC
FL ($\gamma = 1.5, \alpha = 0.25$)	0.857	0.967	0.871	0.909	0.46
FL ($\gamma = 1.5, \alpha = 0.50$)	0.857	0.970	0.868	0.910	0.472
FL ($\gamma = 1.5, \alpha = 0.75$)	0.859	0.966	0.874	0.909	0.46
FL ($\boldsymbol{\gamma = 2.0, \alpha = 0.25}$)	**0.914**	**0.978**	**0.926**	**0.945**	**0.622**
FL ($\gamma = 2.0, \alpha = 0.50$)	0.910	0.977	0.922	0.942	0.61
FL ($\gamma = 2.0, \alpha = 0.75$)	0.898	0.973	0.911	0.934	0.565
FL ($\gamma = 2.5, \alpha = 0.25$)	0.867	0.973	0.876	0.917	0.505
FL ($\gamma = 2.5, \alpha = 0.50$)	0.862	0.973	0.871	0.914	0.498
FL ($\gamma = 2.5, \alpha = 0.75$)	0.890	0.970	0.905	0.928	0.534

Table 11: Classification results with PCA-9 dataset for varying cost functions and respective parameter sets. (patch size parameter S = 100, 3D convolutions).

Loss Function	Accuracy	Precision	Recall	F1-Score	MCC
FL ($\gamma = 1.5, \alpha = 0.25$)	0.897	0.975	0.909	0.934	0.571
FL ($\gamma = 1.5, \alpha = 0.50$)	0.911	0.977	0.923	0.943	0.614
FL ($\gamma = 1.5, \alpha = 0.75$)	0.911	0.974	0.926	0.941	0.602
FL ($\gamma = 2.0, \alpha = 0.25$)	0.945	0.987	0.952	0.966	0.743
FL ($\boldsymbol{\gamma = 2.0, \alpha = 0.50}$)	**0.957**	**0.988**	**0.964**	**0.972**	**0.788**
FL ($\gamma = 2.0, \alpha = 0.75$)	0.927	0.987	0.932	0.956	0.687
FL ($\gamma = 2.5, \alpha = 0.25$)	0.930	0.980	0.941	0.954	0.674
FL ($\gamma = 2.5, \alpha = 0.50$)	0.892	0.981	0.897	0.934	0.583
FL ($\gamma = 2.5, \alpha = 0.75$)	0.882	0.976	0.891	0.927	0.544

Table 12: Classification results with PCA-3 dataset for varying cost functions and respective parameter sets. (patch size parameter S = 100, 3D convolutions).

Loss Function	Accuracy	Precision	Recall	F1-Score	MCC
FL ($\gamma = 1.5, \alpha = 0.25$)	0.873	0.971	0.885	0.919	0.503
FL ($\gamma = 1.5, \alpha = 0.50$)	0.879	0.970	0.894	0.922	0.511
FL ($\gamma = 1.5, \alpha = 0.75$)	0.857	0.941	0.897	0.897	0.336
FL ($\gamma = 2.0, \alpha = 0.25$)	0.899	0.977	0.910	0.936	0.584
FL ($\gamma = 2.0, \alpha = 0.50$)	**0.913**	**0.983**	**0.919**	**0.947**	**0.638**
FL ($\gamma = 2.0, \alpha = 0.75$)	0.906	0.975	0.920	0.939	0.590
FL ($\gamma = 2.5, \alpha = 0.25$)	0.888	0.976	0.897	0.930	0.556
FL ($\gamma = 2.5, \alpha = 0.50$)	0.889	0.974	0.900	0.930	0.550
FL ($\gamma = 2.5, \alpha = 0.75$)	0.923	0.972	0.942	0.947	0.626

Table 13: Classification results with RGB dataset for varying cost functions and respective parameter sets. (patch size parameter S = 100, 3D convolutions).

Loss Function	Accuracy	Precision	Recall	F1-Score	MCC
FL ($\gamma = 1.5, \alpha = 0.25$)	0.852	0.964	0.868	0.905	0.440
FL ($\gamma = 1.5, \alpha = 0.50$)	0.851	0.966	0.865	0.905	0.449
FL ($\gamma = 1.5, \alpha = 0.75$)	0.828	0.936	0.869	0.879	0.269
FL ($\gamma = 2.0, \alpha = 0.25$)	**0.900**	**0.974**	**0.914**	**0.936**	**0.573**
FL ($\gamma = 2.0, \alpha = 0.50$)	0.891	0.972	0.905	0.930	0.544
FL ($\gamma = 2.0, \alpha = 0.75$)	0.882	0.970	0.897	0.924	0.520
FL ($\gamma = 2.5, \alpha = 0.25$)	0.859	0.970	0.871	0.911	0.479
FL ($\gamma = 2.5, \alpha = 0.50$)	0.853	0.968	0.865	0.907	0.458
FL ($\gamma = 2.5, \alpha = 0.75$)	0.878	0.964	0.897	0.919	0.486

We put the comparative results for all spectral resolutions in Table 14. The performance of the model is obviously highest when it is trained with the full 270 bands HIS, boasting an accuracy of 0.970, precision of 0.999, recall of 0.968, F1-Score of 0.984, and MCC of 0.860. This can be explained as the full HSI dataset is the richest in terms of spectral information which is crucial for the accurate classification of HCC and normal liver tissues. As we decrease the number of bands used in the hyperspectral datasets (HSI-90, HSI-30, and HSI-10), there is a clear decline in all the evaluation metrics. This can be attributed to the loss of spectral information that contributes to the discriminatory power of the model. Notably, the model's performance drops significantly when trained with the HSI-10 dataset compared to the full HSI dataset, emphasizing the importance of hyperspectral imaging for a robust classification model. The model trained on the PCA-9 dataset performs relatively close to the HSI-90 dataset, indicating that the first nine principal components contain substantial information for the classification task. However, reducing the number of principal components to three (PCA-3) results in a considerable drop in performance, though it still outperforms the model trained on the RGB data. The RGB model trails behind all the others in terms of performance, indicating the superiority of hyperspectral data or its PCA equivalents over conventional RGB images in the context of tissue classification. Despite the simplicity of RGB data, it misses out on the valuable spectral information encapsulated in the hyperspectral data.

Table 14: Classification results for different spectral datasets (patch size parameter S = 100, 3D convolutions, focal loss parameters optimized)

Dataset	Accuracy	Precision	Recall	F1-Score	MCC
HSI	**0.970**	**0.999**	**0.968**	**0.984**	**0.860**
HSI-90	0.950	0.990	0.955	0.970	0.767
HSI-30	0.931	0.985	0.937	0.957	0.692
HSI-10	0.914	0.978	0.926	0.945	0.622
PCA-9	0.957	0.988	0.964	0.972	0.788
PCA-3	0.913	0.983	0.919	0.947	0.638
RGB	0.900	0.974	0.914	0.936	0.573

4.4.6. Experiment-6: Convolution Operator

In the sixth experiment of our study, we investigated the effect of convolution operator on classification performance. Specifically, we trained another classification model with a 2D convolution operation using the same network topology as the 3D-CNN model. We fine-tuned the hyperparameters for the 2D convolution model, as given in Table 15. We discovered that the best hyperparameter set, $\gamma = 2$ and $\alpha = 0.5$, for the 3D-CNN model also

produced the best results for the 2D-CNN model. Unlike 3D convolution, which slides in three directions of the hyperspectral cube, 2D convolution operates in two directions of the image data. As a result, the feature sets obtained by 2D and 3D convolution operations have different descriptive abilities. Our analysis, presented in Table 16, indicates that the 3D convolution operator outperforms the 2D version in terms of classification performance. We attribute this to the fact that the 3D convolution operator can fully leverage the potential of hyperspectral data, while the 2D convolution operator leads to a decline in classification performance for hyperspectral data.

Table 15: Classification results with HSI dataset for varying cost functions and respective parameter sets. (patch size parameter S = 100, 2D convolutions).

Loss Function	Accuracy	Precision	Recall	F1-Score	MCC
FL ($\gamma = 1.5, \alpha = 0.25$)	0.891	0.974	0.903	0.931	0.554
FL ($\gamma = 1.5, \alpha = 0.50$)	0.885	0.972	0.897	0.926	0.534
FL ($\gamma = 1.5, \alpha = 0.75$)	0.885	0.970	0.900	0.926	0.527
FL ($\gamma = 2.0, \alpha = 0.25$)	0.922	0.983	0.929	0.952	0.663
FL ($\gamma = 2.0, \alpha = 0.50$)	**0.934**	**0.986**	**0.940**	**0.959**	**0.706**
FL ($\gamma = 2.0, \alpha = 0.75$)	0.920	0.980	0.930	0.949	0.646
FL ($\gamma = 2.5, \alpha = 0.25$)	0.932	0.983	0.94	0.957	0.689
FL ($\gamma = 2.5, \alpha = 0.50$)	0.923	0.982	0.931	0.952	0.658
FL ($\gamma = 2.5, \alpha = 0.75$)	0.915	0.975	0.929	0.944	0.615

Table 16: Comparison of classification results of 3D-CNN and 2D-CNN models trained by HSI data.

Model	Accuracy	Precision	Recall	F1-Score	MCC
HSI-3D-CNN	0.970	0.999	0.968	0.984	0.860
HSI-2D-CNN	0.934	0.986	0.940	0.959	0.706

4.4.7. Experiment-7: Rotating Datasets

In our seventh experiment, we aimed to demonstrate that our 3D-CNN model is not prone to overfitting by rotating the split sets - training, validation, and testing - amongst each other. Our dataset comprises three healthy and 27 unhealthy patients. We created three distinct data-splitting configurations by placing one healthy and nine unhealthy patients

in each of the training, validation, and testing sets. We rotated these sets amongst each other and retrained the model for each configuration. As displayed in Table 17, we observed similar classification performance results for all three configurations. This experiment demonstrates that our 3D-CNN model can learn descriptive features from hyperspectral space without overfitting the training data.

Table 17: Comparison of classification results of 3D-CNN when the training, validation, and testing sets are rotated between each other. (patch size parameter $S = 100$, 3D convolutions).

Model	Accuracy	Precision	Recall	F1-Score	MCC
HSI-3D-CNN (configuration-1)	0.970	0.999	0.968	0.984	0.860
HSI-3D-CNN (configuration-2)	0.965	0.997	0.963	0.981	0.836
HSI-3D-CNN (configuration-3)	0.968	0.996	0.968	0.982	0.846

4.5. Qualitative Results and Discussion

In this sub-section, our qualitative results and related discussions are presented. From the perspective of histopathological analysis, normal and HCC liver H&E tissue samples can differ in several ways, as summarized below.

- **Cell Density**: HCC samples often exhibit a higher cell density as a result of rapid and uncontrolled cell proliferation.

- **Cellular and Nuclear Size**: The cells and nuclei in HCC tissues may appear larger and irregular compared to those in normal liver tissues.

- **Nuclear Pleomorphism**: Normal liver cells typically have uniform nuclear shapes, while HCC cells may display nuclear pleomorphism (variation in size, shape, and staining).

- **Cellular Arrangement**: Normal liver tissues have a regular, organized cellular arrangement. In contrast, HCC tissues often show disorganized cell patterns.

- **Portal Tracts**: Normal liver tissues will have visible portal tracts. In HCC samples, these may be absent or significantly disrupted.

- **Necrosis**: HCC tissues may show areas of necrosis or cell death, which are not commonly seen in healthy liver tissues.

- **Stroma**: The stroma, or supportive tissue, may be increased in HCC samples due to the tumor's growth.

- **Vascular Invasion**: HCC tissues may show signs of vascular invasion or growth of the tumor into the blood vessels, which is not present in normal liver tissues.

- **Steatosis**: Normal liver cells may contain small lipid droplets, whereas HCC cells frequently exhibit clear or eosinophilic cytoplasm.

- **Color**: HCC tissues might be more eosinophilic (pink) in H&E staining due to the increased presence of cytoplasm and connective tissues.
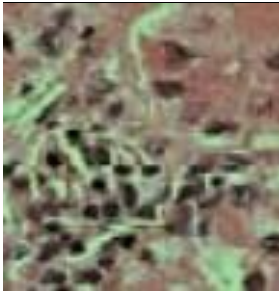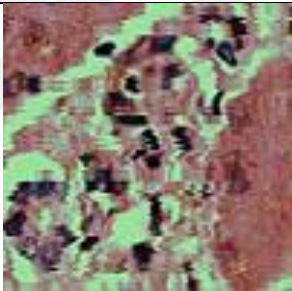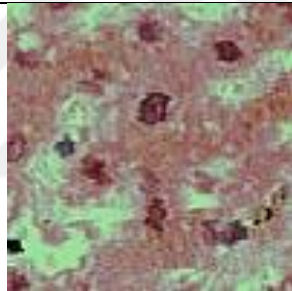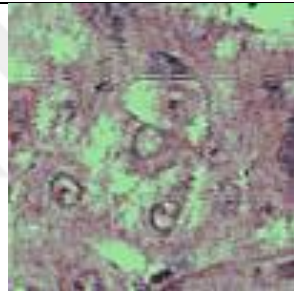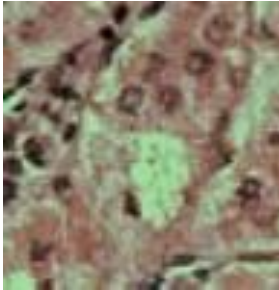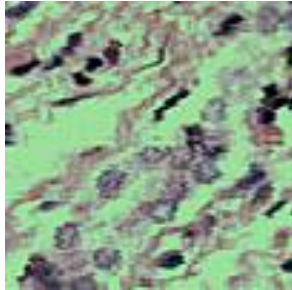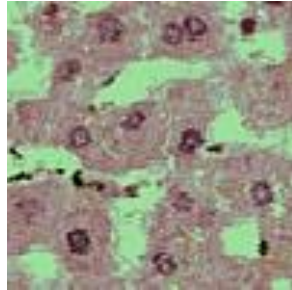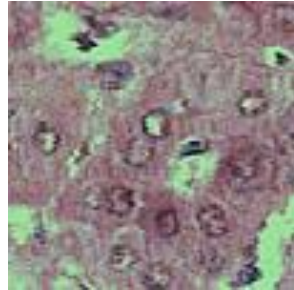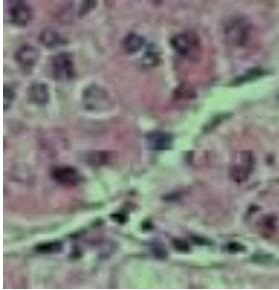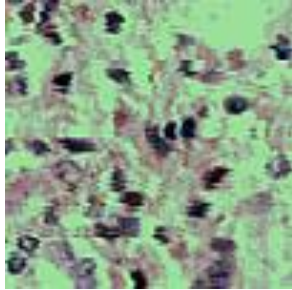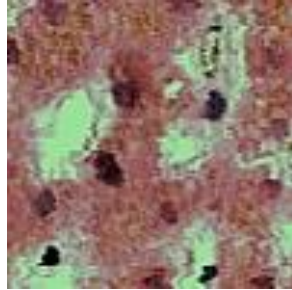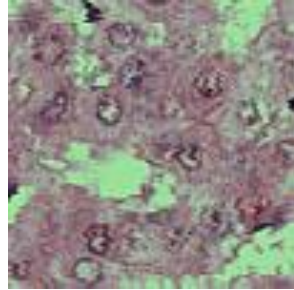
In fact, these differences can vary and may not be present in all samples. While there are obvious differences between normal and HCC liver H&E tissue samples, they also share some similarities due to their origin in the same organ. The similarities can be summarized as follows.

