AUTOMATED METHODS TO DETECT AND CLASSIFY LUNG & COLON CANCER USING IMAGE PROCESSING AND DEEP LEARNING ALGORITHM

BY

Nesma Huda
ID: 221-25-117

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Master of Science in Computer Science and Engineering

Supervised By

Dr. Md Zahid Hasan
Associate Professor
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

JANUARY 2023

APPROVAL

This thesis titled "Automated methods to detect and classify Lung & Colon Cancer using Image processing and Deep Learning algorithm", submitted by Nesma Huda to the Department of Computer Science and Engineering, Daffodil International University, has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of M.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 17 January 2023.

BOARD OF EXAMINERS

17/01/23

Dr. Touhid Bhuiyan, PhD — **Chairman**
Professor and Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Ms. Nazmun Nessa Moon — **Internal Examiner**
Associate Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

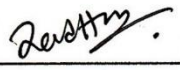
Dr. Fizar Ahmed — **Internal Examiner**
Associate Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Md. Safaet Hossain — **External Examiner**
Associate Professor & Head
Department of Computer Science and Engineering
City University

DECLARATION

I hereby declare that, this thesis has been done by me under the supervision of **Dr**. **Md Zahid Hasan, Associate Professor, Department of CSE,** Daffodil International University. I also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

Supervised by:



Dr. Md Zahid Hasan
Associate Professor
Department of CSE
Daffodil International University

Submitted by:

Nesma Huda
ID: 221-25-117
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First I express my heartiest thanks and gratefulness to Almighty Allah for His divine blessing which makes me possible to complete the final year project/internship successfully.

I really grateful and wish my profound indebtedness to **Dr. Md Zahid Hasan**, **Associate professor**, Department of CSE, Daffodil International University, Dhaka, deep knowledge & keen interest of my supervisor in the field of Machine Learning to carry out this project. His endless patience, scholarly guidance, continual encouragement , constant and energetic supervision, constructive criticism , valuable advice ,reading many inferior draft and correcting them at all stage have made it possible to complete this project.

I would like to express my heartiest gratitude to **Dr. Touhid Bhuiyan**, Head**,** Department of CSE, for his kind help to finish our project and also to other faculty members and the staffs of CSE department of Daffodil International University.

Finally, I must acknowledge with due respect the constant support and patients of my parents.

ABSTRACT

One of the most hazardous and severe diseases that people experience globally is colon and lung cancer, which has spread to become a common medical issue. It is very important to make a reliable and early discovery to reduce the danger of mortality. The difficulty of the work ultimately depends on the histopathologists' experience. Under-prepared histologists may potentially endanger the patient's life. Recent times have seen a rise in the popularity of deep learning, which is now appreciated in the interpretation of medical imaging. In order to diagnose lung and colon cancer utilizing histopathological image dataset and more effective augmentation approaches, this research aims to leverage and transform the present pre-trained CNN-based architecture. In this study, the LC25000 dataset is used to train three different pre-trained Convolutional Neural Network models: EffecientNetB7, ResNet50 and VGG16. The model performances are evaluated based on accuracy, precision, recall & f1-score. The findings illustrate that all three models produced impressive outcomes, ranging from 93% to 98% accuracy.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
INTRODUCTION

1.1 Introduction

Cancer refers to any of the many diseases that are characterized by the expansion of abnormal cells that proliferate uncontrollably and have the potential to infiltrate and destroy healthy body tissue. The propensity of cancer to spread across the human body is a very regular occurrence. According to World Health Organization, an estimated 10 million deaths, or one in every six, is caused by cancer in 2020, making it the second highest cause of death in the globe. [1]. While colon & rectum cancer accounts for 1.93 million cases and 916K deaths, lung cancer accounts for 2.21 million cases and 1.8 million deaths [1]. However, survival rates for different forms of cancer are constantly increasing day by day as a consequence of advances in cancer screening, treatment, and prevention.

Small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC) are the two variations of lungs cancer that manifest and spread abruptly. 15% of instances of lung cancer are Small cell lung cancer (SCLC), a lethal tumor made up of cells with neuroendocrine properties. Another three major pathologic sub-types of non-small cell lung cancer NSCLC are Large cell carcinoma, Squamous cell carcinoma, and Adenocarcinoma, which account for the remaining 85% of cases [3]. The phrase colorectal cancer, which unites colon and rectal cancers by beginning in the rectum, is typically used to refer to colon disease. 96% of all instances of colorectal cancer are adenocarcinomas [4]. In recent times, AI technologies are becoming acclaimed due to their capacity to analyze data more quickly and aid in decision-making process. Artificial intelligence (AI) has a subset known as machine learning (ML) that enables computers to master a specified activity via experience with the data sets they are exposed to, without being explicitly programmed [7]. In biomedical applications, deep learning (algorithms) are utilized for the prediction and classification of a variety of visuals and signals.

Machines can now process massive multi-dimensional data like anatomical images and recordings thanks to deep learning (DL) methods.

DL is a branch of ML that constructs layers of algorithms to simulate the functionality and structure of the human brain in a "artificial neural network." Artificial Intelligence in healthcare has two primary subfields: virtual and physical. The virtual discipline of Artificial Intelligence in medicine is represented by machine learning (ML) and deep learning (DL) [3]. By repeatedly performing certain tasks, ML forms predictive and enumerative models employing computer statistics and analytics. Unsupervised and supervised learning both constitute machine learning. Unsupervised machine learning, like the term suggests, finds groupings based on shared characteristics in the data and requires no prior knowledge when feeding the data. In supervised machine learning, input (individual descriptions) and output (an outcome of interest) data are fed to the system.

Based on the information streams, the computer ultimately defines newer input/output pairs. Generally speaking, there are four types of machine learning tasks: supervised, unsupervised, reinforcement learning, and active learning. Input data with target labels are applied in supervised learning to identify patterns. Unsupervised learning entails picking out patterns from the input data without having target labels already established. Reinforcement learning is the process of training intelligent agents to achieve substantial performance. Multiple data types are received as input by a deep learning system, forming layers of data from which it retrieves the data points of interest.

Deep learning can be both supervised and unsupervised. The most well-known modalities are trained via supervised learning, where datasets are constructed of input data and associated output data labels. There are 2 steps of deep leaning: pre-training and fine-tuning. In the first stage, the Deep Learning model generates outputs unsupervised while attempting to learn the underlying data-distribution. In the second stage, the output produced is adjusted for the particular job to ensure maximizing performance.

The deep neural network (DNN), recurrent neural network (RNN), and convolutional neural network (CNN) are further categories for Deep Learning. DNN uses artificial neural networks in a feedforward fashion. CNNs are particularly crucial for efficiently identifying patterns from visual pixels with minimal pre-processing. RNNs function well for time series data since they are crucial for detecting sequential data in temporal sequence. The second area of artificial intelligence in medicine is physical, which includes robots and medical equipment. In-depth study is being done to employ AI applications in the medical industry, which might offer previously unheard-of chances to enhance healthcare quality.

This article mostly examines non-small cell lung tumors (NSCLC). The goal is to streamline and automate the task of histopathologists, which includes the prime task such as: to examine cells and tissues under a microscope and identify irregularities. In recent years, a great deal of effort has gone into scanning the entire tissue or cell slide and saving this as a digital computerized image. A huge amount of WSI (full slide pictures) are being gathered as a consequence. There have been several attempts to investigate WSIs by leveraging ML algorithms for diagnostic purpose.
However, a great deal of the image categorization activities is shifted towards DL (deep learning).

1.2 Problem Statement

Colon and lung cancers cause the greatest number of fatalities among the organs that are most often afflicted. Worldwide, colon cancer causes 9.2% of all cancer-related fatalities, compared to lung cancer's 18.4% prevalence [1]–[2]. About 17% of cases of lung and colon cancer have simultaneous occurrence. Although this frequency is unusual, cancer cells are extremely prone to spread across these two organs in the absence of an early detection [3]. The only methods for decreasing cancer mortality are adequate care and early detection [4].

Undoubtedly, the prior an individual is detected, the efficient the administration, and the superior opportunity of the patient's recovery and sustenance.
Different tests like imaging sets (x-ray, CT scan), Sputum cytology, and tissue sampling (biopsy) are executed to search for malignant growth cells and preclude further potential circumstances. The inspection of the microscopic histopathology plates by skilled professional pathologists during the biopsy is crucial for founding the diagnosis [5] and defining the different kinds and sub-kinds of malignant growths [6].

This study exclusively relies on histopathological images to detect colon and lung cancers. Healthcare professionals frequently employ histopathological imagery for diagnosis, and these images are crucial for estimating the survival of patients. Traditionally, they have gone through a prolonged procedure to diagnose cancer by reviewing histopathological images; however, this process currently can be accomplished in short time and exertion with the existing deep learning technology [3].

1.3 Research Objectives

a) To assist medical professionals in identifying lung & colon cancer in an effective, precise, and insightful manner, and also to contribute to the new knowledge of AI-based applications.
b) To explore research gaps in the prevailing vision-based deep learning systems to accurately detect various types of colons and lung cancer.
c) To apply a simple, straightforward, and effective vision-based transfer learning method to raise the precision of the lung and colon cancer subclasses

1.4 Research Questions

a) How can we explore the inadequacies of the current deep learning image-based classification systems which can precisely diagnose various classes of colon and lung cancer?

b) How can we build a deep learning image-based method that will increase the accuracy of appropriate differentiating for the classes of various lung and colon cancer-affected cells?

1.5 Report Layout

Chapter 1 presents the research introduction, objectives, and key research questions.

Chapter 2 highlights a detailed review of the related literature.

Chapter 3 describes the proposed methodology with a detailed description.

Chapter 4 explains the result analysis and comparison with existing work.

Chapter 5 concludes the present research along with a direction for future work.

