REPUBLIC OF TÜRKİYE

ALTINBAŞ UNIVERSITY

Institute of Graduate Studies

Electrical and Computer Engineering

MULTI-APPROACHED HISTOPATHOLOGY IMAGES CLASSIFICATION

Ali Hussein Abdulwahhab ABDULWAHHAB

Doctor of Philosophy

Supervisor

Prof. Dr. Oğuz BAYAT

İstanbul, 2025

MULTI-APPROACHED HISTOPATHOLOGY IMAGES CLASSIFICATION

Ali Hussein Abdulwahhab ABDULWAHHAB

Electrical and Computer Engineering

Doctor of Philosophy

ALTINBAŞ UNIVERSITY

2025

The dissertation titled MULTI-APPROACHED HISTOPATHOLOGY IMAGES CLASSIFICATION prepared by ALI HUSSEIN ABDULWAHHAB ABDULWAHHAB and submitted on 01/12/2025 has been **accepted unanimously** for the degree of PhD in Electrical and Computer Engineering.

Prof. Dr. Oğuz BAYAT
Supervisor

Thesis Defense Committee Members:

Prof. Dr. Oğuz BAYAT	Department of Electrical-Electronics Engineering, Yeditepe University	_____
Prof. Dr. Adil Deniz DURU	Department of Sports and Health Sciences, Marmara University	_____
Prof. Dr. Yasa EKŞİOĞLU ÖZOK	Department of Electrical and Electronics Engineering, Altınbaş University	_____
Asst. Prof. Dr. Çağatay AYDIN	Department of Electrical and Electronics Engineering, Ege University	_____
Asst. Prof. Dr. Muhammad ILYAS	Department of Computer Engineering, Altınbaş University	_____

I hereby declare that this dissertation meets all format and submission requirements for a PhD dissertation.

I hereby declare that all information/data presented in this graduation project has been obtained in full accordance with academic rules and ethical conduct. I also declare all unoriginal materials and conclusions have been cited in the text and all references mentioned in the Reference List have been cited in the text, and vice versa as required by the abovementioned rules and conduct.

Ali Hussein Abdulwahhab ABDULWAHHAB

Signature

ABSTRACT

MULTI-APPROACHED HISTOPATHOLOGY IMAGES CLASSIFICATION

ABDULWAHHAB, Ali Hussein Abdulwahhab

Ph.D., Electrical and Computer Engineering, Altınbaş University,

Supervisor: Prof. Dr. Oğuz BAYAT

Date: 12/2025

Pages: 126

Histological images play a crucial role in diagnosing diseases, especially breast cancer, which remains a major health concern for women, and colon and lung for people worldwide. Computer-aided diagnosis tools significantly assist physicians in early detection and treatment planning, helping reduce mortality rates. Although, Convolutional neural networks (CNNs) based on deep learning have proven effective in distinguishing benign from malignant breast cancers, the scarcity of variability and feature diversity in small datasets would steer clear of CNNs and result in overfitting instances, as well as diminished discrimination. Deep learning is very sensitive to the hyperparameter optimization, and for the reason, the learning hierarchies could be the solution for network's comprehension of the relationship subtlety between the of the data types.

In this context, this dissertation introduced a HAFMAB-Net: Hierarchical Adaptive Fusion based on Multilevel Attention-Enhanced Bottleneck Neural Network. The network comprises two pathways utilizing an enhanced Bottleneck architecture with attention mechanisms to extract both global and spatial features. It incorporates a Deeper Spatial Attention Aggregator Module to boost the representation of locative features by focusing on key spatial regions, improving the discriminative power of aggregated features. Additionally, a modified Adaptive Fusion Module combines the enhanced global and boosted spatial features into a comprehensive and enriched feature representation, which is subsequently used for cancer classification based on whole image.

The histopathological image could include more than one type cancer inside one image, to address this issue, we proposed a novel MATERB Net: Multiscale Attention transformation based on Enhanced Residual Block for Patch-level breast histopathology images prediction. The proposed framework employs extraction feature processes based on multiscale to capture more complementary global and local characteristics to enhance the learning process of the model in coping with morphological variability within the tissue. In Additional the model uses Attention transformation techniques based on self-attention and cross-attention is used to learn the relationship between the pixels of the image and focus on the meaningful details and information in calculating the attention weights. Self-attention works on capturing and learning the relation between the extracted features based on multi-magnification factors sub-images, providing various levels of details and information, which feed into the cross-attention beside the extracted features from full images. Cross-attention works to provide meaningful and context-aware representation integrated with the global context and details of each label. Moreover, the proposed net used a Balanced Focal Loss function, addressing the treating all misclassified samples equally, and preventing the model from memorizing easy samples to reduce overfitting

Keywords: Deep Learning, Bottleneck Neural Network, Residual Block, Attention Mechanism, Histopathological Image, Classification, Predication.

TABLE OF CONTENTS

Pages

LIST OF TABLES

Pages

LIST OF FIGURES

Pages

ABBREVIATIONS

ACM	:	Attention Channel Module
AFM	:	Adaptive Fusion Module
AI	:	Artificial Intelligence
AM	:	Attention Mechanism
AP	:	Average Precision
AUC	:	Area Under the Curve
BACH	:	Breast Cancer Histology Challenge
BI-RADS	:	Breast Imaging Reporting and Data System
CAD	:	Computer-Aided Diagnosis
CBAM	:	Convolutional Block Attention Module
CNB	:	Core Needle Biopsy
CNN	:	Convolutional Neural Networks
CT	:	Computed Tomography
DB	:	Deep Broad
DCNN	:	Deep Convolutional Neural Networks
DecT	:	Deconv-Transformer
DL	:	Deep Learning
DSAAM	:	Deeper Spatial Attention Aggregation Module
ER	:	Estrogen Receptor
FNA	:	Fine Needle Aspiration
FPR	:	False Positive Rate

GANs	:	Generative Adversarial Networks
H&E	:	Hematoxylin and Eosin
HAFMAB-Net	:	Hierarchical Adaptive Fusion Based on Multilevel Attention-Enhanced Bottleneck Neural Network
HER2	:	Human Epidermal Growth Factor Receptor 2
K	:	Key
LCM	:	Local Channel Module
MATERB Net	:	Multiscale Attention Transformation Based on Enhanced Residual Block
NSCLC	:	Non-small Cell Lung Cancer
PR	:	Progesterone Receptor
Q	:	Query
RANet	:	Resolution Adaptive Network
SCLC	:	Small Cell Lung Cancer
SVM	:	Support Vector Machines
TL	:	Transfer Learning
TM	:	Attention Transformation
TPR	:	True Positive Rate
TRUS	:	Transrectal Ultrasound
V	:	Value
VAE	:	Variational Autoencoder
WSI	:	Whole Slide Imaging

LIST OF SYMBOLS

x	:	Input Logits
α	:	Weighting Factors for $L_{CEsmooth}$
β	:	Weighting Factors for L_{FL}
C	:	Classes Number
s	:	The Factor of Label Smoothing
y_i	:	Denotes the Actual one-hot Encoded Label for class i
$\hat{y}_i$	:	Refers to the Predicted Probability for class i
μ	:	Correct class of Balancing Factor
p_t	:	True Class Predicted Probability
γ	:	Acts as a Focus Parameter, Regulating the Degree of Emphasis Given to Difficult Examples
y	:	Ground Truth Labels
κ	:	Balancing Factor
λ	:	Focusing Parameter
p_i	:	Model Estimating Probability
$CE\ (x, y)$	:	Cross-entropy Loss

1. INTRODUCTION

1.1 OVERVIEW

Cancer remains a significant global health issue, affecting millions of people worldwide. Its impact on society is profound, leading to a multitude of health-related challenges for those diagnosed, as well as emotional and financial stress for patients and their families [1]. Prominent types of cancer, such as breast cancer, colon cancer, and lung cancer, contribute significantly to morbidity and mortality rates associated with the disease. These conditions not only compromise the physical health of numerous individuals but also place considerable strain on healthcare systems around the globe [2][3].

Traditional diagnostic methods primarily rely on various imaging techniques, including X-rays, CT scans, and MRIs, alongside histopathological evaluations through examining tissue samples [4] [5]. However, these conventional approaches have significant limitations. For example, traditional imaging may lack the precision to distinguish between benign and malignant lesions effectively. At the same time, the manual interpretation of histopathological images can be subject to bias, often depending on the pathologist's level of experience [6].

Histopathology is highly instrumental in diagnosing cancer, as it evaluates the cellular structure and organization of the tissues involved [7]. Though its role is paramount, many problems are faced in analyzing histopathological images that compromise accurate diagnosis. The color tonalities, different subjective readings by pathologists, and inherent data limitations may lead to dispiriting conclusions that could subsequently impact crucial diagnostic and treatment decisions positively or negatively. Additionally, high-resolution images created by contemporary slide scanners may reach sizes up to gigapixels, complicating the analysis and burdening traditional evaluation methods [8].

Deep learning for image analysis, especially histopathology images, will take a giant leap in reducing the subjectivity that plagues the traditional methods of working with such images. Therefore, automated analysis is now attracting keen interest among practitioners and researchers actively trying to incorporate deep learning to analyze histology images [9]. Deep learning algorithms can be leveraged toward enhancing precision and dependability in

tasks related to image classification; they will thus be instrumental in cancer detection and grading through histopathology [10] [11].

The inclusion of attention mechanisms with deep learning models is one of the main developments that they have initiated supports the performance of models in identifying salient characteristics by focusing on important points within tissue images , which in turn supports in solving problems related to the nature of images and storage needs for high-dimensional data, as these developments are changing the diagnostic rules in the world of pathological anatomy through computer-aided diagnostic systems based on deep learning and thus enhancing patient-specific treatment methods.

1.2 MOTIVATION

The main motivation behind the implementation of deep learning technologies in the classification of data related to tissue cancer is to increase the accuracy of diagnosis and understanding of cancer at the cell level, where typical architectural and cellular characteristics appear in many different diseases across the tissue sample, which need to examine the sample correctly under a microscope .These findings will be useful in offering an expedited disease diagnosis and a general overview of the mechanisms cast changes. Thus, it shall remain an easy yet effective alternative supporting traditional diagnostics through biomarker analysis and morphology by focusing on histopathological data without advanced imaging modalities. Deep Learning usually requires big data, but image data has its computational efficiency, cost advantage, ease of storage, and excellent suitability for sharing, making it highly beneficial for all programs based on deep learning. So, deep learning to sort histopathology cancer data can be seen as a hopeful way to improve cancer diagnosis and continue with research.

1.3 PEOPLE STATMENT

Interest in histopathology is increasing, particularly in cancer datasets and deep learning utilizing morphological techniques. Deep convolutional neural networks (DCNNs) are a subset of machine learning and artificial intelligence that have greatly benefited from processing and classifying medical images. Morphological representations, though, may throw a spanner in the works regarding accuracy and robustness in classifications.

Tissue images may be ambiguous, and similar shapes can lead to misclassification of one cancer type from another. This variability within class poses significant challenges for classification, as images from the same class look different due to factors like staining, imaging modalities, or sample preparation. These variations obscure the DCNNs' features, ultimately resulting in reduced performance.

Using just one DCNN model also limits its ability to generalize across different datasets of histopathology. Generalized histopathology datasets take much better advantage of available computational power and larger labeled data collections. On the other hand, a single architecture restricts the optimization of feature representation according to specific dataset characteristics. These issues must be addressed in developing a deep learning method for classifying histopathology images that is robust and reliable.

1.4 AIM OF STUDY

This dissertation aims to develop robust and interpretable two deep learning frameworks for accurately classifying cancer histopathology images (malignant, benign, and normal) under data scarcity and morphological ambiguity and predict sub-image labels based on the overall image context.

1.5 CONTRIBUTIONS AND NOVELTIES OF DISSERTATION

1.5.1 Enhanced Bottleneck Neural Network Based on Multilevel Attention for Histopathological Cancer

We have developed a novelty framework deep-learning model known as “HAFMAB-Net” for classifying histopathology images; it includes an enhanced Bottleneck deep-learning neural network-integrated with an attention mechanism. The proposed model works with a dual-path approach, which processes full image and its relative extracted patches, its increase the efficiency of network in extraction a complementary feature of global and spatial for variations deftly of morphological tissue

The usage of attention mechanism represented by applying multilevel Module via local and channel and local, which improves the representation of the features map through giving the prioritizing for the globally relative information of diagnostic and focusing on fine-grained

regions for the intricated tissues. Moreover, an Aggregation Module of Deeper Spatial Attention is applied for deeper features refinement via directly focus on critical regions. The framework of proposed model is seamlessly integrated with using adaptive fusion model, which facilitates the combination of the global and local level features in dynamic style, hence, driving robust classification across multiple datasets. In addition, a two types of loss function are (focal loss with label smoothing) are used as hybrid loss function for improving the classification accuracy and addressing effectively the imbalanced labels.

HAFMAB-Net outperforms existing models in classification tasks while demonstrating remarkable adaptability across imaging conditions by exhibiting an outstanding balance between computational efficiency, sensitivity, and accuracy. HAFMAB-Net was validated using the BACH dataset, achieving an impressive accuracy rate of 99%. Furthermore, it achieved a classification performance of 98.99% on the BreaKHis dataset and a perfect score of 100% on the and LC25000 datasets and their relative subsets. These results clearly indicate the effectiveness, efficiency, and accuracy of HAFMAB-Net in multiclass and binary classification tasks, hence its potential for wider applications in medical image analysis.

1.5.2 Enhanced Residual Block Based on Multiscale Attention Transformation for Breast Patch-Level Histopathology Image Prediction

We propose a novel MATERB Net: Multiscale Attention transformation based on Enhanced Residual Block for Patch-level breast histopathology image prediction. The proposed framework employs multiscale extraction feature processes to capture more complementary global and local characteristics and enhance the learning process of the model in coping with morphological variability within the tissue. In Addition, the model uses Attention transformation techniques based on self-attention, and cross-attention is used to learn the relationship between the pixels of the image and focus on the meaningful details and information in calculating the attention weights. Self-attention captures and learns the relation between the extracted features based on multi-magnification factors sub-images, providing various levels of details and information, which feed into the cross-attention beside the extracted features from full images. Cross-attention works to provide meaningful and context-aware representation integrated with the global context and details of each label. Moreover, the proposed net used a Balanced Focal Loss function, addressing the treating all

misclassified samples equally and preventing the model from memorizing easy samples to reduce overfitting

1.6 THESIS ORGANIZATION

The remain chapters of this dissertation are organised as follows:

Chapter 2: it provides a background for the histopathological images and the traditional extraction techniques, the role of the CAD based on Deep learning techniques in analysing the histopathology images, deep learning models and their structures, and the types of adopted datasets in this work.

Chapter 3: it discussed the methodology of the proposed model "HAFMAB-Net" with its structure, pre-processing techniques of dataset and how preparing for training process, the proposed hybrid loss function in addressing the issue of imbalanced dataset and how addressing the issue of overfitting of model during training process, results across each dataset, and discussing the findings and the limitation that facing during the training process.

Chapter 4: This will talk about prediction the unlabeled extraction patches from the original labeled full image using the novel MATERB Net, discuss the using of attention mechanism in the MATERB Net and how it will help understanding the pattern of the breast tissue in predicting the images based on patch level. Also, discuss the used of hybrid loss function in improving the efficiency of the MATERB Net's training process across the classification breast cancer using multi-dataset with various magnification factors and addressing the overfitting, and then discuss the findings and the limitation that facing during the training process.

Chapter 5: present the calculation and the future work for the "HAFMAB-Net" and "MATERB Net"

2. BACKGROUND

2.1 INTRODUCTION

In the modern computational environment, such applications based on the deep learning methodology through multilayer artificial neural networks are much more effective and reliable in image processing [13]. Such would enable unmasking yet obscure patterns in the increasingly bloating datasets of medical information, even more so that such information is swelling in both numerical and imaging forms. Progress in computer-aided diagnosis relates to improving clinical evaluation and decreasing unnecessary pathological assessment. Images of histopathology, indeed, are playing an essential role in cancer detection and are, further, the center of enormous research attention [14]. From these images, tumor heterogeneity, an intricate cellular interplay, and distribution over an architectural development form are observed, creating the requirement for a strong CAD system based on reliable image characteristics [15]. Indeed, this poses a major challenge. The diagnostic process previously has been carried out visually by pathologists based on preferred tissue samples. The emergence of CAD systems for pathologists has responded to this labour-intensive and time-consuming process. The development of these systems has been the subject of considerable research over the last couple of decades. Major emerging goals for automated analysis of histopathological images are accurate classifications of tissue samples as malignant or benign. Since the most relevant features for differentiation are not predefined and differ for every pathologist, this task has remained very challenging.

2.2 HISTOPATHOLOGY IMAGES

Histopathology is a subspecialty in medicine that deals with the microscopic study of tissue samples to make diagnostic decisions about the disease conditions [16]. It is essential in this area to identify the cellular aberrations, particularly in the oncological setting, where it enables the detection of the occurrence of the neoplastic processes, as well as their accurate classification. The diagnostic process normally begins with the acquisition of a biopsy, whereby tissues are removed through surgical procedures to undergo a follow-up examination [17]. The process of obtaining digital histopathology images involves a sequence of complex and clearly defined steps, i.e. the purchase of the biopsy, the cut of the tissue, the use of staining procedures and the ultimate digitalization of the ready slides, as

shown in Figure 2.1. After staining, the pathologists examine the sections of the tissues to identify any sign of malignancy, thus highlighting the fundamental importance of histopathology in the diagnosis of the wide range of diseases. The resulting visual data of the histopathological imaging is used to explain the cellular abnormalities at various levels of resolution, including whole-slide-image (WSI) level, patch-level, and cellular-level data [18]. The capability of pathologists to see the changes in tissue architecture and the distinct cell characteristics in their entirety is made possible by such multiscale representation. Figure 2.2 provides examples of the information derived from histopathological images across different analytical levels.

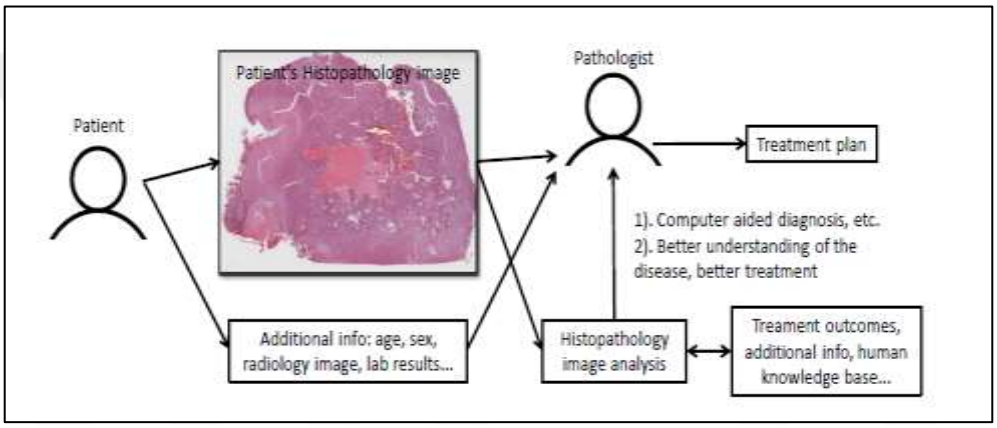

Figure 2.1: Histopathology Image Information [19].

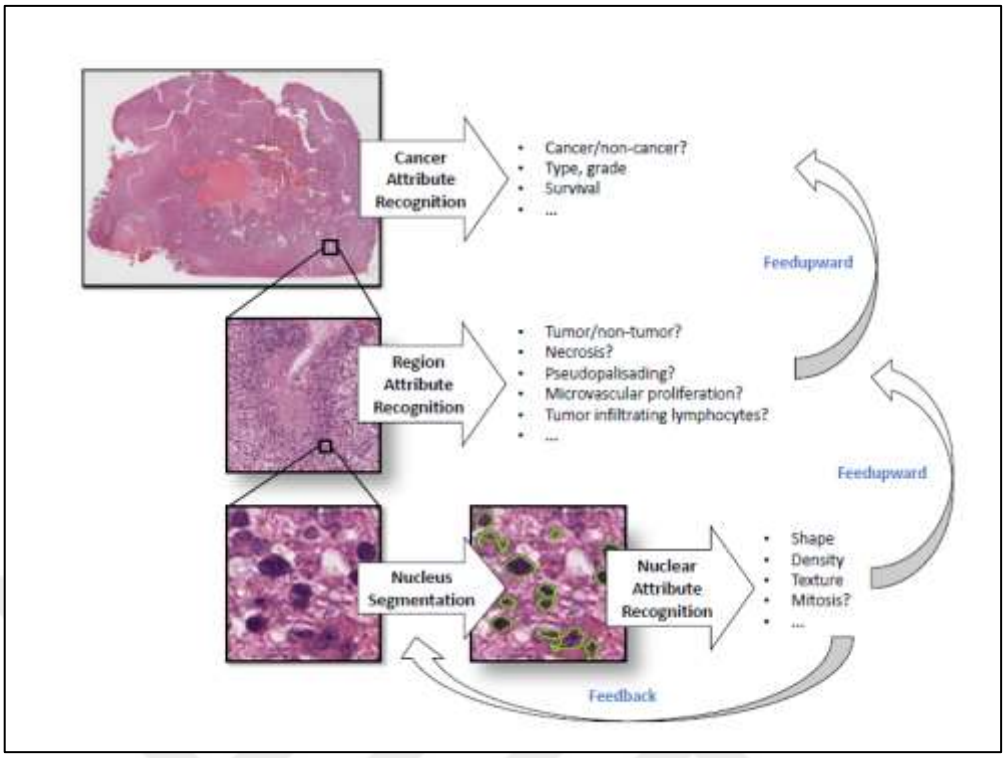

Figure 2.2: Multi-Level Information in Histopathology Images: WSI-Level, Patch-Level, Cellular-Level [19].

By testing these images, pathology can tell normal cells from abnormal ones and identify characteristics typical of different cancers [20]. For instance, lobular and ductal carcinoma show distinct histopathological features that can be identified under a microscope for their diagnosis. There are many challenges, however, in the study of histopathological images. One major problem is that different staining methods may change the appearance of tissue sections and lead to misinterpretation. There is also variability between observers; one pathologist may interpret a given image differently from another [21]. In addition, data scarcity limits training for diagnostic models and available reference material for rare diseases; thus, diagnostic conclusions suffer from lower credibility. Thus, though much technological innovation in image processing has taken place recently, much artificial intelligence has been integrated into disease diagnostics through histopathology, which is perceptibly considered an important input to enhance accuracy and decrease variability in interpretation. These issues must be considered when sharpening diagnostic accuracy, especially in difficult cases. Growing technologies meant to help pathologists' study histopathological images are expected to play a major role in disease detection and strengthening patient care.

2.3 WHOLE SLIDE IMAGING (WSI)

Whole Slide Imaging (WSI) is one of the breakthroughs of digital pathology. It enables conversion of entire histopathological glass slides into high-resolution images with one gigapixel [22]. The arrangement allows detailed computational tissue analysis, which increases diagnostic accuracy and evaluative accuracy in the field of pathology oncology. The shift of the physical to the digital slide increases the availability of the full range of histological information to health professionals, which, in turn, allows making more accurate diagnoses and making more informed clinical decisions. The key strength of WSI is its ability to store histopathological data in a digital library and thus enhance collaboration and information sharing among pathologists and other health care providers at the international level [23]. High-resolution digital image remote viewing is especially beneficial in areas where no specific skills are available. In addition, WSI is expected to help in automating diagnostic processes, which may simplify the steps of preparing and handling slides [24]. Figure 2.3 depicts the workflow that is related to WSI. However, there are still some limitations. One of the main issues includes the high volumes of data obtained with the help of high-resolution scans, which require considerable amounts of computational resources to store and analyze. Image segmentation and annotation are technical issues, because they involve advanced techniques that are needed to accurately outline and identify important areas on tissue images. Figure 2.4 explains the magnification level used in Whole Slide Imaging. As a result, both image segmentation and annotation are faced with technical challenges, and sophisticated methodologies are required to precisely define and delineate important areas in the tissue image.

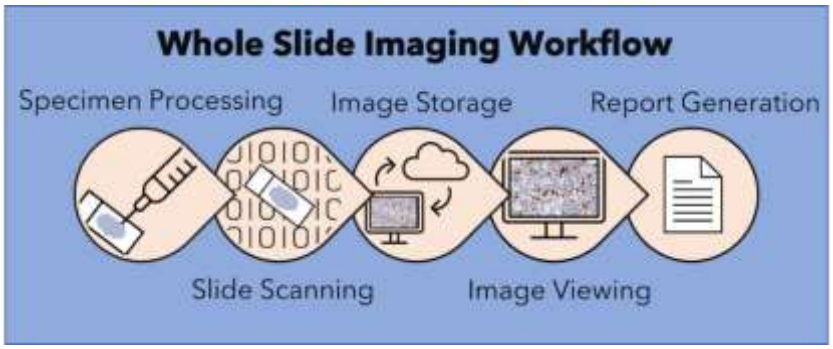

Figure 2.3: WSI Workflow.

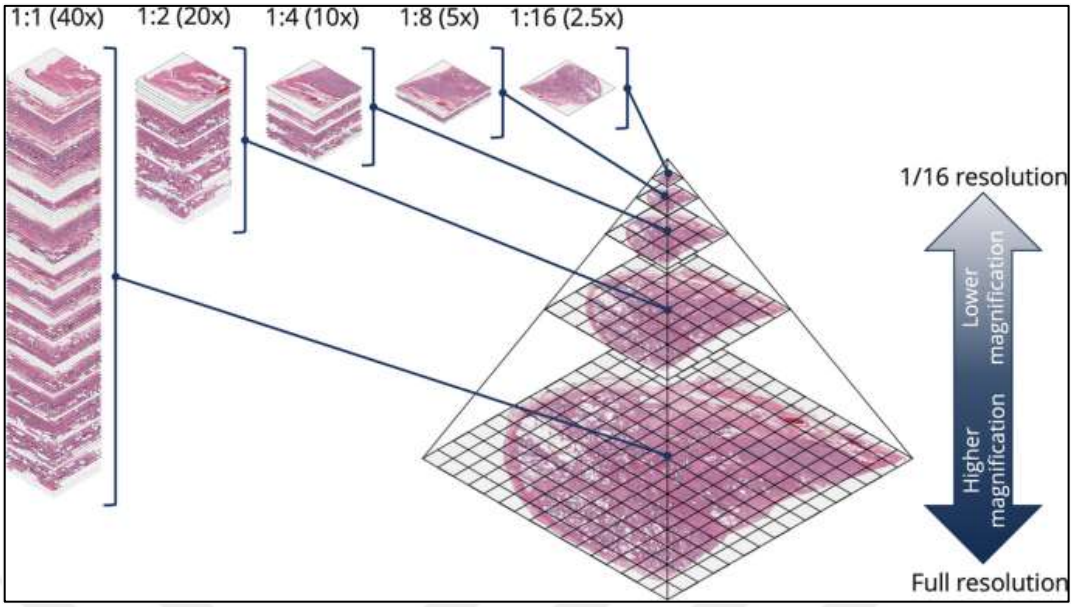

Figure 2.4: Whole Slide Imaging Level Magnification [26].

The application of artificial intelligence to optimize the whole-slide imaging (WSI) in the cancer classification and detection systems through the pattern recognition becomes particularly important as the use of deep learning technologies becomes more prominent. Essentially, deep learning models train neural networks on large datasets to derive patterns and features specific to malignancy, thus increasing diagnostic accuracy and efficiency [27]. An overview of the use of WSI based on deep learning is provided in Figure 2.5. In addition to diagnosing, WSI improves clinical learning for healthcare workers by creating vast repositories of digital, annotated slide images. This material enables learners and professionals to work with a wide spectrum of pathological cases that complement their learning experiences and strengthen their professional development. In its future evolution, therefore, within the health and cancer research fields, WSI can only stand to become even more central with diagnostic features collaborations and educational opportunities integrated.

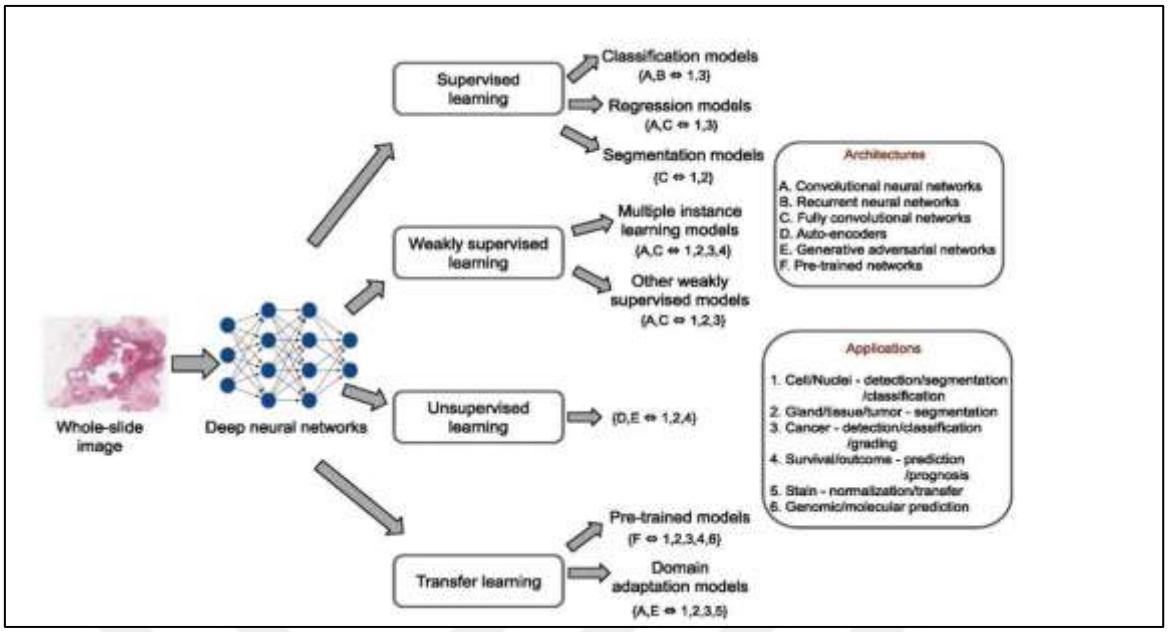

Figure 2.5: WSI Based on Deep Learning Application.

2.4 HISTOPATHOLOGICAL IMAGE ANALYSIS IN CANCER DIAGNOSIS

A precise nuclear pleomorphism and advanced architectural features in WSI have been made indispensable to optimal cancer diagnosis in routine histopathological practice [28]. It provides an excellent source of information regarding different anomalies about patterns and characteristic features of nuclei in malignant tissues, increasing diagnostic capacity to differentiate malignancy from benignity with more precision [29] [30]. The nuclear focus can minimize the existence of subtle differences in the malignant and the benign cellular conditions; however, it is beneficial regarding the latter as well as advancing the diagnostic refinement and the adequate influence on the patient outcome [31].

Computer-aided diagnosis (CAD) is an artificial intelligence that further enhances the examination of histopathology as it detects subtle factors in the morphology of the nucleus and architectural features that the human eye might not notice [32]. The diagnostic processes are speeded up and the accuracy of their results improved by rapid automation offered by CAD, thus allowing pathologists to identify salient patterns that can be used to make therapeutic decisions [33]. Moreover, AI-based on CAD helps classify various neoplasms, which is necessary to support the preparation of treatment programs of a strategic character.

Incorporation of artificial intelligence (AI) technologies and system integration into histopathological diagnostics hastens the development of malignancies detection and management, which leads to a better patient outcome and more effective health interventions [34]. Further understanding of cancer biology can be achieved by the development of AI-based methodologies in parallel with the guidelines of diagnostic procedures. It is expected that further technological progress will result in high-quality diagnostic modalities, which will benefit the clinical outcome of the patients and widen the therapeutic options in the practice [35].

2.4.1 Breast Cancer Diagnosis

Breast cancer is one of the serious problems of population health, as it is a disease of breast tissue and is now considered the second most common cancer diagnosed all over the world [36]. The disease has multiple different subtypes, such as invasive ductal carcinoma, invasive lobular carcinoma, and ductal carcinoma in situ, which have different pathological features that can have a significant impact on the therapeutic decision-making process [37] [38]. Figure 2.6 represents the spectrum of breast cancer subtypes.

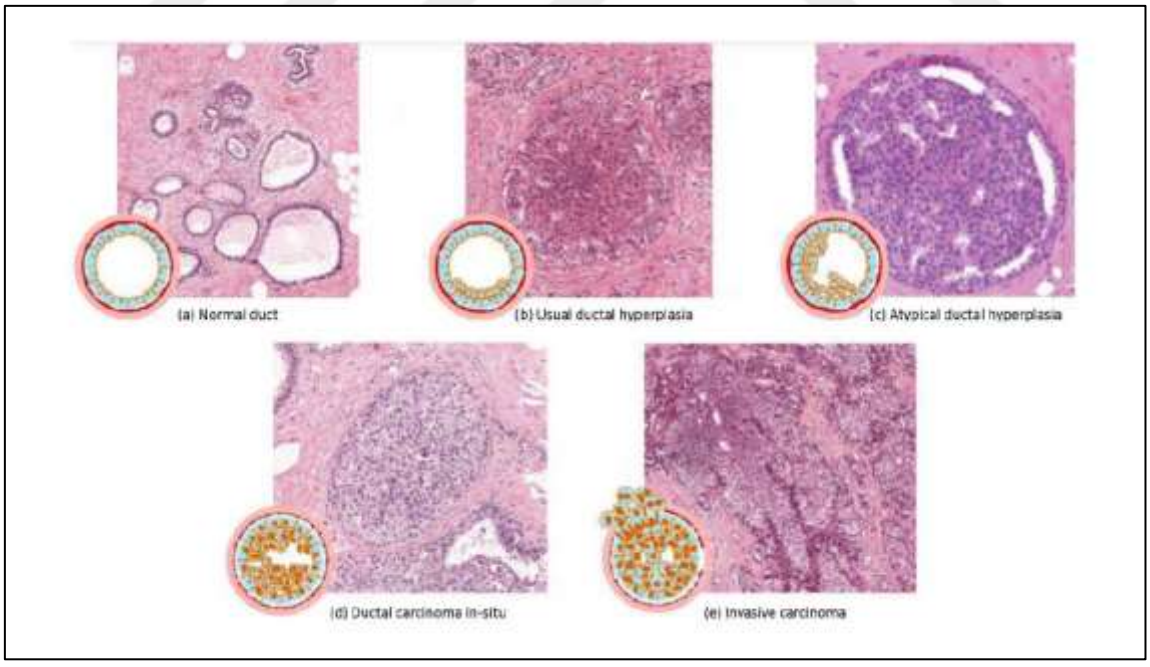

Figure 2.6: Breast Cancer Types.