- **Tissue Structure**: Both normal and HCC liver tissues have a similar basic structural layout, with cells organized around the central vein and portal tracts, though these structures may be distorted or irregular in HCC.

- **Hepatocyte Presence**: Both normal and HCC tissues contain hepatocytes, the main cell type of the liver, though these cells may appear differently due to tumor transformation in HCC.

- **Staining Patterns**: Both types of tissues are stained with H&E in a similar way, highlighting the cellular and extracellular components. The hematoxylin stains cell nuclei blue, while eosin stains the cytoplasm and extracellular matrix pink.

- **Cellular Components**: Both types of tissues contain similar cellular components including cell nuclei, cytoplasm, and extracellular matrix.

- **Blood Supply**: Both types of tissues are vascular, with a blood supply necessary for the organ's function. However, the arrangement and structure of blood vessels may be more irregular in HCC.

- **Inflammation**: Both normal and HCC liver tissues can contain inflammatory cells, especially if there is a background of liver disease like cirrhosis.

- **Bile Ducts**: The presence of bile ducts can be observed in both normal and HCC liver tissues, but they can be deformed or proliferative in HCC.

In the process of evaluating our proposed model for HCC classification, we encountered a series of challenging cases that stemmed from the nuanced differences and similarities between HCC and normal liver tissue samples. Our qualitative results in patch level,

presented in Table 18, highlight the challenge of this task. Particularly noteworthy aspects are the visual similarities between true positive patches (correctly identified tumor regions) and false positive patches (non-tumor regions erroneously classified as tumor). Conversely, we also observed cases where true negatives (correctly identified non-tumor regions) were visually similar to false negatives (tumor regions incorrectly classified as non-tumor). This suggests that certain tumor regions may bear a strong resemblance to normal tissue, resulting in misclassifications. These cases suggest that certain regions of normal tissue may closely resemble tumor tissue, causing our model to mistakenly classify them as such. This insight points to the inherent difficulty in accurately distinguishing tumor regions based solely on visual appearance. Therefore, in order to improve the precision, our classification model needs to capture more complex patterns within the hyperspectral data cube while maintaining the validation error at a low level.

Table 18: Qualitative classification results

True Positive	False Positive	True Negative	False Negative

There are 18 HCC tumor and 2 normal liver samples in scene level in our test set, additionally, there are 200 patches cropped from each scene level sample. When we combine the patch level results into scene level results by using a simple scoring scheme such as majority voting, our method can successfully classify all scene level samples. We have presented our worst performing scene level samples, one for each class, in Table 19. As it can be inferred from the scene level results, there are certain regions that are visually very similar in both classes.

Table 19: Qualitative results from scene level. (A) healthy liver sample. (B) tumor liver sample

(A) Healthy liver in scene level, 3 patches are misclassified out of 200 patches, marked with yellow dashed line.
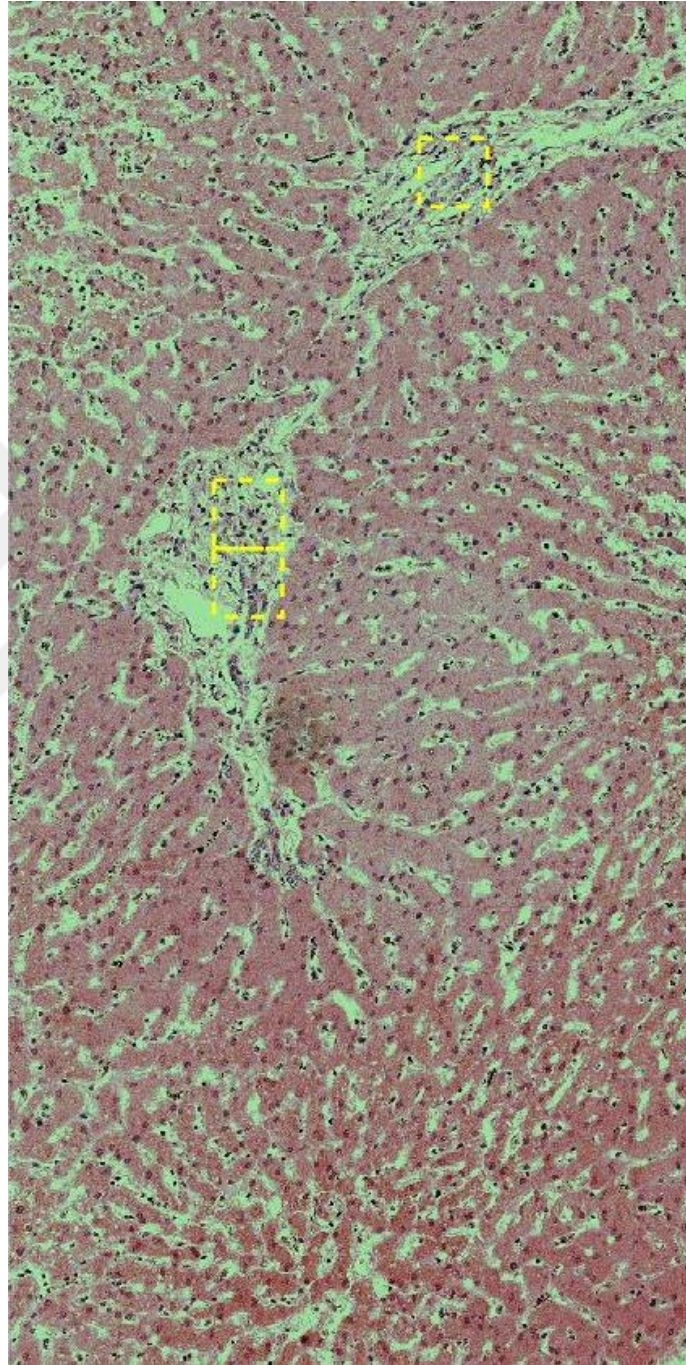
(B) Tumor liver in scene level, 12 patches are misclassified out of 200 patches, marked with yellow dashed line.
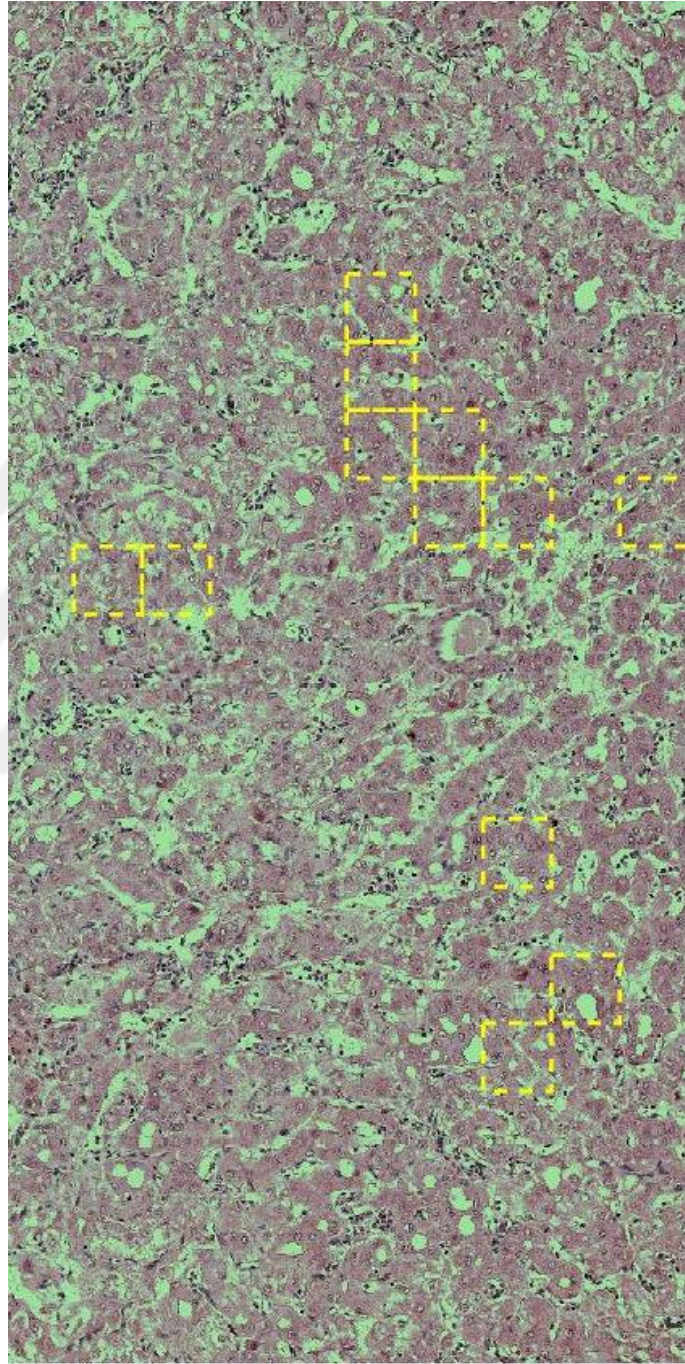
The inherent complexity of liver tissues introduces challenging cases for our CNN learner. This complexity is derived from the structural intricacies of liver tissue at the cellular level, as well as the potential variations in appearance resulting from differing health conditions. That is, normal liver tissues have a distinct cellular structure and arrangement, which can be visibly disrupted in the presence of diseases like HCC. However, other conditions such as inflammation or fibrosis can also result in alterations to the cellular structure of liver tissue. Inflammatory conditions, for instance, can cause tissue swelling

and an influx of immune cells, while fibrosis can lead to the formation of excess fibrous connective tissue, distorting the normal architecture of the liver. Necrosis, the premature death of cells in living tissue, is another condition that can drastically alter tissue appearance. These modifications can complicate the task of differentiating between normal and diseased tissue, as a CNN could mistakenly identify these changes as indicative of HCC. Further complexities arise from the fact that the early stages of HCC can closely resemble normal liver tissue, both visually and structurally. Early-stage HCC is often characterized by small, well-differentiated tumors that can be hard to distinguish from the surrounding liver tissue.

4.6. Comparison with Similar Studies

In this sub-section, we provide a comparative analysis of our experimental results with a similar study [123]. The referenced study used H&E stained HCC liver tissue images to develop a CNN topology for cancer classification. The proposed approach utilizes a custom CNN learner, and it integrates Atrous Spatial Pyramid Pooling (ASPP) blocks to capture multi-scale features present in H&E stained liver histopathology samples. ASPP functions to resample a given feature layer at various rates prior to the convolution process. It operates through the application of multiple filters that probe the sample. This approach allows ASPP to grasp a range of scales of objects along with crucial image context. They derived patches of 224 x 224 from these H&E-stained liver tissues, which were in RGB format and thus contained three bands. Their dataset was composed of 2380 samples from the KMC dataset and 3482 samples from the TCGA dataset, both in patch form after data augmentation. Using their custom CNN topology, the researchers reported an overall accuracy of 90.93%.

In order to perform a fair comparison of their method with ours, we considered several factors. Unfortunately, we were unable to access the exact datasets (KMC and TCGA) that the previous researchers utilized in their study. Consequently, our comparison necessitated the usage of their method on our dataset that also consists of RGB images, since their methodology supports only three-band images. We evaluate the relative performance of their method when applied to a different, though structurally similar, dataset. However, as presented in Table 20, upon the application of the original method, the results of their model fell short of our expectations with an overall accuracy of 0.852. This led us to the process of fine-tuning their method to adapt it to our specific dataset. The original input size utilized by Aatresh et al. was 224x224. We adjusted this parameter and reduced the patch size to 100x100 to better suit the characteristics of our dataset. Beyond this modification, we undertook a parameter tuning process for several hyperparameters within their proposed method. Specifically, we refined the number of filters, batch size, the dropout parameter, and parameters related to ASPP. Following these adjustments, we ran the models again and achieved a better classification performance with their method. However, their methodology did not exceed the classification performance of our method. As presented in Table 20, our method (using RGB data) yielded an accuracy of 0.900, precision of 0.974, recall of 0.914, an F1-score of 0.936,

70

and a MCC value of 0.573. The original method proposed by Aatresh et al., without any refinements, presented an accuracy of 0.852, precision of 0.913, recall of 0.924, an F1-score of 0.881, and an MCC of 0.128. Finally, the refined version of Aatresh et al.'s method, tailored to our dataset, achieved an accuracy of 0.889, precision of 0.913, recall of 0.969, an F1-score of 0.901, and an MCC of 0.202. We attribute our method's superiority to its streamlined CNN topology and 3D convolutions utilized in the model. Additionally, the focal loss function in our classifier has led to focus on hard samples rather than concentrating on only the tumor class. This behavior of model can clearly be seen in the MCC score comparison. Our model has provided much higher MCC score, indicating more robustness to the class imbalance issues.

Table 20: Comparison with Aatresh et al

Model	Accuracy	Precision	Recall	F1-Score	MCC
Proposed Method (RGB)	**0.900**	**0.974**	**0.914**	**0.936**	**0.573**
Aatresh et al (Orig)	0.852	0.913	0.924	0.881	0.128
Aatresh et al (Refined)	0.889	0.913	0.969	0.901	0.202

The second similar study [126] presents an approach for automatic Cholangiocarcinoma (CC) diagnosis using deep learning methodologies on a microscopic hyperspectral pathological dataset. The researchers aimed to exploit fully the spatial-spectral HSI data through a deep CNN. The entire scene is initially divided into several patches, which are then input into the CNN for tumor/non-tumor binary prediction. To leverage the multispectral pixels in the microscopic hyperspectral data, the researchers propose a spectral interval convolution and normalization within the CNN. This approach preserves the spectral information in the channel direction, which is critical for accurate tumor detection. The study utilizes a dataset comprising H&E stained Cholangiocarcinoma and bile duct tissues captured via an Acousto-optic Tunable Filter (AOTF). The images, with 30 hyperspectral bands and patches of size 299x299, consist of 880 sample images, of which 738 contain tumors. The authors utilized two deep learning architectures, ResNet50 and Inception-V3, coupled with interval convolution. This technique involves dividing the spectral dimension into six intervals. Their experiments resulted in an accuracy of 88.20% and 87.90% for ResNet50 and Inception-V3, respectively.