CHAPTER 2
LITERATURE REVIEW

2.1 Related works

Regarding the categorization of the histopathological images of various cancer forms, including lung, breast, prostate, skin and blood cancer, a significant amount of work has been done. Grayscale mean, GR variance, and sixteen texture characteristics were among the eighteen common features that Liping Jiao et al. [9] retrieved using the Gray-Level Cooccurrence Matrix (GLCM) method. The Support Vector Machine [23] based classifier was applied, and on 60 colon tissue pictures that were squarely divided into the 2-two classes, it was able to attain accuracy, F1-score, and recall as 96.67%, 83.33%, and 89.51% respectively. T. Wan et al. [11] produced 3400 characteristics from 48 breast tissue pictures using textural and architectural-based features (18-eighteen benign and30-thirty malignant).

Additionally, spectral clustering was carried out for reducing the feature set's size. Then, breast tissue images with low and high levels of malignancy are binary classified on malignant and non- malignant breast tissue images using an SVM classifier. [3] Results reveal that 95.8% accuracy was achieved in differentiating between cancerous and non-cancerous developments utilizing textural-based features. It was possible to distinguish between cancers with 93.3% accuracy by adding architectural elements. [11] A feature extraction method that statistically describes the geometrical characteristics of the components of colon tissues was proposed by S. Rathore et al. A hybrid feature set is created using traditional features including texture, morphological, Elliptic Fourier descriptors (EFDs), and Scale-Invariant Feature Transform (SIFT), and then, 174 colon biopsy pictures are classified using SVM [12], with a 98% accuracy rate.

The lack of clinical data, which leads to model overfitting [15], is the largest issue that academics throughout the world are dealing with. F. A. Spanhol et al. [13] increased the amount of input photos on the BreakHis dataset using a variety of data augmentation

techniques to prevent overfitting. The AlexNet [23] model was subsequently trained from scratch, yielding a 90% accuracy rate. While training an InceptionV3 [23] model, A. Krizhevsky et al. [7] empirically identify the refinement tuning technique during examining the impact of transfer learning [6]. Combining several datasets, such as the DDSM database, the INbreast database, and the BCDR database, yields a bigger dataset with an accuracy of roughly 98.845%. Assessment of pre-trained CNN models on the LC25000 dataset using overall of four dataset, specifically BreakHis, ICIAR, PCam , and Bioimaging. T. Araújo et al. [15] used well-known Pre-trained Convolutional Neural Network architectures, especially VGG19 [23], MobileNetV2 [23], and DenseNet201 [23], with a purpose of assisted identification of breast cancer cells. With the exception of the Bioimaging dataset, that achieves just 83.1% accuracy, the result demonstrates that greater than 95% accuracy is attained on all datasets.

L. D. Nguyen et al. [18] created a deep convolutional neural network (DCNN) model with three convolutional layers, three pooling layers, and two fully connected layers for multiclass classification on 298 lung pictures. Numerous data augmentation methods, including rotation, zooming, and flipping, were implemented to prevent overfitting [19]. Adenocarcinoma, squamous cell carcinoma, and little cell carcinoma accuracy in the findings obtained using the enhanced pictures is 89.0%, 60.0%, and 70.3%, respectively, for an overall accuracy of 71.1%. We note from the review that only a few sets of research were conducted to categorize unique malignancy sub forms. The more Pre-Trained CNN models that are available, the scarcer models need to be implemented for comparative analytics such as to classify different subclasses of lung / colon / breast cancer, etc.

U. J. Reddy et al. [20] employed a CNN prototype to categorize histopathological imageset of lungs cancer into three categories where 97.2% was the classification success rate. Homology-based techniques and machine learning techniques were utilized by S. Makaju et al. [22] to divide lung tissue images into three categories. 99.43% total classification accuracy was attained. Using a deep learning-based approach, D. Lin et al. [21] categorize pictures of the colon and lung histopathology.

For the purpose of extracting 4 feature sets for image classification, they applied domain transformations of 2 sorts. The final classification results were then determined by combining the characteristics of the two categories. They have achieved 96.33% accuracy rate.

By using a shallow neural network architecture, S. Garg et al. [14] classified colon and lung tumors founded on histopathological imageries. They successfully diagnose lung and colon cancers with accuracy rates of 97% and 96%, respectively. G. A. Singh [24] classified the histopathological pictures of colon and lung cancers by first training the imageset with the Darknet-19 model, then obtaining the feature sets, and then using 2 optimization techniques to the feature sets to remove the ineffective features.

2.2 Scope of the Problem

The automation of assessments and operations by AI, according to researchers, will help to save a substantial amount of time and cost. It will also assist less experienced radiologists in ruling out anything that can be confused for cancer and detecting lung or colon cancer when it is present. Many of the developments have been achieved in cancer imaging techniques, while there are infinite potential applications. Doctors employ imaging tests in a variety of ways, from taking x-rays of whole organs to using microscopes to capture imageries of cancer cells. They can be used to determine the stage of a tumor, determine the cancer's early stages, evaluate the treatment’s functionality, and determine the return of the cancer following treatment.

Over the past several years, researchers have developed AI algorithms which might make cancer imaging more effective, precise, and insightful. And that has caused a great deal of enthusiasm. In addition, AI might simplify and improve the dependability of difficult jobs that need "a human (a radiologist, a dermatologist, or a pathologist) providing an interpretation of an image." Deep learning is making great strides in these areas.

But the potential for AI to surpass individual human capacity is what scientists are most intrigued about. Artificial Intelligence can "see" things that humans cannot with normal eyes and can identify intricate patterns and connections among extremely varied types of

data. This is something that AI excels at, often outperforming humans at. However, it is frequently unclear how the AI comes to its decision, making it challenging for physicians and research professionals to determine whether the tool is functioning efficiently.

In clinical tests, many AI techniques have proven to improve the identification of adenomas, which are precancerous growths that can develop into colon cancer. Since only a small percentage of adenomas progress to cancer, several expert professionals are worried that emerging AI systems may compel many patients to undergo unnecessary therapies and extra testing.

To assist doctors in detecting lung cancer on CT scans and histopathology pictures, many deep learning AI models have been created. There is a significant proportion of false-positive test results that imply a person has lung cancer when they don't because some non-cancerous abnormalities might appear on CT scans to be very identical to cancer.

Theoretically, by better differentiating lung cancer from noncancerous changes on histopathological graphics, Artificial Intelligence may lower the rate of false positives and spare a lot of patients from superfluous stress, additional tests, and treatments. For instance, a group of research professionals created a deep learning algorithm to discover lung cancer and precisely avert other modifications which resemble cancer. In test's conducted in the lab, the algorithm was effective at both at both detecting cancer and disregarding non-cancerous alterations that resembled the ailment.

2.3 Challenges

The following are some research issues specific to this study:

a) **Unfiltered Image Processing**: Occasionally, images that were compiled from numerous mediums had low or high contrast, or they were full of noises. Therefore, the complexity is to accurately provide noise-free, contrast-enhanced images for categorization.

b) **Select Deep Learning Approach**: Several researchers successfully accomplish the challenges by utilizing various deep learning approaches. Therefore, choosing the best deep learning approach that can accurately diagnose many types of lungs and colon cancer was a challenge.

c) **Complex Architecture**: Due to the complex network design and mathematical computations included in the algorithms utilized, deep learning models are exceedingly challenging to comprehend.

d) **Absence of Feature Engineering**: Feature engineering, which is the act of converting unprocessed data into convenient features, assists in improved comprehending the model and boosts its prediction potential, improving the classification model's interpretability. In fact, feature engineering is essential in the medical and diagnostic industry because it enables doctors to understand the significance and influence of each feature on the categorization of cancer subclasses. This knowledge enables them to make life-changing decisions. On the other hand, deep learning models make autonomous decisions throughout the investigation at every level of the neural network without our ability to comprehend what is occurring inside the model schema.

e) **Extensive Time Consumption:** The training process for a deep learning model might take several weeks or just a few hours. This is because there are lots of parameters and hyperparameters. They need to be learnt sturdily in deep learning models

f) **Expensive Computational Resources:** Due to the volume of data processed, deep learning systems demand substantially greater prevailing hardware and computational resources with extremely maximized efficiency. The rising usage of graphics processing units (GPU), which are quite costly, is a result of this requirement for power.

g) **Accuracy Improvement:** Enhancing the machine learning model's accuracy and choosing the best fit deep learning model to achieve performance metrics are two more challenging issues.

CHAPTER 3
RESEARCH METHODOLOGY

3.1 Working Procedure

The entire working procedure of our research methodology is broken down into five steps. Followings are given below:

a) Dataset Collection & Preparation
b) Image Pre-processing
c) Train Test Split
d) Apply Deep Learning Algorithm (CNN based Transfer Learning)
e) Performance Evaluation

The whole working process, from the gathering and processing of picture datasets through performance evaluation. The steps for this process are shown in Figure 1 and described in detail in the following sections:

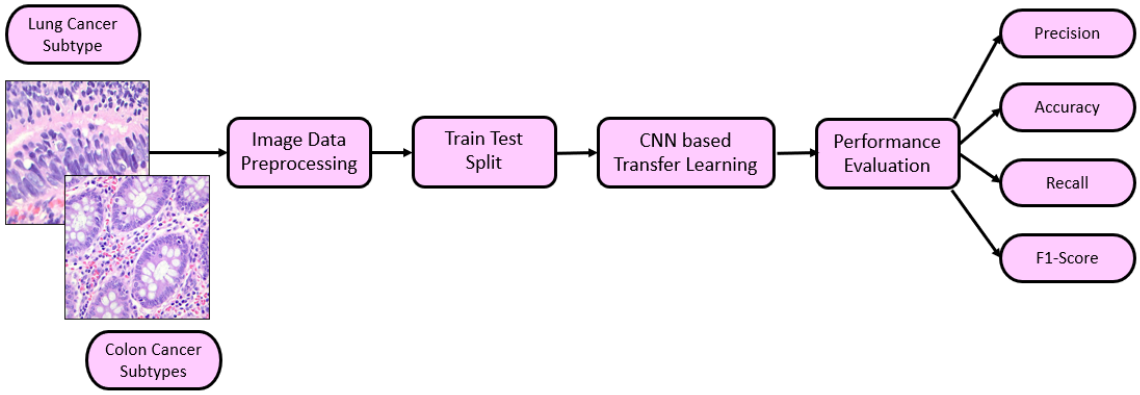


Figure 1: Working method of the proposed model

3.2 Dataset Collection & Preparation

We have used the LC25000 dataset for this research. The LC25000 dataset stands for Lung and Colon Cancer Histopathological Image Dataset, which was circulated in 2020. The LC25000 collection comprises 25,000 pictures, or 5,000 images for each of 5 groups of colon and lung tissues. Each picture has a dimension of 768 by 768 pixels. The five forms include benign colon tissue, lung adenocarcinoma, benign lung tissue, lung squamous cell carcinoma, and colon adenocarcinoma. Colon adenocarcinoma, which makes up more than 95% of all occurrences of colon cancer, is the most common kind. It is created when a polyp known as an anadenoma develops inside the large intestine before turning cancerous. Lung Adenocarcinoma, which account for around 60% of all occurrences of lung cancer, often develop in the glandular cells in the outer part of the lung before spreading to the lung's alveoli. The second most frequent kind of lung cancer, lung squamous cell carcinoma, occurs in the lungs' bronchi or airways and accounts for around 30% of all cases. Figure 2 shows an illustration of histopathological imageries of these 5 types of colon and lung tissues taken specially from the LC25000 dataset.

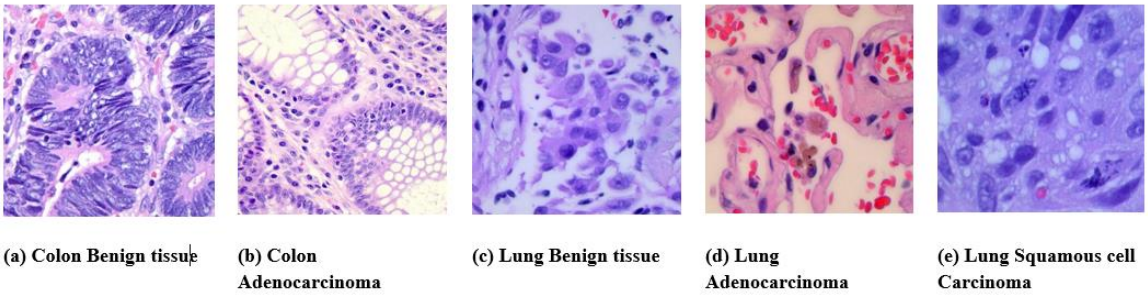


Figure 2. (a) Raw Sample images of: (a) Colon Benign tissue, (b) Colon Adenocarcinoma, (c) Lung Benign tissue, (d) Lung Adenocarcinoma, and (e) Lung Squamous cell Carcinoma collected from the LC25000 dataset.

3.3 Image Pre-processing

Pre-processing is a term used to describe actions on imageries at their most basic level; both the input and output are intensity imageries. These pictures are of the similar type as the primary data the sensor recorded, with an intensity picture typically being represented

by a matrix of imagery functional values (brightness's). Although geometric transformations of imageries (such as rotation, scaling, and translation) are categorized here as pre-processing methods since identical techniques are employed, the objective of pre-processing is an improvement of the imagery data that suppresses unintentional alterations or enhances some photo features crucial for further processing. Although pre-processing techniques are ranked according to how well they handle mathematical changes to pictures (such as rotation, scaling, and translation), these methodologies are used in this case for better classification result.

3.3.1 Image Augmentation

The imageries fetched from LC25000 have size of 768*768. All images are cropped from their original 768*768 pixel sizes to squares of 224*224 pixels. The LC25000 image dataset previously included the augmentations of left and right rotations (up to 25 degrees, 1.0 probability) and horizontal and vertical flips (0.5 probability) for the pictures. However, this study employs an advanced augmentation pipeline from the imgaug library to improve the performance of Pre-trained Convolutional Neural Network models. A colon cancer and lung cancer effected visual and a set of 64 enhanced photos using the imgaug library are shown in Fig. 4, respectively.

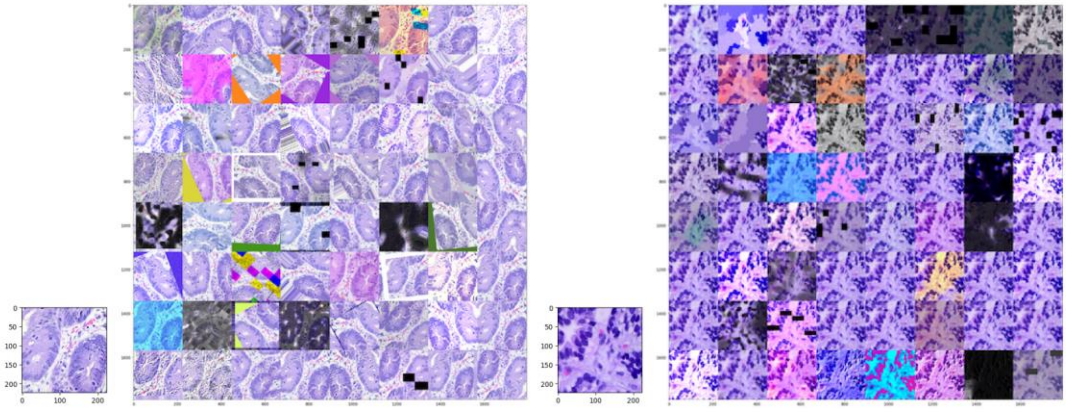


Figure 3. (a) Raw and Augmented images using imgaug library: colon and lung carcinoma

3.4 Train-Test Split

The lung and colon cancer histopathology imaging collection, also known as LC25000, was employed in this study [2]. The dataset was made public in order to meet the demand for and scarcity of freely available datasets in clinical imaging [2]. Each of the five classes - lung adenocarcinoma, lung squamous cell cancer, lung benign, colon carcinoma, and colon benign—contains 10,000 colon and 15,000 lung images. Images were resized to 224x224 pixels. The process entails splitting the dataset into two subgroups where 20% of the dataset was utilized for testing and the residual 80% was utilized for training. These divided train test data were randomly chosen. Then, a transfer learning algorithm performs classification on the image dataset based on the selected features

3.5 Apply Deep Learning Algorithm

Machine learning, which at its core is a neural network with 3 or more layers, includes deep learning as a subset. These neural networks attempt to imitate the functionality of human brain, but incapable to compete with it, allowing it to "learn" from enormous amounts of dataset. Even though a neural network with just one layer may still generate approximate predictions, more hidden layers can assist to tweak and improve for accuracy. These algorithms can administer unstructured textual and visual data. They also ensure automated feature extraction, minimizing the requirement for human expertise.

By using data inputs, weights, and bias, artificial neural networks, often referred to as Deep Learning Neural Networks (DNN), attempt to imitate the human brain functionality. DNNs comprises of several layers of linked nodes. Each of which improves based on the prediction or classification constructed by the layer before it. Forward propagation describes how computations flow through a network. Input and output layers are basically the observable layers of a deep neural network. The deep learning model ingests the data for processing in the input layer. The final prediction or classification is performed in the output layer.

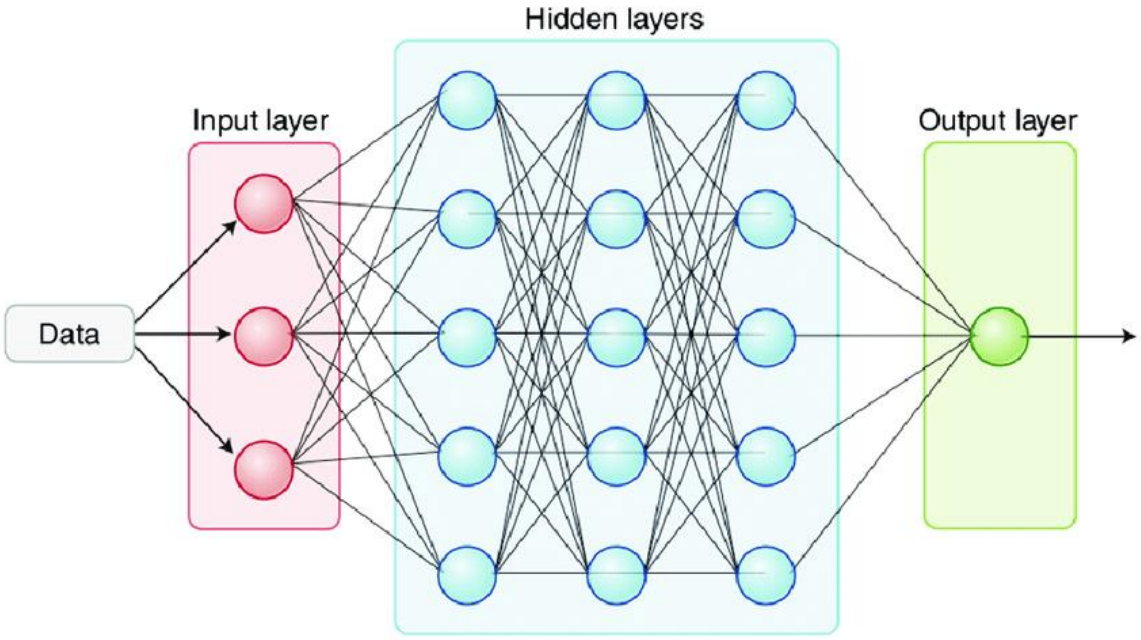


Figure 4: Architecture of Deep Learning Neural Network

3.5.1 Convolutional Neural Network

A deep learning-based neural network that can discover valuable information in both time series and image data is the Convolutional Neural Network (CNN). It is thus very beneficial to image pattern recognition related tasks, object classifications, as well as image recognition. In order to discover patterns in an image, a CNN employs ideas from linear algebra, such as matrix multiplication. CNNs might even categorize audio and signal sets.