Imaging modalities, including mammography and ultrasound, are the most common forms of standard diagnostic procedures and are fundamentally involved in the detection of breast lesions [40] [41]. However, these methods can be faced with difficulties especially when it comes to heterogeneous interpretations. As an example, an ultrasound scan can classify a lesion as highly suspicious thus showing a high likelihood of malignancy. At the same time, a mammogram can refer to the same lesion as indeterminate according to the guidelines provided by the Breast Imaging Reporting and Data System (BI-RADS) [42] [43] [44]. After the lesion is detected, biopsy procedures are used to make a diagnosis of breast cancer. There are numerous types of biopsy methods with different merits as shown in Figure 2.7.

i. Fine Needle Aspiration (FNA): Involves a thin needle to obtain a small tissue sample for cytological examination.

ii. Core Needle Biopsy (CNB): Uses a larger needle to extract a tissue core, providing more comprehensive information than FNA.

iii. Surgical Biopsy: Consists of a more invasive procedure to excise a larger portion of tissue, which can be classified as excisional (removing the entire tumor) or incisional (removing part of the tumor).

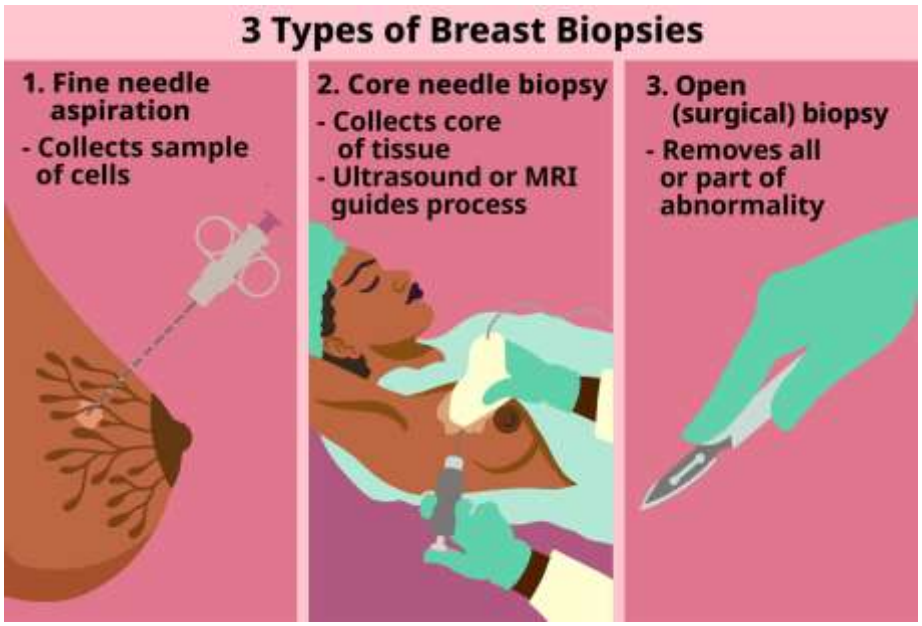

Figure 2.7: Breast Cancer Biopsy Techniques [45].

To clarify the peculiarities of a neoplasm, it is critical to conduct a histopathological examination of the obtained tissue samples. The main characteristic of this kind of assessment is the tumour grading that measures the degree of deviation of malignant cellular differentiation and the normal one [46]. Figure 2.8 represents a grading system that ranges between 1 (well-differentiated) and 3 (poorly differentiated) and provides essential data on the biological behaviour of the tumour and its possible aggressiveness; the data, in turn, lead to the decision-making in therapy and management [47]. Biomarker assessment is an inseparable part of cancer classification. The main biomarkers that are considered in this regard are Human Epidermal Growth Factor Receptor 2 (HER2), Estrogen Receptor (ER), and Progesterone Receptor (PR) [48]. These indicators play a critical role in choosing the right treatment such as treatment that focuses on oncogenic signalling pathways or hormone receptors. Histopathological parameters that are important to the diagnosis of breast cancer include tumour grade and biomarker assessment. Tumour grade is an indicator of the extent of cellular change that has been accrued in the course of tumour development, whereas the expression of HER2, ER, and PR give crucial prognostic and predictive data on probable treatment response. Computer-aided diagnosis is an emerging technology using deep-learning that can be used to enhance breast-cancer diagnostics. This advanced algorithm enhances the accuracy of medical image analysis and, therefore, allow radiologists to detect and describe tumours with better accuracy and efficiency [49]. Thus, CAD has a significant potential to improve the accuracy of the diagnosis, which will eventually be reflected in patient outcomes in the management of breast-cancer.

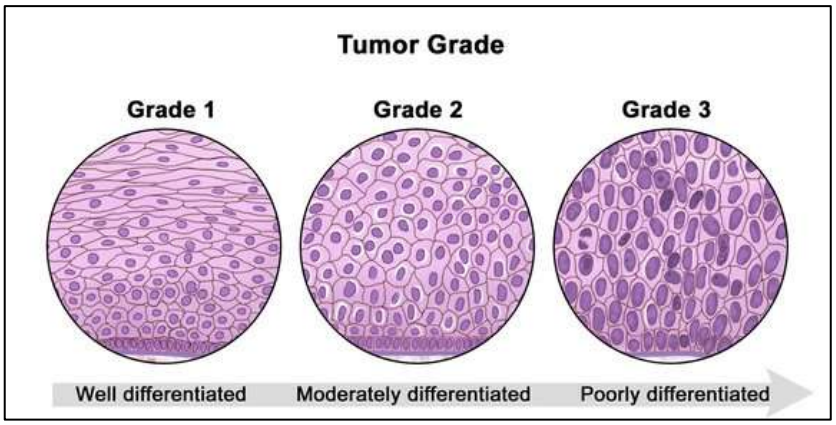

Figure 2.8: Breast Cancer Grading Scale [50].

2.4.2 Colon Cancer Diagnosis

Colon cancer is characterized by the unregulated growth of colonic epithelial cells, recurrence and is a significant issue in the field of oncology because of its prevalence and mortality [51]. The necessity to find effective screening and treatment is urgent with the increased prevalence of this disease in the world. Adenocarcinoma comprises 95 per cent of colorectal neoplasms; other less frequent entities are carcinoid tumours, lymphomas and gastrointestinal stromal tumours [52] [53]. Diagnosis is usually based on colonoscopy, imaging and stool tests. These modalities are however not without problems: colonoscopies are invasive, patient compliance is not always the same and imaging can not always help to detect the disease at an early stage. The gold standard of diagnosis of colorectal carcinoma is histopathology. It involves the microscopic analysis of tumour tissue in order to establish malignancy, subtyping, and tumour burden. The results inform the decision-making process in therapy. Various methods of biopsy have been created with this purpose in mind:

i. Endoscopic Biopsy is a minimally invasive procedure that involves collecting tissue samples during a colonoscopy, allowing for direct inspection of lesions.

ii. Transrectal Ultrasound (TRUS)- guided Biopsy: This procedure targets lesions in the rectum or lower colon, utilizing ultrasound for precise tissue localization.

iii. Surgical Biopsy: This method is employed when other diagnostic approaches yield inconclusive results. It enables the acquisition of a larger tissue sample.

The cell shapes of colon cancer are not typical; their cell structures are destroyed, and their mitotic activity is not normal; all aspects are evidence of malignancy. The more cautiously and unbiasedly these features are assessed, the more accurate the staging and prognosis will be. Until most recently, computer-aided diagnosis (CAD) systems, which involve deep learning technology, have been very helpful in improving diagnostics related to colon cancer [57]. The systems undertake an in-depth analytical process of histopathologic images to enhance the sensitivity and specificity of detection and generate quantifiable data that human pathologists can use to make informed, intelligent diagnostic assessments. The error of human fallibility is also virtually minimized in the CAD-driven image analysis process, which speeds up the duration for eventually coming up with medically more correct diagnoses on the way for better outcomes for patients under treatment for colon cancer.

2.4.3 Lung Cancer Diagnosis

Lung cancer is, therefore, a serious public health problem both in terms of incidents and deaths, ranking among the leading types of cancer in mortality in the world [49]. Millions of new cases are reported each year. Small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC) have been shown in Figure 2.9 as the two major distinctly different types of lung cancer, NSCLC being the most common as well as the more heterogeneous form, accounting for approximately 85% of cases [58]. The basic concept of the traditional approach to diagnosing lung cancer includes various imaging modalities such as chest X-ray and computed tomography (CT) scans for sampling tissues either by bronchoscopy or needle biopsy [58]. Despite being highly effective, these ways suffer their limitations and, therefore, express false negatives that complicate the diagnosis and unnecessary treatments of lesions in distinguishing between benign and malignant. Lung cancer diagnosis depends exclusively on histopathology. Pathologists look at tissue samples under the microscope to find cancer cells and identify the specific subtype of lung cancer because both features are required to decide on a treatment plan and prognosis for a patient [59]. SCLC is histopathologically characterized by small, rapidly dividing oval cells with a tendency to early metastasis. NSCLC encompasses several subtypes, including particularly different cytopathologic features for adenocarcinoma, squamous cell carcinoma, and large cell carcinoma. The procedures associated with the routine biopsy in the diagnosis of lung cancer include [60]:

i. Bronchoscopy: A minimally invasive technique that allows for direct visualization and sampling of lung tissue using a thin tube with a camera.

ii. CT-guided needle biopsy: This approach employs imaging technology to target lung lesions for tissue collection without surgery precisely.

iii. Mediastinoscopy: A surgical procedure that enables examination and biopsy of lymph nodes in the mediastinum, aiding in staging and diagnosis.

The lung histopathological features generally depict atypical cell morphology, mitotic abnormality, and lost normal tissue architecture. This would further escalate the tumor grade and its aggressiveness assessment by pathologists [61]. Until recently, another prophylactic approach relating to the latest diagnostic facilities, such as the deep learning-based computer-aided detection (CAD) system, has been rushed into the market. This has been

done to raise the diagnostic accuracy of lung cancer. Studies have also indicated that CAD can increase nodule detection sensitivity at CT but is accompanied by a lower false positive rate than before; this enhances the potential for earlier diagnoses and outcomes of such findings for patients. The above, therefore, marks it as a continuing trend in the lung cancer diagnosis process.

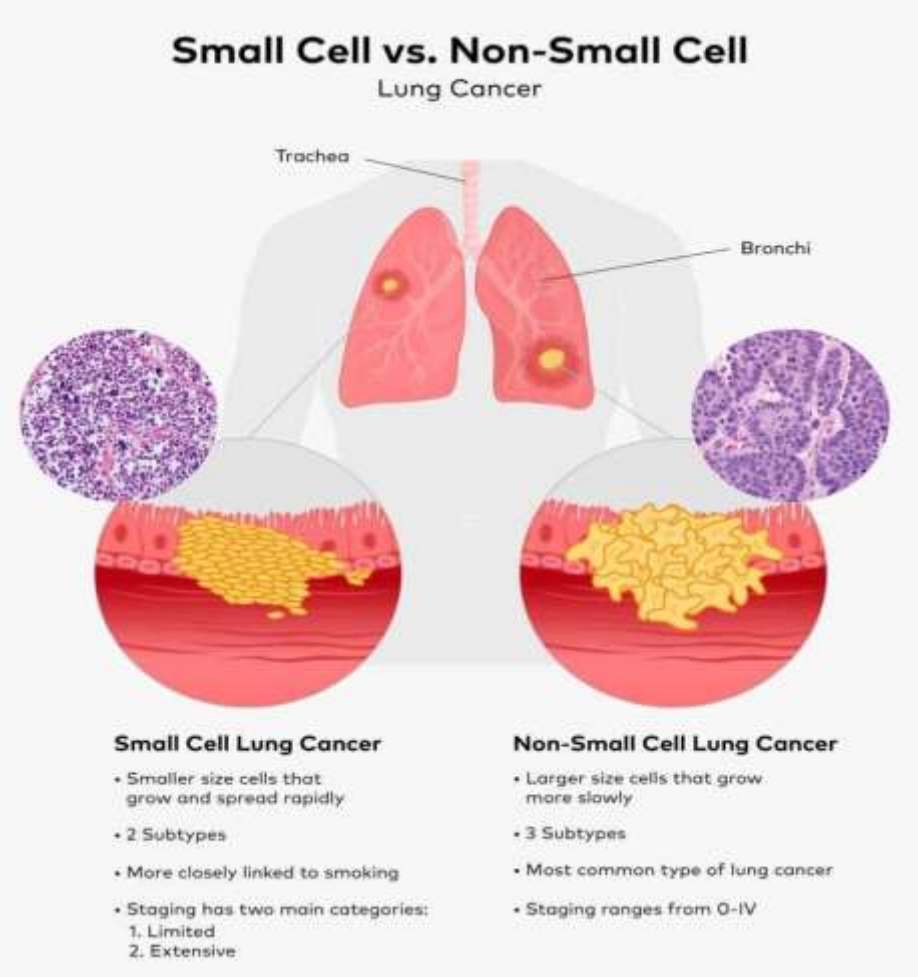

Figure 2.9: Lung Cancer Based on The Cell Size [62].

2.5 DATASETS

2.5.1 Dataset of BACH

As part of the 15th International Conference on Image Analysis and Recognition (ICIAR 2018), the Grand Challenge on Breast Cancer Histology Images (BACH) acknowledged the complexity of the issue and the high emend for automated diagnosis and detection tools [63]. The dataset comprises microscopic and whole-slide breast histopathology images stained

with H&E. It forms part of the Classification Grand Challenge that was presented at the International Conference on Image Analysis and Recognition (ICIAR). The dataset is composed of 400 images, distributed in four categories as indicated in Figure 2.10. The BACH images dataset is of size 2040x1536 and it is scanned at a magnification factor of 200x with a pixel size of 0.42um x 0.42um.

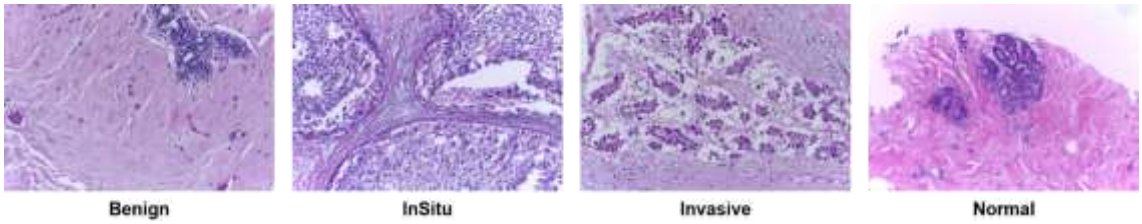

Figure 2.10: BACH Dataset Samples.

2.5.2 Dataset of BreaKHis

The BreakHis dataset originates from the Robotics and Imaging Laboratory at the Federal University of Parana, Brazil, and it has two major classes of breast tumors: malignant and benign. The magnifications of the images are 40x, 100x, 200x, and 400x. Each image has a resolution of 700x460. There are 5429 samples of malignant and 2480 of benign images [64]. The distribution of images in the dataset for each cancer subtype is provided in Table 2.1.

Table 2.1: Overview Distribution of BreaKHis-Dataset.

Type	**40X-magnifications**	**100X-magnifications**	**200X-magnifications**	**400X-magnifications**
Benign				
Adenosis	114	113	111	106
Fibro	253	260	264	237
Tubular	109	121	108	115
Phyllodes	149	150	140	130
All	625	644	623	588

Table 2.1: Overview Distribution of BreaKHis-Dataset (Table Continued).

Type	40X-magnifications	100X-magnifications	200X-magnifications	400X-magnifications
Malignant				
Ductal	864	903	896	788
Lobular	156	170	163	137
Mucinous	205	222	196	169
papillary	145	142	135	138
All	1370	1437	1390	1232

2.5.3 Dataset of LC25000

The source of the LC25000 dataset was from confirmed origins complying with HIPAA regulations; Figure 2.11 depicts the five different categories of LC25000-datset. Initially, there were just 250 samples per class, but now they have all been augmented to 5,000 samples, providing 25,000 samples in this entire dataset. The images within this dataset measure 768 by 768 pixels in dimension [65].

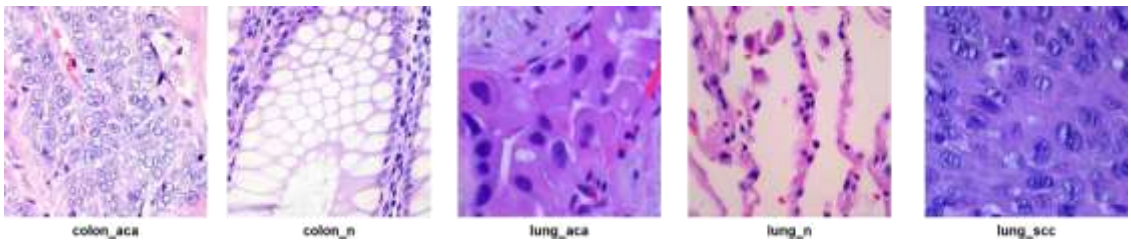

Figure 2.11: LC 25000 Dataset Samples.

2.6 PRE-PROCESSING AND AUGMENTATION TECHNIQUES

The role of data augmentation techniques is to increase the size and variability of datasets so that models can be trained on these images more robustly [66]. For histopathology images, detecting whether there is a normal or malignant part is very hard. Augmentation, for instance, rotation, flipping, scaling, and transformations in colors, among others, boosts even

more model performance where such an image is presented in Figure 2.12. Such methods help to reduce overfitting by modeling the same incident and the same appearance into tissue samples and ultimately improve the generalization ability of the model [67]. Augmentation drives diversity into the training data without the need to gather more images, which proves useful especially in domains like pathology where data acquisition is expensive and time-consuming [68].

Pre-processing of data presents the very prime methodology in using images before entering analysis and classification tasks [69]. Under histopathology, normalization can act as the key pre-processing step to bring the consistency of the overall distribution of images in the data set. Therefore, this factor greatly enhanced the performance and accuracy of model training [70]. For example, speedy and robust unification of the pixel values provides enhanced resolution in experimental images. Techniques like input layer normalization with models like the U-Net and VGG have peaked the model accuracies to around 80%. Another major pre-processing method is Z-score normalization that further splits the color images into smaller patches to normalize each image [71] properly. More recent studies promise to deliver substantial general improvement on the performance measure from the ten-fold normalization process within the phase of pre-training datasets. Furthermore, dealing with imbalances of classes, such as undersampling, improved classification performance in histopathological images [72]. For example, in classifying colon and lung tissues, there have been successes in achieving high accuracies. At the same time, challenges remain in that aspect of classifying breast cancer samples due to the clustering of tissue pixels [73]. Misclassification can be reduced and held at acceptable levels by selecting a strategic undersampling approach of the majority class and a portion of the remaining data. In brief, pre-processing and augmentation strategies cooperate in improving the process of training by creating more precise diagnostic models in pathology.

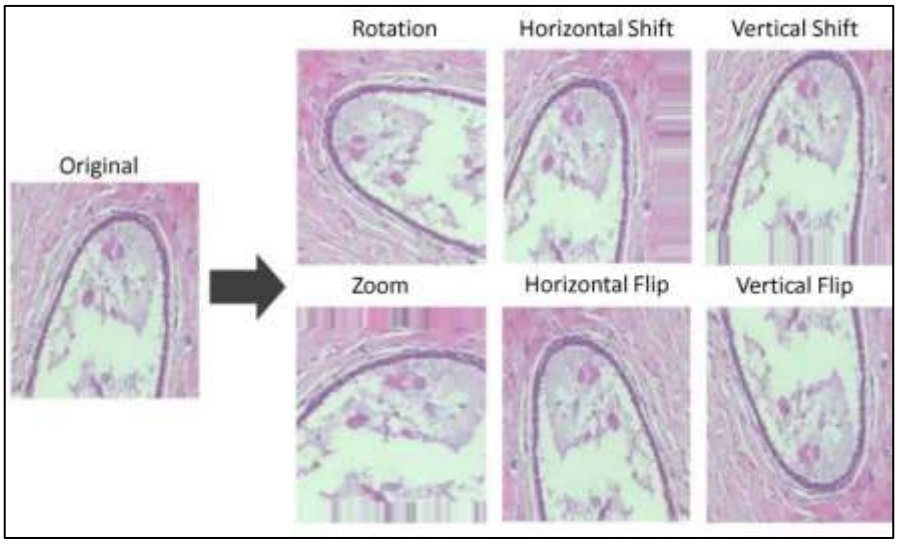

Figure 2.12: Histopathological Augmentation Techniques.

2.7 EVOLUTION OF COMPUTER-AIDED DIAGNOSIS (CAD) IN MEDICAL IMAGING

CAD has recently gained great momentum in medical imaging for the detection of cancers. If thoughtfully designed to push for early and highly accurate diagnosis, such tools hold great promise for meaningfully improving patient outcomes and survival [74]. The CAD systems integrate image processing, machine learning, and computer vision that deliver detailed analysis of histopathological images across various tissue samples [75].

CAD has recently been extensively employed in histopathology to assist in the diagnosis of common cancers such as cancers of the breast, colon, and lung. It allows for a more systematic analysis of tissues. For example, CAD has been useful in breast cancer for detecting and characterizing lesions with more accuracy, hence assisting pathologists toward appropriate management sooner [76]. The same approach in CAD is being utilized in colon and lung histopathology to highlight abnormalities that might represent malignancies, leading to improved clinical outcomes [77].

By incorporating deep learning into CAD, diagnostic capability is empowered through enhanced pattern recognition, which ensures more informed patient care. Studies illustrate this: Lee et al. [79], reduced false positives, suggesting that CAD, combined with expertise, strengthens diagnostic reliability. Gonçalves et al. [80], found that CNNs can identify cancerous cells with a higher degree of accuracy. Yoon et al. [81], showed that CAD, if

adapted to standard pathology practice, improves the ability of pathologists to highlight subtle tumors that might otherwise have been overlooked. Chen et al. [82], extended the application range of CAD to different types of cancer. Garcia and colleagues proposed the adaptation of such systems to the local clinical setting to maximize usefulness.

Despite remarkable achievements, a major challenge persists: the use of multisource images with differing degrees of clarity impairs CAD performance in histopathology. Improvement is possible through training by more representative and better-annotated datasets [83]. Besides that, ever-changing medical knowledge requires revising algorithms on a regular basis, something very resource-consuming and resistant to simplification [84][85].

Although CAD holds a promise for the solution of many diagnostic challenges, there are a number of obstacles for its real clinical-world adoption. For example, Coudray et al. [86], illustrated the dependency of high-performance systems on high-quality training data. Secondly, it was not until 2020 that Zhenfeng et al. highlighted the much-required adjustment in algorithms perpetually to catch up with the dynamic standards of clinical diagnoses [87]. The more recent Wang et al. in 2021 and also Weiming et al. in 2021 voiced their concern over the quality of generalizability of CAD systems limited by the lack of specialized training datasets essential for effective integration [88] [89]. The study further goes on to consider the work of Xiao et al. in 2022 which went into evaluating the scalability of those systems in different healthcare settings applying for generalizable flexible frameworks catering to the variable requirements of real clinical practice [90].

2.8 DEEP LEARNING TECHNIQUES

2.8.1 VGGNet 16

VGG 16 is actually an architecture that was developed by the Visual Geometry Group at Oxford [91]. Simplicity and depth were the design principles behind VGG 16 because, as the name implies, it has 16 layers consisting of a large number of convolutional layers followed by max-pooling layers. It's emphasized the use of small 3x3 convolutional filters; when stacked on top of one another, this allows the model to learn very complex features [92]. Figure 2.13 present the structure of VGG16. This model further includes fully connected layers towards the end that increase the capacity to make an analysis and a classification from an image.

Since it originates from a pre-training ImageNet model, VGG 16 can satisfactorily facilitate a whole range of computer-based vision duties. It can very well stand as feature extraction in the preliminary layers and thereby support a plethora of applications across domains [93]. Practically, VGG 16 might find immense use in image classification tasks, wherever researchers or practitioners from applicable industries wish to discover objects within an image and categorize them with higher precision. In this very vein of medical imaging, VGG 16 is used for feature extraction to help diagnose diseases based on histopathological images. This proves the applicability and significance of VGG 16 thereby stamping its effect on powering research and realistic applications in computer vision [94].

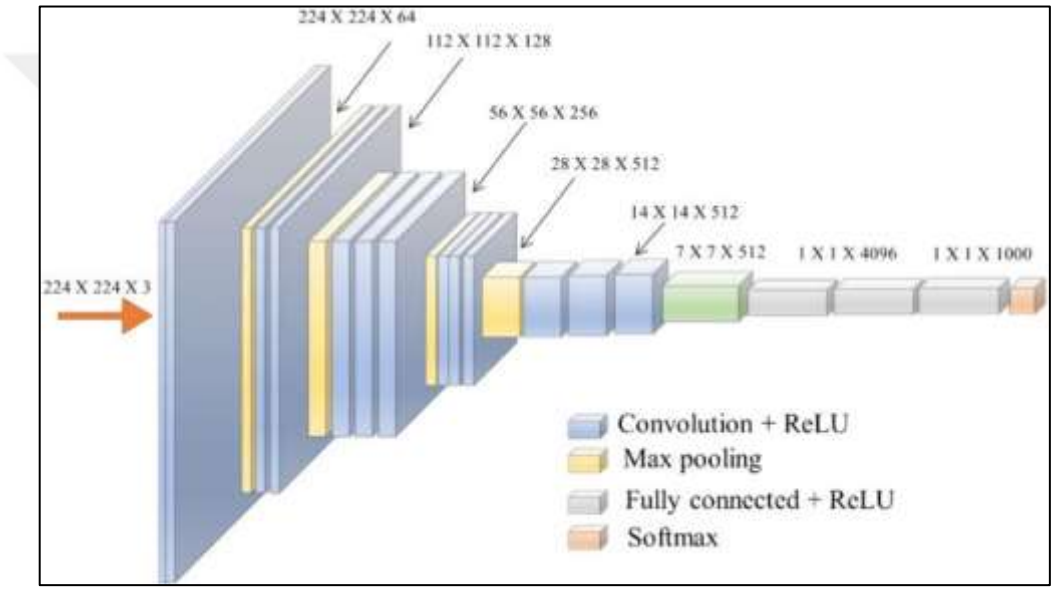

Figure 2.13: VGG 16 Architecture [92].

2.8.2 ResNet 18

ResNet 18 is a deep network architecture developed by Microsoft [95]. It is actually designed to try to enable training deep networks in a new way which is the underpinning skip connections. It has 18 layers, with lots of convolutional layers, activation functions, and pooling layers just so that the model can visibly get good features all the while removing the vanishing gradient problem. Figure 2.14 shows the ResNet 18 architecture.

Initial ResNet 18, since it is pre-trained per the ImageNet dataset, is found to be good enough in extracting features. In this approach, however, it may be difficult to see it right away if we do not first view it on such general data as ImageNet data [96].

ResNet 18 is very popular and applied in many applications. In the application of image classification, it is used between researchers and corporations to find objects and render them into categories inside images. Another example can be seen in medical imaging; that is, ResNet 18 is very good at feature extraction for diagnostic processes about illnesses from histopathological images this shows its importance to computer vision and improving diagnoses. The fact that valuable features can be extracted from data speaks to the important role that ResNet 18 can play in many areas [97].

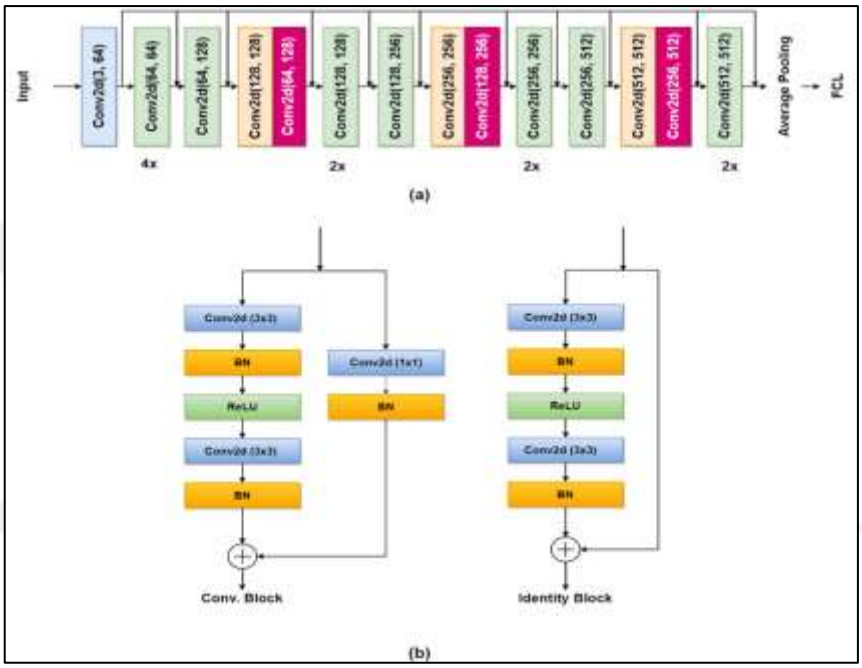

Figure 2.14: (a) Resnet 18 Architecture, (b) Identity Block - Skip Connections.

2.8.3 ResNet 50

ResNet 50 is based on a network architecture termed deep residual learning. Microsoft introduced the network to try to ease and encourage the learning of ultra-deep networks; it operates under a unified theme of identity mappings by making use of skip connections [98]. It has a total of 50 Convolutional Layers Plus Activation and Pooling Layers, which are responsible for digging deeper to come up with much more complex representations of the data and address the vanishing gradient problem well enough. Figure 2.15 shows the ResNet 50 architecture. ResNet 50 is pre-trained on ImageNet, which is more of a standard benchmark for visual recognition tasks; this helps it gain a feature representation foundation for extracting features and illuminating critical features in images [99].

As a model that is widely used and accepted, ResNet 50 finds application in many spheres. Image classification holds promise for research as well as industry by making available the feature to detect objects in images accurately and classify them with good precision. In turn, in the domain of medical imaging, ResNet 50 stands to deliver excellent support in feature extraction for histopathological image analysis and, thereby, disease diagnosis. The appropriateness of extracting meaningful features from quite different datasets proves the importance of ResNet 50 in the domain of computer vision [100].

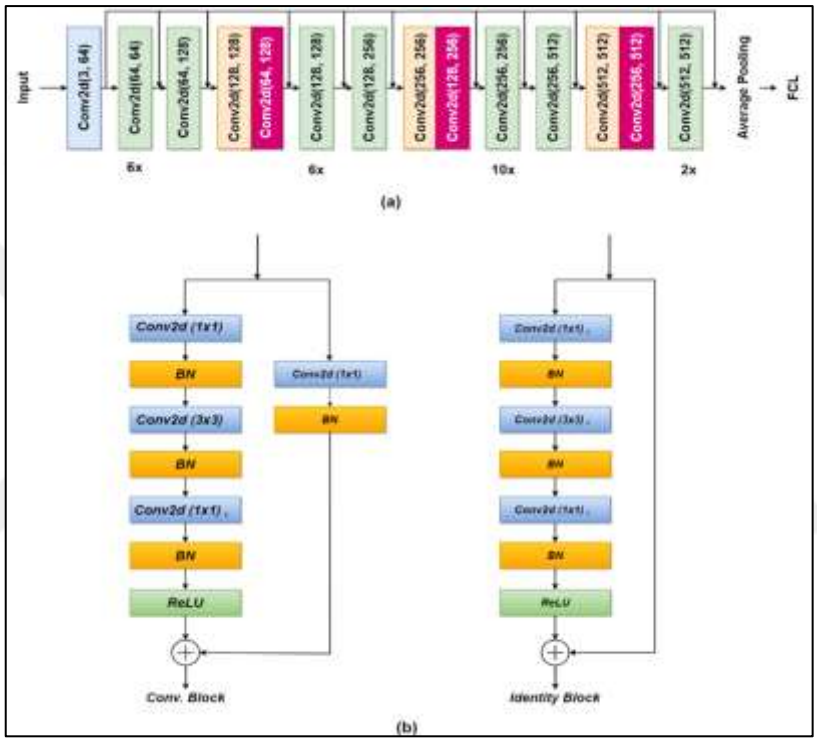

Figure 2.15: (a) Resnet 50 Architecture, (b) Identity Block - Skip Connections.

2.8.4 InceptionNet V1

InceptionNet V1 is a Google-developed architecture of convolutional neural networks, highly innovative in structure for extracting features from an image at different scales simultaneously [101]. The architecture includes a good number of layers containing conventional convolutional, pooling, and inception modules that use diversity in kernel sizes. InceptionNet V1 normally unfolds in 22 layers, which enable the model to grasp complicated patterns in data while computationally quite effective [102]. Figure 2.16 shows the InceptionNet V1 architecture.

It plays a major role in different applications because it can effectively extract features from images. Its early layers are very good at extracting the basic features from an image, so it can be used in different domains [103].

