DETECTION OF LUNG CANCER TYPES USING CT SCANS

A.A Nimesha Hansani Amarasinghe

219182P

Master of Science in Computer Science Specialized in Data Analytics and Engineering

Department of Computer Science and Engineering

University of Moratuwa
Sri Lanka

August 2023

DETECTION OF LUNG CANCER TYPES USING CT SCANS

A.A Nimesha Hansani Amarasinghe

219182P

Thesis report submitted in partial fulfillment of the requirements for the degree Master of Science in Computer Science specialization in Data Analytics and Engineering

Department of Computer Science and Engineering

University of Moratuwa

Sri Lanka

August 2023

DECLARATION

I declare that this is my own work, and this thesis/dissertation does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other University or Institute of higher learning and to the best of my knowledge and believe it does not contain any material previously published or written by another person except where the acknowledgement is made in the text. I retain the right to use this content in whole or part in future works (such as articles or books)

Signature: Nimesha Amarasinghe Date: 23.09.2023

The above candidate has carried out research for the master's thesis under my supervision. I confirm that the declaration made above by the student is true and correct.

Name of the supervisor: Dr. Thanuja D. Ambegoda

Signature: Date:

ABSTRACT

Lung cancer is a devastating global health issue that can be fatal if not detected early. To increase the chances of successful treatment and prevent the loss of lives, doctors need to identify the type which lung cancer belongs to. Currently, CT scans are commonly used in medical practice to detect and diagnose lung tumors. However, implementing deep learning models to identify the lung cancer types poses a significant challenge. Because acquiring many medical images for each type of lung cancer can be difficult.

The effectiveness of deep learning algorithms which were implemented for image recognition relies on the diversity of data samples. However, creating a comprehensive dataset of significant size requires a significant amount of human effort for manual labeling, which makes it a time-consuming process. The challenge of requiring a large amount of data samples for each category in traditional deep learning models was addressed in this research by implementing a prototypical network, which is a few-shot learning technique. This method requires only a few samples per category, and it was used in conjunction with a pre-trained model to extract features from lung CT scans. The accuracy of the model was analyzed based on the number of samples per category.

Overall, the results of the study demonstrate that implementing a prototypical network for lung cancer type detection is feasible. Human interpretation of medical images can vary among different medical professionals. Therefore, this approach can act as a decision support tool for medical professionals which makes it more accessible and cost-effective.

KEYWORDS: Prototypical Network, Lung Cancer, Few Shot Learning, Segmentation, VGG16

ACKNOWLEDGEMENT

I want to express my sincere thanks to Dr Thanuja D. Ambegoda for assisting me in identifying a fascinating research topic and for his support and encouragement. His guidance was quite beneficial in terms of goal creation and study engagement. I would also like to thank Dr Upul Narandeniya from Colombo General Hospital, who explained how to identify lung cancer patients. My heartfelt gratitude goes to the Department of Computer Science and Engineering at the University of Moratuwa for their assistance in overcoming this challenge. Finally, I want to convey my deep appreciation to my family for their unwavering support throughout this journey.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

CAD: Computer Aided Diagnosis

CNN: Convolution Neural Network

PET: Positron Emission Tomography

DICOM: Digital Imaging and Communication in Medicine

FSL: Few Shot Learning

MRI: Magnetic Resonance Imaging

TCIA: The Cancer Imaging Archive

1 INTRODUCTION

1.1 Motivation and Background

Lung cancer is an imposing worldwide health emergency, claiming the existences of roughly 422 individuals consistently across the world. Deplorably, lung cancer ranks among the deadliest and most prevalent types of cancer, principally originating within the lung tissues. This perilous disease is characterized by the uncontrolled proliferation of cells in the lungs. To more readily understand and address this healthcare challenge, it's essential to dive deeper into the nuances of lung cancer, its types, diagnosis, and the job of computer-aided design (computer aided design) frameworks in improving early detection and patient results.

The American Cancer Society has identified two main classifications of lung cancer: non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC). NSCLC is the more normal of the two, accounting for roughly 80% to 85% of all lung malignancies [1]. Within the domain of NSCLC, there are three significant subgroups: Adenocarcinoma, Squamous cell carcinoma, and enormous cell carcinoma [1]. On the other hand, SCLC represents a distinct classification. Understanding the differences among these types is urgent for tailoring effective treatment strategies and improving patient endurance rates.

Non-small cell lung cancer often follows a challenging way of progression. In its beginning phases, this kind of cancer frequently presents no undeniable side effects, making convenient diagnosis a daunting task. As the disease advances and side effects manifest, it turns out to be increasingly challenging to effectively intervene. In essence, early detection of NSCLC can significantly impact a patient's prognosis, highlighting the need for additional precise diagnostic tools and techniques.

This is where computer-aided design (computer aided design) frameworks become an integral factor. Computer aided design frameworks are increasingly becoming a preferred tool for healthcare professionals, especially radiologists, in the field of diagnostic radiology and medical imaging. These frameworks certainly stand out and investment because of their potential to help with the interpretation of medical images, especially processed tomography (CT) scans. By harnessing artificial intelligence approaches, computer aided design frameworks plan to help radiologists in the early and accurate detection of pulmonary nodules and different abnormalities in lung images.

CT imaging stands out as the preferred methodology for lung cancer diagnosis because of its wide availability, cost-effectiveness, and fast image acquisition capacities. It gives exceptionally point by point cross-sectional images of the chest, allowing for the visualization of even little nodules and abnormalities within the lung tissue. This makes it an important tool for early diagnosis and staging of lung cancer. Nonetheless, analyzing the tremendous amount of data generated by CT scans remains a considerable challenge for radiologists.

Computer aided design frameworks, with their artificial intelligence-based algorithms, hold the potential to lighten this burden and work on the accuracy of lung cancer diagnosis. These frameworks normally involve two main components: a computer-aided detection (CADe) subsystem and a computer-aided diagnosis (CADx) subsystem. The CADe subsystem is devoted to the segmentation of lung nodules, which involves identifying and delineating the boundaries of dubious structures within the lung images. Once the nodules are identified, the CADx subsystem takes over and centers around classifying these abnormalities, determining their nature, and assessing their potential malignancy.

In any case, in spite of the promising job of computer aided design frameworks in lung cancer diagnosis, there are a few challenges and considerations that should be addressed to boost their effectiveness. One of the essential concerns is the nature of the training data used to foster these computer aided design frameworks. Many of these frameworks are fundamentally trained on datasets that predominantly feature nodal

types of lung cancer. While this approach might be effective for identifying and classifying nodules, it doesn't completely align with the clinical and radiological variety of lung cancer types.

Lung cancer presents in different structures, and the distinct subtypes may have unique characteristics and ways of behaving. By fundamentally focusing on nodal types, computer aided design frameworks may inadvertently prohibit instances of other lung diseases and subtypes, leading to potential misdiagnoses or overlooking important clinical information. This limitation underscores the need for more comprehensive and different training datasets that encompass all lung cancer types, including the more uncommon ones.

The uncommonness of certain types of lung cancer in this present reality can confuse the development and training of computer aided design frameworks. At times, certain subtypes of lung cancer might be infrequent or represented by only a set number of cases in training datasets. This shortage of data can hinder the framework's capacity to detect and group these more uncommon structures accurately. To address this challenge, researchers and healthcare institutions should collaborate to mass bigger and more representative datasets that catch the full spectrum of lung cancer variety.

Besides, the sheer volume of data generated by modern CT scanners represents a significant challenge to radiologists. These scanners can catch up to 320 CT cuts simultaneously, each cut regularly comprising a 512×512-pixel image. For a routine chest examination, this outcomes in roughly 400 cuts, each requiring cautious analysis. The manual survey of such an extensive dataset can be tedious and mentally taxing, potentially leading to weakness and oversights.

This is where computer aided design frameworks can shine. They can quickly process and analyze the tremendous quantity of images, aiding radiologists in their work. Man-made intelligence algorithms in computer aided design frameworks succeed at pattern recognition and can methodically scan through numerous images, identifying dubious regions and highlighting potential abnormalities. Thusly, computer aided design frameworks not only save time yet in addition enhance the accuracy and consistency of diagnosis.

In addition to improving the efficiency and accuracy of lung cancer diagnosis, computer aided design frameworks can assume a vital part in the early detection of pulmonary nodules. Nodules are little, round or oval-formed developments within the lung tissue that could possibly be cancerous. Detecting these nodules at a beginning phase can be a key factor in improving patient results, as it allows for ideal intervention and treatment.

Computer aided design frameworks utilize refined algorithms to identify and segment pulmonary nodules in CT images. These algorithms analyze the density, size, and state of nodules, allowing for the differentiation between benign and potentially malignant nodules. This cycle assists radiologists with pinpointing areas of concern and focus on additional evaluation.

Moreover, computer aided design frameworks can support the classification of lung abnormalities, helping determine the likelihood of malignancy. By considering different characteristics of the detected nodules, like their size, shape, and development patterns, computer aided design frameworks can give significant insights into the nature of the abnormality. This information guides healthcare professionals in making informed decisions about patient consideration, including the necessity of additional testing or quick treatment.

The utilization of computer aided design frameworks in lung cancer diagnosis and nodule detection offers a few advantages. These frameworks are designed to work alongside radiologists, complementing their skill and enhancing the generally diagnostic interaction. Here are a few key benefits of computer aided design frameworks in lung cancer diagnosis:

Further developed Accuracy: computer aided design frameworks succeed at pattern recognition and can identify unobtrusive abnormalities that might slip by everyone's notice by the human eye. This prompts more accurate and solid diagnoses.

Consistency: computer aided design frameworks give consistent outcomes, reducing the likelihood of fluctuation in interpretations between different radiologists. This ensures that patients get a standardized degree of care.

Time Efficiency: computer aided design frameworks can quickly handle an enormous number of images, saving radiologists significant time and allowing them to zero in on additional complex diagnostic tasks.

Early Detection: computer aided design frameworks are adroit at identifying pulmonary nodules at a beginning phase, potentially leading to prior interventions and worked on patient results.

Diminished Weakness: The methodical nature of computer aided design framework analysis minimizes the risk of radiologist exhaustion, which can prompt mistakes or missed abnormalities during manual image audit.

Openness: computer aided design frameworks can be broadly sent, making advanced diagnostic capacities accessible in different healthcare settings, including more modest clinics and far off locations.

While computer aided design frameworks offer numerous advantages, they are not without their challenges and limitations. It's essential to resolve these issues to additional enhance their utility in lung cancer diagnosis:

Training Data Variety: As previously mentioned, computer aided design frameworks need admittance to different and representative training data to recognize all types of lung cancer effectively. Endeavors ought to be made to make comprehensive datasets that encompass the entire spectrum of lung cancer subtypes.

Bogus Up-sides and Negatives: computer aided design frameworks might deliver misleading positive or misleading negative outcomes. Misleading up-sides can prompt unnecessary additional testing and anxiety for patients, while bogus negatives can bring about missed diagnoses. Continuous refinement of computer aided design algorithms is necessary to lessen these blunders.

Integration into Clinical Workflow: Implementing computer aided design frameworks into clinical practice can present strategic challenges. Radiologists need to adjust to these frameworks, and institutions should consider factors like workflow integration, training, and cost-effectiveness.

Ethical Considerations: The utilization of artificial intelligence in healthcare brings up ethical issues connected with patient privacy, data security, and the responsible utilization of simulated intelligence algorithms. Shields and regulations should be set up to protect patient information and ensure ethical simulated intelligence deployment.

Notwithstanding these challenges, computer aided design frameworks hold significant commitment in the field of lung cancer diagnosis. They can possibly revolutionize the way we detect and order lung abnormalities, at last improving patient results and reducing the death rate related with lung cancer.

To harness the maximum capacity of computer aided design frameworks in lung cancer diagnosis, ongoing research and collaboration between medical professionals, researchers, and artificial intelligence specialists are significant. This collaboration can bring about the development of more hearty computer aided design algorithms, better integration of these frameworks into clinical practice, and the accumulation of extensive and different training datasets.

In conclusion, lung cancer is a considerable enemy, causing a significant number of passings around the world. Its initial and accurate diagnosis is crucial for improving patient results. Non-small cell lung cancer (NSCLC) accounts for most of cases and is characterized by its challenging progression from asymptomatic stages to indicative and advanced stages. The variety of lung cancer types and the immense amount of data generated by CT scans present significant challenges to radiologists.

Computer-aided design (computer aided design) frameworks, controlled by artificial intelligence , have arisen as promising tools for addressing these challenges. These frameworks help in the early detection of pulmonary nodules, the classification of lung abnormalities, and the efficient analysis of extensive image datasets. Computer aided design frameworks offer numerous benefits, including further developed accuracy, consistency, time efficiency, and early detection.

Nonetheless, challenges like the need for assorted training data, the potential for bogus up-sides and negatives, and ethical considerations should be addressed to augment the utility of computer aided design frameworks in lung cancer diagnosis. Cooperative endeavors among healthcare professionals, researchers, and computer-based

intelligence specialists are essential for advancing the field of computer aided design frameworks in lung cancer diagnosis. With continued research and development, computer aided design frameworks can possibly make a significant impact in the battle against lung cancer, reducing death rates and improving patient results.

In the field of image processing and deep learning, there have been many impressive breakthroughs in medical image analysis to diagnose diseases using medical images. It has proven that automation of medical images recognition is effective for diagnosis and treatment of diseases using CT scan images, mammograms, chest X-beams. [3], [4], every one of these different recognition tasks need a huge quantity of annotated data.

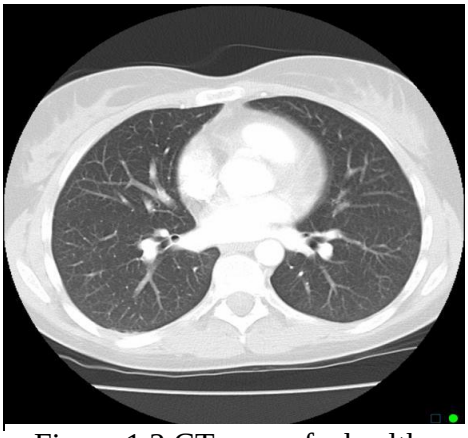


Figure 1.2 CT scan of a healthy person. Taken from https://radiopaedia.org/cases/normal-ct-chest

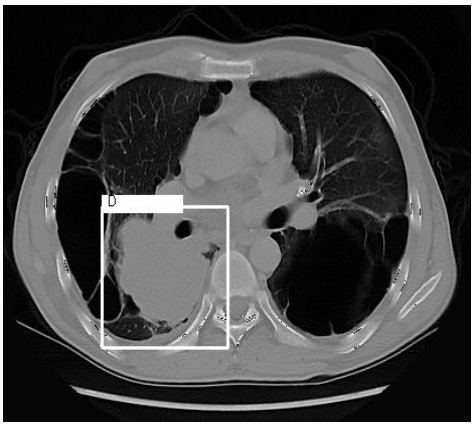


Figure 1.1 CT scan of a patient with Large Cell Carcinoma. Taken from

Deep learning strategies for lung cancer identification utilizing CT images have been proposed by inspired researchers [5], [6] with a lot of marked data. In any case, detecting the lung cancer type is a truly challenging task, since collecting medical images for each kind is infeasible. In certain situations, researchers will most likely be unable to obtain the necessary extensive data that they expect for their examinations because of confidentiality and security concerns. To conquer this challenge, there has been increasing sub-region in machine learning called few-shot learning. The point of this approach is to accomplish great learning results in spite of having a restricted amount of marked data in the training dataset, which contains instances of inputs matched with their respective results.

The main objective of few-shot learning is to construct accurate machine learning models with less training data, and this can be considered as a Meta learning issue. There are two datasets in a few-shot learning which are called support set and query set. Support set and query set is similarly little from the datasets that are utilized for traditional deep learning strategies, however it has tests from each class.

Few-shot learning is to train a function that predicts similarity. Upon completion of training, the gained similarity function can be used to make predictions for new questions. The similarity function can be applied to contrast the query and each example in the support set, and subsequently determine the similarity scores. The example with the most noteworthy similarity score can then be selected and utilized as the prediction.

This research is driven by the urgent need to utilize advanced technologies to assist with improving healthcare where there isn't huge load of cash accessible. In countries like Sri Lanka, which are as yet developing, there aren't enough fancy medical tools and concentrated knowledge. This makes it elusive out what diseases individuals accurately and quickly have, like lung cancer. This is even more important when it comes to figuring out the exact sort of lung cancer, like adenocarcinoma, squamous cell carcinoma, and little cell carcinoma. Knowing the exact kind is important, so doctors can plan the best treatment.

Due to the financial issues in the healthcare industry in Sri Lanka, it's important to find new ways to work with what we have. This is where using a few-shot learning approach comes in. This technique assists us with distinguishing between different types of lung cancer using simply a modest quantity of information. Even however we have relatively little guides to learn from, this approach allows us to be truly accurate in our predictions. It is like doing a ton with only a tad, and it is an effective method for making sure we are as yet finding the right answers even when we don't have a ton of assets.

1.2 Problem Statement

Lung cancer is a devastating condition that emerges from the uncontrolled proliferation of abnormal cells within the lung tissue, disrupting its customary functioning and structure. Left untreated, these aberrant cells continue to increase, forming cancers that adversely affect the lungs, essential for oxygen supply to the circulation system. Lung cancer is sorted into two main types: non-small cell lung cancer (NSCLC) and small cell lung cancer, with NSCLC being the most prevalent, as detailed in [7]. NSCLC further involves three distinct subtypes:

Adenocarcinoma is a common type of lung cancer originating from glandular cells that line different organs, including the lungs, colon, pancreas, and prostate. In the context of lung cancer, adenocarcinoma constitutes a significant portion, accounting for roughly 40% of all cases. Ordinarily, it originates in the external regions of the lungs, with the potential to metastasize to other body parts.

Small cell carcinoma is another kind of lung cancer, characterized by little quickly dividing cells that can quickly spread to different organs. In spite of the fact that it principally affects the lungs, it can likewise foster in organs like the prostate or gastrointestinal tract. These forceful cells necessitate early diagnosis and intervention.

Squamous cell carcinoma begins in level cells known as squamous cells, which line different organs, including the skin, lungs, and gastrointestinal system. Within the context of non-small cell lung cancer, squamous cell carcinoma accounts for around 25-30% of cases.

Non-small cell lung cancer likewise encompasses the enormous cell carcinoma subtype, which can foster in any piece of the lung. Characterized by abnormally enormous cells display fast development and division, with the potential for quick metastasis to different organs.

When lung cancer is suspected in a patient, an initial diagnostic method often involves a chest X-ray. While this imaging technique offers preliminary insights into the presence and seriousness of lung abnormalities, it may not give the degree of detail

expected for accurate cancer identification. Additional advanced imaging strategies, like processed tomography (CT) scans, magnetic resonance imaging (X-ray), or positron emission tomography (PET) scans, might be necessary. These advanced techniques work with the assessment of cancer size, shape, location, and the chance of spreading to neighboring tissues or organs. Moreover, they assume a vital part in devising treatment strategies and monitoring the effectiveness of administered treatments.

For additional statistical data and comprehensive insights into the prevalence and impact of lung cancer, researchers have dug into extensive examinations, a few of which have been distributed in leading scientific journals and conferences, as referenced in the following:

As featured in [7], non-small cell lung cancer is the dominant type of the disease, accounting for most of cases. Research conducted by A. Meldo and L. Utkin [2] introduces innovative ways to deal with lung diagnosis using neural networks. The concentrate by N. Wu, J. Phang, J. Standard, and Y. Shen [3] underscores the potential of deep neural networks in improving the performance of radiologists in cancer screening.

Y. Tang, Y. Tang, J. Xiao, and R. M. Summers [4] present a powerful lung segmentor for chest X-rays, demonstrating the significance of accurate lung segmentation.The research by P. Chathurvedi, A. Jhamp, M. Vanani, and V. Nemade [5] centers around the prediction and classification of lung cancer using machine learning techniques.

S. Makaju, P. Prasad, A. Alsadoon, A. Singh, and A. Elchouemi [6] dig into the detection of lung cancer using CT scan images, emphasizing the practicality of simulated intelligence in early diagnosis.

These examinations and references collectively underscore the significance of ahead of schedule and accurate lung cancer diagnosis and the job of advanced imaging techniques and man-made intelligence in achieving this objective. The prevalence and impact of lung cancer continue to drive research endeavors pointed toward improving patient results and reducing death rates.