The structure of a CNN is comparable to the interconnection structure of the human brain. CNNs contain billions of neurons, much like the brain, but they are organized differently. CNN's neurons are somehow assembled to resemble the frontal lobe of the brain, which infers visual information. This design resolves the problem of piecemeal image processing of standard neural networks, which requires feeding them pictures in low-resolution pieces. It ensures the whole visual area coverage. In comparison to the

prior networks, CNN gives better performance with photo and audio inputs.

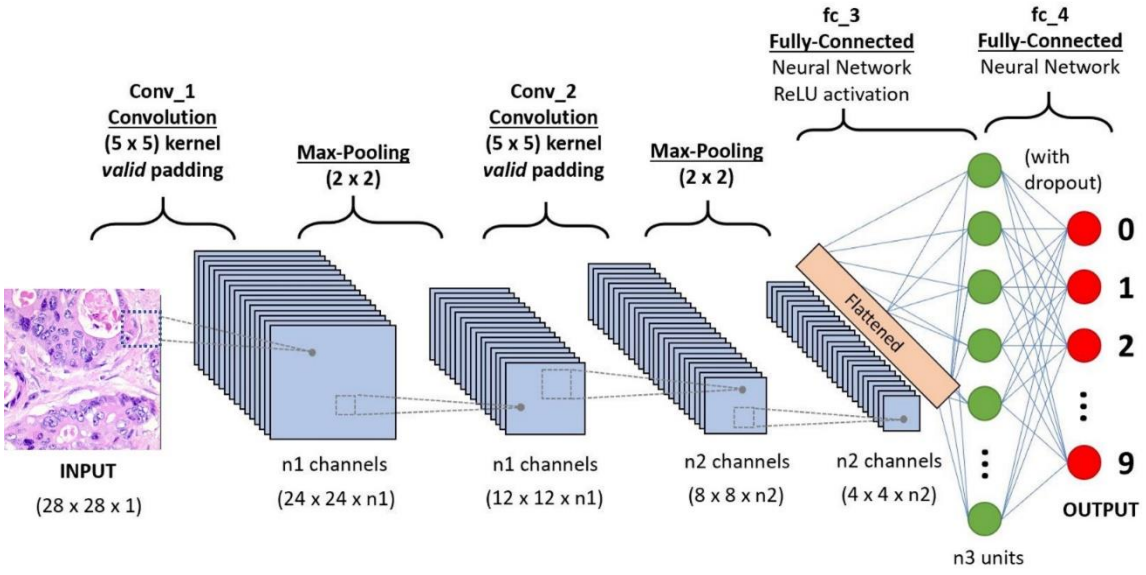


Figure 5: Architecture of Convolutional Neural Network.

3.5.1.1 Convolutional Layer

As its name suggests, the convolutional layer is crucial to CNN's functionality. The boundaries of the layers are based on the use of comprehensible kernels. Despite these kernels are generally small, they might vary depending on the profundity of the input. While data hits a convolutional layer, the layer convolves each channel throughout the spatial dimensions of the input in order to produce a 2-D enactment map. The network will take in "firing" kernels, also known as activations, when a given characteristic on an assumed spatial location is detected.

Each kernel has an identical activation plot, which is weighted along with the complexity measurement to form the convolutional layer's whole output dimensions. Three hyperparameters—stride + depth + zero-padding setting, can improve these. The number of neurons that are given priority in the layer relative to the input's equivalent area can significantly affect how deeply the output size is spread across the convolutional layers.

In order to describe the step, we may also select the depth around the input's spatial dimensions and ask how to place the open field. If the stride was set to 1, as an illustration, we would then take a receptive field that was heavily covered and generate extraordinarily enormous activations. Additionally, positioning the step in a more prominent quantity will result in a result with fewer spatial dimensions and less overlapping.

Zero-padding is regarded as an efficient way to further regulate the dimensionality of the output dimensions. It is a straightforward approach for padding the input border. These techniques aid in changing the spatial dimensionality of the convolutional layer's outcome. This computation is performed utilizing the following equation:

$$\frac{(V - R) + 2Z}{S + 1} \quad (1)$$

Here, V denotes the input volume size (depth x height x width), R and Z denote the size of the accessible field, and the volume of zero padding is adjusted independently. S standing for stride. Since the neurons can't properly correct through the planned input, the stride has been incorrectly fixed if the deliberate result of this equation is not identical to a whole integer.

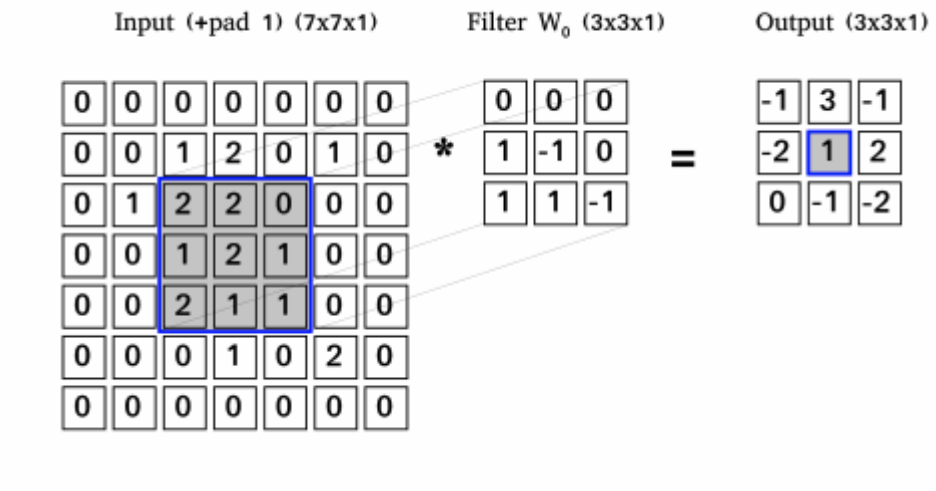


Figure 6: A 7x7x1 input image convolved with a 3x3x1 kernel filter to be a 3x3x1 output feature

3.5.1.2 Maxpool Layer

The primary goal of the max-pooling layer is mainly the gradual reduction of the demonstration's dimensionality as well as the number of parameters and computing complexity of the program. Every activation plot in the input is activated by the pooling layer, which then uses the "MAX" function to calculate the dimensionality of each plot. These are generated using max-pooling layers and 2 x 2 dimensional kernels that are spread out throughout the input's spatial dimensions with a 2-step lateral distance in maximum CNNs. This maintains the depth size at the standard size while measuring the activation plot down to 25 % of the original scope.

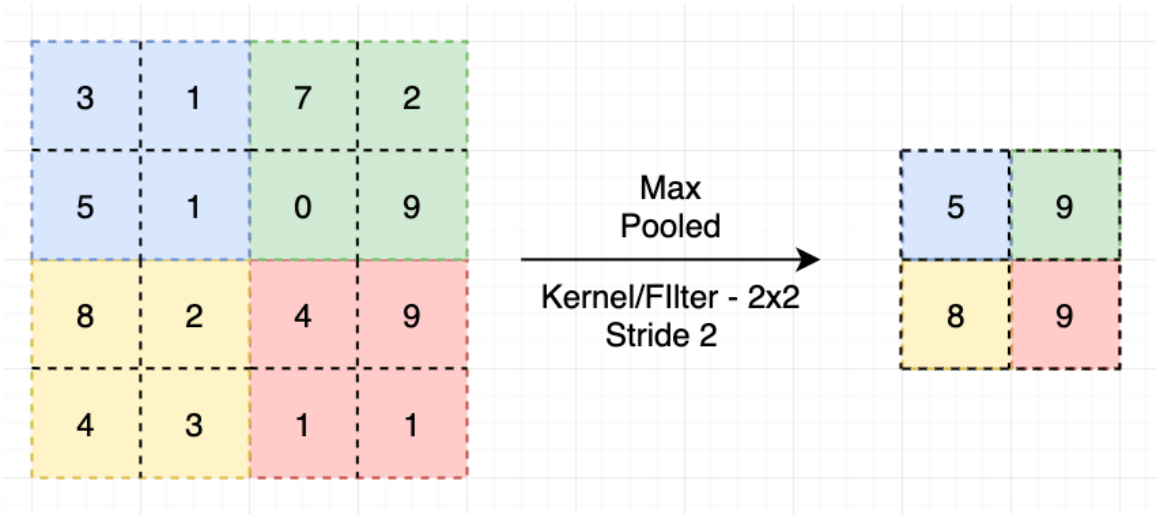


Figure 7: Max Pooling Layer with 2x2 Kernel / Filter and Stride size 2

There are only two methods for maximum pooling that have received much attention because of the pooling layer's destructive tendency. Pooling layers frequently have their stride and filters set to 2 X 2 at the same time, allowing the layer to spread out throughout the whole input spatial dimensions. Where the stride is specified to two by setting the kernel size equal to three, overlapping pooling must also be used. Pooling has a destructive nature, therefore using more than three kernels will often cause the archetypal to perform noticeably worse. It is important to realize that in excess to max-pooling,

CNN structures may also use general-pooling. Normalization: L1+ L2 and average pooling are two frequent activities that may be carried out by the pooling neurons in generic pooling layers.

3.5.1.3 Fully Connected Layer

The fully connected layer consists of neurons inextricably linked to the neurons in the dual adjacent layer's and are not linked to any neurons in those layers. The pattern of neurons in this is identical to that in traditional varieties of ANN.