InceptionNet V1 is widely leveraged in image classification, most popularly in the research and business spheres, to identify objects in images and label them with high precision. In another use case, medical imaging seems to find its best application in feature extraction, thereby helping diagnose diseases from histopathological images. This speaks for its place in the field of computer vision and, further on diagnostic errors, will accentuate the imperative contributions the net makes across a range of applications [104].

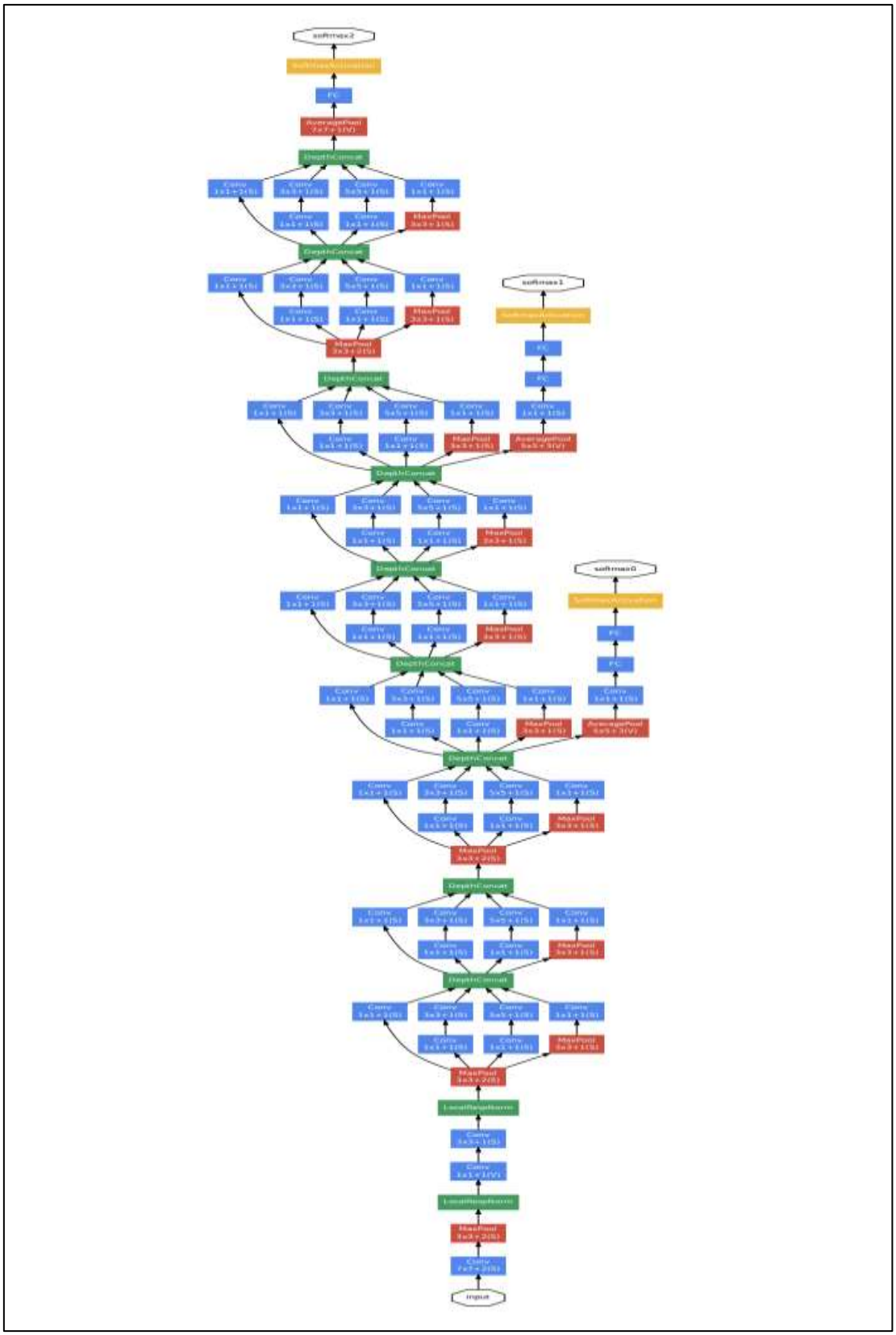

Figure 2.16: InceptionNet V1 Architecture.

2.8.5 MobileNetV3

Google developed MobileNetV3 as a lightweight convolutional neural network architecture to run on mobile and edge devices [105]. It comprises 18 layers of depthwise separable convolutions plus linear bottlenecks and is constructed in a way that minimizes the model's size but not at the cost of performance. This model uses standard and asymmetric convolutions to increase feature extraction capabilities, always keeping computational efficiency in mind [106]. Figure 2.17 shows the MobileNetV3 architecture.

Originally introduced to the world through training on the ImageNet dataset, MobileNetV3 plays a crucial role in different computer vision tasks. One place where its architecture might be most useful is in advanced feature extraction. The advantage of having the first few layers capable of capturing the vital features from the input images makes it a valuable asset for many different applications [107].

In both research and industry settings for object identification and classification problems based on image content, MobileNetV3 finds massive deployment in carrying out image classification tasks for identifying objects in images. Equally potent would be its application in medical imaging, where it extracts features needed for diagnosing ailments from histopathological images. This proves its malleability and significance in advancing computer vision to advance headway and apply it academically [108].

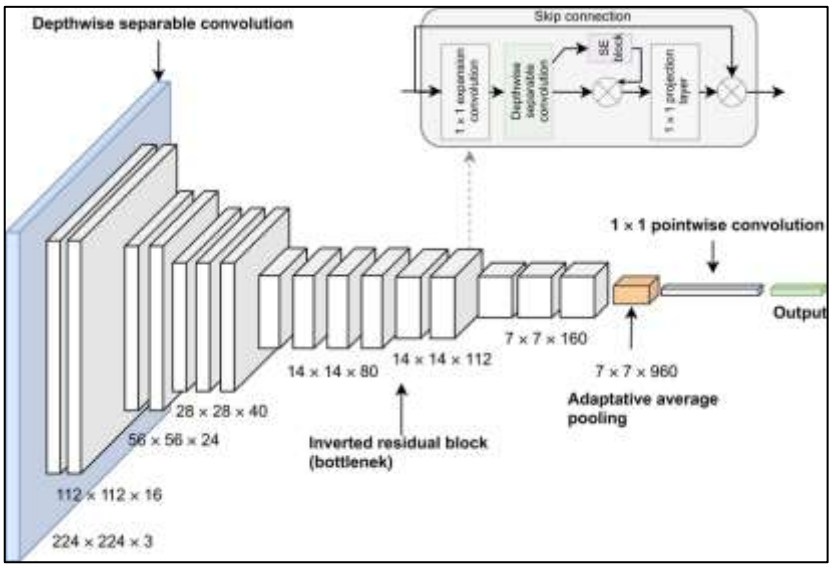

Figure 2.17: MobileNetV3 Architecture.

2.9 TRANSFER LEARNING

Transfer learning strategies are leveraged toward deep learning progress in low data and resource scenarios. Such an approach entails fine-tuning pre-trained models (based on learning-related tasks) to improve their performance for a particular application, thus increasing learning effectiveness and accuracy [109]. This is central to something like medical imaging because, of course, getting labeled data can be a real pain.

In the deep learning, transfer learning often works with DNNs because only some layers are retrained from the pre-trained model. Normally, the first few layers in these networks are created to pick up general features that work well on many tasks. Then, high-level abstraction layers are constructed to make features particular to the target dataset. When most researchers talk about fine-tuning, they typically mean all the pre-trained layers being set as untouchable and only the classification layer being adjusted [110]. For example, while working out the model to carry out the classification of images based on histopathology, the output layer is such that it takes the class labels "malignant" and "benign" [111]. Figure 2.18 presents the transfer learning strategy.

Multiple carefully created datasets have become indispensable in training and testing transfer learning-sensitive models, with the epicenter being the domain of applications in medical imaging [112]. Among those available are the Camelyon and BreakHis datasets, flooded with histopathology images that MobileNetV3 uses to collect features, which should ultimately be properly identified as cancer of the tissues.

It thus appears that most newly trained studies would produce very high performance, as confirmed by the very high performance attained by models optimized on histopathology datasets. Newly trained models have made a vast improvement in the precision of classifying malignant and benign tissues, but these are significantly inferior compared to optimized histopathology models. This finding has repeatedly been found in other fields be it speech recognition, not to mention any aspect of natural language processing, among many others [113].

The applications of transfer learning are wide and still expanding. In addition to medical imaging, the approach is also applied in the field of autonomous driving. For example, pre-trained models serve object detection and scene understanding. Another example is

sentiment analysis, which leverages text data from different datasets to improve model performance [114]. So, this basically means it is a promise, and hence, it attracts lots of interest.

Conclusively, transfer learning is a concept in deep learning by which pre-trained models can smoothly fine-tune themselves very effectively for new tasks and, occasionally, data-poor tasks. Exactly designating fine-tuning layers and handling the classification results would go a long way in leveraging the potential of already existing models to outperform applications in a wide variety of fields, most particularly in the challenging domain of medical imaging.

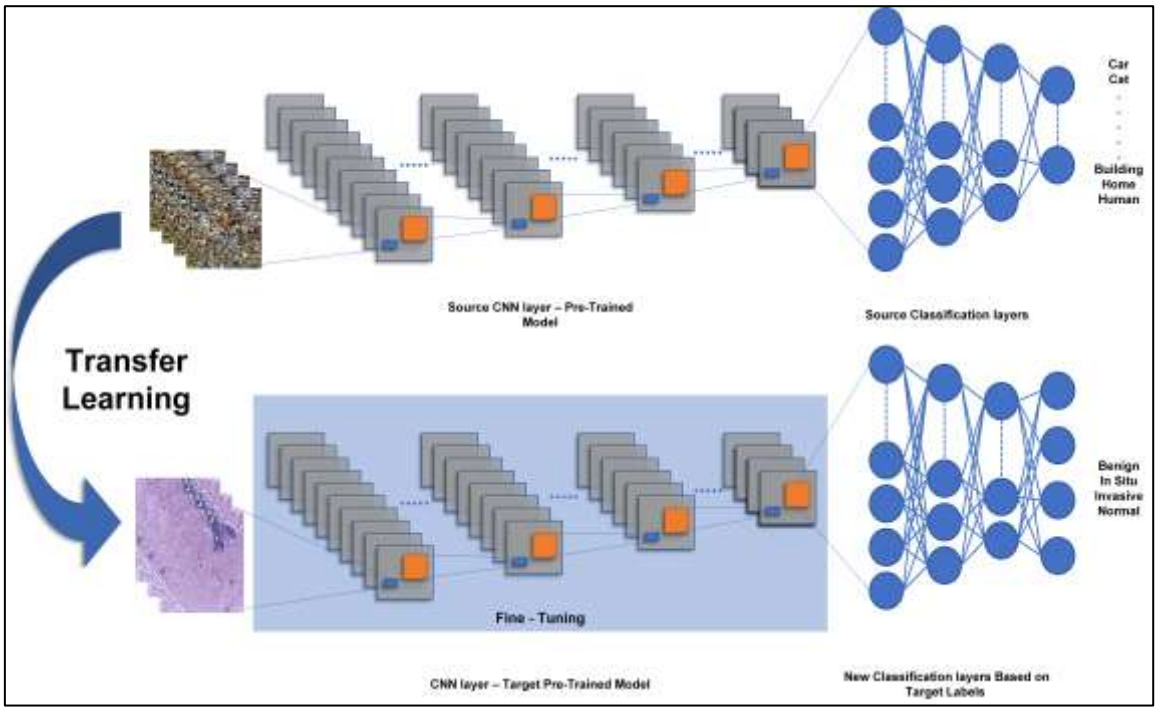

Figure 2.18: Transfer Learning Strategy.

3. HAFMAB-NET: HIERARCHICAL ADAPTIVE FUSION BASED ON MULTILEVEL ATTENTION-ENHANCED BOTTLENECK NEURAL NETWORK

3.1 INTRODUCTION

Diagnostic tools that utilize computer-aided technology are essential in reducing mortality rates, as they enable physicians to identify diseases at an early stage and facilitate accurate diagnosis and treatment planning. Deep learning-based convolutional neural networks have demonstrated their effectiveness in differentiating benign and malignant tumors.

However, conventional CNN architectures often fail to recognize fine morphological details and are inadequate for complex medical image analysis. Common designs have limited receptive fields and use fixed-scale filters, leading to the potential loss or over-extraction of important feature information that is critical for optimal model performance. In addition, typical CNNs tend to perform less well when applied to new datasets or unfamiliar domains due to the generalization of medical imaging difficulties associated with Radiology Scans or Common Pathology Slides.

In 2015, He et al. [115] presented the concept of deep residual learning, which facilitates effective training of very deep neural networks. Their methodology incorporates shortcut connections that enhance gradient propagation, thereby addressing the vanishing gradient problem. A significant contribution of their work is the introduction of bottleneck blocks, an architectural design aimed at optimizing computational efficiency. This architecture consists of three layers: the first layer uses 1 × 1 convolutions for dimensionality reduction, followed by feature extraction via 3 × 3 convolutions, and ends with another 1 × 1 convolution for dimensionality restoration. In this framework, depth does not hinder generalization, and the computational requirements are lower than those typically demanded by networks with similar generalization performance. Furthermore, recent advances in deep learning have leveraged improved constraints on convolutional neural networks (CNNs), particularly in the context of medical applications.

In complicated imaging domains, recent advancements have improved classification and detection efficiency through the use of ensemble learning and hierarchical feature integration

techniques [116] [117]. Transfer learning, which uses pre-trained models to lessen the need for large training datasets, is also becoming more and more acknowledged as a viable strategy for small medical datasets [118]. Nevertheless, notable inter-domain differences across medical imaging modalities negatively impact the advantages of transfer learning, emphasizing the necessity of domain-specific modifications to facilitate generalization during inference.

Attention techniques have shown remarkable efficacy in directing the model's focus to image regions that are significant for diagnostic purposes. These processes have been included in CNN designs to increase sensitivity and resilience, especially with regard to small variation patterns that are crucial for diagnosis [119]. This section, which discusses the task of cancer classification, will introduce our framework, "HAFMAB-Net: Hierarchical Adaptive Fusion based on Multilevel Attention-Enhanced Bottleneck Neural Network," as an introduced solution to this challenge.

3.2 RELATED WORK

The diagnosis of breast cancer via histopathological images has received significant attention; this is due to the advances in deep learning and computer-aided diagnosis. Several methods have been attempted to increase the accuracy of classification, address the similarity between classes, reduce the variance within classes, and improve the model's stability across diverse datasets. These strategies include learning-based transfer, sophisticated structures with attention processes, and hybrid frameworks. Despite their success, many have issues such as restricted flexibility, computational complexity, and the challenge of recording morphological differences. Table 3.1 lists the recent works based on the method and their limitations.

Table 3.1: Summarizes Previous Works Proposed for Cancer Diagnosis Based on Deep Learning Techniques.

	Method	**Limitation**
Alaa et al. [120]	This paper puts forward a deep learning hybrid that combines both unsupervised and supervised approaches to make breast cancer diagnosis more accurate. The histopathology images are segmented into nine parts, which will be further exercised on a set of features. In this research, patterns are created by VAE for a k-means clustering-based process, while the main stage of image classification is carried out by a methodology amalgamating VGG19 and SVM models.	The method merged unsupervised representation learning with a supervised classification technique. Some limitations existed, like using fixed segmentation, which could miss out on morphological variation, the constraints of k-means in spotting complicated links between features, and the lack of an optimization method among parts, which could all bring about less-than-ideal performance.
Musfequa et al. [121]	In this paper, ADBNet, a wide-and-deep convolutional neural network architecture, is presented to enhance the accuracy of breast cancer classification using histopathology images. The model incorporates a newly proposed convolutional attention module for optimizing feature importance extraction and also includes a special block that overcomes scale change problems due to magnification factors, therefore recognizing patterns at different resolution levels. A diffusion methodology applied within an adversarial network for generating realistic synthetic images is also leveraged here so that classification performance can benefit by mitigating training data limitations.	Though ADBNet has advantages in detection and classification, its complex architecture requirements increase computational power, which can be a constraint for deployment on devices with limited resources. Using images generated by GANs also risks biasing the training data; thus, it may affect the model's performance when generalizing to real-world data.
Daniel et al. [122]	The analysts developed a low-cost BCHI-CovNet AI model for sorting histopathological breast cancer. The model employs multi-depth-wise separable convolution scales of various kernel dimensions for catching low— and high-detail patterns, an aggregation module for second-order feature statistics, and long-range centering mechanism feature representation. All these enhancements reduce computation complexity to be run on resource-constrained devices while keeping performance high.	Fixed-sized kernels could inhibit the adaptation to irregular patterns, and pooling that focuses on global features may overlook important classification local patterns.

Table 3.1: Summarizes Previous Works Proposed for Cancer Diagnosis Based on Deep Learning Techniques (Table Continued).

	Method	Limitation
Saif et al. [123]	The authors developed GLNET to participate in detecting breast cancer using a pre-developed ResNet101 and a global-local mechanism for convolution. The method employs a global attribute to describe the color and shape of slide-level features. In contrast, a local feature extractor focuses on important areas to produce dense predictions that improve the classification performance.	Fixed-sized patches cannot recognize small details, while global-local feature combinations would negatively affect the model's ability to enter different datasets.
Chen et al. [124]	The authors proposed an MDFF-Net model of the CNN that combines 1-D with 2-D attributes via a mechanism that promotes enhanced breast pathology classification. Multiscale channel hopping and attention mechanisms that are intelligent enough for capturing the important features of images and ensure that the global and local characteristics are incorporated into the classifier increase the accuracy of the diagnosis.	The finite nature of the extraction process diminishes the model's flexibility to different datasets, and the way the visual data is treated may become complex; this will lead to a limited application of the model in environments with limited resources.
Vivek et al. [125]	The authors investigate GARL-Net, a graph-based regularized learning framework for classifying the cancer of the breast via pathological analysis. They utilize the DenseNet121 model, which brings the benefits of transfer-learning along an `in novel' loss- a function that considers not just the contribution from cross-entropy but also a graph-based term intended to enhance regularization. This grants a way for detecting errors more carefully and pushing more learning unbrokenly because of the feature graph structure kept.	This model uses a static graph Laplacian, which limits its ability to adjust for different datasets. Also, the large computation needed for the graph-based convolutional layers limits its use, mostly in systems with low resources.

Table 3.1: Summarizes Previous Works Proposed for Cancer Diagnosis Based on Deep Learning Techniques (Table Continued).

	Method	**Limitation**
Mohammad et al. [126]	The authors designed a transfer learning scheme using histopathology images as a pre-training task for the binary classification of samples into cancer and normal breast tissue. The scheme aims to overcome limitations regarding cytology data annotation since the features from different domains complement each other and give promising results with different fine-tuning strategies.	Using fixed data for pre-training, such as BreakHis, could limit the model from adjusting to other domains. Also, not having multi-scale feature integration might lower the system's performance, i.e., dealing with resolution changes as well as the morphology looking simple or tricky.
Zhu et al. [127]	Authors designed the Deconv-Transformer (DecT) model for breast cancer classification by using convolutional layers that address color effects. It combines RGB and HED color spaces via residual connections, which keeps important information on different colorations. It applies an attention mechanism for the better extraction of particular features from each channel which eventually increases the accuracy in breast cancer subtype classification.	While residual connections are fixed and serve to preserve information, they may as well do harm in a way that it may limit the model to adapt to different datasets. On the other hand, the applied dual-path architecture increases the computational requirement, which may be a constraint for resource-limited systems.
Pratik et al. [128]	This paper introduces GARL-Net, a fuzzy logic-based ensemble method incorporating the Choquet integral, coalition game theory, and information theory to improve the accuracy of classifying breast cancer. The work aggregates the outputs of five deep learning architectures VGG16, VGG19, Xception, Inception V3, and InceptionResNet V2, taking advantage of the strengths of each in a collaborative manner. This technique tries to solve the problems of computational complexity by introducing new heuristic measures based on fuzzy logic.	Although this technique is capable of improving classification performance, its reliance on a fixed set of classifiers and the computational complexity of the fuzzy system limit its flexibility and scalability. Furthermore, the lack of a mechanism for dynamically emphasizing significant diagnostic features may make it difficult for the system to handle noisy or highly variable data.

Table 3.1: Summarizes Previous Works Proposed for Cancer Diagnosis Based on Deep Learning Techniques (Table Continued).

	Method	Limitation
Yiping et al. [129]	The authors present an Adaptive Resolution Network RANet that works in conjunction with a Support Vector Machine-based anomaly detection approach, which can also be referred to as ADSVM. The mislabeling patches are lessened under the supervision of ADSVM so that the model could be trained more effectively. With RANet, images will be processed through different sub-networks at different levels of resolution but balanced depth. In this particular processing scheme, speed and accuracy of classification have been well compromised, wherein the easily recognizable images were classified at lower resolution and the more complex ones required higher resolution for classification.	Though it gives multiple benefits, this method has restrictions in dealing with data groups that have high levels of change because of set resolution and depth settings. So, in cases where the data layout is not even, the SVM style used to find unusual actions might be likely biased which can impact the total correctness of the system.

3.3 CONTRIBUTIONS

In this chapter, we introduce HAFMAB-Net, a novel approach for classifying histopathology images related to cancer. This method effectively addresses several major limitations noted in previous studies, as discussed in Section 3.2. The introduced model "HAFMAB-Net" leverages an attention mechanism with multiple to focus on diagnostically important regions, which will help overcome the challenges associated with static feature extraction techniques and fixed region analysis.

Adaptive multiscale feature fusion is incorporated to capture global and local morphological features accurately, guaranteeing boosted performance across different datasets and scales. As a result, HAFMAB-Net emerges as an effective algorithm that achieves high accuracy without incurring excessive computational costs or relying on synthetic data.

Additionally, a hybrid loss function combines focus loss with label smoothing to address class-imbalance and enhance classification accuracy. We believe that the proposed model, stands out as a premier solution in its field, improving diagnostic precision and

demonstrating robustness and adaptability in tackling the challenges existing methodologies face.

3.4 METHODOLOGY

3.4.1 Overview

Histopathology images are one of the most critical resources in cancer diagnosis; notwithstanding, they are structurally complex with high cellular composition and ionic details. Meanwhile, even in patients with the same type of cancer, the patterns in the corresponding tissues may greatly differ. Information regarding rare cancers is often limited, which presents challenges in developing a highly accurate and efficient machine-learning model.

Deep learning methods have emerged as one of the most fruitful techniques for cancer detection using histopathology images. The method uses deep architecture and attention-based CNN layers, which help the model focus on important features while downplaying less important ones. It is a dynamic method implementation that combines features across various sources.

The learn features across the deep architecture models can, at the lowest levels, be directly linked to some basic properties of objects- for example, edges and textures. Features at the next level of abstraction- for example, patterns and shapes- are learned at much deeper levels. Putting attention mechanisms into these models makes the model focus on important parts and improves tissue-type discrimination. It makes the model more robust against the noise and variability within the dataset.

Also, the rewards of this way are larger through change joining, which mixes levels of a structure to make a full info display covering low, middle, and high parts of global and local layouts. This joining helps feature reset and model reweight, thus boosting the model's ability to generalize across different types of histopathological tissues.

This study introduces HAFMAB-Net; the HAFMAB-Net protocol is illustrated in Figure 3.1. Several strategies have been incorporated into HAFMAB-Net to mitigate the risk of overfitting. A dropout layer is applied after each convolutional layer at a rate of 0.25,

meaning neurons are randomly deactivated during training. This approach helps to avoid overfitting by reducing the mutual adaptation of neurons.

Moreover, an L1 regularization technique is employed across layers to encourage a sparse distribution of learned weights, which improves the model's generalization capacity. During the pre-processing stage, input images are transformed to standardized scale values; this confirms coherence and minimizes the impact of data variability. All these strategies enhance the proposed model's effectiveness across various datasets.

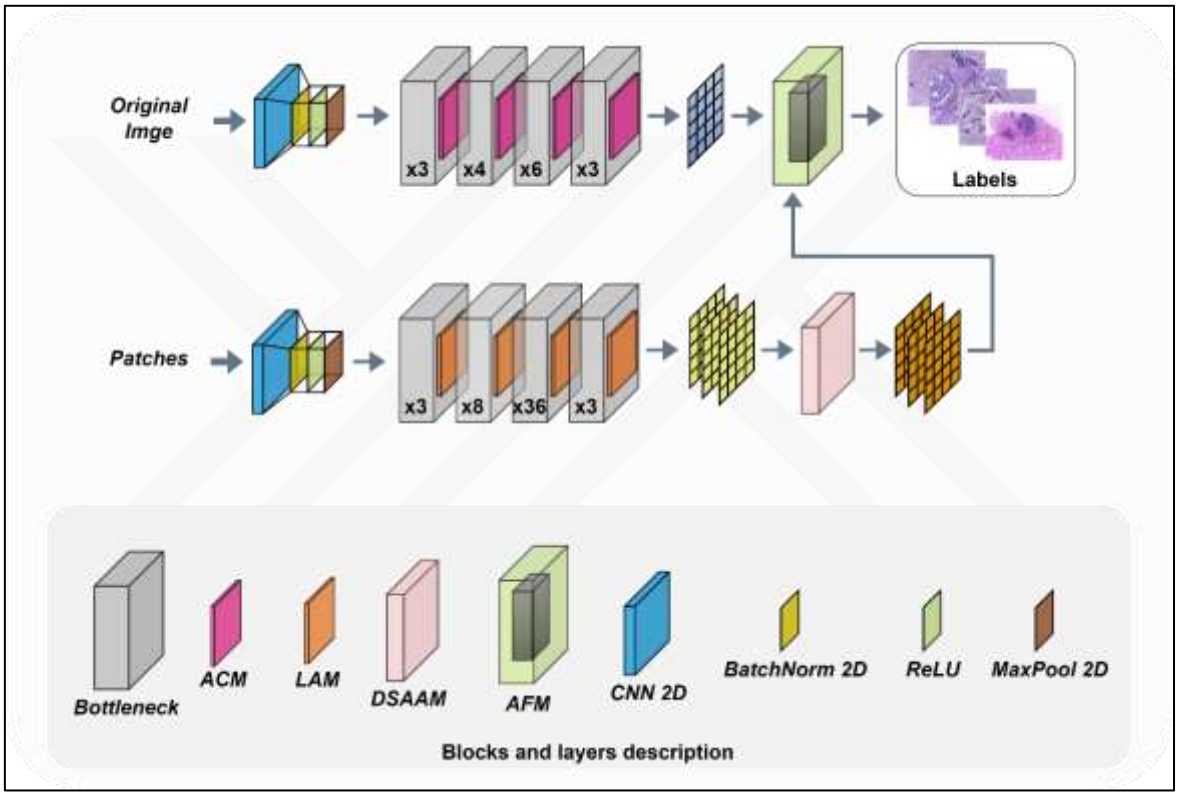

Figure 3.1: Architecture of Proposed HAFMAB-Net.

3.4.2 Bottleneck Architecture

The bottleneck is a crucial structural element; it enhances computational regulation and optimizes resource utilization in deep learning models while maintaining performance levels. Its mechanism significantly reduces feature dimensionality by adjusting the number of channels / spatial resolution using convolutional neural network (CNN) layers [130].

As illustrated in Figure 3.2, an adaptation of the bottleneck structure incorporates an attention module. It consists of two parallel paths. The first path features three layers of 2D

CNN (kernel sizes of 1x1, 3x3, and 1x1, respectively). These layers sequentially reduce dimensionality, perform fundamental arithmetic operations to recover the original feature dimensions, and generate feature maps with dimensions represented as $\left(M_{Image}^{CNN} / M_{Patch}^{CNN}\right) \in R^{BxCxHxW}, [B, C, H, W]\right)$, resulting in a shape of $R^{BxCxHxW}$, where $[B, C, H, W]$ denote batch size, channels, height, and width.

After the third 2D CNN layer, the attention module (which includes either Channel Attention or Local Attention) is integrated to enhance the accuracy of feature map representation by assigning significant weights to features based on their location or channel. This is represented as $(M_{attention})$ yielding shape $\in R^{BxC_{attenion}xHxW}, [B, C_{attention}, H, W]$.

The second parallel path consists of one 2D CNN layer with a 1x1 kernel size, serving as a downsampling mechanism that reduces the spatial resolution while effectively processing the hierarchical feature representation. This results in a feature map with dimensions $\left(M_{Image}^{\text{downsampling}} / M_{Patch}^{\text{downsampling}}\right)$, also yielding a shape of $\in R^{BxCxHxW}, [B, C, H, W]$.

An element-wise summation process integrates the enriched contextual information from the attention mechanism with the downsampled representation, thereby enhancing feature discrimination and preserving salient information. This is expressed as information $\left(M_{Image}^{Bottleneck} / M_{Patch}^{Bottleneck}\right)$, and results in a shape of $\in R^{BxCxHxW}, [B, C, H, W]$.

This strategy balances between (local and global) features, allowing the network to leverage attention-driven specificity alongside general representations. Eliminating the introduction of additional parameters enhances computational efficiency and facilitates smooth gradient flow. This contributes to greater training stability and helps alleviate the vanishing gradient problem. Consequently, the resulting merged output raises the feature space, enhancing the network's expressiveness and generalization capacity.

In summary, this mechanism contributes to improved accuracy and performance in tasks requiring detailed feature analysis. The formula for the improved bottleneck is expressed as follows:

$$M_{Image}^{CNN} = BN\left(Conv2D\left(ReLU\left(BN\left(Conv2D\left(ReLU\left(BN\left(Conv2D\,(X)\right)\right)\right)\right)\right)\right)\right) \quad (3.1)$$

$$M_{Image}^{\text{downsampling}} = BN\left(Conv2D(x)\right) \tag{3.2}$$

$$M_{attention} = M_{Image}^{CNN} \rightarrow Attention\ Module \tag{3.3}$$

$$M_{forward} = M_{Image}^{CNN} \otimes M_{attention} \tag{3.4}$$

$$M_{image}^{Bottleneck} = ReLU\left(M_{forward} \oplus M_{image}^{\text{downsampling}}\right) \tag{3.5}$$

For Patch

$$M_{Patch}^{CNN} = BN\left(Conv2D\left(ReLU\left(BN\left(Conv2D\left(ReLU\left(BN\left(Conv2D\left(\tilde{x}\right)\right)\right)\right)\right)\right)\right)\right) \tag{3.6}$$

$$M_{Patch}^{\text{downsampling}} = BN\left(Conv2D(x)\right) \tag{3.7}$$

$$M_{attention} = M_{Patch}^{CNN} \rightarrow Attention\ Module \tag{3.8}$$

$$M_{forward} = M_{Patch}^{CNN} \otimes M_{attention} \tag{3.9}$$

$$M_{Patch}^{Bottleneck} = ReLU\left(M_{forward} \oplus M_{Patch}^{\text{downsampling}}\right) \tag{3.10}$$

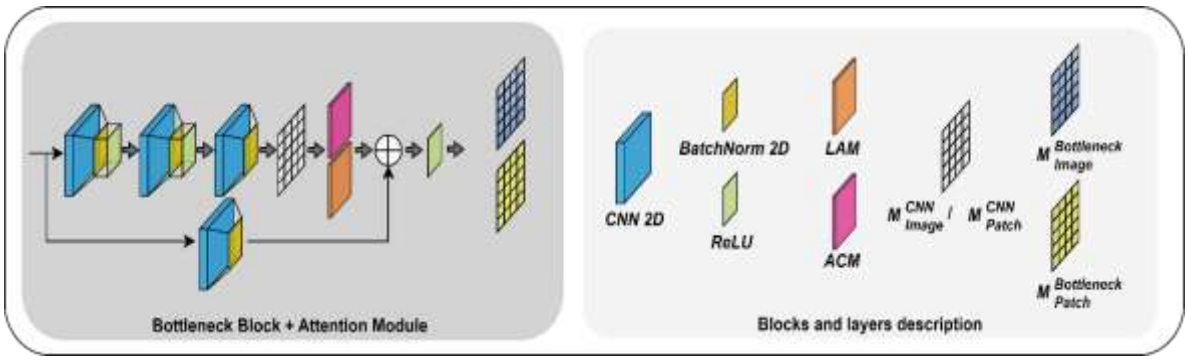

Figure 3.2: Architecture of Bottleneck Block Integrated with Attention Module.

3.4.3 Attention Channel Module (ACM)

Figure 3.3 presents the neural network architecture of ACM. It consists of three layers of CNN in 2D with a filter 1x1 to reduce the channel dimension, improve the transformation between channels, and restore the original channel dimensions. The ACM architecture works

on improving the representation of feature maps by emphasizing the most informative channels through allocating attention to different input channels, enabling the network to selectively amplify or diminish features based on their relevance to the specific task [131].