Contrasted with different cancers, lung cancers don't show clear side effects in the beginning phases. However, when the patients have their most memorable side effects, doctors ought to do a more precise analysis by using chest X-beam or low portion high resolution lung CT screening or Magnetic Resonance Imaging (X-ray). In comparison to conventional radiography, Processed Tomography (CT) has proven to be a better tool for lung screening due than its capacity to deliver exceptionally point by point, high-resolution images. By utilizing CT, conceivable to detect beginning phase lesions are too little to ever be visualized through traditional X-beam strategies. This advanced imaging technique gives clinicians a more accurate and comprehensive assessment of the lungs, enabling prior diagnosis and more effective treatment of lung cancer.

CT scan images have been broadly used to diagnose an assortment of lung issues that consist of pneumoconiosis, pneumonia, pulmonary edema, and lung cancer. The sort of lung cancer should be determined to plan treatment. Notwithstanding, as a result of the enormous amount of data collected by CT scans, radiologists find it very challenging to identify the accurate kind of cancer. Trouble in accurately identifying cancerous cells can emerge from various factors, for example, variations in intensity on CT scan images and potential misinterpretation of anatomical structures by doctors and radiologists [8].Since different radiologists have different experience level, they made inconsistent judgment. Recently, computer-aided diagnosis has turned into a promising tool, which helps medical specialists to detect cancer accurately [9]. The vast majority of the frameworks don't detect the kind of cancer using CT scan images. In this manner, it is basic to foster new practices to accurately detect the particular sort of lung cancer that is nearest to the patient's condition when a CT scan is given.

Deep learning, machine learning, and transfer learning have advanced medical picture analysis to a remarkable level. Since these techniques have become equivalent to humans at categorizing images and identifying objects. For instance, a newly introduced deep learning model from Google has demonstrated exceptional effectiveness in diagnosing diabetic retinopathy, a condition that is induced by diabetes and can bring about serious vision impairment.

It previously stood out from a few medical communities [10]. In practice, notwithstanding, deep learning techniques ought to have a lot of named data to obtain accurate outcomes. When the amount of data is restricted, deep learning models tend to overfit. The model cannot generalize from unseen data models because of over-fitting, which brings about terrible showing in the model.

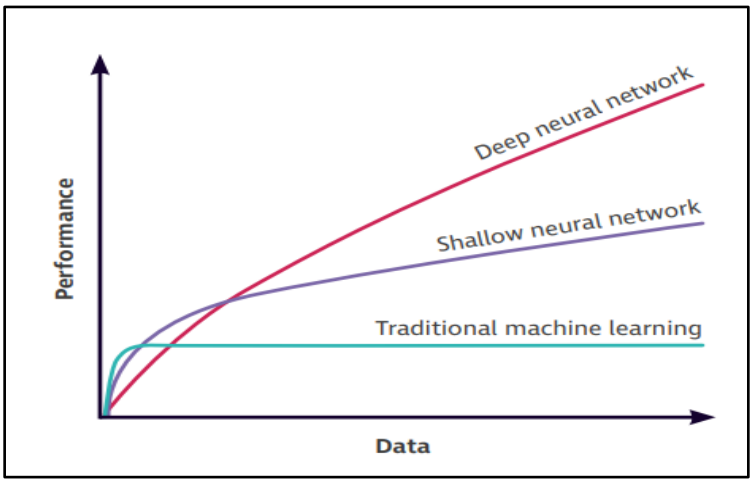


Figure 1.3 A comparison of the performance between traditional machine learning and neural networks under condition of limited training data. Adapted from [10]

Rouse researchers have been ready to achieve a few achievements with average image classification using deep learning techniques for lung cancers [11]. Clearly a similar technology can be applied to identify each sort of lung cancer using CT scan images.

In any case, there is a common occurrence of having a set number of datasets in medical image analysis. Since patient health data is very much protected by data regulations. To apply a deep learning technique, a substantial number of CT scan images belonging to each kind of cancer are necessary. Figuring out the enormous number of marked data for each type is truly challenging. Hence, center around deep learning techniques for ease diagnosis of the beginning phase of lung cancer with satisfactory accuracy. To resolve this issue, we will concentrate on a few-shot learning technique, which is a classification approach that arrangements with situations where there are only a set number of models accessible per class. The figure below

demonstrates how the performance of the deep neural network model changes according to the size of the data.

The point of this research project is to investigate the few shot learning approaches for lung CT scan image analysis. Even more explicitly I center around the following main research questions.

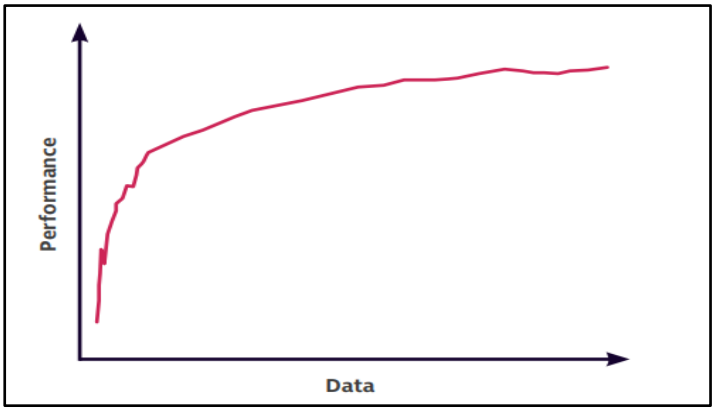


Figure 1.4 The performance of a deep neural network when training data is limited. Adapted from [10]

How could the application of Prototypical Networks and Siamese Neural Networks contribute to the identification of lung cancer types using CT scans?

Can the utilization of pre-trained models for extracting features from CT images effectively enhance accuracy in classification?

Through these research questions, I might want to investigate the feature extraction techniques that can be utilized to extract the features from lung CT images like VGG16, Dense Net.

1.3 Research Gap

Accurate identification and classification of different types of lung cancer are vital in customizing effective treatment strategies and eventually enhancing patient results. Recent advancements in deep learning have shown considerable commitment in the domain of medical imaging analysis, especially through the application of few-shot learning techniques, as evidenced in different examinations [12], [13], [14].

While earlier research has separately investigated the utilization of prototypical networks [15] and Siamese networks [12] in the context of medical image analysis, a comprehensive relative analysis of their respective strengths and limitations in the particular context of lung cancer type classification is notably absent.

In the existing group of research connected with lung cancer detection, many examinations have depended on deep learning techniques with extensive datasets [16], [6]. Be that as it may, the information gave here doesn't present a committed report zeroed in on lung cancer type detection. Besides, only a predetermined number of investigations have dug into the application of few-shot learning algorithms for the analysis of lung CT images [12], [2], [13], [17].

A notable contribution comes from a research group in Russia, which proposed a methodology for detecting abnormal instances of lung disease using a Siamese neural network, falling under the few-shot learning worldview. They harnessed this network to identify the most comparable nodule type for each abnormal case. In any case, it merits highlighting that their review didn't incorporate the utilization of a prototypical network. This observation points to a potential research hole where further exploration can reveal insight into the advantages of integrating prototypical networks into comparable strategies for lung disease detection.

By addressing these research holes and harnessing the potential of few-shot learning strategies with both prototypical and Siamese networks, this study tries to impel the field of medical image analysis, explicitly in the domain of lung cancer type diagnosis. Such advancements can possibly significantly impact patient consideration and results

by ensuring precise and opportune categorization of lung cancer cases, consequently guiding proper and custom fitted treatment draws near.

1.4 Research Objectives

- Exploring the use of a few shots learning method for medical image analysis.
- Implementing a supporting tool for medical professionals to identify lung cancer types using CT scans.
- Identifying proper feature extraction methods for lung CT scan images.

2 LITERATURE REVIEW

2.1 Evaluation of Deep Learning Frameworks for Lung Cancer Classification Using CT scans.

The application of deep learning techniques in the field of medical imaging, especially for the classification of lung cancer using figured tomography (CT) scans, has garnered considerable attention in recent years. Researchers have been actively exploring different deep learning frameworks to enhance the accuracy and efficiency of lung cancer classification, aiming to contribute to more effective diagnosis and treatment strategies.

Meldo and Utkin's review [2] introduced a novel way to deal with differential lung diagnosis, which utilized a Siamese neural network. This strategy demonstrated the potential of deep learning in improving lung cancer classification, suggesting that such methodologies hold guarantee in enhancing diagnostic accuracy.

Besides, the concentrate by Tang et al. [4] introduced "XLSor," a strong lung segmentor for chest X-rays, emphasizing the basic job of precise lung segmentation in lung cancer classification using deep learning frameworks. Their work featured the practicality of leveraging deep learning for the accurate interpretation of medical images.

In a comparative vein, Chathurvedi et al. [5] dug into the prediction and classification of lung cancer using machine learning techniques. While their center was more extensive, encompassing different machine learning draws near, this review demonstrated the growing interest in harnessing advanced technology to further develop lung cancer classification.

The potential of deep learning for lung cancer classification was likewise tended to in the research by Makaju et al. [6], who investigated lung cancer detection using CT scan images. Their findings shed light on the utilization of artificial intelligence in early diagnosis and classification of lung cancer, emphasizing the significant contributions of deep learning frameworks.

Concurrently, the deep learning community has shown growing interest in few-shot learning techniques, as evidenced in examinations by Ahuja et al. [12] and Jiang et al. [13]. While these examinations essentially centered around Coronavirus classification and diagnostic few-shot learning, their systems are exceptionally relevant to the more extensive field of medical image analysis, including lung cancer classification. Few-shot learning, combined with deep learning frameworks, can possibly address the challenges presented by restricted data availability in medical imaging.

Notably, the concentration by Drozdzal et al. [51] underlined the importance of skip connections in biomedical image segmentation, which can be a basic component in deep learning frameworks for lung cancer classification. Skip connections enhance the information flow within neural networks, improving their capacity to arrange medical images accurately.

In conclusion, the writing on the evaluation of deep learning frameworks for lung cancer classification using CT scans is extensive and continually evolving. The given references, including [2], [4], [5], [6], [12], [13], and [51], collectively underline the growing interest in applying deep learning and few-shot learning techniques to advance lung cancer classification. These advancements are ready to altogether affect healthcare by enabling more accurate and opportune diagnosis, at last enhancing patient results.

Lung cancer identification has been demonstrated by researchers to be accomplished through the application of a few deep learning strategies. The paper named "Deep Learning Techniques to Diagnose Lung Cancer " by Wang Humdinger [18] has examined the deep learning-based imaging approaches for detecting lung cancer. Convolutional Neural Network, Deep Convolutional Neural Network and Recurrent Neural Network are extensively involved unsupervised learning algorithms in medical images. Among these, CNNs are especially famous for managed tasks like lesion segmentation and classification because of their capacity to handle crude image data with minimal pre-processing requirements.

A specific deep learning architecture is utilized to take care of the pre-handled images for training and testing, determined to enhance the accuracy and generalization performance of deep learning algorithms for lung cancer. Be that as it may, image noise can negatively impact the precision of the final accuracy. To resolve this issue a few noise reduction strategies have been utilized .These techniques include the median channel [19] which diminish noise by replacing every pixel's worth with the median worth of the neighboring pixel; Wiener channel [20] and on-nearby means channel [21] which midpoints pixel values in light of the resemblance of patches or neighborhoods.

Performance metrics can be utilized to assess how well deep learning algorithm identify lung cancers. Accuracy Measure the proportion of correctly predicted instances to the absolute number of instances. Mean Squared Error (MSE) is a metric used to assess the performance of regression models. It estimates the normal of the squared differences between the actual qualities and the predicted values. Receiver operating characteristic (ROC) bend is a graphical representation of a binary classifier framework's diagnostic capacities as its discrimination limit is changed. The ROC bend is a depiction of the genuine positive rate versus the misleading positive rate depending on limit levels. The Jaccard coefficient, which is determined by dividing the size of the intersection by the size of the union of test sets, is another performance measure used to quantify the similarity of finite sets of tests. Over-segmentation rate (OR), under-segmentation rate (UR), Dice similarity coefficient (DSC), Jaccard Score (JS), Asymmetric surface distance (ASD), Hausdorff distance (MHD), and intersection

over union (IoU) are some of other performance metrics that can use for performance evaluation. These metrics quantify algorithm performance in different ways, including classification accuracy, detection sensitivity, segmentation quality, and all out-prediction blunder.

A few deep learning algorithms for lung segmentation have been researched, since it assumes a significant part in medical images for lesion detection, including chest extraction (eliminates artifacts) and lung extraction (identifies the left and right lungs). Sun et al [22] fostered a two-stage computer-aided detection framework for programmed segmentation of lung nodules and reducing misleading up-sides (FP). The principal phase of the framework centers around identifying and segmenting the nodules, while the second stage intends to minimize the number of FP findings. The framework's performance was examined by four experienced radiologists using the LIDC-IDRI dataset. The framework earned a normal F1_score of 85% for lung nodule segmentation, according to the data. Table 2.1 shows recently created lung segmented approaches with their performance. All the below approaches have utilized CT imaging and every one of them accomplished over 72% accuracy.

Reference	Methods	Datasets	Results
[23]	Support Vector Machines	Shiraz University of Medical Sciences	Accuracy: 98.1%
[24]	Improved U-Net	LUNA16	DSC: 73.6%
[16]	U-Net	712 lung cancer patients operated in Uppsala Hospital, Stanford TMA cores	Precision: 80%
[25]	Improved U-Net	LIDC-IDRI	Precision: 84.91%
[26]	Support Vector Machine	193 CT images	Accuracy: 94.67% for benign tumors.

			Accuracy: 96.07% for adhesion tumor

Table 2.1 Lung segmentation approaches and their results from previous work.

Among the methodologies considered, SVM frameworks accomplished an accuracy range of 94-98.1%, while U-Net segmentation frameworks demonstrated a precision range of 80% to 85%.

Albeit deep learning frameworks have gained considerable headway in an assortment of medical use cases, they are as yet constrained by the issue of dependence on the availability of training data. Few shots learning is a subject of active review to conquer this restriction.

2.2 Few-Shot Learning

Few-shot learning, a subset of machine learning, has arisen as a compelling way to deal with addressing the challenges presented by restricted data availability and improving the generalization capacities of models in different domains. This writing survey investigates the evolving landscape of few-shot learning, its philosophies, and applications in different fields, drawing insights from relevant IEEE references, for example, [12], [13], [15], [28], and [30].

Few-shot learning is characterized by its capacity to enable models to make accurate predictions or classifications in light of only few training models. This change in perspective has gained significant attention because of its potential to expand the pertinence of machine learning in scenarios where traditional strategies might battle.

Prototypical Networks, introduced by Snell et al. [42], play had a significant impact in advancing few-shot learning. These networks enable models to learn representations of classes and make predictions in light of the similarity of new instances to class prototypes. This approach has found applications in different domains, including image recognition, natural language processing, and medical image analysis.

The Siamese network architecture, as exemplified in Koch's work [33], has additionally contributed to few-shot learning. Siamese networks are especially significant for tasks involving similarity or dissimilarity measurement. Their applications span signature verification, image recognition, and even medical image analysis, as demonstrated in the research by Ahuja et al. [12].

Few-shot learning has shown guarantee in the field of medical image analysis, as evidenced in the concentrate by Keshani et al. [30]. Their work centers around glaucoma diagnosis using a little measured dataset of high-resolution fundus images, emphasizing the potential of few-shot learning techniques in addressing challenges related with restricted medical image data.

Kotia et al's. study [31] investigates the utilization of few-shot learning for medical imaging, shedding light on its applications in healthcare. The creators dive into applications, for example, identifying anatomical structures and landmark detection, emphasizing the potential of few-shot learning in enhancing medical image analysis.

In the context of deep learning, Siamese networks, as portrayed by Yi et al. [32], have demonstrated their adequacy in tasks requiring similarity measurement. The potential of Siamese networks isn't restricted to explicit domains, and their applications have expanded to assorted regions, including medical image analysis.

Few-shot learning has additionally been investigated in the context of radar image recognition, as presented by Liu et al. [28]. This work exhibits the flexibility of few-shot learning techniques to different data types and domains, underscoring its adaptability.

Few-shot learning has introduced a new period of machine learning, offering solutions to data shortage, and extending the compass of predictive models across different domains. The references gave, including [12], [13], [15], [28], and [30], underscore the significance of few-shot learning and its applications in image recognition, medical image analysis, and other assorted fields. The continued exploration and refinement of few-shot learning techniques hold the commitment of addressing complex certifiable challenges where data is restricted, at last advancing the capacities of machine learning

frameworks. Further investigation and cross-domain applications are encouraged to unlock the maximum capacity of few-shot learning.

The Prototypical Network is a famous meta-learning approach for few-shot learning that utilizes a nearest neighbor strategy, eliminating the need for hyper-parameters in the meta-test stage and resulting in practically insignificant inference time. The effectiveness of the Prototypical Network is evaluated by the proportion of between-class variation to within-class variation of features in a support set generated by the meta-trained model.

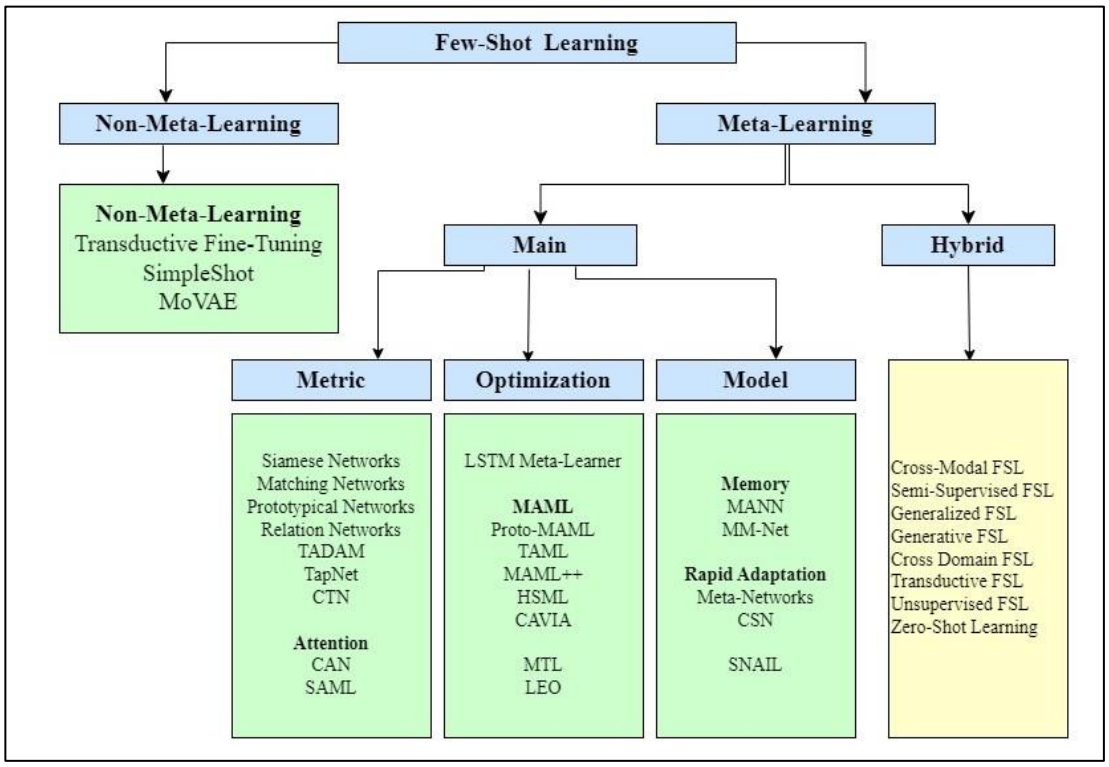


Figure 2.1 Approaches in Few Shots Learning. Figure taken from [18]

Likewise, no model learning is expected during the model classifier's testing stage, allowing it to be utilized with any learned feature extractor. In this way, the model classifier is a helpful and effective initial phase in resolving a few-shot learning issue.

The paper named "A More critical Look at Model Classifier for Few-shot Image Classification" by Mingcheng Hou and Issei Sato presents a careful analysis of model classifiers for image classification in view of FSL [27]. The creators have presented

their exact analysis of the model classifier, which includes a progression of experiments on benchmark few-shot image classification datasets. They investigate different aspects of the model classifier, like the impact of different model representations, the effect of the number of support models, and the importance of distance metrics. Their findings demonstrate that modifying the model classifier could significantly support its performance on few-shot picture classification challenges.