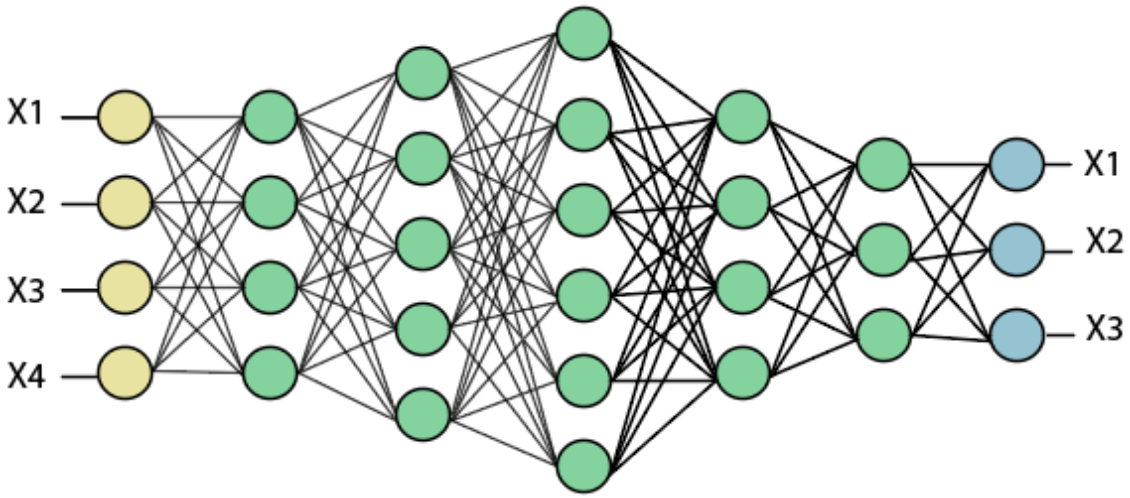


Figure 8: Sample architecture of Fully Connected Layer

3.5.2 Transfer Learning

With more people using the internet, a large amount of information is available for rigorous analysis. It would be challenging in the upcoming years to manage such massive volumes of data. Unfortunately, due to security and confidentiality considerations, there are extremely low number of healthcare field datasets. Histopathological graphical images, on the contrary, are incredibly volatile. It will be challenging and problematic to construct a model from start containing characteristics (features) that were randomly initialized and few data points. In this study, pre-trained Convolutional Neural Network models are employed to tackle these difficulties. The aim is to transfer the weights from a

prevailing model to dissimilar but related problems which is basically the main method of Transfer Learning. It contributes to lessen the computational load and boost performance, which leads to outputs that are faster and more efficient. In this study, the LC25000 dataset is used to train three different pre-trained CNN models: VGG16, ResNet50 and EffecientNetB7.

3.5.2.1 VGG16

VGG16, a popular Convolution neural network (CNN) architecture, is considered as one of the greatest visualization model schemes available till date. The most idiosyncratic feature of VGG16 is that it prioritized containing convolution layers of a 3x3 filter with a stride 1 and all-time utilized the same padding and maxpool layer of a 2x2 filter with a stride 2. Across the whole design, convolution and max pool layers are organized in the similar fashion. Two FC (completely connected) layers are present at the very last, trailed by a SoftMax for output. The 16 in VGG16 denotes that there are sixteen layers with weights. This network comprises over 138 million parameters, making it a huge network.

3.5.2.2 ResNet50

He Kaiming, Zhang Xiangyu, Ren Shaoqing, and Sun Jian's 2016 publication "Deep Residual Learning for Image Recognition" proposed a convolutional neural network (CNN) variant known as ResNet [24]. ResNet50 is widely and most frequently utilized in computer vision applications. ResNet, or Residual Networks, is a generic name for a neural network which forms the basis of numerous computer vision solutions. ResNet's key breakthrough was its potential to train extremely sophisticated neural networks containing higher than 150 layers. Every two-layer block in the thirty-four-layer net is replaced by this three-layer bottleneck block, creating a fifty-layer ResNet with a fixed input photo resolution with 224 by 224 pixels. In this architecture, there are almost twenty-three million parameters.

3.5.2.3 EfficientNetB7

EfficientNetB7 is also a widely and most frequently utilized in computer vision applications. In a research article titled "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," Mingxing Tan and Quoc V. Le of the Google Research, Brain team suggested a variation of the CNN.

They created a new baseline network using the neural architecture search and ramped it up to create a family of deep learning models dubbed EfficientNet's. These models use far fewer parameters while outperforming earlier CNN models in terms of accuracy and effectiveness. The researchers scaled the network's scope by employing a compound scaling method. The implemented grid search technique was utilized to establish the affiliation between the several scaling dimensions of the baseline network using a static resource limitation.

By employing the method, they may regulate the proper scaling coefficients for each of the to-be-scaled-up dimensions using this method. The baseline network was scaled to the requisite size utilizing this factor. The experts initially constructed a baseline network utilizing the neural architecture search, a technique for automatically building neural networks. It augments accurateness and effectiveness on a foundation of floating-point operations per second (FLOPS). The lately developed design makes advantage of the moveable inverted bottleneck convolution (MBConv). The EfficientNet's family of deep learning models was originated after the researchers ramped up this original network. They also portrayed a comparison of EfficientNet's performance against other potent transfer learning models while applied to the ImageNet dataset. It is proven that the most recent pre-trained model, EfficientNet-B7, has the maximum accuracy of all with the minimum parameters.

Any network's stem comes first, after that all experimentation with the architecture universal to all eight models, and the final layers begin. Following that, each of them has

seven blocks. The number of these blocks' sub-blocks rises from EfficientNetB0 to EfficientNetB7, with a distinct quantity present in each block. There are 813 layers in EfficientNet-B7 overall.