For comparative analysis, while we were unable to access their dataset, we applied their methodology on our HSI dataset. When we tested Sun et al.'s approach on our dataset, we found that their method's performance improved compared to their reported results. It is notable that their original study was based on 30 band hyperspectral images, whereas our dataset contains 270 hyperspectral bands. This provides more depth and complexity to the data and might have contributed to the improved performance of their method on our

dataset. However, our method still outperformed both the Resnet50 and InceptionV3 models of Sun et al. in terms of all metrics. The results are presented in Table 21. Additionally, our method's MCC value was much higher than the competitive methods, indicating our model's versatility to resolve class imbalance problem.

Table 21: Comparison of with Sun, L et al

Model	Accuracy	Precision	Recall	F1-Score	MCC
Proposed Method (HSI)	**0.970**	**0.999**	**0.968**	**0.984**	**0.860**
Sun, L et al (Resnet50)	0.91	0.926	0.978	0.918	0.39
Sun, L et al (InceptionV3)	0.902	0.923	0.972	0.912	0.335

4.7. Experiments with Other Machine Learning Methods

Throughout our research, we have tested numerous machine learning methods with different configurations on our dataset. In this sub-section, we present the machine learning methods different from our main 3D-CNN model that we experimented and found notable to include in this dissertation.

4.7.1. K-Nearest Neighbor

The K-Nearest Neighbors (K-NN) algorithm, considered one of the simplest and highly interpretable machine learning algorithms, belongs to the instance-based learning category of supervised learning techniques [184]. K-NN employs a distance metric such as Euclidean or Manhattan distance to identify the 'K' training examples nearest to an unseen instance. The algorithm then predicts the unseen instance's class based on the most frequent class among these nearest neighbors. The primary advantage of the K-NN algorithm is its simplicity and straightforward implementation. It requires no explicit training phase. Being non-parametric, K-NN does not make any underlying assumptions about the data's distribution, which can be beneficial when the distribution is unknown [185]. The K-NN algorithm is also adaptive to the addition of new data because it utilizes the entire dataset for prediction. It is robust to noisy data due to its reliance on majority voting, and this attribute typically leads to fair performance with large training datasets. However, the K-NN algorithm does present several challenges. K-NN's performance can be adversely affected by redundant or irrelevant features because they contribute equally to the distance calculation. This problem can potentially be alleviated through feature selection or feature weighting strategies [186]. Deciding the optimal 'k' value can also be difficult; a small 'k' can make the algorithm sensitive to noise and outliers, while a large 'k' may excessively smooth out decision boundaries, leading to inaccurate classifications.

Most importantly, the K-NN algorithm is particularly prone to the curse of dimensionality, meaning that the distances in high-dimensional spaces become less meaningful, which can lead to deteriorating performance [187].

We have adapted a pixel-based classification method with KNN approach. The feature vector fed to KNN algorithm contains PCA-9 data and 40 Gabor filter features, resulting in 49 total features per pixel. We have experimented several parameter configurations with KNN. We found out that K=32 with Euclidean distance and uniform weighting performs best for the classification task. The classification result of KNN approach is available in Table 22.

4.7.2. Support Vector Machines

Support Vector Machines (SVMs) are a versatile method for classification and regression, offering an effective technique for handling high-dimensional data. SVMs implement a learning method that performs a binary classification by defining a hyperplane that optimally separates two classes of data in a multidimensional space [188]. The fundamental principle behind SVMs is the construction of an optimal separating hyperplane in a high-dimensional feature space, such that the margin, or the distance between the hyperplane and the nearest data point from either class, is maximized. This hyperplane is constructed with the help of support vectors, which are the data points that lie closest to the decision boundary. By maximizing the margin, SVMs aim to enhance the model's generalization ability and prevent overfitting. One of the significant advantages of SVMs lies in their ability to handle high-dimensional data effectively. This is achieved through the "kernel trick", which allows SVMs to operate in a high-dimensional, implicit feature space without computing the coordinates of the data in that space [189]. Despite their strengths, SVMs do have several limitations. For one, they can be computationally intensive and inefficient to train, particularly for large datasets. The complexity of training an SVM typically scales between quadratic and cubic in the number of samples, which makes it less suitable for very large datasets [190]. Another challenge with SVMs lies in the selection of an appropriate kernel and the tuning of associated parameters, which requires domain expertise and can have a significant impact on the performance of the model.

Similar to the KNN approach, we devised a pixel-level classification method using the SVM model. The feature vector, provided to the SVM algorithm, includes data from PCA-9 and 40 Gabor filter features, culminating in a total of 49 features per pixel. We experimented with multiple parameter configurations within the SVM model, eventually determining that Radial Basis Function (RBF) kernels with C=1.0 and gamma=0.1 were the most effective for this classification task. The results from the KNN approach can be found in Table 22.

4.7.3. Stacked Autoencoders

Stacked Autoencoders (SAE) represent a specific type of autoencoder, which is a neural network model that is utilized for unsupervised learning tasks. The principle goal of an autoencoder is to learn a compressed, distributed representation (encoding) for a set of data, typically for the purpose of dimensionality reduction [191]. A stacked autoencoder consists of multiple layers of autoencoders in which the output of each layer serves as the input for the next. The operation of a stacked autoencoder is divided into two primary phases: encoding and decoding. In the encoding phase, the model learns to reduce the dimensionality of the input data and compress it into an encoded representation. Each layer of the encoder typically applies a transformation function to the input data to reduce its dimensionality. This is generally achieved by using a nonlinear activation function such as a sigmoid or a rectified linear unit (ReLU). The decoding phase, conversely, operates to reconstruct the input data from the encoded representations. The aim is to minimize the reconstruction error, which is usually defined as the difference between the input data and the reconstructed data [192]. One of the main advantages of SAEs is their ability to learn complex, hierarchical representations of the input data. They can be used to learn features from raw input data in an unsupervised manner, which can then be employed to improve the performance of other machine learning tasks such as classification or regression. However, there are also certain drawbacks associated with the use of SAEs. One of the main challenges is that they can be difficult to train, particularly when the network architecture is deep. This is largely due to the problem of vanishing gradients, where the gradients of the loss function become increasingly small as they are propagated back through the layers of the network. This can slow down the learning process and can lead to suboptimal solutions. Moreover, SAEs, like other deep learning models, can be prone to overfitting, especially when the amount of training data is limited [193].

We devised a pixel-level classification method using the SAE model. We utilized HSI dataset with a stacked topology of 100 hidden layers at the first autoencoder and 50 hidden layers at the second autoencoder. The results from the SAE approach can be found in Table 22.

4.7.4. Resnet50

The Residual Network (ResNet) is a popular deep learning model that provided significant contribution to the field of computer vision. ResNet50, specifically, is a variant of the ResNet model that is 50 layers deep, and it is commonly used for a variety of computer vision tasks such as image classification and object detection. The main aspect of ResNet lies in its unique "skip connections" or "shortcut connections" that allow the gradient to be directly backpropagated to earlier layers [194]. In traditional deep learning models, each layer attempts to learn the underlying mapping from inputs to outputs. However, in ResNet, each layer is designed to learn the residual (or difference) between its input and output. This residual learning strategy is effective for avoiding the vanishing gradient problem and helps train very deep networks. One of the significant advantages of ResNet50 is its ability to train extremely deep networks by using residual learning. It helps

to avoid overfitting by reusing previously learned features, thereby reducing the need for a larger number of parameters. The model also addresses the degradation problem (where accuracy saturates and then degrades rapidly) when training deep networks [195]. However, while Resnet50 works well on large and complex datasets, it may not be the best choice for smaller and simpler datasets due to its complexity. When comparing ResNet50 with its 3D variant, ResNet50-3D, the main difference lies in the nature of the data they process. ResNet50-3D extends the original model to handle volumetric (3D) data, making it suitable for tasks such as video processing or medical imaging where temporal or depth information is crucial. This is achieved by replacing the 2D convolutions in the original ResNet50 with 3D convolutions [196].

We utilized Resnet50 and its 3D variant (Resnet50-3D) with our HSI dataset with patch size 100 pixels to train an HCC classification model. The results from the Resnet50 approach can be found in Table 22.

4.7.5. DenseNet121

Densely Connected Convolutional Networks, or DenseNets, represent a significant model in the design of convolutional networks. DenseNet121, in particular, refers to a variant that has 121 layers [197]. DenseNets introduce a unique architecture where each layer is directly connected to every other layer in a feed-forward manner, meaning that each layer receives the feature maps of all preceding layers and passes on its own feature maps to all subsequent layers. This design results in a significant reduction in the number of parameters, and it enhances feature propagation throughout the network. It is also beneficial in strengthening feature reuse, alleviating the vanishing gradient problem, and providing regularization to prevent overfitting. DenseNet121 has shown remarkable performance in many computer vision tasks. Its main advantage is the substantial improvement in computational efficiency due to parameter reduction without compromising the learning capability. The dense connections also enforce feature reuse, which is beneficial for learning compact and robust models [198]. As for the 3D version of DenseNet121, it extends the network to handle 3D data. In DenseNet121-3D, the 2D convolutions are replaced with 3D convolutions to process volumetric data.

We employed DenseNet121 and its 3D variant (DenseNet121-3D) with our HSI dataset with patch size 100 pixels to train an HCC classification model. The results from the DenseNet121 approach can be found in Table 22.

4.7.6. Results and Discussion

The results of the supervised learning models constructed by different machine learning approaches including KNN, SVM, SAE, Resnet50 and DenseNet121are presented in Table 22. For the CNN based topologies (Resnet50 and DenseNet121), we have included both 2D and 3D versions by utilizing 2D and 3D convolutions at distinct configurations.

Table 22: Comparison of other machine learning methods

Model	Accuracy	Precision	Recall	F1-Score	MCC
Proposed Method (3D-CNN)	**0.970**	**0.999**	**0.968**	**0.984**	**0.860**
K-NN	0.805	0.911	0.868	0.855	0.088
SVM	0.875	0.921	0.942	0.897	0.238
SAE	0.894	0.929	0.955	0.911	0.343
Resnet50-2D	0.913	0.946	0.958	0.929	0.491
Resnet50-3D	0.938	0.962	0.969	0.95	0.642
DenseNet121-2D	0.883	0.942	0.927	0.912	0.392
DenseNet121-3D	0.909	0.965	0.933	0.936	0.559

The empirical results presented in the table clearly demonstrate the superior performance of our proposed method, 3D-CNN, over a range of commonly used machine learning and deep learning models for HCC classification. Our method achieved the highest performance across all the key metrics - accuracy (0.970), precision (0.999), recall (0.968), F1-Score (0.984), and MCC (0.860). This high level of performance is indicative of the effectiveness of 3D convolutions in capturing both spatial and spectral features in hyperspectral images, which is critical for accurate HCC classification. KNN model, on the other hand, yielded the lowest results in all metrics. This lower performance could be due to the high dimensional nature of the data, where distance-based methods like K-NN may struggle to identify meaningful patterns. The SVM model exhibited decent performance, scoring higher than K-NN, but still lagged behind our proposed method. This could be attributed to the limitations in SVM's ability to handle high-dimensional hyperspectral data and the requirement for manual feature engineering, which may limit its capacity to capture the complex feature relationships inherent in our dataset. Both the SAE and the 2D versions of the Resnet50 and DenseNet121 models posed moderate performances. Their lower scores, relative to our method, might be attributed to the use of 2D convolutions, which are less capable of capturing the interplay between spatial and spectral features inherent in the hyperspectral cube. The 3D versions of Resnet50 and DenseNet121 showed an improvement over their 2D counterparts, reflecting the added value of considering the spectral dimension in the convolution process. Despite this, their performance still fell short compared to our proposed 3D-CNN model, further emphasizing the effectiveness of our specific design choices, such as the streamlined topology and the use of focal loss function for dealing with class imbalance.

During our study, we have experimented numerous machine learning methods on our dataset in order to devise a robust classification model. We have presented the empirical

results for various classification methods including KNN, SVM, SAE, Resnet50 and Densenet121. We observed that the methods requiring engineered feature vectors including KNN and SVM performs worse than the automatic methods. Manual feature engineering essentially requires domain knowledge and an understanding of which features are relevant for the prediction task. Therefore, the selection and creation of these features are subjective and may not capture all relevant information. In high-dimensional data, such as hyperspectral images, the relationships between variables can be complex and non-linear. Attempting to manually define features in such a context is a challenging task that can easily lead to suboptimal results. This complexity is further amplified in the context of biological data, where interactions among features can be complex and not well-understood. This is why models that rely on manual feature engineering, such as KNN and SVM, may underperform compared to methods that automatically learn features from data, such as deep learning models. Deep learning models such as Resnet50 and DenseNet121 have exhibited lesser performance in our specific context. In fact, those models are known for their robust architecture and proven efficiency in generic image recognition tasks involving hundreds of classes. These models are developed to handle complex image recognition challenges, which implies their complexity is a necessary trait to solve a wide array of tasks. However, this complexity can be a disadvantage when applied to simpler or more specific problems, like classifying H&E stained tissue samples into either HCC or normal. The method we proposed in our study, on the contrary, is a streamlined and optimized model tailored specifically for discerning patterns in H&E stained tissue images. We invested significant effort in fine-tuning and experimenting with our architecture to ensure it is well-suited to the unique characteristics of tissue classification. This approach resulted in a network consisting of just four convolution layers, tailored for our specific task, outperforming the more complex architectures of Resnet50 and DenseNet121. Moreover, these deep network architectures are more susceptible to overfitting, especially when trained on smaller datasets like ours. Overfitting is a common problem in machine learning where a model learns the noise and detailed specifics in the training data to an extent that it negatively impacts its performance on new, unseen data. As our dataset is relatively small, a complex model might fit too well to the training data and fail to generalize the learnt features to unseen data, resulting in poorer performance. Hence, our model showed better results, demonstrating the importance of optimizing model complexity to align with the specificity and size of the task at hand.