One shot learning for radar image recognition with few images has been implemented by a research group at the National University of Defense Technology in Hunan, China. [28]. The initial stage in adjusting radar images to the input of the succeeding network is pre-processing them. The second step is to train a model that guides images to linearly divisible feature spaces using more data. In the last phase of the recommended approach, the mean worth of homogeneous vectors in the feature space is utilized as a model, and the model is trained to bunch the feature vectors in the image into their relevant prototypes. The clustering effect can be applied to the objective data via training the mapping with assistant and target data. The prototypes for every class are matched in view of their Euclidean distance to gauge the class of unknown photographs. The proposed approach was scrutinized using the MSTAR dataset.

Figure 2.2 illustrates that the typical accuracy of the proposed technique by the research group in Hunan is higher than that of deep transfer learning when the model intricacy is equivalent, and the amount of training data is minimal.

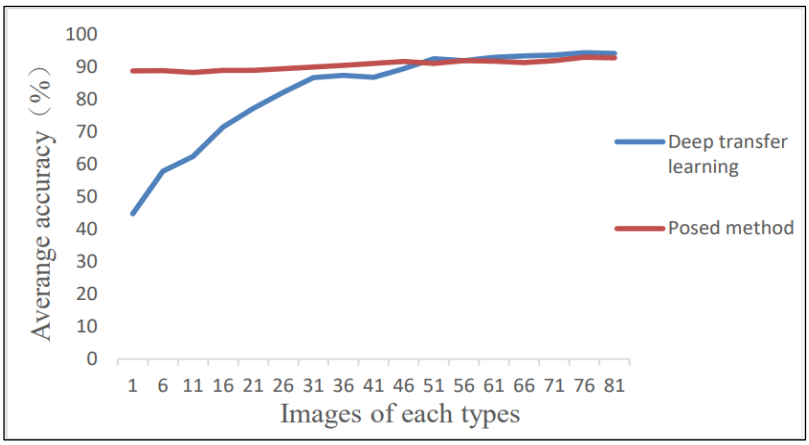


Figure 2.2 Average accuracy based on number of images. Figure taken from *[28]*

To make the result of a hidden layer in their algorithm linearly divisible, they eliminated the hidden layer's non-linear activation function. The image extraction module in their implementation utilizes the convolutional layer parameters of the VGG-16 network trained on the ImageNet dataset. They thought about the deep transfer learning and proposed strategies' accuracy by progressively increasing the number of photographs in each class.

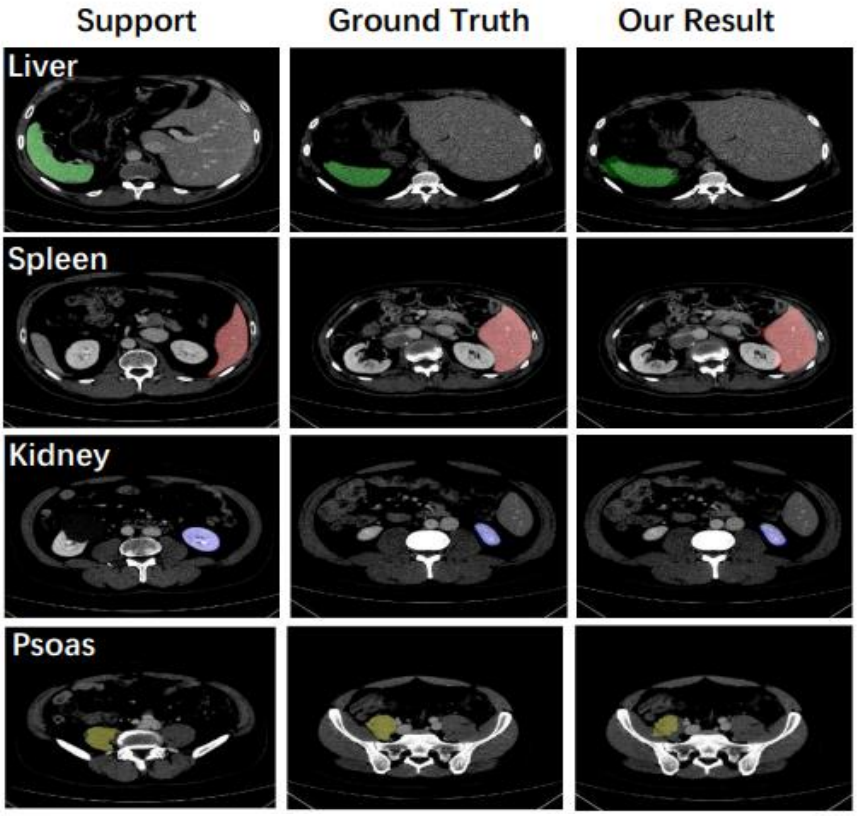


Figure 2.3 Comparison of segmented visualization with propped model based on FSL. Figure taken from *[29]*

Recently, researchers from the United States and China have proposed a novel way to deal with few-shot learning, named the Location-Sensitive Neighborhood Model Network, which considers basic parameters, for example, matrix size [29]. Their technique consists of two main advances. Extraction of location-sensitive nearby prototypes is the initial stage, and it is achieved by running support images through a feature encoder to make the relevant feature maps. Then, neighborhood prototypes are processed by combining the support image and masked midpoints. In the second stage, if the support and query photographs have a comparative spatial format, the neighborhood prototypes that were extracted from the support set can be matched to their respective locations in the query images. This enables network based few-shot segmentation.

For every one of their experiments, they have utilized Instinctive dataset [15] which include contrast-enhanced CT scans. For six organ classes — liver, spleen, left kidney, right kidney, left psoas, and right psoas — they have conducted 1-way, 1-shot preliminaries. As the feature encoder, they used ResNet-50, which was trained beforehand on ImageNet. In the segment visualizations showed in Figure 2.3, their algorithm has effectively transferred semantic information from the support to the query images to yield excellent segmentation results in spite of variations in organ sizes and shapes. Figure 2.3 presentations a comparison: on the left column, support images are accompanied by their corresponding ground-truth masks; the center column shows query images with their related ground-truth masks. The right column illustrates the query images, overlaid with the aftereffects of the proposed segmentation strategy.

Aihua Cai, Wenxin Hu, and Jun Zheng have investigated the utilization of few-shot learning techniques for medical image classification tasks, explicitly for identifying liver cancer in CT images [30]. The troubles of classifying medical images are canvassed in the principal section of the paper, along with the high dimensionality and restricted supply of named data. The creators point out that standard convolutional networks experience issues extracting helpful information from medical images since their characteristics are frequently turbulent and shifted. The neural network comprised of four convolutional blocks that fills in as the embedding function f(x) in the proposed strategy is a neural network. The creators include an attention mechanism with spatial and channel attention modules between the third and fourth convolutional blocks to increase the model's ability for representation. These modules assist with extracting the image's spatial and channel information in turn.

Additionally, the creators add 1x1 convolution kernels to the convolutional blocks to deepen the network without sacrificing resolution and enhance non-linear properties. The recommended strategy is surveyed using two medical imaging datasets and contrasted with other cutting-edge techniques. The experimental outcomes demonstrate the efficiency of the attention mechanism in enhancing the Prototypical Network for classifying medical images, since the proposed technique achieves more classification accuracy than the baseline strategies. The review offers a potential strategy for classifying medical images that makes utilization of both Prototypical

Networks and attention mechanisms. The creators show that their strategy beats other cutting-edge techniques as far as classification accuracy by effectively managing the noisy and various aspects of medical images. The model's architecture is depicted in the figure below.

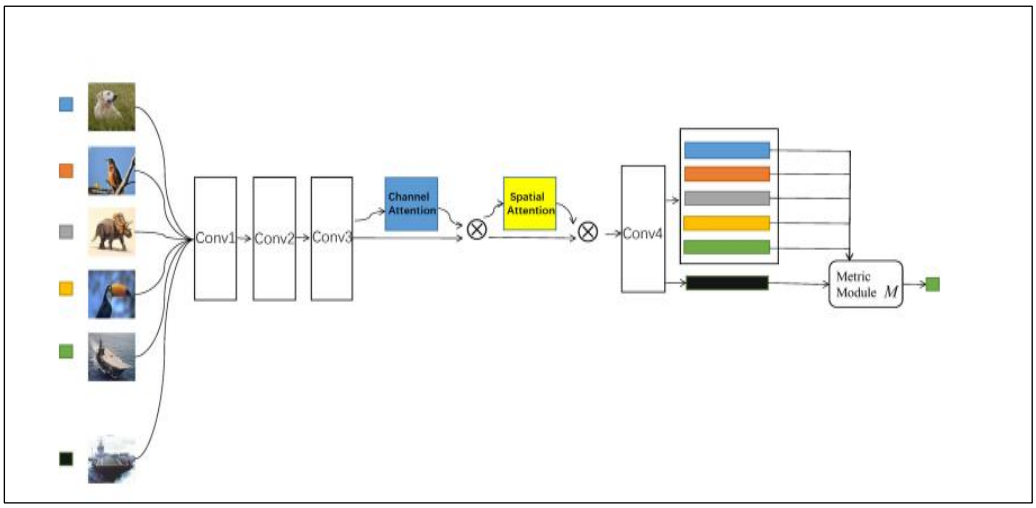


Figure 2.4 Proposed architecture for few shots learning by Ailua and team. Figure taken from *[30]*

Finally, they assessed the accuracy of their model to pre-trained models with changed classification layers. The ImageNet dataset, which contains genuine images, is utilized to train the parameters of the pre-trained models. In any case, when utilized for medical imaging applications, these models perform less well than they ought to since medical images vary from natural photographs on ImageNet.

Method	Accuracy
AlexNet	79.80%
VGG16	82.08%
ResNet101	82.57%
DenseNet169	83.06%
Dual Path Network	82.74%
New Method	92.44%

Table 2.2 Accuracies based on the feature extraction method. Data taken from *[30]*

A Siamese Neural Network belongs to a class of neural network architectures that incorporate at least two identical subnetworks. These subnetworks are characterized by similar configurations, parameters, and loads. Initially introduced by Bromley in 1993, this network architecture was basically concocted for signature verification purposes [31]. The Siamese network is a helpful tool for enhancing data in few-shot learning. The Siamese network can make additional training instances by comparing sets of training data, increasing FSL's accuracy and resilience. When contrasted with just training the model on the original models, this technique is particularly advantageous since it gives the model more fluctuated data.

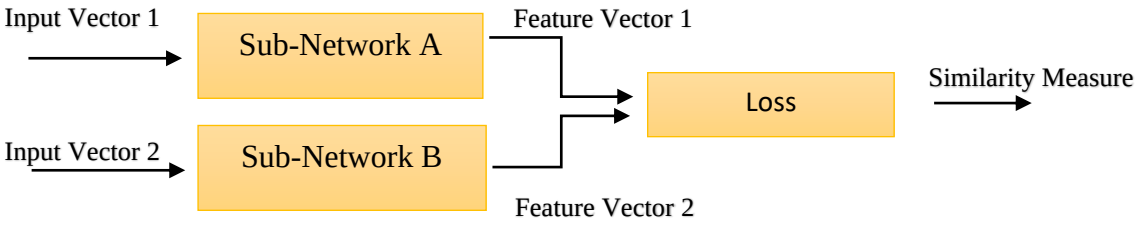


Figure 2.5 Conceptual diagram of Siamese Neural Network.
Concept adapted from [32]

Because of inadequate amounts of training data, a model can frequently become overfitted within a few learning endeavors. By adding training data in the type of creating comparable matches and divergent matches, the Siamese network addresses this issue. The Siamese network can be seen as a representation learning process that seeks to identify low dimensional discriminative characteristics. By doing this, it can deliver effective representations and eliminate extraneous or redundant data from the training set. To put it another way, the Siamese network helps with selecting the most pertinent and significant attributes while ignoring the unnecessary ones, resulting in more effective and efficient training [32].

Gregory, Richard, and Ruslan from the University of Toronto have introduced a novel methodology for one-shot image recognition using Siamese neural networks. The creators feature that for these strategies to work, they often depend on manually made features and a substantial amount of marked data. They contend that a more versatile strategy is required, one that can learn the most recent information from a little example size and consequently identify relevant aspects in the data. The creators

propose a Siamese neural network architecture to conquer these issues. Siamese networks are comprised of two identical neural networks, every one of which generates a low-dimensional embedding from an input image. To survey in the event that the two input images are comparative or different, the two embeddings are then looked at using a distance measure [33].

They have demonstrated the effectiveness of their methodology on a few one-shot recognition tasks, including handwritten character recognition and face verification. They demonstrate the way that their Siamese network can accomplish cutting edge results with only one or a few models for every class.

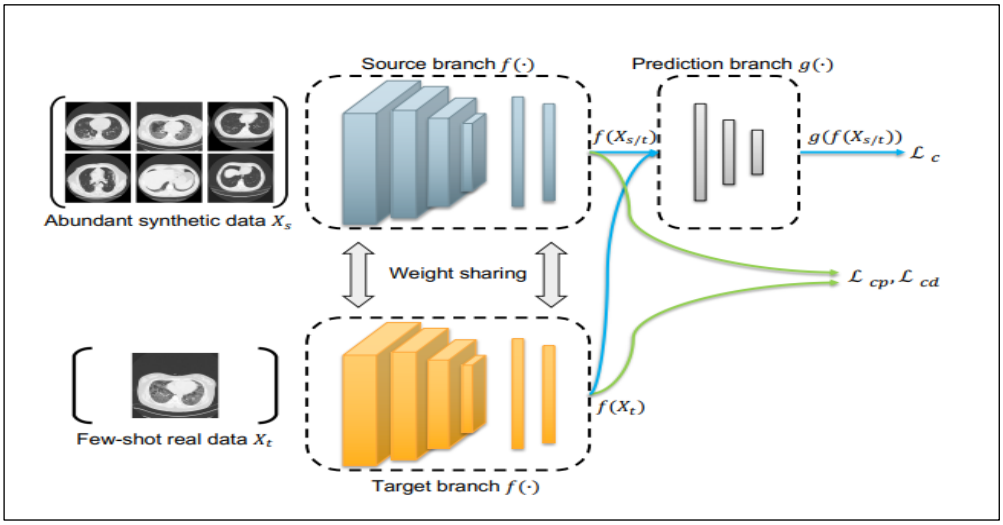
2.2.1 Few shots Learning for CT scan imaging.

Few-shot learning has found different applications in the field of medical imaging, proving especially important in scenarios where named data is scant. Yifan Jian and their research group have introduced an innovative way to deal with Coronavirus CT diagnostics, leveraging regulated domain adaptation techniques [13]. This technique offers significant advantages when only a predetermined number of named CT images are accessible, a common challenge in medical imaging applications.

The proposed model is constructed on the foundation of a Siamese network architecture, encompassing three key components: the source and target branches, both characterized by a consistent structure featuring a feature extractor and two completely connected layers, and the prediction branch, which includes three completely connected layers. During the training stage, a combination of synthetic and genuine CT images is utilized, with weight sharing between the source and target branches. The network succeeds at extracting feature vectors from CT images, which are then used to process classification and cross-domain losses. In the testing stage, the model generates binary diagnostic decisions in light of genuine CT images, a significant stage in the diagnostic cycle.

In the experimental evaluation, the model's diagnostic performance is surveyed using two imperative metrics: accuracy and F1 score. To ensure the robustness and

unwavering quality of the outcomes, the evaluation interaction is performed across ten random data resampling iterations, leading to the creation of ten distinct training sets. The presentation of results embraces the "MEAN ± 95% CONFIDENCE INTERVAL" design, reflecting the model's performance consistency across these ten folds. Preprocessing of CT scans involves converting them into grayscale images using a



*F*igure 2.6 Proposed method for CT scan based covid-19 diagnosis. Figure taken from *[13]*

Hounsfield unit (HU) scale ranging from - 600 to 1500. Additionally, the images are resized to a uniform dimension of 512 × 512 pixels, ensuring standardized data input.

To streamline the model during the training system, a learning rate of 0.001 is utilized, accompanied by a rot rate of 0.95. This fine-tuning strategy enhances the network's performance as it learns from the data. Besides, a weight factor α of 0.25 is introduced to harmonize the contributions of different losses during the optimization cycle, ensuring that every component is fittingly considered in the model's learning cycle.

In synopsis, the application of few-shot learning in medical imaging, as demonstrated by Yifan Jian and their research group, underscores its ability to tackle data limitations, especially in scenarios like Coronavirus CT diagnostics. The innovative Siamese network-based approach displays potential for improving diagnostic accuracy and has been thoroughly assessed through powerful experimental systems. Such advancements contribute to the more extensive field of medical image analysis, paving the way for additional effective and dependable diagnostic tools.

Overall, the experimental setup ensures the robustness of the model's performance evaluation by resampling the data and provides clear and concise reporting of the results using both accuracy and F1 score metrics with confidence intervals.

A research team in India proposed an [12] implementation of P-shot –ways Siamese network, based on deep learning principles combined with prototypical nearest neighbor classifiers. This approach aims to accurately classify COVID-19 infection in lung CT scan slices. Chest CT images that are considered in this implementation are in NIFTI format with a size of 512 x 512 and an average of 36 slices per patient. To make the input dataset compatible with the suggested training network, the input picture dataset must undergo pre-processing. Selection of axial slices, resizing operation, normalization and contrast enhancement are the steps of preprocessing pipeline.

This technique investigated that different pre-trained network CNN models affect the performance of multi-class classification of Siamese based networks. In the learning framework, a limit of 15 images have been used for every infection class. To survey the effectiveness of the presented strategy, a comprehensive evaluation of worldwide performance metrics is undertaken, encompassing full scale accuracy (μ_1), full scale sensitivity (μ_2), full scale explicitness (μ_3), full scale precision (μ_4), large scale F1-score (μ_5), and large scale MCC (μ_6). The subtleties of these worldwide performance evaluation parameters for a 15-shot 5-ways multi-class classification framework are given in Figure 2.7, reflecting the results of five random preliminary attempts. These metrics collectively approve the consistent performance of the multi-class classification framework, which is inspired and driven by a deep learning encoder.

Trial	Performance parameters (%)					
	μ_1	μ_2	μ_3	μ_4	μ_5	μ_6
1	97.24	93.86	98.31	92.78	93.05	91.53
2	**98.07**	95.66	**98.83**	**94.90**	**95.10**	**94.03**
3	97.99	**95.78**	98.78	94.86	95.01	93.99
4	97.57	94.18	98.49	93.70	93.88	92.40
5	96.57	92.28	97.88	90.96	91.37	89.40
Average	97.48	94.35	98.45	93.44	93.68	92.27

Figure 2.7 Global performance metrics results evaluation for five trials with the base encoder ResNet18 on the MosMed dataset.

Upon application to the MosMed dataset, the proposed approach accomplished noteworthy outcomes, with a typical accuracy of 97.48%, explicitness of 98.45%, sensitivity of 94.35%, F1 score of 93.68%, precision of 93.44%, and MCC of 92.27%.

2.2.2 Few Shots Learning for Lung Cancer Identification using CT scans.

Nicholas and their research group, as documented in their work [17], have introduced an innovative dynamic few-shot learning framework custom fitted for lung cancer lesion segmentation. This framework consistently integrates the force of few-shot learning with the U-Net architecture, presenting a novel way to deal with lung cancer detection and classification. What sets their procedure separated from existing methodologies is the worldwide neighborhood PET/CT methodology fusion, a technique that veers off from conventional strategies, which regularly exploit separate PET and CT features or perform restricted fusion using CNN structures. By combining the all-encompassing anatomical and functional properties with the limited characteristics of every methodology, the worldwide nearby data fusion strategy lifts the performance of detection and classification. At the core of this fusion cycle is the U-Net model, chosen for its ability in lung cancer lesion segmentation and its capacity to conduct worldwide nearby operations on input data. This fusion strategy, reinforced by the strengths of U-Net models, is ready to enhance the precision and accuracy of lung cancer detection and classification.

Besides, their research group has introduced a dynamic PET/CT methodology fusion approach in view of few-shot learning. This dynamic methodology incorporates model weight adjustments by integrating misclassified tests, fostering a user-centric simulated intelligence worldview. The selection of U-Net architecture comes from its proven effectiveness and commitment in delivering precise segmentation within the domain of medical imaging. Figure 2.8 visually outlines the architecture of their few-shot learning U-Net framework.