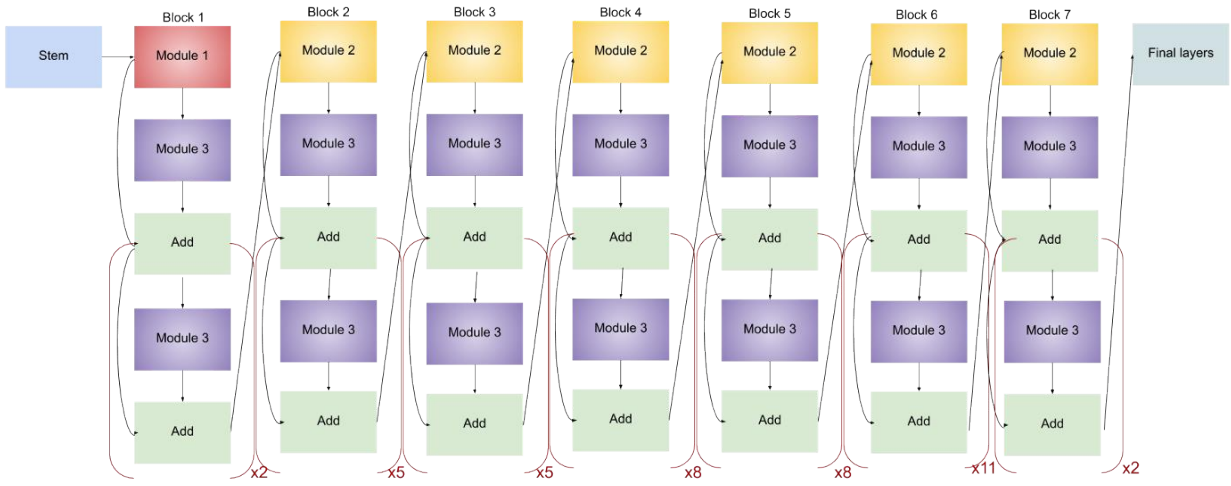


Figure 9: Architecture of EffecientNetB7

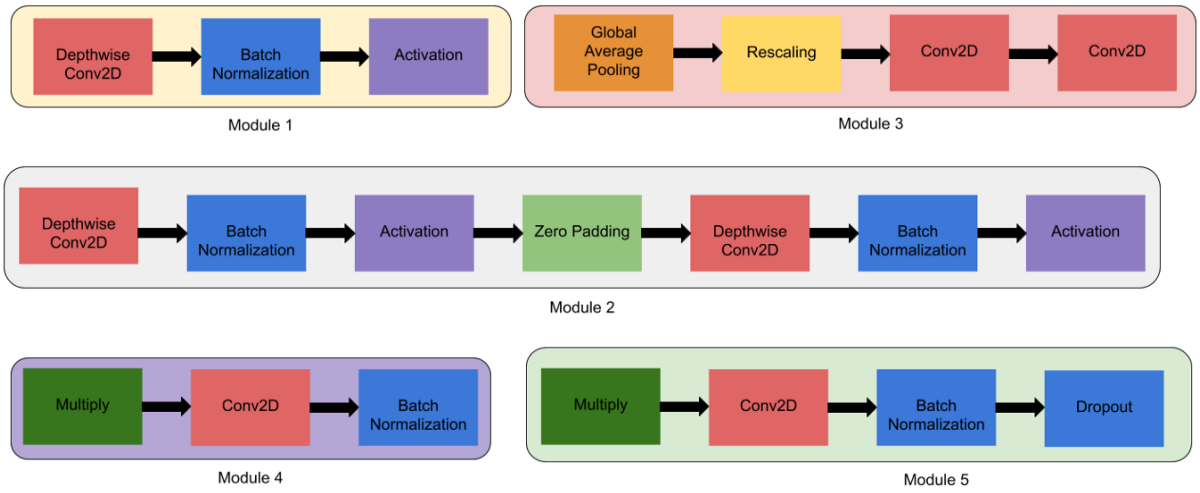


Figure 9.1: Models that make up the EffecientNetB7 network architecture

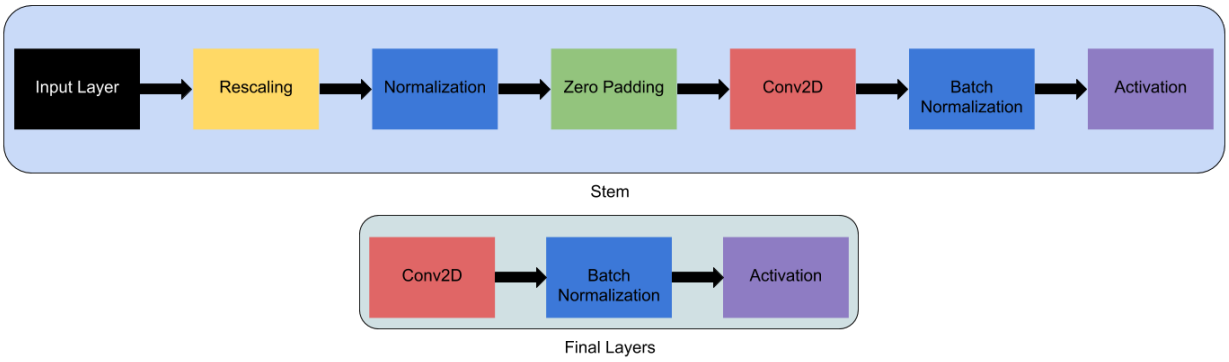


Figure 9.2: Stem and the final layers

3.5.3 Training and Testing

In LC25000 histopathological image dataset, three transfer learning algorithms are applied classified the image based on the selected features. The process entails splitting the dataset into two subgroups where 20% of the dataset was harnessed for testing and the remaining 80% was used for training. These divided train test data were randomly chosen. The training dataset is the initial subset, which is used to fit the model. The training set's imagery features are utilized to train the algorithms. The input element of the dataset is supplied to the model instead of the test subset, and its predictions are then generated and compared to the predicted values. Finally, the testing set's photo features are utilized to evaluate the model's performance. Python has been selected as the chosen programming language in this study and the following libraries have been implemented: numpy, pandas, scikit-learn, scikit-image matplotlib, itertools, cv2, os, time, tensorflow, and glob. All the models were trained using a transfer learning method, and the loss function stated in equation was categorical cross-entropy. The learning rate was adjusted equal to 0.001.

$$L_{CE} = -\sum_{i=1}^{n} t_i \log(p_i) \tag{2}$$

In this study, we suggest a framework for our model's training that can be applied for all three binary classification tasks. The structure is divided into three sections:

a) Feature Extraction Part - Using the three well-known pre-trained CNN architectures, the models are adjusted in order to extract the most crucial properties. After this, the transformed portion would receive these attributes for further categorization...

b) Modified Part - In order to produce a vector of features, Max-pooling 2D, Average-pooling 2D, and a flatten layer were combined into an append layer, which is a combination of three coatings. A dropout layer has also been developed with a 0.5 dropout rate.

c) Output Section - Appended a vector of features with the following sigmoid activation function to an output layer:

$$f\ (x) = \frac{1}{1 + e^{-x}} \tag{3}$$

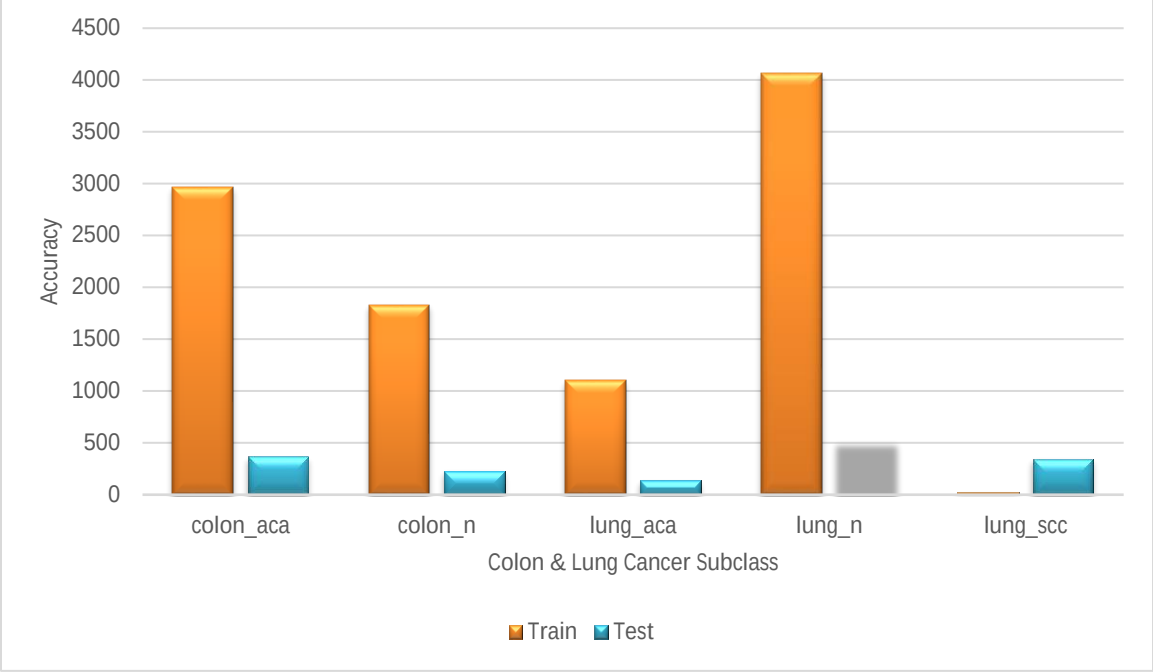


Figure 10: Train Test split among different Lung & Colon cancer classes

3.5.4 Compilation

An optimization algorithm is necessary for model compilation. Adam, RMSProp, Mini-Batch Gradient Descent, AdaMax, Learning Rate Decay etc are examples of optimization techniques. Adam optimizer is chosen because it meets the most standards and performs well for all deep learning models where SoftMax was employed by means of the activation function for all the topologies. Learning rate = 0.0001, epochs = 30, and batch size = 128 are the hyperparameter values used to train the models.

Adam optimizer shown in equation (1).

$$f_i(\overrightarrow{a}) = \frac{e^{a_i}}{\sum_k e^{a_k}} \quad (4)$$

3.6 Performance Evaluation

Machine learning models are assessed using a variety of measures. The confusion matrix and relevant performance indicators including Accuracy, Precision, Recall, and F1-score are utilized in this study to evaluate the performance. Let the letter be:

(Tp) = Number of instances the positive class accurately predicted by the model.

(Tn) = Number of instances the negative class accurately predicted by the model.

(Fp) = Number of instances the positive class inaccurately predicted by the model

(Fn) = Number of instances the negative class accurately predicted by the model

Precision, recall, accuracy, and F1-score can be represented as below:

Precision

Precision is a measure of how many of the positive predictions made are correct (true positives). The formula for it is:

$$Precision = \frac{Tp}{Tp + Fp} \quad (5)$$

Recall

Recall is a measure of how many of the positive cases the classifier correctly predicted, over all the positive cases in the data. It is sometimes also referred to as Sensitivity. The formula for it is:

$$Recall = \frac{Tp}{Tp + Fn} \quad (6)$$

Accuracy

The base metric used for model evaluation is often Accuracy, describing the number of correct predictions over all predictions. The formula for it is:

$$Accuracy = \frac{Tp + Tn}{Tp + Tn + Fp + Fn} \quad (7)$$

F1-Score

F1-Score is a measure combining both precision and recall. It is generally described as the harmonic mean of the two. Harmonic mean is just another way to calculate an "average" of values, generally described as more suitable for ratios (such as precision and recall) than the traditional arithmetic mean. The formula used for F1-score in this case is:

$$F1\ Score = 2\,x\frac{Precision.Recall}{Precision + Recall} \quad (8)$$

CHAPTER 4
EXPERIMENTAL RESULTS AND DISCUSSION

4.1 Results and Discussion

This section summarizes the learning outcomes for the three potent CNN architectures that were tested in this experiment for the detecting the classes of colon and lung cancer. The three classifiers being investigated in this paper's obtained results are presented in this part.

The classifiers EfficientNetB7, ResNet50, and VGG16 on the test data produced accuracy of 98%, 93%, and 95% as well as F1-scores of 98%, 20%, and 95%, respectively. Table I displays the performance of the classification models on the identical test dataset and Figure 11 compares the test accuracy among different models. The obtained findings demonstrate that the machine learning models can detect lung and colon cancer subcategories with high accuracy.

The EfficientNetB7model fared best, scoring 98% on the F1 scale and 98% on accuracy. Table II displays the EffecientNetB7 model's precision, recall, and f1-score for the multiple segments of histopathological images on the testing dataset.

A confusion matrix is used to evaluate the test viability of the various models. The confusion matrix comprises the true positive, true negative, false positive, and false negative values. The values in the confusion matrix's diagonal position assess the model's precise prediction. Depending on the confusion matrix; accuracy, precision, recall, sensitivity and the f1-score are computed.