The attention weight is calculated after the third-CNN layer by applying Sigmoid function, represented by $(M_{ACM\ attention})$. These feature map attention weights are then multiplied element-wise by the input feature map $\left(M_{Image}^{CNN} \in R^{BxCxHxW}, [B, C, H, W]\right)$, producing the modified feature map $\left(M_{Image}^{ACM} \in R^{BxCxHxW}, [B, C, H, W]\right)$. This reweighting process for the feature map channels dynamically enhances the representation for further processing, thereby improving the network's capacity to identify significant patterns in the image while minimizing noise and irrelevant features. The formula for the ACM expression is as follows:

$$M_{ACM\ attention} = sigmoid(Conv2D(\ SELU\ (Conv2D\ (SELU(Conv2D(M_{Image}^{CNN})))) \quad (3.11)$$

$$M_{Image}^{ACM} = M_{Image}^{CNN} \otimes M_{ACM\ attention} \quad (3.12)$$

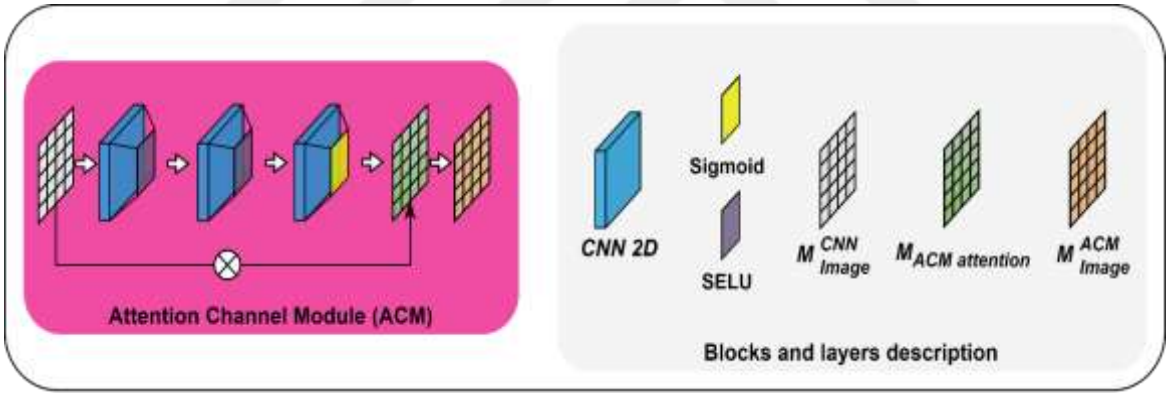

Figure 3.3: Architecture of Attention Channel Module (ACM).

3.4.4 Local Attention Module (LAM)

Figure 3.4 presents the architecture of the LAM, which is focused on the specific regions of the (M_{Patch}^{CNN}) to capture different pattern and context-specific details accurately. The architecture of the LAM focuses on locative area according to sliding windows, and CNN layers through applying the attention weight to each region [132]. The block of LAM uses two pooling operation layers Average and Max along the Channel of the (M_{Patch}^{CNN}), respectively. The operations of the Averaged Max pooling layers capture the spatial information into two feature maps concatenated via element-wise-summation operation

generating a new feature map, denoted as $(sum\ M_{Patch}^{CNN})$ with dimensions of size $\in R^{Bx2xHxW}, [B, 2, H, W]$.

The $(sum\ M_{Patch}^{CNN})$ is processed through parallel 2D CNN layer with filter of (3x3, 7x7, and 11x11) followed with Sigmoid function, these parallel CNN layers Sigmoid work on capturing the localized spatial dependencies features and determining the attention weights along multiple scales; low, medium, and high, which later combined using element-wise addition, resulting $(M_{LAM\ attention})$. This $(M_{LAM\ attention})$ is multiplied by the (M_{Patch}^{CNN}) via element-wise multiplication, resulting the (M_{Patch}^{LAM}) with dimensions $\in R^{BxCxHxW}, [B, C, H, W]$.

The LAM mechanism improves the Bottleneck network's robustness by focusing attention on significant features, and reduces the impact of irrelevant or noisy elements in particular areas of the data. This approach improves generalization by focusing on intricate patterns while capturing contextual information within a confined scope. The corresponding formula for representing LAM is provided as well:

$$Max\ M_{Patch}^{CNN} = Adaptive\ MaxPool\ (\ M_{Patch}^{CNN}) \tag{3.13}$$

$$Avg\ M_{Patch}^{CNN} = Adaptive\ AvgPool\ (\ M_{Patch}^{CNN}) \tag{3.14}$$

$$sum\ M_{Patch}^{CNN} = Max\ M_{Patch}^{CNN} + Avg\ M_{Patch}^{CNN} \tag{3.15}$$

$$M_{LAM\ attention} = For\ i\ in\ [3,7,11] \rightarrow \sum_i Sigmoid\ (Conv2D^i\ (\ sum\ M_{Patch}^{CNN}) \tag{3.16}$$

$$M_{Patch}^{LAM} = M_{Patch}^{CNN} \otimes M_{LAM\ attention} \tag{3.17}$$

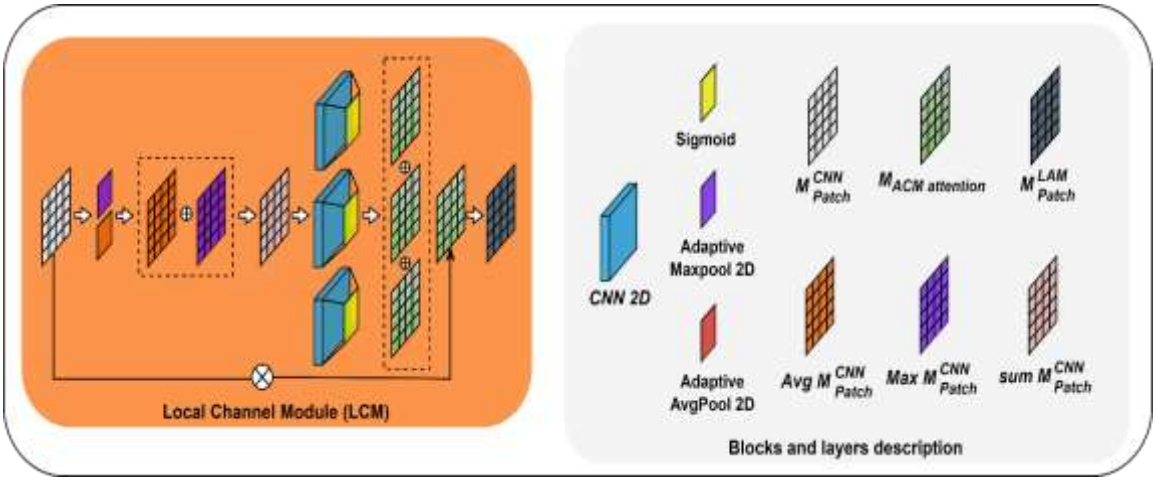

Figure 3.4: Architecture of Local Attention Module (LAM).

3.4.5 Deeper Spatial Attention Aggregator Module (DSAAM)

Figure 3.5 presents the enhanced neural network architecture of the DSAAM, which is designed to integrate multiple levels of attention within a single patch. This approach enhances the patch's ability to represent various features. The DSAAM network comprises three 2D sequential CNN layers, which work together to progressively reduce the channels' dimensionality. This procedure obtains neighbourhood information associated with ($M_{Patch}^{Bottleneck}$) regarding the refined fine-grained details while preserving the spatial resolution. The output of the third CNN layer is then fed to a Sigmoid layer that computes the attention weights as ($M_{DSAAM\ attention} \in R^{Bx1xHxW}, [B,1,H,W]$). The DSAAM process involves ($M_{Patch}^{Bottleneck}$) generating three individual attention patches, these are then added together across channels to produce the spatial version of DSAAM attention ($M_{spatial\ DSAAM\ attention} \in R^{Bx1xHxW}, [B,1,H,W]$). This spatial attention ($M_{DSAAM\ attention}$) multiplied by the number of elements in the set, which is 3, results in a more effective representation in enhancing important spatial regions while reducing less significant regions ($M_{Patch}^{DSAAM} \in R^{BxCxHxW}, [B,C,H,W]$). DSAAM offers several key benefits that enhance its effectiveness. First, it improves the ability to represent features by concentrating on important spatial regions, which helps to increase the distinctiveness of aggregate features. Second, its inherent flexibility and modularity enable easy integration into existing feature extraction processes and facilitate the development of more complex architectural designs. Furthermore, the hierarchical pipeline reduction in the framework leads to a decrease in computational cost while maintaining the attention focus on the spatial

domain. Finally, the ability to provide multiple attention maps with different spatial resolutions enables a versatile capacity for context fusion, this allows the module to effectively capture different spatial patterns associated with larger scale input tasks or larger patches. The formula for expressing DSAAM is given as follows:

$$M_{spatial\ DSAAM\ attention} = \sum_{i=1}^{3} sigmoid(Conv2D(\ SELU\ (Conv2D\ (SELU(Conv2D(M_{Patch}^{Bottleneck}{}_{i}))))) \quad (3.18)$$

$$M_{Patch}^{DSAAM} = M_{Patch}^{Bottleneck} \otimes M_{spatial\ DSAAM\ attention} \quad (3.19)$$

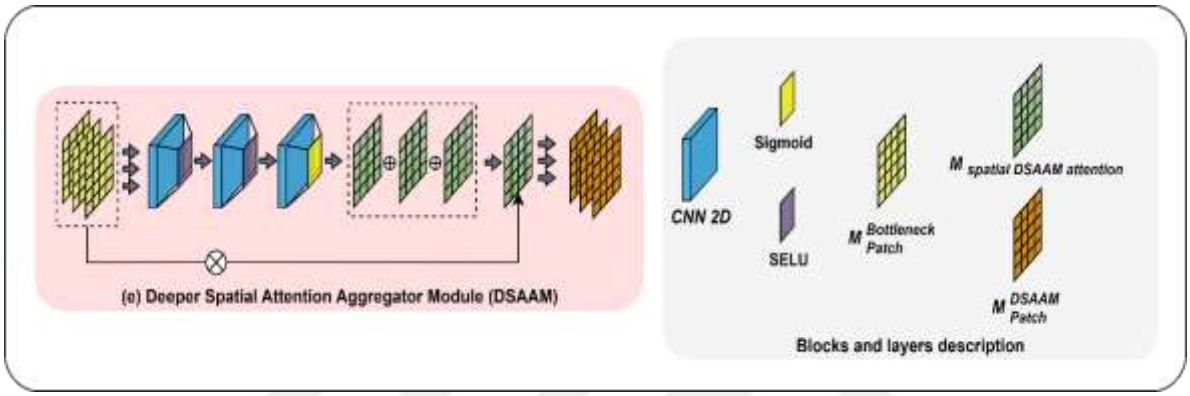

Figure 3.5: Architecture of Deeper Spatial Attention Aggregator Module (DSAAM).

3.4.6 Adaptive Fusion Module (AFM)

A dynamically designed Adaptive Fused Module (AFM) is applied in the proposed network, it works on integrating three components: (M_{Patch}^{DSAAM}) and ($M_{Image}^{Bottleneck}$). These components work together to enhance the network's performance. This integration results in comprehensive attention weights that focus on the most important total characteristics. The block diagram of the designed AFM is depicted in Figure 3.6.

The initial phase of the AFM involves two aggregation procedures: first, the combination of the three (M_{Patch}^{DSAAM}) creates ($M_{aggregation-Patch}^{DSAAM} \in R^{BxCxHxW}, [B, C, H, W]$); second, the aggregation of ($M_{aggregation-Patch}^{DSAAM}$) with ($M_{Image}^{Bottleneck}$) creates ($M_{\text{aggregative}}^{AFM} \in R^{BxCxHxW}, [B, C, H, W]$). Both steps are implemented through element-by-element addition.

The computation of comprehensive attention weights is achieved through multiple layers, concluding with a Sigmoid layer that produces ($M_{Comprehensive\ attention}^{AFM} \in R^{BxCx1x1}, [B, C, 1,1]$). This ($M_{Comprehensive\ attention}^{AFM}$) is used to augment

$\left(M_{Image}^{Bottleneck} \ and \ M_{aggregation-Patch}^{DSAAM}\right)$ through element-by-element multiplication, resulting in ($M_{Comprehensive}^{AFM} R^{BxCxHxW}$, $[B, C, H, W]$), which are then combined to produce ($M_{Comprehensive}^{AFM}$). The proposed method for the Aggregated Feature Model (AFM) offers several advantages. Firstly, the aggregation process enhances feature representation by broadening local and global coverage. Furthermore, using a series of fully connected layers followed by SELU activation functions improves the quality of feature imitation, introduces nonlinearity, and increases the model's learning capacity.

he ($M_{Comprehensive}^{AFM}$) is passed to the classifier block, it includes two sequence pooling layers (Max and Adaptive Max pooling layers) to reduce the dimension of channels, followed by Flatten layer works on reshape the feature map into two 2D [B, C], to be ready for the FC layers, followed by the SoftMax to calculate the probability classification of the class label.

Additionally, combining the enhanced aggregated patch-DSAMM with the enhanced image bottleneck selectively amplifies the most critical components. This integration merges global context with intricate spatial pieces of information. Consequently, this developed design fosters dynamic-integration, enhances computational capacity, promotes scalability, and improves interpretively crucial benefits for tasks requiring a robust and context-aware learning approach. The formula for expressing the Aggregated Feature Model (AFM) is as follows:

$$M_{\text{aggregative}}^{AFM} = M_{aggregation-Patch}^{DSAAM} \oplus M_{Image}^{Bottleneck} \tag{3.20}$$

$$M_{Comprehensive\ attention}^{AFM} = M_{\text{aggregative}}^{AFM} \rightarrow AFM\ block_{[C]} \tag{3.21}$$

$$M_{Image}^{AFM} = M_{Comprehensive\ attention}^{AFM} \otimes M_{Image}^{Bottleneck} \tag{3.22}$$

$$M_{Patch}^{AFM} = M_{Comprehensive\ attention}^{AFM} \otimes M_{aggregation-Patch}^{DSAAM} \tag{3.23}$$

$$M_{Comprehensive}^{AFM} = M_{Patch}^{AFM} \oplus M_{Image}^{AFM} \tag{3.24}$$

And, the expression formula of the classifier can be given as follows;

$$\hat{y} = Softmax\left(W * flatten\left(Adaptive\ Max\ Pooling^{2D}\left(Max\ Pooling^{2D}(x)\right)\right)\right) \tag{3.25}$$

where, $\hat{y}$, and W are output; weight of the FCL, respectively.

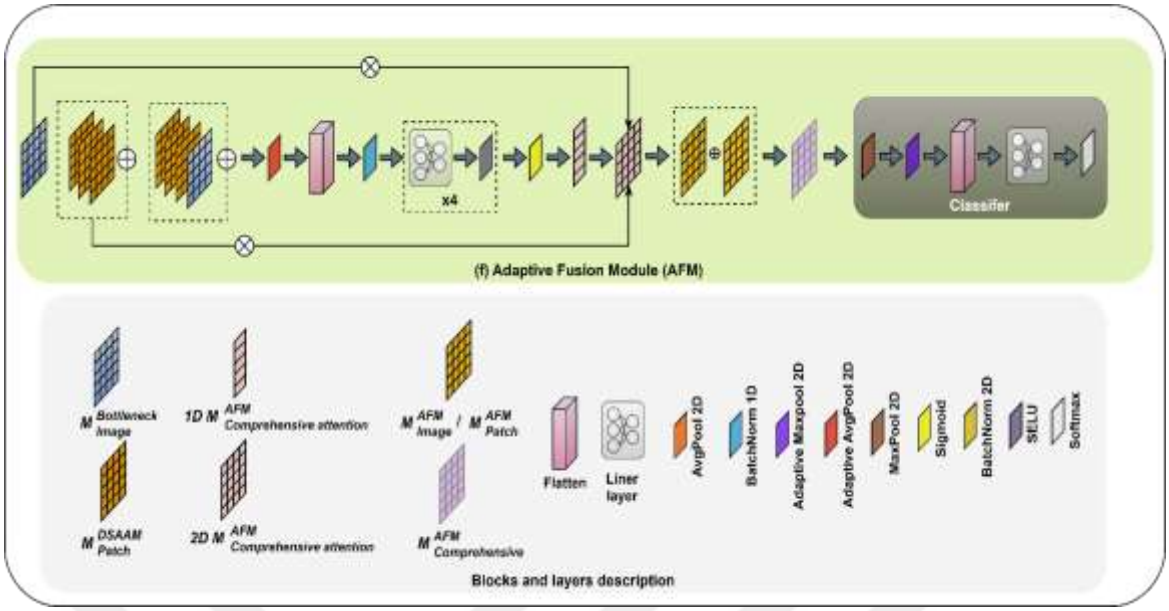

Figure 3.6: Architecture of Dynamic Adaptive Fusion Module (AFM).

3.5 EXPERIMENTS RESULTS AND DISCUSSION

3.5.1 Dataset Pre-processing

The first step during the training and evaluation phase is data preparation. For the BACH-data, it is split into training and evaluation at 80% and 20%, respectively. For the BreaKHis-data; all four amplification factors of the sub-labels are merged according to their main labels, and it is split into training, validation, and testing at 80%, 10%, and 10%, respectively. The same BreaKHis-data split is applied to the colon, lung, and LC25000-datasets. More information concerning how the datasets were divided can be found in Figure 3.7.

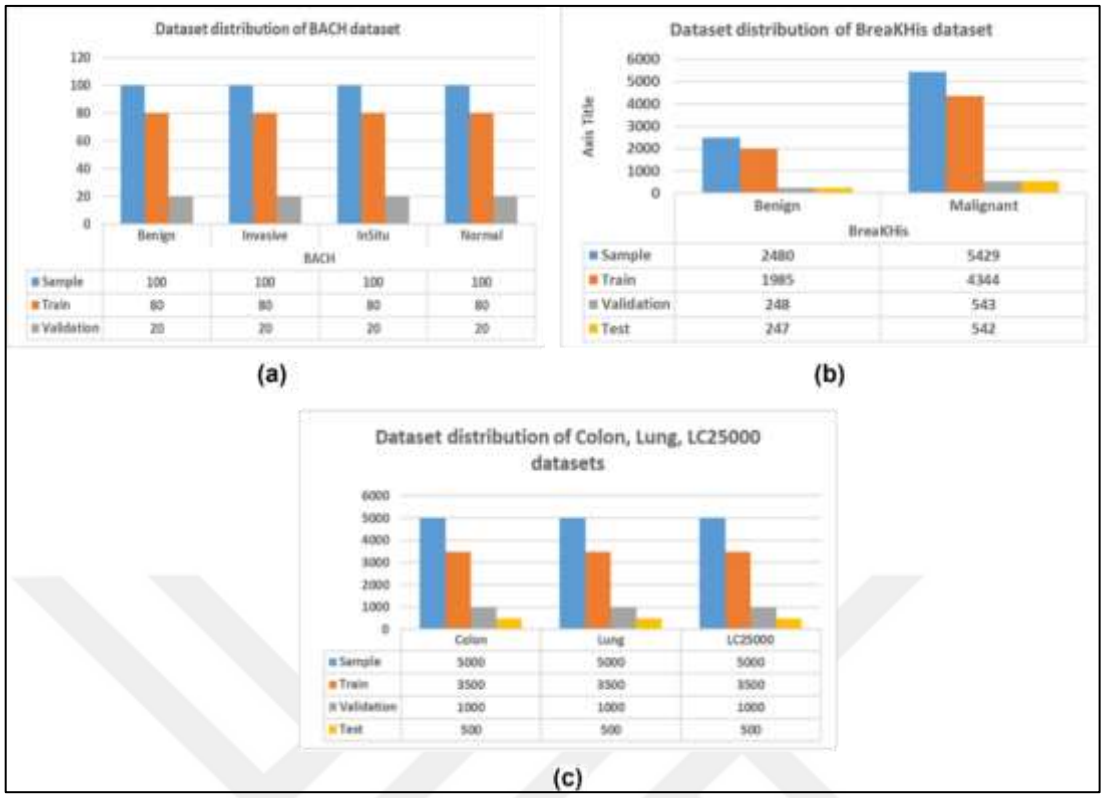

Figure 3.7: Distribution of Dataset (a) BACH-Dataset, (b) BreaKHis-Dataset, (c) Colon-Dataset, Lung-Dataset, And LC25000-Dataset.

After the partitioning phase, three key steps were taken in the pre-processing that contributed greatly to the whole process. The first step is data augmentation, such as mirror-flipping samples, rotation angles, and adding noise introduction to training samples. It increases diversity and robustness within training data by simulating real-world variations, thus enhancing the model's generalization capabilities. Due to the augmentation applied herein, at least eight times have been ensured to increase the size of the training dataset. It should be noted that all datasets apart from LC25000 have borne this particular means of augmentation.

The next step is patch extraction, which takes three different patches from the chosen pictures using a set window size. This method is important to find areas of interest in the original pictures so that the model can see spatial patterns linked to particular classes. More views help patch extraction improve the model's understanding of each class because there is more variability. To keep the data safe, the sample patches are worked on separately, ensuring they do not overlap.

The last step includes changing and resizing the data to one side, keeping the pixel values within a certain range, and turning them into the tensor format. This process makes the data uniform for entering the neural network, ready for successful training and matching with what is needed by the artificial network structure. The pre-processing procedure is shown in Figure 3.8.

It should be noted that the patch addition and extraction procedures are carried out only on the training samples. To enable an unbiased assessment of the model's performance, the validation and test samples must remain untouched for the entire cross-validation sequence.

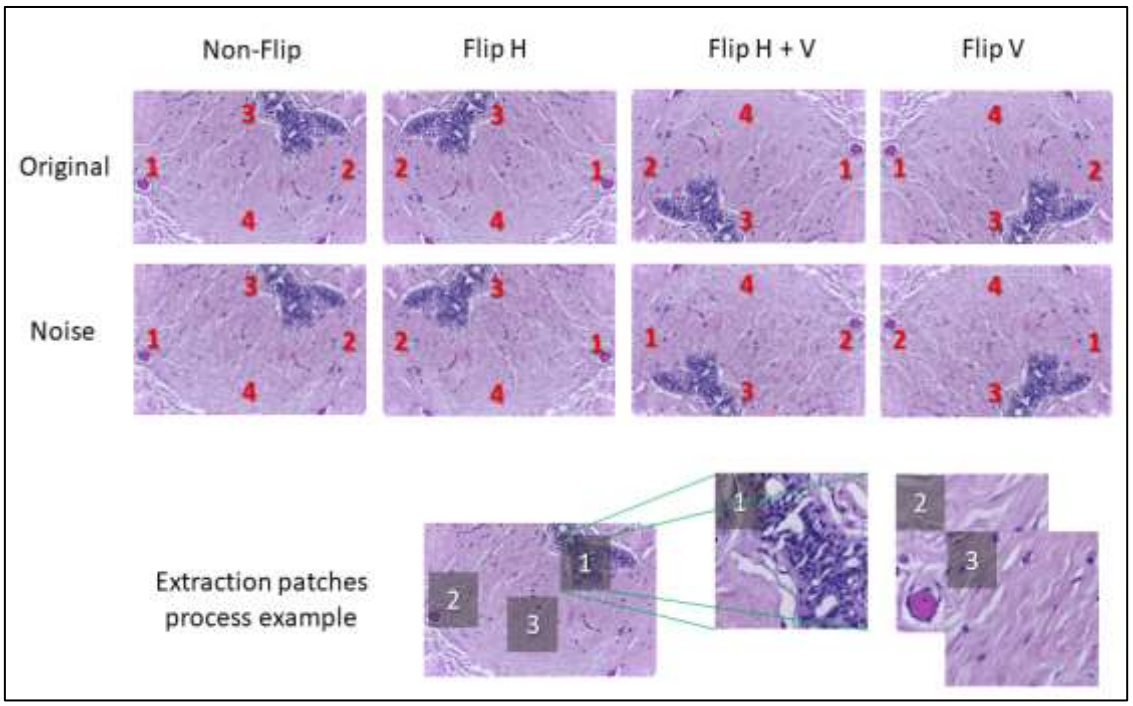

Figure 3.8: Pre-processing Process.

3.5.2 Loss Function and Parameters Setting

The proposed model "HAFMAB-Net" uses a combination loss function to improve the generalization capacity and the performance of the training process in classification samples. The loss function uses a combination of label-smoothing cross-entropy loss $\boldsymbol{L_{CEsmooth}}$ and focal loss $\boldsymbol{L_{FL}}$ according to the modifiable weights to equilibrate their contributions. The (3.26) equation provides the mathematical expression of the combined loss function

$$\boldsymbol{L_{combined}} = \alpha . L_{CEsmooth} + \beta . L_{FL} \tag{3.26}$$

In this equation, the parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ serve as weighting factors for $\boldsymbol{L_{CEsmooth}}$ and $\boldsymbol{L_{FL}}$, respectively, which determine their relative impacts.

The $\boldsymbol{L_{CEsmooth}}$ incorporates smooth factors that soften the rigidity of target labels. This approach can help mitigate overfitting and decrease excessive confidence in model predictions. Redistributing a portion of the probability mass to less likely classes contributes to achieving more stable outcomes. The (3.27) equation provides the mathematical expression of the $\boldsymbol{L_{CEsmooth}}$.

$$\boldsymbol{L_{CEsmooth}} = -\sum_{i=1}^{C}\left[(1-s).y_i + s.\frac{1}{c}\right].\log(\hat{y}_i) \tag{3.27}$$

this formulation, $\boldsymbol{C}$ denotes the total number of classes, $\boldsymbol{s}$ denotes the label smoothing factor (usually a small value such as 0.1), $\boldsymbol{y_i}$ denotes the actual one-hot encoded label for class $\boldsymbol{i}$, $\boldsymbol{\hat{y}_i}$ refers to the predicted probability for class$\boldsymbol{i}$.

The design of this function helps mitigate the network's tendency to focus overly on predictions and improves robustness and generalization on small datasets.

Focal Loss effectively addresses the class imbalance issue by adaptively adjusting the loss based on the classification difficulty of each sample. Samples that are more challenging to classify receive a higher weight, while those that are easier to classify are assigned a lower weight. This approach has proven to be highly effective for datasets that exhibit class imbalance. The (3.28) equation provides the mathematical expression of the $\boldsymbol{L_{FL}}$.

$$\boldsymbol{L_{FL}} = -\mu * (1-p_t)^{\gamma} * \log(p_t) \tag{3.28}$$

In this equation, the variable μ denotes the balancing factor for the correct class, while p_t denotes the predicted probability of the true class (where $p_t = \hat{y}_i$ when $y_i = 1$). The parameter γ acts as a focus parameter, regulating the degree of emphasis given to difficult examples, which is usually set to a value of 2.

The $\boldsymbol{L_{FL}}$ prioritizes the network's learning from difficult examples, allowing it to improve its performance in handling minority classes.

This combination loss offers a magnificent approach to make a balance between performance and generalization; the Cross-Entropy Loss with Label Smoothing addresses

the overfitting by reducing it via promoting stable learning for the network, and the Focal Loss focuses on the enhancement of the network's adaptability with the classification difficult sample via emphasizing challenging samples

3.5.3 Experimental Setup

This study uses Python version of 3.9 and employs the library of PyTorch, which is considered an efficient environment for building and developing a high-performance application such as the proposed model. The proposed model has been implemented using a system with (Nvidia GeForce GTX with GPU of 1660 Ti and VRAM of 6GB providing efficient matrix operations execution and computations in DNN. Also, the system has an Central Processing Unit with Intel Core i7-9750H and 16 GB of RAM, thus, offering an efficient pre-processing data and batch loading. Table 3.2 shows a summary of the hyperparameters that were used to train HAFMAB-Net on each dataset: learning rate, batch size, stride size epochs, and many others, which were tuned painstakingly for each dataset so that it could be robust and reliable.

Table 3.2: Hyperparameters Training of HAFMAB-Net Based on Datasets.

<table>
<tr><th>Dataset</th><th>Optimization</th><th>Learning Rate</th><th>Batch Size</th><th>Alpha</th><th>Beta</th><th>Label Smooth Factor</th><th>Step Size</th><th>Epoch</th></tr>
<tr><td>BACH</td><td rowspan="3">Adam</td><td rowspan="2">10^{-4}</td><td>10</td><td>0.5</td><td>0.4</td><td rowspan="3">0.1</td><td rowspan="2">5</td><td rowspan="3">150</td></tr>
<tr><td>BreaKHis</td><td>15</td><td>0.6</td><td rowspan="2">0.5</td></tr>
<tr><td>LC 25000</td><td>10^{-3}</td><td>40</td><td>0.5</td><td>10</td></tr>
</table>

3.5.4 Evaluation Matrix

The terms (TP), (FP), (TN), and (FN) are used to calculate the evaluation metrics, such as accuracy, precision, recall, and F1 score, that are used to evaluate the proposed model's performance. These metrics offer a clear report of the performance across various datasets [46].

i. $Accuracy = \frac{TP+TN}{TF+FN+TP+FP}$ (3.29)

ii. $Precision = \frac{TP}{TP+FP}$ (3.30)

iii. $Recall = \frac{TP}{TP+FN}$ (3.31)

iv. $\mathrm{F1-Score} = 2\,.\ \frac{(Precision \times Recall)}{(Precision + Recall)}$ (3.32)

3.5.5 Results

To confirm HAFMAB-Net's strength and efficiency in classifying histopathology images, it was independently assessed with conventional deep-learning models on all datasets. This approach enables HAFMAB-Net to evolve based on each dataset's peculiar requirements and challenges, thus leading to a more precise and contextual appraisal. With customized training applied to each dataset, our goal is to reach maximum performance while keeping an element of objectivity in the assessment process. The next section will discuss and present the results of the independent evaluations.

3.5.5.1 Results of BACH-Dataset

The accuracy curves for training and validation of the proposed model "HAFMAB-Net" for 150 epochs on the BACH dataset are shown in Figure 3.9. During the training process, model accuracy was about 40%, increasing until it reached 100% at the end of training. On the other hand, the validation rate started at close to 20% and then shot up very quickly in the initial phase before finally stabilizing to around 99% at the end of training. Following this initial period of assessment, a thorough analysis of student performance and validity results indicates effective instruction with minimal overfitting. In addition to strengthening generalization, this study highlights the model's ability to provide accurate classifications based on the BACH In-Instance dataset.

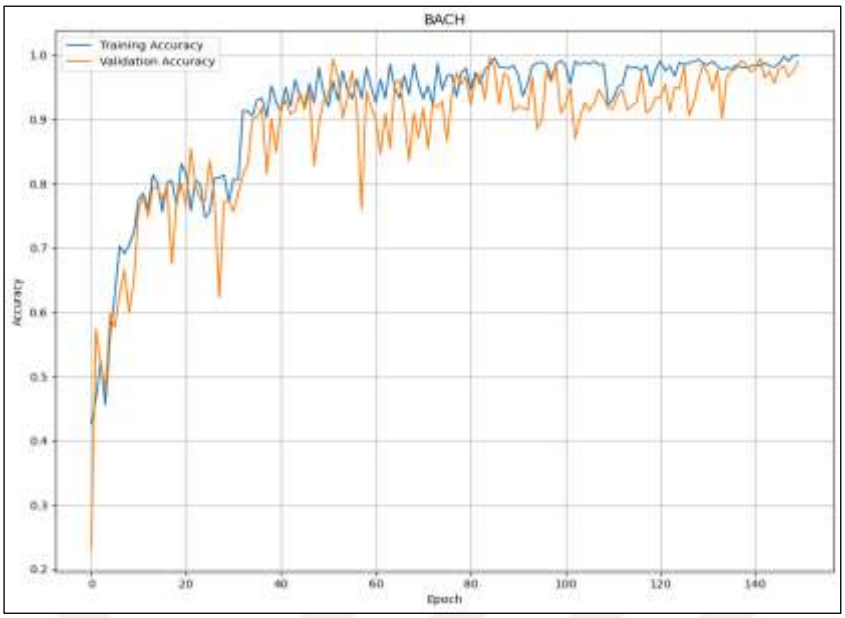

Figure 3.9: The Accuracy Curves for The HAFMAB-Net During The Training Process on The BACH- Dataset.

The training and validation losses for HAFMAB-Net over 150 epochs are characterized in Figure 3.10. By the end of the last epoch, the initial training loss, which is roughly 1.1, steadily drops to about 0.1. Likewise, right before training ends, the validation loss drops to about 0.2 from a starting point slightly above 1.2. This pattern identifies a distinct initial peak and the point at which the model begins to use fine-tuning. The training loss and validation loss exhibit little variation for the remainder of the training period, suggesting efficient learning with little sign of overfitting. Remarkable performance in dependable classification is reflected in the convergence of both losses, which shows a great ability to capture the majority of the features in the BACH dataset as well as an effective generalization to previously unseen datasets.

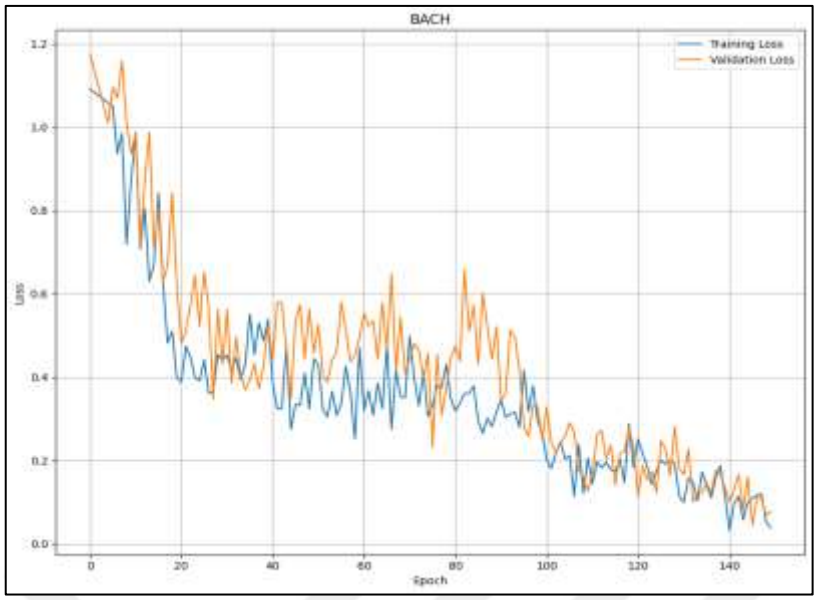

Figure 3.10: The Loss Curves for The HAFMAB-Net During The Training Process on The BACH- Dataset.

Table 3.3 shows the evaluation measures of the proposed model "HAFMAB-Net" and traditional models of deep-learning on the BACH-dataset. The proposed model achieved near perfection in all metrics (precision, recall, F1-score, and accuracy) with an outclassed value of 99.00%. In contrast, the deep-learning models such as ResNet 18 and MobileNetV3 performed significantly lower with accuracy values of 71.00% and 60.00%, respectively. While the ResNet 50 could not perform better than HAFMAB-Net while achieving a fairly competitive accuracy of 80.00%. The results evidenced that HAFMAB-Net could detect complex patterns within the BACH-dataset, showing robustness against shallower and deeper pre-trained architectures on all measures considered.