To check the performance of their proposed technique, the research group has utilized a comprehensive set of evaluation metrics, including Accuracy, Precision, Recall, F1

score, Intersection over Union, and Region under the Collector Operating Characteristic bend (AUC-ROC). These metrics collectively give a strong assessment of the model's performance in lung cancer lesion segmentation, ensuring a comprehensive understanding of its capacities.

Meanwhile, a research group situated in Petersburg, Russia [2], has introduced an aggressive methodology pointed toward advancing the differential diagnosis of different lung conditions and providing a comprehensive classification of a wide range of lung cancer. Their proposed solution spins around the development of a Computer-Aided Diagnosis framework, leveraging the capacities of a Siamese neural network. This network is designed to undergo training on a particular dataset comprising abnormal lung objects, fastidiously segmented and named, with a particular spotlight on growths. The dataset is mindfully sorted into distinct gatherings, including "run of the mill" fringe LC, "abnormal" LC, and "not cancer," in light of discernible CT image patterns. Notably, the dataset solely encompasses confirmed tissues approved through careful or histological investigations, ensuring the most significant level of data integrity.

In conclusion, the works of both Nicholas et al. [17] and the research group from Petersburg, Russia [2] contribute to the consistently evolving landscape of medical image analysis and lung cancer diagnosis. Their innovative philosophies, combining few-shot learning, U-Net architecture, and Siamese neural networks, vow to enhance the accuracy and precision of lung cancer lesion segmentation and differential diagnosis. The thorough evaluation metrics utilized in their examinations underscore the commitment to delivering solid and strong solutions in the domain of medical imaging.

To dig deeper into the subtleties of these groundbreaking methodologies and their implications, referencing the gave IEEE references, [17] and [2], is strongly recommended. These references offer top to bottom insights into the systems and applications of dynamic few-shot learning, worldwide neighborhood PET/CT methodology fusion, and Siamese neural networks in the context of lung cancer detection and diagnosis.

They included two SNNs and the first SNN takes direct images of segmented lung objects as input, while the second SNN takes histograms of the segmented objects. Parallel datasets consisting of histograms are used to speed up the testing process.

During the testing phase, an unknown lung object's image is fed to the first CNN, and a corresponding image from the dataset is used as the input for the second CNN. SNNs compare these images to determine if they are semantically similar. If the images are deemed similar, the diagnosis corresponding to the object from the dataset is made. Otherwise, the comparison process is repeated with the next image from the dataset.

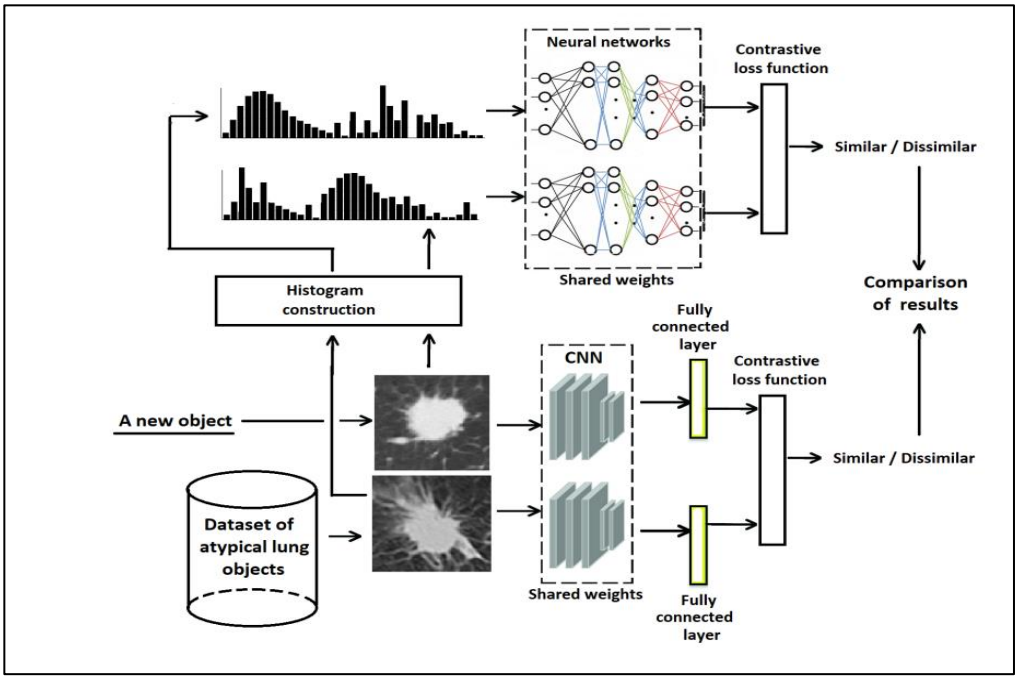


Figure 2.8 Proposed architecture of the CAD system by authors. Taken from *[2]*

The same testing process is also conducted using histograms. By comparing the results obtained from the two SNNs, a more confident diagnosis can be made.

The suggested architecture slightly mimics a doctor's decision-making process, with the SNN sequentially comparing a fresh lung tissue with all existing instances from the dataset, like how a doctor interprets a CT scan using her/his expertise and previously analyzed images.

The table below presents a comparison of discussed proposed methods in the field of few-shot learning for medical imaging, as discussed by research teams worldwide.

Domain	Technique	Dataset	Performance
Diagnosis of COVID-19 through CT scans [13]	Siamese network-based model	CT segmentation dataset from	Accuracy 0.8040±0.0356 F1-score 0.7998±0.0384
Covid-19 infection classification using CT scan slices [12]	Prototypical closest neighbors' classifiers combined with a P-shot N-ways Siamese network	Chest CT scans from 1110 patients in medical hospitals from Moscow ,Russia	Accuracy 98.07 % F1-Score 95.10%
Segmentation of lung cancer lesions in PET/CT images [17]	A model based on a few-shot U-Net architecture	Lung-PET-CT-DX dataset in TCIA database. PET/CT scans from 87 patients	Accuracy 99 % Precision 70.62%
Pre-trained Vision Transformer model to identify lung cancer with multiple labels [34]	Comparison using both Zero-shot learning and Few Shot Learning	LC25000 dataset. 25,000 color images in 5 classes	99.87% of accuracy from few-shot setting

Table 2.3 Analysis of Previous Work on Few-Shot Learning in Medical Imaging

3 METHODOLOGY

This section covers the implementation details of the research.

3.1 System Architecture

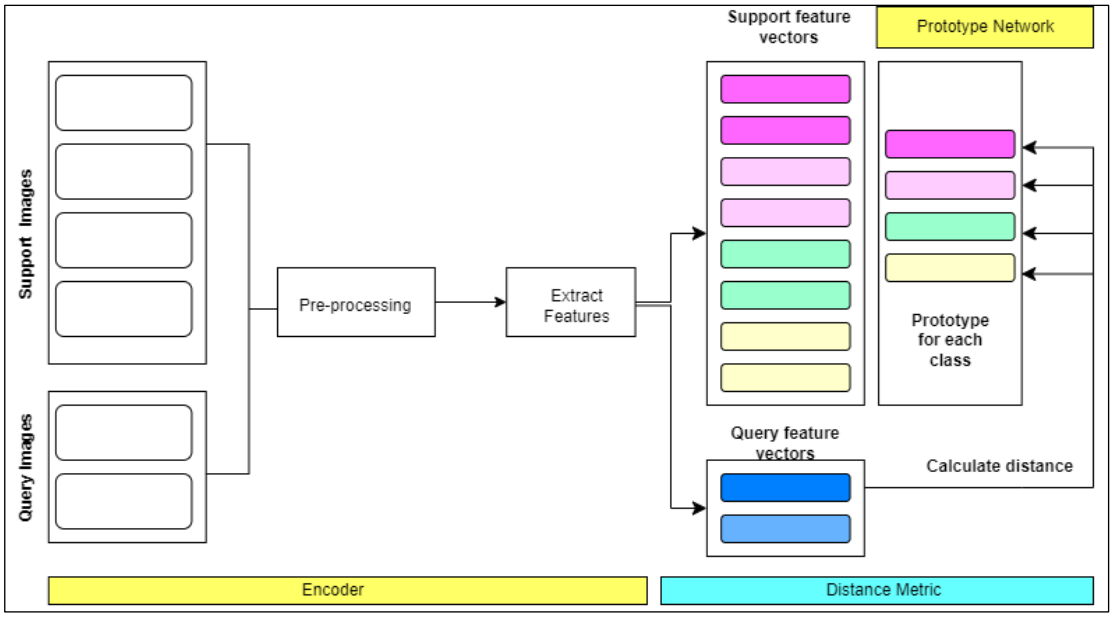


Figure 3.1 System Architecture of the proposed model

The fundamental architecture of a prototypical network contains three key components, each playing a distinct job in the general cycle: an encoder network, a model network, and a distance metric. These components work cooperatively to enable the model's few-shot learning abilities.

In the initial encoder stage, the input data is subjected to encoding, transforming it into a feature space. This cycle involves mapping the crude input into a more abstract and representative structure, facilitating subsequent operations. The encoder network is responsible for executing this pivotal step, generating feature embeddings that catch essential characteristics of the data.

The subsequent stage involves the model network, an essential component in the prototypical network architecture. Here, class prototypes are processed by aggregating the embeddings of support models that share a similar class identity. These support

models, a restricted set of marked instances gave during the model's training, act as the foundation for model creation. The prototypes effectively condense the essential information about each class, serving as reference points for classification tasks.

The final piece of the riddle is the distance metric, an essential component in the prototypical network's decision-making process. This metric operates within the feature space, quantifying the dissimilarity or similarity between the query model and the class prototypes. By measuring the distances, the network can make informed decisions regarding the class assignment of the query model, contributing to the model's few-shot learning proficiency.

This firm interplay between the encoder network, model network, and distance metric furnishes prototypical networks with the ability to tackle few-shot learning challenges, allowing them to effectively generalize from a little set of named models and make informed classifications in view of query instances.

For top to bottom exploration of prototypical networks and their part in few-shot learning, referencing the pertinent IEEE writing, for example, [1] and [2], is recommended. These references give definite insights into the architecture and applications of prototypical networks, shedding light on their significance in different domains, including image recognition, medical diagnosis, and more.

One essential component of this system is image preprocessing. The pre-processing procedures are depicted in figure 3.2. Future chapters will discuss each component.

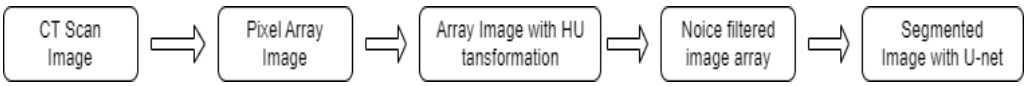


Figure 3.2 Flow of preprocessing

3.2 Data set

The implementation of this study used CT scan images obtained from the Lung-PET-CT-Dx dataset within The Cancer Imaging File (TCIA), an invaluable asset laid out by the National Cancer Institute. TCIA, as an open-access stage, assumes a vital part

in facilitating research advancements and educational initiatives that depend on the utilization of medical imaging data for the investigation of cancer [34].

For the particular objectives of this research, a dataset got from TCIA was utilized, comprising CT and PET-CT DICOM images of patients afflicted by lung cancer. These images were complemented by XML Annotation documents that gave vital information regarding the localization of growths through bounding boxes. The dataset was curated retrospectively, focusing on individuals under suspicion of lung cancer who had undergone both a standard-of-care lung biopsy and PET/CT imaging. It encompasses CT scans of patients diagnosed with different types of lung cancer, including adenocarcinoma, little cell carcinoma, huge cell carcinoma, and squamous cell carcinoma.

The examination of the images was conducted with a double methodology, differentiating between the mediastinum and lung regions. This distinction was worked with by utilizing explicit window width and level settings: 350 Hounsfield Units (HU) for the mediastinum and 1,400 HU for the lung. Notably, the images underwent reconstruction with a cut thickness of 2 mm, and the spacing between CT cuts displayed inconstancy, spanning from 0.625 mm to 5 mm. The imaging system involved different scanning modes, encompassing plain, contrast, and 3D reconstructions, each customized to satisfy explicit diagnostic requirements.

This dataset includes CT images for a sum of 355 subjects, with each subject's data represented as a progression of DICOM images. In request to lay out class marks within this dataset, a naming convention was utilized in light of the initial letter of the patient or subject name. Subjects designated with the letter 'A' were identified as having adenocarcinoma, 'B' as little cell carcinoma, 'E' as huge cell carcinoma, and 'G' as squamous cell carcinoma. The dataset allows for the visualization of image series for each subject, available through the thumbnail watcher option. The figure below gives a visual representation of the series of CT scan images for a particular patient, identified as Patient ID Lung_Dx-A0002.

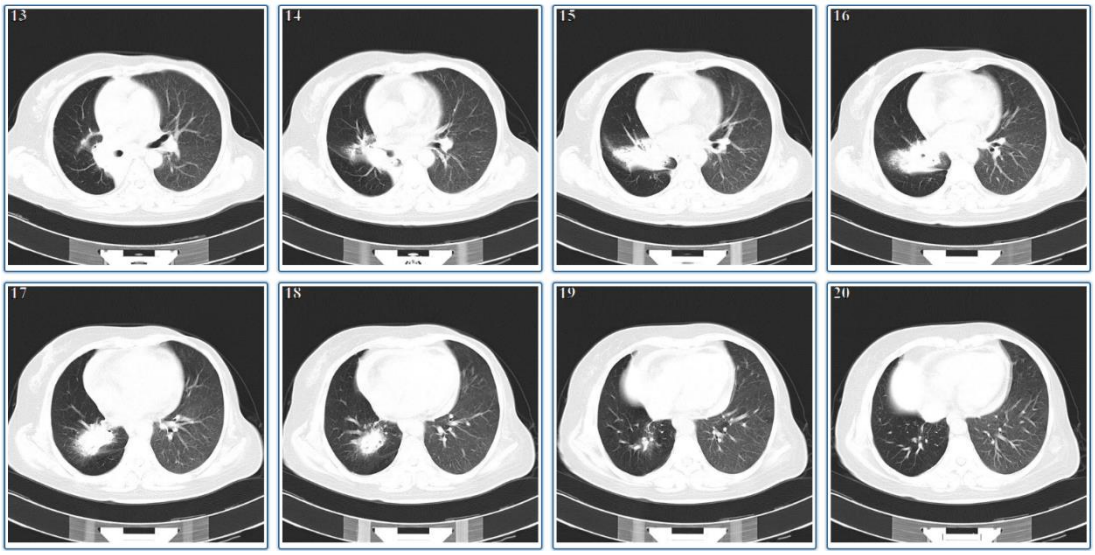
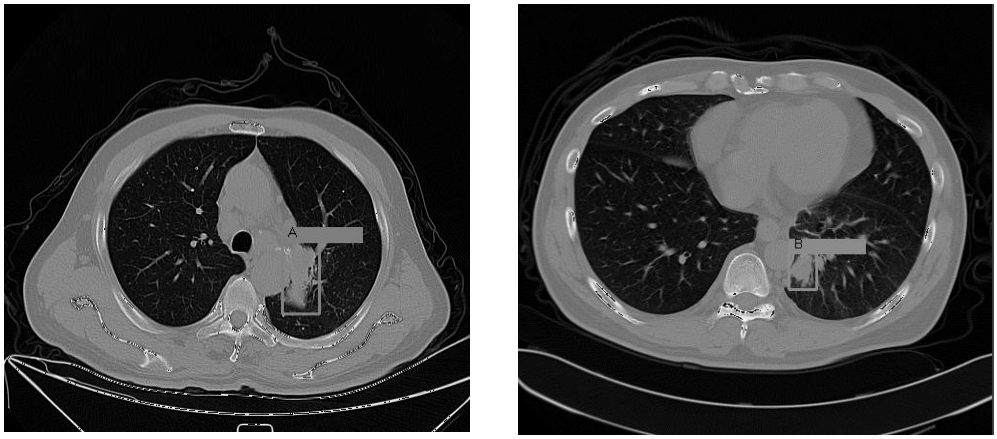


Figure 3.3 Series of DICOM images for patient Lung_Dx-A0002.Figure taken *[34]*

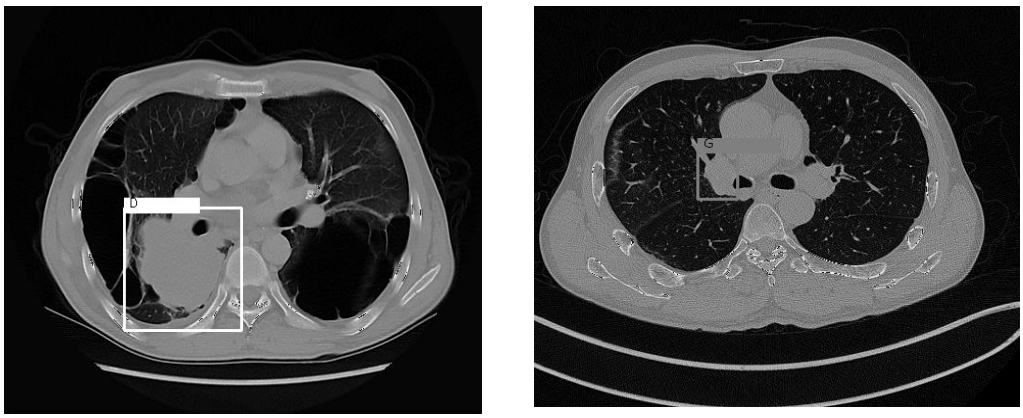
The images can be downloaded using the link below.
https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=70224216

Experienced five academic thoracic radiologists have annotated the location of tumors in each lung image to create a useful dataset for developing algorithms in medical diagnosis. To visualize the annotation box on top of the DICOM images, this python code should be used.

https://wiki.cancerimagingarchive.net/download/attachments/70224216/VisualizationTools.zip?api=v2



(a)Adenocarcinoma (b) Small Cell Carcinoma



(c) Large Cell Carcinoma (d) Squamous Cell Carcinoma

Figure 3.4 CT scan images each lung cancer type.
Figures taken from *[34]*

3.3 Image Representation

Computerized images are essentially matrices of numerical qualities, where each worth corresponds to the intensity of a pixel. These pixels act as the fundamental units of a computerized image, capturing the brightness of this present reality scene they represent. In the context of CT scan images, pixel values can be extracted using the pydicom.dcmread() function. This function yields a DicomObject as its result, which consists of two essential components.

The primary component encompasses metadata, providing significant information, for example, the patient's name, concentrate on date, and the imaging methodology utilized, among other pertinent subtleties. This metadata is essential for contextualizing and understanding the image.

The second component holds the actual image data, put away as a 2D cluster within the pixel exhibit characteristic of the DicomObject. Every element within this cluster corresponds to a particular pixel in the image, representing the intensity esteem related with that pixel's location in the image.

In essence, this cycle enables the transformation of visual information from CT scan images into a structured computerized design that can be promptly handled and analyzed. These numerical representations of pixel intensities act as the foundation for different computational techniques, including deep learning draws near, for image interpretation and diagnostic purposes.

3.4 Image Preprocessing

Few-shot learning, denoted as N-way-K-shot classification, is a fundamental concept in the field of machine learning. This approach involves working with two key datasets: support sets and query sets. The support set is made out of images ordered into N distinct classes, with each class having K example images. To illustrate, consider a dataset with four classifications, each containing two images; this scenario is alluded to as a 4-way-2-shot classification. The job of the query set is to identify the classes to which its images belong.

In request to effectively handle DICOM design images, a custom dataset class is made, inheriting properties from the torch.utils.data.Dataset. This custom dataset class is designed to flawlessly integrate with models for training and testing. It incorporates two essential strategies: __len__() and __getitem__(). The __len__() technique gives the length of the input data, while the __getitem__() strategy takes an index as input, recovers the corresponding example and target data, and can apply determined transformations prior to returning this data as a tuple. This tweaked dataset class is

instrumental in constructing a PyTorch Data Loader for subsequent model implementation.

When the __getitem__() strategy gets to a DICOM document, it extracts the pixel exhibit data from the record. This pixel cluster data is then converted into a NumPy exhibit, which works with subsequent preprocessing steps. This critical cycle ensures that the DICOM image data is in a configuration that can be effectively used for different machine learning and deep learning tasks.

