Artificial Intelligence-based Computational Pathology (AI-CPath) Platform for Lung Cancer

Ingkar Chegedekova, Master

**Submitted in fulfillment on the requirements
for the degree of Master of Science
in Electrical and Computer Engineering**

**School of Engineering and Digital Science
Department of Electrical and Computer Engineering
Nazarbayev University**

53 Kabanbay Batyr Avenue,
Astana, Kazakhstan, 010000

Lead Supervisor: Professor Prashant Jamwal
Co-supervisor: Professor Muhammad Akhtar

Date of submission: 11.04.2025

DECLARATION

I hereby declare that this manuscript, entitled “Artificial Intelligence-based Computational Pathology (AI-CPath) Platform for Lung Cancer”, is the result of my own work except for quotations and citations which have been duly acknowledged.

To the best of my knowledge and belief, I affirm that this manuscript has neither been previously submitted nor concurrently submitted, either wholly or partially, for any degree or diploma at Nazarbayev University or any other academic institution nationally or internationally.

Name:
Date:

Abstract

Lung cancer stands as a malignant tumor which demonstrates both the highest frequency of occurrence and the highest death rate worldwide. Delayed diagnosis and scarce early detection strategies lead to this high incidence so researchers demand new diagnostic tools that would improve effectiveness and efficiency. Medical experts utilize histopathology as the current diagnostic standard for lung cancer yet slide examination by human pathologists takes considerable time along with showcasing inconsistent evaluation methods and demanding specialized training. Artificial intelligence (AI) combined with depth learning technology now provides a promising approach to improve diagnostic accuracy and decrease analyst workloads throughout pathology analysis.

A digital pathology platform utilizes AI techniques for automating the classification procedure of lung cancer through histopathology image analysis. This platform utilizes Convolutional Neural Networks (CNNs) to distinguish between two NSCLC subtype groups which include adenocarcinoma and squamous cell carcinoma. The work investigates different CNN network designs to achieve maximum feature detection and classification results. One major obstacle in AI pathology emerges from insufficient high-quality annotated data that hampers generalization capabilities of models. The project addressed its issues by implementing methods to overcome the challenge of limited annotated medical data, contrastive learning (SimCLR) was integrated into the pipeline, enabling robust feature representation learning from unlabelled histopathology images. These self-supervised representations were then fine-tuned using a lightweight supervised classifier, demonstrating strong classification performance.

This research delivers a reliable, accurate, and efficient AI-based classification tool that

addresses the constraints of limited annotated data through contrastive learning, while also demonstrating the viability of both custom and pre-trained CNN models in digital pathology workflows for lung cancer diagnosis. Additionally, the careful composition of data augmentations was found to play a critical role in defining effective predictive tasks, especially within the contrastive learning framework, enabling the model to learn robust and transferable representations from unlabelled histopathological images.

Table of Contents

List of Abbreviations

CNN	Convolutional Neural Network
NSCLC	Non-small cell lung cancer
SCLC	Small cell lung cancer
DL	Deep Learning
CAD	Computer-Aided Detection
AC	Adenocarcinoma
WSI	Whole-slide images
CT	Computed tomography
CXR	X-rays
AI	Artificial Intelligence
AC	Adenocarcinoma
ReLU	Rectified Linear Unit
AUC	Area Under the Receiver Operating Characteristic Curve
FCN	Fully Convolutional Network
LeakyReLU	Leaky Rectified Linear Unit
SCC	Squamous Cell Carcinoma
SSL	Self-Supervised Learning
SimCLR	Simple Framework for Contrastive Learning of Visual Representations
AvgPool	Average Pooling
VGG	Visual Geometry Group
ResNet	Residual Network
GPU	Graphics Processing Unit

List of Tables

List of Figures

CHAPTER 1 – Introduction

1.1. Importance of Early Lung Cancer Diagnosis

Lung cancer maintains its position as the world's foremost cancer-related cause of death since it generates substantial numbers of annual cancer deaths worldwide. Lung cancer has a high mortality level because people typically do not discover it until advanced stages when treatment options become less effective and survival becomes less probable. Several risk elements such as smoking and carcinogenic exposure and air pollution and radiation and genetic heredity and chronic infections elevate lung cancer development possibilities [1]. The ongoing advances in oncology combined with public health prevention measures fail to address the difficulty of diagnosing lung cancer at initial stage along with distinguishing among various cancer types.

Statistical research demonstrates how urgent it is to discover lung cancer at its early stages. According to current statistics lung cancer survival exceeds only 16.6% which places this disease among the most fatal type of cancer. The posibility of survival among patients depends significantly on what stage their disease reached during diagnosis. Stage I detection leads to close to 70% five-year survival but this number declines to 5% when diagnosis happens at stage IV [2]. The survival statistics indicate that swift and precise detection enhances treatment outcomes and patient recovery along with boosting survival rates. Early detection enables medical teams to use surgical resection and targeted therapies for treatment because these methods achieve better control of disease evolution compared to advanced stage interventions.

Standard lung cancer diagnosis requires histopathology analysis which comprises tissue examination of samples collected from bronchoscopy biopsies and fine-needle aspirations and surgical excisions. The successful diagnosis of lung cancer through histopathology encounters multiple difficulties. The process of examining slides manually requires extensive time that leads

to diagnostic delays which delays the start of treatment. Pathological assessments demonstrate a known subjectivity that results in classification and staging inconsistencies among specialists when performing assessments.

Patients need histopathology analysis for lung cancer diagnosis because experts must examine tissue collected from biopsies such as bronchoscopy and fine needle aspiration together with tissue removal during surgery. Multiple essential difficulties exist in histopathology identification procedures even though it produces exact results. Manual slide analysis of cancer tests requires extensive time requirements to successfully identify cancer cells thus causing analysis and treatment delays. Many experts document the problems linked to inter-observer variability because subjective pathology assessments often cause inconsistent staging across specialists [3].

The differential diagnosis between adenocarcinoma and squamous cell carcinoma becomes more difficult for pathologists because the two cancer subtypes possess similar identifiable characteristics. Manual diagnosis of lung cancer through histopathological assessment faces growing limitations because it slows down the diagnosis process for a rising worldwide number of patients [4].

Lung cancer diagnostics need efficient and accurate diagnostic solutions which simplify testing procedures and decrease human failure rates due to current limitations. Pathology will achieve improved lung cancer diagnostic capabilities through the implementation of advanced technology systems. The research in the field develops new pathology approaches that should work together with current histopathology techniques to enhance subtype classification accuracy for personalized treatments.

1.2. Challenges in Traditional Histopathology-Based Diagnosis

Despite the fact that the methods like bronchoscopy and percutaneous core needle biopsy give the most important tissue samples for lung cancer diagnosis, they have their limitations [5].

Diagnostic errors are not unusual even among expert pathologists, particularly in the late stages of the disease. Studies have shown that differences in staging of lung cancer among pathologists are not rare, and they still cannot detect early-stage pulmonary neoplasms even by using histological markers during advanced stages [6]. These methods of diagnosis guarantee precision in therapeutics as well as provide the physician with the necessary clinical data for the decision-making process.

Diagnostic inconsistency stands as a significant problem in the current medical practice. Professional pathologists obtain varying interpretations about cancer staging results because tumor heterogeneity combined with overlapping histological features makes advanced case assessments complex [6]. Early-stage pulmonary neoplasms sometimes remain hidden because their morphological traits are difficult for pathologists to detect despite the presence of histological indicators. More objective and reproducible procedures for diagnosis become essential because of these existing challenges.

Histopathology evaluation suffers from manual processing that leads to an extensive time requirement in its diagnostic procedures. The analysis of individual biopsy slides under a microscope demands specialized knowledge which results in extended delays before starting patient treatment. The rapid growth of biopsy sample volumes combined with digital pathology scans creates an urgent need for faster diagnostic solutions that are easy to scale.

The problems faced by histopathology diagnostics can be managed through the implementation of artificial intelligence and machine learning-based CAD systems. Deep learning algorithms help these systems to improve diagnostics by assisting pathologists through faster and more accurate detection of lung cancer while reducing their workload. The majority of artificial intelligence approaches have focused on classification instead of segmentation which provides better precision and interpretability for lung cancer histopathology analysis.

1.3. AI-Powered Image Analysis for Lung Cancer Diagnosis

Artificial intelligence stands today as a powerful technology which successfully improves lung cancer diagnostic procedures specifically within histopathology image analysis operation. The deployment of Artificial Intelligence models proves beneficial for disease progression detection together with tumor morphological disorder identification and better classification outcomes [7]. Machine learning along with deep learning allows AI to process large datasets with exceptional accuracy and reliability to help pathologists make rapid and uniform diagnosis decisions.

CNNs represent among the most prominent artificial intelligence techniques for medical imaging because these models effectively identify intricate patterns in histopathological images [8]. The field of oncology employs reinforcement learning (RL) [9] and recurrent neural networks (RNNs) [10] together with another ML methods for predictive analysis. CNN-based deep learning models lead as the primary selection in medical practice because they outperform other methods in image recognition along with feature extraction.

Deep learning algorithms have established themselves in radiology-based lung cancer identification through their analysis of high-resolution CT scans and X-ray images to detect abnormal lesions [11]. The systems produce medical image abstract representations which enable better identification of diseases together with enhanced risk evaluation capabilities. The application of AI technologies for improving histopathologic diagnostics remains a current focus for research development even after radiology diagnostics have significantly benefited from these advancements.

AI-based histopathology research primarily concentrated on assigning overall image labels to samples (cancerous vs non-cancerous) through its classification approach. The valuable information obtained from classification methods does not provide enough detailed spatial

information required to precisely locate tumors and identify subtypes correctly. The recent interest surrounding segmentation-based AI models results from their ability to classify tissue pixels which produces detailed interpretable tissue analysis.

1.4. Research Objectives and Significance

The research focuses on developing and testing an artificial intelligence-based detection method for lung cancer in histopathological images by classification for higher diagnostic precision. Lung cancer represents one of the main causes of mortality from cancer worldwide because patients receive late diagnosis as well as physicians encounter difficulties in distinguishing cancer subtypes. Routine histopathological evaluation by manual examination produces slow results along with inconsistent evaluation between observers and results in delayed diagnosis and unstandardized classification outcomes. Medical diagnostic technology requires AI computational pathology to meet increasing needs for improved competency and accuracy in treating oncological conditions.

Efforts to overcome these challenges in this study involve exploring the use of deep learning methods especially Convolutional Neural Networks (CNNs) to automatic subtype classification. The work presents a two-week approach: first, the researchers used the development of several custom CNNs with varying degrees of complexity in order to identify the best hyper-parameter trade-off between accuracy and resource efficiency, respectively and second, they employed the contrastive learning to improve representation learning of unlabeled histopathological data as a self-supervised pre-training strategy. This is particularly beneficial in the medical domain where the acquisition of high quality annotated dataset is both expensive and laborious.

The study also includes transfer learning in examining common pre-trained models including VGG19, EfficientNet, and ResNet50 to compare performance to custom-designed networks. By doing this it not only shows that it is viable to train performant models on rather

small datasets but also it also shows the complementary strengths of self-supervised and transfer learning approaches in digital pathology.

This work seeks to offer strong and interpretable AI-enabled diagnostic tools. capable of supporting clinical decision making, reducing diagnostic variability and promoting, early detection potential for lung cancer with the help of a classification driven framework..

CHAPTER 2 – Literature review

2.1. AI in lung cancer diagnosis and management

Lung cancer currently stands as one of the most cancer-related cause of death worldwide since it contributes to substantial mortality numbers in global cancer statistics. Lung cancer exists in two primary forms that include non-small cell lung cancer (NSCLC) as well as small cell lung cancer (SCLC) which represent 85% and 15% of total cases [12]. The successful treatment of early-stage lung cancer is difficult because most patients receive their diagnosis after their condition has progressed to advanced stages which substantially lowers their survival rate and reduces treatment possibilities.

Doctors usually employ CT scans together with CXR imaging and histopathological tests to detect lung cancer. The evaluation techniques require time-consuming manual assessment together with diagnostic delays and variable diagnostic outcomes between observers [13]. Although histopathology serves as the definitive method to identify lung cancer subtypes it depends on pathologist expertise which creates inconsistencies during the classification process and staging procedures. The utilization of LDCT for early diagnosis advancement remains challenging by false positive results and imaging artifact presence along with interpretation differences [13].

The combination of artificial intelligence technology and digital pathology systems helps enhance automated diagnostic systems which lead to better lung cancer identification rates. Using AI-based decision-support systems that include multi-parametric clustering models

enables medical experts to prevent errors by identifying uncommon patterns which traditional screening systems miss [14]. CNNs as part of deep learning algorithms deliver exceptional results by providing advanced tumor assessment for histopathological images.

The advancement of artificial intelligence surpasses diagnostic testing because it succeeds in determining tumor behaviors and creating personalized treatment approaches. Healthcare professionals leverage machine learning tumor structure assessments to acquire better clinical data that improves their decisions about tumor development and treatment responses. The utilization of AI-based models creates automated segmenting functions that improve both tumor segmentation processes and histopathological image clarity. The incorporation of AI technology into oncology medical practitioners enables them to direct their attention to patient care while obtaining improved therapeutic strategies from AI-generated insights.

The initial part of this chapter examines standard histopathological diagnostic procedures for lung cancer while moving toward explaining AI-powered training methodologies for disease classification and precise tissue segmentation. This section analyzes how AI supports medical professionals in predicting disease outcomes while examining obstacles that emerge when implementing AI solutions for healthcare delivery.

2.2. Convolutional Neural Networks in Medical Image Analysis

The Convolutional Neural Network (CNN) represents an advanced deep learning network that experts use commonly to process images and recognize their contents along with standard classifications. The two-dimensional spatial structure of images becomes an explicit feature in CNNs compared to MLPs because this enhancement enables higher accuracy in pattern recognition for vision-based applications [15].

CNN architectures depend mainly on their structural elements comprising convolutional and pooling layers and activation functions and fully connected layers to extract stable features

for accurate classification purposes. The core component of CNN systems are convolutional layers that apply numerous filters to images through mathematical operations of convolution [16]. Weights form trainable matrix elements that exist in each kernel which identifies specific visual elements such as edges and textures and shapes. The computational process of convolution leads to generating feature maps from input images through kernel sliding operations and subsequent point-by-point calculation and summation steps which identify important data patterns. The subsequent pooling layers serve to reduce feature map spatial dimensions while advancing important visual patterns to cut down computational requirements. The two principal pooling operations for localized regions are MaxPooling which identifies maximum values from local features and AveragePooling which calculates average values from local features [17]. Through these operations the CNN develops stronger resistance against modifications including translation and scaling and rotation of images which results in better generalization across different datasets. The fully connected (dense) layers begin their evaluation of high-level features which convolution and pooling layers presented in earlier stages. The dense layers use the extracted features to create output predictions through mappings that assign each prediction neuron to a different image class or category. Location-specific images map their raw outputs at the last layer through softmax activation to generate probabilities which define the probability of image-match to specific classes. The prime objective of CNN training involves maximizing accuracy in these probability predictions by means of repeated weight adjustments and training cycles. Current research in computer vision showcases CNNs together with their adaptations as leaders in object detection and segmentation while also excelling at classification tasks particularly through medical imaging applications which foster diagnostic and decision processes today.