CHAPTER 5

CONCLUSION

This study proposes a robust method for detecting HCC tumors using hyperspectral imaging and a custom deep-learning model. A biological tissue imaging system is devised by integrating a VNIR hyperspectral camera with a light microscopy device, which is employed to collect hyperspectral images of tumor and healthy liver tissues. A custom 3D-CNN classification topology is designed to fully utilize the potential of HSI data. The topology includes four convolution layers with max-pooling and batch normalization layers between them to down-sample the data and reduce model complexity. The use of 3D convolution layers enables both textural and spectral features to be leveraged in a single training pipeline. This way, we eliminate the need for separate feature extraction stage. The network topology is optimized by replacing the traditional cross-entropy cost function with the focal loss cost function, which significantly overcomes the class imbalance problem in the dataset. The proposed method is shown to be effective in detecting HCC tumors, and the results demonstrate the potential of hyperspectral imaging and deep learning in the field of biomedical imaging.

A significant contribution of this research lies in the successful construction of a biological tissue image capture system within our laboratory environment. This system's design and implementation have enabled us to capture high-resolution hyperspectral images of biological tissues, substantially enhancing the quality of histopathological assessments. Hyperspectral imaging technique captures light across a broad spectrum, resulting in multidimensional, high-resolution images that reveal more detailed information about tissue structures. Consequently, the utilization of the hyperspectral data has the potential to improve diagnostic accuracy significantly. Furthermore, the system presents an affordable alternative to commercial systems that are often financially prohibitive for many research labs and medical institutions. Therefore, our system not only advances the field through enhanced imaging capabilities, but also broadens the potential user base for hyperspectral imaging in histopathological assessments by providing a more economically viable solution. This cost-effective approach opens doors for smaller or economically disadvantaged institutions to benefit from hyperspectral imaging's advanced capabilities, thus democratizing access to high-quality, detailed biological tissue imaging.

The use of hyperspectral data in our research has played a pivotal role in enhancing the classification performance of our machine learning models. The vast majority of CAD studies rely on RGB data obtained using CCD or CMOS sensors [199-201]. However, our study employs a more detailed range of the electromagnetic spectrum. Specifically, we utilize a hyperspectral dataset that consists of 270 contiguous bands spanning from 400 to 800 nm in the spectrum. In comparison, the RGB dataset only includes three individual bands at 630, 540, and 480 nm. With hyperspectral imaging, we are able to analyze the

chemical composition of the subject material in addition to conventional spatial attributes like size, shape, and texture. This abundance of spectral information presents a significant advantage as it allows for the capture of a more comprehensive representation of tissue samples, which in turn, contributes to a more accurate and precise classification.

The hyperspectral cube is ideal for our classification task since it can capture the fine details of material properties as spectral signatures. Different from RGB dataset, the descriptive features along the spectral dimension can be effectively captured using a 3D convolution operation. Our supervised learner, based on 3D CNN is capable of capturing the complex relationships between the features in both spectral and spatial dimensions. Specifically, features such as corners, edges, and textures in the spatial plane can be associated with peaks, dips, slopes, and valleys in the spectral signatures of pixels. By adopting a 3D CNN approach, we were able to consolidate both spectral and spatial feature extraction into a single convolution operation. This is a major improvement from previous studies requiring separate steps for spectral and spatial feature extraction. Such an integrated approach not only streamlines the feature extraction process but also allows for the extraction of hybrid spectral-spatial features that better capture the complexities of hyperspectral data. One of the key advantages of our approach is that it eliminates the need for manual feature engineering or pre-processing steps, which are often time-consuming and can introduce bias or inaccuracies. Instead, our method utilizes 3D convolutions that automatically learn the most discriminative features from the hyperspectral data. By feeding hyperspectral data directly into our 3D CNN model, we enable the model to learn directly from the spectral-spatial complexities of the tissue samples, thus ensuring a more accurate and robust representation of the data. Moreover, our method is characterized by a simpler network topology compared to general purpose image recognition CNN topologies. By adopting a streamlined CNN topology and incorporating 3D convolutions, our model is less complex and easier to train, yet capable of capturing the complex spectral-spatial interactions present in hyperspectral data. This not only reduces the computational resources required for model training but also makes the model less prone to overfitting, thus resulting in more reliable and generalizable results.

With the help of the vast amount of information present in the hyperspectral cube, our deep learning model can create a strong classifier with its strong feature extraction capabilities. Additionally, we have observed that by subsampling the original hyperspectral dataset into lower dimension datasets, such as HSI-90, HSI-30, and HSI-10, compared to RGB dataset, we can improve classification accuracy by introducing more spectral bands to the model. This suggests that the deep learning model's predictive power can be boosted by utilizing more spectral bands in classification. In addition to using hyperspectral imaging, we explored the effectiveness of principal component analysis for reducing the dimensionality of our original hyperspectral dataset containing 270 bands. We generated two new datasets using PCA, one with nine principal components (PCA-9) and another with three principal components (PCA-3). The PCA method has the advantage of reducing data complexity and improving the signal-to-noise ratio, making it easier for the learner to converge. By operating 3D convolutions on PCA data, our learner

captures patterns across the different PCA bands as well as spatial patterns within each band, as coined by previous studies [202-206]. Our CNN models trained with PCA data showed lower classification accuracy than the CNN model trained with the original HSI dataset. The PCA-9 dataset, which had a maximum variance of 98.60%, performed almost as well as the HSI dataset. Considering the simplicity of PCA-9 compared to the original HSI dataset, PCA provides a cost-effective way to utilize hyperspectral data for our classification task. However, it's important to note that PCA, being a linear transformation technique, may not always capture the full complexity of the spectral signatures present in the hyperspectral data. Furthermore, the PCA-3 dataset outperformed the RGB dataset, suggesting that compressing hyperspectral data into three bands contains more useful information for classifying tissue samples than RGB. Our experimental results demonstrate that the HSI dataset is more useful for accurate classification than its RGB and PCA counterparts.

Class imbalance, where the number of samples in different classes are significantly disproportionate, is a common challenge in machine learning and particularly in medical imaging tasks. In our dataset, there are 54 tumor tissue samples and 6 healthy tissue samples. In other words, only 10% of the data belongs to the healthy class. Running a malformed classifier predicting tumor class for all the given samples would result in 90% accuracy for our dataset. In order to identify such anomalies in the classification performance we have utilized MCC metric in addition to traditional metrics such as accuracy, precision and recall. MCC considers all four values of the confusion matrix (true positives, false positives, true negatives, and false negatives) and generates a score that lies between -1 and 1. A value of +1 represents a perfect prediction, 0 indicates no better than random prediction, and -1 demonstrates total disagreement between prediction and the actual class. The reason MCC is particularly suited to our dataset is that it considers both under-predicted and over-predicted classes at the same time. In our case, classification errors often appeared as false positives, in fact, healthy tissues classified as tumor. In our work, we tackled the class imbalance problem effectively by utilizing the focal loss function instead of the conventional cross-entropy loss function. The focal loss function is specifically designed to increase the model's sensitivity to hard-to-classify instances. It does this by down-weighting the contribution of easy-to-classify examples and focusing more on the misclassified samples. This approach is particularly beneficial when dealing with imbalanced datasets, as it encourages the model to pay more attention to the underrepresented class that typically contains harder to classify instances. In contrast, the traditional cross-entropy loss function treats all errors equally, regardless of whether they come from majority or minority classes. This can result in a model that is biased towards the majority class and performs poorly on the minority class. The hyperparameters of the focal loss function, γ (focusing parameter) and α (weighting factor), are empirically optimized for each experiment configuration. It is worth noting that the γ parameter has a critical impact on the classification performance, while α has a minor effect on the results. By using focal loss instead of cross-entropy loss, we were able to more effectively handle the class imbalance in our dataset, resulting in a robust learner that performs well across all classes.

Our work in HCC classification resonates with two comparable studies in the literature, even though with key differences in methodologies and results. One study [123] utilized RGB histopathological images to achieve HCC classification. The competitive study employed a custom CNN topology with ASPP blocks. The other study [126] was based on 30 bands hyperspectral imaging of bile duct tissues. The authors have proposed two distinct classification models based on two popular topologies (Resnet50 and Inceptionv3) and interval convolutions. We have implemented the competitive methods and run their algorithms on our own datasets. The empirical results confirm the superiority of our proposed method over the previous studies. In fact, 3D convolutions and focal loss function are the key elements in our methodology to achieve high accuracy and MCC scores. The competitive studies are lack of essential modules to deal with class imbalance. Furthermore, as the experimental results imply, their feature extraction layers (ASPP and interval convolution) are not functioning to collect sufficient descriptive features from the hyperspectral cube.

Limitations

Despite the promising results obtained from our proposed 3D-CNN classification model, our study has some limitations that need to be addressed. First, the dataset used in our study needs to be expanded by incorporating more tissue samples. By increasing the amount of data used during the training phase, we expect to obtain a classifier that is less susceptible to overfitting and has a greater ability to generalize. It is advisable to perform additional validation of our model using a larger tissue sample dataset collected from various laboratories and labeled by different pathologists. This approach would increase the variability of the dataset and enable the classification model to cover a larger area in the solution space. This diversification would better represent the broad spectrum of histopathological variations seen in real-world settings.

The system we have developed has promising potential applications not only in aiding HCC diagnosis but also in several other areas of pathology. Given its ability to accurately classify tissue samples, it can be used to support pathologists during their examinations. By providing a quantitative, data-driven second opinion, our system could help reduce the risk of human error and facilitate more accurate diagnoses. Furthermore, our model could be an invaluable resource for training purposes. Medical students and early-career pathologists could use the system to learn about and familiarize themselves with the features of HCC and normal liver tissues. It can provide instant feedback, allowing users to self-correct and reinforce learning effectively. Similarly, the system could be integrated into the initial screening process. The ability of our model to analyze hyperspectral images and swiftly categorize them as 'normal' or 'tumor' could expedite the preliminary screening process. By identifying suspicious cases early on, the system could facilitate timely intervention and potentially improve patient outcomes. However, it should be noted that while our model can be a powerful tool, it is not intended to replace human expertise. Rather, it should be used to supplement and support the decision-making process, ultimately leading to better clinical outcomes. While our system can effectively differentiate between HCC and normal liver tissues based on hyperspectral images, it

currently does not have the ability to inspect other essential attributes that a trained pathologist would consider in a comprehensive diagnosis. For instance, the system is not capable of assessing features such as inflammation, necrosis, and the presence or condition of blood vessels. These elements are essential for a complete and accurate evaluation of liver tissues and are routinely considered by pathologists during examination. These limitations highlight the current confines of our system and underline the necessity of the human element in the pathology field. Trained pathologists have a solid understanding and expertise that encompasses a wider range of pathological indicators beyond the simple HCC versus normal tissue binary. As it stands, our system is a tool to assist and enhance the work of pathologists rather than a replacement.

Future Work

Our model's utility extends beyond its current application in classifying HCC and normal liver tissues. In fact, the methodologies we have developed could be broadly applicable to a variety of other tissue types such as lung, breast, brain, colon and cervix. The principles and methodologies that our study embodies have the potential to make significant contributions to the broader field of histopathology and cancer research, demonstrating the potential to transform diagnostic processes across a spectrum of tissue types and diseases.

Generative Adversarial Networks (GANs), composed of two neural networks, a generator and a discriminator, can be used to generate synthetic images of pathological tissues, which can be valuable for augmenting our dataset and improving the robustness of our model. That is, GANs can be employed to generate synthetic hyperspectral images of tissue samples that closely represent real samples. This could help in overcoming the limitations posed by a shortage of certain types of tissue samples and lead to more robust and generalizable models [207-208]. In order to develop new image classification approaches within hyperspectral histopathology field, there are a few promising alternatives that we spot. For instance, the image classification task can be modelled as an object detection task, and the methods for identifying tumor regions from scene level images can be studied. For this purpose, You Only Look Once (YOLO) models can help to improve the speed and accuracy of histopathological assessments [209]. By identifying and isolating regions of interest within tissue samples, YOLO could significantly enhance the efficiency of our classification process. This is especially important in pathological analysis, where quick and accurate identification of cancerous regions can significantly improve patient prognosis. YOLO's real-time object detection capabilities could also aid in automated biopsy targeting, thereby increasing the precision of tissue sampling and reducing potential harm to patients. On the other hand, the incorporation of the attention mechanism into our computational model has the potential to enhance its performance by facilitating targeted processing of input tissue samples. This mechanism offers the model the capability to prioritize specific regions in the hyperspectral images, especially those regions bearing the most diagnostic relevance. Consequently, the overall effectiveness of our model can be improved as it is enabled to focus more on these informative and critical areas in the hyperspectral images, rather than processing the entire image uniformly.