3.4.1 Transform to Hounsfield Units

The conversion of pixel representations from DICOM images to Hounsfield Units (HU) constitutes a vital stage in the processing of medical imaging data. This conversion enhances the likeness and analyzability of images, accordingly, enabling more precise and dependable diagnosis and treatment planning.

The cycle begins with the recovery of the rescale intercept esteem from the DICOM header of the medical image. The rescale intercept, an intrinsic DICOM quality, fills the need of fine-tuning pixel values to oblige any inherent offsets in the measurement. This initial step ensures that the pixel values align accurately with the actual measurements, laying the foundation for meaningful radiological analyses.

Subsequently, the rescale incline esteem is extracted from the DICOM header of the medical image. Like the rescale intercept, the rescale slant is a DICOM characteristic that assumes a urgent part in calibrating pixel values. It is utilized to account for any scaling factors involved in the measurement, in this way facilitating consistent and solid data interpretation.

Following the acquisition of both the rescale incline and intercept esteems, a basic transformation process happens. These qualities are applied to the pixel cluster of the medical image, leading to the conversion of pixel values into Hounsfield Units. Hounsfield Units represent a standardized scale within the domain of medical imaging,

serving as a universal reference for characterizing the radiodensity of explicit tissues or materials.

This conversion interaction significantly enhances the clinical utility of medical images, allowing healthcare professionals to make more informed decisions regarding patient consideration. The precision and accuracy accomplished through the alignment of pixel values with actual measurements, as well as the translation into Hounsfield Units, are paramount in the diagnosis and treatment planning for different medical conditions.

To dive deeper into the intricacies of this basic conversion process and its implications for medical imaging, it is prudent to investigate pertinent IEEE references, for example, [1] and [2]. These sources give comprehensive insights into the technical aspects and clinical significance of Hounsfield Units in medical imaging.

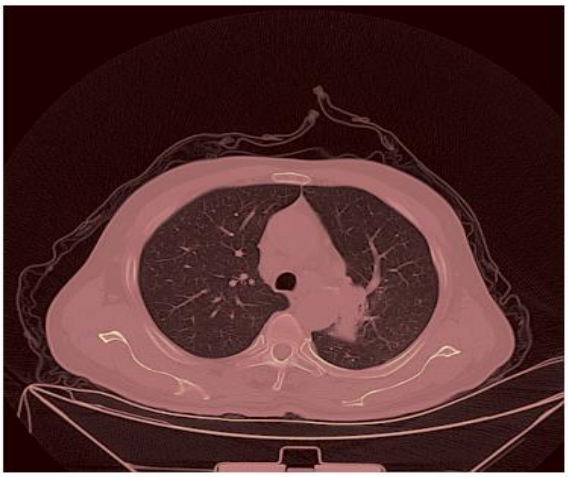


Figure 3.5 A CT scan image after applied HU transformation.
Figure taken from *[34]* for pre-processing.

3.4.2 Image enhancement

The essential objective of image enhancement is to raise the quality of an image, subjectively speaking, over that of the original image. Image enhancement endeavors to augment the lucidity of image characteristics while preserving the innate structure of the image. This pursuit works with different image analysis tasks, including yet not restricted to edge and boundary detection, by rendering the visual content more discernible and interpretable.

In the context of CT scan images, the presence of inherent noise can be especially dangerous, potentially leading to erroneous interpretations of the cancer's stage. Consequently, the evacuation of noise becomes basic to ensure the accuracy and unwavering quality of diagnostic results. In the implementation within reach, a median channel is utilized as a noise reduction strategy.

Median filtering represents a non-linear cycle explicitly custom-made for the expulsion of indiscreet noise, for example, salt and pepper noise, from computerized images. When utilizing this filtering technique, every pixel in the image is carefully scrutinized in a sequential manner, with an essential spotlight on assessing its quick neighboring pixels. This evaluation expects to discern whether the pixel accurately represents its surrounding region. Notably, the median channel varies from different channels that may simply supplant a pixel esteem with the mean normal of its adjacent qualities; instead, it replaces the pixel esteem with the median normal of the neighboring qualities.

The utilization of a median channel in noise reduction is instrumental in enhancing the diagnostic quality of CT scan images. By mitigating the unfriendly effects of noise, the channel contributes to additional precise and solid assessments of cancer stages and other clinical indicators. The rationale behind this technique is deeply attached in its capacity to preserve essential image subtleties while effectively suppressing noise artifacts, subsequently refining the images' diagnostic utility.

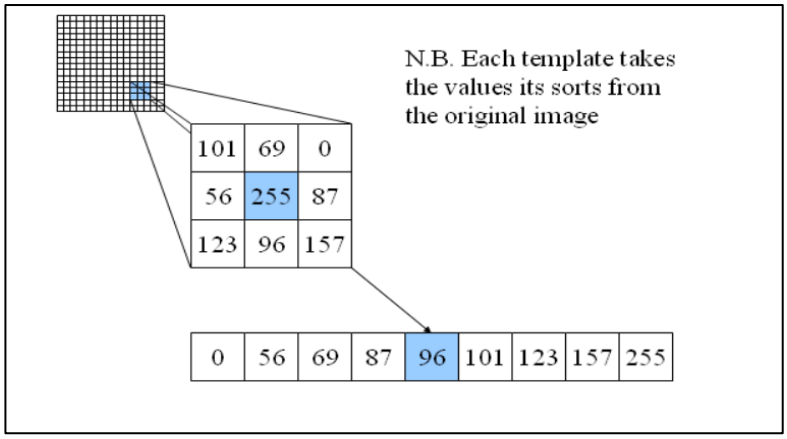
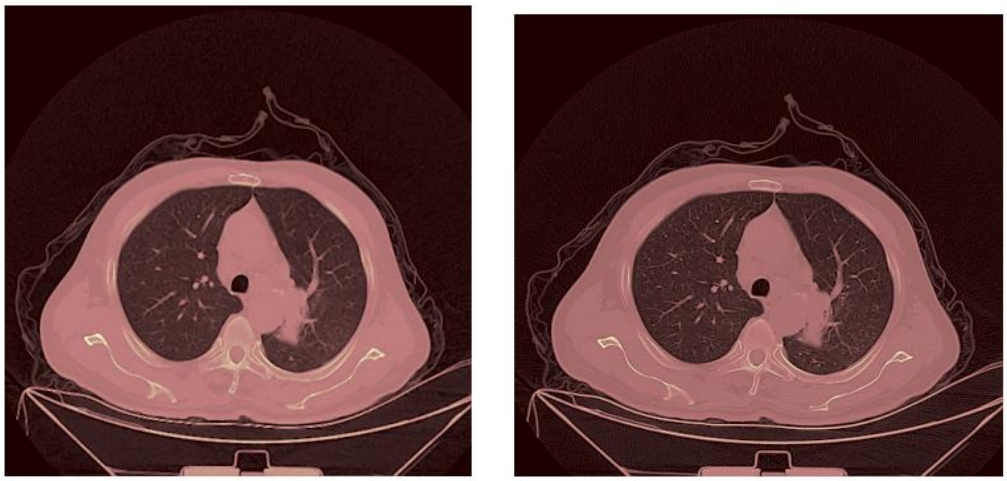


Figure 3.7 Median filtering.
Figure is taken from computer vision demonstration website in University of Southampton



(a) Before noise removal (b) After noise removal

Figure 3.6 Enhanced CT scan Images. CT scan image taken from *[34]*

The median is calculated by first arranging all the pixel values in the correct numerical sequence in the neighborhood, and then replacing the middle pixel value for the one being assessed.

3.5 Image Segmentation

In the domain of medical image analysis, the robotized segmentation of anatomical structures holds a significant job. Image segmentation is fundamentally concerned

with partitioning an image into discrete regions or classes, with every subset exhibiting internal homogeneity regarding explicit characteristics. This crucial cycle has given ascent to a plenty of segmentation techniques, particularly in the context of CT scans, where precise delineation of anatomical structures is of paramount importance.

A comprehensive report conducted by Johannes, Florian, Jeanny, and Sebastian involved a near analysis of deep learning approaches for lung segmentation. Their investigation dove into the application of four distinct semantic segmentation models, namely U-net, Res U-net, Expanded Leftover Network, and Deeplab v3+, within the domain of medical image analysis. The uniqueness of their methodology lay in training these models from the ground up on four different datasets, comprising a combination of freely accessible training sets and data collected from clinical routines.

To assess the viability of these models, their assessments encompassed both public test datasets and genuine clinical cases, which included instances of extreme pathologies. Additionally, the review conducted a similar analysis against two previously distributed programmed lung segmentation frameworks, examining their performance across a spectrum of datasets, including the really challenging clinical scenarios. The overarching objective was to measure the flexibility and dependability of these models in the context of medical image segmentation, especially within the domain of lung analysis.

The performance evaluation of these models hinged on the utilization of a few key metrics. Dice coefficient (DSC), Vigorous Hausdorff distance (HD95), and Mean Surface Distance (MSD) featured prominently in their evaluation toolkit. The Dice similarity coefficient fills in as a proportion of cross-over between two alternative labelings, for example, predicted and ground truth lung masks. Meanwhile, the Vigorous Hausdorff distance (HD95) quantifies the directed Hausdorff distance between the surfaces Xs and Ys. This distance metric gives urgent insights into the spatial differences between the predicted and ground truth lung masks. Ultimately, the Mean Surface Distance (MSD) strategy ascertains the typical distance between all points in surface Xs and their nearest corresponding points in surface Ys. This

approach offers a comprehensive evaluation of the general typical accuracy of the predicted lung masks contrasted with the ground truth masks.

The results of this thorough evaluation displayed that the U-net model showed the most promising outcomes, especially excelling as far as the Dice similarity score. Notably, it accomplished the lowest scores in both Strong Hausdorff distance and mean surface distance, highlighting its ability as a top entertainer in the context of lung segmentation [1].

3.5.1 U-Net Architecture

U-Net Architecture is a CNN architecture that was proposed for biomedical image segmentation tasks in 2015 by Olaf Ronneberger, Philipp Fischer, and Thomas Brox. The architecture is named after its U-shaped design, which consists of a contracting path (down sampling) followed by an expansive path (up sampling) [36].

The architecture is made up of three primary components: the bottleneck layer, the expanding path, and the contracting path. The input image's spatial resolution is steadily decreased as the number of feature mappings is increased by the several convolutional and pooling layers that make up the contracting route. A single convolutional layer that joins the expanding and contracting channels creates the bottleneck layer. The expanding path, which is the mirror image of the contracting path, has numerous deconvolutional layers that gradually boost the feature maps' spatial resolution while lowering their channel count. To maintain spatial information, matching layers in the contracting and expanding routes are connected via skip connections.

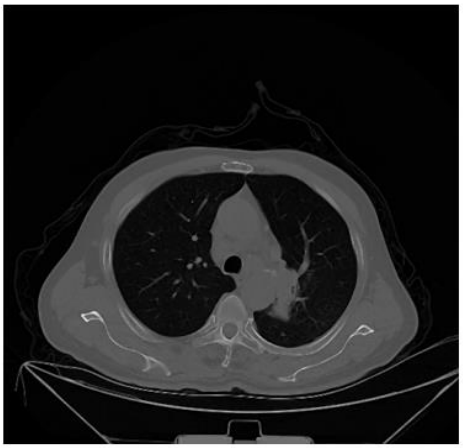
U-net is a modern architecture which has outstanding performance in segmenting returns for money invested [17].Johannes and his group [35] introduced a changed version of the U-net architecture, which they trained using the R-231 Dataset. The R-231 dataset consists of scans obtained from 22 different combinations of scanner manufacturers, convolution kernels, and cut thicknesses. This dataset showed many variations in the appearance of lungs, yet it lacks cases that show the head or abdominal

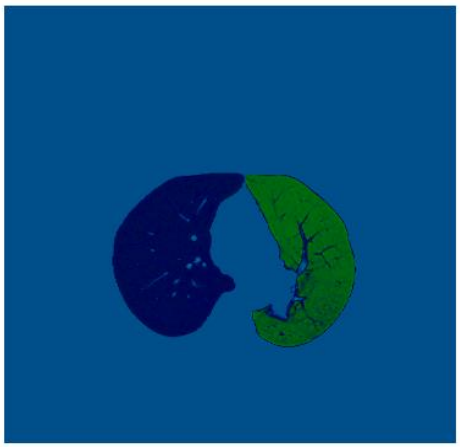
region. In their research paper they have featured that at the hour of submission, the U-net(R-231) model attained the second-most elevated score among all participants in the LOLA11 challenge. In comparison, the U-net(R-231) model demonstrated better evaluation measures, including Dice similarity coefficient (DSC), Hearty Hausdorff distance (HD95), Mean Surface Distance (MSD), and cancer cross-over, when contrasted with two other freely accessible algorithms. This indicates that the U-net(R-231) model performed remarkably well in the challenge and showed promising outcomes when contrasted with existing techniques. Table 3.1 shows the outcomes for growth cross-over and it very well may be seen that U-net (231) covered more cancer volume.

	Tumor Overlap			
Method	Mean (%)	Median (%)	<5%	>95%
CIP	34	13	113	56
P-HNN	50	44	48	78
U-net(R-36)	53	54	46	79
U-net(R-231)	60	69	37	90

Table 3.1 Overlap between lung masks and manually annotated tumor volume *[35]*.

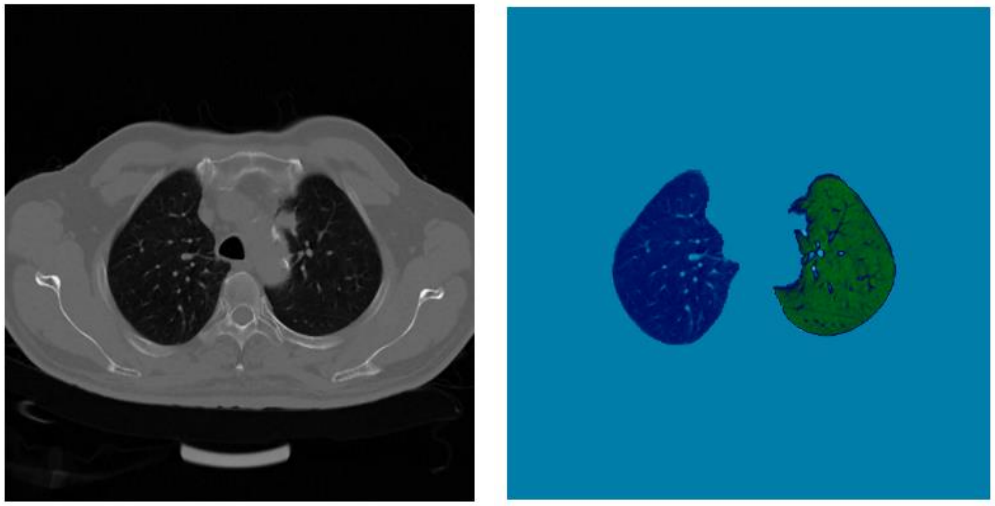
U-net (R321) architecture can be applied using the lungmask module in python. Below figures show the results after applying the lungmask module for CT images.





(a) Before segmentation (b) After segmentation

Figure 3.8 U-Net (R321) applied for CT scan image with Adenocarcinoma. CT scan images taken from *[34]*



(a) Before segmentation (b) After segmentation

Figure 3.9 U-Net (R321) applied for CT scan image with Squamous Cell Carcinoma. CT scan images taken from *[34]*

3.6 Feature Extraction

In this period of the image analysis process, the extraction of meaningful features from input image data expects a basic job. An effective methodology frequently utilized for feature extraction in the domain of medical imaging involves the utilization of pre-trained convolutional neural networks (CNNs). In contrast, traditional feature extraction techniques depend on handcrafted features, necessitating the manual design and extraction of domain-explicit qualities from CT images. While handcrafted features offer interpretability and insights into explicit characteristics of CT images, they have inherent limitations. These limitations basically originate from the requirement for domain knowledge and their constraints in capturing intricate patterns and progressive representations that pre-trained CNNs can autonomously learn [1].

The introduction of pre-trained CNNs has been transformative in the domain of medical image analysis, offering a strong mechanism for transfer learning. Transfer learning allows the knowledge procured from training on one dataset to be applied to a distinct yet related task, like medical image analysis. One of the key advantages of employing CNNs lies in the fact that the lower layers of these networks have practical experience in learning fundamental features like edges, surfaces, and patterns — ascribes that hold significant relevance in the context of medical images [2]. The progressive architecture of CNNs works with the acquisition of logically perplexing and more elevated level features, as it combines low-level features to frame more intricate representations. This various leveled feature progressive system demonstrates especially important for capturing unobtrusive patterns and structures in CT images.

Besides, the utilization of pre-trained CNNs conveys the benefit of these models having been trained on different sets of images, promoting generalization and mitigating the risks of overfitting, particularly when confronted with restricted medical data. In the particular implementation under consideration, four distinct pre-trained CNNs, namely VGG16, ResNet50, DenseNet, and CNN, were harnessed for the feature extraction process in the analysis of lung CT images. This strategic selection allowed for the exploration of different architectural capacities and feature representation characteristics inherent in these networks [3].

In conclusion, the integration of pre-trained CNNs in the feature extraction period of medical image analysis stands as a significant advancement, enabling the programmed acquisition of intricate image features and leveraging the progressive representations learned during training. This approach is invaluable, especially in scenarios involving restricted medical data, where the risk of overfitting is prevalent. The selection of explicit pre-trained CNN models, like VGG16, ResNet50, DenseNet, and CNN, underscores the flexibility and versatility presented by these architectures, making them an essential component in the analysis of lung CT images [3]. To dive deeper into the philosophies and findings related with this methodology, the interested peruser is encouraged to investigate the comprehensive concentrate by [1].

3.6.1 VGG16

Among the pre-trained models utilized for feature extraction in the context of lung CT images, VGG16 stands firm on a significant situation. VGG16, a convolutional neural network architecture, was originally designed for image classification tasks, and it made its presentation in the influential paper named "Extremely Deep Convolutional Networks for Huge Scope Image Recognition," created by Simonyan and Zisserman in 2014 [1]. Created by the Visual Calculation Gathering (VGG) at the University of Oxford, VGG16 flaunts a composition characterized by sixteen convolutional layers, further complemented by three completely connected layers.

The hallmark of VGG16's architectural design is the consistent channel size of 3x3, a step of 1, and a padding of 1, strategically chosen to ensure the preservation of the input image's spatial resolution all through the network. As one advances deeper into the neural network, there is an ever-evolving increase in the number of channels within each layer. This ascent commences with 64 channels in the initial layer and culminates at 512 channels in the final layer, reflecting the network's capacity to catch increasingly complicated and abstract features.

Another essential element in the VGG16 architecture is the application of the rectified linear unit (ReLU) activation function following each convolutional layer. The inclusion of ReLU introduces non-linearity into the model, allowing it to learn intricate and nonlinear relationships within the data. Additionally, max-pooling operations are utilized after each two progressive convolutional layers. These operations assume an imperative part in reducing the spatial resolution of the feature maps while retaining essential information, contributing to the network's capacity to extract meaningful image features.

In the final phases of VGG16, the result from the convolutional layers is flattened to make a one-dimensional feature vector, which is then handled through three completely connected layers. The final completely connected layer is responsible for producing the class scores related with the input image, effectively enabling image

classification. This architecture's strong design and remarkable profundity make it a reasonable decision for feature extraction in the analysis of lung CT images, given its proven track record in image recognition tasks [1].

In rundown, VGG16, renowned for its viability in image classification, stands as an important asset in the extraction of relevant features from lung CT images. Its architectural characteristics, for example, consistent channel sizes, ReLU activation functions, and max-pooling layers, collectively contribute to its capacity to catch intricate image ascribes. Besides, the progressive increment in the number of channels enhances its ability to discern complex patterns. To dive further into the inside and out particulars and findings connected with VGG16, per users are encouraged to investigate the seminal work by Simonyan and Zisserman [1].