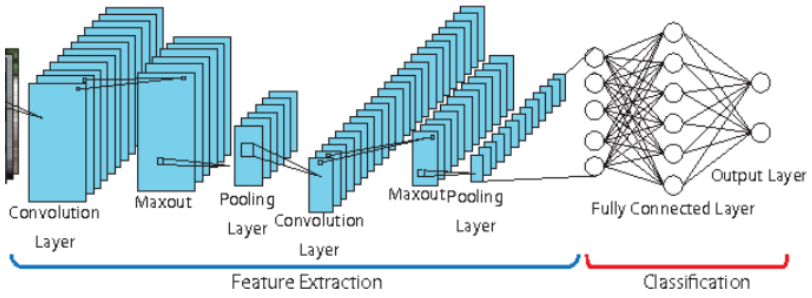


Figure 1: Architecture of a Convolutional Neural Network: Feature Extraction and Classification Layers

2.2.1. CNN Architectures in Medical Image Analysis

Medical image analysis technology advances primarily from the development and implementation of strong convolutional neural network (CNN) structures. Different established versions of CNN demonstrate exceptional performance capabilities and operating efficiency while providing flexibility for various medical applications that both classification and segmenttaion.

A distinctive aspect of VGGNet emerges when its designers choose uniform structural simplicity through the use of small 3x3 convolutional filters over its entire architecture [18]. With its basic design VGGNet reaches deep layers up to 19 which makes it successful in drawing detailed hierarchical image representations from medical data effectively. VGGNet continues to impact medical diagnostics through feature extraction despite being resource-intensive because of its numerous parameters.

The Residual Networks (ResNet) include ResNet50 which employs residual connections to resolve the gradient vanishing problem present in deep CNN structures. ResNet50 achieves gradient propagation efficiency through its inclusion of skip connection features known as residual connections. The new design enables deeper networks to operate through stable training processes that keep accuracy levels high. The medical imaging classification field heavily uses ResNet50 because its 50 layers capture detailed pathological characteristics that produce

excellent performance results when processing complex histopathological images [19].

A newer version of the GoogLeNet inception family network models named InceptionV4 builds upon previous models through its innovative design of inception modules which perform parallel operations on various scale blocks [20]. The architecture enables efficient achievement of major features at diverse scales with unchanged computational needs. Research investigators improved the algorithms of previous work and added many inception modules and training optimization elements to InceptionV4. The system is efficient in medical image analysis, adopting varied pattern scales with contextual changes to achieve strong diagnosis outputs.

The MobileNet gives users unprecedented computational speed and efficiency which enables its deployment on limited-resource devices for embedded systems and mobile solutions [21]. Due to its depth-wise separable convolutions and NAS architectural search methods it accomplishes efficient performance improvements. MobileNetV3 delivers lightning-fast inference operations with preserved accuracy levels that make it suitable for time-sensitive medical imaging diagnostics needs.

Finally, the compound scaling method EfficientNet optimizes all components including depth and width and resolution in parallel to strike a balance between model complexity and computational resources [22]. The EfficientNet architecture delivers exceptional recognition function in various tasks with minimal computational requirements. The efficient scaling formula of EfficientNet delivers excellent results for medical imaging due to its ability to achieve high accuracy while using limited resources effectively which makes it ideal for medical diagnostics tasks particularly cancer classification and segmentation.

These architectural frameworks span various computational requirements which offer researchers and clinical practitioners flexible diagnostic tools for multiple clinical applications.

2.3. Cytopathological diagnosis of AI

The development of AI-augmented cytopathology has been successful recently.Atsushi Teramoto [23] created a deep convolutional neural network (DCNN) to distinguish different types of lung cancer from a series of microscopic images, including adenocarcinoma, squamous cell carcinoma, and small cell carcinoma. The DCNN could achieve a classification accuracy over 71%,which is very much the same grade as a trained cytotechnologist or pathologist, demonstrating the fact that this is a robust diagnostic technology in cytopathology. Additionally, Taehee Kim [24] carried out a retrospective, multi-center study using the deep learning model transferred to respiratory tract cytological images. The DenseNet121 model operating in this study reflected the best diagnostic performance with the accuracy exceeding the experienced pathologist's considerable by 0.355 and increasing inter-observation agreement from 0.553 to 0.908 [24]. This demonstrates the function of deep learning models in improving the diagnostic outcome of cytopathology and its benefit to pathologists in being able to provide more reliable decisions.

2.4. AI in Histopathological Image Analysis

AI serves as one of the disciplines which has substantially enhanced histopathological diagnostic capabilities. Histopathological analysis through manual methods faces major drawbacks because it depends on inconsistent expert readings and takes long analysis times and creates challenges when diagnosing tumors that present similar morphological patterns. There exists evidence showing pathologists with extensive experience still make errors while classifying tumors especially when differentiating between adenocarcinoma and squamous cell carcinoma because these cancers share comparable histological features . The diagnostic process becomes more time-consuming and expensive through the requirement for additional molecular marker tests including EGFR, ALK, and PD-L1 markers. Computational pathology driven by AI presents itself as an effective way to optimize diagnostic processes while minimizing human

interpretation inconsistencies. The deep learning technology consisting of convolutional neural networks (CNNs) analyzes whole-slide images (WSIs) to detect tumor-specific patterns with an impressive accuracy rate. Research by Coudray et al. [25] showed the Inception v3 deep learning model reached an AUC measurement of 0.97 for detecting between adenocarcinoma, squamous cell carcinoma and normal lung tissue at levels equal to expert clinical pathologists. Research by Ahmed et al. [26] demonstrated XGBoost reached 91% precision when identifying SCLC and NSCLC while showing superior performance than SVM and KNN traditional machine learning approaches. The study demonstrates that AI systems might establish standardized lung cancer diagnostic categories and lower diagnostic mistakes and simultaneously speed up pathology examination processes.

Deep learning models serve as tools to handle cases of challenging pathology classifications which occur rarely. Research by JanBen et al. [27] established that utilizing large annotated datasets for AI training enhances classification reliability by minimizing variations between pathologists which indicates AI decision support systems are beneficial for histopathology applications. AI technologies focus on segmentation work alongside classification in histopathology which enables researchers to detect tumor boundaries with precision and spot unusual regions in WSIs. Research conducted by Zhang Li [28] proved that multiple deep learning models create better segmentations which results in rapid more precise medical diagnosis outcomes.

AI in pathology evolves beyond diagnostic tasks for predicting tumors behavior as well as determining patient outcomes and treatment reactions. Researchers Yongjun Wu et al. [29] have developed artificial intelligence models to assess lung cancer biomarkers by using artificial neural networks (ANNs) as they outperform traditional statistical analysis techniques. Research conducted by Teramoto et al. [23] as well as Khosravi et al. [30] demonstrates how CNNs effectively predict tumor progression while distinguishing benign and malignant cases

accurately for improved lung cancer prognosis and individualized treatments.

2.5. Contrastive Learning in Medical Imaging

Modern research demonstrates that contrastive learning operates as an efficient self-supervised technique which generates versatile features through unlabelled data [31]. Getting large amounts of labeled data in medical imaging is often expensive and difficult, mainly because it requires a lot of time and effort from trained medical experts. Contrastive learning helps get around this issue by learning useful features from the natural patterns and relationships found in unlabeled images.

Today's healthcare depends heavily on medical imaging technologies because they provide essential diagnostic tools as well as crucial therapeutic planning features along with patient care applications. Medical imaging maintains essential status in contemporary healthcare because of its speedy development and growing dependency on this essential technology. Medical imaging uses image classification as its primary application to detect diseases while also designing treatments and monitoring patient health dynamics. The growing requirement for precise analytical methods and efficient classification techniques becomes challenging because medical image classification lacks sufficient labeled datasets. The growing challenges in the field led to increased attention on self-supervised learning methods (SSL) by researchers. SSL techniques make use of large volumes of unlabeled medical images to automatically generate useful features and representations through which medical organizations reduce their requirement for annotated data [31].

2.5.1. Fundamentals of Contrastive Learning

Through self-supervised learning methods contrastive models acquire meaningful data patterns directly from unmarked datasets. Models learn to identify distinct data points depending on their interrelationships through this teaching method. This is especially useful in areas like medical imaging, where labeled data can be hard to get or expensive to produce.

The process starts by creating multiple slightly different versions of the same image using techniques like cropping, rotating, color changes, or resizing. Even though these versions look different, they still represent the same original image and are treated as “positive pairs”. These are then passed through a neural network separately, and the goal is to make their outputs as close to each other as possible. At the same time, images that come from different sources “negative pairs” are pushed apart in the embedding space. This way, the model learns to focus on the most important and stable features of the images, which makes it better at recognizing patterns in new, unseen data.

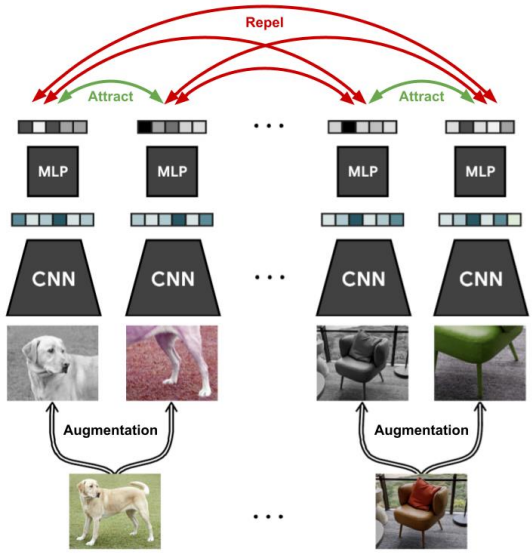


Figure 2: Contrastive Learning with Positive and Negative Pairs

As shown in Figure 2, contrastive learning works by bringing similar images closer together in the feature space, while pushing dissimilar ones apart. For example, when two augmented views of the same image are passed through the model (shown as green arrows), they end up with similar representations. In contrast, unrelated images (red arrows) are mapped farther apart. This setup allows the model to learn useful visual features without requiring labeled data.

To achieve this, each input image x is transformed into two different augmented versions, $\tilde{x}_i$ and $\tilde{x}_j$, using a set of random augmentation techniques. These are then processed by a shared encoder network, typically a convolutional neural network (CNN), which produces intermediate

feature vectors. External images passing through the shared encoder neural network $f(\cdot)$, built using CNN technology create intermediate represents h_i and h_j.

A nonlinear projection head $g(\cdot)$ transforms the features for contrastive loss $g(\cdot)$ is applied. The projection function expressed as multilayer perceptron (MLP) transforms intermediate representations into final embeddings z_i and z_j defined as:

$$\begin{aligned} h_i &= f(\tilde{x}_i), \quad z_i = g(h_i), \\ h_j &= f(\tilde{x}_j), \quad z_j = g(h_j). \end{aligned} \tag{2.1}$$

2.5.2. Contrastive Learning Architectures and Loss Functions

The performance of self-supervised learning has improved through multiple introduced contrastive learning frameworks which specialize in visual representation tasks. SimCLR (Simple Framework for Contrastive Learning) [32], MoCo (Momentum Contrast) [33], and SimSiam [34] while being among the most prominent structures in this domain..

SimCLR, in particular, stands out for its simplicity and effectiveness. By combining strong image augmentations with contrastive learning principles the teaching process helps models develop useful features without needing labeled dataset information. The method applies random augmentation techniques to individual input images which produces two related versions of the same content. These two views are then passed through an encoder typically a ResNet model to produce high-dimensional feature representations [32]. The goal of this training is to bring the representations of positive pairs closer together in the latent space, while simultaneously pushing apart representations of negative pairs, i.e., views from different images. This process encourages the model to learn discriminative features that are invariant to augmentations and can later be transferred to downstream classification tasks.

Expanding on the earlier description of SimCLR, Figure 3 offers a clear visual representation of how the contrastive learning process unfolds. It shows how two differently augmented versions of the same image are processed through the same encoder and projection head, resulting in

feature embeddings. The model is then trained to increase the similarity between these paired representations, which come from the same original image.

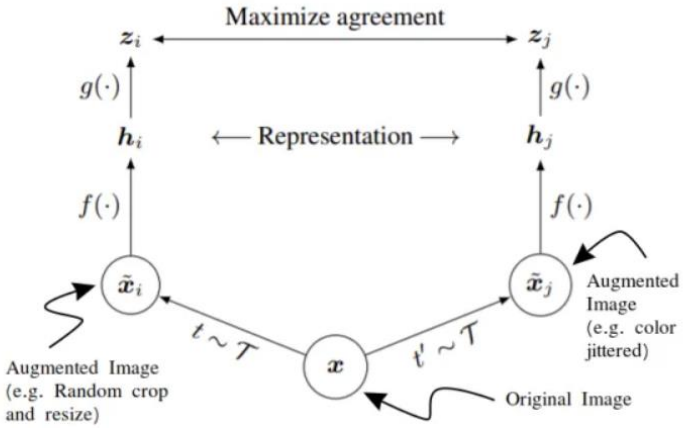


Figure 3: A Simple Framework for Contrastive Learning (SimCLR) [32].

To formalize this training strategy, SimCLR employs a loss function known as the normalized temperature-scaled cross-entropy loss (NT-Xent), which encourages alignment of similar image embeddings [35] :

$$L_{i,j} = -\log \frac{exp\left(sim\left(z_i, z_j\right)/\tau\right)}{\sum_{k=1}^{2N} 1_{[k \neq i]} exp(sim(z_i, z_k)/\tau)}$$

,where $sim\left(z_i, z_j\right)$ represent the projected embeddings of a positive pair, τ is a temperature scaling parameter , and the denominator includes all negative examples in the batch. This formulation ensures that the model focuses on distinguishing between similar and dissimilar instances efficiently during training.

The MoCo system creates momentum encoding along with a mutable memory queue to achieve its objectives. The MoCo framework avoids using large batch sizes to produce diverse negative pairs by keeping an automatically updated storage of encoded image features. The queue mechanism lets the model retrieve a wide range of diverse negative samples which dramatically decreases memory consumption and computational needs. The essential feature of MoCo involves the momentum encoder parameter update mechanism where the query encoder

weights serve to define this process through the following equation:

$$\theta_k \leftarrow m\,\theta_k + (1-m)\theta_q \tag{2.3}$$

,where parameters θ_k and θ_q control both the momentum and query encoder whereas m represents the coefficient for momentum updates [33].