Table 3.3: Evaluation of Performance Metrics of The BACH-Dataset.

	HAFMAB-Net	**Deep learning - ResNet 18**	**Deep learning - ResNet 50**	**Deep learning - VGG16**	**Deep learning - MobileNetV3**	**Deep learning - nceptionV1**
Precision	99%	60.22%	81.92%	75.32%	61.32%	97.90%
Recall	99%	77.50%	80%	72.50%	67.50%	67.50%
F1-score	99%	67.72%	78.97%	71.14%	61.67%	61.92%
Accuracy	99%	91%	80%	92%	60%	62.50%

The confusion matrix for the test subset of the BACH dataset is shown in Figure 3.11, indicating HAFMAB-Net's classification performance over Benign, In Situ, Invasive, and Normal. The proposed model perfectly classified all 20 samples in each category—Normal, Invasive, and Benign. Only one sample was wrongly labelled as Invasive; 19 were correctly classified as In Situ. Therefore, total classification accuracy is close to the real maximum possible value, implying the strong ability of the model to distinguish between these four histopathological categories with no appreciable bias either for or against any one class. These results show the model's efficacy and reliability in breast cancer multi-classification tasks.

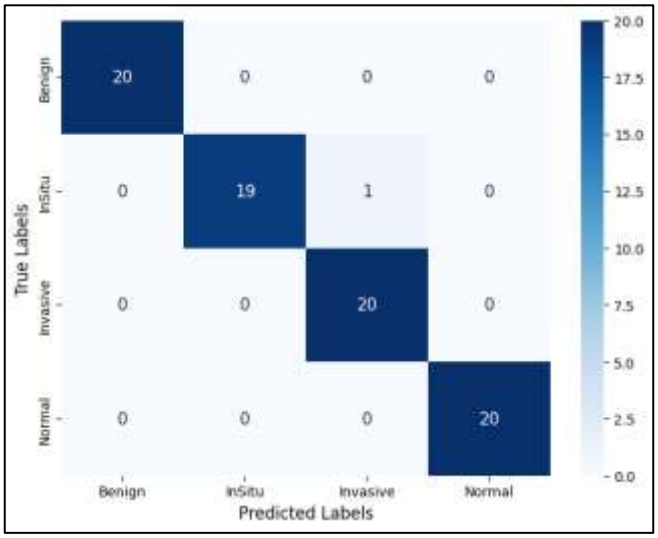

Figure 3.11: HAFMAB-Net Confusion Matrix of BACH-Dataset.

The AUC plot in Figure 3.12 shows HAFMAB-Net's performance against other deep-learning classification models on the BACH dataset. The True Positive and False Positive rates in this plot explain how much the model can discriminate between different classes. The proposed “HAFMAB-Net” reported AUC to be 1.00, outperforming all other tested architectures. MobileNetV3 recorded the lowest AUC, 0.68, whereas ResNet50 and ResNet18 obtained AUC values of 0.82 and 0.78, respectively. Class discrimination by the proposed model, as reflected in the near-ideal AUC value, speaks to its efficacy and reliability in classifying histopathology images from the BACH dataset with high accuracy at multiple decision thresholds.

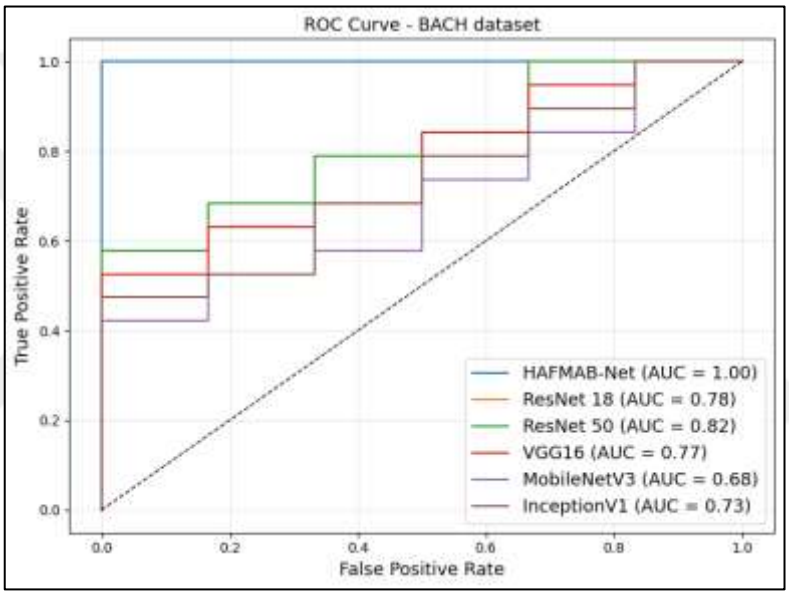

Figure 3.12: ROC Curve Clarifies Each Model's Area Under The Curve (AUC) Based on The BACH-Dataset.

3.5.5.2 Results of BreakHis-Dataset

The accuracy of HAFMAB-Net's training and validation throughout 150 epochs on the BreakHis dataset is shown in Figure 3.13. Starting at roughly 65.00% in the early epochs and continuously reaching values of at least 100.00% in the following epochs, the training accuracy shows a notable improvement from the beginning. Similarly, the validation accuracy shows a steady increase, surpassing 95.00% shortly before the end of training, albeit it does experience some oscillations in previous periods. This implies that while maintaining a high level of accuracy, the model exhibits a considerable capacity for

generalization. These outcomes demonstrate the model's adaptability and robustness in categorizing photos of breast tumors from the BreakHis dataset.

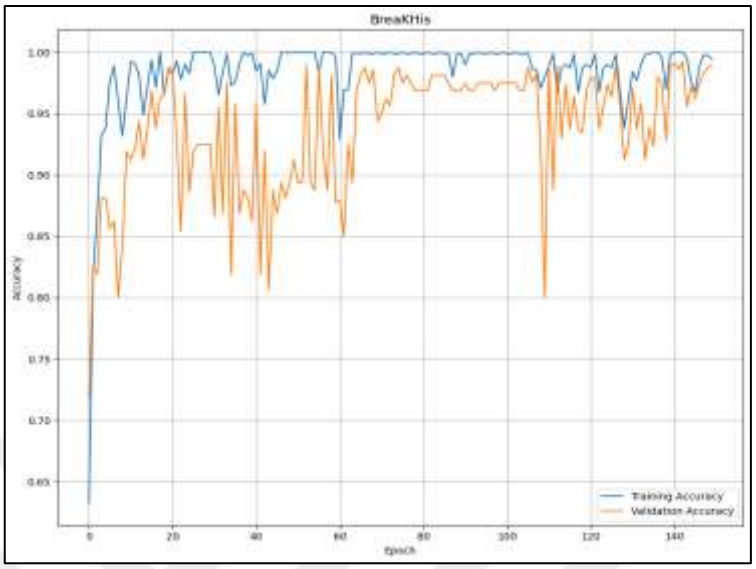

Figure 3.13: The Accuracy Curves for The HAFMAB-Net During The Training Process on The BreaKHis-Dataset.

Figure 3.14 shows the curves of training loss and validation loss of “HAFMAB-Net” on the BreakHis dataset for 150 epochs. In the beginning, the training loss value at 1.25 starts to decrease continuously till almost zero by the time of completion of training. The validation loss starts around 1.5, decreases progressively, and stabilizes at 0.1. A good learning process is shown by the correlation between training and validation curves, though lower values in the validation set are evidenced by slight overfitting. This finding proves that the model can optimize parameters very well while keeping a generalization capacity, which ensures robustness in classification tasks within the BreakHis dataset.

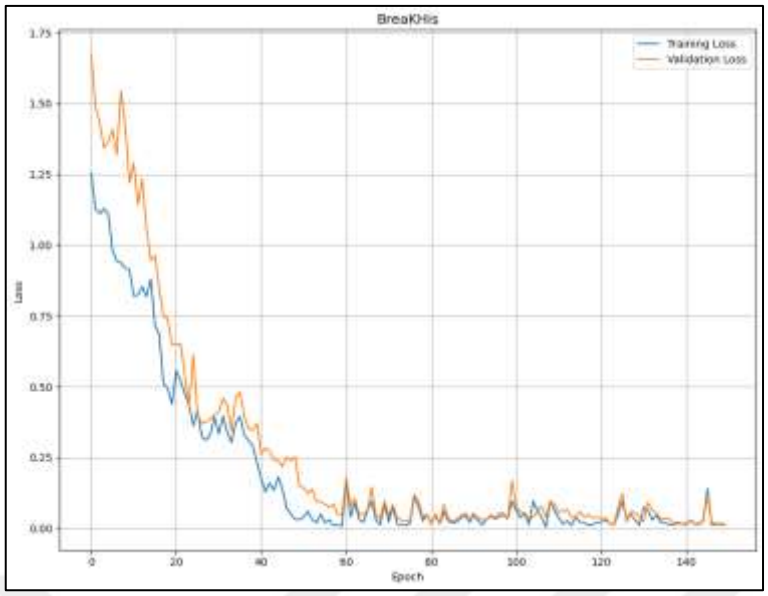

Figure 3.14: The Loss Curves for The HAFMAB-Net During The Training Process on The BreaKHis- Dataset.

Table 3.4 illustrates the performance evaluation of the proposed model “HAFMAB-Net” against conventional deep-learning pre-trained models using the BreakHis dataset. From the results obtained, HAFMAB-Net performed better with a precision of 98.00%, a recall of 98.50%, an F1 score of 98.50%, and an accuracy of 98.99%. In turn, VGG16 and ResNet18, as competing models, reported lesser accuracies that were pegged at 91.31% and 91.72%, respectively. These results have proven that the HAFMAB-Net can reduce false positives, proving that it is possible to detect malignant samples under such imbalanced dataset conditions. Such a high value for the F1 score indicates that this model can be trusted for binary classification, where it is presumed to give an optimal equal consideration for both precision and recall measured, plus its robustness concerning the distribution of classes, which was a major(malignant)and minor(benign).

Table 3.4: Evaluation of Performance Metrics of BreaKHis-Dataset.

	HAFMAB-Net	**Deep learning - ResNet 18**	**Deep learning - ResNet 50**	**Deep learning - VGG16**	**Deep learning - MobileNetV3**	**Deep learning - InceptionV1**
Precision	98%	90.12%	88.4%	91.36%	86.54%	89.33%
Recall	98.5%	93.35%	90.45%	9234%	92.66%	91.45%
F1-score	98.85%	91.28%	89.73%	91.64%	89.41%	90.41%
Accuracy	98.99%	91.72%	88.95%	91.31%	89.63%	90.35%

The classification performance of the HAFMAB-Net on the BreakHis-dataset is presented and analyzed using a confusion matrix, as shown in Figure 3.15. The model showed an outstanding accuracy of 99.39% in distinguishing benign and malignant breast tissues. Out of a total of 247 benign samples, 245 were correctly classified, confirming the model's reliability in the classification task with a very low error rate. Similarly, 532 out of 542 malignant samples were successfully diagnosed, indicating a 98.14% classification accuracy for malignant samples. The findings show the model is remarkably well despite the dataset's intrinsic class imbalance. The small misclassification cases prove that HAFMAB-Net is highly effective at correctly detecting the minority Benign class with a negligible bias towards the majority. Although the differences between the classes may provide difficulties, the FP and FN support the model's resilience and dependability in class distinction.

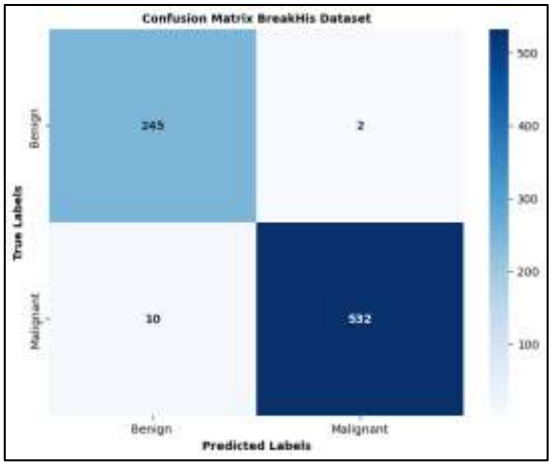

Figure 3.15: HAFMAB-Net Confusion Matrix of BreaKHis-Dataset.

Further proving the effectiveness of the HAFMAB-Net over conventional deep learning methods is the examination of the ROC curve obtained from the BreaKHis dataset, as shown in Figure 3.16. With an AUC of 1.00, HAFMAB-Net fared better than any other model. The competing models with the greatest AUC scores were ResNet 18 at 0.94, InceptionV1 at 0.95, and VGG16 at 0.96. The AUC scores for MobileNetV3 and ResNet 50 were both 0.91. The curve of ROC shows that the proposed model "HAFMAB-Net" keeps the FPR close to zero while maintaining a significant TPR across all decision thresholds. This confirms that the model can correctly classify most samples across all threshold levels, irrespective of class imbalance.

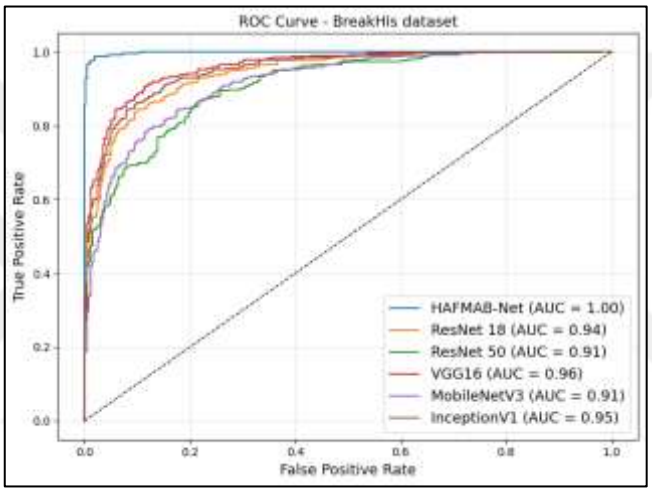

Figure 3.16: ROC Curve Clarifies Each Model's Area Under The Curve (AUC) Based on The BreaKHis-Dataset.

3.5.5.3 Results of LC25000-Dataset

LC25000 is a dataset comprising two separate parts: The Colon and Lung datasets, which are first analyzed individually and then together to assess how well HAFMAB-Net and a classification model based on deep learning perform.

The training and validation accuracies of HAFMAB-Net on Colon-data over 150 epochs are illustrated in Figure 3.17. In the initial phases, the training accuracy was about 30.00%. Later, it progressed towards nearly 100.00%. Validation accuracy started at about 20.00% and also showed a steady rise, reaching 95.00% before the completion of training. This finding shows that the model has an output in generalization with a high accuracy level, proving its potential in correctly classifying images of colon tissues.

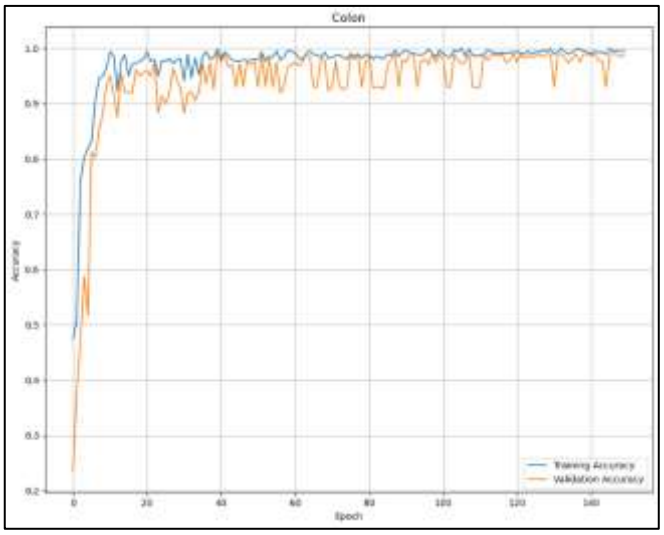

Figure 3.17: The Accuracy Curves for The HAFMAB-Net During The Training Process on The Colon- Dataset.

Figure 3.18 illustrates the HAFMAB-Net training and validation loss curves across 150 epochs using Colon data. It begins at about 0.60 and slowly falls to approach 0.05 in the final epochs. About 0.80 is the starting point of the validation loss. However, it declines more nonlinearly until it approaches 0.10 in the final training epochs. The curves of loss for HAFMAB-Net demonstrate effective learning with little overfitting. These outcomes indicate that the model has undergone significant optimization to guarantee resilience in subsequent tasks of categorizing colon tissue.

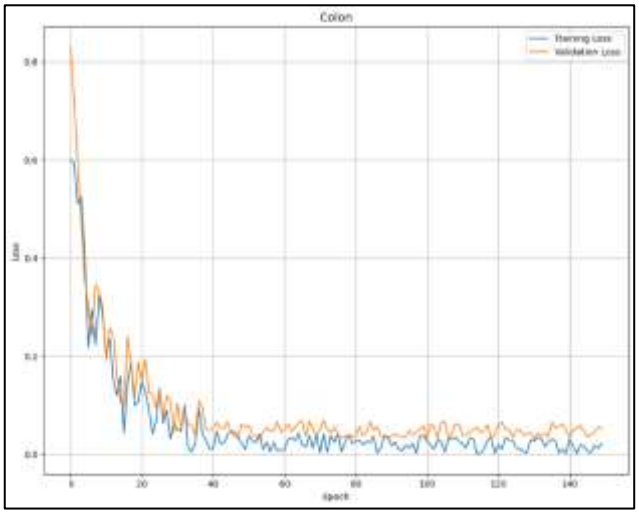

Figure 3.18: The Loss Curves for the HAFMAB-Net During the Training Process on the Colon-Dataset.

Table 3.5 shows the performance evaluation of the proposed model "HAFMAB-Net" on the Colon-dataset compared to other deep-learning models. The proposed model, "HAFMAB-Net," achieves maximum scores on all metrics, precision, recall, F1 score, and accuracy; hence, it is considered the best model. ResNet18, ResNet50, and VGG16 results are 99.00%, a strong performance but still slightly lesser than HAFMAB-Net's. These results prove that HAFMAB-Net has very high classification accuracy on the Colon dataset.

Table 3.5: Evaluation of performance metrics of the Colon-dataset.

	HAFMAB-Net	**Deep learning - ResNet 18**	**Deep learning - ResNet 50**	**Deep learning - VGG16**
Precision	100%	99%	99%	99%
Recall	100%	99%	99%	99%
F1-score	100%	99%	99%	99%
Accuracy	100%	99%	99%	99%

The classification performance of the two classes colon_aca and colon_n by the HAFMAB-Net on a test subset of the Colon-dataset is described by the confusion matrix, displayed in Figure 3.19. For colon_aca, the model provides 100.00% accuracy, with 500 correctly categorized. Likewise, 500 colons_n samples were accurately categorized with a 100.00% accuracy rate. Given these findings and their interpretations, we can say that the HAFMAB-Net has supported the dataset's balance and performed admirably in a discriminative capacity for these two classes, functioning effectively for binary classification.

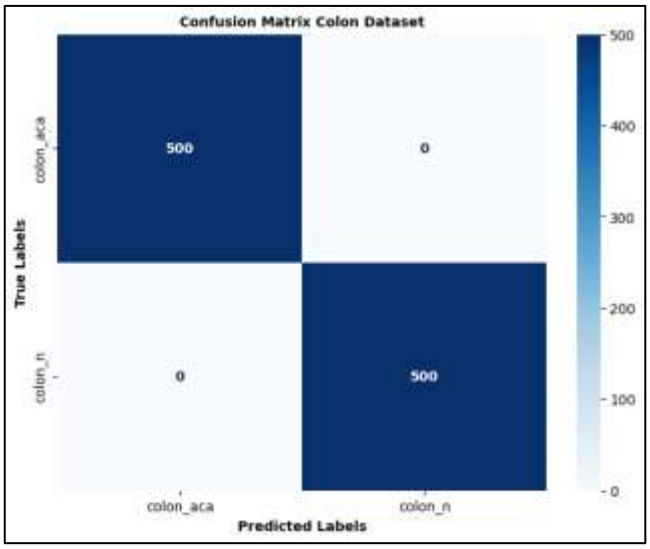

Figure 3.19: HAFMAB-Net Confusion Matrix of Colon-Dataset.

Compare the proposed model “HAFMAB-Net” with several other deep learning models in classifying the Colon-dataset using the ROC curves shown in Figure 3.20. Ahh-Net, ResNet18, ResNet50, and VGG16 attained the maximum AUC score equal to 1.00. This results in HAFMAB-Net correctly classifying colon tissues, which is consistent with the very clear level of class separation in the dataset.

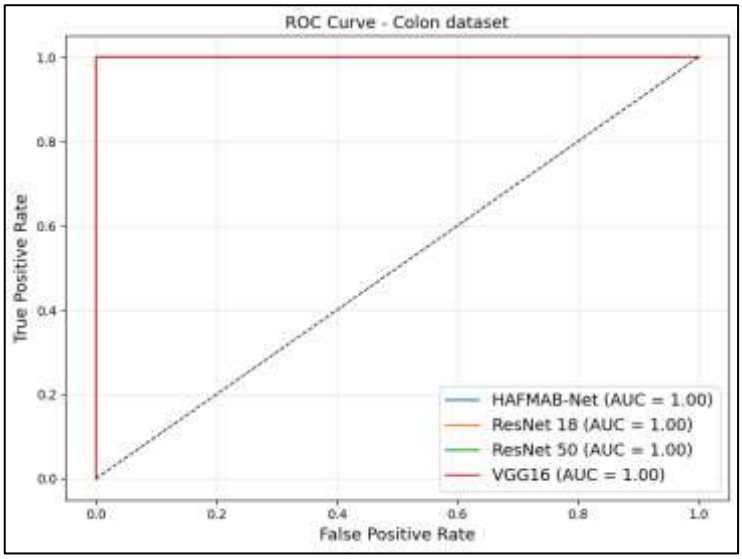

Figure 3.20: ROC Curve Clarifies Each Model's Area Under The Curve (AUC) Based on The Colon-Dataset.

Figure 3.21 shows the HAFMAB-Net for Lung-dataset training and validation accuracy over 150 epochs. Initially, the training accuracy is nearly 50.00%, and it reaches almost 100.00% within the initial epochs and then stabilizes. Validation accuracy follows almost the same trend, starting at around 40.00%, rising progressively, and finally stabilizing at close to 98.00% by the end of the training process. Excellent capacity for generalization and accuracy in classifying lung tissue images is evidenced by the similarity between training and validation accuracies, coupled with a slight degree of variation.

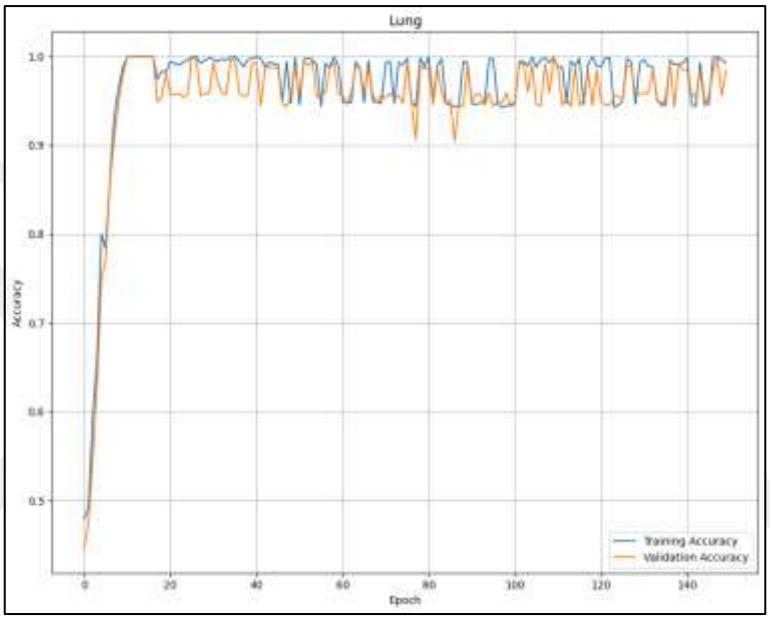

Figure 3.21: The Accuracy Curves for The HAFMAB-Net During The Training Process on The Lung- Dataset.

Figure 3.22 shows the curves of training and validation loss for the proposed model "HAFMAB-Net" according to Lung-dataset over 150 epochs. Starting approximately from 0.60, the training loss kept decreasing until it reached a stable value of 0.05 at the end of the epochs. Contrarily, validation loss started at about 0.70 and showed an incessant decline, reaching about 0.10 by the close of the training session. The close alignment between training and validation loss curves suggests that learning was effective, with minimal overfitting observed. Such a result means that the model has been well-optimized and demonstrates strong generalization, which is needed for the efficient classification task on the lung dataset.

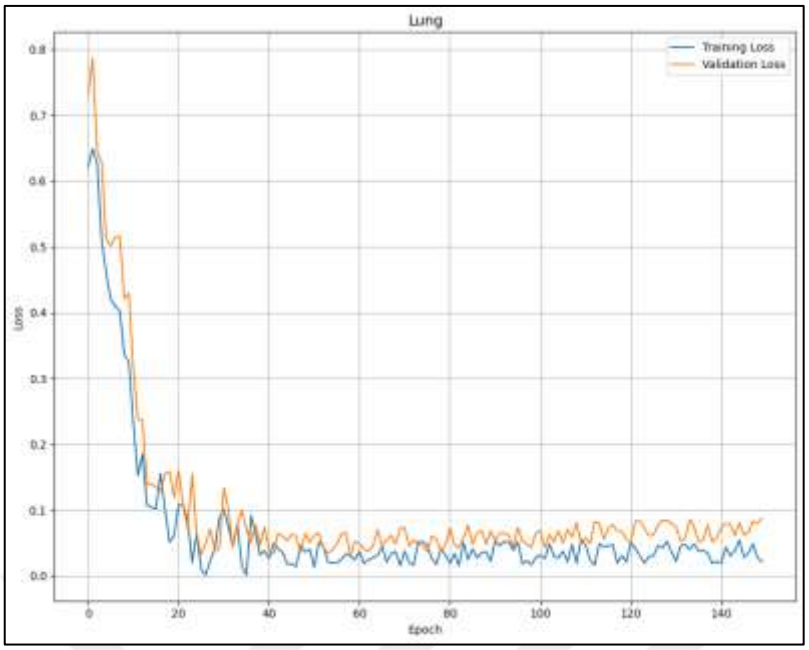

Figure 3.22: The Loss Curves for The HAFMAB-Net During The Training Process on The Lung-Dataset.

Table 3.6 shows the performance evaluation of the proposed model "HAFMAB-Net" on the Lung-dataset compared to other deep-learning models. HAFMAB-Net outperforms its rival models with flawless scores of 100.00% on all evaluated parameters, including accuracy, precision, recall, and F1-score. On the other hand, models with scores of 99.00% in these criteria, including ResNet 18, ResNet 50, and VGG16, demonstrate strong performance, albeit marginally worse. These findings support HAFMAB-Net's exceptional performance and highlight its exceptional capacity to accurately and flawlessly classify images of lung tissue.

Table 3.6: Evaluation Performance Metrics of Lung-Dataset.

	HAFMAB-Net	**Deep learning-ResNet 18**	**Deep learning-ResNet 50**	**Deep learning-VGG16**
Precision	100%	99%	99%	99%
Recall	100%	99%	99%	99%
F1-score	100%	99%	99%	99%
Accuracy	100%	99%	99%	99%

Figure 3.23 illustrates the confusion matrix of HAFMAB-Net-Net performance in classifying lung tissues that fell under three categories in the test subset of the Lung dataset: lung_aca, lung_n, and lung_scc, with 100.00% accuracy as recorded by the model in identifying all 500 samples from each category. This fact further underlines the great accuracy and classification power that HAFMAB-Net-Net brings to type lung tissues; it has been validated to support AI-based pathology analysis.

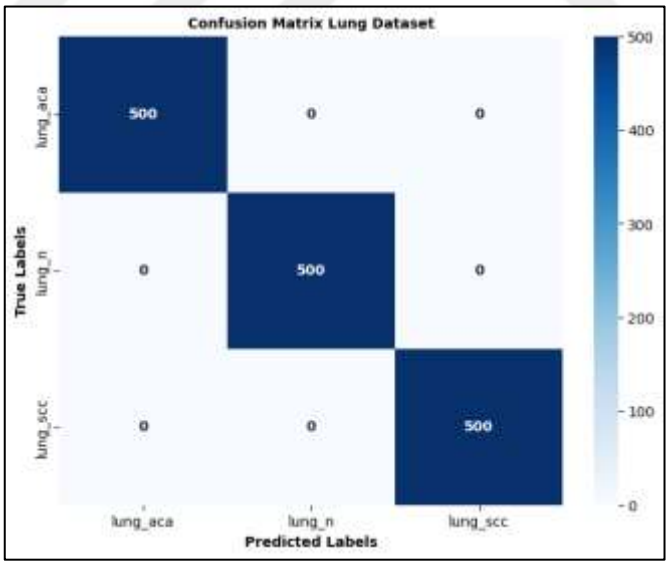

Figure 3.23: HAFMAB-Net Confusion Matrix of Lung-Dataset.

The ROC curves for the "HAFMAB-Net" model and some conventional deep-learning models are being compared on the Lung-dataset in Figure 3.24. Each model (VGG16, ResNet 18, ResNet 50, and HAFMAB-Net) attained an AUC score of 1.00, therefore, acing

it. The extraordinary performance of each and every model shows the highest-class separability level observed in the Lung-dataset. This, therefore, validates the effectiveness and dependability of HAFMAB-Net for lung tissue classification tasks.

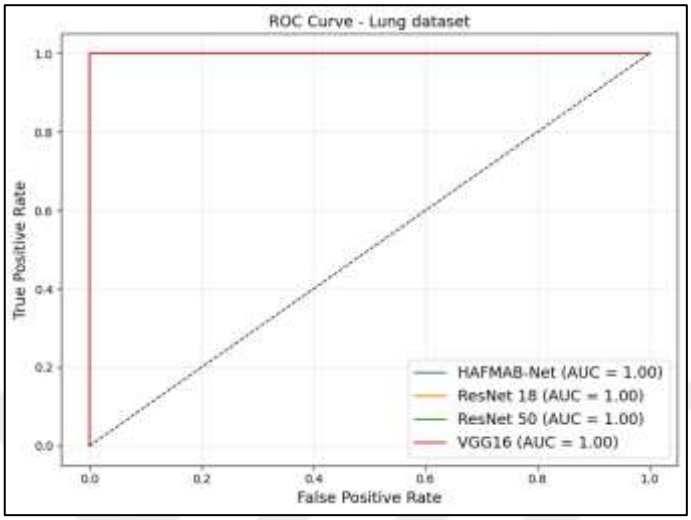

Figure 3.24: ROC Curve Clarifies Each Model's Area Under The Curve (AUC) Based on The Lung-Dataset.

Figure 3.25 illustrates the trend of the HAFMAB-Net training and validation accuracies over 150 epochs on the LC25000 dataset. The training accuracy began at 50.00% and gradually increased to nearly 100.00%, where it held stable until the finishing point. Validation accuracy was a little above 40% at the beginning, increasing progressively and reaching around 99% by the end of that epoch. A near one-to-one relationship between the two accuracies would confirm the model's good generalizability of the classification patterns within the LC25000-dataset.

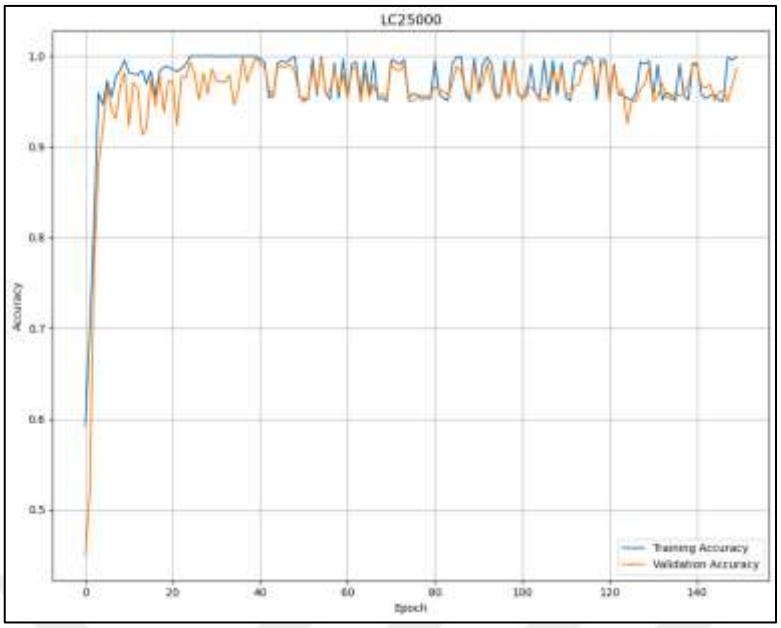

Figure 3.25: The Accuracy Curves for The HAFMAB-Net During The Training Process on The LC25000- Dataset.

Figure 3.26 displays the HAFMAB-Net's training and validation loss curves over 150 epochs for the LC25000-dataset. Beginning at about 0.80, the training loss drops precipitously in the early epochs before leveling at about 0.01 at the conclusion. Starting at almost 1.00, the validation loss progressively dropped until it settled at 0.10 in the last training epochs. With very little indication of overfitting, the two curves' similar forms demonstrate that learning was comparatively effective. These findings demonstrate the proposed model's obvious optimization performance potential and generalizability via the LC25000-dataset, enabling reliable performance on classification tasks.