Complementing the attention mechanism, Vision Transformers (ViT) present a promising technique in image classification tasks. ViT can effectively capture long-range dependencies within the hyperspectral cube by establishing correlations between both spatial and spectral features in a compact and efficient manner. The utilization of ViT in our dataset could thus pave the way for a more comprehensive analysis of hyperspectral images. By this way, the learner takes into account the interconnectedness of spatial and spectral features and the model's ability to differentiate between normal and diseased tissues will be streamlined. [210-211]. The self-attention mechanism, employed in ViT, provides a global view of the tissue sample, enabling the model to consider relationships and dependencies between distant pixels. Incorporating ViT into our methodology has the potential to enhance the depth of feature understanding, by integrating the spatial and spectral inter-dependencies that traditional CNNs might overlook.
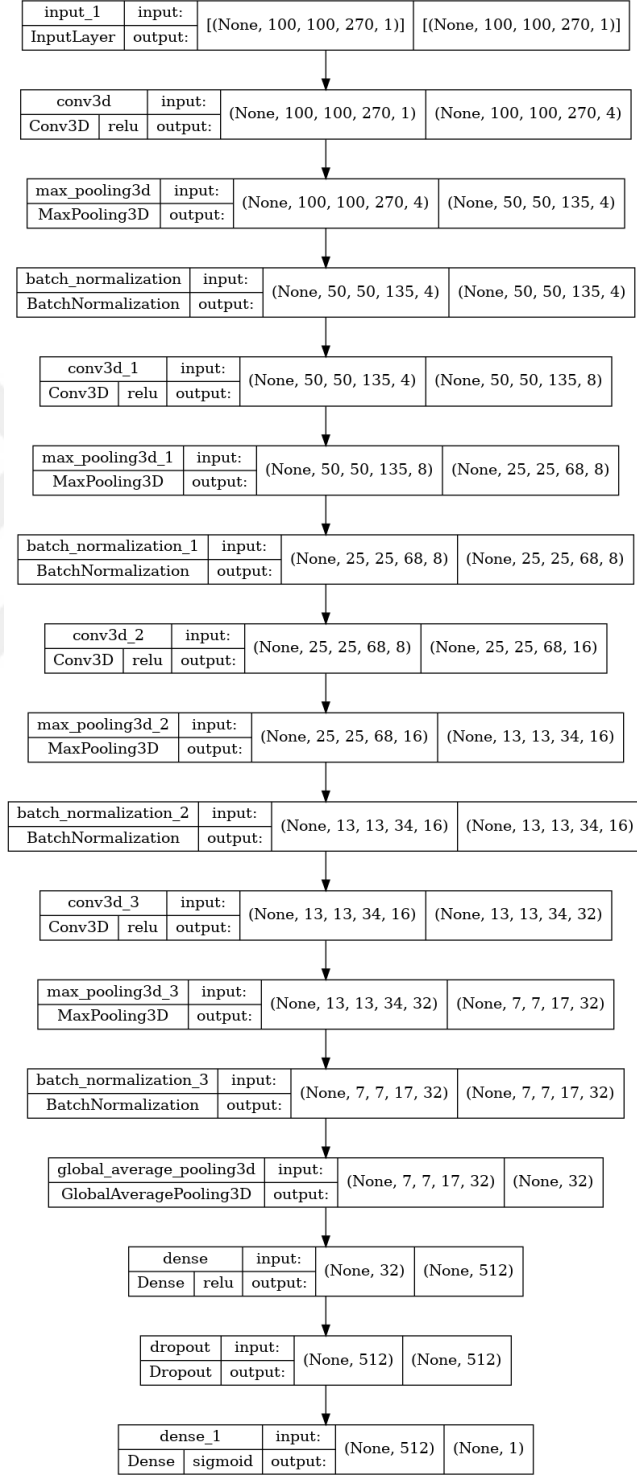
APPENDICES

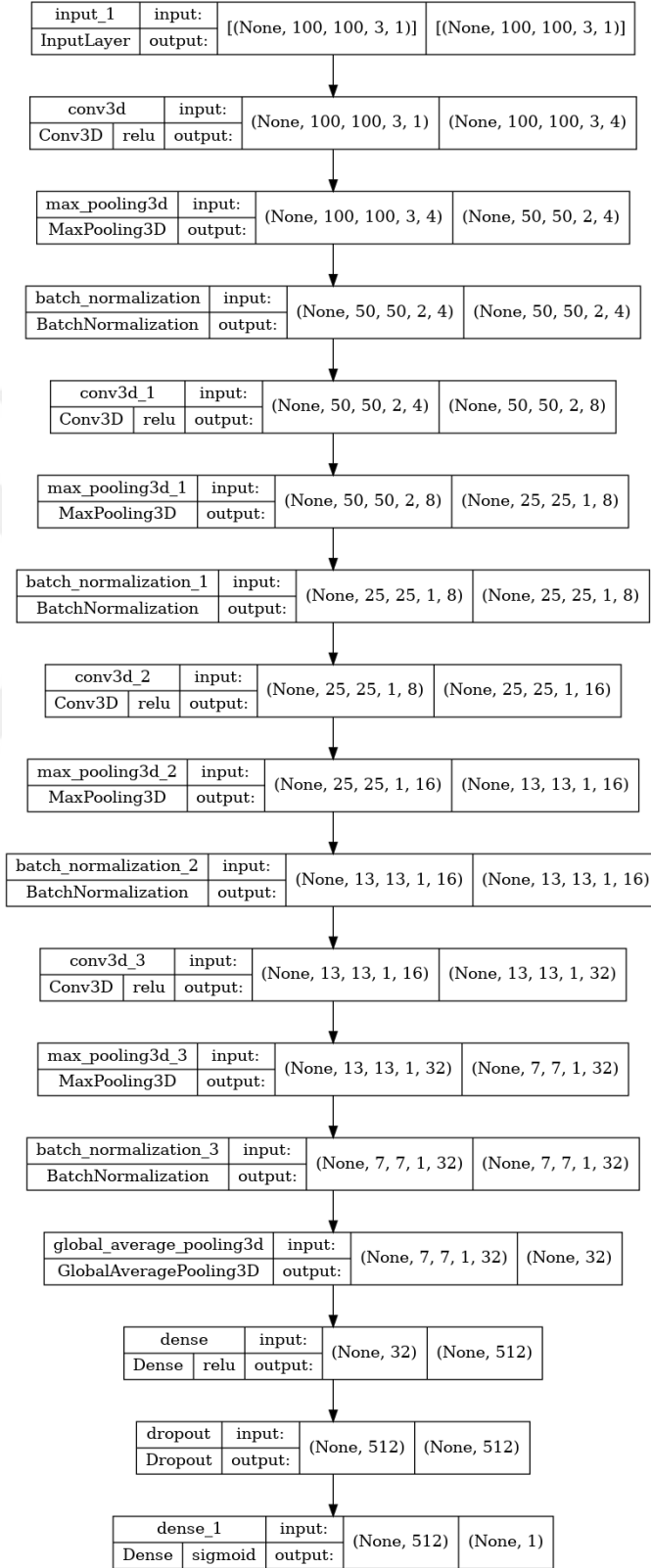
Figure A1. Topology of 3D-CNN utilized for HSI data.

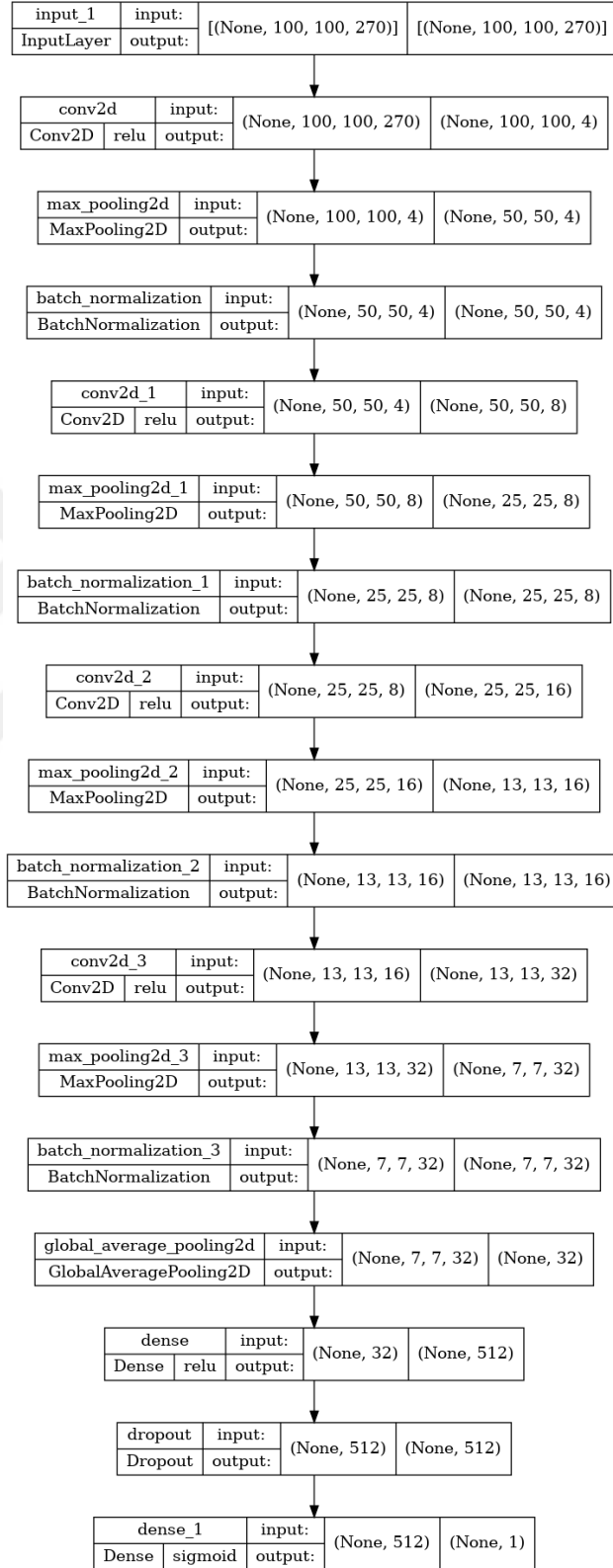**Figure A2.** Topology of 3D-CNN utilized for Three-Channel data.

86

Figure A3. Topology of 2D-CNN utilized for HSI data.

REFERENCES

[1] M. Govender, K. Chetty, and H. Bulcock, "A review of hyperspectral remote sensing and its application in vegetation and water resource studies," Water SA, vol. 33, no. 2, pp. 145–151, 2007.

[2] T. M. Lillesand, R. W. Kiefer, and J. W. Chipman, Remote Sensing and Image Interpretation. John Wiley & Sons, 2014.

[3] Gu, Yanfeng, Chen Wang, and Xian Li. "An Intensity-Independent Stereo Registration Method of Push-Broom Hyperspectral Scanner and LiDAR on UAV Platforms." IEEE Transactions on Geoscience and Remote Sensing 60 (2022): 1-14.

[4] Goetz, A. F. H., Vane, G., Solomon, J. E., & Rock, B. N. (1985). Imaging Spectrometry for Earth Remote Sensing. Science, 228(4704), 1147-1153.

[5] Thenkabail, P. S., Lyon, J. G., & Huete, A. (Eds.). (2011). Hyperspectral Remote Sensing of Vegetation. CRC Press.

[6] Chang, C. I. (Ed.). (2003). Hyperspectral Imaging: Techniques for Spectral Detection and Classification. Kluwer Academic Publishers.

[7] Zabalza, Jaime. (2015). Feature extraction and data reduction for hyperspectral remote sensing Earth observation.

[8] J. R. Goldblum, L. W. Lamps, and J. L. McKenney, Rosai and Ackerman's Surgical Pathology, 11th ed. Elsevier, 2018.

[9] S. Suster and C. Moran, Histopathology Specimens: Clinical, Pathological and Laboratory Aspects. Springer, 2006.

[10] Burchette, Jim. "Theory and Practice of Histological Techniques." (2009): 514.

[11] D. L. Longo and A. T. Kasper, Harrison's Principles of Internal Medicine, 19th ed. McGraw Hill Education, 2015.

[12] More, S. S., Beach, J. M., McClelland, C., Mokhtarzadeh, A., & Vince, R. (2019). In vivo assessment of retinal biomarkers by hyperspectral imaging: early detection of Alzheimer's disease. ACS chemical neuroscience, 10(11), 4492-4501.

[13] Bajema, I. M., Hagen, E. C., Hansen, B. E., Hermans, J., Noël, L. H., Waldherr, R., ... & Bruijn, J. A. (1996). The renal histopathology in systemic vasculitis: an international survey study of inter-and intra-observer agreement. Nephrology Dialysis Transplantation, 11(10), 1989-1995.

[14] L. Pantanowitz, A. J. Glassy, A. J. Valenstein, and J. E. Henricks, "Digital pathology: Current status and future perspectives," Histopathology, vol. 61, no. 1, pp. 1-9, 2012.

[15] Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians, 71(3), 209-249.

[16] Professional, C. C. M. (n.d.). Hepatocellular Carcinoma (HCC). Cleveland Clinic. https://my.clevelandclinic.org/health/diseases/21709-hepatocellular-carcinoma-hcc

[17] R. Lencioni and J.M. Llovet, "Modified RECIST (mRECIST) Assessment for Hepatocellular Carcinoma," Seminars in Liver Disease, vol. 30, no. 01, pp. 052-060, 2010.

[18] Vij M, Calderaro J. Pathologic and molecular features of hepatocellular carcinoma: An update. World J Hepatol 2021; 13(4): 393-410 [PMID: 33959223 DOI: 10.4254/wjh.v13.i4.393]

[19] Yang, J. D., and H. H. Roberts, "Hepatocellular carcinoma: A global view," Nature Reviews Gastroenterology & Hepatology, vol. 7, no. 8, pp. 448-458, Aug. 2010.

[20] Wu, Bocheng & Sodji, Quaovi & Oyelere, Adegboyega. (2022). Inflammation, Fibrosis and Cancer: Mechanisms, Therapeutic Options and Challenges. Cancers. 14. 552. 10.3390/cancers14030552.

[21] Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. Cell, 144(5), 646-674.

[22] O'Neil, Maura, Ivan Damjanov, and Ryan M. Taylor. Liver pathology for clinicians. Springer, 2015.

[23] Stevens, W. Ross; Gulino, Samuel P.; Batts, Kenneth P.; Stephens, David H.; Johnson, C. Daniel. Mosaic Pattern of Hepatocellular Carcinoma: Histologic Basis for a Characteristic CT Appearance. Journal of Computer Assisted Tomography 20(3):p 337-342, May 1996.

[24] Bishayee, A., & Darvesh, A. S. (2012). Angiogenesis in hepatocellular carcinoma: a potential target for chemoprevention and therapy. Current cancer drug targets, 12(9), 1095-1118.

[25] Fattovich, G., Stroffolini, T., Zagni, I., & Donato, F. (2004). Hepatocellular carcinoma in cirrhosis: incidence and risk factors. Gastroenterology, 127(5), S35-S50.

[26] Harris, A. L. (2002). Hypoxia—a key regulatory factor in tumor growth. Nature reviews cancer, 2(1), 38-47.

[27] Carmeliet, P. (2005). Angiogenesis in life, disease and medicine. Nature, 438(7070), 932-936.

[28] Kuraishy, A., Karin, M., & Grivennikov, S. I. (2011). Tumor promotion via injury- and death-induced inflammation. Immunity, 35(4), 467-477.

[29] National Cancer Institute, Normal And Cancer Cells Structure - NCI Visuals Online. (n.d.). https://visualsonline.cancer.gov/details.cfm?imageid=2512

[30] Eldridge, L., MD. (2023). Cancer Cells vs. Normal Cells: How Are They Different? Verywell Health. https://www.verywellhealth.com/cancer-cells-vs-normal-cells-2248794

[31] Bruix J., Reig M., Sherman M. (2016). Evidence-Based Diagnosis, Staging, and Treatment of Patients With Hepatocellular Carcinoma. Gastroenterology. 150(4):835–853.

[32] Yao F.Y., Ferrell L., Bass N.M., Watson J.J., Bacchetti P., Venook A., Ascher N.L., Roberts J.P. (2001). Liver transplantation for hepatocellular carcinoma: Expansion of the tumor size limits does not adversely impact survival. Hepatology. 33(6):1394–1403.