SimSiam implements contrastive learning through a simple strategy because it eliminates the requirement for negative samples. The network architecture features two branches that operate on differently augmented versions of the same image following the Siamese-style design. A prediction training occurs between one network branch while the other network branch generates the target output [34]. The configuration of SimSiam enables high performance alongside reduced computational needs because the model does not need to compare images from unrelated data. The training process remains stable through the implementation of the stop-gradient technique which prevents unproductive or trivial representation learning.

Each contrastive learning method presented in SimCLR, MoCo, and SimSiam addresses different demanding problems in the field through enhancing representation quality, optimizing training efficiency, and establishing learning stability. Their advantageous characteristics enable effective use in medical imaging applications because these settings typically face restricted labeled data while demanding efficient feature extraction.

2.5.3. Contrastive Learning Applications in Histopathology

The analysis of histopathology images proves to be best performed with contrastive learning techniques due to their ability to extract generalized features from unlabeled data. Technological experts face major challenges when trying to acquire labeled datasets because acquiring these datasets takes a long time and expert labor. The practical method of contrastive learning enables the discovery of meaningful patterns through unlabeled images to eliminate the requirement of time-consuming manual annotation.

Research has proved that contrastive learning achieves successful results while classifying tissue types and identifying cancer subtypes in conjunction with detecting rare pathological conditions. Models trained by this self-supervised approach achieve better results in processing usual variations found in histopathological images. The discrepancies stem from differences between staining protocols and methodology of preparing slides or imaging devices [36]. Contrastive learning enables neural networks to discover essential biological and clinical information by allowing them to learn from multiple augmented versions of each image thus preventing them from assessing superficial variations that affect supervised models.

A research study used contrastive learning methods for lung cancer histopathology to differentiate between adenocarcinoma and squamous cell carcinoma together with benign tissue types. The self-supervised contrastive method generated results demonstrating successful subtype identification of over 89 % accuracy for these subtypes beyond traditional unsupervised and semi-supervised evaluation methods. The model achieved improved accuracy because it learned high-quality features which enhanced its capability to detect the crucial distinctions within similar groups of cancer [37]. The research shows that contrastive learning methods handle medical data labeling inconsistencies with high effectiveness.

It is possible to call out the novelty of the strong potential of the contrastive learning technique because it assists in elevating the subsequent supervised tasks like the tumor segmentation and classification, as well as make it gives the pre-training. The use of contrastive techniques for training non labeled data to breast and lung cancer histopathology experiments yielded notably better results.The technique delivers superior results after performing fine-tuning operations on datasets with limited label information. Performance improvements exceeded 28 % when measuring F1-score above pre-trained models that used ImageNet [38] as their standard dataset. The pre-training process using contrastive methods enables models to share important information across datasets and thus decreases the demand for human-generated annotations.

Contrastive learning has achieved notable success through its applications in different pathology fields. The research on colorectal cancer histopathology demonstrated how SimCLR-trained models provided outstanding generalization abilities. The research models achieved superior performance than standard supervised approaches across various external testing datasets that incorporated various staining methods and imaging systems [39]. Real clinical settings benefit from contrastive learning because the models demonstrate consistent performance despite the standardization differences commonly observed in slide preparation procedures.

Overall, the field of computational pathology sees contrastive learning develop into a crucial approach for current research. The method generates dependable adaptable image features which help resolve crucial problems connected to the restricted amount of available expert annotations. Medical imaging AI tools will become more dependable and scalable through improved precision and speed of diagnostic processes enabled by evolving self-supervised learning methods.

2.5.4. Challenges and Future Directions in Contrastive Learning

Although contrastive learning has shown great promise in medical imaging, especially in histopathology, there are still important challenges that need to be addressed before it can be widely adopted in clinical practice.

One major challenge lies in choosing and fine-tuning the right data augmentation techniques. The success of contrastive learning depends heavily on creating meaningful pairs of augmented images, where each version still preserves the key biological and pathological features. In histopathology, however, common augmentations like cropping, rotation, or color adjustments can sometimes distort crucial details, which may hurt the model's ability to learn useful representations. Recent studies have highlighted the importance of developing specialized augmentation strategies that protect diagnostic features while still introducing enough variety to

support effective learning [37].

The high complexity of contrastive learning methods represents a major challenge for their implementation due to their steep computational requirements. Both SimCLR and MoCo struggle in clinical settings due to their need for sizable batch sizes and memory banks even though these healthcare facilities typically have constrained computer resources [40]. Researchers should explore advanced versions of contrastive learning techniques or develop mixed models incorporating the advantages of contrastive learning methods with basic yet accessible architectural designs for future work. Medical implementation benefits from these efforts because they establish methods which increase the productivity and scalability of these advanced techniques [41].

The interpretability of features derived from contrastive learning requires stronger investigation in order to advance understanding. The models demonstrate great ability to generate valuable generalized representations though the actual biological and pathological features they detect remain elusive to interpretation. Such unclear decision processes limit clinical professionals in their trust toward adopting these models. Explainable AI features including saliency maps and Grad-CAM visualizations alongside attention mechanisms should be implemented to improve the healthcare providers' understanding of these models' clinical decision-making capabilities [42].

Research shows that future applications can achieve successful results by integrating supervised methods with contrastive learning. Research demonstrates that hybrid approaches utilizing small labeled data sets with large unlabeled data collections enhance both accuracy along with generalization effectiveness. These methods provide an alternative solution to using supervised or unsupervised learning independently. The emergence of federated contrastive learning represents a prospective novel method for handling these issues. Healthcare facilities can develop flexible systems by jointly training models using this process because patient

information remains protected during the training process [43].

Tackling these challenges through focused research and innovation will be essential for making contrastive learning more interpretable, reliable, and useful in clinical practice. This progress will help pave the way for wider adoption of AI-driven tools that are not only effective but also transparent and trustworthy in real-world medical imaging.

2.6. AI in the prognosis of lung cancer

Apart from prognosis, AI is also effective in forecasting lung cancer progression and treatment effectiveness. Research done by Teramoto et al. [23] where the morphological images of the cytological were used to train a CNN model in order to extract features and had a recognition accuracy rate of 71.1% which was close to the pathologists. Coudray's study [25] employed a deep learning model for classifying different types of tissues on TCGA images with a classification accuracy of 97%. Furthermore, Khosravi et al. [30] proposed a deep convolutional neural network (DCNN) to classify between benign and malignant tumor with a high distinction for the squamous cell carcinoma and adenocarcinoma. These studies demonstrate lung cancer diagnosis and prognosis that depicts AI capabilities especially the use of deep learning.

Shuyin Duan [44] proposed a complex three-layer diagnostic system of lung cancer using various machine learning techniques which include decision trees (C5. 0), artificial neural networks (ANN) and support vector machines (SVM). This study demonstrated the effectiveness of combining various models, with the ANN achieving the highest area under the curve (AUC) score of 0.908, just ahead of the SVM at 0.910. These considerations support the use of multimodal AI systems for improving diagnosis productivity and stability.

Other contributions include Wei Choi's [45] for example who designed a radiomics prediction model that can be used to distinguish between a small pulmonary nodule in a low dose CT scan, and Marliese Alexander who [46] designed a system that determines cancer eligible

patients for clinical trials. Nasrullah Nasrullah [47] presented a 3D Customized Mixed Link Network (CMixNet) for early identification of malignant lung nodules and combining the diseases symptoms and clinical indexes to decrease misdiagnosis and false positive. Finally, Sijia Cui [48] implemented DL algorithms for the detection of pulmonary nodules in LDCT scans which confirm the extensive involvement of AI in cancer.

Table 1. Methodological approaches in pathology studies.

Source	Year	Method	Sample size
Zhang L *et al.* [28]	2021	ResNet and InceptionV2 (IncRes model)	150
Gazi Muhammed *et al.* [49]	2022	Hybrid Artificial Intelligence Method	1097
Surajit Das *et al.* [50]	2023	Pre-trained VGG19 CNN architecture	14600
Ayah Alomar *et al.* [51]	2023	ResNet50 and InceptionV3	1688
Gajera *et al.* [52]	2021	Generative Adversarial Network (GAN)	82
Dua Hişam *et al.* [53]	2021	DarkNet-53 , ResNet50, and VGG19	1600
Md. Sabbir Ahmed *et al.* [54]	2023	Decision Tree, Logistic Regression, Random Forest, Naive Bayes classifier	238
Muntasir Mamun *et al.* [55]	2022	XGBoost, LightGBM, Bagging, and AdaBoost	309
R. P. Ram Kumar *et al.* [56]	2023	CNN(Kaggle Domain)	1250
P Shyamala Bharathi and Praveen Kumar Reddy.T [57]	2022	LAMNN with Image Pruning Strategy	320
Francesco Ciompi *et al.* [58]	2017	CNN:Vector machine	943
Chao Zhang *et al.* [59]	2019	CNN	2285
Diego Ardila *et al.* [60]	2019	Deep Learning(three dimensional CNN)	6716
Panayiotis Petousis *et al.* [61]	2019	Machine Leaarning,DBN	5402
Mitsuru Koizumi *et al.* [62]	2020	ANN and bone scan index (BSI)	54
Jing Li *et al.* [63]	2020	The DECT-based deep learning	204
Xiaojun Yang MD *et al.* [64]	2020	Deep learning signatures	348
Yuan Gao *et al.* [65]	2019	FR-CNN deep learning	1371
M. Dohopolski et al [66]	2019	CNN(AlexNet-like or UNET)	129
Yoshiko Ariji, *et al.* [67]	2019	Deep learning image classification system	127
Zhou YP *et al.* [68]	2019	The HD MRI lymph automatic recognition system	301

Table 2. Overview of pathology imaging data

Source	Dataset	Release Date	Sample size	Number of Images	Image Modality	Image Dimension	Image Format	Ground Truth Availability
Nusraat Nawreen *et al.* [69]	Cancer Imaging Archive	2021	NS	NS	CT	256 x 256	DICOM	No
Zhang Li *et al*. [28]	The ACDC@LungHP	2021	200	200	WSI	768 x 768 x 3	H&E	Yes
Ahmed S. Sakr [70]	LC25000	2022	25	25000	Histopathological images	768 x 768	JPEG	Yes
Xi Wang *et al.* [71]	The Cancer Genome Atlas (TCGA)	2020	112	112	WSI	244 x 244 x 3	JPEG RGB	Yes
Nowshin Tasnim *et al.* [72]	IQ-OTH lung cancer	2024	1097	32717	CT	512 × 512, 512×623, 404 × 511	NS	Yes
Yoganand Balagurunathan *et al.* [73]	ISBI 2018 Lung Nodule Malignancy Prediction test set	2021	1593	100	LDCT	3D scans	DICOM,DICOM-RT,NIfTI	Yes
D. Jayaraj and S. Sathiamoorthy [74]	Lung Image Database Consortium (LIDC)	2019	NS	1018	CT	512 x 512	JPEG Grayscale	Yes
Sayyada Hajera Begum *et al.* [75]	Lung Cancer CT Scan Images Dataset	2022	110	800	CT	256x256	NS	Yes
Worawate Ausawalaithong *et al.*, [76]	JSRT Dataset	2018	247	247	X-ray	2048 x 2048	NS	Yes
Worawate Ausawalaithong *et al.*, [76]	ChestX-ray14 Dataset	2018	112120	112120	X-ray	1024 x 1024	NS	Yes
Juan Cañada *et al.* [77]	LC25000 dataset	2019	25	1200	Histopathological images	512 x 512	IMG	Yes
Massion, P. P., *et al.* [78]	NLST dataset	2020	463	14761	CT	0.25x 0.25 x 1	NS	Yes
Rohit Y. Bhalerao *et al.* [79]	LIDC (Lung Image Database Consortium)	2019	910	910	CT	256x256	DICOM,JPEG,PNG	Yes

CHAPTER 3 – RESEARCH METHODOLOGY

3.1 Histopathological Image Classification

3.1.1 Dataset description

This research uses a starting dataset consisting of 15,000 histopathological images from lung scans. The database contains three equal parts for the classification of images that include Benign as well as Adenocarcinoma (AC) and Squamous Cell Carcinoma (SCC). A total of 5000 images exist in each section because this balanced styling ensures unbiased training and model evaluation becomes possible.

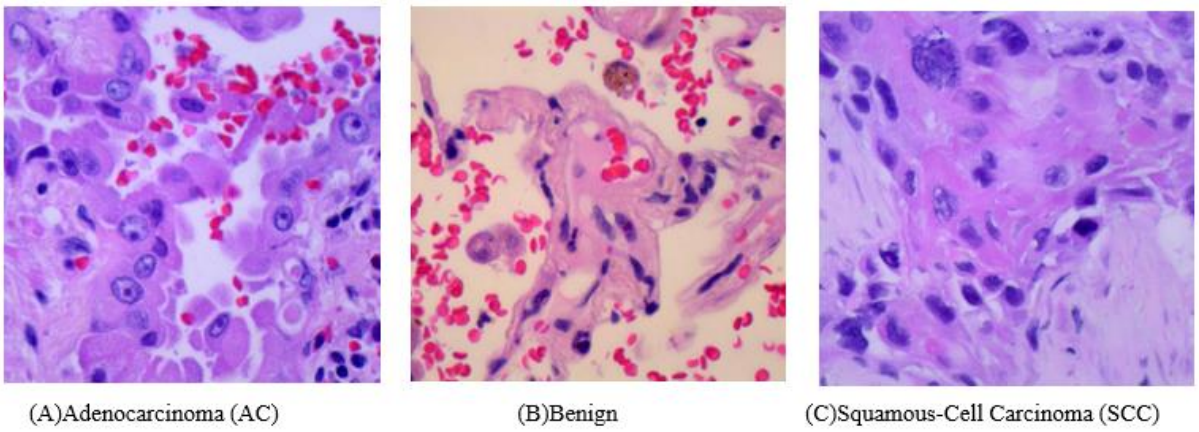


Figure 4: Dataset classes:Adenocarcinoma (AC),BenignSquamous Cell Carcinoma (SCC)

3.1.2 Data preparation

The image processing used uniform resizing to achieve 224 × 224 pixels dimensions to enable consistent processing and model compatibility. The sample received uniform resizing then became separated into training (80%) along with validation (10%) and testing (10%) subsamples through stratified sampling which kept class balance in all subsets.

Data augmentation techniques that applied extensive methods to the training set addressed both generalization capability and overfitting risks. Random rotations along with horizontal and vertical flips as well as zoom variations and brightness adjustments expanded data diversity and

this made CNN models excel when processing common histopathological image variations.

3.1.3 Model Architectures

To investigate the effect of architectural choices on the performance of CNNs for lung cancer histopathological image classification, five custom architectures were developed and evaluated. The models differentiated through their individual network configuration elements of depth and activation functions alongside pooling mechanisms together with parameter count to reveal trade-offs between information accuracy and processing efficiency.