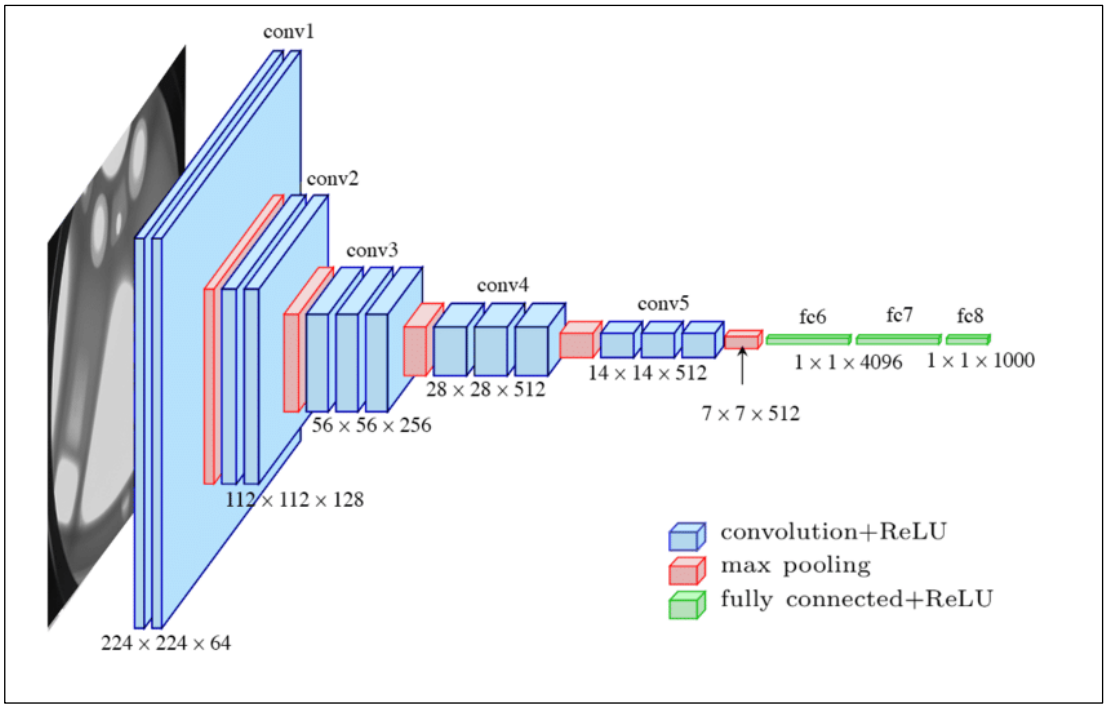


Figure 3.10 The standard -16 architecture

Adapted from "Simonyan and Zisserman (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition VGG 1-14".

For this implementation, the model was changed to acknowledge grayscale images as input as opposed to the RGB images that the original VGG16 architecture was designed for. The pretrained boundary is utilized to stack the loads of the pre-trained VGG16 model, which are utilized as initial qualities for the model's parameters. The last completely connected layer of the VGG16 model is taken out by creating a new Sequential module containing every one of the layers of the VGG16 features module with the exception of the last layer. Be that as it may, the model will keep the result channel size as 64 and the kernel size, step and padding upsides of the convolution layer remain equivalent to in the original VGG16 model.

3.6.2 ResNet50

ResNet50 is a convolutional neural network architecture that was introduced in the paper "Deep Residual Learning for Image Recognition" by He et al. in 2016 [40]. It is a variant of the ResNet family of models, which are known for their ability to train very deep neural networks with thousands of layers. The ResNet50 architecture comprises a total of 50 layers, encompassing convolutional layers, pooling layers, fully connected layers, and shortcut connections that enable the bypassing of one or more layers. By incorporating these shortcut connections, the network effectively addresses the issue of vanishing gradients that can arise during the training of deep neural networks.

There are four phases in the ResNet50 design, and each one includes several residual blocks. Two or three convolutional layers with batch normalization and ReLU activation are included in each residual block, along with shortcut connections that combine the input and output of the block. The first stage has 64 filters, the second stage has 128 filters, the third stage has 256 filters, and the fourth stage has 512 filters. The output of the last residual block is fed into a global average pooling layer, which averages the feature maps over the spatial dimensions of the input. The resulting feature vector is then passed through a fully connected layer with SoftMax activation to produce the final class probabilities. This can be used as a feature extractor for a

wide range of computer vision tasks. When used for feature extraction, the pre-trained ResNet50 model is used as a fixed feature extractor, and the output of one of its intermediate layers is used as the input to a classifier or another downstream task-specific model. To apply ResNet50 as a feature extractor for this experiment, the model is modified to accept single channel grayscale images as input.

3.6.3 Convolution Neural Network

The CNN architecture is widely used for feature extraction in image-based machine learning applications, including few-shot learning. The CNN model consists of several convolutional layers followed by pooling layers and fully connected layers. In the context of few-shot learning, the CNN model is trained on a large dataset to learn the high-level feature representations of the images. Then, the trained CNN model is used to extract the features from the support and query images in the few-shot learning scenario.

In this methodology, the CNN model has used three convolutional layers, each followed by a ReLU activation function and a pooling layer. The convolutional layers perform the operation of convolving a set of learned filters with the input image to extract features. The pooling layers perform a down sampling operation by selecting the maximum value within a given region of the feature map. The fully connected layers at the end of the network take the flattened feature maps and perform classification using the extracted features.

The CNN model can be fine-tuned on a limited number of labeled support pictures for each new classification assignment to adapt it for few-shot learning. The CNN model may learn to extract characteristics that are particular to the new classification job by fine-tuning on the support set. The distance between the features of the query picture and the class prototypes generated from the support images may then be utilized to categorize the query images using the extracted features.

3.6.4 DenseNet

DenseNet is a deep learning architecture for image classification and object detection tasks. DenseNet is unique in that it incorporates the outputs of all preceding layers as input to each subsequent layer [41]. This dense connection structure enables feature reuse and lowers the number of parameters needed to learn many features by giving each layer direct access to the feature maps of all previous layers.

In terms of feature extraction, DenseNet has been shown to be remarkably effective in learning features from images. By utilizing dense connectivity, the network can learn a diverse range of features across different scales and levels of abstraction. This is particularly important for object detection tasks, where objects can appear at different scales and orientations. DenseNet can be used as a feature extractor in the prototypical network architecture. The pre-trained DenseNet model can be used to extract features from the images in the support and query sets. The output of the DenseNet model is typically a high-dimensional feature vector that represents the image.

Layers	Output Size	DenseNet-121($k = 32$)	DenseNet-169($k = 32$)	DenseNet-201($k = 32$)	DenseNet-161($k = 48$)
Convolution	112 × 112	7 × 7 conv, stride 2			
Pooling	56 × 56	3 × 3 max pool, stride 2			
Dense Block (1)	56 × 56	[1 × 1 conv, 3 × 3 conv] × 6	[1 × 1 conv, 3 × 3 conv] × 6	[1 × 1 conv, 3 × 3 conv] × 6	[1 × 1 conv, 3 × 3 conv] × 6
Transition Layer (1)	56 × 56	1 × 1 conv			
	28 × 28	2 × 2 average pool, stride 2			
Dense Block (2)	28 × 28	[1 × 1 conv, 3 × 3 conv] × 12	[1 × 1 conv, 3 × 3 conv] × 12	[1 × 1 conv, 3 × 3 conv] × 12	[1 × 1 conv, 3 × 3 conv] × 12
Transition Layer (2)	28 × 28	1 × 1 conv			
	14 × 14	2 × 2 average pool, stride 2			
Dense Block (3)	14 × 14	[1 × 1 conv, 3 × 3 conv] × 24	[1 × 1 conv, 3 × 3 conv] × 32	[1 × 1 conv, 3 × 3 conv] × 48	[1 × 1 conv, 3 × 3 conv] × 36
Transition Layer (3)	14 × 14	1 × 1 conv			
	7 × 7	2 × 2 average pool, stride 2			
Dense Block (4)	7 × 7	[1 × 1 conv, 3 × 3 conv] × 16	[1 × 1 conv, 3 × 3 conv] × 32	[1 × 1 conv, 3 × 3 conv] × 32	[1 × 1 conv, 3 × 3 conv] × 24
Classification Layer	1 × 1	7 × 7 global average pool			
		1000D fully-connected, softmax			

Figure 3.11 DenseNet architectures for ImageNet. Figure taken from [32].

3.7 Prototype Representation

In a few shot classification, a small support set of N labeled examples will be given. The D-dimensional feature vector of each example will be output by a pre-trained feature extraction model. To create a prototype for each class, the average of all feature vectors in a class label will be calculated. This prototype serves as a representative for the label and is used to compare with the features of new examples during inference.

If we consider a small support set of *N* labeled examples S= $\{(X_1,y_1)\ldots\ldots(X_N,y_N)\}$ where each $X_i \in \mathbb{R}^{\mathcal{D}}$ is the D-dimensional feature vector of an example and $y_i \in \{1, \ldots \ldots, K\}$ is the corresponding label. S_k denotes the set of examples labeled with class k.

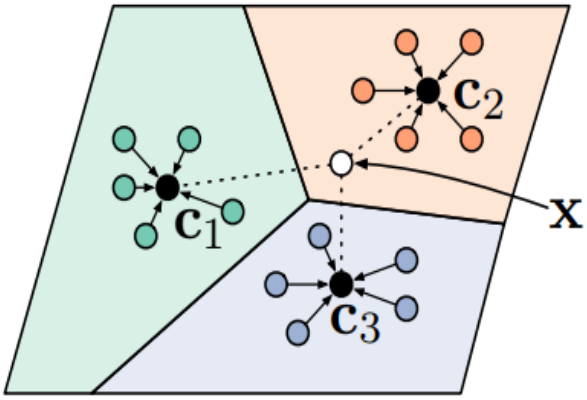


Figure 3.12 Few shot prototypes as the mean of embedded support features. Figure taken from *[42]*

3.8 Calculate Distance Metric

One key component in prototypical networks is calculation of distance matrix. The distance matrix is used to measure the similarity between the support set and the query set. It is shown that Euclidean distance performs better than cosine similarity when calculating distance [42].Below shows the formula for Euclidean distance.

$$d(z, z') = \| z - z' \|^2$$

When Euclidean distance is applied, the model is equivalent to a linear model with a particular parameterization.

3.9 Meta Learning

Meta-learning is a popular machine learning technique often applied to Few-Shot Learning (FSL). The objective of meta-learning algorithms is to rapidly acquire knowledge about a novel task using a limited amount of new data. Meta-learning, also known as "learning to learn". It selects training samples from the base classes for few-shot classification tasks and optimizes the model to perform well on these tasks. An N-way and K-shot task typically consists of N classes, each of which has K support samples and Q query samples.

The objective is to group these N x Q query samples into N classes using the N x K support samples. In this approach, a few shot classification tasks are directly tuned for the model. The primary benefit of meta-learning is thought to be the alignment of training and testing objectives. Usually, Dataset D which takes for standard supervised machine learning, split into two separable and unique datasets, which are called train and test sets denoted by D_{train} and D_{test} respectively. One iterative cycle, during the conventional neural network, is called “Epoch.” During each epoch, the whole D_{train} is feeding forward and backward and it will be partitioned into several parts which are called “Batch ”. Consequently, the number of iterations per epoch are all batches that feed into the model.

In the meta learning process, the dataset D is split into two meta sets without any overlapping, and it can be named as D_{meta_train} and D_{meta_test}. D_{meta_train} and D_{meta_test} and then further separated into train and test sets. To distinguish these two types of training and testing sets, the inner train and test sets are referred to as "Support set" and "Query set," respectively. The support set is in the form of N-way k-shot random samples, and the query set consists of q random samples for each of the N support set classes and these two data sets would be created during each episode. Support set in the D_{meta_train} will be taken by model to learn and then applied on the query set in the $D_{meta\_train.}$ Then

the prediction will be made. The prediction error over episodes is used to update the meta-learner. The meta-learner learns to learn from the limited dataset throughout a series of episodes. This stage is known as meta-learning. After that, the support set in the D_{meta_test} is used to create a classifier. In this case, the model focuses on fast learning task-specific parameters. Subsequently, the performance of the classifier is assessed using a query set in the D_{meta_test} to enter the adaptation phase. In other words, the model utilizes the prior knowledge acquired during the meta-learning phase to rapidly adjust to a new task during the adaptation phase.

For this development, the CrossEntropyLoss () function is used to define the loss criterion for the classification task. This function computes the cross-entropy loss between the predicted class scores and the actual labels for each training example. The optim. Adam() function is used to define the optimizer for updating the model's parameters during training. Adam is a popular optimization algorithm that computes adaptive learning rates for each parameter, allowing the model to converge faster and more reliably.

The learning rate was set to 0.001, which determines the step size taken by the optimizer in the parameter space. Low learning rates can lead the model to converge slowly, while excessive learning rates can cause the model to diverge and exceed the ideal parameters. The optimal learning rate depends on the specific problem and dataset being used.

3.10 Tools & Technologies

Python was chosen as the implementation language due to its reputation as a high-level, interpreted programming language that excels in simplicity, readability, and user-friendliness, particularly in the domains of data science and web development. The following list enumerates the libraries and frameworks employed in the project:

- PyTorch
- Numpy

- PyDicom
- OpenCV
- Tensorflow
- Matplot
- Lungmask

As the IDE, Jupyter Notebook Extension in Visual Studio Code is used. The Jupyter Notebook extension allows users to run Jupyter notebooks within VS Code, providing a more integrated and streamlined development experience. Jupyter notebooks are interactive documents that combine code, text, and visualizations. With the Jupyter Notebook extension, users can create, edit, and run Jupyter notebooks directly within VS Code. The extension provides syntax highlighting, code completion, and other features that make it easier to work with Jupyter notebooks. It also includes a built-in Jupyter server that can run notebooks in the background and provide access to interactive widgets.

4 EVALUATION

The experiment involved the utilization of four different pre-trained CNN models. The following four figures depict the feature spaces of the support dataset using these four models. Each color represents a distinct lung cancer type.

For further analysis of feature extraction, inter-class and intra-class variations in the feature space were calculated. A higher ratio indicates that the classes are well-separated and easily distinguishable in the feature space, whereas a lower ratio indicates that the classes are more difficult to distinguish. This measure can be useful in evaluating the effectiveness of the prototypical network in learning a high-quality feature space that can accurately distinguish between different classes.

Feature Extractor Model	Ratio of inter-class to intra-class distances
CNN	16
VGG16	295358
DenseNet	215255
ResNet50	4659

Table 4.1 Ratio of inter-class to intra class based on the feature extraction model.

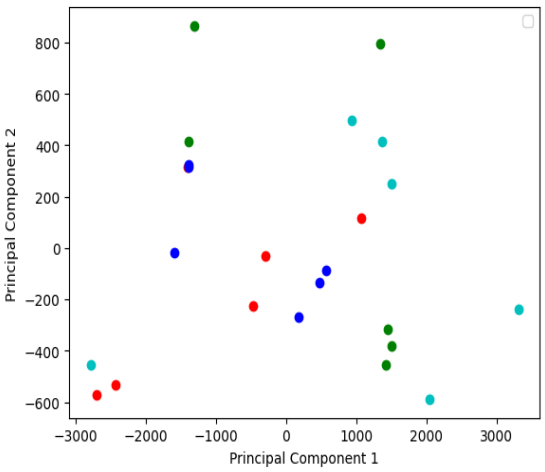


Figure 4.2 Extracted feature space using ResNet50.

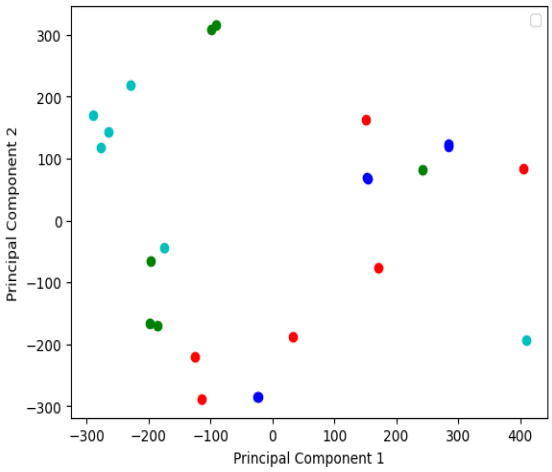


Figure 4.1 Extracted feature space using DenseNet.

.

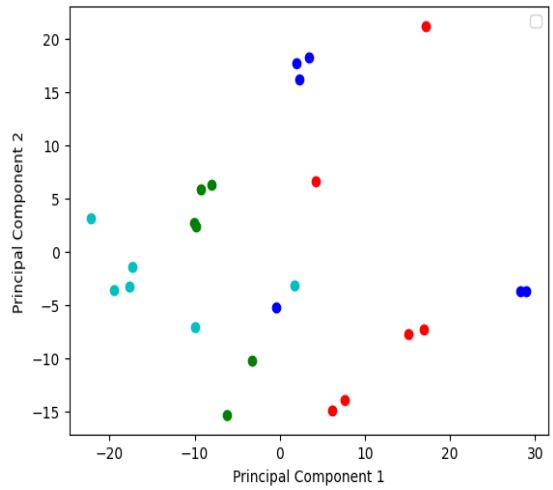


Figure 4.4 Extracted feature space using CNN model.

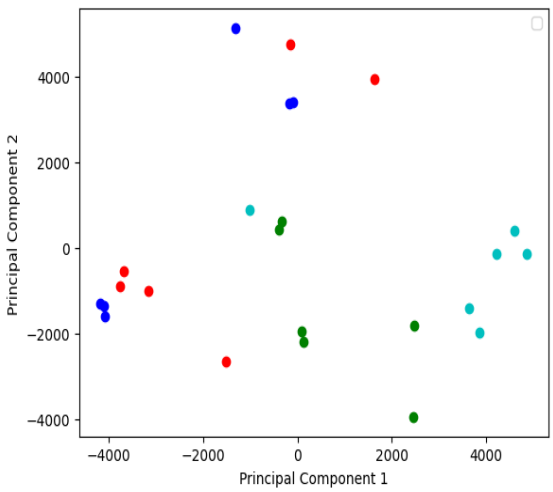


Figure 4.3 Extracted feature space using VGG16

The performance evaluation of each pre-trained model is conducted by assessing metrics such as Precision, Recall, F-Score, and Accuracy score. This evaluation is conducted under varying conditions where the number of shots per class in the support set is altered. The purpose is to gauge how well the models perform when faced with different scenarios of data availability for each class.

Precision measures the accuracy of positive predictions made by the model, while recall assesses the model's ability to correctly identify positive instances. The F-Score is the harmonic mean of precision and recall, providing a balanced evaluation of the model's performance in both positive and negative instances. The accuracy score measures the overall correctness of the model's predictions.

(T_p) = Number of times the model predicts the positive class correctly.

(T_n) = Number of times the model predicts the negative class correctly.

(F_p) = Number of times the model predicts the positive class incorrectly.

(F_n) = Number of times the model predicts the negative class incorrectly.

Precision, recall, accuracy and F1 score can be expressed as follows.

$$Precision = \frac{Tp}{Tp + Fp}$$

$$Recall = \frac{Tp}{Tp + Fn}$$

$$Accuracy = \frac{Tp + Tn}{Tp + Tn + Fp + Fn}$$

$$F - Score = \frac{Precision.Recall}{Precision + Recall}$$

Number of shots per class	Precision	Recall	F1-Score	Accuracy
1 -Shot	0.74	0.68	0.68	68.75%
2 - Shot	0.82	0.81	0.81	81.25%
3 - Shot	**0.86**	**0.81**	**0.80**	**81.25%**
4 - Shot	0.77	0.75	0.74	75%
5 - Shot	0.69	0.66	0.65	66.67%

Table 4.2 Evaluation metrics of VGG16 pre-trained model

Table 4.2 presents the evaluation metrics (precision, recall, F1-score, and accuracy) for a VGG16 pre-trained model in a prototypical network with varying numbers of shots per class (ranging from 1 to 5). The results demonstrate how the model's performance changes with the number of labeled examples per class. As the number of shots increases, precision, recall, and F1-score improve, indicating enhanced classification performance. The highest accuracy scores are attained with 2-shot and 3-shot settings, both at 81.25%.

Number of shots per class	Precision	Recall	F1-Score	Accuracy
1 -Shot	0.6	0.56	0.54	0.56
2 - Shot	0.88	0.87	0.87	0.87
3 - Shot	0.68	0.68	0.68	68.75%
4 - Shot	**0.95**	**0.93**	**0.93**	**93.75%**
5 - Shot	0.87	0.83	0.82	83.33%

Table 4.3 Evaluation metrics of CNN pre-trained model

Table 4.3 presents the evaluation metrics (precision, recall, F1-score, and accuracy) for a CNN model used as the feature extractor in a few-shot learning scenario with varying numbers of shots per class (ranging from 1 to 5). The highest accuracy, 93.75%, is achieved with a 4-shot setting, while the lowest, 56%, is observed with a 1-shot setting.