[33] Llovet, J. M., and J. Bruix, "Systematic review of randomized trials for unresectable hepatocellular carcinoma: Chemoembolization improves survival," Hepatology, vol. 37, no. 2, pp. 429-442, Feb. 2003.

[34] Llovet J.M., Real M.I., Montaña X., Planas R., Coll S., Aponte J., Ayuso C., Sala M., Muchart J., Solà R., Rodés J., Bruix J. (2002). Arterial embolisation or chemoembolisation versus symptomatic treatment in patients with unresectable hepatocellular carcinoma: A randomised controlled trial. The Lancet. 359(9319):1734–1739.

[35] Kudo M. (2018). Systemic Therapy for Hepatocellular Carcinoma: Latest Advances. Cancers. 10(11):412.

[36] Forner A., Reig M., Bruix J. (2018). Hepatocellular carcinoma. The Lancet. 391(10127):1301–1314.

[37] Heimbach, J. K., Kulik, L. M., Finn, R. S., Sirlin, C. B., Abecassis, M. M., Roberts, L. R., ... & Marrero, J. A. (2018). AASLD guidelines for the treatment of hepatocellular carcinoma. Hepatology, 67(1), 358-380.

[38] Caturelli, E., Solmi, L., Anti, M., Fusilli, S., Roselli, P., Andriulli, A., Fornari, F., & De Sio, I. (2004). Ultrasound guided fine needle biopsy of early hepatocellular carcinoma complicating liver cirrhosis: a multicentre study. Gut, 53(9), 1356–1362.

[39] Llovet, J. M., Ricci, S., Mazzaferro, V., Hilgard, P., Gane, E., Blanc, J. F., De Oliveira, A. C., Santoro, A., Raoul, J. L., Forner, A., Schwartz, M., Porta, C., Zeuzem, S., Bolondi, L., Greten, T. F., Galle, P. R., Seitz, J. F., Borbath, I., Häussinger, D., Giannaris,

T., Shan, M., Moscovici, M., Voliotis, D., & Bruix, J. (2008). Sorafenib in advanced hepatocellular carcinoma. New England Journal of Medicine, 359(4), 378–390.

[40] Komuta, M., Govaere, O., Vandecaveye, V., Akiba, J., Van Steenbergen, W., Verslype, C., ... & Roskams, T. (2012). Histological diversity in cholangiocellular carcinoma reflects the different cholangiocyte phenotypes. Hepatology, 55(6), 1876-1888.

[41] Cancio, L. C., Batchinsky, A. I., Mansfield, J. R., Panasyuk, S., Hetz, K., Martini, D., ... & Freeman, J. E. (2006). Hyperspectral imaging: a new approach to the diagnosis of hemorrhagic shock. Journal of Trauma and Acute Care Surgery, 60(5), 1087-1095.

[42] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436-444, 2015.

[43] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, 2015, pp. 234-241.

[44] Fan, L., Mac, M. T., Frishberg, D. P., Fan, X., Dhall, D., Balzer, B. L., ... & Wang, H. L. (2010). Interobserver and intraobserver variability in evaluating vascular invasion in hepatocellular carcinoma. Journal of gastroenterology and hepatology, 25(9), 1556-1561.

[45] Brill, M. H. (2021). Trichromatic Theory. In Computer Vision: A Reference Guide (pp. 1288-1291). Cham: Springer International Publishing.

[46] Niall, K. K. (1988). On the trichromatic and opponent-process theories: An article by E. Schrödinger. Spatial Vision, 3(2), 79-95.

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

[48] Vila-Francés, J. Analysis of Acousto-Optic Tunable Filter Performance for Imaging Applications. Opt. Eng. 2010, 49, 113203. https://doi.org/10.1117/1.3509243.

[49] Calpe-Maravilla, J. 400– to 1000–Nm Imaging Spectrometer Based on Acousto-Optic Tunable Filters. J. Electron. Imaging 2006, 15, 023001. https://doi.org/10.1117/1.2201057.

[50] Xu, Z.; Zhao, H.; Jia, G.; Sun, S.; Wang, X. Optical Schemes of Super-Angular AOTF-Based Imagers and System Response Analysis. Opt. Commun. 2021, 498, 127204. https://doi.org/10.1016/j.optcom.2021.127204.

[51] Santos, M. K., Ferreira, J. R., Wada, D. T., Tenório, A. P. M., Barbosa, M. H. N., & Marques, P. M. D. A. (2019). Artificial intelligence, machine learning, computer-aided

diagnosis, and radiomics: advances in imaging towards to precision medicine. Radiologia brasileira, 52, 387-396.

[52] Yanase, J., & Triantaphyllou, E. (2019). A systematic survey of computer-aided diagnosis in medicine: Past and present developments. Expert Systems with Applications, 138, 112821.

[53] Fujita, H. (2020). AI-based computer-aided diagnosis (AI-CAD): the latest review to read first. Radiological physics and technology, 13(1), 6-19.

[54] Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., ... & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. Medical Image Analysis, 42, 60-88.

[55] Abbas, Abbas K., and Rasha Bassam. "Phonocardiography signal processing." Synthesis Lectures on Biomedical Engineering 4.1 (2009): 1-194.

[56] Acharya, U. R., Fujita, H., Lih, O. S., Hagiwara, Y., Tan, J. H., & Adam, M. (2017). Automated detection of arrhythmias using different intervals of tachycardia ECG segments with convolutional neural network. Information sciences, 405, 81-90.

[57] Abbas, Qaisar, M. Emre Celebi, and Irene Fondón. "Computer-aided pattern classification system for dermoscopy images." Skin Research and Technology 18.3 (2012): 278-289.

[58] Anthimopoulos, M., Christodoulidis, S., Ebner, L., Christe, A., & Mougiakakou, S. (2016). Lung pattern classification for interstitial lung diseases using a deep convolutional neural network. IEEE transactions on medical imaging, 35(5), 1207-1216.

[59] Ayer, T., Chhatwal, J., Alagoz, O., Kahn Jr, C. E., Woods, R. W., & Burnside, E. S. (2010). Comparison of logistic regression and artificial neural network models in breast cancer risk estimation. Radiographics, 30(1), 13-22.

[60] Fazakis, N., Kocsis, O., Dritsas, E., Alexiou, S., Fakotakis, N., & Moustakas, K. (2021). Machine learning tools for long-term type 2 diabetes risk prediction. IEEE Access, 9, 103737-103757.

[61] Freer TW, Ulissey MJ. Screening mammography with computer aided detection: prospective study of 12,860 patients in a community breast center. Radiology. 2001;220(3):781–6.

[62] Fenton JJ, Taplin SH, Carney PA, Abraham L, Sickles EA, D'Orsi C, Berns EA, Cutter G, Hendrick RE, Barlow WE, Elmore JG. Influence of computer-aided detection on performance of screening mammography. N Engl J Med. 2007;356(14):1399–409.

[63] Lehman CD, Wellman RD, Buist DS, Kerlikowske K, Tosteson AN, Miglioretti DL. Breast cancer surveillance consortium. Diagnostic accuracy of digital screening mammography with and without computer-aided detection. JAMA Intern Med. 2015;175(11):1828–37.

[64] Frid-Adar M, Diamant I, Klang E, Amitai M, Goldberger J, Greenspan H. GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. Neurocomputing. 2018;321:321–31.

[65] Ardila D, Kiraly AP, Bharadwaj S, Choi B, Reicher JJ, Peng L, Tse D, Etemadi M, Ye W, Corrado G, Naidich DP, Shetty S. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. Nat Med. 2019;25:954–61.

[66] Ma, L., Tan, G., Luo, H., Liao, Q., Li, S., & Li, K. (2022). A Novel Deep Learning Framework for Automatic Recognition of Thyroid Gland and Tissues of Neck in Ultrasound Image. IEEE Transactions on Circuits and Systems for Video Technology, 32(9), 6113-6124.

[67] Shin, H., Choi, G. S., Shon, O. J., Kim, G. B., & Chang, M. C. (2022). Development of convolutional neural network model for diagnosing meniscus tear using magnetic resonance image. BMC Musculoskeletal Disorders, 23(1), 1-9.

[68] El-Dahshan, E. S. A., Mohsen, H. M., Revett, K., & Salem, A. B. M. (2014). Computer-aided diagnosis of human brain tumor through MRI: A survey and a new algorithm. Expert systems with Applications, 41(11), 5526-5545.

[69] Cortese, L., Aranda, G., Buttafava, M., Contini, D., Dalla Mora, A., de Fraguier, S., ... & Durduran, T. (2018). The LUCA project-Laser and Ultrasound Co-Analyzer for Thyroid Nodules: Overview and Current Status. Optical Tomography and Spectroscopy, OTh4D-5.

[70] Pantanowitz, L., Sharma, A., Carter, A. B., Kurc, T., Sussman, A., & Saltz, J. (2018). Twenty years of digital pathology: An overview of the road travelled, what is on the horizon, and the emergence of vendor-neutral archives. Journal of Pathology Informatics, 9, 40.

[71] Elmore, J. G., Longton, G. M., Carney, P. A., Geller, B. M., Onega, T., Tosteson, A. N., ... & Nelson, H. D. (2015). Diagnostic concordance among pathologists interpreting breast biopsy specimens. JAMA, 313(11), 1122-1132.

[72] Gurcan, M. N., Boucheron, L. E., Can, A., Madabhushi, A., Rajpoot, N. M., & Yener, B. (2009). Histopathological image analysis: A review. IEEE Reviews in Biomedical Engineering, 2, 147-171.

[73] Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., ... & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. Medical Image Analysis, 42, 60-88.

[74] Sun, H., Zeng, X., Xu, T., Peng, G., & Ma, Y. (2019). Computer-aided diagnosis in histopathological images of the endometrium using a convolutional neural network and attention mechanisms. IEEE journal of biomedical and health informatics, 24(6), 1664-1676.

[75] Meng, Z., Zhao, Z., Li, B., Su, F., & Guo, L. (2021). A cervical histopathology dataset for computer aided diagnosis of precancerous lesions. IEEE Transactions on Medical Imaging, 40(6), 1531-1541.

[76] Budak, Ü., Cömert, Z., Rashid, Z. N., Şengür, A., & Çıbuk, M. (2019). Computer-aided diagnosis system combining FCN and Bi-LSTM model for efficient breast cancer detection from histopathological images. Applied Soft Computing, 85, 105765.

[77] Yang, H., Kim, J. Y., Kim, H., & Adhikari, S. P. (2019). Guided soft attention network for classification of breast cancer histopathology images. IEEE transactions on medical imaging, 39(5), 1306-1315.

[78] Xu, Y., Jia, Z., Wang, L. B., Ai, Y., Zhang, F., Lai, M., & Chang, E. I. (2017). Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features. BMC bioinformatics, 18(1), 1-17.

[79] Xu, Y., Zhu, J. Y., Eric, I., Chang, C., Lai, M., & Tu, Z. (2014). Weakly supervised histopathology cancer image segmentation and classification. Medical image analysis, 18(3), 591-604.

[80] Xu, Y., Zhu, J. Y., Eric, I., Chang, C., Lai, M., & Tu, Z. (2014). Weakly supervised histopathology cancer image segmentation and classification. Medical image analysis, 18(3), 591-604.

[81] Chikontwe, P., Kim, M., Nam, S. J., Go, H., & Park, S. H. (2020). Multiple instance learning with center embeddings for histopathology classification. In Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part V 23 (pp. 519-528). Springer International Publishing.

[82] Ker, J., Bai, Y., Lee, H. Y., Rao, J., & Wang, L. (2019). Automated brain histology classification using machine learning. Journal of Clinical Neuroscience, 66, 239-245.

[83] Iizuka, O., Kanavati, F., Kato, K., Rambeau, M., Arihiro, K., & Tsuneki, M. (2020). Deep learning models for histopathological classification of gastric and colonic epithelial tumours. Scientific reports, 10(1), 1504.

[84] Kieffer, B., Babaie, M., Kalra, S., & Tizhoosh, H. R. (2017, November). Convolutional neural networks for histopathology image classification: Training vs. using pre-trained networks. In 2017 Seventh International Conference on Image Processing Theory, Tools and Applications (IPTA) (pp. 1-6). IEEE.

[85] Talo, Muhammed. "Automated classification of histopathology images using transfer learning." Artificial intelligence in medicine 101 (2019): 101743.

[86] Golatkar, Aditya, Deepak Anand, and Amit Sethi. "Classification of breast cancer histology using deep learning." Image Analysis and Recognition: 15th International Conference, ICIAR 2018, Póvoa de Varzim, Portugal, June 27–29, 2018, Proceedings 15. Springer International Publishing, 2018.

[87] Coudray, N., Ocampo, P. S., Sakellaropoulos, T., Narula, N., Snuderl, M., Fenyö, D., ... & Tsirigos, A. (2018). Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning. Nature medicine, 24(10), 1559-1567.

[88] Nguyen, H. G., Blank, A., Dawson, H. E., Lugli, A., & Zlobec, I. (2021). Classification of colorectal tissue images from high throughput tissue microarrays by ensemble deep learning methods. Scientific reports, 11(1), 1-11.

[89] N. Gat, "Imaging spectroscopy using tunable filters: a review," in Proc. SPIE 4056, Wavelet Applications VII, 2000, pp. 50–64.

[90] M. E. Martin, D. M. Wabuyele, and S. K. Sharma, "Applications of Hyperspectral Imaging in Medical Diagnosis and Treatment," in Proc. SPIE 9855, Next-Generation Spectroscopic Technologies IX, 2016, pp. 98550A.

[91] S. Lu and B. Fei, "Medical hyperspectral imaging: a review," J. Biomed. Opt., vol. 19, no. 1, 2014, Art. no. 010901.

[92] A. Savitzky, E. R. Dougherty, and M. K. Kim, "Hyperspectral Imaging and Machine Learning in Medical Applications," in Proc. SPIE 10989, Big Data: Learning, Analytics, and Applications, 2019, pp. 1098909.

[93] S. E. Campbell, B. Fei, and M. J. P. de Graaf, "Hyperspectral imaging for noninvasive tissue characterization and diagnosis," J. Biophotonics, vol. 12, no. 10, 201

[94] G. Lu and B. Fei, "Medical Hyperspectral Imaging for Surgical Guidance and Tissue Assessment: A Review," IEEE Rev. Biomed. Eng., vol. 12, pp. 218-233, 2019.

[95] Malkoff, Donald B., and William R. Oliver. "Hyperspectral imaging applied to forensic medicine." Spectral Imaging: Instrumentation, Applications, and Analysis. Vol. 3920. SPIE, 2000.