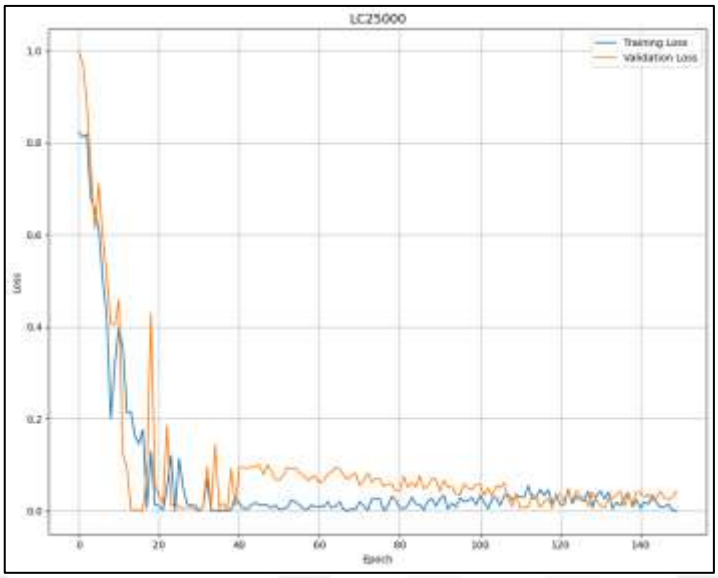

Figure 3.26: The Loss Curves for The HAFMAB-Net During The Training Process On The LC25000- Dataset.

As depicted in Table 3.7, evaluation metrics for the HAFMAB-Net and conventional deep learning models are based on the LC25000 dataset. A 100.00% perfect precision, recall, F1-score, and accuracy were found in HAFMAB-Net, which outperformed all its rival models. VGG16, ResNet 18, and ResNet 50 came at nearly 99.00%. This summary proves that HAFMAB-Net is capable of very complex classification tasks within the LC25000 dataset and does it without making a single error and with 100% assurance.

Table 3.7: Evaluation Performance Metrics of LC25000 Dataset.

	HAFMAB-Net	**Deep learning-ResNet 18**	**Deep learning-ResNet 50**	**Deep learning-VGG16**
Precision	100%	99%	99%	99%
Recall	100%	99%	99%	99%
F1-score	100%	99%	99%	99%
Accuracy	100%	99%	99%	99%

An evaluation of the HAFMAB-Net on a subset of the LC25000 dataset, which focuses on classification performance across five different categories (colon_aca, colon_n, lung_aca,

lung_n, and lung_scc), is shown in Figure 3.27. There were 500 samples in each category, and all were classified with perfect accuracy, giving the remaining five classes a classification accuracy of 100.00%. This outcome demonstrates the exceptional performance of the proposed model on the LC25000-dataset by highlighting its extraordinary precision, stability, and generalization abilities in differentiating between various tissue types and situations.

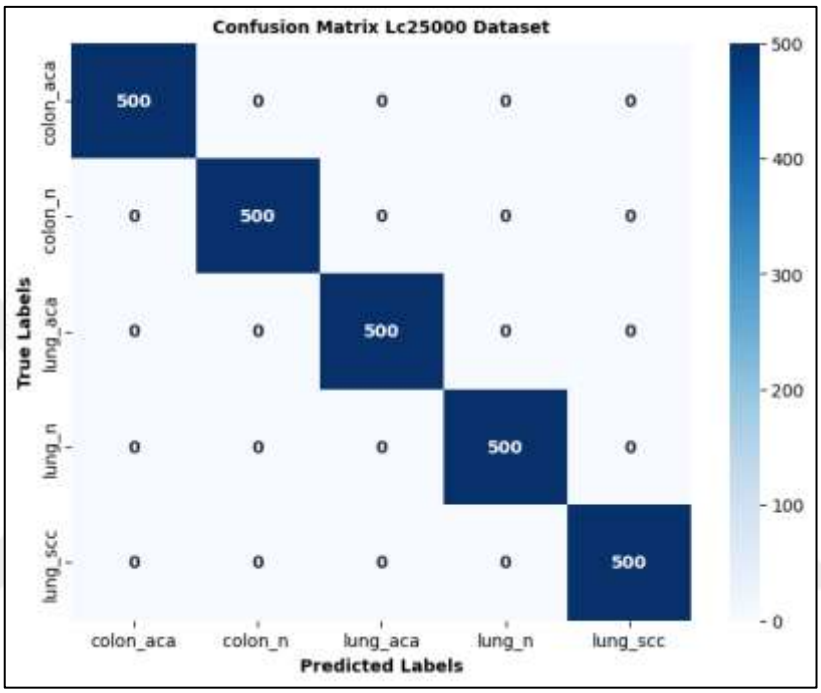

Figure 3.27: HAFMAB-Net Confusion Matrix of LC25000 Dataset.

Figure 3.28 illustrates the AUC curves of the HAFMAB-Net model proposed, together with those of some conventional deep learning frameworks, all using the LC25000-dataset. The area under the curve for each model HAFMAB-Net, ResNet 18, ResNet 50, and VGG16 is 1.00. This great performance proves the model's value and shows its ability for multi-class classification tasks by indicating a high degree of separability inside the LC25000-dataset.

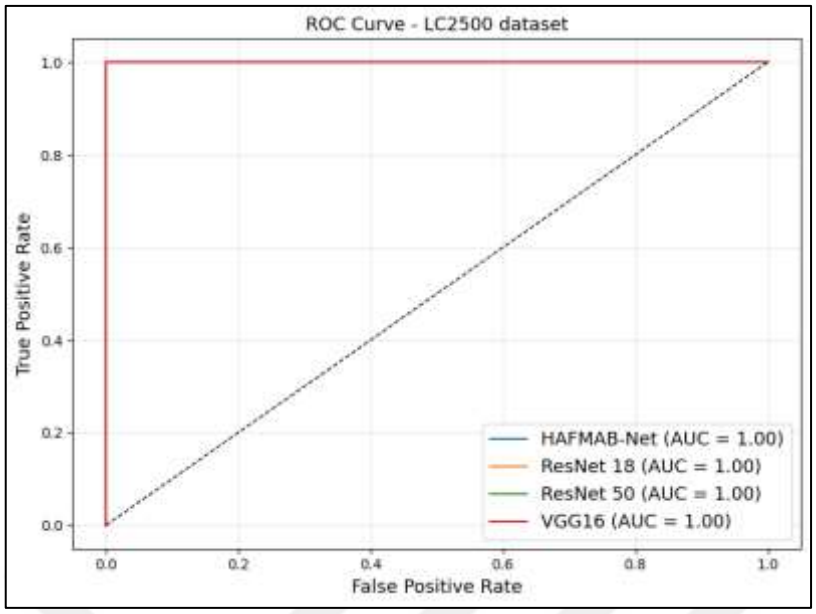

Figure 3.28: ROC Curve Clarifies Each Model's Area Under the Curve (AUC) Based on the LC25000-Dataset.

3.5.6 Comparison with Recent Works

The comparative assessment of the proposed model's "HAFMAB-Net" performance against other top approaches on the BACH and BreaKHis datasets is listed in Table 3.8. The table shows that though some of these latest methodologies achieved near-perfect accuracy, HAFMAB-Net attained the best accuracy, about 99%, higher than all other models in the Proposed Method in the classification BACH dataset. Regarding the BreaKHis dataset, HAFMAB-Net is presented to have obtained the best classification accuracy of 98.99%, again superior to existing approaches where the top models registered accuracies of 98.86% and 98.26% [124] [135]. Clinically, such differences, when analyzed in terms of practicality, translate into significant benefits since they aid in enhancing diagnostic accuracy as well as reducing misclassifications related to breast cancer. The great success of HAFMAB-Net is due to its smart design, which uses deep learning methods like attention tools and adaptive fusion to improve feature extraction. These features help the model beat problems of change and imbalance in the small BreaKHis and BACH datasets, making classification more reliable and stable.

Table 3.8: Comparison of The Proposed Model "HAFMAB-Net" Against State-of-The-Art Models.

References	Dataset	Accuracy	Dataset	Accuracy
[123]	BACH	90.27%	BreaKHis	91.45%
[124]		86.25%		98.86%
[125]		97.50%		97.70%
[127]		79.06%		93.02%
[128]		-		96.30%
[129]		97.75%		-
[131]		93.62%		87.6%
[134]		97.80 %		95.29%
[135]		94%		98.26%
[136]		91.25%		91.14%
[137]		88.75%		98.13 %
[138]		94.80%		-
[139]		92%		97%
[140]		90%		97%
[141]		93.57%		96.095%
[142]		-		98.13 %
HAFMAB-Net		99%		98.99%

3.5.7 Discussion

The classification capabilities of medical images under the CAD paradigm have been boosted by integrative deep learning, offering supporting diagnosis with sharper precision and efficacy. HAFMAB-Net is an impactfully performing model that is robust in answering

the challenges of classification of histopathological images. It strategically architecture deep convolutional networks, attention mechanisms, and adaptive fusion ways to bring multiscale feature extraction and synthesis into different effectiveness. Where traditional CNN frameworks would small-involve only the fine details of tissue images and therefore underperform, HAFMAB-Net performs exceptionally well in recognizing coarse and fine features. The reason for this lies jointly in the AMFM attention modules through which the framework can recognize pathology-important features while keeping all spatial, and hierarchical, representation thorough.

The results in the above sections titled “BACH dataset Results,” “BreakHis dataset Results,” and “LC25000 dataset Results” prove the great classification ability HAFMAB-Net has to show through many datasets. The proposed “HAFMAB-Net” attained an excellent accuracy of 99% on the BACH dataset with AUC scores of 1, which means it can separate classes with near-perfect intensity. Also, HAFMAB-Net performed well on independent and LC25000 datasets, achieving 100% perfect accuracy and AUC scores 1, indicating impossible classification performance. On the other hand, the BreaKHis-dataset achieved a very high accuracy of 98.99% with nearly perfect confusion matrix AUC scores of 1, which means that HAFMAB-Net can deal with minor classes best without any major bias towards the bigger class.

The above results make the method reliable, flexible, and steady in different classification tasks. So, it differentiates itself from typical deep learning models and other top methods, as shown in Tables 3.9 The greatness of HAFMAB-Net comes from its novel feature extraction and aggregation method. The model smartly captures complex patterns and tiny discrepancies within histopathological images through adaptive fusion at both the global and patch levels, which boosts classification accuracy and mitigates dataset variation, imbalance, and noise problems, making it very applicable in clinical and research settings.

The proposed model “HAFMAB-Net” finishes the hunt for a method that helps fast, right, and easy finding and sorting breast cancer. Its great accuracy and the ability to deal with both two-way and many-way sorts make it very good for real test systems, and it is best in clinical places where quick and exact checks are needed. Also, its skill in handling uneven data sets gives a big plus for use in places with few resources, like health care centers in country areas.

Though HAFMAB-Net has been greatly improved, some minor issues remain. The feature extraction at the patch level could be one reason for the minor reduction in sensitivity to wide context information. Moreover, while the computational costs of bringing in several attention and fusion modules can be handled, this may impede real-world usage in a resource-limited setting. These are small concerns and will not bring down the total efficacy of the model.

This creates an environment for future efforts to develop and extend HAFMAB-Net's functions further. One possible way would be to add new unsupervised learning methods that will make it easier to spot small internal variations and improve how things are sorted out in different data sets. Future versions of HAFMAB-Net will also be more efficient in places with limited resources, like rural healthcare facilities. Also, new ways of domain adaptation could be pursued, and strong data augmentation strategies would boost generalization and effectiveness across different medical imaging datasets.

3.6 CONCLUSION

This chapter presents the HAFMAB-Neta novel model introduced here that stands for Hierarchical Adaptive Fusion and is based on a Multilevel Attention-Enhanced Bottleneck Neural Network for cancer classification from histopathology images. It integrates an improved version of the Bottleneck Channel module with attention at both global and local levels for feature extraction from general image content and spatial patches, respectively. The representation of features at the patch level is further enriched through a Deeper Spatial Attention Aggregator Module that emphasizes fine-grained discriminatory information, thus allowing the network to better deal with more intricate elements. An Adaptive Fusion Module will be added to this design to synthesize these discriminative features and effectively exploit different levels of hierarchal multi-scale information. HAFMAB-Net has been validated extensively on BACH, BreaKHis, and LC25000-datasets to prove this capability. In the second dataset, known test results were attained by gaining 99% accuracy and 1 in AUC, which signifies almost perfect class separation. On the BreaKHis dataset, with significant data imbalance, HAFMAB-Net achieved an accuracy of 98.99% at all zoom levels and an AUC score of 1.00; thus, it can be explained how it performs well in generalization despite challenges in data distribution. In the LC25000-dataset, this model attained 100% accuracy and an AUC of 1.00, showing the extraordinary dependability and

sturdiness of HAFMAB-Net in exceeding conventional deep learning methods and even other advanced techniques.

4. MATERB NET: MULTISCALE ATTENTION TRANSFORMATION BASED ON ENHANCED RESIDUAL BLOCK FOR BREAST HISTOPATHOLOGY IMAGES PREDICTION BASED ON PATCH-LEVEL

4.1 INTRODUCTION

Histopathological analysis is the most commonly used method for breast cancer diagnosis among various molecular markers. Pathologists visually examine samples under a microscope to assess and classify breast cancer [145].

Histopathological analysis performed by a pathologist is a time-consuming task, and its accuracy depends on the pathologist's experience. Therefore, the development of a Computer-Based Detection (CAD) system that can differentiate cancerous (malignant) tissue from non-cancerous (benign) tissue is critical [146]. This system works to help pathologists simplify the diagnosis process, besides saving time and providing high efficiency in diagnosis.

In recent years, technologies such as machine learning, deep learning, and image processing have been used in the field of pathology to overcome these problems; advances in machine learning and digital image processing have allowed researchers to develop more sophisticated CAD systems, which can make pathologists more productive, objective, and consistent in diagnosis [147]. The creation of automated image processing systems for breast cancer diagnosis has been considered a challenging task in recent years due to the natural complexity of histopathological images [148]. An effective CAD system is necessary for early detection of breast cancer and prevention of complications [149].

The majority of automated CAD systems rely on conventional machine-learning techniques to extract features, morphology, and structural characteristics from images [150]. These features are then used to classify images and build automated systems. Some researchers have assessed cell nuclei by extracting nuclear features [151], thereby obtaining the information necessary for the Classification of benign and malignant cells [152]. Based on traditional machine learning algorithms, various core separation algorithms, such as Support Vector Machines (SVM) and Random Forests, are used to determine features [153]. This

algorithm determines discriminative features by extracting information to distinguish normal, benign, and malignant slides.

Deep learning models identify features such as size, color, and shape of cells during the training stage using many examples [154]. They test these models on never-before-seen images to assess their accuracy. Due to the large number of pixel and matrix operations performed, these models require computers with powerful hardware, and training can take days [153] [154]. Despite requiring considerable resources, these models sometimes produce incorrect predictions due to the difficulty of separating tumor cells from the background and other tissue, as well as the loss of features due to image processing. In order to overcome this shortcoming and obtain better results, pre-trained models have been introduced [155]. In order to improve breast cancer diagnosis, Various studies have addressed fundamental problems in automatic breast cancer classification using Deep Learning models (DL)-based Transfer learning (TL) techniques with integration of Attention Mechanism (AM) on histopathological images.

4.2 RELATED WORK

The classification of breast histopathological medical images plays a crucial role in the field of computer-aided diagnosis (CAD) because it provides essential diagnostic assistance to healthcare professionals. This section will be highlighting strengths and limitations of the recent contributions in breast histopathological medical image classification and outline the establishes the foundation for the current study to address these gaps.

Gul et al [156], presented a hybrid framework combining Local Binary Patterns (LBP) with a CNN for classification breast histopathological images. The proposed approach captures local texture features using LBP, with a designed CNN model for classifying these features. Although, the effectively combined framework, the study does not integration the multi-scale spatial and temporal features to provide more classification context

Jackson et al [157], introduced a 5-B network based on patch level to classification breast cancer. The network uses incorporates multiple convolutional layers and block-wise attention mechanisms to capture a critical feature for patches-based image. Although, the network successfully demonstrated capture local features, it could be limited to capture larger-scale leading to loss the contextual relationships across the entire histopathology slide.

Balasubramanian et al [158], proposed ensemble model based on deep learning models to classification breast cancer using whole slide histopathology images. The proposed model reduces the individual architectures' weaknesses, which enhance robustness and generalization of the classification performance. Despite the advantages of ensemble CNNs, the model focused only on extraction of the spatial feature, instead of incorporating attention mechanisms focusing on the different parts of the input data, to compute the relationships between different across different regions of histopathology images.

Mewada et al [159], discussed an integrates residual features into a deep neural network architecture for classification breast histopathological images. The proposed approach focused only on enhancing the network's depth, without enhancing the model's ability on critical regions of interest within the histopathological images

Maleki et al [160], developed a hybrid approach integration XGBoost with DCNNs to classifcation breast cancer. The DCNNs works on extraction high-level features, and the XGBoost serves as a classifier. However, the focusing on feature-level fusion without incorporating multi-scale feature fusion reducing the model comprehensiveness.

Mondol et al [161], proposed deep learning architecture called "hist2RNA" inspired by bulk RNA sequencing techniques to predict gene expression profiles directly from breast cancer histopathological images. the model extracts complex pattern translating to molecular information, eliminating the traditional gene expression assays. However, still some challenges in capturing the complex pattern from different scales of images, which is critical for capturing complex tumor architectures.

4.3 CONTRIBUTIONS

The MATERB Net model introduces novel approach for classification breast cancer histopathological image and prediction the unknown patch-level of full image, and addressing the limitation in the Section 4.2. The proposed MATERB Net model uses the SE-ResB and CBAM-ResB for extraction features from full image and sub-images based on multi-magnification factors, respectively, and self-attention and cross-attention to learn the relationship between the pixels of the images and focus on the meaningful details and information from both full image and sub-images in calculating the attention weights. Self-attention works on capturing and learning the relation between the extracted features of the

extraction features of sub-images, providing various levels of details and information, which feed into the cross-attention beside the extracted features from full image. Cross-attention works to provide meaningful and context-aware representation integrated with the global context and details of each label. Also, the Balanced Focal Loss function used in the proposed method enhancing model robustness by focusing more on hard-to-classify samples, leading to improve the prediction reliability for the unknow labels based on patch-level

4.4 PRE-PROCESSING DATASET

A pre-processing is applied on the dataset before initiating the training process, in order to optimize the dataset for the valid model training and evaluation. Since the BreaKHis dataset has two class (Benign, Malignant) with sub-classes as compared to the BACH with Four classes, and due to their sub-classes considered as a type of the Invasive, and InSitu classes, the BreaKHis-based malignant has been referred as Invasive, and InSitu based on types [155]. Moreover, extracted patches process is applied on Normal class-based BACH via different window size to generate equivalent magnify images for the BreaKHis. A new referred dataset has been created via the referred and extracted patches processes with four classes equal to the BACH classes, composes of four magnification factors. The datasets are splitted into 80% training, and 20% validation. After the splitting dataset, augmentation dataset is applied including mirror-flipping, rotation, and adding noise, on the training samples. Note the augmentation is applied only on BACH set because of the small size. Due to the imbalanced new referred dataset as compared with the balanced BACH, the DataLoader of the new refiled dataset has been set to select randomly samples equal to the BACH's DataLoader after augmentation process because the model deals with multi-dataset in same time during training process.

4.5 METHODOLOGY

4.5.1 Overview

The framework of the proposed MATERB Net model is shown in Figure 4.1(a). The proposed model composes of deep residual block integrated with Squeeze and Excitation (SE) using full image as input, deep residual block integrated Convolutional Block Attention

Module (CBAM) fed by four different magnification of image, and attention transformation multi-head model included self-attention and cross-attention. Several strategies have been incorporated into MATERB-Net to mitigate the risk of overfitting. A dropout layer is applied after each convolutional layer at a rate of 0.25, meaning neurons are randomly deactivated during training. This approach helps to avoid overfitting by reducing the mutual adaptation of neurons.

Moreover, an L1 regularization technique is employed across layers to encourage a sparse distribution of learned weights, which improves the model's generalization capacity. During the pre-processing stage, input images are transformed to standardized scale values; this confirms coherence and minimizes the impact of data variability. All these strategies enhance the proposed model's effectiveness across various datasets.

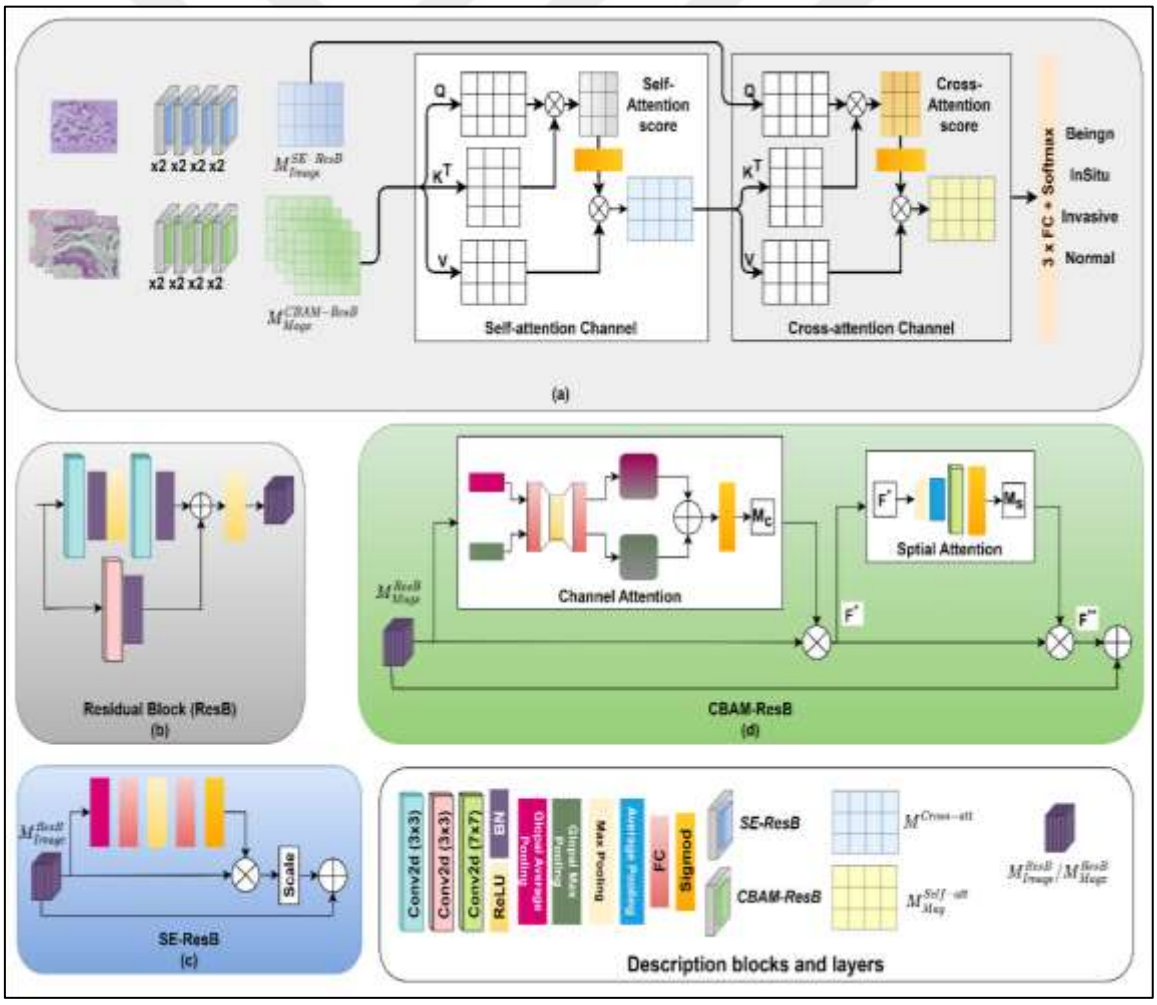

Figure 4.1: (a) MATERB Net Framework, (b) ResB, (c) SE-ResB, (d) CBAM-ResB.

4.5.2 Residual Block (SE-ResB)

The Residual Block (ResB) is a part of ResNet and it was designed to fight the vanishing gradient problem while trying to make the deep networks easily trainable. Figure 4.1(b) presents the ResB block diagram, it has two paths, the main path includes two 3×3 Conv2D layers with Batch Normalization and ReLU after each layer, and shortcut has single 1×1 Conv2D layer with Batch Normalization for matching the dimensions. These two paths are summed up element-wise and finally ReLU is applied to the added result. This creates a skip connection and allows gradients to flow undiluted; therefore, it offers very effective training to deeper networks. The mathematical expression for ResB is shown below;

$$M_{Image}^{ResB} \,/\, M_{mag\,x}^{ResB} = ReLU((CNN(x) \rightarrow BN \rightarrow ReLU \rightarrow CNN \rightarrow BN) \oplus (CNN(x) \rightarrow BN)) \quad (4.1)$$

4.5.3 Squeeze and Excitation - Residual Block (SE-ResB)

In Residual layers, filters capture local spatial relationships, forming feature maps. These are typically treated equally without assigning any importance to them. Such kind of an attempt can indirectly propagate irrelevant features through the network eventually leading to performance degradation. To overcome this, the Squeeze-and-Excitation (SE) block was developed in [162] for better representing features in Residual networks. Figure 4.1(c) presents the block diagram of SE; SE blocks re-calibrate feature maps $\left(M_{Image}^{ResB} \in R^{BxCxHxW}, [B, C, H, W] \right)$ from the ResB using the global information to exaggerate the most informative features. It follows the subsequent two steps: first is squeeze, which global pooling for created channel descriptor, and second is excitation, which learning channel-wise weights. The generic squeeze operation occurs by performing global average pooling (GAP) over the spatial dimensions, which brings about a descriptor that summarizes the channel-wise feature responses regarding how they are positioned within the image. This applies every layer to exploit global receptive field information. Excitation of this serialization is used to create modulation weights for each feature map to adjust their importance. The M_{Image}^{ResB} passes through GAP to become a C × 1 × 1 vector. This vector is further reduced by a factor of r through a ($fully\ connected\ layer\ (FC) \rightarrow ReLU\ activation$) to make the squeeze step. This output is then upsampled back to C × 1 × 1 through another FC, followed by a sigmoid activation to generate channel-wise weights.

The excitation step then rescales the input M_{Image}^{ResB} to generate the weighted feature map the SE-ResNet module further improves the M_{Image}^{ResB} by integrating SE blocks via summation-elements -wise process into each residual mapping, resulting in $\left(M_{Image}^{SE-ResB} \in R^{BxCxHxW}, [B, C, H, W] \right)$. The mathematical expression for SE-ResB is shown below;

$$M_{Image}^{SE-ResB} = Rescale\left((GAP\left(M_{Image}^{ResB}\right) \rightarrow FC \rightarrow ReLU \rightarrow FC \rightarrow Sigmod) \otimes M_{Image}^{ResB}\right) \oplus M_{Image}^{ResB} \quad (4.2)$$

4.5.4 Convolutional Block Attention Module - Residual Block (CBAM-ResB)

The CBAM involves two sequential submodules to enhance the features map via incorporating both channel-wise and spatial attention mechanisms, the channel-wise attention highlighting the most critical features via assigns different weights to channels based on their relevance, and Spatial attention emphasizes important regions by focusing on high-saliency locations [163]. Figure 4.1 (d), illustrates the CBAM-ResB. The $\left(M_{mag\,x}^{ResB} \in R^{BxCxHxW}, [B, C, H, W] \right)$, first pass to the channel attention to remove the spatial dimension via GAP and global max pooling (GMP), which pass to the $(FC \rightarrow ReLU\ activation \rightarrow FC)$, followed by Sigmoid generating the channel attention weight, which multiply with the M_{Image}^{ResB} to generate the M_c refining each channel based on its learned importance resulting in F^*. The spatial attention works on enhanced the specific regions within the F^* via Max pooing (MP) and Average pooling (AP) removing the spatial dimension, following by CNN layer with kernel size 7x7 works on refine the features map F^*. The output of the CNN passes through Sigmoid to calculate the spatial attention score M_s, which multiply by the F^* via summation-elements-wise to generate the F^{**} then integrating with the $\left(M_{mag\,x}^{ResB} \right)$ to enhance the M_{Image}^{ResB}, resulting in $\left(M_{mag\,x}^{CBAM-ResB} \in R^{BxCxHxW}, [B, C, H, W] \right)$. The mathematical expression for CBAM-ResB is shown below;

$$F^* = M_c \otimes M_{Image}^{ResB} \quad (4.3)$$

$$F^{**} = M_s \otimes F^* \quad (4.4)$$

$$M_{magnification\,x}^{CBAM-ResB} = \left(M_{magnification\,x}^{ResB} \oplus F^{**} \right) \quad (4.5)$$

4.5.5 Attention Transformation Multi-Head Model

The attention transformation models (TM) play a crucial role in modern deep learning models, especially in Transformer structure-based image processing [164]. The TM can learn the contribution between the elements of data based on the attention mechanism. The TM has two primary attention mechanisms, self-attention, and cross-attention, and these two attentions represent the connection between the encoder and decoder part of the TM, as shown in Figure 4.2 (a) [164]. In general, after embedding the input sequence, the Query "Q," Key "K," and Value "V" vectors are created using three-liner transformation. These three vectors pass the multi-head attention, as shown in Figure 4.2 (b). The multi-head attention works on determining the score attention using the Scaled Dot-Product Attention by calculating the similarity between Q vector and K vectors by multiplying (Q * KT) resulting attention are scores normalized using Softmax, and then multiplying by V vector resulting the features maps as shown in Figure 4.2 (c). Multi-head attention has gotten its name because Scaled Dot-Product Attention is repeated in a parallel process. This leads to an update of the Q, K, and V vectors in each head, resulting in different attention scores combined and normalized to be stable and practical. Since the Q, K, and V vectors are generated from the same input sequence, it is known as self-attention, represented by the encoder part if the Key (K) and Value (V) are generated from different input sequences of the Query (Q), it known as cross-attention represented by the encoder-decoder parts [164].

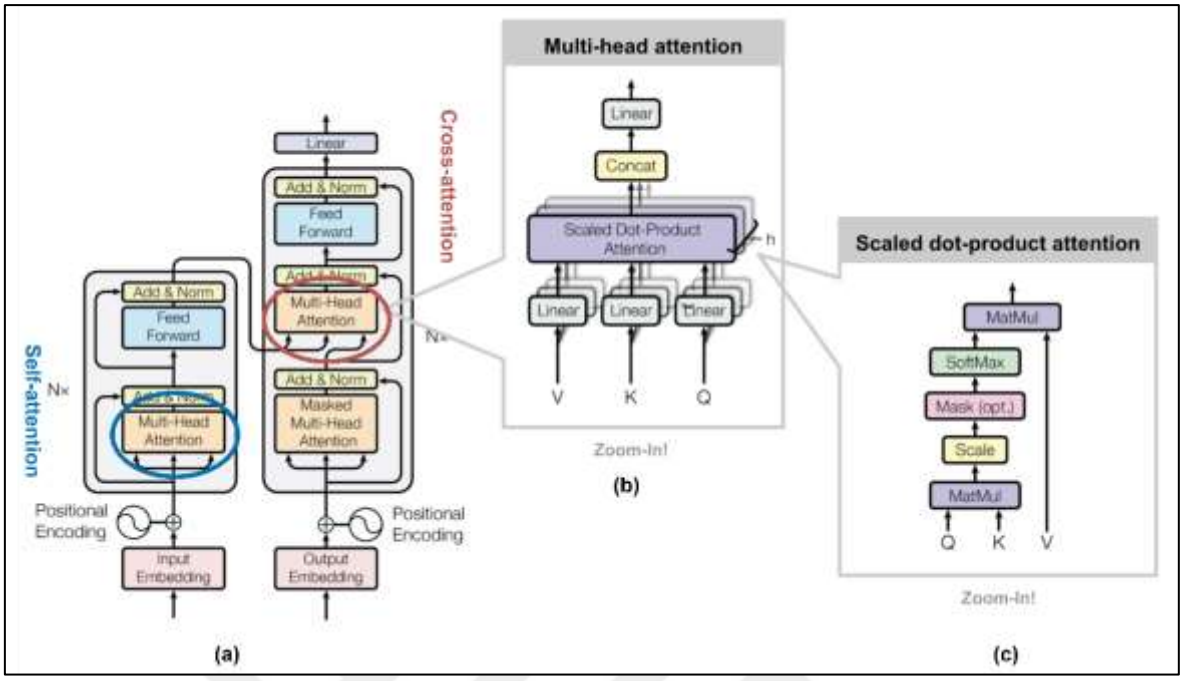

Figure 4.2: Attention Transformation Model [164].

4.5.6 Evaluation Matrix

The terms (TP), (FP), (TN), and (FN) are used to calculate the evaluation metrics, such as accuracy, precision, recall, and F1 score, that are used to evaluate the proposed model's performance. These metrics offer a clear report of the performance across various datasets [46].

i. $Accuracy = \frac{TP+TN}{TF+FN+TP+FP}$ (4.6)

ii. $Precision = \frac{TP}{TP+FP}$ (4.7)

iii. $Recall = \frac{TP}{TP+FN}$ (4.8)

iv. $\mathrm{F1-Score} = 2\,.\,\frac{(Precision \times Recall)}{(Precision + Recall)}$ (4.9)

4.6 EXPERIMENTAL RESULTS

4.6.1 Experimental Settings

The proposed MATERB Net model has been implemented in Python using PyTorch framework and performed on a Windows 64-bit system (Intel Core i7 9750H CPU 3.2 GHz, NVidia GTX 1660 Ti GPU, and 1TB - 256GB). The proposed MATERB Net model uses The Balanced Focal Loss function is used in the proposed MATERB Net model to improve the performance and generalization of the training process. The Balanced Focal Loss function integrates the Focal loss and weight cross-entropy to focus on small, hard-to-classify regions in the full images, and prevent the model from biased by dominant tissue structures while missing subtle malignant patterns. The math expression of the Loss function is given as follows:

$$FacalLoss = \kappa(1 - p_t)^{\lambda} * CE(x, y) \tag{4.6}$$

$$p_t = exp\left(- CE\ (x, y)\right) \tag{4.7}$$

$$CE\ (x, y) = -\sum_{i=1}^{c} y_i\ log\ p_i \tag{4.8}$$

$$BaclancedFacalLoss = \frac{1}{N}\sum_{i=1}^{N} FL\ (x_i\ \ y_i) \tag{4.9}$$

Where x is the input logits, y is the ground truth labels, κ is the balancing factor, λ is the focusing parameter, p_i is the model estimating probability, $CE\ (x, y)$ is the cross-entropy loss, and C is the number of classes. Table 4.1 provides a summary of the hyperparameters employed for training the proposed MATERB Net. To attain consistent and dependable performance, these parameters comprising learning rate, batch size, step size, epoch, and various settings were optimized for each dataset.