The first model focused on creating a design that balanced efficiency with simplicity. The network contained three convolutional layers in sequence which were activated by ReLU followed by MaxPooling implementation. The implementation enabled effective training times coupled with good initial results that made it a perfect benchmark model. The network utilized dropout regularization as an extra layer before full connection layers to prevent overfitting and boost generalization capabilities.

The second architecture incorporated a substantial depth with thirteen convolutional layers into its design. Through structuring its layers with ReLU activation and MaxPooling operations the model attained superior feature extraction ability through its advanced learning of abstract patterns. The deeper framework excelled at detecting elaborate histological structures that define cancer subtypes.

The third model used LeakyReLU as an alternative to ReLU since this activation increased gradient flow while reducing non-responsive neurons. The AveragePooling operation replaced MaxPooling to enable the model to draw information from wider areas spanning input features. These two techniques strengthened the learning stabilization process and protected important spatial element features common to medical imaging applications.

The fourth architecture used the strengths of its predecessor to build a model with optimal computational efficiency. AveragePooling combined with LeakyReLU activation remained in the

network framework while decreasing the number of convolutional layers. The designed architecture applies to situations with limited availability of computational power while keeping activated robust features enabled through pooling mechanics.

The fifth model incorporated LeakyReLU activation by using a modified pooling operation. This network design featured a middle-depth architecture along with an optimal parameter count to enhance generalization capabilities while optimizing feature discrimination abstraction. The new architectural design focused on creating a system that adjusted well to different learning conditions.The detailed hyperparameters and structural attributes of all five custom CNN models are summarized in Table 3.

Table 3: Hyperparameters and Architectural Details of Custom CNN Models

Hyperparameter	Model 1	Model 2	Model 3	Model 4	Model 5
Optimizer	Adamax	Adamax	Adam	Adam	Adam
Batch Size	64	128	128	64	64
Activation function	ReLU	ReLU	LeakyReLU	LeakyReLU	LeakyReLU
Dropout	0.3	0.3	0.4	0.4	0.4
Epochs	5-20	5-20	5-20	5-20	5-20
Learning rate	0.001	0.001	0.0001	0.0001	0.0001
Pooling type	MaxPooling	MaxPooling	AvgPooling	AvgPooling	-
Parameters	2.87 M	21.15 M	11.61 M	10.83 M	11.18 M
Conv layers	3	13	7	5	6

3.1.4 Activation Functions and Pooling Operations

Activation functions drive neural network training through their mechanism of generating non-linear features during the feature extraction sequence. The non-linear nature of CNNs helps them understand complex data relationships which proves essential for medical image recognition tasks. ReLU activation functions have found extensive use in deep learning models because they effectively mitigate gradients decay while offering a simple and productive solution to the vanishing gradients issue of sigmoid and tanh activation functions. The use of ReLu function in the network greatly shortens the learning cycle and improves speed and efficiency. The mathematical representation of ReLU :

$$ReLU(x) = max(0, x) \tag{3.1}$$

The function transforms all negative inputs into zeros to selectively activate them through positive input values. Through its implementation ReLU makes some neurons so inactive that they become permanently nonresponsive which is known as the "dying ReLU" issue. The persistent negative inputs to neurons drive them to produce zero gradients that stop their learning process. The LeakyReLU activation function serves as a substitute solution for standard ReLU since it handles its key limitations. Through its introduction of a little negative slope LeakyReLU makes neurons capable of learning from negative input values thus minimizing their potential inactivity. Its mathematical representation is:

$$LeakyReLU(x) = \{x \quad if\ x > 0, \alpha x \quad otherwise \tag{3.2}$$

,where α is is normally set to 0.01 to regulate the degree of negative gradient learning which helps enhance network robustness.

The use of Pooling operations helps CNN architectures reduce their computational demands and extract key translation-invariant features through downsampling of feature maps. Medical image analysis benefits greatly from CNN generalization through pooling because it decreases spatial variation sensitivity. The pooling methods used in most applications include

MaxPooling and AveragePooling techniques. The MaxPooling method finds and maintains the most important pixel value throughout a specific spatial area as it selects discriminative features with the highest emphasis. Mathematically, it is represented as:

$$MaxPooling = (x_i) \tag{3.3}$$

,where x is the values of the input image in the pooling region.

The network becomes spatially invariant because this approach enables pattern emphasis and noise suppression thus allowing small input image position shifts to be disregarded. Medical images benefit from this technique since their essential diagnostic indicators show small differences in spatial positions.

AveragePooling determines the computed arithmetic mean from the pixel values contained inside selected pooling areas. Its mathematical definition is:

$$AveragePooling = \frac{1}{N}\sum_{i\,\epsilon\,region}^{1} x \tag{3.4}$$

,where x is the values of the input image in the pooling region.

AveragePooling method applies smoothing to feature maps while maintaining a broader context composition which surpasses the focus on primary features. Model learning stability improves through AveragePooling because it achieves balanced feature aggregation which benefits diagnostic applications requiring spatial consistency between medical image elements.

3.1.5 Training procedure

To train the five custom-designed convolutional neural network (CNN) architectures, a standardized procedure was implemented, tailored to the multi-class classification of lung cancer histopathological images. Categorical cross-entropy loss function was used to train all models because it is a suitable loss function for mode- class problems as it provides a means of measuring divergence in predicted probability distributions from the actual class label. This loss enables stable convergence by imposing penalties to misclassifications in relation to confidence in prediction.

Optimizations were performed with adaptive gradient based algorithms. The Adam optimizer was used on Models 3 to 5 and the AdaMax optimizer on Models 1 and 2. Both are variations of the stochastic gradient descent mechanism (SGD), equipped with dynamic learning introspective mechanisms. Adam estimates first and second moments of gradients, which make it robust to noisy and sparse updates properties favorable when working in the medical domain where pixel level variations occur often. AdaMax, a variant under the infinity norm, provides more stability and convergence behavior especially in deeper architectures or if the gradient variance is high.

Initially several learning rates were tried and a final value was chosen empirically. For Models 1 and 2, a learning rate of 0.001 was applied, while inModels 3 to 5 we adopted a learning rate of 0.0001 because of their deeper/ more sensitive configurations. A batch size of 64 or 128 was used for training all models, as was found to be a good compromise between computational efficiency and the stability of the gradient estimates. Even though larger batch sizes may accelerate training process, they can decrease generalization performance, while smaller batches can pollute the weight updates with excessive noise.

The number of training epochs varied between 5-20, depending on architectural depth and validation dynamics of the model in question. As to the models which were both simpler and not much complex, they converged very quickly as compared with deeper or more complex architectures that needed further iterations for their learning. Early stopping was used in order to prevent overfitting, stopping training when the validation accuracy and loss stabilized.

Dropout regularization was included to improve generalization taking advantage of co-adaptation pruning away. 0.3 dropout rate was assumed in Models 1 and 2 and 0.4 dropout rate under Models 3 to 5 that suggests their greater parametrization and depth.

During training, model performance was assessed by validation accuracy and loss so the networks maintained high generalization ability over the histopathological dataset. The training

pipeline was developed such that each architecture obtained a trade-off between classification accuracy and computational efficiency that was optimal. Summary details for the five architectures are listed in Table 3 (see Section 3.1.3).

3.1.6 Classification Using Pre-trained CNN Architectures

Building upon the custom CNN models generated in this work, the set of well-known pre-trained architectures; VGG19, ResNet50 and EfficientNet have been used. These are well established for reliability in medical image analysis, and were integrated through a transfer learning pipeline. Each model was initially trained on a large-scale ImageNet dataset before subsequent fine-tuning with a curated lung histopathology dataset for this study.

Before training all images were uniformly resized to the size of 768x768 pixels and split into training (85%) and test (15%) sets. In adaptation, convolutional backbones of pre-trained models were left unchanged while only fully connected layers were re-initialized and trained. For VGG19, for the classifier three dense layers of sizes 4096, 4096, 3 were used. ResNet 50 and EfficientNet were also trained to deliver predictions for three output classes. The training process incorporated Adadelta optimizer with 1e-4 learning rate, and MSE function was used to calculate loss and ensured a variety across all comparative experiments.

Finally, ensemble model using soft voting mechanism was constructed which would combine predictions from all three architectures. The ensemble strategy greatly improved the levels of accuracy and robustness, with every model contributing complementary strengths. The outputs were contrasted against those of the custom built CNN models in the resulting analysis section that followed.

3.2 Contrastive Learning in Histopathological Image Analysis

3.2.1 Dataset Description

The dataset of 15,000 unlabeled histopathological images described previously was also utilized for the contrastive learning component. The massive unlabeled dataset serves contrastive learning by allowing it to derive effective robust feature representation through data-inherent relationships. For this self-supervised learning part of the dataset each image carries the category value of Benign or Adenocarcinoma (AC) or Squamous Cell Carcinoma (SCC) but they were unlabeled.

The substantial amount of unlabeled data improves the contrastive learning model's capacity to apply learned patterns to different histopathological patterns. The large collection of unlabeled images allows the extraction of relevant discriminative features that enhance the subsequent supervised classification model efficiency. The practical value of contrastive learning becomes evident through this method especially in medical imaging fields since it takes a long time and costs money to obtain expert pathologist-labeled data.

3.2.2 Data Augmentation and Preparation

The effectiveness of contrastive learning largely depends on the high quality and diverse character of augmentation images which stem from primary data. The augmentation process requires semantic preservation together with sufficient variability to expose neural networks to robust and generalized features. It was implemented SimCLR framework along with specialized modifications to fulfill the requirements of histopathological images in our dataset. The main purpose of implementing data augmentation in SimCLR involves making the network optimize matching representations between different image augmentations from the same original picture while preventing overlap between different images' representations.

The objective required a combination of data augmentation approaches which fit specific needs of analyzing histopathological images. The first data processing step included random

cropping followed by resizing operations. Random cropping proves useful in histopathological images because it teaches the network spatial independence which allows it to detect diagnostic elements anywhere in the image. The cropped areas received standard image dimensions of 224 by 224 pixels after random selection. Whatever dimension alterations occur through this preprocessing improve dataset diversity while maintaining consistent input sizes needed for effective processing by the convolutional neural network which results in unbiased feature extraction.

Additionally, color distortion served as an essential augmentation method in the process. Learning models can get distracted by irrelevant superficial variations in histopathological images because of their significant variability in staining protocols and intensities. Color distortion reductions obtained through modifications of brightness values alongside saturation adjustments and hue modifications and contrast adjustments help minimize this issue. By executing this approach the neural network acquires representation patterns that ignore color variations so the model can effectively detect significant structural tissue elements used for medical diagnosis identification.

The augmentation approach included grayscale transformation as one of its components. The model develops an increased focus on identifying morphological and structural features through grayscale conversions which apply randomly. The method encourages the model to learn stain-independent features that matter significantly in medical imaging because color fluctuations result from various sample processing methods.

The augmentation process included two duplicate applications per image of random cropping and resizing as well as color distortion and grayscale conversion techniques. The augmented views function as positive pairs for SimCLR training to help the model develop effective representations that are crucial for later histopathological image analysis tasks. Real-time GPU-transformations based on Kornia library powers efficient processing for training. The

dataset required images to undergo normalization through specific mean values of [0.4503, 0.3645, 0.2949] and standard deviation values of [0.2490, 0.2213, 0.1984].

3.2.3 Contrastive Learning Architecture and Implementation (SimCLR)

To build the contrastive learning framework for this study, it was adapted a customized version of SimCLR (Simple Framework for Contrastive Learning of Visual Representations) that was specifically adjusted to suit the context of histopathological image analysis. The architecture is structured around two primary modules: a convolutional encoder and a projection head.

The encoder plays a critical role by extracting informative and abstract visual features from input images. For this purpose, ResNet-18 was chosen due to its effective trade-off between computational speed and representational power. The final classification layer of the original ResNet-18 was removed and replaced with a projection head made up of two fully connected layers. There was a ReLU activation in the first layer of 256 neurons and an outputting linear layer to map features into a 128-dimensional embedding space on the second layer.

This projection head provides a condensed version of high-dimensional features upon which the contrastive loss can be derived. By separating the tasks of feature extraction and optimization, this setup enhances the model's ability to generate representations that are more transferable and effective for downstream classification tasks closely following the original SimCLR architecture design.

During training, each image is augmented twice to create two distinct but related views. Such views are separately processed by the shared encoder and projection layer, which produce two respective feature embeddings. The following are compared to other samples included in the training batch given the normalized temperature-scaled cross-entropy loss (NT-Xent loss). The purpose is to reduce the distance of similar pair embeddings (ways of seeing the same image) and increase the distance between unconnected image pairs.

The choice of ResNet-18 over another variants (ResNet-50) is guided by computational

constraints and the dataset scale. Despite its reduced depth, ResNet-18 is capable of learning discriminative histological features while maintaining training efficiency, which is critical in environments with limited GPU resources.

The architectural setup creates a solid foundation for contrastive pretraining which allows the model to acquire reliable and versatile features from unlabelled data.

3.2.4 Training Procedure for Contrastive Learning

This study incorporated a contrastive learning strategy based on the SimCLR framework, which uses contrastive loss as the foundation for representation learning. The key goal of training was to enhance similarity between different augmented views of the same image often referred to as positive pairs,while reducing similarity between representations of distinct images, or negative pairs. This learning strategy encourages the model to extract features that remain stable despite common variations such as image transformations and perturbations.

To implement this, we used a ResNet-18 model as the encoder. The network was initialized with ImageNet weights, and its final classification layer was removed. Instead, a custom projection head was connected, made up of two thick layers. The first layer shrunk the feature space from 512 to 256 dimensions through the process of ReLU activation, after which the second layer shrunk the output to 128 dimensions taking it to the embedding space used in contrastive training. This dimensionality reduction is consistent with the essence of the SimCLR, and allows for optimum application of contrastive loss in a compact feature space.

The training process was implemented through the use of Adam optimizer with a learning rate of 0.001 and a weight decay (5e-4) to reduce over-fitting. A batch size of 512 was chosen to ensure a balanced and diverse mix of positive and negative examples during each training iteration. Since the performance of contrastive learning is highly sensitive to the variety of negative pairs in a batch, this choice was critical for stable learning.

3.2.5 Evaluation of Learned Representations

To evaluate the effectiveness of the features learned during contrastive pretraining, a fine-tuning stage was performed on a labeled subset of the same histopathological image dataset. The encoder originally trained using the SimCLR framework was repurposed as a feature extractor. This was achieved by detaching its projection head and adding a simple linear classification layer to distinguish among the three target classes.

The fine-tuning was conducted in a supervised manner using categorical cross-entropy loss, which is appropriate for multi-class classification tasks. Optimization was done by Adam optimizer with learning rate being 0.001 and a batch size of 64. In order to encourage the process of generalization and decrease the risks of overfitting, early stopping was applied according to the performance ob prima validation.