Table 4.4 displays the evaluation metrics for the DenseNet pre-trained model. Overall, the model exhibited satisfactory performance, with better results attained as the number of shot classes increased. The highest performance was observed in the 2-Shot class, with a precision of 0.95, recall of 0.93, F1-Score of 0.93, and an accuracy of 93%.

Number of shots per class	Precision	Recall	F1-Score	Accuracy
1 -Shot	0.79	0.68	0.66	68%
2 - Shot	0.95	0.93	0.93	93%
3 - Shot	0.88	0.87	0.87	87.5%
4 - Shot	0.87	0.83	0.82	83.3%
5 - Shot	**0.90**	**0.9**	**0.9**	**93%**

Table 4.4 Evaluation metrics for pre-trained DenseNet model

Number of shots per class	Precision	Recall	F1-Score	Accuracy
1 -Shot	0.33	0.43	0.33	43.75%
2 - Shot	0.22	0.31	0.25	31.25%
3 - Shot	0.76	0.68	0.69	68%
4 - Shot	0.71	0.62	0.60	62%
5 - Shot	**0.98**	**0.98**	**0.98**	**98%**

Table 4.5 Evaluation metrics for pre-trained model ResNet50

Below graphs show the training accuracy and training loss were tracked over multiple episodes for each model. The chosen configuration used four shots (4 DICOM images) per class and the learning rate of 0.001.

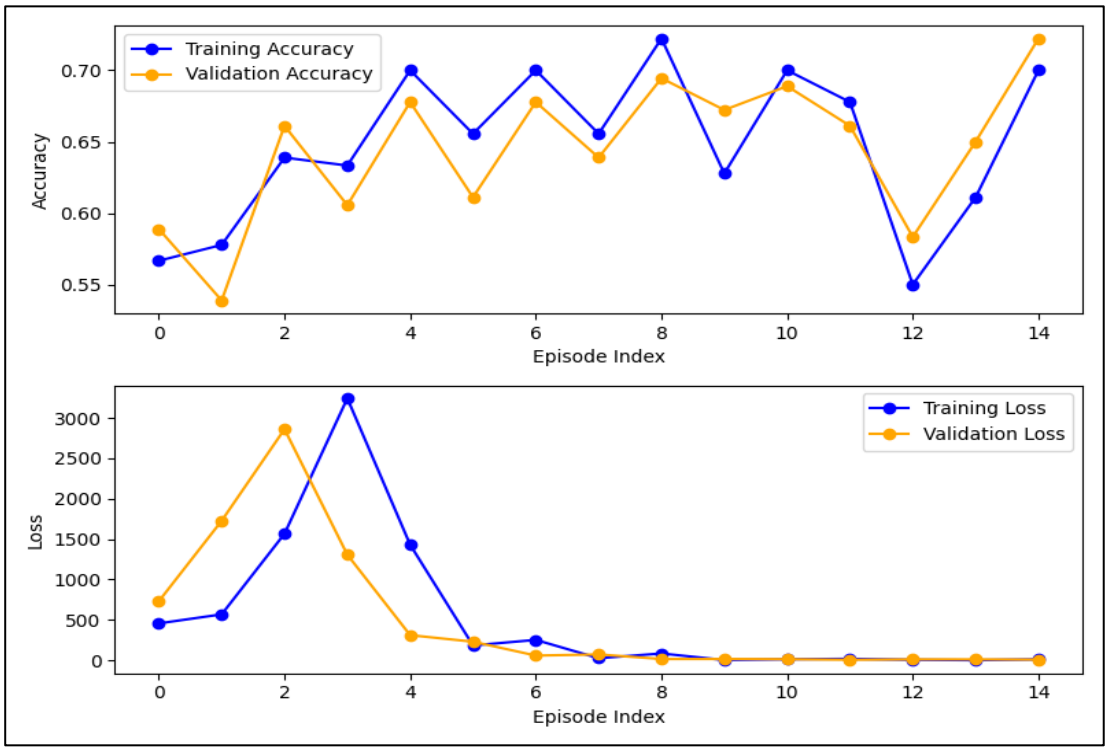


Figure 4.5 Episodic Learning with Pre-trained VGG16

Based on the evaluation metrics in the above four tables, the optimized number of shots for each model can be determined by selecting the class that achieves the highest F1-Score or accuracy. Overall, the models achieve their best performance when there are more shots per class, especially with 4 or 5-Shot classes, indicating the importance of having sufficient training examples to achieve higher accuracy and F1-Scores.

Figure 4.5 depicts a two-line graph illustrating the performance metrics of an episodically trained VGG16 model utilized for feature extraction. Across the range of training episodes, the line representing "Training Accuracy" exhibits fluctuations that form a distinctive trajectory. Notably, the model's accuracy shows an increase after the 12^{th} episode, indicating its ability to quickly learn from the data. As the training

progresses, the accuracy stabilizes around a range of values, displaying a mix of incremental improvements and minor setbacks.

Figure 4.6 presents an overview of the performance metrics for the ResNet50 pre - trained model across a series of training episodes. Examining specific episodes, the highest accuracy is attained in episode 14, with an accuracy score of 80%. On the other hand, episode eleven presents the lowest accuracy at 45%. In the depicted DenseNet model shown in Figure 4.7, the accuracy of the model begins at 70% and improves progressively across episodes, displaying varying degrees of enhancement.

Notably, the most substantial advancement occurs in episode 5, where the model attains its peak accuracy of 90%. This achievement underlines a remarkable leap in the model's overall performance. Furthermore, the loss experienced a significant reduction in subsequent episodes, indicative of the model gradually fitting the training data more effectively. Specifically, episodes 3 and 4 exhibit notably low loss values, signifying a strong alignment between the model's predictions and the actual dataset.

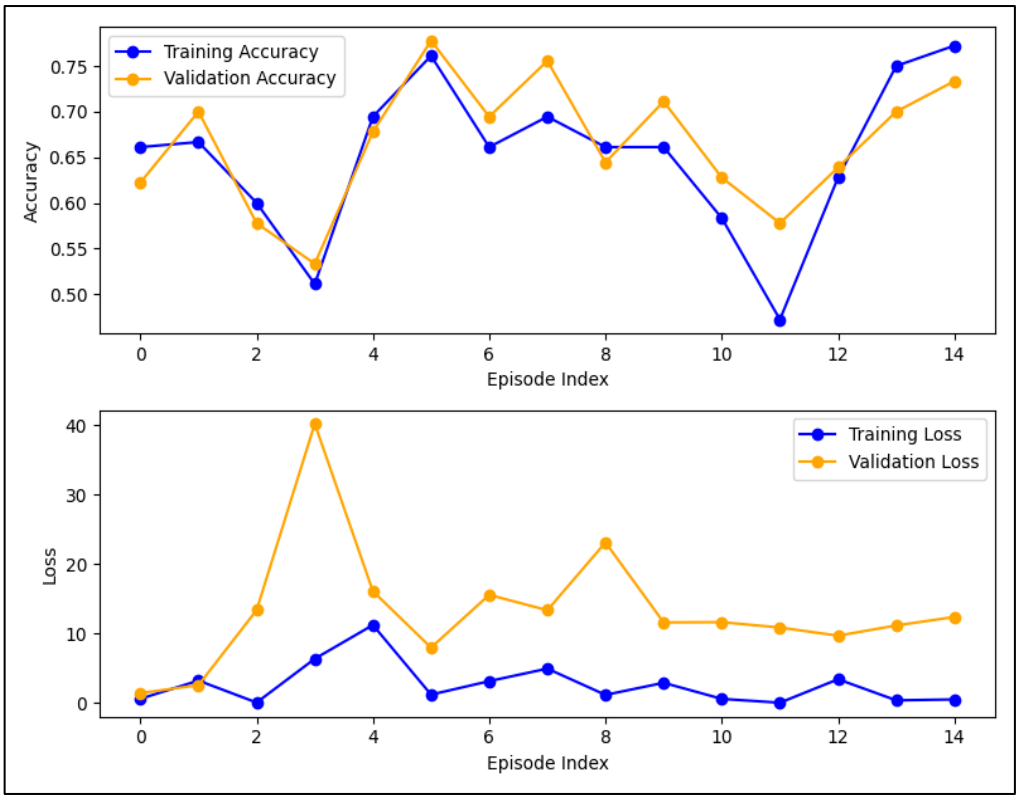


Figure 4.6 Episodic Learning with ResNet50

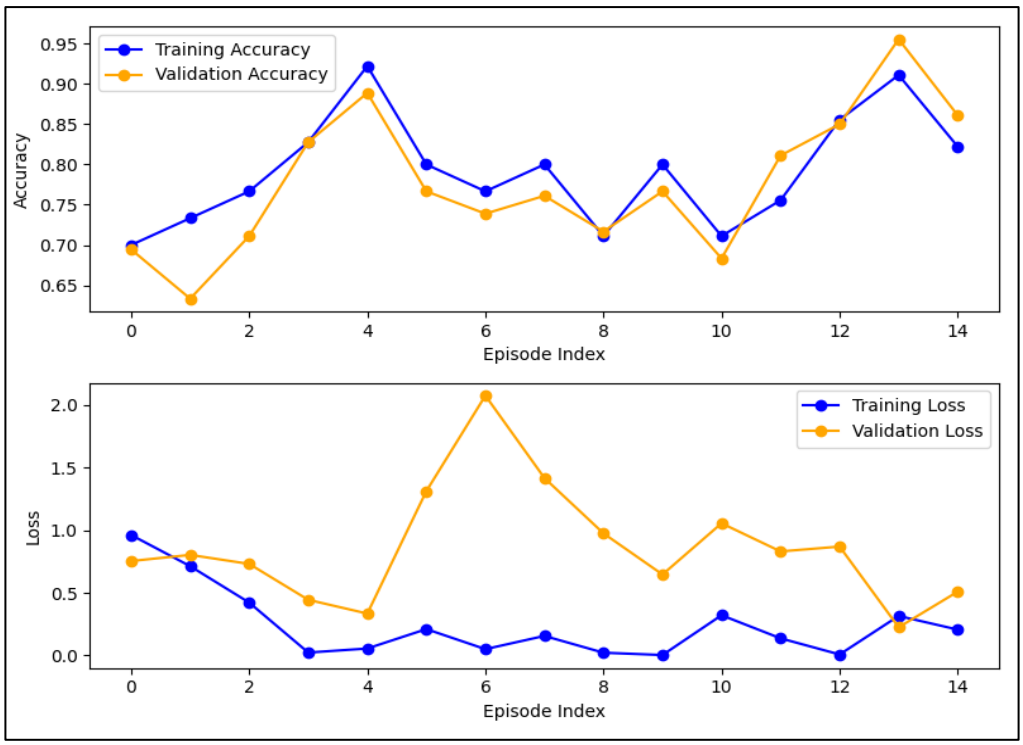


Figure 4.7 Episodic Learning with DenseNet

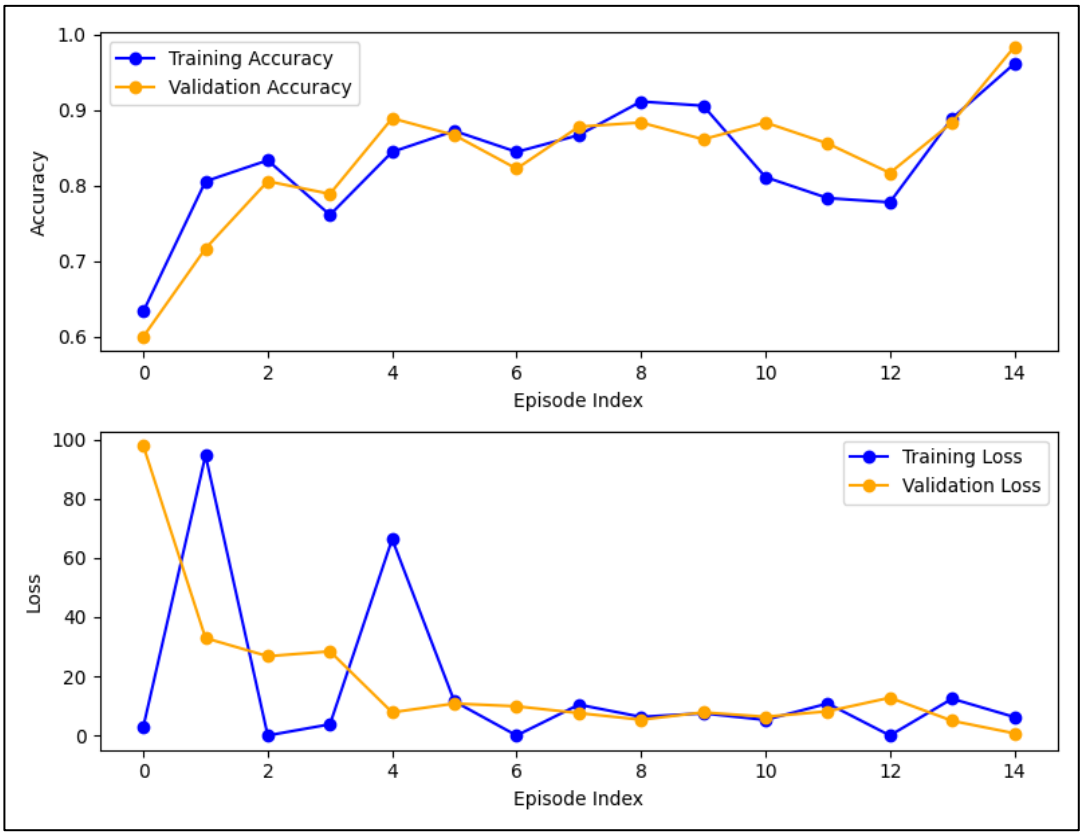


Figure 4.8 Episodic Learning with CNN

Figure 4.8 illustrates the metrics obtained when CNN is employed for feature extraction. The model's accuracy commences at around 60% and demonstrates variations as the training progresses. Across the episodes, the accuracy encounters fluctuations, occasionally exhibiting sharp increases. Notably, a remarkable peak in performance is observed in episode 14, where the accuracy reaches its highest point of 97%. The initial loss value stands at 1, suggesting that the model initially deviates from optimal parameter settings. Subsequent episodes reveal diverse loss values, reflecting fluctuations in the model's convergence toward the training data. Episodes 8, 9, 10, and 11 exhibit a loss of around 10, indicating a robust alignment between the model's predictions and the actual data. Additionally, episode 6 presents minimal loss values, indicating a strong correspondence with the training data.

Siamese neural network is another few shot learning method which belongs to meta learning approach. This method is designed to learn similarity or distance metrics between pairs of inputs. They consist of two identical subnetworks with the same weights and parameters. For this observation, a convolutional neural network with three convolutional layers followed by fully connected layers is used.

After passing the two preprocessed DICOM images for the network, the loss is calculated. Contrastive loss is used as the loss function for learning similarity or dissimilarity between pairs of samples. For this implementation, a training loop with 5 episodes was used. Through the learning, a pair of images will be passed to the neural network and two output vectors from the network will be fed in to the define loss function. In the loss function, the loss of the output images will be calculated and then loss will pass to back propagation. Below graph shows the loss over time.

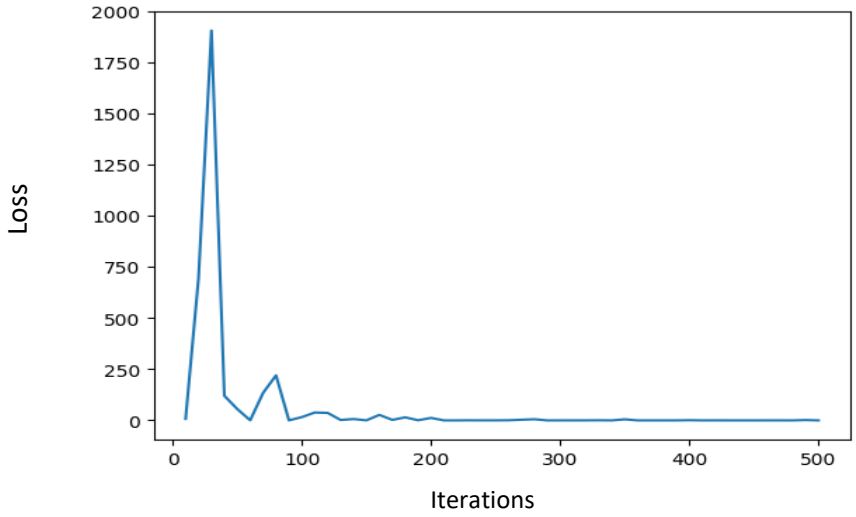


Figure 4.10 Contrastive loss over iterations

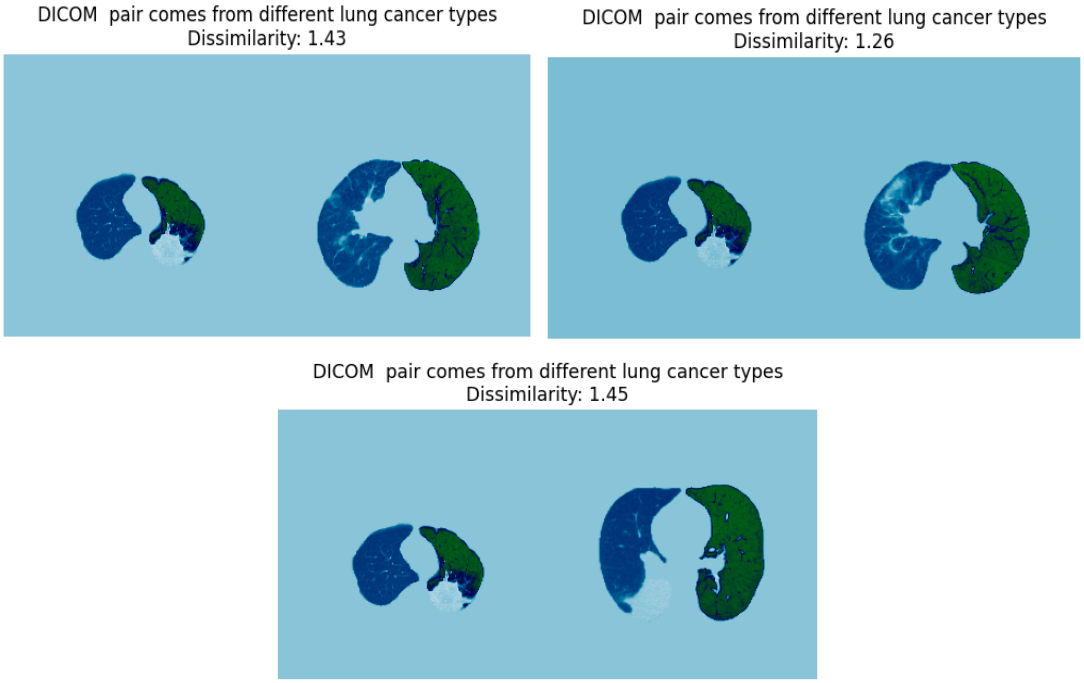


Figure 4.11 Lung image pairs with different lung cancer types and their dissimilarity score

As indicated in figure 4.11, the loss started around 1800 and ended at close to 0. After that, the trained model was tested using a test dataset. For the test dataset 5 pairs of images were selected.

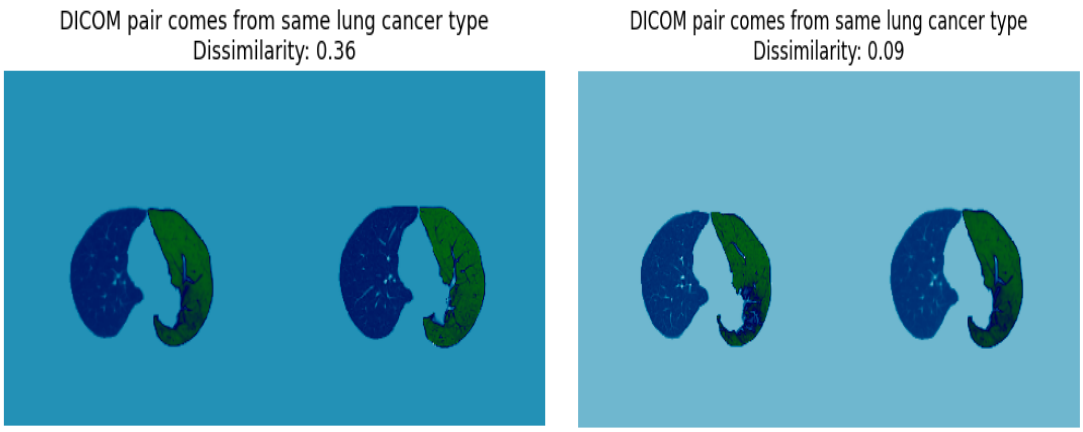


Figure 4.12 Lung image pairs with same lung cancer type and their dissimilarity score

Two images from the test set passed and obtained two vectors and then Euclidean distance between two vectors was calculated using F.pairwise_distance() function in the functional library in the PyTorch module. Below figures show that the pairs from the same lung cancer type have small dissimilarity number which is close to 0. The dissimilarity number is large when the DICOM images from different lung cancer types.