TABLE I. PERFORMANCE METRICS OF APPLIED CNN TRANSFER LEARNING ARCHITECTURES

Architectures	Accuracy	Precision	Recall	F1-Score
VGG16	.95	.95	.95	.95
ResNet50	.93	.20	.20	.20
EffieientNetB7	.98	69	70	69

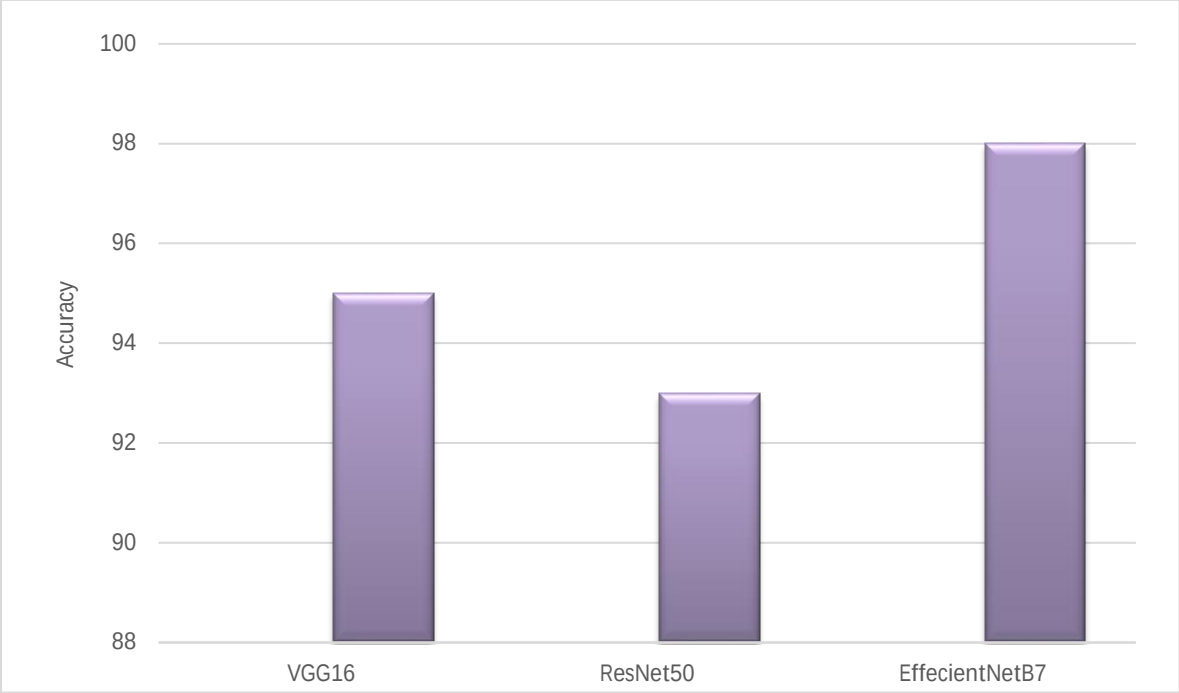


Figure 11: Test Accuracy of different Transfer Learning Architecture

TABLE II. PERFORMANCE METRICS OF EFFECIENTNETB7 FOR FIVE DIFFERENT CLASSES

Class	Precision	Recall	F1-Score
colon_aca	1.00	.96	.98
colon_n	0.98	1.00	0.99
lung_aca	0.94	0.98	.96
lung_n	1.00	1.00	1.00
lung_scc	0.98	0.96	0.97
weighted avg	.98	.98	.98

Figure 12 & Figure 13 portrays not only the training & validation accuracy but also training and validation loss respectively. It shows that incremental epochs maximizes accuracy and limiting loss.

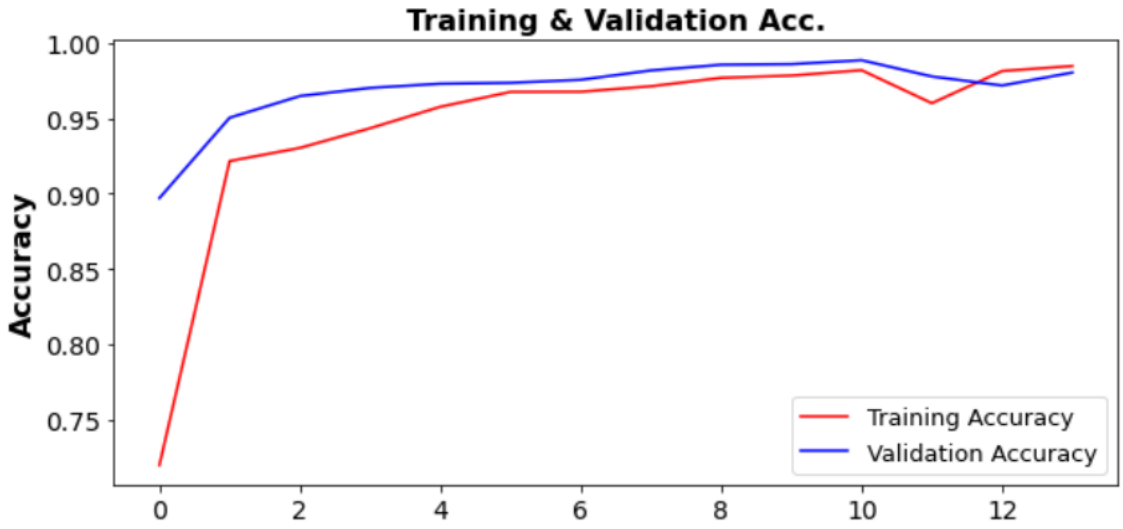


Figure 12: Training & Validation Accuracy of EffecientNetB7

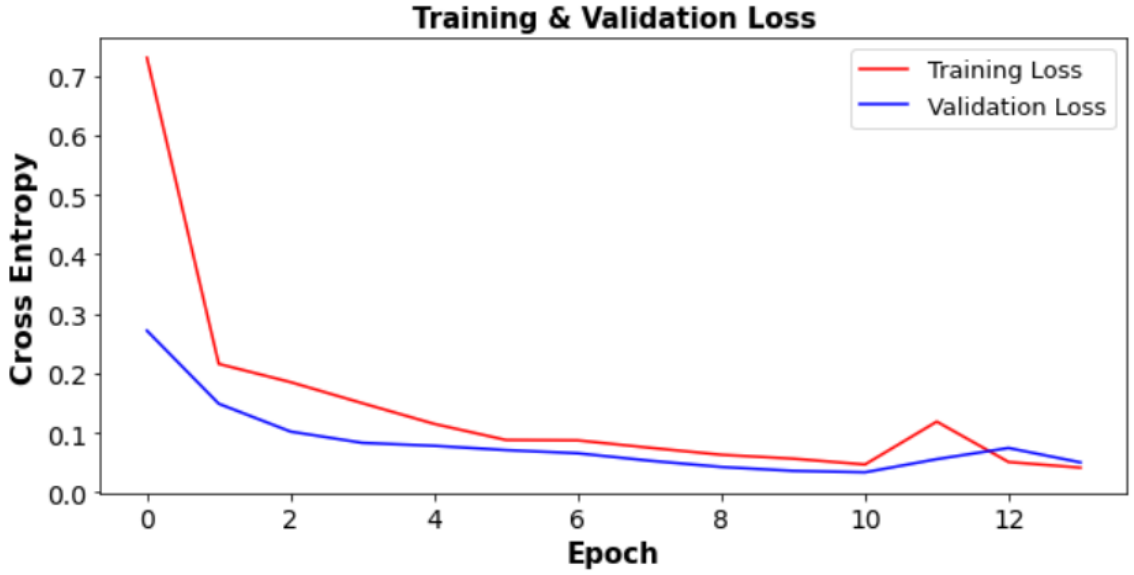


Figure 13: Training & Validation Loss of EffecientNetB7

The confusion matrix for each approach applied to the identical dataset is depicted in Figures 14 and 15.The actual (true) label vs the expected (predicted) label of the visuals

regarding the testing data in the specified labeled classifications is represented by the confusion matrix.

TABLE III. CONFUSION MATRIX OF EFFECIENTNETB7

	Predicted				
Actual	**colon_aca**	**colon_n**	**lung_aca**	**lung_n**	**lung_scc**
colon_aca	.96	.02	.02	0.0	0.0
colon_n	0.0	1.0	0	0	0
lung_aca	0.0	0.0	.98	0.0	0.02
lung_n	0.0	0.0	0.0	1.0	0.0
lung_scc	0.0	0.0	0.04	0.0	0.96

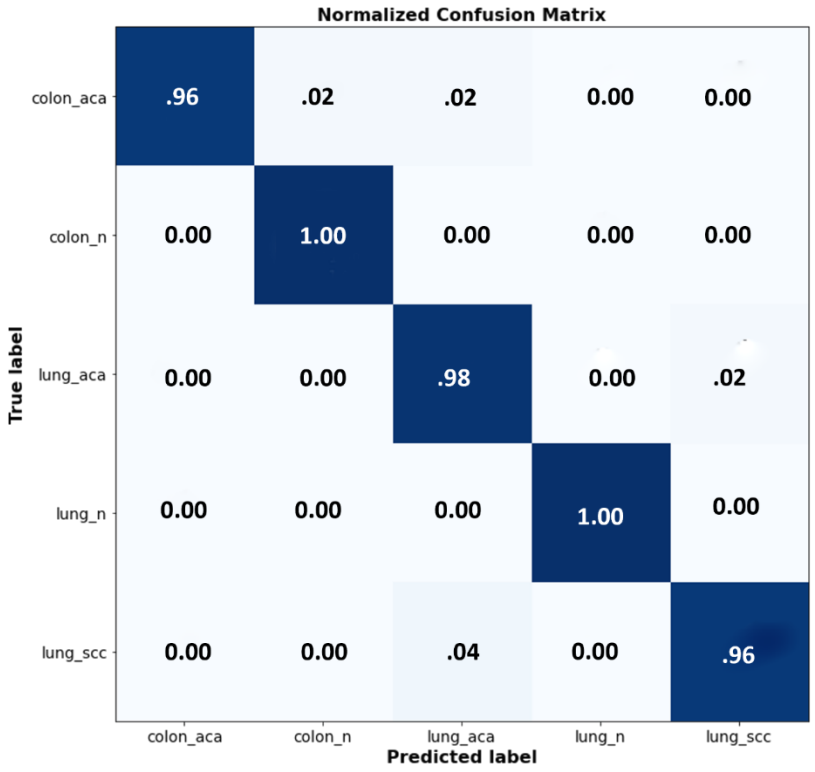
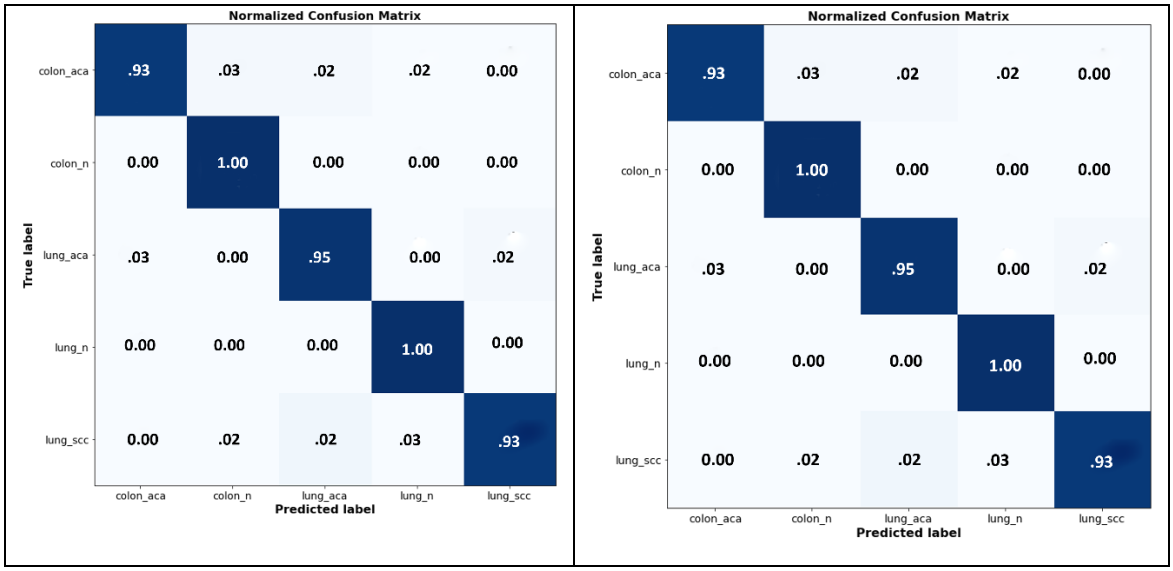