Model performance was evaluated on the basis of widely accepted classification metrics (accuracy,precision, recall) which gave quantitative appreciation of diagnostic reliability. In addition, a t-SNE visualization was computed for the encoder's high dimensional outputs to visualize the two-dimensional projection of the outputs and a qualitative view of class separation in the learned representation space. This supported visual confirmation of the model's capability to differentiate among various tissue categories.

The pretrained encoder showed great generalization and needed fewer epochs to converge than training from scratch. These results support the claim that contrastive pretraining enabled extraction of semantically meaningful and class discriminative features such that it is a promising approach for medical image classification in weakly-supervised settings.

CHAPTER 4 – Simulation results and analysis

4.1 Evaluation of Proposed CNN Architectures

Model 1: Performance Evaluation

Model 1, which had three layers of convolution and regularization by dropout, had a validation accuracy of 0.93 and a validation F1 Score of 0.92. While it showed small variance during training, its generalization capability was strong though. As is shown in Figure 5, the training and validation accuracies show an upward trend consistently resulting in a minimal generalization gap and signifying stable learning dynamics.

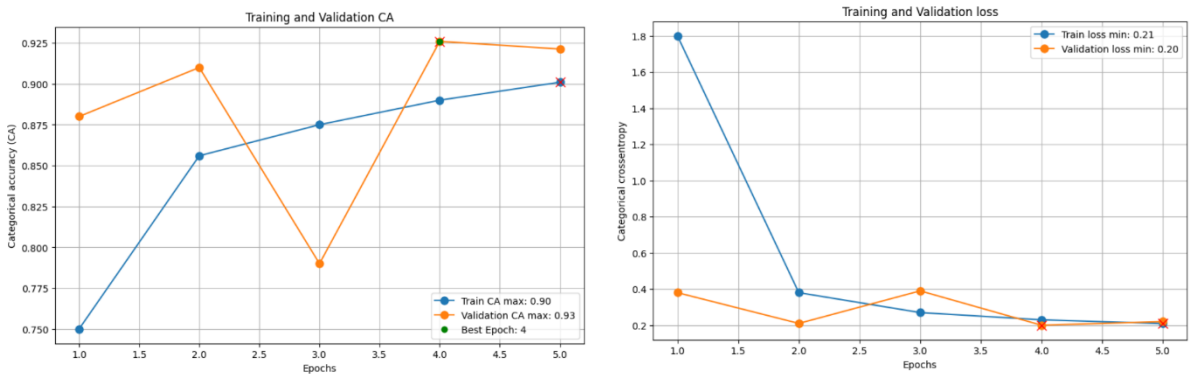


Figure 5: Training and validation accuracy (left) and loss (right) curves for Model 1

From an initial high point of 1.8, loss analysis in Figure 5 demonstrates a fast drop-off of training loss to a stabilized point of 0.21, while validation loss plateaued at 0.20. Based on these curves, there is efficient convergence and little to no overfitting, which supports the validity of learned representations.

The strong performance in classification of all three tissue types is demonstrated in the confusion matrix in Figure 6. Benign and squamous cell carcinoma were classified correctly in 450 and 486 cases respectively with minimal misclassification between adenocarcinoma and the other two classes.

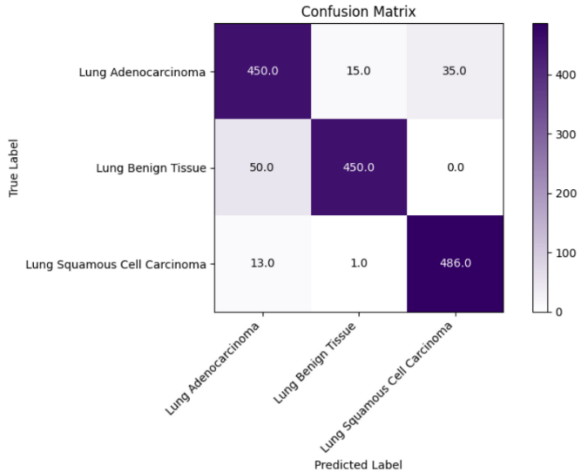


Figure 6: Confusion matrix for Model 1

Model 2: Performance Evaluation

Model 2 with an added convolutional layer and larger filter size demonstrated a better result with validation accuracy of 0.98 % and validation f1 Score of 0.98 %. This improved architecture led to rapid convergence and high generalization as can be observed from Figure 7 where both accuracy and loss curves demonstrate steady progress and low overfitting.

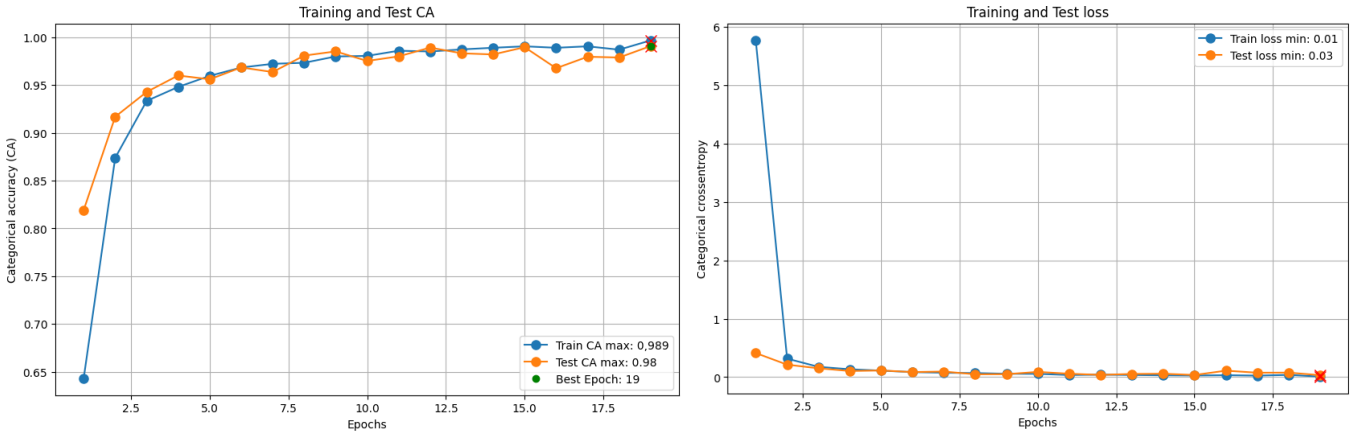


Figure 7: Training and validation accuracy (left) and loss (right) curves for Model 2

Loss convergence demonstrated in Figure 7 shows very fast convergence as training loss rapidly decreases from an initial value of 5.8 to 0.01 and test loss converges at the low minimum value of 0.03. The small stable value maintained over the epochs' ranges verifies the model's stability and effective optimization when being learned.

The confusion matrix of Figure 8 which endorses this robust form of performance, with very good separation among all classes. The model misclassified to a very small extent and for very rare disease subtypes and proved to be nearly perfect for benign tissue and squamous cell carcinoma.

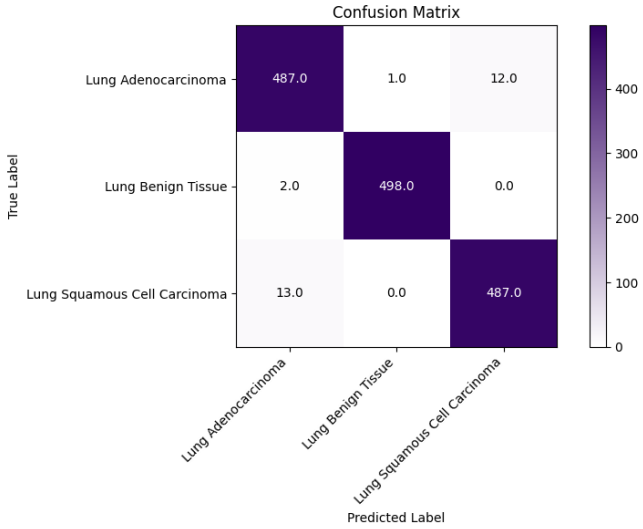


Figure 8: Confusion matrix for Model 2

Model 3: Performance Evaluation

Model 3 with the LeakyReLU activation and the AveragePooling steps produced an excellent result with validation accuracy of 0.94 and the F 1 score of 0.93. As it appears in Figure 9, stable learning and efficient convergence are achieved, as well as low generalization error after 20 epochs on training and validation curves.

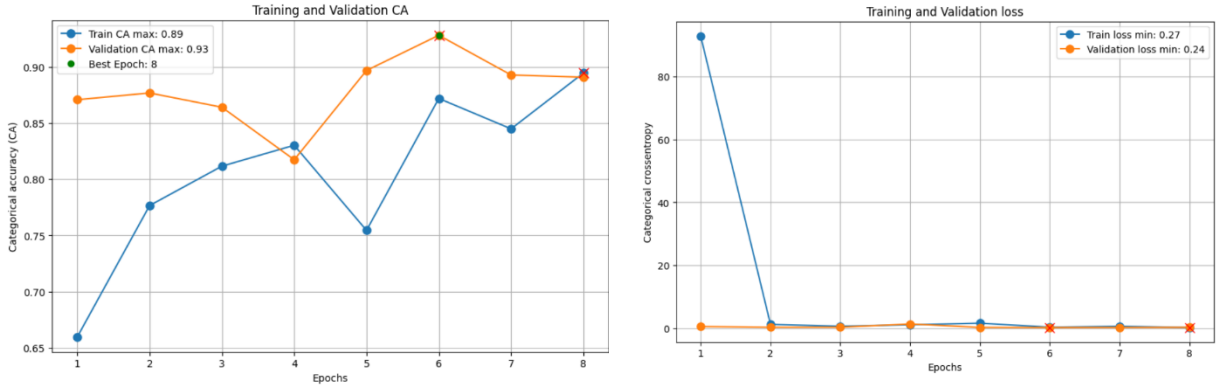


Figure 9: Training and validation accuracy (left) and loss (right) curves for Model 3

The reliable separation among classes, especially for benign and squamous cell carcinoma, is demonstrated by the model's confusion matrix shown in Figure 10. Although some miss-classifications took place between adenocarcinoma and squamous samples, the model remained consistent in detecting the core tissue patterns with high precision.

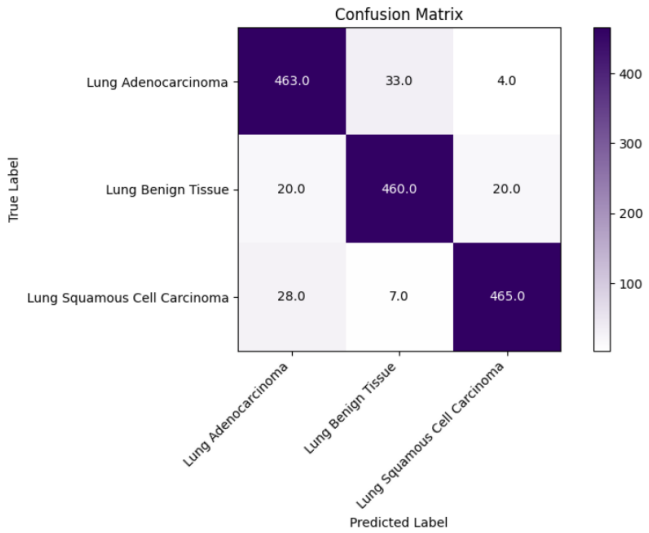


Figure 10: Confusion matrix for Model 3

Model 4: Performance Evaluation

Model 4, performed optimally in balancing between classification accuracy and training speed. From the figure 11, the model achieved a training accuracy of 0.88 and a validation F1 Score of 0.83. Although the model is relatively shallow in comparison with the best performing model architecture, the accuracy curves demonstrate consistent growth throughout epochs with little drift implying stable learning with no serious overfitting.

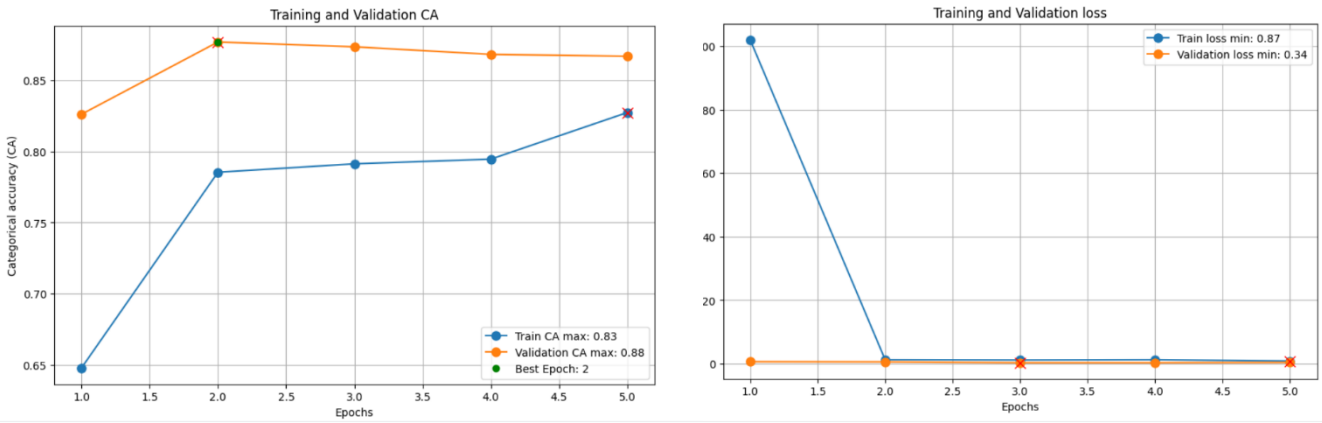


Figure 11: Training and validation accuracy (left) and loss (right) curves for Model 4

Based on Figure 12 the confusion matrix gives deeper information about the classifying behavior of the model. The model classified 430 samples for . adenocarcinoma and squamous cell carcinma classes correctly, whereas, 422 benign tissue samples were accurately predicted.

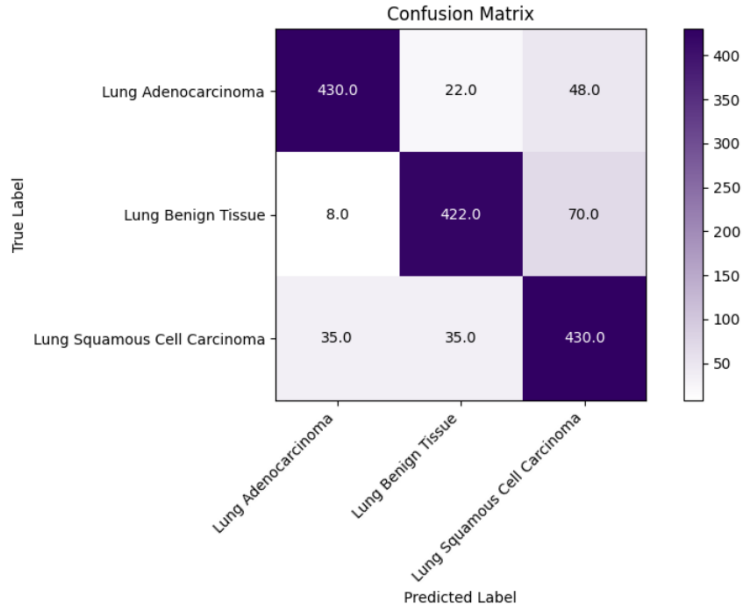


Figure 12: Confusion matrix for Model 4

Model 5: Performance Evaluation

Model 5, which has a streamlined architecture and fewer number of parameters, ensured a great balance between performance as well as simplicity. Its validation accuracy rose up to a value of 0.89 while its validation F1 Score up to 0.89 and thereby showing reliable generalization. As illustrated in Figure 13, both training and validation accuracy curves progressed in a learning curve to a steady increasing curve with negligible divergence, with indication of stable training and low risk of overfitting. This stability is further confirmed by the loss curves, as validation loss achieves a minimum value of 0.20.