[96] Vo-Dinh, Tuan. "A hyperspectral imaging system for in vivo optical diagnostics." IEEE Engineering in Medicine and Biology Magazine 23.5 (2004): 40-49.

[97] Panasyuk, S. V., Yang, S., Faller, D. V., Ngo, D., Lew, R. A., Freeman, J. E., & Rogers, A. E. (2007). Medical hyperspectral imaging to facilitate residual tumor identification during surgery. Cancer biology & therapy, 6(3), 439-446.

[98] Wehner, E., Thapa, A., Livingston, E., & Zuzak, K. (2011, February). NIR DLP hyperspectral imaging system for medical applications. In Emerging Digital Micromirror Device Based Systems and Applications III (Vol. 7932, pp. 30-38). SPIE.

[99] Liu, Zhi, Hongjun Wang, and Qingli Li. "Tongue tumor detection in medical hyperspectral images." Sensors 12.1 (2011): 162-174.

[100] Schweizer, J., Hollmach, J., Steiner, G., Knels, L., Funk, R. H., & Koch, E. (2012). Hyperspectral imaging-A new modality for eye diagnostics. Biomedical Engineering/Biomedizinische Technik, 57(SI-1-Track-P), 293-296.

[101] Fei, Baowei, Hamed Akbari, and Luma V. Halig. "Hyperspectral imaging and spectral-spatial classification for cancer detection." 2012 5th International Conference on BioMedical Engineering and Informatics. IEEE, 2012.

[102] Akbari, H., Halig, L. V., Schuster, D. M., Osunkoya, A., Master, V., Nieh, P. T., ... & Fei, B. (2012). Hyperspectral imaging and quantitative analysis for prostate cancer detection. Journal of biomedical optics, 17(7), 076005-076005.

[103] Kiyotoki, S., Nishikawa, J., Okamoto, T., Hamabe, K., Saito, M., Goto, A., ... & Sakaida, I. (2013). New method for detection of gastric cancer by hyperspectral imaging: a pilot study. Journal of biomedical optics, 18(2), 026010-026010.

[104] Gerstner, A. O., Laffers, W., Bootz, F., Farkas, D. L., Martin, R., Bendix, J., & Thies, B. (2012). Hyperspectral imaging of mucosal surfaces in patients. Journal of biophotonics, 5(3), 255-262.

[105] Calin, M. A., Parasca, S. V., Savastru, R., & Manea, D. (2015). Characterization of burns using hyperspectral imaging technique–a preliminary study. Burns, 41(1), 118-124.

[106] Wahabzada, M., Besser, M., Khosravani, M., Kuska, M. T., Kersting, K., Mahlein, A. K., & Stürmer, E. (2017). Monitoring wound healing in a 3D wound model by hyperspectral imaging and efficient clustering. PloS one, 12(12), e0186425.

[107] Chung, H., Lu, G., Tian, Z., Wang, D., Chen, Z. G., & Fei, B. (2016, March). Superpixel-based spectral classification for the detection of head and neck cancer with hyperspectral imaging. In Medical Imaging 2016: Biomedical Applications in Molecular, Structural, and Functional Imaging (Vol. 9788, pp. 260-267). SPIE.

[108] Lu, G., Qin, X., Wang, D., Muller, S., Zhang, H., Chen, A., ... & Fei, B. (2016, March). Hyperspectral imaging of neoplastic progression in a mouse model of oral carcinogenesis. In Medical Imaging 2016: Biomedical Applications in Molecular, Structural, and Functional Imaging (Vol. 9788, pp. 252-259). SPIE.

[109] Li, Q., Zhou, M., Liu, H., Wang, Y., & Guo, F. (2015). Red blood cell count automation using microscopic hyperspectral imaging technology. Applied spectroscopy, 69(12), 1372-1380.

[110] Huang, Q., Li, W., Zhang, B., Li, Q., Tao, R., & Lovell, N. H. (2019). Blood cell classification based on hyperspectral imaging with modulated Gabor and CNN. IEEE journal of biomedical and health informatics, 24(1), 160-170.

[111] Wang, Q., Wang, J., Zhou, M., Li, Q., Wen, Y., & Chu, J. (2021). A 3D attention networks for classification of white blood cells from microscopy hyperspectral images. Optics & Laser Technology, 139, 106931.

[112] Verebes, G. S., Melchiorre, M., Garcia‑Leis, A., Ferreri, C., Marzetti, C., & Torreggiani, A. (2013). Hyperspectral enhanced dark field microscopy for imaging blood cells. Journal of biophotonics, 6(11‑12), 960-967.

[113] Fabelo, H., Ortega, S., Szolna, A., Bulters, D., Piñeiro, J. F., Kabwama, S., ... & Sarmiento, R. (2019). In-vivo hyperspectral human brain image database for brain cancer detection. IEEE Access, 7, 39098-39116.

[114] Weijtmans, P. J. C., Shan, C., Tan, T., De Koning, S. B., & Ruers, T. J. M. (2019, April). A dual stream network for tumor detection in hyperspectral images. In 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019) (pp. 1256-1259). IEEE.

[115] Lin, H.; Wei, C.; Wang, G.; Chen, H.; Lin, L.; Ni, M.; Chen, J.; Zhuo, S. Automated Classification of Hepatocellular Carcinoma Differentiation Using Multiphoton Microscopy and Deep Learning. J. Biophotonics 2019, 12, e201800435. https://doi.org/10.1002/jbio.201800435.

[116] J. Yoon, "Hyperspectral imaging for clinical applications," BioChip Journal, vol. 16, no. 1, pp. 1-12, 2022.

[117] Li, Q., Zhou, M., Liu, H., Wang, Y., & Guo, F. (2015). Red blood cell count automation using microscopic hyperspectral imaging technology. Applied spectroscopy, 69(12), 1372-1380.

[118] T. Vo-Dinh, "A hyperspectral imaging system for in vivo optical diagnostics," IEEE Engineering in Medicine and Biology Magazine, vol. 23, no. 5, pp. 40-49, 2004.

[119] Halicek, M., Fabelo, H., Ortega, S., Callico, G. M., & Fei, B. (2019). In-vivo and ex-vivo tissue analysis through hyperspectral imaging techniques: revealing the invisible features of cancer. Cancers, 11(6), 756.

[120] Huang, Q., Li, W., Zhang, B., Li, Q., Tao, R., & Lovell, N. H. (2019). Blood cell classification based on hyperspectral imaging with modulated Gabor and CNN. IEEE journal of biomedical and health informatics, 24(1), 160-170.

[121] Lu, G., Qin, X., Wang, D., Muller, S., Zhang, H., Chen, A., ... & Fei, B. (2016, March). Hyperspectral imaging of neoplastic progression in a mouse model of oral carcinogenesis. In Medical Imaging 2016: Biomedical Applications in Molecular, Structural, and Functional Imaging (Vol. 9788, pp. 252-259). SPIE.

[122] Leavesley, S. J., Annamdevula, N., Boni, J., Stocker, S., Grant, K., Troyanovsky, B., ... & Alvarez, D. F. (2012). Hyperspectral imaging microscopy for identification and quantitative analysis of fluorescently‑labeled cells in highly autofluorescent tissue. Journal of biophotonics, 5(1), 67-84.

[123] Aatresh, A.A.; Alabhya, K.; Lal, S.; Kini, J.; Saxena, P.P. LiverNet: Efficient and Robust Deep Learning Model for Automatic Diagnosis of Sub-Types of Liver Hepatocellular Carcinoma Cancer from H&E Stained Liver Histopathology Images. Int. J. CARS 2021, 16, 1549–1563. https://doi.org/10.1007/s11548-021-02410-4.

[124] Schultz, R.A.; Nielsen, T.; Zavaleta, J.R.; Ruch, R.; Wyatt, R.; Garner, H.R. Hyperspectral Imaging: A Novel Approach for Microscopic Analysis. Cytometry 2001, 43, 239–247. https://doi.org/10.1002/1097-0320(20010401)43:4<239::AID-CYTO1056>3.0.CO;2-Z.

[125] Song, J.; Hu, M.; Wang, J.; Zhou, M.; Sun, L.; Qiu, S.; Li, Q.; Sun, Z.; Wang, Y. ALK Positive Lung Cancer Identification and Targeted Drugs Evaluation Using Microscopic Hyperspectral Imaging Technique. Infrared Phys. Technol. 2019, 96, 267–275. https://doi.org/10.1016/j.infrared.2018.12.001.

[126] Sun, L.; Zhou, M.; Li, Q.; Hu, M.; Wen, Y.; Zhang, J.; Lu, Y.; Chu, J. Diagnosis of Cholangiocarcinoma from Microscopic Hyperspectral Pathological Dataset by Deep Convolution Neural Networks. Methods 2022, 202, 22–30. https://doi.org/10.1016/j.ymeth.2021.04.005.

[127] Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. Rethinking the Inception Architecture for Computer Vision. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 2818–2826.

[128] He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.

[129] Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. In Advances in Neural Information Processing Systems (pp. 1097-1105).

[130] Goodfellow, I., Bengio, Y., and Courville, A. (2016). Deep Learning. MIT Press.

[131] LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., and Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. Neural Computation, 1(4), 541-551.

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

[133] Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. The Journal of Physiology, 160(1), 106-154.

[134] Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. The Journal of Physiology, 148(3), 574-591.

[135] Daugman, J. G. (1985). Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. Journal of the Optical Society of America A, 2(7), 1160-1169.

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

[137] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems, 25, 1097-1105.

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

[139] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015 (pp. 1-9).

[140] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016 (pp. 770-778).

[141] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017 (pp. 4700-4708).

[142] Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). MobileNets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861.

[143] Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. In International Conference on Machine Learning (pp. 6105-6114).

[144] Zunair, H., Rahman, A., Mohammed, N., & Cohen, J. P. (2020). Uniformizing techniques to process CT scans with 3D CNNs for tuberculosis prediction. In Predictive Intelligence in Medicine: Third International Workshop, PRIME 2020, Held in Conjunction with MICCAI 2020, Lima, Peru, October 8, 2020, Proceedings 3 (pp. 156-168). Springer International Publishing.

[145] Bishop, C. M. (1995). Neural networks for pattern recognition. Oxford university press.

[146] Hochreiter, S., Bengio, Y., Frasconi, P., & Schmidhuber, J. (2001). Gradient flow in recurrent nets: the difficulty of learning long-term dependencies. In A field guide to dynamical recurrent networks (pp. 237-244). John Wiley & Sons.

[147] LeCun, Y., Bottou, L., Orr, G. B., & Müller, K. R. (2012). Efficient backprop. In Neural networks: Tricks of the trade (pp. 9-48). Springer, Berlin, Heidelberg.

[148] Nair, V., & Hinton, G. E. (2010). Rectified linear units improve restricted boltzmann machines. In Proceedings of the 27th International Conference on Machine Learning (ICML-10) (pp. 807-814).

[149] Bridle, J. S. (1990). Probabilistic interpretation of feedforward classification network outputs, with relationships to statistical pattern recognition. In Neurocomputing (pp. 227-236). Springer, Berlin, Heidelberg.

[150] Bottou, L. (2010). Large-scale machine learning with stochastic gradient descent. In Proceedings of COMPSTAT'2010 (pp. 177-186). Springer, Heidelberg.

[151] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436-444.

[152] Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. In Proceedings of the 32nd International Conference on Machine Learning (ICML-15) (pp. 448-456).

[153] Santurkar, S., Tsipras, D., Ilyas, A., & Madry, A. (2018). How does batch normalization help optimization?. In Advances in Neural Information Processing Systems (pp. 2483-2493).

[154] Luo, Y., Li, W., Urtasun, R., & Zemel, R. (2018). Understanding the effective receptive field in deep convolutional neural networks. In Advances in Neural Information Processing Systems (pp. 4898-4906).

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

[156] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," Journal of Machine Learning Research, vol. 15, pp. 1929-1958, 2014.

[157] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov, "Improving neural networks by preventing co-adaptation of feature detectors," arXiv preprint arXiv:1207.0580, 2012.

[158] Abramowitz, M., Spring, K. R., Keller, H. E., & Davidson, M. W. (2002). Basic principles of microscope objectives. Biotechniques, 33(4), 772-781.

[159] S. Inoué and K. R. Spring, Video Microscopy: The Fundamentals, 2nd ed. New York: Plenum Press, 1997.

[160] D. B. Murphy and M. W. Davidson, Fundamentals of Light Microscopy and Electronic Imaging, 2nd ed. Hoboken, NJ: Wiley-Blackwell, 2013.

[161] Budinger, T.F. Absorbed Radiation Dose Assessment from Radionuclides. In Comprehensive Biomedical Physics; Elsevie: Amsterdam, The Netherlands, 2014; pp. 253–269 ISBN 978-0-444-53633-4.

[162] Abe, T.; Murakami, Y.; Yamaguchi, M.; Ohyama, N.; Yagi, Y. Color Correction of Pathological Images Based on Dye Amount Quantification. OPT REV 2005, 12, 293–300. https://doi.org/10.1007/s10043-005-0293-6.

[163] Tuer, A.E.; Tokarz, D.; Prent, N.; Cisek, R.; Alami, J.; Dumont, D.J.; Bakueva, L.; Rowlands, J.A.; Barzda, V. Nonlinear Multicontrast Microscopy of Hematoxylin-and-Eosin-Stained Histological Sections. J. Biomed. Opt. 2010, 15, 026018. https://doi.org/10.1117/1.3382908.

[164] Wang, R.; He, Y.; Yao, C.; Wang, S.; Xue, Y.; Zhang, Z.; Wang, J.; Liu, X. Classification and Segmentation of Hyperspectral Data of Hepatocellular Carcinoma Samples Using 1-D Convolutional Neural Network. Cytometry 2020, 97, 31–38. https://doi.org/10.1002/cyto.a.23871.

[165] Aref, M.H.; Aboughaleb, I.H.; El-Sharkawy, Y.H. Tissue Characterization Utilizing Hyperspectral Imaging for Liver Thermal Ablation. Photodiagnosis Photodyn. Ther. 2020, 31, 101899. https://doi.org/10.1016/j.pdpdt.2020.101899.

[166] Rocha, A.D.; Groen, T.A.; Skidmore, A.K.; Darvishzadeh, R.; Willemen, L. The Naïve Overfitting Index Selection (NOIS): A New Method to Optimize Model Complexity for Hyperspectral Data. ISPRS J. Photogramm. Remote Sens. 2017, 133, 61–74. https://doi.org/10.1016/j.isprsjprs.2017.09.012.