Table 4.1: MATERB Net Training Hyperparameters.

Type	
Optimization	Adam
learning rate	10^{-4}
batch size	10
Alpha	0.65
γ	3
step size	5
epoch	120
Time	16h

4.6.2 Feature Extraction

Five datasets are used to train the proposed method, BACH, and the four magnification factors of the magnification referred dataset. The SE-ResB progressively extracts BACH features providing an enhanced global feature map $\left(M_{Image}^{SE-ResB} \in R^{BxCxHxW}, [B, C, H, W] \right)$, SE blocks adaptively recalibrate the feature responses channel-wise even in the interdependencies channel feature associations; thus, it gives the network the possibility of actively keying onto informative features. It shares a hyperparameter with the former called 'r' (reduction ratio); during the squeeze operation, this hyperparameter controls the dimensionality reduction factor. In contrast, the four magnification factors, i.e., (40x, 100x, 200x, 400x) of the magnification referred dataset are fed to the CBAM-ResB for extracts features, results in four enhanced locative feature maps $\left(M_{mag\ x}^{CBAM-ResB} \in R^{BxCxHxW}, [B, C, H, W] \right)$. The first part of the CBAM attention mechanism is applied channel-wise to select the channels carrying more relevance considered independently of spaces, while second part of the CBAM attention mechanism is applied along the two spatial dimensions; thus, choosing the relevant places in the feature maps considered independently of channels. The (40x, 100x, 200x, 400x) - based $M_{mag\ x}^{CBAM-ResB}$ are processed sequentially using the self-attention part of the attention transformation model. The self-attention process

allows the attention transformation model to learn the relationship between the extracted features and calculate the attention weight between the features. This results in creating a new self-attention map with more affluent and more contextual information for each (40x, 100x, 200x, 400x)- $\left(M_{Mag\ x}^{self-attention\ transofrmation} \in R^{BxCxHxW}, [B, C, H, W]\right)$. These self-attention maps $M_{Mag\ x}^{self-attention\ transofrmation}$ are combined and used as Value (V) - $\left(M_{mag}^{self-att} \in R^{BxCxHxW}, [B, C, H, W]\right)$ and Key (K) - $\left(M_{mag}^{self-att} \in R^{BxCxHxW}, [B, C, H, W]\right)$ beside the Query (Q) - $\left(M_{Image}^{SE-ResB} \in R^{BxCxHxW}, [B, C, H, W]\right)$ for the cross-attention. This combination process presents different levels of features' detail, with more diverse representation improving the ability of the cross-attention to generalize and capture relative information. Moreover, The using of the extracted feature from the BACH full image as Query potentials the general information from the entire image, boosting the detailed information provided by the Key (K) and Value (V) from the magnified images of the magnification referred dataset, resulting in creating a cross-feature map with more meaningful and context-aware representation integrated the global context and fine-details of the labels ($M^{cross-att} \in R^{BxCxHxW}, [B, C, H, W]$) . In the classifier part, three fully connected layers were used, each followed by the normalized layer and ReLU activation function, followed by the SoftMax layer.

4.6.3 Results

The performance classification report of the proposed method is given in Table 4.2. According to the evaluation report in Table 4.2, The Normal class highlighted the perfect percentages cross the evaluation terms (Precision, Recall, F1-score) with 100%, while the Invasive class achieved the lower values with 95% for the Precision, Recall, F1-score, respectively. The Benign and InSitu classes shown a Precision and Recall between 100%-95% for each, and F1-score values equal to 97%, and 98%, respectively. In general, the model reported an overall accuracy equal to 97%, Precision equal to 98%, Recall and F1-score equal to 97%, respectively, indicating a high accuracy classification all the classes. These results highlighted the MATERB-Net model's ability via capturing the complex characteristic through the full image and the various magnification factors.

Table 4.2: Performance Classification Report of The MATERB-Net.

labels	Precision	Recall	F1-score	Accuracy
Benign	100%	95%	97%	97%
InSitu	95%	100%	98%	
Invasive	95%	95%	95%	
Normal	100%	100%	100%	
All	98%	97%	97%	

In additionally, the confusion matrix provides a visual and numerical representation for the model's predictions against true labels as shown in Figure 4.3. For example, the Normal and InSitu shown a correctly predicted for all labels with 20 out of 20 correct predications, in contrast, the benign and Invasive shown 19 out of 20 correct predications with one misclassification label for each. In overall, the confusion matrix reveals only 2 samples misclassified out of 80, and low error rate equal to 3% indicting the superior performance of the MATERB-Net model in distinguish between the four class of the BACH full image.

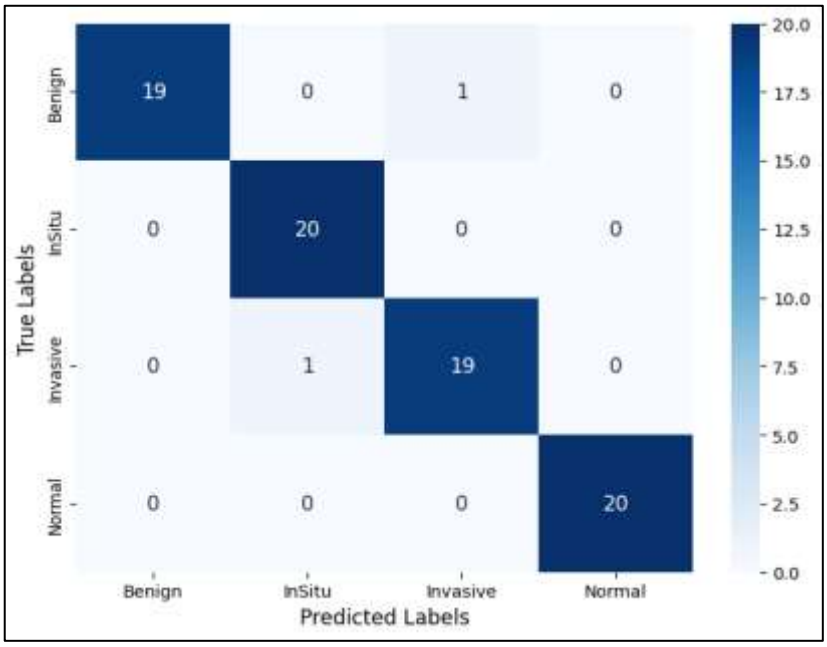

Figure 4.3: Confusion Matrix of The MATERB-Net Model.

For further evaluation performance, Figure 4.4 presents a visual evaluation of the models' ability to distinguish between classes, employing the True Positive Rate (TPR) and False Positive Rate (FPR) through the Area Under Curve (AUC) metric. The MATERB-Net model has shown AUC equal to 0.95, which means the model has a 95% probability of correctly ranking a randomly chosen positive instance higher than a negative one, and this AUC likely represents an average across all classes using the one-vs-all method, reinforcing the model's robustness.

Moreover, The Precision-Recall (PR) curve in Figure 4.5 provides a clear idea of the model trade-off between precision (positive predictive value) and recall (sensitivity) across different classification thresholds. The MATERB-Net model achieved AP equal to 0.98. This suggesting that the model maintains strong precision even as it increases recall, hence, reinforces the model's reliability in maintaining accuracy across different thresholds. Figure 4.6 presents the prediction process of the unlabeled patch level of full image

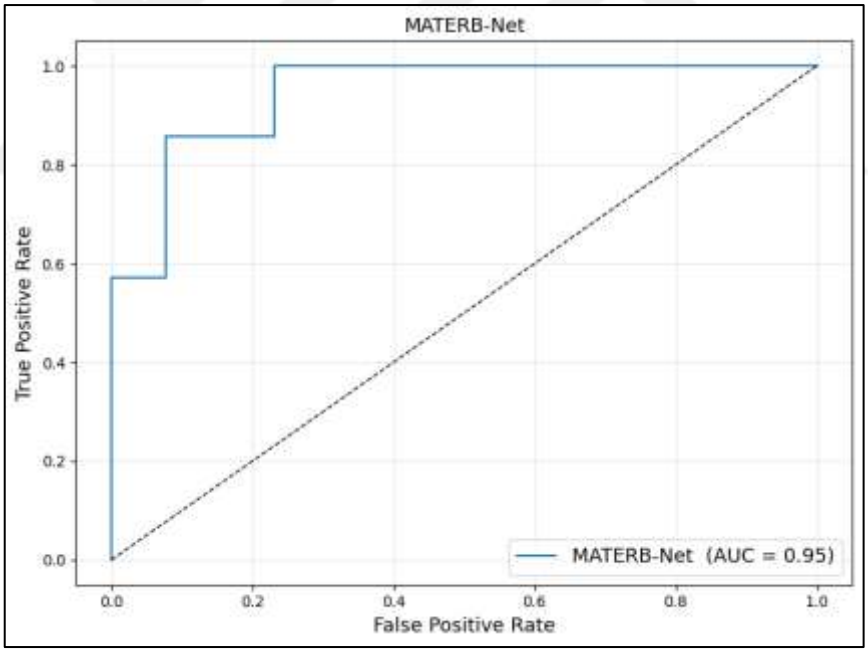

Figure 4.4: ROC Curve of The MATERB-Net Model.

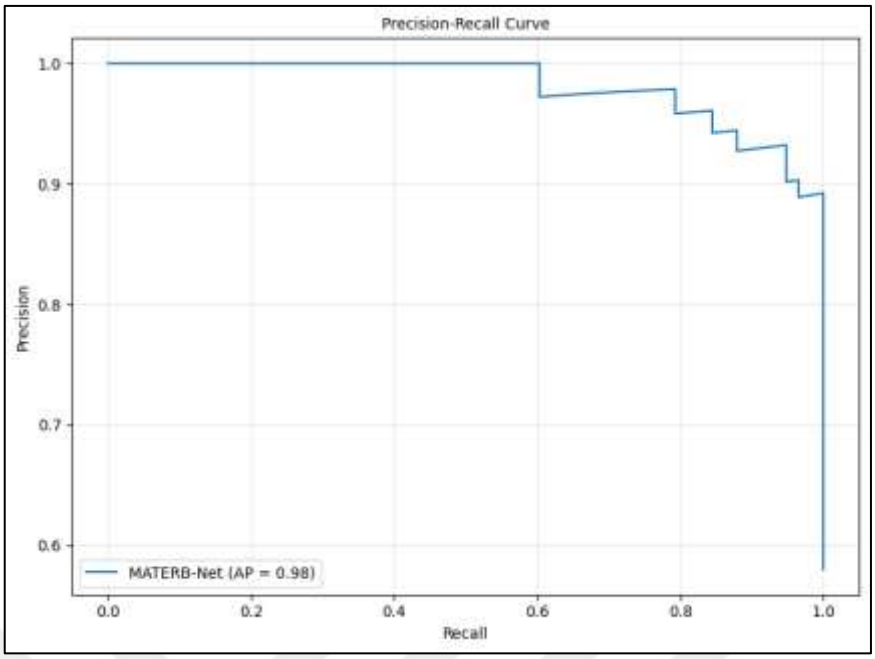

Figure 4.5: Precision-Recall Curve of The MATERB-Net Model.

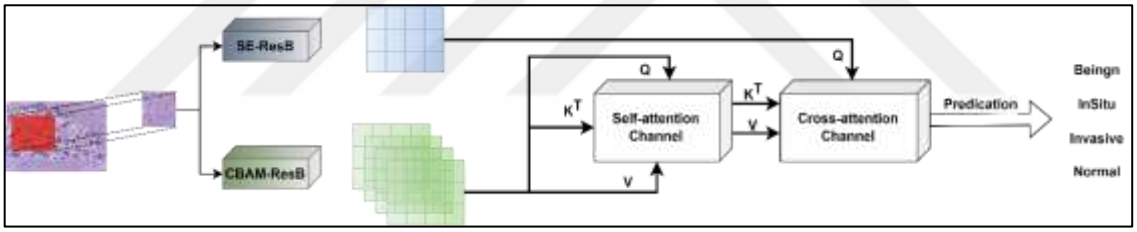

Figure 4.6: Prediction of Unknown Labels For The Patch-Level.

Table 4.3 outlines a summary of the comparison between the proposed MATERB Net model and recent approaches. The proposed MATERB-Net shows the superiority in the performance by achieving accuracy classification equal to 97%, as compared to the mentioned state-of-art; MSMV-PFENet shows with an accuracy classification of 94.8%, followed successively by DL-AM at 94.11%, MA-MIDN at 93.57%, Ensemble Model at 92.58T, and adaptive threshold based on CNN at 91.25%. In addtionally, MDFF-Net and AHoNet are quite inferior with accuracies of 86.25% and 85%, respectively. This supertority demonstrates the efficiency of the proposed MATERB-Net in histopathological image analysis.

Table 4.3: An Analysis of The Proposed MATERB-Net Model in Comparison to State-of-The-Art Methods Based on The BACH Dataset.

Reference	Method	Results
[123]	MDFF-Net	86.25%
[136]	MSMV-PFENet	94.8%
[155]	DL-AM	94.11%
[158]	Ensemble Model	92.58%
[165]	MA-MIDN	93.57%
[166]	AHoNet	85%
[167]	adaptive threshold based on CNN	91.25%
Proposed model	MATERB-Net	97%

4.7 DISCUSSION

The proposed MATERB Net is compared with recent innovations in the field of breast pathology. It has a superior capacity to address the complexity of classifying breast histopathology images. its innovative design that involves integrating SE and CBAM with ResB has had a significant impact on the way the hierarchical feature map is derived from the enlarged images and the full image. Diverse patterns, styles, and structural variations across different magnifications serve to inject the feature space with rich diversity of information dynamics about cancer. This will give the network a much wider view about how different cancer types operate. Through this multi-source learning approach, the model gets to see more varieties of morphological features from different sources and thereby improves its knowledge concerning variations in tissue patterns and styles. The results obtained underscore the robustness, adaptability, and reliability of the MATERB Net in breast classification tasks, distinguishing it from other leading methods presented in Table 4.3. The novelty of MATERB Net lies in it identifying a holistic feature based on the full images including the cellular structure, edges, patterns, tissue morphology, and texture, hence, providing a global vision. In additionally, providing a comprehensive locative pattern

for each class via the four magnification factors of the magnification refiled dataset by enhanced the features map through focus on the most relative channels and region, hence, enable the model to understand the pattern and the spatial characteristics of each class, and learning the relation between the enhanced extracted features, and provide meaningful and context-aware representation via attention transformation mechanism. Although our model has achieved remarkable and effectiveness results, there are still some limitations. first, the use of multiple modules, increases the computational complexity, which can bring infeasibility towards resource-limited environments, second, its practical feasibility since due to different types of features, such as spatial, global, and concentrated attention features, need to be carefully tuned and optimized so that they work well together, which is challenges under practical implementation conditions. Finally, although the model shows prediction equal to 94% at the patch-level based on the full image of the BACH dataset, it still some challenges in capturing the correct prediction of the unknown actual labels, might be led to incorrect diagnosis for the patch label, and requires to Doctor interference to confirm the predicted label.

4.8 CONCLUSION

In this chapter, MATERB Net: Multiscale Attention transformation based on Enhanced Residual Block for unknown Patch-level breast histopathology images prediction is proposed. Two histopathology datasets in training the proposed model are; BACH, and magnification refiled dataset (MRD) (40x, 100x, 200x, 400x) referred based on the sub-types of the BreakHis and extracted patches of the BACH (Normal class) are used in the training model. A Squeeze and Excitation (SE) and Convolutional Block Attention Module (CBAM) are used in Residual blocks to enhance the hierarchical extracted features. Each enhanced extracted feature map from (40x, 100x, 200x, 400x) - MRD via SE-ResB is processed using a self-attention mechanism to understand the correlation between extracted features in the feature map and calculated the attention score in order to determine the attention weights, resulting in new self- locative feature map with richer and contextual information for each (40x, 100x, 200x, 400x). These self-feature maps (S-FMs) are combined creating spatial-SFM, which uses with the BACH-global enhanced extracted feature map (B-GEEFM) as input to the cross-attention mechanism, the B-GEEFM as Query potentials the general information from the entire image, boosting the detailed information provided by the Key

(K) and Value (V) from the spatial-SFM, resulting in creating a cross-feature map (CFM)with more meaningful and context-aware representation integrated the global context and fine-details of the labels. The MATERB Net has achieved 97%-total accuracy classification, 98% at Precision, and 97% Recall, and F1-score, respectively. Moreover, it shows a 94% prediction probability across the unlabeled patch-level of the image. In the future work, the focus will be on implementing semi-supervised deep learning techniques and integrate a grad cam attention map to identify and utilize the most discriminative regions within each microscopy image to enhance the selection of feature in order to improve the accuracy prediction.

5. CONCLUSION

The present dissertation deals with the most limitation problems in the use of deep learning in the classification of histopathology images, such as the availability of insufficient numbers of labeled images for training, micro-level variations in histopathological images that affect the training process.

In Chapter 3, we introduce a framework titled "HAFMAB-Net: Hierarchical Adaptive Fusion based on Multilevel Attention-Enhanced Bottleneck Neural Network" for the classification of histopathological images. A new architecture presented by the model combines a bottleneck network supported by Channel and local attention modules in order to extract local and global features and characteristics. In addition, the deep spatial clustering attention module improves and enhances the delivery of debugging features by focusing on accurate and distinctive information, which in turn improves the network's ability to handle complex details. These features are dynamically integrated into the adaptive integration module, which in turn offers a powerful integration process and scalability of the network in focusing on features in a multi-scale and hierarchical manner

Three types of data were used to evaluate the model and the model achieved excellent performance during all tasks. In BACH dataset, the proposed model achieved a rating accuracy equal to 99% and an AUC score of 1.00 reflecting the distinctive performance of the model, while the model achieved a division accuracy of 98.99% and an AUC score of 1.00 during the unbalanced BreaKHis data, demonstrating the generalization ability of the model even with difficult data distribution, and achieved a perfect accuracy of 100% and an AUC score of 1.00 for LC25000. These results indicate the effectiveness, reliability and ability of the proposed model to adapt and overcome the problems faced by traditional deep learning models and other modern methods

In future work, the focus will be on extending the capabilities of HAFMAB-Net by integrating unsupervised learning techniques to identify and classify different regions within a single image rather than analyzing the entire image. This approach can potentially predict the unknown labels of sub-images within histopathological slides, enabling more granular and precise diagnoses. Such advancements could pave the way for broader applications of the model in multi-type cancer detection and more personalized diagnostic tools.

In Chapter 4, we discuss the predication of the unknown labeled of the sub-images within histopathological slides via proposed a Multiscale Attention transformation based on Enhanced Residual Block with multi-head transformation self-attention integrated with cross-attention. A refiled dataset titled “Magnification refiled dataset (MRD)" using the Robotics and Imaging Laboratory at the Federal University of Parana in Brazil (BreakHis) and the BACH dataset provided by Grand Challenge are both used in the training proposed model. A Squeeze and Excitation (SE) and Convolutional Block Attention Module (CBAM) are used in Residual blocks to enhance the hierarchical extracted features. Each enhanced extracted feature map from (40x, 100x, 200x, 400x) - MRD via SE-ResB is processed using a self-attention mechanism to understand the correlation between extracted features in the feature map and calculated the attention score to determine the attention weights, resulting in new self- locative feature map with richer and contextual information for each (40x, 100x, 200x, 400x). These self-feature maps (S-FMs) are combined, creating spatial-SFM, which uses the BACH-global enhanced extracted feature map (B-GEEFM) as input to the cross-attention mechanism, the B-GEEFM as Query potentials the general information from the entire image, boosting the detailed information provided by the Key (K) and Value (V) from the spatial-SFM, resulting in creating a cross-feature map (CFM)with more meaningful and context-aware representation integrated the global context and fine-details of the labels. The MATERB Net has achieved 97%-total accuracy classification, 98% at Precision, 97% Recall, and F1-score, respectively. Moreover, it shows a 94% prediction probability across the unlabeled patch level of the image. In future work, the focus will be on implementing semi-supervised deep learning techniques and integrating a grad cam attention map to identify and utilize the most discriminative regions within each microscopy image to enhance the selection of features and improve accuracy prediction.

In the future, the focus will be on implementing semi-supervised deep learning techniques and integrating a grad cam attention map to identify and utilize the most discriminative regions within each microscopy image to enhance feature selection and improve accuracy prediction.

REFERENCES

[1] H. Abdel-Nabi et al., “A comprehensive review of the deep learning-based tumor analysis approaches in histopathological images: segmentation, classification and multi-learning tasks,” Cluster Computing, vol. 26, no. 5, pp. 3145–3185, Jan. 2023, doi: 10.1007/s10586-022-03951-2.

[2] A. A. Joseph, M. Abdullahi, S. B. Junaidu, H. H. Ibrahim, and H. Chiroma, “Improved multi-classification of breast cancer histopathological images using handcrafted features and deep neural network (dense layer),” Intelligent Systems With Applications, vol. 14, p. 200066, Feb. 2022, doi: 10.1016/j.iswa.2022.200066.

[3] A. Merabet, A. Saighi, M. A. Ferradji, and Z. Laboudi, “Enhancing Colon Cancer Prediction in Histopathology with Integrated Deep Learning Models: A Comparative Study on the LC25000 Dataset,” 2024 6th International Conference on Pattern Analysis and Intelligent Systems (PAIS). IEEE, pp. 1–7, Apr. 2024, doi: 10.1109/pais62114.2024.10541219.

[4] Y. Wu et al., “Recent Advances of Deep Learning for Computational Histopathology: Principles and applications,” Cancers, vol. 14, no. 5, p. 1199, Feb. 2022, doi: 10.3390/cancers14051199.

[5] A. A. Ahmed, M. Abouzid, and E. Kaczmarek, “Deep learning approaches in histopathology,” Cancers, vol. 14, no. 21, p. 5264, Oct. 2022, doi: 10.3390/cancers14215264.

[6] Y. Zhao, J. Zhang, D. Hu, H. Qu, Y. Tian, and X. Cui, “Application of Deep Learning in Histopathology Images of Breast Cancer: a review,” Micromachines, vol. 13, no. 12, p. 2197, Dec. 2022, doi: 10.3390/mi13122197.

[7] P. Saxena, S. K. Singh, G. Srivastav, and R. Mamoon, “Computer-Assisted analysis of histopathological images,” in Advances in medical technologies and clinical practice book series, 2024, pp. 77–120. doi: 10.4018/979-8-3693-5226-7.ch004.

[8] S. Banerji and S. Mitra, “Deep learning in histopathology: A review,” Wiley Interdisciplinary Reviews Data Mining and Knowledge Discovery, vol. 12, no. 1, Nov. 2021, doi: 10.1002/widm.1439.

[9] S. U. R. Khan, A. Raza, I. Shahzad, and G. Ali, “Enhancing Concrete and Pavement Crack Prediction through Hierarchical Feature Integration with VGG16 and Triple Classifier Ensemble,” 2024 Horizons of Information Technology and Engineering (HITE). IEEE, Oct. 2024, doi: 10.1109/hite63532.2024.10777242.

[10] U. S. Khan and S. U. R. Khan, “Boost diagnostic performance in retinal disease classification utilizing deep ensemble classifiers based on OCT,” Multimedia Tools and Applications, Jul. 2024, doi: 10.1007/s11042-024-19922-1.

[11] U. S. Khan, M. Ishfaque, S. U. R. Khan, F. Xu, L. Chen, and Y. Lei, “Comparative analysis of twelve transfer learning models for the prediction and crack detection in concrete dams, based on borehole images,” Frontiers of Structural and Civil Engineering, Sep. 2024, doi: 10.1007/s11709-024-1090-2.

[12] S. U. R. Khan and S. Asif, “Oral cancer detection using feature-level fusion and novel self-attention mechanisms,” Biomedical Signal Processing and Control, vol. 95, p. 106437, May 2024, doi: 10.1016/j.bspc.2024.106437.

[13] S. Boumaraf et al., “Conventional Machine Learning versus Deep Learning for Magnification Dependent Histopathological Breast Cancer Image Classification: A Comparative Study with Visual Explanation,” Diagnostics, vol. 11, no. 3, p. 528, Mar. 2021, doi: 10.3390/diagnostics11030528.

[14] A. H. Abdulaal et al., “CUTTING-EDGE CNN APPROACHES FOR BREAST HISTOPATHOLOGICAL CLASSIFICATION: THE IMPACT OF SPATIAL ATTENTION MECHANISMS,” ShodhAI: Journal of Artificial Intelligence, vol. 1, no. 1, Oct. 2024, doi: 10.29121/shodhai.v1.i1.2024.14.

[15] A. B. Hamida et al., “Deep learning for colon cancer histopathological images analysis,” Computers in Biology and Medicine, vol. 136, p. 104730, Aug. 2021, doi: 10.1016/j.compbiomed.2021.104730.

[16] R. Rashmi, K. Prasad, and C. B. K. Udupa, “Breast histopathological image analysis using image processing techniques for diagnostic purposes: A methodological review,” Journal of Medical Systems, vol. 46, no. 1, Dec. 2021, doi: 10.1007/s10916-021-01786-9.

[17] J. T. Ash, G. Darnell, D. Munro, and B. E. Engelhardt, “Joint analysis of expression levels and histological images identifies genes associated with tissue morphology,” Nature Communications, vol. 12, no. 1, Mar. 2021, doi: 10.1038/s41467-021-21727-x.

[18] M. M. Abdelsamea, U. Zidan, Z. Senousy, M. M. Gaber, E. Rakha, and M. Ilyas, “A survey on artificial intelligence in histopathology image analysis,” Wiley Interdisciplinary Reviews Data Mining and Knowledge Discovery, vol. 12, no. 6, Jul. 2022, doi: 10.1002/widm.1474.

[19] L. Hou, Deep Learning for Digital Histopathology across Multiple Scales. State University of New York at Stony Brook, 2019.

[20] M. M. Abdelsamea, U. Zidan, Z. Senousy, M. M. Gaber, E. Rakha, and M. Ilyas, “A survey on artificial intelligence in histopathology image analysis,” Wiley Interdisciplinary Reviews Data Mining and Knowledge Discovery, vol. 12, no. 6, Jul. 2022, doi: 10.1002/widm.1474.

[21] S. Taheri and Z. Golrizkhatami, “Magnification-specific and magnification-independent classification of breast cancer histopathological image using deep learning approaches,” Signal Image and Video Processing, vol. 17, no. 2, pp. 583–591, Jun. 2022, doi: 10.1007/s11760-022-02263-7.

[22] D. Sambyal and A. Sarwar, “Recent developments in cervical cancer diagnosis using deep learning on whole slide images: An Overview of models, techniques, challenges and future directions,” Micron, vol. 173, p. 103520, Jul. 2023, doi: 10.1016/j.micron.2023.103520.

[23] Y. Chen et al., “A whole-slide image (WSI)-based immunohistochemical feature prediction system improves the subtyping of lung cancer,” Lung Cancer, vol. 165, pp. 18–27, Jan. 2022, doi: 10.1016/j.lungcan.2022.01.005.

[24] W.-W. Hsu et al., “A weakly supervised deep learning-based method for glioma subtype classification using WSI and mpMRIs,” Scientific Reports, vol. 12, no. 1, Apr. 2022, doi: 10.1038/s41598-022-09985-1.

[25] K. Lee et al., “Deep learning of histopathology images at the single cell level,” Frontiers in Artificial Intelligence, vol. 4, Sep. 2021, doi: 10.3389/frai.2021.754641.

[26] D. S. McClintock, J. T. Abel, and T. C. Cornish, “Whole slide imaging hardware, software, and infrastructure,” in Cham: Springer International Publishing, 2021, pp. 23–56. doi: 10.1007/978-3-030-83332-9_2.

[27] H. Kim et al., “Deep learning-based histopathological segmentation for whole slide images of colorectal cancer in a compressed domain,” Scientific Reports, vol. 11, no. 1, Nov. 2021, doi: 10.1038/s41598-021-01905-z.

[28] C. Mercan et al., “Deep learning for fully-automated nuclear pleomorphism scoring in breast cancer,” Npj Breast Cancer, vol. 8, no. 1, Nov. 2022, doi: 10.1038/s41523-022-00488-w.

[29] H. Peregrina-Barreto, V. Y. Ramirez-Guatemala, G. C. Lopez-Armas, and J. A. Cruz-Ramos, “Characterization of nuclear pleomorphism and tubules in histopathological images of breast cancer,” Sensors, vol. 22, no. 15, p. 5649, Jul. 2022, doi: 10.3390/s22155649.

[30] A. Nadeem, R. Ashraf, T. Mahmood, and S. Parveen, “Automated CAD system for early detection and classification of pancreatic cancer using deep learning model,” PLoS ONE, vol. 20, no. 1, p. e0307900, Jan. 2025, doi: 10.1371/journal.pone.0307900.

[31] A. Shmatko, N. G. Laleh, M. Gerstung, and J. N. Kather, “Artificial intelligence in histopathology: enhancing cancer research and clinical oncology,” Nature Cancer, vol. 3, no. 9, pp. 1026–1038, Sep. 2022, doi: 10.1038/s43018-022-00436-4.

[32] Abdulwahhab, Ali H., Noof T. Mahmood, et al. “A Review on Medical Image Applications Based on Deep Learning Techniques.” Journal of Image and Graphics, vol. 12, no. 3, Jan. 2024, pp. 215–27. https://doi.org/10.18178/joig.12.3.215-227.

[33] P. Jannatdoust et al., “Computer-Aided Detection (CADE) and segmentation methods for breast cancer using magnetic resonance imaging (MRI),” Journal of Magnetic Resonance Imaging, Jan. 2025, doi: 10.1002/jmri.29687.

[34] B. Hunter, S. Hindocha, and R. W. Lee, “The role of artificial intelligence in early cancer diagnosis,” Cancers, vol. 14, no. 6, p. 1524, Mar. 2022, doi: 10.3390/cancers14061524.

[35] M. Sufyan, Z. Shokat, and U. A. Ashfaq, “Artificial intelligence in cancer diagnosis and therapy: Current status and future perspective,” Computers in Biology and Medicine, vol. 165, p. 107356, Aug. 2023, doi: 10.1016/j.compbiomed.2023.107356.

[36] M. Arnold et al., “Current and future burden of breast cancer: Global statistics for 2020 and 2040,” The Breast, vol. 66, pp. 15–23, Sep. 2022, doi: 10.1016/j.breast.2022.08.010.

[37] R. Sharma, “Global, regional, national burden of breast cancer in 185 countries: evidence from GLOBOCAN 2018,” Breast Cancer Research and Treatment, vol. 187, no. 2, pp. 557–567, Jan. 2021, doi: 10.1007/s10549-020-06083-6.

[38] J. Huang et al., “Global incidence and mortality of breast cancer: a trend analysis,” Aging, vol. 13, no. 4, pp. 5748–5803, Feb. 2021, doi: 10.18632/aging.202502.

[39] G. Forrai et al., “Use of diagnostic imaging modalities in modern screening, diagnostics and management of breast tumours 1st Central-Eastern European Professional Consensus Statement on Breast Cancer,” Pathology & Oncology Research, vol. 28, Jun. 2022, doi: 10.3389/pore.2022.1610382.

[40] E. A. Abeelh and Z. AbuAbeileh, “Comparative effectiveness of mammography, ultrasound, and MRI in the detection of breast carcinoma in dense breast tissue: a systematic review,” Cureus, Apr. 2024, doi: 10.7759/cureus.59054.

[41] R. Iacob et al., “Evaluating the Role of Breast ultrasound in early detection of breast Cancer in Low- and Middle-Income Countries: A Comprehensive Narrative Review,” Bioengineering, vol. 11, no. 3, p. 262, Mar. 2024, doi: 10.3390/bioengineering11030262.

[42] P. Steyerova and A. Burgetova, “Current imaging techniques and impact on diagnosis and survival —a narrative review,” Annals of Breast Surgery, vol. 6, p. 25, Jul. 2021, doi: 10.21037/abs-21-22.

[43] R. Gerami et al., “A literature review on the imaging methods for breast cancer,” International Journal of Physiology, Pathophysiology and Pharmacology, vol. 14, no. 4, p. 171, Jun. 2022.

[44] M. Daniaux et al., “Preoperative staging by multimodal imaging in newly diagnosed breast cancer: Diagnostic performance of contrast-enhanced spectral mammography compared to conventional mammography, ultrasound, and MRI,” European Journal of Radiology, vol. 163, p. 110838, Apr. 2023, doi: 10.1016/j.ejrad.2023.110838.

[45] “Breast biopsy and cancer: What results do and don’t mean,” Jan. 25, 2025. https://www.verywellhealth.com/breast-biopsy-7966144 (accessed Mar. 10, 2025).

[46] T. H. Rafi, R. M. Shubair, F. Farhan, Md. Z. Hoque, and F. M. Quayyum, “Recent advances in Computer-Aided medical diagnosis using machine learning algorithms with optimization techniques,” IEEE Access, vol. 9, pp. 137847–137868, Jan. 2021, doi: 10.1109/access.2021.3108892.