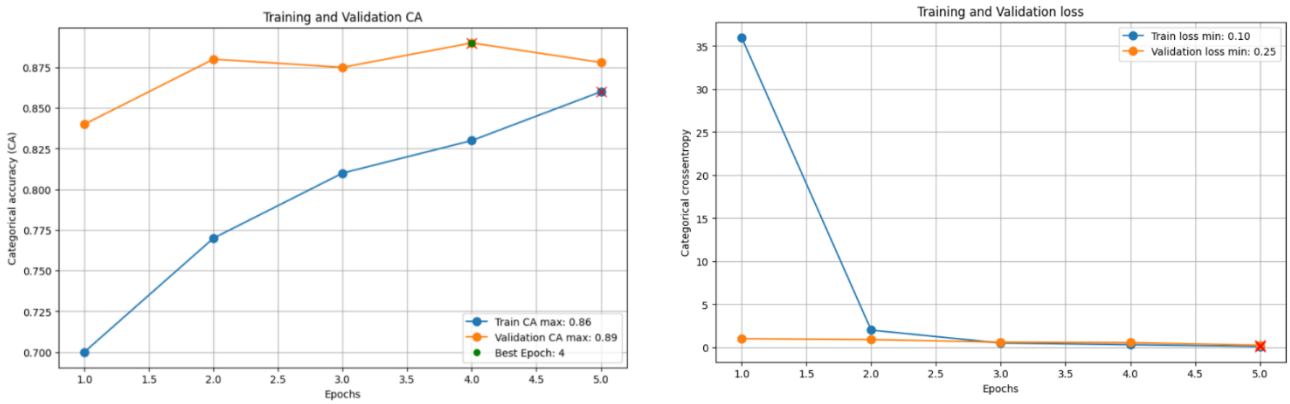


Figure 13: Training and validation accuracy (left) and loss (right) curves for Model 5

A broadly robust classification performance from the confusion matrix presented in Figure 14 is indicated as most errors in classification (55 samples) occur between adenocarcinoma and squamous carcinoma and between benign tissue and squamous (40 samples). In spite of these, the model has been able to predict more than 430 samples for each class into which it classified the samples in questions making it a good candidate for applications having moderate model complexities and reliable accuracy.

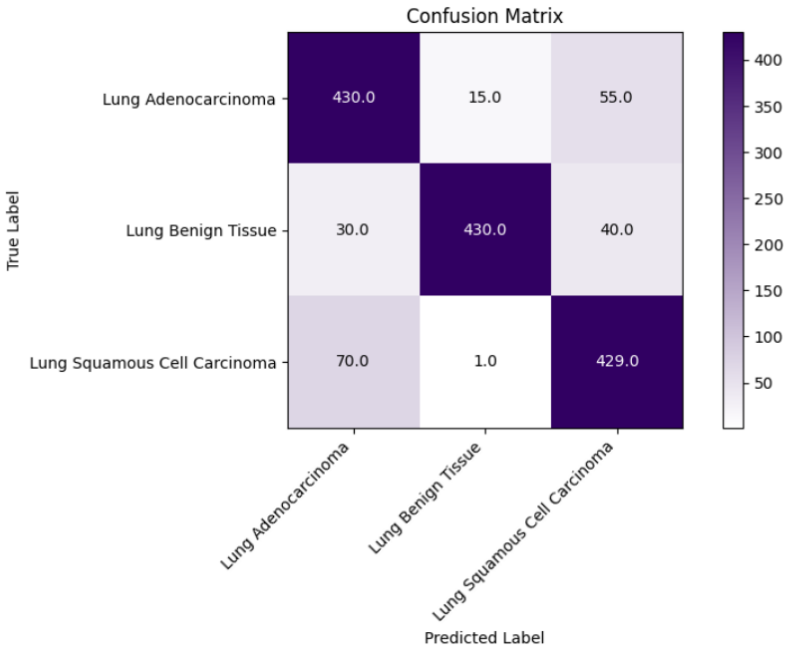


Figure 14: Confusion matrix of Model 5

To make a clear comparison of the models' classification abilities, Table 4 concludes the important performance metrics: Accuracy, F1 Score, Recall, and AUROC; for all five suggested CNN architectures. The integrated vision allows for a rational comparison of their comparative strengths and reveals the best solution for histopathological image classification.

Table 4: Performance metrics for custom CNN models

Model	Accuracy	F1 Score	Recall	AUROC
1	0.93	0.92	0.92	0.92
2	0.98	0.98	0.978	0.97
3	0.93	0.88	0.92	0.97
4	0.88	0.83	0.86	0.84
5	0.89	0.89	0.87	0.89

4.2 Evaluation of Pre-Trained CNN Architectures

To examine the effect of transfer learning in histopathological image classification, three well-known CNN frameworks were trained for transfer learning on a dataset considering 15,000 lung histopathology images, namely: VGG19, EfficientNetB4, and ResNet50. Each model was optimized with the help of the Adadelta algorithm, the learning rate set to 1e-4, and trained for 20 epochs.

The assessments were undertaken using four indicators of performance: accuracy, F1 score, recall and area under the receiver operating characteristic curve (AUROC). These metrics provided a strong basis against which to compare the pre-trained networks with the custom-built CNN models described previously.

VGG19, of the three models, produced the best results by achieving 99.91% accuracy, and high scores in the other values of evaluation. The fast convergence and low overfitting of its low-level representations demonstrate the benefit of borrowing pretrained weights from large-scale data sets such as ImageNet. ResNet50 also did commendable work especially in recall of which could be traced to the fact that it had a residual learning structure that would allow for better flow of gradients and better feature retention. EfficientNetB4 with fewer parameters still had competitive performance and reliable generalization during the training process.

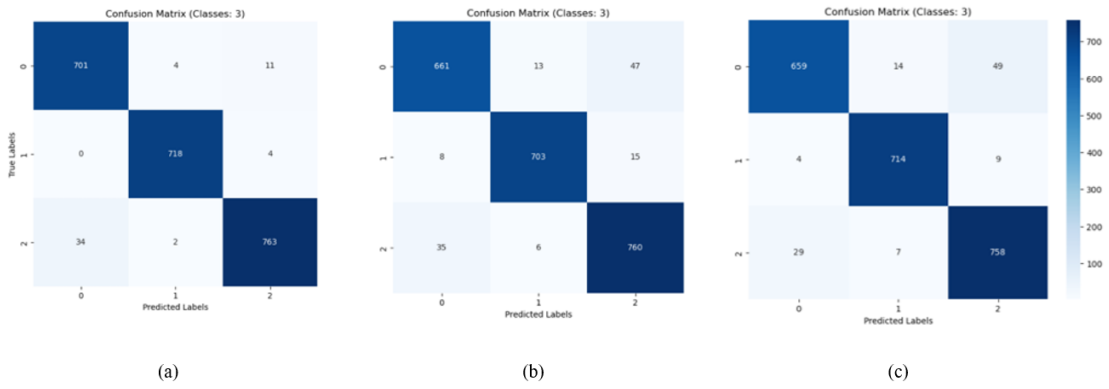
To enhance classification robustness two ensemble strategies were studied. A soft voting ensemble which averages the predicted class probabilities of all three models resulted in 99.95% accuracy. At the same time, a hard voting ensemble based on majority class predictions reached 99.91% accuracy. A combined summary of all these findings is presented as part of Table 5.

Table 5: Classification Performance Metrics for CNN models

Model	Accuracy	F1 Score	Recall	AUROC
VGG19	0.9991	0.9964	0.9964	0.9964
EfficientNet	0.9756	0.9622	0.9622	0.9610
ResNet	0.9964	0.9889	0.9889	0.9817
Ensemble (soft)	0.9995	0.9845	0.9845	0.9845
Ensemble (hard)	0.9991	0.9950	0.9950	0.9951

To complement the quantitative results, each pre-trained model had their corresponding confusion matrices generated, providing a visual representation of what happened in terms of classification throughout the three diagnostic views (see Figure 10): benign, adenocarcinoma and squamous cell carcinoma. The three models demonstrated robust diagonal dominance, driving high true positive rates and low misclassifications.

To analyze the performance of pre trained models on the classification, confusion matrices for VGG19, EfficientNetB4, and ResNet50 were generated. As can be seen through Figure 15, all three models exhibit good class discrimination for three classes.



(a) (b) (c)

Figure 15: Confusion matrices for pre-trained models: (a) VGG19, (b) EfficientNetB4, (c) ResNet50

Subfigure (a) displays the confusion matrix for VGG19 that recorded the best performance, with least off-diagonal misclassifications, a fact that indicates that it has the highest accurecy and recall metrics. EfficientNetB4 in subfigure, (b) performed competitively but moderately misclassified, especially between benign and squamous classes. Balanced results are produced by ResNet50 (subfigure (c)) with fairly similar predictions and high margins between all the three classes.

4.3 Comparative Analysis

A comparative study of the custom designed CNNs and the pre-trained architectures shows significant differences in the level of accuracy of classification, train behavior, generalization ability and computational efficiency. Model 2, among the custom models, showed the highest performance having attained a validation accuracy of 98.0%, given an architecture that is deep in structure with a total of 13 convolutional layers. However, this was at expense of increased training time and computational requirements.

In comparison, pre-trained models, especially VGG19 and the ensemble setups, performed just a little better than the best custom model, achieving validation accuracies above 99%. These findings highlight the benefit of transfer learning where models that have been pretrained on large datasets like ImageNet will adapt further using limited datasets better generalizing to medical imaging tasks.

In terms of training dynamic, pre-trained models demonstrated less fluctuation behavior and were less sensitive to tuning hyperparameter thanks to the strong initialization. In the meantime, the lightweight custom models (Models 3, 4, 5) were a compelling alternative as they delivered decent performance at the reduced cost of architectural complexity and were therefore more realistic for the real-time or resource-constrained environment.

In addition, confusion matrices indicated that, although both custom and pre-trained

models, in general, did a good job in distinguishing between benign and malignant tissue types, they demonstrated a lower number of misclassifications on all three classes for pre-trained networks. In ensemble models, this was particularly noticeable with an ensemble of the strengths of the individual classifiers to increase robustness.

In general, the results indicate that although deep custom models can provide competitive results, the pre-trained ones are of much better performance and generalization and very effective, thus, suitable for histopathological image classification tasks. However, the fact that the flexibility and reduced overhead of custom CNNs still may render them appropriate in cases where hardware limitations or deployment constraints do play a role.

4.4 Contrastive learning results

The SimCLR based contrastive learning method applied on histopathology dataset proved to be good in feature representation and seen to improve downstream classification accuracy. The process started by unsupervised pretraining of a ResNet-18 encoder using unlabeled histological image data.

As can be seen on Figure 16, the training loss of the model systematically reduced throughout 100 epochs from roughly 5.7 to slightly below 5.0 indicating stable learning behavior and network convergence. This decline in NT-Xent loss function, normalized temperature-scaled cross-entropy, is an indication of increasing power with the most meaningful image patterns being differentiated by the encoder.

The NT-Xent loss favors the model to increase alignment of augmented versions of the same image (positive pairs) but meanwhile push away unrelated images (negative pairs). The slowness of the decline in loss suggests that the model did adjust its internal representations in a reasonable manner, which means that positive samples were tightened together in the space of features while negative samples were taken further from each other. This behavior allows to assume that the encoder learned how to accentuate biologically significant structures under

different image augmentations.

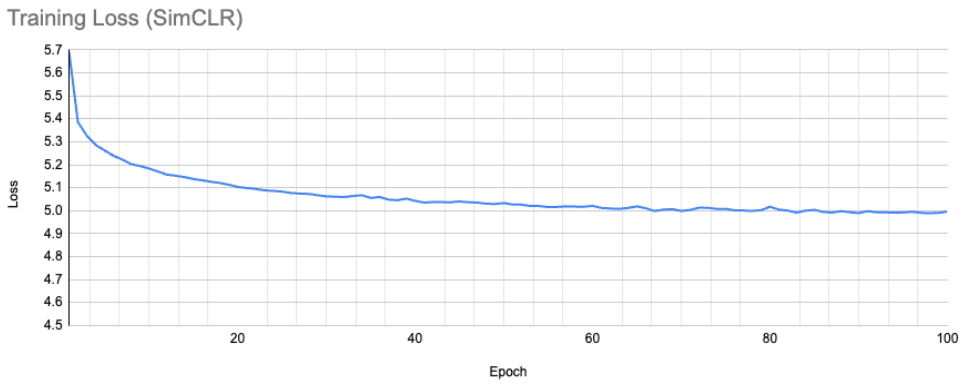


Figure 16: SimCLR Training Loss Curve

After contrastive pre-training the frozen ResNet-18 encoder was fine-tuned with a linear classification head for evaluation of the transferability of the learned representations. Figure 17 shows training and test loss curves during 20 epochs. Both the curves have a steady and consistent decline over the course of training without any signs of divergence or overfitting.

This mirror reduction of loss is indicative of the strength and excellence of the features that were extracted in the self-supervised SimCLR phase. The well learned by the encoder from unlabeled data, the linear classifier had all it needed to adapt without so much supervision. The lack of sharp oscillations and early plateaus shows that the representations were not only meaningful, but also linearly separable, which verifies the power of the contrastive pretraining strategy.

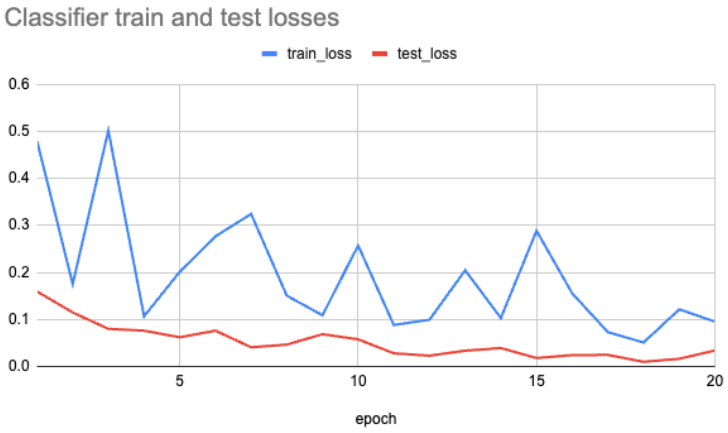


Figure 17: Classifier Training and Test Loss Curves

The final performance metrics out of best epoch (epoch 20) are summarized in table 6. With the classifier, 96.58% accuracy is achieved along with 96.56% precision and 96.57% recall. These results suggest that the classification performance over all classes are strong and balanced.