Figure 14: Confusion Matrix of EffecientNetB7



(a) (b)

Figure 15: Confusion Matrix of a) ResNet50 & b) VGG16

4.2 Comparative Analysis

It is crucial to quantify the research's significance. Part of its mechanisms are compared to supplementary study works that are shown in table IV for this goal. For the classification of lung and colon cancer, several machine learning techniques were used, including K-Nearest Neighbour (K-NN), Decision Tree, Multinomial Naive Bayes classifier, Random Forest classifier, Multi-Layer Perceptron Classifier (MLP), XGBoost, Stochastic Gradient Descent classifier, Support Vector Machine (SVM) and Fuzzy Neural Network.

There are lots of techniques that have been developed recently to identify lung cancer, most of which use CT scan pictures and some of which use x-ray images. Additionally, several classifier techniques are coupled. With a variety of segmentation methods, lung cancer nodules may be located using image recognition.

TABLE IV. COMPARATIVE ANALYSIS

Ref.	Year	Preprocessing	Methods	Datasets	Results
(Reddy et al. 2019) [20]	2019	Image securing, pre-handling, binarization, thresholding, division, and feature extraction are implemented.	To evaluate the neural network system using machine learning, the fuzzy neural system is employed.	UCI repository dataset	Accuracy 96.67 %.
(Bhatia et al., 2019) [21]	2019	Segmentation trailed by normalization, and zero centering constitute this stage.	Several classifiers as XGBoost and Random Forest are employed.	Dataset of Lung Image Database Consortium (LIDC-IDRI)	Accuracy 84%
(Makaju et al., 2018) [22]	2018	On the CT images, median and Gaussian filters are implemented.	Watershed segmentation for identification, SVM to classify nodule as Malignant or benign.	Database of Lung Image Database Consortium (LIDC)	Accuracy 92%
(Singh and Gupta, 2018) [23]	2018	Employing thresholding techniques to turn the grayscale imagery into a binary image after applying denoising techniques like Median, Gaussian, and Bilateral blur.	SVM,K-NN, Decision Tree, Random Forest, Stochastic Gradient Descent, Multinomial Naive Bayes, and Multi-Layer Perceptron (MLP) classifier are implemented	Database of Lung Image Database Consortium (LIDC)	Accuracy 88.55%.
Proposed Model		**Image augmentation pipeline from the imgaug library & Unsharp Masking**	**Pre-trained model EffecientNetB7 is applied as the main model**	**Lung and Colon Cancer Histopathological Image Dataset (LC25000)**	**Accuracy 98%.**

CHAPTER 5
CONCLUSION AND FUTURE WORK

5.1 Conclusion

In this study, we addressed the necessity of automation in clinical imaging, which can significantly reduce the demand for experts. Next, we discuss similar studies that progress from employing fundamental models to cutting-edge models like deep learning. We then experimented with information enlargement techniques on our histology images. Three well-known pre-trained CNN models: VGG16, ResNet50, and EfficientNetB7 were used, and they produced astounding results with accuracy ranging from 93% to 98%. In all classification tasks, EfficientNetB7 achieved 98% precision, accuracy, recall, and f1-score. All remaining models have similarly shown remarkable performance in all assessment criteria, ranging from 93% to 95%. These high levels of accuracy are the consequence of combining a sophisticated image augmentation pipeline (imgaug) with sophisticated pre-trained Convolutional Neural Network Models. ResNet50 is performing in an unexpected way when compared to other models, as noticed by the accuracy and training + validation loss of several Pre-trained Convolutional Neural Network Architectures.

5.2 Future Work

To enhance model performance, it is intended to investigate additional deep learning pre-trained models that offer pertinent characteristics for detecting colon and lung cancer subtypes from histopathological image datasets in future. In order to assess the effectiveness of our suggested strategy, more histopathology images of colon and lung cancer will be used. We will also combine several cancer datasets and perform multi-class classification between various cancer subtypes using all the approaches that have been mentioned. Increasingly, the genetic information of the patient's cancer will be explored to determine the best course of action.

REFERENCES

[1] “Cancer,” World Health Organization. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/cancer.

[2] A. Borkowski, M. Bui, L. B. Thomas, C. P. Wilson, L. Deland, and S. Mastorides, "Lung and colon cancer histopathological image dataset (lc25000)." arXiv preprint arXiv:1912.12142, 2019.

[3] D. Komura and S. Ishikawa, “Machine learning methods for histopathological image analysis,” Computational and Structural Biotechnology Journal, vol. 16, pp. 34–42, 2018.

[4] Cruz, C.S.D., Tanoue, L.T. and Matthay, R.A., “Lung cancer: epidemiology, etiology, and prevention. Clinics in chest medicine”, 32(4), pp.605-644, 2011

[5] P. Rawla, T. Sunkara, and A. Barsouk, “Epidemiology of colorectal cancer: Incidence, mortality, survival, and risk factors,” Gastroenterology Review, vol. 14, no. 2, pp. 89–103, 2019.

[6] “Imgaug,” imgaug. [Online]. Available: https://imgaug.readthedocs.io/en/latest/

[7] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional Neural Networks,” Communications of the ACM, vol. 60, no. 6, pp. 84–90, 2017.

[8] D.Powers, "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation." arXiv preprint arXiv:2010.16061, 2020.

[9] L.Jiao, Q.Chen, “Colon Cancer Detection Using Whole Slide Histopathological Images”. IFMBE Proceedings. 39. 1283-1286, 2013

[10] S. Mehmood et al., "Malignancy Detection in Lung and Colon Histopathology Images Using Transfer Learning With Class Selective Image Processing," in IEEE Access, vol. 10, pp. 25657-25668, 2022.

[11] T. Wan, J. Cao, J. Chen, and Z. Qin, “Automated grading of breast cancer histopathology using cascaded ensemble with combination of multi-level image features,” Neurocomputing, vol. 229, pp. 34–44, 2017.

[12] S. Rathore, M. Hussain, and A. Khan, “Automated Colon Cancer Detection using hybrid of novel geometric features and some traditional features,” Computers in Biology and Medicine, vol. 65, pp. 279–296, 2015.

[13] F. A. Spanhol, L. S. Oliveira, C. Petitjean and L. Heutte, "Breast cancer histopathological image classification using Convolutional Neural Networks," 2016 International Joint Conference on Neural Networks (IJCNN), pp. 2560-2567, 2016

[14] S. Garg and S. Garg, “Prediction of lung and colon cancer through analysis of histopathological images by utilizing pre-trained CNN models with visualization of class activation and saliency maps,” 2020 3rd Artificial Intelligence and Cloud Computing Conference, 2020.

[15] T. Araújo, G. Aresta, E. Castro, J. Rouco, P. Aguiar, C. Eloy, A. Polónia, and A. Campilho, “Classification of breast cancer histology images using convolutional neural networks,” PLOS ONE, vol. 12, no. 6, 2017.

[16] 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. 1–6, 2017.

[17] M. Raghu, C. Zhang, J. Kleinberg, and S. "Transfusion: Understanding transfer learning for medical imaging." Advances in neural information processing systems 32, 2019.

[18] L. D. Nguyen, D. Lin, Z. Lin and J. Cao, "Deep CNNs for microscopic image classification by exploiting transfer learning and feature concatenation," 2018 IEEE International Symposium on Circuits and Systems (ISCAS), pp. 1-5, 2018

[19] D. Soydaner, “A comparison of optimization algorithms for Deep Learning,” International Journal of Pattern Recognition and Artificial Intelligence, vol. 34, no. 13, p. 2052013, 2020.

[20] U. J. Reddy, B. R. Reddy, and B. E. Reddy, “Recognition of lung cancer using machine learning mechanisms with fuzzy neural networks,” Traitement du Signal, vol. 36, no. 1, pp. 87–91, 2019.

[21] S. Bhatia, Y. Sinha, and L. Goel, “Lung cancer detection: A deep learning approach,” Advances in Intelligent Systems and Computing, pp. 699–705, 2018.

[22] S. Makaju, P. W. C. Prasad, A. Alsadoon, A. K. Singh, and A. Elchouemi, “Lung cancer detection using CT scan images,” Procedia Computer Science, vol. 125, pp. 107–114, 2018.

[23] G. A. Singh and P. K. Gupta, “Performance analysis of various machine learning-based approaches for detection and classification of lung cancer in humans,” Neural Computing and Applications, vol. 31, no. 10, pp. 6863–6877, 2018.

[24] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90.