[47] Y. A. Kadhim, M. U. Khan, and A. Mishra, “Deep Learning-Based Computer-Aided Diagnosis (CAD): applications for medical image datasets,” Sensors, vol. 22, no. 22, p. 8999, Nov. 2022, doi: 10.3390/s22228999.

[48] R. Guetari, H. Ayari, and H. Sakly, “Computer-aided diagnosis systems: a comparative study of classical machine learning versus deep learning-based approaches,” Knowledge and Information Systems, vol. 65, no. 10, pp. 3881–3921, May 2023, doi: 10.1007/s10115-023-01894-7.

[49] N. Coudray et al., “Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning,” Nature Medicine, vol. 24, no. 10, pp. 1559–1567, Sep. 2018, doi: 10.1038/s41591-018-0177-5.

[50] A. S. N. Raju, K. Jayavel, and T. Rajalakshmi, “Intelligent recognition of colorectal cancer combining application of computer-assisted diagnosis with deep learning approaches,” International Journal of Power Electronics and Drive Systems/International Journal of Electrical and Computer Engineering, vol. 12, no. 1, p. 738, Nov. 2021, doi: 10.11591/ijece.v12i1.pp738-747.

[51] M. Tharwat, N. A. Sakr, S. El-Sappagh, H. Soliman, K.-S. Kwak, and M. Elmogy, “Colon cancer diagnosis based on Machine learning and Deep Learning: Modalities and analysis techniques,” Sensors, vol. 22, no. 23, p. 9250, Nov. 2022, doi: 10.3390/s22239250.

[52] Q. Jiang, Y. Yu, Y. Ren, S. Li, and X. He, “A review of deep learning methods for gastrointestinal diseases classification applied in computer-aided diagnosis system,” Medical & Biological Engineering & Computing, Sep. 2024, doi: 10.1007/s11517-024-03203-y.

[53] Q. N. Van Bokhorst, B. B. Houwen, Y. Hazewinkel, P. Fockens, and E. Dekker, “Advances in artificial intelligence and computer science for computer-aided diagnosis

of colorectal polyps: current status," Endoscopy International Open, vol. 11, no. 08, pp. E752–E767, Aug. 2023, doi: 10.1055/a-2098-1999.

[54] X. Pang, Z. Zhao, and Y. Weng, "The role and Impact of Deep Learning Methods in Computer-Aided Diagnosis using Gastrointestinal Endoscopy," Diagnostics, vol. 11, no. 4, p. 694, Apr. 2021, doi: 10.3390/diagnostics11040694.

[55] S. Nazarian, B. Glover, H. Ashrafian, A. Darzi, and J. Teare, "Diagnostic Accuracy of Artificial Intelligence and Computer-Aided diagnosis for the detection and characterization of colorectal polyps: Systematic review and Meta-analysis," Journal of Medical Internet Research, vol. 23, no. 7, p. e27370, May 2021, doi: 10.2196/27370.

[56] M. Masud, N. Sikder, A.-A. Nahid, A. K. Bairagi, and M. A. AlZain, "A machine learning approach to diagnosing lung and colon cancer using a Deep Learning-Based Classification framework," Sensors, vol. 21, no. 3, p. 748, Jan. 2021, doi: 10.3390/s21030748.

[57] Y. Yang, Y. Zhou, C. Zhou, and X. Ma, "Novel computer aided diagnostic models on multimodality medical images to differentiate well differentiated liposarcomas from lipomas approached by deep learning methods," Orphanet Journal of Rare Diseases, vol. 17, no. 1, Apr. 2022, doi: 10.1186/s13023-022-02304-x.

[58] H. Tekchandani, S. Verma, N. D. Londhe, R. R. Jain, and A. Tiwari, "Computer aided diagnosis system for cervical lymph nodes in CT images using deep learning," Biomedical Signal Processing and Control, vol. 71, p. 103158, Sep. 2021, doi: 10.1016/j.bspc.2021.103158.

[59] Y. Zhang et al., "Histopathology images-based deep learning prediction of prognosis and therapeutic response in small cell lung cancer," Npj Digital Medicine, vol. 7, no. 1, Jan. 2024, doi: 10.1038/s41746-024-01003-0.

[60] M. Zacharias et al., "Expanding broad molecular reflex testing in Non-Small cell lung Cancer to squamous histology," Cancers, vol. 16, no. 5, p. 903, Feb. 2024, doi: 10.3390/cancers16050903.

[61] R. Javed, T. Abbas, A. H. Khan, A. Daud, A. Bukhari, and R. Alharbey, "Deep learning for lungs cancer detection: a review," Artificial Intelligence Review, vol. 57, no. 8, p. 197, Jul. 2024, doi: 10.1007/s10462-024-10807-1.

[62] M. French, “Small Cell vs. Non-Small Cell Lung Cancer: How Do They Differ?” Healthgrades, Mar. 23, 2023. [Online]. Available: https://resources.healthgrades.com/right-care/lung-cancer/small-cell-vs-non-small-cell-lung-cancer-whats-the-difference. Accessed: Mar. 15, 2025.

[63] G. Aresta et al., “BACH: Grand challenge on breast cancer histology images,” Medical Image Analysis, vol. 56, pp. 122–139, May 2019, doi: 10.1016/j.media.2019.05.010.

[64] A. Spanhol, L. S. Oliveira, C. Petitjean and L. Heutte, "Breast cancer histopathological image classification using Convolutional Neural Networks," 2016 International Joint Conference on Neural Networks (IJCNN), Vancouver, BC, Canada, 2016, pp. 2560-2567, doi: 10.1109/IJCNN.2016.7727519.

[65] A. A. Borkowski, M. M. Bui, L. B. Thomas, C. P. Wilson, L. A. DeLand, and S. M. Mastorides, “Lung and colon Cancer Histopathological Image Dataset (LC25000),” arXiv.org, Dec. 16, 2019. https://arxiv.org/abs/1912.12142

[66] K. Faryna, J. Van Der Laak, and G. Litjens, “Automatic data augmentation to improve generalization of deep learning in H&E stained histopathology,” Computers in Biology and Medicine, vol. 170, p. 108018, Jan. 2024, doi: 10.1016/j.compbiomed.2024.108018.

[67] F. Li, Z. Hu, W. Chen, and A. Kak, “A Laplacian Pyramid Based Generative H&E Stain Augmentation Network,” IEEE Transactions on Medical Imaging, vol. 43, no. 2, pp. 701–713, Sep. 2023, doi: 10.1109/tmi.2023.3317239.

[68] B. Muñoz, R. Pezoa, and H. Gutierrez, “Histopathology image augmentation through StyleGAN2-ADA,” in Communications in computer and information science, 2025, pp. 216–228. doi: 10.1007/978-3-031-80084-9_15.

[69] N. Marini et al., “Data-driven color augmentation for H&E stained images in computational pathology,” Journal of Pathology Informatics, vol. 14, p. 100183, Jan. 2023, doi: 10.1016/j.jpi.2022.100183.

[70] A. M. Alhassan, “Enhanced pre-processing technique for histopathological image stain normalization and cancer detection,” Multimedia Tools and Applications, Oct. 2024, doi: 10.1007/s11042-024-20216-9.

[71] Md. Z. Hoque, A. Keskinarkaus, P. Nyberg, and T. Seppänen, “Stain normalization methods for histopathology image analysis: A comprehensive review and experimental comparison,” Information Fusion, vol. 102, p. 101997, Sep. 2023, doi: 10.1016/j.inffus.2023.101997.

[72] D. Murcia-Gómez, I. Rojas-Valenzuela, and O. Valenzuela, “Impact of image preprocessing methods and deep learning models for classifying histopathological breast cancer images,” Applied Sciences, vol. 12, no. 22, p. 11375, Nov. 2022, doi: 10.3390/app122211375.

[73] K. D. McCombe et al., “HistoClean: Open-source software for histological image pre-processing and augmentation to improve development of robust convolutional neural networks,” Computational and Structural Biotechnology Journal, vol. 19, pp. 4840–4853, Jan. 2021, doi: 10.1016/j.csbj.2021.08.033.

[74] X. Xing, X. Zhao, H. Wei, and Y. Li, “Diagnostic accuracy of different computer-aided diagnostic systems for prostate cancer based on magnetic resonance imaging,” Medicine, vol. 100, no. 3, p. e23817, Jan. 2021, doi: 10.1097/md.0000000000023817.

[75] Z. Guo et al., “A review of the current state of the computer-aided diagnosis (CAD) systems for breast cancer diagnosis,” Open Life Sciences, vol. 17, no. 1, pp. 1600–1611, Jan. 2022, doi: 10.1515/biol-2022-0517.

[76] T. Habuza et al., “AI applications in robotics, diagnostic image analysis and precision medicine: Current limitations, future trends, guidelines on CAD systems for medicine,” Informatics in Medicine Unlocked, vol. 24, p. 100596, Jan. 2021, doi: 10.1016/j.imu.2021.100596.

[77] O. Debellotte et al., “Artificial intelligence and Early detection of breast, lung, and colon Cancer: A Narrative review,” Cureus, vol. 17, no. 2, Feb. 2025, doi: 10.7759/cureus.79199.

[78] Chugh, Gunjan, et al. “Survey on Machine Learning and Deep Learning Applications in Breast Cancer Diagnosis.” Cognitive Computation, vol. 13, no. 6, Jan. 2021, pp. 1451–70. https://doi.org/10.1007/s12559-020-09813-6.

[79] Lee, Jeongmin, et al. “Evaluation of Computer-Aided Detection (CAD) in Screening Automated Breast Ultrasound Based on Characteristics of CAD Marks and False-

Positive Marks." Diagnostics, vol. 12, no. 3, Feb. 2022, p. 583. https://doi.org/10.3390/diagnostics12030583.

[80] Gonçalves, Caroline Barcelos, et al. "CNN Architecture Optimization Using Bio-inspired Algorithms for Breast Cancer Detection in Infrared Images." Computers in Biology and Medicine, vol. 142, Jan. 2022, p. 105205. https://doi.org/10.1016/j.compbiomed.2021.105205.

[81] Yoon, Jung Hyun, et al. "Artificial Intelligence-based Computer-assisted Detection/Diagnosis (AI-CAD) for Screening Mammography: Outcomes of AI-CAD in the Mammographic Interpretation Workflow." European Journal of Radiology Open, vol. 11, July 2023, p. 100509. https://doi.org/10.1016/j.ejro.2023.100509.

[82] B. Jones, D. Hildreth, D. Chen, I. Baran, V. G. Kim, and A. Schulz, "AutoMate," ACM Transactions on Graphics, vol. 40, no. 6, pp. 1–18, Dec. 2021, doi: 10.1145/3478513.3480562.

[83] R. Krishnan, P. Rajpurkar, and E. J. Topol, "Self-supervised learning in medicine and healthcare," Nature Biomedical Engineering, vol. 6, no. 12, pp. 1346–1352, Aug. 2022, doi: 10.1038/s41551-022-00914-1.

[84] X. Li et al., "A comprehensive review of computer-aided whole-slide image analysis: from datasets to feature extraction, segmentation, classification and detection approaches," Artificial Intelligence Review, vol. 55, no. 6, pp. 4809–4878, Jan. 2022, doi: 10.1007/s10462-021-10121-0.

[85] Phung, Justin. "Artificial Intelligence in Diagnosing Lung Cancer: Applications and Future Developments." The Meducator, no. 39, Mar. 2021, https://doi.org/10.35493/medu.39.18.

[86] Coudray, Nicolas, et al. "Classification and Mutation Prediction From Non–small Cell Lung Cancer Histopathology Images Using Deep Learning." Nature Medicine, vol. 24, no. 10, Sept. 2018, pp. 1559–67. https://doi.org/10.1038/s41591-018-0177-5.

[87] Lei, Zhenfeng, et al. "A Novel Data-driven Robust Framework Based on Machine Learning and Knowledge Graph for Disease Classification." Future Generation Computer Systems, vol. 102, Sept. 2019, pp. 534–48. https://doi.org/10.1016/j.future.2019.08.030.

[88] Xu, Lixue, et al. “Diagnostic Accuracy and Generalizability of a Deep Learning-Based Fully Automated Algorithm for Coronary Artery Stenosis Detection on CCTA: A Multi-Centre Registry Study.” Frontiers in Cardiovascular Medicine, vol. 8, Nov. 2021, https://doi.org/10.3389/fcvm.2021.707508.

[89] Mi, Weiming, et al. “Deep Learning-Based Multi-Class Classification of Breast Digital Pathology Images.” Cancer Management and Research, vol. Volume 13, June 2021, pp. 4605–17. https://doi.org/10.2147/cmar.s312608.

[90] Xiao, Zuoxiu, et al. “Emerging Early Diagnostic Methods for Acute Kidney Injury.” Theranostics, vol. 12, no. 6, Jan. 2022, pp. 2963–86. https://doi.org/10.7150/thno.71064.

[91] mrgrhn. (2021, Jan. 23). VGGNet with TensorFlow (Transfer Learning with VGG16 Included). [Online]. Available: https://ai.plainenglish.io/vggnet-with-tensorflow-transfer-learning-with-vgg16-included-7e5f6fa9479a. Accessed: Mar. 15, 2025.

[92] S. Sharma, K. Guleria, S. Tiwari, and S. Kumar, “A deep learning based convolutional neural network model with VGG16 feature extractor for the detection of Alzheimer Disease using MRI scans,” Measurement Sensors, vol. 24, p. 100506, Oct. 2022, doi: 10.1016/j.measen.2022.100506.

[93] S. Sharma and K. Guleria, “A Deep Learning based model for the Detection of Pneumonia from Chest X-Ray Images using VGG-16 and Neural Networks,” Procedia Computer Science, vol. 218, pp. 357–366, Jan. 2023, doi: 10.1016/j.procs.2023.01.018.

[94] N. Gouda and J. Amudha, "Skin Cancer Classification using ResNet," 2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA), Greater Noida, India, 2020, pp. 536-541, doi: 10.1109/ICCCA49541.2020.9250855.

[95] M. Wu et al., “Detecting neonatal acute bilirubin encephalopathy based on T1-weighted MRI images and learning-based approaches,” BMC Medical Imaging, vol. 21, no. 1, Jun. 2021, doi: 10.1186/s12880-021-00634-z.

[96] L. Alzubaidi et al., “Review of deep learning: concepts, CNN architectures, challenges, applications, future directions,” Journal of Big Data, vol. 8, no. 1, Mar. 2021, doi: 10.1186/s40537-021-00444-8.

[97] B. Koonce, “ResNet 50,” in Convolutional Neural Networks with Swift for TensorFlow, Apress, Berkeley, CA, 2021, pp. 63–72. doi: 10.1007/978-1-4842-6168-2_6.

[98] Md. B. Hossain, S. M. H. S. Iqbal, Md. M. Islam, Md. N. Akhtar, and I. H. Sarker, “Transfer learning with fine-tuned deep CNN ResNet50 model for classifying COVID-19 from chest X-ray images,” Informatics in Medicine Unlocked, vol. 30, p. 100916, Jan. 2022, doi: 10.1016/j.imu.2022.100916.

[99] R. Anand, S. V. Lakshmi, D. Pandey, and B. K. Pandey, “An enhanced ResNet-50 deep learning model for arrhythmia detection using electrocardiogram biomedical indicators,” Evolving Systems, vol. 15, no. 1, pp. 83–97, Dec. 2023, doi: 10.1007/s12530-023-09559-0.

[100] Y. Chuang, S. Zhang, and X. Zhao, “Deep learning-based panoptic segmentation: Recent advances and perspectives,” IET Image Processing, vol. 17, no. 10, pp. 2807–2828, Jun. 2023, doi: 10.1049/ipr2.12853.

[101] V. N, M. Chinta, K. E and S. V. L, "Classification of Pneumonia using InceptionNet, ResNet and CNN," 2022 6th International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS), Bangalore, India, 2022, pp. 1-6, doi: 10.1109/CSITSS57437.2022.10026402.

[102] X. Wang et al., “A recognition method of ancient architectures based on the improved Inception V3 model,” Symmetry, vol. 14, no. 12, p. 2679, Dec. 2022, doi: 10.3390/sym14122679.

[103] N. K. Ratha, V. M. Patel, and R. Chellappa, Deep Learning-Based Face Analytics. 2021. doi: 10.1007/978-3-030-74697-1.

[104] B. Koonce, “MobileNetV3,” in Convolutional Neural Networks with Swift for TensorFlow, Apress, Berkeley, CA, 2021, pp. 125–144. doi: 10.1007/978-1-4842-6168-2_11.

[105] A. Howard et al., “Searching for MobileNetV3,” in Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV), 2019, pp. 1314–1324.

[106] L. Zhao and L. Wang, “A new lightweight network based on MobileNetV3,” KSII Transactions on Internet and Information Systems, vol. 16, no. 1, Jan. 2022, doi: 10.3837/tiis.2022.01.001.

[107] C. Bi, S. Xu, N. Hu, S. Zhang, Z. Zhu, and H. Yu, “Identification method of corn leaf disease based on improved MobileneTV3 model,” Agronomy, vol. 13, no. 2, p. 300, Jan. 2023, doi: 10.3390/agronomy13020300.

[108] A. A. Mukhlif, B. Al-Khateeb, and M. A. Mohammed, “An extensive review of state-of-the-art transfer learning techniques used in medical imaging: Open issues and challenges,” Journal of Intelligent Systems, vol. 31, no. 1, pp. 1085–1111, Jan. 2022, doi: 10.1515/jisys-2022-0198.

[109] A. W. Salehi et al., “A study of CNN and transfer learning in Medical imaging: Advantages, challenges, future scope,” Sustainability, vol. 15, no. 7, p. 5930, Mar. 2023, doi: 10.3390/su15075930.

[110] S. H. Kassani, P. H. Kassani, M. J. Wesolowski, K. A. Schneider, and R. Deters, “Deep transfer learning based model for colorectal cancer histopathology segmentation: A comparative study of deep pre-trained models,” International Journal of Medical Informatics, vol. 159, p. 104669, Dec. 2021, doi: 10.1016/j.ijmedinf.2021.104669.

[111] S. J. Shin et al., “Style transfer strategy for developing a generalizable deep learning application in digital pathology,” Computer Methods and Programs in Biomedicine, vol. 198, p. 105815, Oct. 2020, doi: 10.1016/j.cmpb.2020.105815.

[112] J. Gupta, S. Pathak, and G. Kumar, “Deep Learning (CNN) and Transfer Learning: A review,” Journal of Physics Conference Series, vol. 2273, no. 1, p. 012029, May 2022, doi: 10.1088/1742-6596/2273/1/012029.

[113] C. Desai, “Image classification using transfer learning and deep learning,” International Journal of Engineering and Computer Science, vol. 10, no. 9, pp. 25394–25398, Sep. 2021, doi: 10.18535/ijecs/v10i9.4622.

[114] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2016, pp. 770–778.

[115] H. Zhou, Z. Liu, T. Li, Y. Chen, W. Huang, and Z. Zhang, “Classification of precancerous lesions based on fusion of multiple hierarchical features,” Computer Methods and Programs in Biomedicine, vol. 229, p. 107301, Dec. 2022, doi: 10.1016/j.cmpb.2022.107301.

[116] M. Moradi, Y. Chen, X. Du, and J. M. Seddon, “Deep ensemble learning for automated non-advanced AMD classification using optimized retinal layer segmentation and SD-OCT scans,” Computers in Biology and Medicine, vol. 154, p. 106512, Jan. 2023, doi: 10.1016/j.compbiomed.2022.106512.

[117] Abdulwahhab, Ali H., et al. “HAFMAB-Net: Hierarchical Adaptive Fusion Based on Multilevel Attention-enhanced Bottleneck Neural Network for Breast Histopathological Cancer Classification.” Signal Image and Video Processing, vol. 19, no. 5, Mar. 2025, https://doi.org/10.1007/s11760-025-04001-1.

[118] M. Iman, H. R. Arabnia, and K. Rasheed, “A review of deep transfer learning and recent advancements,” Technologies, vol. 11, no. 2, p. 40, Mar. 2023, doi: 10.3390/technologies11020040.

[119] A. H. Abdulaal et al., “Unsupervised Histopathological Sub-Image analysis for breast cancer diagnosis using variational autoencoders, clustering, and supervised learning,” Journal of Engineering and Sustainable Development, vol. 28, no. 6, pp. 729–744, Nov. 2024, doi: 10.31272/jeasd.28.6.6.

[120] M. Rahman, K. Deb, P. K. Dhar, and A. T. Shimamura, “ADBNET: An Attention-Guided Deep Broad convolutional neural Network for the classification of breast cancer histopathology images,” IEEE Access, vol. 12, pp. 133784–133809, Jan. 2024, doi: 10.1109/access.2024.3419004.

[121] D. Addo et al., “A hybrid lightweight breast cancer classification framework using the histopathological images,” Journal of Applied Biomedicine, vol. 44, no. 1, pp. 31–54, Dec. 2023, doi: 10.1016/j.bbe.2023.12.003.

[122] S. U. R. Khan, M. Zhao, S. Asif, X. Chen, and Y. Zhu, “GLNET: global–local CNN’s-based informed model for detection of breast cancer categories from histopathological slides,” The Journal of Supercomputing, vol. 80, no. 6, pp. 7316–7348, Nov. 2023, doi: 10.1007/s11227-023-05742-x.

[123] C. Xu, K. Yi, N. Jiang, X. Li, M. Zhong, and Y. Zhang, “MDFF-Net: A multi-dimensional feature fusion network for breast histopathology image classification,” Computers in Biology and Medicine, vol. 165, p. 107385, Aug. 2023, doi: 10.1016/j.compbiomed.2023.107385.

[124] M. A. Shamshiri, A. Krzyżak, M. Kowal, and J. Korbicz, “Compatible-domain Transfer Learning for Breast Cancer Classification with Limited Annotated Data,” Computers in Biology and Medicine, vol. 154, p. 106575, Jan. 2023, doi: 10.1016/j.compbiomed.2023.106575.

[125] Z. He et al., “Deconv-transformer (DecT): A histopathological image classification model for breast cancer based on color deconvolution and transformer architecture,” Information Sciences, vol. 608, pp. 1093–1112, Jun. 2022, doi: 10.1016/j.ins.2022.06.091.

[126] P. Bhowal, S. Sen, J. D. Velasquez, and R. Sarkar, “Fuzzy ensemble of deep learning models using choquet fuzzy integral, coalition game and information theory for breast cancer histology classification,” Expert Systems With Applications, vol. 190, p. 116167, Nov. 2021, doi: 10.1016/j.eswa.2021.116167.

[127] Y. Zhou, C. Zhang, and S. Gao, “Breast cancer classification from histopathological images using resolution adaptive network,” IEEE Access, vol. 10, pp. 35977–35991, Jan. 2022, doi: 10.1109/access.2022.3163822.

[128] K. Jabeen, M. A. Khan, M. A. Hameed, O. Alqahtani, M. T.-H. Alouane, and A. Masood, “A novel fusion framework of deep bottleneck residual convolutional neural network for breast cancer classification from mammogram images,” Frontiers in Oncology, vol. 14, Feb. 2024, doi: 10.3389/fonc.2024.1347856.

[129] I. Nissar, S. Alam, S. Masood, and M. Kashif, “MOB-CBAM: A dual-channel attention-based deep learning generalizable model for breast cancer molecular subtypes prediction using mammograms,” Computer Methods and Programs in Biomedicine, vol. 248, p. 108121, Mar. 2024, doi: 10.1016/j.cmpb.2024.108121.

[130] X. Xu et al., “GLA-Net: A global-local attention network for automatic cataract classification,” Journal of Biomedical Informatics, vol. 124, p. 103939, Nov. 2021, doi: 10.1016/j.jbi.2021.103939.

[131] J. Li, K. Wang, and X. Jiang, “Robust Multi-Subtype identification of breast cancer pathological images based on a Dual-Branch frequency domain fusion network,” Sensors, vol. 25, no. 1, p. 240, Jan. 2025, doi: 10.3390/s25010240.

[132] V. Sreelekshmi, K. Pavithran and J. J. Nair, "SwinCNN: An Integrated Swin Transformer and CNN for Improved Breast Cancer Grade Classification," in IEEE Access, vol. 12, pp. 68697-68710, 2024, doi: 10.1109/ACCESS.2024.3397667.

[133] A. Fiaz, B. Raza, M. Faheem, and A. Raza, “A deep fusion-based vision transformer for breast cancer classification,” Healthcare Technology Letters, vol. 11, no. 6, pp. 471–484, Oct. 2024, doi: 10.1049/htl2.12093.

[134] R. Maurya, N. N. Pandey, M. K. Dutta, and M. Karnati, “FCCS-Net: Breast cancer classification using Multi-Level fully Convolutional-Channel and spatial attention-based transfer learning approach,” Biomedical Signal Processing and Control, vol. 94, p. 106258, Mar. 2024, doi: 10.1016/j.bspc.2024.106258.

[135] L. Liu et al., “Collaborative Transfer Network for Multi-Classification of Breast Cancer Histopathological Images,” IEEE Journal of Biomedical and Health Informatics, vol. 28, no. 1, pp. 110–121, Jun. 2023, doi: 10.1109/jbhi.2023.3283042.

[136] L. Liu, W. Feng, C. Chen, M. Liu, Y. Qu, and J. Yang, “Classification of breast cancer histology images using MSMV-PFENet,” Scientific Reports, vol. 12, no. 1, Oct. 2022, doi: 10.1038/s41598-022-22358-y.

[137] S. Garg and P. Singh, “Transfer learning based lightweight ensemble model for imbalanced breast cancer classification,” IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol. 20, no. 2, pp. 1529–1539, May 2022, doi: 10.1109/tcbb.2022.3174091.

[138] S. Tripathi, S. K. Singh, and H. K. Lee, “An end-to-end breast tumour classification model using context-based patch modelling – A BiLSTM approach for image classification,” Computerized Medical Imaging and Graphics, vol. 87, p. 101838, Dec. 2020, doi: 10.1016/j.compmedimag.2020.101838.

[139] G. Li, C. Li, G. Wu, D. Ji, and H. Zhang, “Multi-View Attention-Guided multiple instance detection network for interpretable breast cancer histopathological image diagnosis,” IEEE Access, vol. 9, pp. 79671–79684, Jan. 2021, doi: 10.1109/access.2021.3084360.

[140] A. Ijaz et al., “Modality specific CBAM-VGGNET model for the classification of breast histopathology images via transfer learning,” IEEE Access, vol. 11, pp. 15750–15762, Jan. 2023, doi: 10.1109/access.2023.3245023.

[141] D. Wang, Z. Chen, and H. Zhao, “Prototype transfer generative adversarial network for unsupervised breast cancer histology image classification,” Biomedical Signal Processing and Control, vol. 68, p. 102713, May 2021, doi: 10.1016/j.bspc.2021.102713.

[142] X. Y. Liew, N. Hameed, and J. Clos, “An investigation of XGBoost-based algorithm for breast cancer classification,” Machine Learning With Applications, vol. 6, p. 100154, Sep. 2021, doi: 10.1016/j.mlwa.2021.100154.

[143] C. Wild, World Cancer Report 2014, C. P. Wild and B. W. Stewart, Eds. Geneva, Switzerland: World Health Organization, 2014.

[144] S. R. Lakhani, I. O. Ellis, S. J. Schnitt, P. H. Tan, and M. J. Van De Vijver, WHO classification of tumours of the breast. 2012. [Online]. Available: http://ci.nii.ac.jp/ncid/BB0963818X

[145] N. M. U. Din, R. A. Dar, M. Rasool, and A. Assad, “Breast cancer detection using deep learning: Datasets, methods, and challenges ahead,” Computers in Biology and Medicine, vol. 149, p. 106073, Aug. 2022, doi: 10.1016/j.compbiomed.2022.106073.

[146] S. Shafi and A. V. Parwani, “Artificial intelligence in diagnostic pathology,” Diagnostic Pathology, vol. 18, no. 1, Oct. 2023, doi: 10.1186/s13000-023-01375-z.

[147] A. Das, M. S. Nair, and S. D. Peter, “Computer-Aided Histopathological Image Analysis Techniques for Automated Nuclear Atypia Scoring of Breast Cancer: a Review,” Journal of Digital Imaging, vol. 33, no. 5, pp. 1091–1121, Jan. 2020, doi: 10.1007/s10278-019-00295-z.

[148] S. Dabeer, M. M. Khan, and S. Islam, “Cancer diagnosis in histopathological image: CNN based approach,” Informatics in Medicine Unlocked, vol. 16, p. 100231, Jan. 2019, doi: 10.1016/j.imu.2019.100231.

[149] M. Kowal, P. Filipczuk, A. Obuchowicz, J. Korbicz, and R. Monczak, “Computer-aided diagnosis of breast cancer based on fine needle biopsy microscopic images,” Computers in Biology and Medicine, vol. 43, no. 10, pp. 1563–1572, Aug. 2013, doi: 10.1016/j.compbiomed.2013.08.003.

[150] N. K.-M. Lee and W. N. Street, “An adaptive resource-allocating network for automated detection, segmentation, and classification of breast cancer nuclei topic area: image processing and recognition,” IEEE Transactions on Neural Networks, vol. 14, no. 3, pp. 680–687, May 2003, doi: 10.1109/tnn.2003.810615.

[151] M. M. Jadoon, Q. Zhang, I. U. Haq, S. Butt, and A. Jadoon, “Three-Class mammogram classification based on descriptive CNN features,” BioMed Research International, vol. 2017, pp. 1–11, Jan. 2017, doi: 10.1155/2017/3640901.

[152] D. M. Vo, N.-Q. Nguyen, and S.-W. Lee, “Classification of breast cancer histology images using incremental boosting convolution networks,” Information Sciences, vol. 482, pp. 123–138, Jan. 2019, doi: 10.1016/j.ins.2018.12.089.

[153] T. Ching et al., “Opportunities and obstacles for deep learning in biology and medicine,” Journal of the Royal Society Interface, vol. 15, no. 141, p. 20170387, Apr. 2018, doi: 10.1098/rsif.2017.0387.

[154] S. Robertson, H. Azizpour, K. Smith, and J. Hartman, “Digital image analysis in breast pathology—from image processing techniques to artificial intelligence,” Translational Research, vol. 194, pp. 19–35, Nov. 2017, doi: 10.1016/j.trsl.2017.10.010.

[155] A. Rakhlin, A. Shvets, V. Iglovikov, and A. A. Kalinin, “Deep Convolutional Neural Networks for Breast Cancer Histology Image Analysis,” in Image Analysis and Recognition, A. Campilho, F. Karray, and B. ter Haar Romeny, Eds. Cham: Springer, 2018, vol. 10882, pp. 900–908, doi: 10.1007/978-3-319-93000-8_83.

[156] Gül, Mehmet. “A Novel Local Binary Patterns-based Approach and Proposed CNN Model to Diagnose Breast Cancer by Analyzing Histopathology Images.” IEEE Access, Jan. 2025, p. 1. https://doi.org/10.1109/access.2025.3545052.

[157] Jackson, Jehoiada, et al. “A Patch-based Deep Learning Framework With 5-B Network for Breast Cancer Multi-classification Using Histopathological Images.” Engineering Applications of Artificial Intelligence, vol. 148, Mar. 2025, p. 110439. https://doi.org/10.1016/j.engappai.2025.110439.

[158] Balasubramanian, Aadhi Aadhavan, et al. “Ensemble Deep Learning-Based Image Classification for Breast Cancer Subtype and Invasiveness Diagnosis From Whole Slide

Image Histopathology.” Cancers, vol. 16, no. 12, June 2024, p. 2222. https://doi.org/10.3390/cancers16122222.

[159] Mewada, Hiren. “Extended Deep-Learning Network for Histopathological Image-Based Multiclass Breast Cancer Classification Using Residual Features.” Symmetry, vol. 16, no. 5, Apr. 2024, p. 507. https://doi.org/10.3390/sym16050507.

[160] Maleki, Alireza, et al. “Breast Cancer Diagnosis From Histopathology Images Using Deep Neural Network and XGBoost.” Biomedical Signal Processing and Control, vol. 86, June 2023, p. 105152. https://doi.org/10.1016/j.bspc.2023.105152.

[161] Mondol, Raktim Kumar, et al. “hist2RNA: An Efficient Deep Learning Architecture to Predict Gene Expression From Breast Cancer Histopathology Images.” Cancers, vol. 15, no. 9, Apr. 2023, p. 2569. https://doi.org/10.3390/cancers15092569.

[162] Hu, Jie, et al. “Squeeze-and-Excitation Networks.” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 42, no. 8, May 2019, pp. 2011–23. https://doi.org/10.1109/tpami.2019.2913372.

[163] Woo, Sanghyun, et al. “CBAM: Convolutional Block Attention Module.” Lecture notes in computer science, 2018, pp. 3–19. https://doi.org/10.1007/978-3-030-01234-2_1.

[164] Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, et al. “Attention Is All You Need.” arXiv (Cornell University), vol. 30, June 2017, pp. 5998–6008. arxiv.org/pdf/1706.03762v5.

[165] C. Wang, F. Xiao, W. Zhang, S. Huang, W. Zhang and P. Zou, "Transfer Learning and Attention Mechanism for Breast Cancer Classification," 2021 17th International Conference on Computational Intelligence and Security (CIS), Chengdu, China, 2021, pp. 75-79, doi: 10.1109/CIS54983.2021.00024.

[166] R. Karthik, R. Menaka, and M. V. Siddharth, “Classification of breast cancer from histopathology images using an ensemble of deep multiscale networks,” Journal of Applied Biomedicine, vol. 42, no. 3, pp. 963–976, Jul. 2022, doi: 10.1016/j.bbe.2022.07.006.

[167] Liu, Yujian, et al. “Adaptive Threshold Learning in Frequency Domain for Classification of Breast Cancer Histopathological Images.” International Journal of Intelligent Systems, vol. 2024, Mar. 2024, pp. 1–13. https://doi.org/10.1155/2024/9199410.