Table 6: Evaluation Metrics of the SimCLR

Performance Metric	Value (%)
Accuracy	96.58
Precision	96.56
Recall	96.57
Train Loss	0.0954
Test Loss	0.0337

These high metrics are further illustrated by the confusion matrix presented in Figure 18 that shows good class segregation and low misclassifications among adenocarcinoma, squamous cell carcinoma and benign tissue. The majority of the predictions are found in the diagonal and this is to prove that the classifier does indeed label correctly across all of the classes.

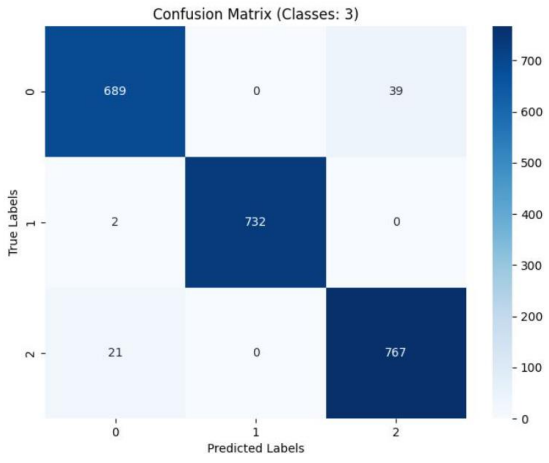


Figure 18: Confusion matrix of the linear classifier trained on SimCLR-pretrained embeddings

A two-dimensional t-SNE embedding of the vectors from the encoder that results in a visualization of how well learned representations capture class-specific structure in feature space as shown in Figure 19. This plot is showing three well distinct clusters associated with the three tissue types having very little overlap. The spatial separation observed in the t-SNE projection is a strong qualitative indicator that the SimCLR encoder has been able to learn, without use of labels during optimization, semantically useful and separable representations.

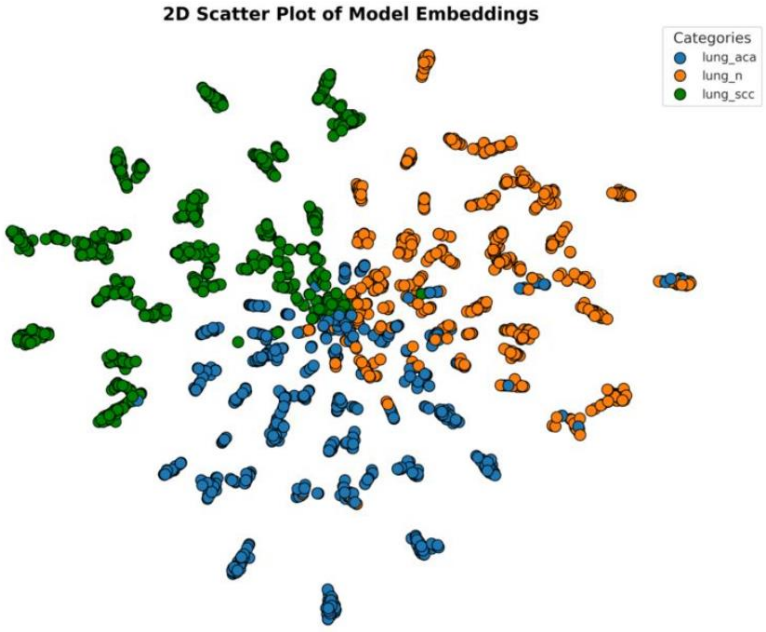


Figure 19:2D t-SNE projection of SimCLR encoder features for the three histopathological classes

All results taken together show that the SimCLR-based method of contrastive learning is effective for histopathological classification. The approach yielded high quality features, which were transferable to the downstream task, even though it was trained without label access. Such findings emphasize the value of self-supervised contrastive methods in medical image analysis, especially when contacting data is scarce or expensive to procure.

CHAPTER 5 – Conclusion and future work

5.1 Summary of Work Done

This study proposed a holistic deep learning based framework to promote the diagnostic abilities of lung cancer histopathological studies under conditions of scarce annotated data. The research integrated two key strategies. designing and evaluating several custom CNN architectures for classification, with the conformity of contrastive learning as self-supervised pretraining method to enhance the ability to extract features from unlabeled data.

The classification task required the design of five unique CNN architectures, distinguished by different tiers, activation functions and pooling structures. This variability made it possible for a closer look at how model complexity and design choices influence classification accuracy, training stability, and computational efficiency. Interestingly, VGG-like deep architecture used in Model 2 results in the highest validation accuracy (98.6%) showing the benefits of depth in capturing complex histological patterns. Meanwhile, lighter models such as Model 4 yielded equivalent accuracy towards better training times and less parameters, which offered an alternative for deployment in resource-constrained environments.

However, with the practical constraints on the availability of annotated data in medical imaging, the current study also used contrastive learning with the SimCLR framework. This element was critical in assisting the model to learn strong and transferable feature representations from unlabeled histopathological images. The contrastive training phase, which was conducted before the supervised classification, has shown that there are valuable patterns that can be captured even when there are no direct links built between data scarcity and high performance modeling.

Besides, pre-trained models including EfficientNetB4, VGG19, and ResNet50 were also compared for comparative purposes. Based on their training with large datasets, such as ImageNet, in these cases, these models were used as a means of contextualization for the

effectiveness of customized architecture. Though VGG19 topped the table by achieving the highest classification accuracy of 98.9% it got there at the expensive of high computational cost that saw it use a lot of memory with high number of parameters. Forms of ensemble methods further propelled performance limits to achieve upto 99.96% accuracy.

The combination of contrastive learning was not an addition here but was essential to this study's purpose – it showed that using self-supervision, we can boost model generalization and performance in scenarios characterized by a lack or high cost of obtaining labeled medical data. It provided a particularly useful way of obtaining decent representations of complex histopathological slides, which helped in more reliable downstream classification even for small datasets.

To conclude, this research gives a practical and scalable pipeline to the lung cancer histopathological image analysis by fusing classification, contrastive learning, and transfer learning in a unified system. It demonstrates that high-accuracy diagnostic tools can be developed without being fully relying on big annotated datasets which means that AI-based pathology can be applied in reality on a real-world clinical setting much more readily and practicably.

5.2 Future Work

In the future, there are many ways this study can be expanded from. An approach that holds promising potential is the combination of hybrid transformer-CNN architectures to capture short-range and long-range feature relationships between the component of tissue morphology, improving knowledge of complex histopathological structures.

Furthermore, future work may take the current patch-based method and apply it to whole slide image classification (WSI) which will offer richer diagnostic context, while requiring more scalable architectures and more efficient memory management. While contrastive learning was useful in exploiting unlabeled data, it was also possible to enhance model robustness if semi-supervised or few-shot learning strategies were adopted in situations with a small availability of

annotated data.

Finally, availability of larger expertly labelled datasets would facilitate stronger supervised finetuning and better validation. Alternative effort may also investigate model interpretability and domain adaptation to support clinical reliability and deploy ability.

References

[1] C. S. Dela Cruz, L. T. Tanoue, and R. A. Matthay, "Lung cancer: epidemiology, etiology, and prevention," *Clin Chest Med,* vol. 32, no. 4, pp. 605-44, Dec, 2011.

[2] D. Yessenbayev, and Z. Khamidullina, "Epidemiology of Lung Cancer in Kazakhstan: Trends and Geographic Distribution," vol. 24, no. 5, pp. 1521-1532, May 1, 2023.

[3] J. A. Barta, C. A. Powell, and J. P. Wisnivesky, "Global Epidemiology of Lung Cancer," *Ann Glob Health,* vol. 85, no. 1, Jan 22, 2019.

[4] G. J. Amir, and H. P. Lehmann, "After Detection:: The Improved Accuracy of Lung Cancer Assessment Using Radiologic Computer-aided Diagnosis," *Academic Radiology,* vol. 23, no. 2, pp. 186-191, 2016/02/01/, 2016.

[5] M. J. van den Bent, "Interobserver variation of the histopathological diagnosis in clinical trials on glioma: a clinician's perspective," *Acta Neuropathologica,* vol. 120, no. 3, pp. 297-304, 2010/09/01, 2010.

[6] H.-Y. Chiu, H.-S. Chao, and Y.-M. Chen, "Application of Artificial Intelligence in Lung Cancer," *Cancers,* 14, 2022].

[7] D. Li, Z. Li, S. Li, H. Zhang, S. Yao, Y. Li, and J. Chen, "Development and Validation of a Prediction Model for Positive Findings of Preoperative Flexible Bronchoscopy in Patients with Peripheral Lung Cancer," *Current Oncology,* 30, 2023].

[8] R. S. Winokur, B. B. Pua, B. W. Sullivan, and D. C. Madoff, "Percutaneous lung biopsy: technique, efficacy, and complications," *Semin Intervent Radiol,* vol. 30, no. 2, pp. 121-7, Jun, 2013.

[9] X. Yin, H. Liao, H. Yun, N. Lin, S. Li, Y. Xiang, and X. Ma, "Artificial intelligence-based prediction of clinical outcome in immunotherapy and targeted therapy of lung cancer," *Seminars in Cancer Biology,* vol. 86, pp. 146-159, 2022/11/01/, 2022.

[10] G. G, J. Thimmiaraja, C. J. Shelke, G. Pavithra, V. K. Sharma, and D. Verma, "Deep Learning with Unsupervised and Supervised Approaches in Medical Image Analysis." pp. 1580-1584.

[11] M. Šarić, M. Russo, M. Stella, and M. Sikora, "CNN-based Method for Lung Cancer Detection in Whole Slide Histopathology Images." pp. 1-4.

[12] Y. Wang, Q. Zhang, L. Ying, and C. Zhou, "Deep Reinforcement Learning for Early Diagnosis of Lung Cancer," *Proceedings of the AAAI Conference on Artificial Intelligence,* vol. 38, no. 20, pp. 22410-22419, 03/24, 2024.

[13] D. Moitra, and R. K. Mandal, "Automated AJCC (7th edition) staging of non-small cell lung cancer (NSCLC) using deep convolutional neural network (CNN) and recurrent neural network (RNN)," *Health Inf Sci Syst,* vol. 7, no. 1, pp. 14, Dec, 2019.

[14] L. Wang, "Deep Learning Techniques to Diagnose Lung Cancer," *Cancers (Basel),* vol. 14, no. 22, Nov 13, 2022.

[15] J. R. Molina, P. Yang, S. D. Cassivi, S. E. Schild, and A. A. Adjei, "Non-small cell lung cancer: epidemiology, risk factors, treatment, and survivorship," *Mayo Clin Proc,* vol. 83, no. 5, pp. 584-94, May, 2008.

[16] S. Blandin Knight, P. A. Crosbie, H. Balata, J. Chudziak, T. Hussell, and C. Dive, "Progress and prospects of early detection in lung cancer," *Open Biol,* vol. 7, no. 9, Sep, 2017.

[17] I. Castiglioni, F. Gallivanone, P. Soda, M. Avanzo, J. Stancanello, M. Aiello, M. Interlenghi, and M. Salvatore, "AI-based applications in hybrid imaging: how to build smart

and truly multi-parametric decision models for radiomics," *European Journal of Nuclear Medicine and Molecular Imaging,* vol. 46, no. 13, pp. 2673-2699, 2019/12/01, 2019.

[18] M. Aharonu, and L. Ramasamy, "A Multi-Model Deep Learning Framework and Algorithms for Survival Rate Prediction of Lung Cancer Subtypes With Region of Interest Using Histopathology Imagery," *IEEE Access,* vol. 12, pp. 155309-155329, 2024.

[19] W. Ayadi, W. Elhamzi, I. Charfi, and M. Atri, "Deep CNN for Brain Tumor Classification," *Neural Processing Letters,* vol. 53, no. 1, pp. 671-700, 2021/02/01, 2021.

[20] E. A. Mohamed, T. Gaber, O. Karam, and E. A. Rashed, "A Novel CNN pooling layer for breast cancer segmentation and classification from thermograms," *PLOS ONE,* vol. 17, no. 10, pp. e0276523, 2022.

[21] P. S. Thakur, T. Sheorey, and A. Ojha, "VGG-ICNN: A Lightweight CNN model for crop disease identification," *Multimedia Tools and Applications,* vol. 82, no. 1, pp. 497-520, 2023/01/01, 2023.

[22] N. Behar, and M. Shrivastava, "ResNet50-Based Effective Model for Breast Cancer Classification Using Histopathology Images," *CMES-Computer Modeling in Engineering & Sciences,* vol. 130, no. 2, 2022.

[23] M. A. S. Al Husaini, M. H. Habaebi, T. S. Gunawan, M. R. Islam, E. A. A. Elsheikh, and F. M. Suliman, "Thermal-based early breast cancer detection using inception V3, inception V4 and modified inception MV4," *Neural Computing and Applications,* vol. 34, no. 1, pp. 333-348, 2022/01/01, 2022.

[24] M. Kim, Y. Kwon, J. Kim, and Y. Kim, "Image Classification of Parcel Boxes under the Underground Logistics System Using CNN MobileNet," *Applied Sciences,* 12, 2022].

[25] V. Ravi, H. Narasimhan, and T. D. Pham, "EfficientNet-Based Convolutional Neural Networks for Tuberculosis Classification," *Advances in Artificial Intelligence, Computation, and Data Science: For Medicine and Life Science,* T. D. Pham, H. Yan, M. W. Ashraf and F. Sjöberg, eds., pp. 227-244, Cham: Springer International Publishing, 2021.

[26] A. Teramoto, T. Tsukamoto, Y. Kiriyama, and H. Fujita, "Automated Classification of Lung Cancer Types from Cytological Images Using Deep Convolutional Neural Networks," *BioMed Research International,* vol. 2017, pp. 4067832, 2017/08/13, 2017.

[27] T. Kim, H. Chang, B. Kim, J. Yang, D. Koo, J. Lee, J. W. Chang, G. Hwang, G. Gong, N. H. Cho, C. W. Yoo, J. Y. Pyo, and Y. Chong, "Deep learning-based diagnosis of lung cancer using a nationwide respiratory cytology image set: improving accuracy and inter-observer variability," *Am J Cancer Res,* vol. 13, no. 11, pp. 5493-5503, 2023.