[167] Geman, S., Bienenstock, E., & Doursat, R. (1992). Neural networks and the bias/variance dilemma. Neural Computation, 4(1), 1-58. doi:10.1162/neco.1992.4.1.1

[168] Hastie, T., Tibshirani, R., & Friedman, J. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction (2nd ed.). Springer. doi:10.1007/978-0-387-84858-7

[169] Domingos, Pedro. "A unified bias-variance decomposition." Proceedings of 17th international conference on machine learning. Stanford: Morgan Kaufmann, 2000.

[170] Belkin, M., Hsu, D., Ma, S., & Mandal, S. (2019). Reconciling modern machine-learning practice and the classical bias–variance trade-off. Proceedings of the National Academy of Sciences, 116(32), 15849-15854. doi:10.1073/pnas.1903070116

[171] Nakkiran, P., Kaplun, G., Bansal, Y., Yang, T., Barak, B., & Sutskever, I. (2019). Deep Double Descent: Where Bigger Models and More Data Hurt. arXiv preprint arXiv:1912.02292.

[172] Frankle, Jonathan, and Michael Carbin. "The lottery ticket hypothesis: Finding sparse, trainable neural networks." arXiv preprint arXiv:1803.03635 (2018).

[173] Yang, Tianle, and Joe Suzuki. "CA1-4 Dropout for Mitigating Double Descent in Deep Learning." Proceedings of the Annual Meeting of the Behavior metric Society of Japan 49, 2021.

[174] Xiang, Y.; Kim, W.; Chen, W.; Ji, J.; Choy, C.; Su, H.; Mottaghi, R.; Guibas, L.; Savarese, S. ObjectNet3D: A Large Scale Database for 3D Object Recognition. In Computer Vision—ECCV 2016; Leibe, B., Matas, J., Sebe, N., Welling, M., Eds.; Lecture Notes in Computer Science; Springer International Publishing: Cham, Switzerland, 2016; Volume 9912, pp. 160–176, ISBN 978-3-319-46483-1.

[175] Ji, S.; Xu, W.; Yang, M.; Yu, K. 3D Convolutional Neural Networks for Human Action Recognition. IEEE Trans. Pattern Anal. Mach. Intell. 2013, 35, 221–231. https://doi.org/10.1109/TPAMI.2012.59.

[176] Kleesiek, J.; Urban, G.; Hubert, A.; Schwarz, D.; Maier-Hein, K.; Bendszus, M.; Biller, A. Deep MRI Brain Extraction: A 3D Convolutional Neural Network for Skull Stripping. NeuroImage 2016, 129, 460–469. https://doi.org/10.1016/j.neuroimage.2016.01.024.

[177] Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet Classification with Deep Convolutional Neural Networks. Commun. ACM 2017, 60, 84–90. https://doi.org/10.1145/3065386.

[178] O'Shea, K.; Nash, R. An Introduction to Convolutional Neural Networks; ArXiv e-prints, 2015. https://doi.org/10.48550/ARXIV.1511.08458.

[179] Lin, M.; Chen, Q.; Yan, S. Network in Network; ArXiv e-prints, 2013. https://doi.org/10.48550/ARXIV.1312.4400.

[180] Kingma, D.P.; Ba, J. Adam: A Method for Stochastic Optimization; ArXiv e-prints, 2014. https://doi.org/10.48550/ARXIV.1412.6980.

[181] Li, D.-C.; Liu, C.-W.; Hu, S.C. A Learning Method for the Class Imbalance Problem with Medical Data Sets. Comput. Biol. Med. 2010, 40, 509–518. https://doi.org/10.1016/j.compbiomed.2010.03.005.

[182] Boughorbel, S.; Jarray, F.; El-Anbari, M. Optimal Classifier for Imbalanced Data Using Matthews Correlation Coefficient Metric. PLoS ONE 2017, 12, e0177678. https://doi.org/10.1371/journal.pone.0177678.

[183] Ross, D.A.; Lim, J.; Lin, R.-S.; Yang, M.-H. Incremental Learning for Robust Visual Tracking. Int. J. Comput. Vis. 2008, 77, 125–141. https://doi.org/10.1007/s11263-007-0075-7.

[184] T. Cover and P. Hart, "Nearest Neighbor Pattern Classification," in IEEE Transactions on Information Theory, vol. 13, no. 1, pp. 21-27, Jan. 1967.

[185] S. B. Kotsiantis, "Supervised Machine Learning: A Review of Classification Techniques," in Informatica, vol. 31, no. 3, pp. 249-268, 2007.

[186] Huang, J., Wei, Y., Yi, J., & Liu, M. (2018, February). An improved knn based on class contribution and feature weighting. In 2018 10th international conference on measuring technology and mechatronics automation (ICMTMA) (pp. 313-316). IEEE.

[187] Kouiroukidis, Nikolaos, and Georgios Evangelidis. "The effects of dimensionality curse in high dimensional knn search." 2011 15th Panhellenic Conference on Informatics. IEEE, 2011.

[188] Cortes, Corinna, and Vladimir Vapnik. "Support-vector networks." Machine learning 20 (1995): 273-297.

[189] B. Scholkopf and A. J. Smola, "Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond," MIT Press, Cambridge, MA, USA, 2002.

[190] S. S. Keerthi and C. J. Lin, "Asymptotic behaviors of support vector machines with Gaussian kernel," in Neural Computation, vol. 15, no. 7, pp. 1667-1689, July 2003.

[191] Hinton, G. E., and Salakhutdinov, R. R., "Reducing the Dimensionality of Data with Neural Networks," Science, vol. 313, no. 5786, pp. 504–507, Jul. 2006.

[192] Vincent, P., Larochelle, H., Bengio, Y., and Manzagol, P. A., "Extracting and Composing Robust Features with Denoising Autoencoders," Proceedings of the 25th international conference on Machine learning - ICML '08, Helsinki, Finland, 2008.

[193] Bengio, Y., "Practical Recommendations for Gradient-Based Training of Deep Architectures," Neural Networks: Tricks of the Trade, Springer Berlin Heidelberg, 2012, pp. 437–478.

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

[195] A. Veit, M. J. Wilber, and S. Belongie, "Residual networks behave like ensembles of relatively shallow networks," Advances in Neural Information Processing Systems, pp. 550–558, 2016.

[196] C. Chen, Q. Dou, H. Chen, and P.-A. Heng, "3D fully convolutional networks for co-segmentation of tumors on PET-CT images," IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), pp. 533–537, 2017.

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

[198] A. Veit, M. J. Wilber, and S. Belongie, "Residual Networks Behave Like Ensembles of Relatively Shallow Networks," Advances in Neural Information Processing Systems, pp. 550-558, 2016.

[199] Mosquera-Lopez, C.; Agaian, S.; Velez-Hoyos, A.; Thompson, I. Computer-Aided Prostate Cancer Diagnosis From Digitized Histopathology: A Review on Texture-Based Systems. IEEE Rev. Biomed. Eng. 2015, 8, 98–113. https://doi.org/10.1109/RBME.2014.2340401.

[200] Chen, J.-M.; Li, Y.; Xu, J.; Gong, L.; Wang, L.-W.; Liu, W.-L.; Liu, J. Computer-Aided Prognosis on Breast Cancer with Hematoxylin and Eosin Histopathology Images: A Review. Tumour. Biol. 2017, 39, 101042831769455. https://doi.org/10.1177/1010428317694550.

[201] Saxena, S.; Gyanchandani, M. Machine Learning Methods for Computer-Aided Breast Cancer Diagnosis Using Histopathology: A Narrative Review. J. Med. Imaging Radiat. Sci. 2020, 51, 182–193. https://doi.org/10.1016/j.jmir.2019.11.001.

[202] De Lucia, Gianluca, Marco Lapegna, and Diego Romano. "A GPU Accelerated Hyperspectral 3D Convolutional Neural Network Classification at the Edge with Principal Component Analysis Preprocessing." International Conference on Parallel Processing and Applied Mathematics. Cham: Springer International Publishing, 2022.

[203] Mei, S., Ji, J., Geng, Y., Zhang, Z., Li, X., & Du, Q. (2019). Unsupervised spatial–spectral feature learning by 3D convolutional autoencoder for hyperspectral classification. IEEE Transactions on Geoscience and Remote Sensing, 57(9), 6808-6820.

[204] Hossain, M. M., Hossain, M. A., Musa Miah, A. S., Okuyama, Y., Tomioka, Y., & Shin, J. (2023). Stochastic Neighbor Embedding Feature-Based Hyperspectral Image Classification Using 3D Convolutional Neural Network. Electronics, 12(9), 2082.

[205] Fırat, H., Asker, M. E., Bayındır, M. İ., & Hanbay, D. (2023). Hybrid 3D/2D complete inception module and convolutional neural network for hyperspectral remote sensing image classification. Neural Processing Letters, 55(2), 1087-1130.

[206] Feng, Y., Zheng, J., Qin, M., Bai, C., & Zhang, J. (2021). 3D octave and 2D vanilla mixed convolutional neural network for hyperspectral image classification with limited samples. Remote Sensing, 13(21), 4407.

[207] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2020). Generative adversarial networks. Communications of the ACM, 63(11), 139-144.

[208] Frid-Adar, M., Diamant, I., Klang, E., Amitai, M., Goldberger, J., & Greenspan, H. (2018). GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. Neurocomputing, 321, 321-331.

[209] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (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).

[210] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. Advances in neural information processing systems, 30.

[211] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., ... & Houlsby, N. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. arXiv preprint arXiv:2010.11929.

CURRICULUM VITAE

PERSONAL

Name	:	Umut Çinar
Birth Date/Place	:	1987/Ankara
E-mail	:	
GSM	:	

EDUCATION

PhD. Degree	:	Middle East Technical University
		Graduate School of Informatics, Information Systems
		2012 – 2023
MSc. Degree	:	Middle East Technical University
		Graduate School of Informatics, Information Systems
		2010 – 2012
Undergraduate	:	Middle East Technical University
		Faculty of Engineering, Computer Engineering
		2005 – 2010

WORK EXPERIENCE

Partner, Chief Solution Officer,	:	Link Retail (https://www.linkretail.com) 2021 – present

Video Analytics		- Link Retail is a technology pioneer company developing innovative software solutions for retailers by harmonizing its retail expertise, - Responsibilities include leading and developing video analytic product portfolio.
Founder	:	RetailFlux Inc. (https://www.retailflux.com) 2016 – 2021 - RetailFlux is a company specialized in retail analytic and business intelligence technologies, - Responsibilities include leading product development activities, creating new partnership relations and promoting the business.
General Manager	:	3Y Ltd. (http://www.3yteknoloji.com.tr) 2012 – present - 3Y is a company specialized in video analytic technologies, - Responsibilities include leading R&D team, leading marketing and sales activities
Research Assistant	:	Middle East Technical University 2010 – 2016 - Worked in R&D projects related to defense industry, - Studied image processing, pattern recognition, statistics and machine learning
Software Developer	:	Worked for various IT companies 2005 – 2010 - Obtained hands on experience with nearly all aspects of software technologies.

PUBLICATIONS

[1] Cinar, U., Cetin Atalay, R., & Cetin, Y. Y. (2023). Human Hepatocellular Carcinoma Classification from H&E Stained Histopathology Images with 3D Convolutional Neural Networks and Focal Loss Function. Journal of Imaging, 9(2), 25.

[2] Cinar, U.; Yardimci, Y.; Atalay, R.; Cetin, E; Classification of Human Carcinoma Cells Using Multispectral Imagery, SPIE Medical Imaging, 2016

[3] Kılıçkaya, M.; Cinar, U.; Ugurluoglu, S.; Efficient Marble Slab Classification using Simple Features, International Conference on Computer Vision Theory and Applications, 2016

[4] Yılmaz, K.; Karaman E.; Cinar U.; An Evaluation on the Usage of Intelligent Video Analysis Software for Marketing Strategies, Journal of Business Research, p. 146-158 , 2013

[5] Umut Çinar, Road Extraction from High Resolution Sattelite Images Using Adaptive Boosting with Multi-resolution Analysis, Master Thesis, 2012

[6] Gedik, E.; Cinar, U.; Karaman, E.; Yardimci, Y.; Halici, U.; Pakin, K.; , "A rule based approach for detecting bridges from high resolution electro-optic satellite images," *Signal Processing and Communications Applications Conference (SIU), 2012 20th* , vol., no., pp.1-4, 18-20 April 2012

[7] Karaman, E.; Cinar, U.; Gedik, E.; Yardimci, Y.; Halici, U.; , "Automatic road network extraction from multispectral satellite images," *Signal Processing and Communications Applications Conference (SIU), 2012 20th* , vol., no., pp.1-4, 18-20 April 2012

[8] Gedik, E.; Cinar, U.; Karaman, E.; Yardimci, Y.; Halici, U.; Pakin, K.; , "A new robust method for bridge detection from high resolution electro-optic satellite images," Geographic Object-Based Image Analysis Conference(GEOBIA), 2012 4th, vol., no., p.298, May 7-9, 2012

[9] Karaman, E.; Cinar, U.; Gedik, E.; Yardimci, Y.; Halici, U.;, "A new algorithm for automatic road network extraction in multispectral satellite images," Geographic Object-Based Image Analysis Conference(GEOBIA), 2012 4th, vol., no., p.298, May 7-9, 2012

[10] Cinar, U.; Karaman, E.; Gedik, E. ; Yardimci, Y. ; , "Yüksek çözünürlüklü multispektral uydu görüntülerinde kuvvetlendirilmiş sınıflandırıcılar kullanılarak otomatik yol çıkarımı," Savtek 2012, Savunma Teknolojileri Kongresi, 2012, vol., no., p.79, June 20-22, 2012

[11] Karaman, E. ; Cinar, U. ; Gedik, E. ; Yardimci, Y., "Yüksek çözünürlüklü multispektral uydu görüntülerinde su alanlarının otomatik sınıflandırılması ile nehir ve su kanalı tespiti ," MASFOR, 2012,

[12] Karaman, E.; Cinar, U.; Gedik, E.; Yardimci, Y.; Halici, U.;, "Fourier based feature descriptors for railroad extraction from aerial images," IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2012, July 22-27, 2012

[13] Cinar, U.; Karaman, E.; Gedik, E. ; Yardimci, Y. ;, "A new approach to automatic road extraction from satellite images using boosted classifiers," SPIE Image and Signal Processing for Remote Sensing Conference, 2012, September 24-26, 2012

[14] Cinar, U.; Karaman, E.; Gedik, E. ; Halici, U.; Yardimci, Y. ;, "Junction extraction on road masks by pruned skeletons," SPIE Image and Signal Processing for Remote Sensing Conference, 2012, September 24-26, 2012