[28] N. Coudray, P. S. Ocampo, T. Sakellaropoulos, N. Narula, M. Snuderl, D. Fenyö, A. L. Moreira, and N. Razavian, "Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning," vol. 24, no. 10, pp. 1559-1567, Oct, 2018.

[29] S. Ahmed, B. Raza, L. Hussain, A. Aldweesh, A. Omar, M. S. Khan, E. Tageldin, and M. Nadim, "The Deep Learning ResNet101 and Ensemble XGBoost Algorithm with Hyperparameters Optimization Accurately Predict the Lung Cancer," *Applied Artificial Intelligence,* vol. 37, 06/03, 2023.

[30] C. Janßen, T. Boskamp, J. Le'Clerc Arrastia, D. Otero Baguer, L. Hauberg-Lotte, M. Kriegsmann, K. Kriegsmann, G. Steinbuß, R. Casadonte, J. Kriegsmann, and P. Maaß, "Multimodal Lung Cancer Subtyping Using Deep Learning Neural Networks on Whole Slide Tissue Images and MALDI MSI," *Cancers,* 14, 2022].

[31] Z. Li, J. Zhang, T. Tan, X. Teng, X. Sun, H. Zhao, L. Liu, Y. Xiao, B. Lee, Y. Li, Q. Zhang, S. Sun, Y. Zheng, J. Yan, N. Li, Y. Hong, J. Ko, H. Jung, Y. Liu, Y. c. Chen, C. w. Wang, V. Yurovskiy, P. Maevskikh, V. Khanagha, Y. Jiang, L. Yu, Z. Liu, D. Li, P. J. Schüffler, Q. Yu, H. Chen, Y. Tang, and G. Litjens, "Deep Learning Methods for Lung Cancer Segmentation in Whole-Slide Histopathology Images—The ACDC@LungHP Challenge 2019," *IEEE Journal of Biomedical and Health Informatics,* vol. 25, no. 2, pp. 429-440, 2021.

[32] Y. Wu, Y. Wu, J. Wang, Z. Yan, L. Qu, B. Xiang, and Y. Zhang, "An optimal tumor marker group-coupled artificial neural network for diagnosis of lung cancer," *Expert Systems with Applications,* vol. 38, no. 9, pp. 11329-11334, 2011/09/01/, 2011.

[33] P. Khosravi, E. Kazemi, M. Imielinski, O. Elemento, and I. Hajirasouliha, "Deep Convolutional Neural Networks Enable Discrimination of Heterogeneous Digital Pathology Images," *EBioMedicine,* vol. 27, pp. 317-328, 2018/01/01/, 2018.

[34] G. V. S. Sudhamsh, S. Girisha, and R. Rashmi, "Semi-supervised tissue segmentation from histopathological images with consistency regularization and uncertainty estimation," *Scientific Reports,* vol. 15, no. 1, pp. 6506, 2025/02/22, 2025.

[35] R. Azad, M. Heidari, M. Shariatnia, E. Khodapanah Aghdam, S. Karimijafarbigloo, E. Adeli, and D. Merhof, "TransDeepLab: Convolution-Free Transformer-Based DeepLab v3+ for Medical Image Segmentation," pp. 91-102, 2022.

[36] J. Wu, Y. Pan, Q. Ye, J. Zhou, and F. Gou, "Intelligent cell images segmentation system: based on SDN and moving transformer," *Scientific Reports,* vol. 14, no. 1, pp. 24834, 2024/10/22, 2024.

[37] Y. Liu, H. Shi, Q. He, Y. Fu, Y. Wang, Y. He, A. Han, and T. Guan, "Invasive carcinoma segmentation in whole slide images using MS-ResMTUNet," *Heliyon,* vol. 10, no. 4, pp. e26413, 2024/02/29/, 2024.

[38] P. Bándi, M. Balkenhol, B. Ginneken, J. van der Laak, and G. Litjens, "Resolution-agnostic tissue segmentation in whole-slide histopathology images with convolutional neural networks," *PeerJ,* vol. 7, pp. e8242, 12/17, 2019.

[39] F. Hou, W. Lei, S. Li, J. Xi, M. Xu, and J. Luo, "Improved Mask R-CNN with distance guided intersection over union for GPR signature detection and segmentation," *Automation in Construction,* vol. 121, pp. 103414, 2021/01/01/, 2021.

[40] D. Karimi, and S. E. Salcudean, "Reducing the Hausdorff Distance in Medical Image Segmentation With Convolutional Neural Networks," *IEEE Transactions on Medical Imaging,* vol. 39, no. 2, pp. 499-513, 2020.

[41] L. Lei, S. Sun, Y. Zhang, H. Liu, and W. Xu, "PSIC-Net: Pixel-Wise Segmentation and Image-Wise Classification Network for Surface Defects," *Machines,* 9, 2021].

[42] C. P. Jayapandian, Y. Chen, A. R. Janowczyk, M. B. Palmer, C. A. Cassol, M. Sekulic, J. B. Hodgin, J. Zee, S. M. Hewitt, J. O'Toole, P. Toro, J. R. Sedor, L. Barisoni, and A. Madabhushi, "Development and evaluation of deep learning-based segmentation of histologic structures in the kidney cortex with multiple histologic stains," *Kidney Int,* vol. 99, no. 1, pp. 86-101, Jan, 2021.

[43] G. M. Kalkan, B. Dönmez, Y. C. Tok, S. Gürel, E. AkdaĞliŞ, H. Çapkan, M. Z. Kaya, D. Kandaz, M. K. Uçar, and S. N. Şenol, "Diagnosis of Lung Cancer with Hybrid Artificial Intelligence Method." pp. 1-4.

[44] S. Das, G. L, S. J. Prakash, N. S. Dey, J. Panuganti, and R. Poojitha, "Lung Cancer Detection and Classification using Transfer Learning with Pre-trained VGG19 Convolutional Neural Networks." pp. 1-6.

[45] A. Alomar, M. Alazzam, H. Mustafa, and A. Mustafa, "Lung Cancer Detection Using Deep Learning and Explainable Methods." pp. 1-4.
[46] B. Gajera, S. R. Kapil, D. Ziaei, J. Mangalagiri, E. Siegel, and D. Chapman, "CT-Scan Denoising Using a Charbonnier Loss Generative Adversarial Network," *IEEE Access,* vol. 9, pp. 84093-84109, 2021.
[47] D. Hişam, and E. Hişam, "Deep learning models for classifying cancer and COVID-19 lung diseases." pp. 1-4.
[48] M. S. Ahmed, K. N. Iqbal, and M. G. R. Alam, "Interpretable Lung Cancer Detection using Explainable AI Methods." pp. 1-6.
[49] M. Mamun, A. Farjana, M. A. Mamun, and M. S. Ahammed, "Lung cancer prediction model using ensemble learning techniques and a systematic review analysis." pp. 187-193.
[50] R. P. R. Kumar, S. Polepaka, D. Likithasree, and S. Keerthika, "An Investigation on CNN-based Lung Cancer Prediction Method." pp. 1-5.
[51] P. S. Bharathi, and T. P. K. R, "Neural Network based Earlier Stage Lung Cancer Prediction Scheme with Differential Learning Assistance." pp. 01-06.
[52] F. Ciompi, K. Chung, S. J. van Riel, A. A. A. Setio, P. K. Gerke, C. Jacobs, E. T. Scholten, C. Schaefer-Prokop, M. M. W. Wille, A. Marchianò, U. Pastorino, M. Prokop, and B. van Ginneken, "Towards automatic pulmonary nodule management in lung cancer screening with deep learning," *Scientific Reports,* vol. 7, no. 1, pp. 46479, 2017/04/19, 2017.
[53] C. Zhang, X. Sun, K. Dang, K. Li, X. w. Guo, J. Chang, Z. q. Yu, F. y. Huang, Y. s. Wu, Z. Liang, Z. y. Liu, X. g. Zhang, X. l. Gao, S. h. Huang, J. Qin, W. n. Feng, T. Zhou, Y. b. Zhang, W. j. Fang, M. f. Zhao, X. n. Yang, Q. Zhou, Y. l. Wu, and W. z. Zhong, "Toward an Expert Level of Lung Cancer Detection and Classification Using a Deep Convolutional Neural Network," *The Oncologist,* vol. 24, no. 9, pp. 1159-1165, 2019.
[54] D. Ardila, A. P. Kiraly, S. Bharadwaj, B. Choi, J. J. Reicher, L. Peng, D. Tse, M. Etemadi, W. Ye, G. Corrado, D. P. Naidich, and S. Shetty, "End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography," *Nature Medicine,* vol. 25, no. 6, pp. 954-961, 2019/06/01, 2019.
[55] P. Petousis, A. Winter, W. Speier, D. R. Aberle, W. Hsu, and A. A. T. Bui, "Using Sequential Decision Making to Improve Lung Cancer Screening Performance," *IEEE Access,* vol. 7, pp. 119403-119419, 2019.
[56] M. Koizumi, K. Motegi, M. Koyama, M. Ishiyama, T. Togawa, T. Makino, Y. Arisaka, and T. Terauchi, "Diagnostic performance of a computer-assisted diagnostic system: sensitivity of BONENAVI for bone scintigraphy in patients with disseminated skeletal metastasis is not so high," *Annals of Nuclear Medicine,* vol. 34, no. 3, pp. 200-211, 2020/03/01, 2020.
[57] J. Li, D. Dong, M. Fang, R. Wang, J. Tian, H. Li, and J. Gao, "Dual-energy CT–based deep learning radiomics can improve lymph node metastasis risk prediction for gastric cancer," *European Radiology,* vol. 30, no. 4, pp. 2324-2333, 2020/04/01, 2020.
[58] X. Yang, L. Wu, W. Ye, K. Zhao, Y. Wang, W. Liu, J. Li, H. Li, Z. Liu, and C. Liang, "Deep Learning Signature Based on Staging CT for Preoperative Prediction of Sentinel Lymph Node Metastasis in Breast Cancer," *Academic Radiology,* vol. 27, no. 9, pp. 1226-1233, 2020/09/01/, 2020.
[59] Y. Gao, Z.-D. Zhang, S. Li, Y.-T. Guo, Q.-Y. Wu, S.-H. Liu, S.-J. Yang, L. Ding, B.-C. Zhao, S. Li, Y. Lu, and Y.-Y. Ji, "Deep neural network-assisted computed tomography diagnosis of metastatic lymph nodes from gastric cancer," *Chinese Medical Journal,* vol. 132, no. 23, pp. 2804-2811, 2019/12/05, 2019.

[60] M. Dohopolski, L. Chen, D. J. Sher, and J. Wang, “Predicting Lymph Node Metastasis in Patients with Oropharyngeal Cancer by Convolutional Neural Networks with associated Epistemic Uncertainty,” *International Journal of Radiation Oncology, Biology, Physics*, vol. 105, no. 1, pp. S122, 2019.

[61] Y. Ariji, M. Fukuda, Y. Kise, M. Nozawa, Y. Yanashita, H. Fujita, A. Katsumata, and E. Ariji, “Contrast-enhanced computed tomography image assessment of cervical lymph node metastasis in patients with oral cancer by using a deep learning system of artificial intelligence,” *Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology*, vol. 127, no. 5, pp. 458-463, 2019/05/01/, 2019.

[62] Y. P. Zhou, S. Li, X. X. Zhang, Z. D. Zhang, Y. X. Gao, L. Ding, and Y. Lu, “[High definition MRI rectal lymph node aided diagnostic system based on deep neural network],” *Zhonghua wai ke za zhi [Chinese journal of surgery]*, vol. 57, no. 2, pp. 108-113, 2019/02//, 2019.

[63] N. Nawreen, U. Hany, and T. Islam, "Lung Cancer Detection and Classification using CT Scan Image Processing." pp. 1-6.

[64] A. S. Sakr, "Automatic Detection of Various Types of Lung Cancer Based on Histopathological Images Using a Lightweight End-to-End CNN Approach." pp. 141-146.

[65] X. Wang, H. Chen, C. Gan, H. Lin, Q. Dou, E. Tsougenis, Q. Huang, M. Cai, and P. A. Heng, “Weakly Supervised Deep Learning for Whole Slide Lung Cancer Image Analysis,” *IEEE Transactions on Cybernetics*, vol. 50, no. 9, pp. 3950-3962, 2020.

[66] N. Tasnim, K. R. Noor, M. Islam, M. N. Huda, and I. H. Sarker, "A Deep Learning Based Image Processing Technique for Early Lung Cancer Prediction." pp. 1060-1064.

[67] Y. Balagurunathan, A. Beers, M. Mcnitt-Gray, L. Hadjiiski, S. Napel, D. Goldgof, G. Perez, P. Arbelaez, A. Mehrtash, T. Kapur, E. Yang, J. W. Moon, G. B. Perez, R. Delgado-Gonzalo, M. M. Farhangi, A. A. Amini, R. Ni, X. Feng, A. Bagari, K. Vaidhya, B. Veasey, W. Safta, H. Frigui, J. Enguehard, A. Gholipour, L. S. Castillo, L. A. Daza, P. Pinsky, J. Kalpathy-Cramer, and K. Farahani, “Lung Nodule Malignancy Prediction in Sequential CT Scans: Summary of ISBI 2018 Challenge,” *IEEE Transactions on Medical Imaging*, vol. 40, no. 12, pp. 3748-3761, 2021.

[68] D. Jayaraj, and S. Sathiamoorthy, "Random Forest based Classification Model for Lung Cancer Prediction on Computer Tomography Images." pp. 100-104.

[69] S. H. Begum, M. I. Baig, M. A. Hussain, and M. A. Muqeet, "A Lightweight Deep Learning Model for Automatic Diagnosis of Lung Cancer." pp. 1-5.

[70] W. Ausawalaithong, A. Thirach, S. Marukatat, and T. Wilaiprasitporn, "Automatic Lung Cancer Prediction from Chest X-ray Images Using the Deep Learning Approach." pp. 1-5.

[71] J. Cañada, E. Cuello, L. Téllez, J. M. García, F. J. Velasco, and J. Cabrera, "Assistance to lung cancer detection on histological images using Convolutional Neural Networks." pp. 1-4.

[72] P. P. Massion, S. Antic, S. Ather, C. Arteta, J. Brabec, H. Chen, J. Declerck, D. Dufek, W. Hickes, T. Kadir, J. Kunst, B. A. Landman, R. F. Munden, P. Novotny, H. Peschl, L. C. Pickup, C. Santos, G. T. Smith, A. Talwar, and F. Gleeson, “Assessing the Accuracy of a Deep Learning Method to Risk Stratify Indeterminate Pulmonary Nodules,” *Am J Respir Crit Care Med*, vol. 202, no. 2, pp. 241-249, Jul 15, 2020.

[73] R. Y. Bhalerao, H. P. Jani, R. K. Gaitonde, and V. Raut, "A novel approach for detection of Lung Cancer using Digital Image Processing and Convolution Neural Networks." pp. 577-583.