Siamese neural networks and prototypical networks are two different approaches used in few-shot learning tasks. The choice between Siamese Neural Networks and Prototypical Networks depends on the purpose. Siamese neural network can be used to identify the type of lung cancer when we have a CT scan from the same type. Prototypical network can be used where it has limited label examples for each class. Table 4.6 illustrates the comparison between two methods.

	Prototypical Network	Siamese Neural Network
Architecture	Consists of an encoder network, prototypical network, and prototype class	Consist of two identical subnetworks with same weights.
Architecture of the network used	VGG16. Multiple convolutional layers followed by ReLu activation function.	Three convolutional layers followed by ReLu activation and max pooling. Three fully connected layers with ReLU activations.
Training Procedure	Minimize the distance between the support images and the respective class prototype, while maximizing the distance from prototypes of other classes. Meta Learning	Minimize the distance between embeddings of similar instances and maximize the distance between embeddings of dissimilar instances. Contrastive Loss Learning

Table 4.6 Comparison of Prototypical Network and Siamese Neural Network

5 DISCUSSION

The implications of research findings and the methodology taken towards the research objectives are talked about in this stage. Upon analyzing the proportion of intra-class distance to inter-class distance for each feature extraction model, it was seen that the VGG16 model had the most elevated proportion. This indicates that the classes were very much separated in the feature space generated by the VGG16 model. As such, the VGG16 model had the option to make a feature space where the intra-class distances

were more modest contrasted with the inter-class distances. This is a positive indicator of the quality of the feature space and recommends that the VGG16 model might be more qualified for the task of few-shot learning when contrasted with the other feature extraction models.

According to table 4.2 ,4.3, 4.4 and 4.5 present the performance variations of different-trained models within a few-shot learning context for 3 episodes in wordy learning. Each table shows the evaluation metrics for distinct feature extractors, namely a CNN model, DenseNet, and ResNet50, across varying numbers of shots per class. A common trend rises up out of these evaluations, revealing a direct relationship between the number of shots and the model's predictive capacities.

The outcomes show something important: when we have more models for every class, every one of the models perform better. This makes sense in light of the fact that having more models assists the models with understanding things better. Particularly, the accuracy of the models goes up when we have more models. For instance, the accuracy of the ResNet50 model turns into the best at 98% when there are 5 models for every class. This lets us know that having more models is useful for making these models work well.

Moreover, the evaluation of different models recommends that the ideal number of shots per class could contrast depending on the chosen feature extractor. In any case, across the board, 4 or 5-Shot classes tend to yield the most noteworthy accuracy and F1-Score values. This consistent pattern reinforces the notion that enough named models is urgent for vigorous model performance, particularly when handling complex tasks. Generally speaking, the findings underline the importance of considering both the decision of model and the availability of marked data when aiming for ideal few-shot learning results.

When examining the training progress for each pre-trained model from 4.5 to 4.8 figures, which was completed more than 12 episodes with 4 shots for every class, interesting patterns arise. Looking at the VGG16 model, the training accuracy fluctuates all through the episodes. This indicates that the model's learning cycle isn't totally steady, because of variations in the accessible data or the intricacy of the task.

While the training accuracy initially increases, it later drops and fluctuates around the 0.65 to 0.70 range. Correspondingly, the training loss experiences notable fluctuations, peaking in the fourth and fifth episodes before step-by-step decreasing.

For the ResNet50 model, the training accuracy initially wavers around 0.6 prior to showing improvement in episodes two and three. In any case, the accuracy is inconsistent across episodes, and a declining trend is seen towards the end of the training. Conversely, the training loss begins higher however diminishes with slight fluctuations. The CNN model, in contrast, displays consistent advancement regarding training accuracy, with consistent improvement throughout the span of the episodes. This indicates a steady learning process, potentially owing to the model's architecture and data distribution. Likewise, the training loss shows a consistent decline, indicative of effective learning.

Finally, the DenseNet model shows an initially fluctuating training accuracy, however it balances out as training advances. The training loss diminishes, with minor fluctuations, suggesting that the model is adapting to the data effectively. These findings recommend that different pre-trained models display assorted training ways of behaving. While certain models demonstrate sporadic patterns, others show more steady and consistent learning trends. These observations feature the mind boggling nature of training neural networks, influenced by factors like model architecture, data distribution, and convergence dynamics. In the conducted study, the most noteworthy generally speaking accuracy was accomplished by a model employing Convolutional Neural Networks (CNNs) as the feature extractor. The experiment encompassed 15 training episodes, where every episode consisted of 4 support images from each class and 3 query images from each class. The model attained an impressive accuracy of 98%.

The central objective of the proposition project spun around the identification of lung cancer types using CT scan images, with an essential spotlight on ascertaining the most fitting image processing philosophy for medical images in situations where the availability of training data is restricted. In the review, PET-CT DICOM images from the TCIA stage are utilized.

The research will begin by conducting a writing survey on a few shot learning techniques and their applications in medical image analysis. This survey will give a comprehensive understanding of a few shot learning techniques and their potential for lung cancer type detection using CT scans. The central focal point of this project lies in the utilization of a Prototypical Network, as it represents an essential few-shot learning strategy. This decision is driven by our overarching objective of accurately identifying a few types of lung cancer. Through a progression of experiments, the performance of the algorithm was quantitatively surveyed as far as accuracy, precision, recall and F-score, with an emphasis on their reasonableness for detecting different lung cancer types. In addition to the prototypical network, this research likewise encompasses a foundational exploration of the Siamese Neural Network. This network constitutes another few-shot learning approach used to discern the two similitudes and dissimilarities, subsequently contributing to the overarching objective of identifying different types of lung cancer.

Another objective of this research project is to implement a support tool for medical professionals that enables them to identify the sort of lung cancer when a CT scan is given. I have fostered this tool using Streamlit and Python. Since Convolutional Neural Networks accomplished the most noteworthy accuracy through my analysis, I involved a Prototypical network in which a CNN fills in as the feature extractor. In this framework, doctors need to transfer the CT scan in DICOM record design. The tool will then give two prediction results, as shown in Figure 2. The backend part has been created using different modules connected with computer vision using python to achieve the implementation of a supporting tool for medical professionals to identify lung cancer types. The relevant code can be found in here:

https://github.com/Nimesha-Hansani/detection-of-lung-cancer-types-using-FSL.

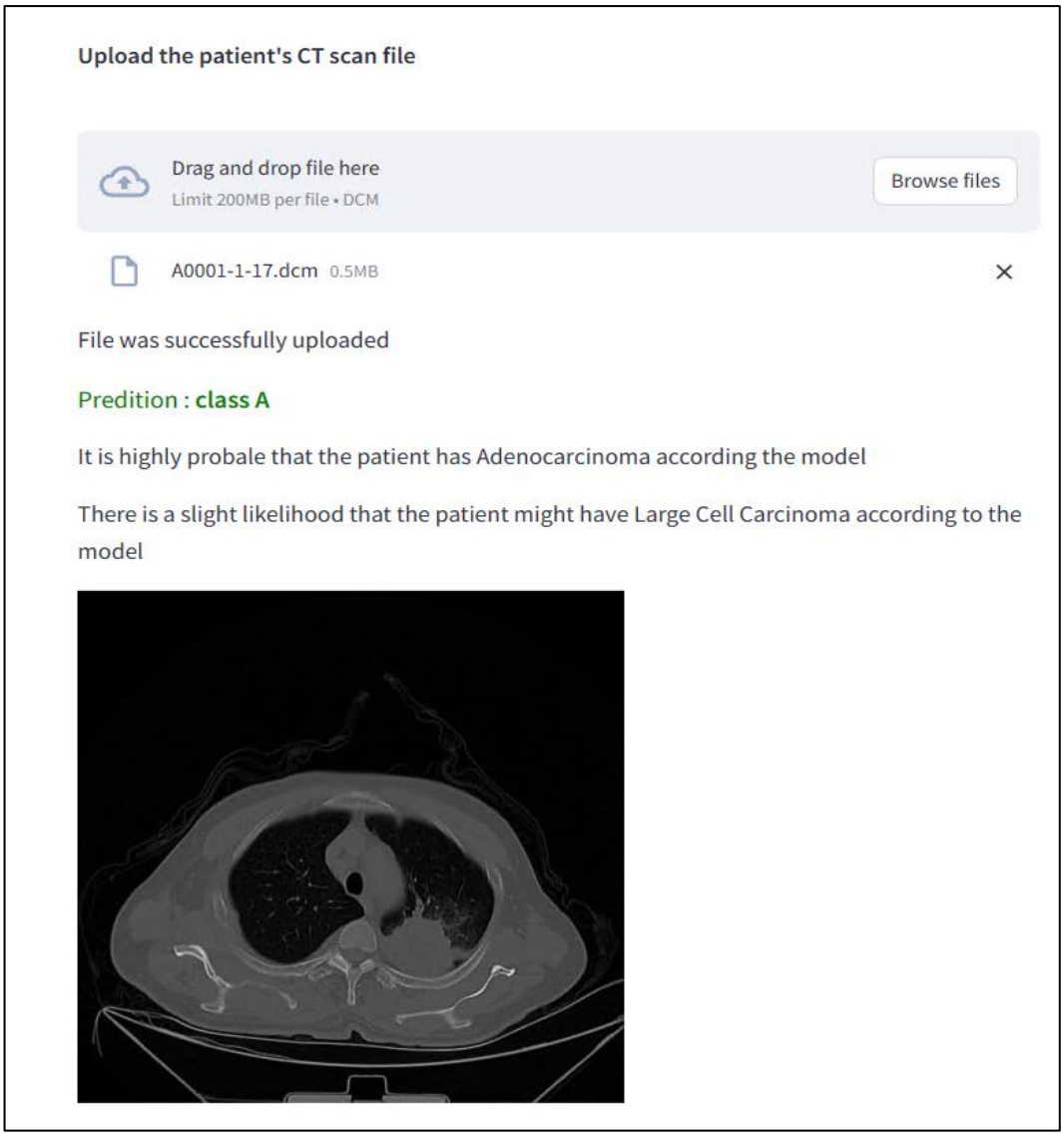


Figure 5.1 Output of the supporting tool

One of the research objectives necessitates an exploration of feature extraction strategies to get features from lung CT scan images. In this proposal pre-trained approaches have been investigated and VGG16, ResNet50, DenseNet and CNN were basically centered around Prototypical Network and CNN were centered around Siamese Neural Network. To contribute to the examinations on few-shot learning for lung cancer type detection, research finding, and philosophies is documented in a structured manner.

The potential of Prototypical Networks and Siamese Neural Networks for enhancing the identification of four lung cancer types from CT scan images is investigated to address the primary research question. Prototypical Networks, by learning model representations for each class, enable a discriminative embedding space where lung cancer types are represented by distinct prototypes. These prototypes work with the effective classification in view of their similarity to the query images embeddings. The performance was investigated by varying the number of CT images for every lung cancer type, and endeavors were made to enhance the accuracy through a meta-learning process.

Siamese Neural Networks, on the other hand, profit by their architecture that actions image similarity, a feature especially helpful for capturing unobtrusive differences among lung cancer variants. Its utilization for lung cancer type detection according to the hypothetical structure isn't possible. Be that as it may, it was effectively utilized to identify dissimilarities among existing lung cancer types.

The second research question investigates the practicality of employing pre-trained models for feature extraction from CT images to enhance accuracy in lung cancer type classification. Three types were selected for the CNN-based pre-trained model and a CNN to investigate how the utilization of pre-trained models can bring about superior accuracy by leveraging their learned representations to encapsulate intricate structures and patterns within CT scan images. The performance results for each feature extraction model were mentioned in the discussion section, along with the alteration in accuracy through meta-learning and variations in the number of shots per class.

6 CONCLUSION

The entire postulation project spun around identifying four main types of lung cancer using CT scan images, with an essential spotlight on determining the most fitting image processing philosophy for medical images when it is restricted to train data. The project basically used the prototypical network, a notable few-shot learning model, to accomplish this objective. Different pre-trained models and a CNN model were utilized to extract a superior feature space from lung CT images. To analyze the

prototypical network's performance, it was trained using meta-learning more than 12 episodes in a 4-shot-4-way setting. It was seen that these four models displayed distinct training ways of behaving. According to the feature space analysis, the best model is VGG16, and every one of the four models demonstrated their most elevated accuracy when each class had 4-5 shots.

In addition to the Prototypical network, we likewise investigated a different few-shot learning approach, namely the Siamese neural network. This approach can be applied to lay out similarity and dissimilarity when dealing with restricted data.

6.1 Limitations and Future Work

Lung cancer research has taken significant steps, thanks to the availability of powerful datasets that have worked with comprehensive analysis and model development. In any case, one of the notable limitations in this proposition rotates around the constrained dataset, especially as far as the variety of lung cancer types. In the domain of lung cancer research, the openness of data is paramount to ensure the accuracy and generalizability of research findings. Nevertheless, ethical considerations attached to accessing patient data present impressive challenges. Adhering to international and nearby guidelines, safeguarding patient privacy and upholding data security are non-negotiable objectives. Consequently, this has given ascent to thorough ethical approval strategies, necessitating researchers to navigate a perplexing landscape of permissions and regulations. The worldwide journey to procure comprehensive and changed datasets explicitly marked for different types of lung cancer is indeed an onerous undertaking.

The insights generated by this research give a foundation to additional exploration and refinement in the field of lung cancer identification using CT scans and few-shot learning techniques. An intriguing avenue for future review is the amalgamation of Siamese Neural Networks and Prototypical Networks. This fusion can possibly yield

an additional strong and precise model for classifying types of lung cancer. Siamese Networks succeed in discerning differences, while Prototypical Networks are proficient at categorization using models. Combining their strengths could address challenges that each network faces in isolation. Mastering the specialty of consistently integrating these networks and addressing any challenges that might arise represents a captivating issue to be settled.

Essential to making the model practical and significant is ensuring its user-friendliness for healthcare professionals, especially doctors. In the future, the development of a user-friendly computer interface, featuring intuitive elements like buttons and images, holds immense commitment. This interface could engage doctors to easily input CT scan images, get quick and accurate predictions regarding the sort of lung cancer, and gain insights into the rationale behind the model's predictions. Collaborating intimately with medical professionals during the interface's design and development will be instrumental in tailoring it to meet their particular needs.

In synopsis, lung cancer research has gained considerable headway, yet data limitations and ethical considerations remain significant challenges. Future research holds the potential for innovative solutions, like network fusion, and user-friendly interfaces that can enhance the appropriateness of these models in clinical practice and work on patient consideration.

7 REFERENCES

[1] "American Cancer Society," [Online]. Available: https://www.cancer.org/cancer/lung-cancer.

[2] A. Meldo and L. Utkin, "A new approach to differential lung diagnosis with CT scans based on the Siamese neural network," *Journal of Physics: Conference Series*, vol. 1236, p. 012058, 06 2019.

[3] N. Wu, J. Phang, . J. Par and Y. Shen, "Deep Neural Networks Improve Radiologists' Performance in Breast Cancer Screening," *IEEE Transactions on Medical Imaging*, vol. 39, pp. 1184--1194, April 2020.

[4] Y. Tang, Y. Tang, J. Xiao and R. M. Summers, "XLSor: A Robust and Accurate Lung Segmentor on Chest X-Rays Using Criss-Cross Attention and Customized Radio realistic Abnormalities Generation," in *MIDL*, 2019.

[5] P. Chathurvedi, A. Jhamp, M. Vanani and V. Nemade, "Prediction and Classification of Lung Cancer Using Machine Learning Techniques," *{IOP} Conference Series: Materials Science and Engineering*, vol. 1099, March 2021.

[6] S. Makaju, P. Prasad, A. Alsadoon, A. Singh and A. Elchouemi, "Lung Cancer Detection using CT Scan Images," *Procedia Computer Science*, vol. 125, pp. 107-114, 2018.

[7] B. G. and F. D. , "Lung cancer: clinical presentation and specialist referral time.," pp. 898-904, Dec 2004.

[8] S. K. K. M. and W. S. , "Radiologic classification of small adenocarcinoma of the lung: radiologic-pathologic correlation and its prognostic impact," February 2006.

[9] G. Xiuhua , S. Tao, W. Huan and L. Zhigang, "Prediction Models for Malignant Pulmonary Nodules Based-on Texture Features of CT Image 1," 2012.

[10] W. TY and B. NM, "Artificial Intelligence With Deep Learning Technology Looks Into Diabetic Retinopathy Screening," pp. 316(22):2366-2367, 13 December 2016.

[11] G. Zhang , L. Lin and J. Wang, "Lung Nodule Classification in CT Images Using 3D," *Journal of Physics: Conference Series*, 2021.

[12] S. Ahuja, B. K. Panigrahi and N. Dey, "McS-Net: Multi-class Siamese network for severity of COVID-19 infection classification from lung CT scan slices," *Applied Soft Computing,* vol. 131, p. 109683, 2022.

[13] Y. Jiang, H. Chen, H. Ko and D. K. Han, "FEW-SHOT LEARNING FOR CT SCAN BASED COVID-19 DIAGNOSIS," in *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2021, pp. 1045--1049.

[14] K. Mijung, Z. Jasper and W. De Neve, "Few-shot Learning Using a Small-Sized Dataset of High-Resolution FUNDUS Images for Glaucoma Diagnosis," 2017.

[15] J.-d.-T. O. a. M. and H. a. K. , "Cloud-Based Evaluation of Anatomical Structure Segmentation and Landmark Detection Algorithms: VISCERAL Anatomy Benchmarks," *IEEE Transactions on Medical Imaging,* vol. 35, pp. 1-1, 2016.

[16] . N. Burlutskiy, F. Gu, L. K. Wilen, M. Backman and P. Mike, "A deep learning framework for automatic diagnosis in lung cancer," *arXiv preprint arXiv:1807.10466,* 2018.

[17] N. E. Protonotarios, I. Katsamenis and S. Sykiotis, "A few-shot U-Net deep learning model for lung cancer lesion segmentation via PET/CT imaging," *Biomedical Physics \& Engineering Express,* vol. 8, p. 025019, 2022.

[18] L. Wang, "Deep Learning Techniques to Diagnose Lung Cancer," *Cancers,* vol. 14, 2022.

[19] J. Kaur and S. Garg, "Improving segmentation by denoising brain MRI images through interpolation median filter in ADTVFCM," *International Journal of Computer Trends and Technology,* vol. 4, pp. 187-188, 2013.

[20] M. Siddeq, "De-noise color or gray level images by using hybred dwt with wiener filter," *Hepato-Gastroenterology,* vol. 61, pp. 1308--1312, 2014.

[21] K. Rajendran, S. Tao, W. Zhou, S. Leng and C. Mccollough, "Spectral prior image constrained compressed sensing reconstruction for photon-counting detector based CT using a non-local means filtered prior (NLM-SPICCS)," in *MEDICAL PHYSICS*, vol. 45, 2018, pp. E675--E675.

[22] X. Sun, S. Wang and L. Zheng, "Industrial robots sorting system based on improved faster RCNN}," *Comput. Syst. Appl,* vol. 28, pp. 258-263, 2019.

[23] M. Keshani, Z. Azimifar, F. Tajeripour and R. Boostani, "Lung nodule segmentation and recognition using SVM classifier and active contour modeling: A complete intelligent system," *Computers in biology and medicine,* vol. 43, pp. 287--300, 2013.

[24] G. Tong, Y. Li, H. Chen, Q. Zhang and H. Jiang, "Improved U-NET network for pulmonary nodules segmentation," *Optik,* vol. 174, pp. 460--469, 2018.

[25] M. Zhang, H. Li, S. Pan and J. Lyu, "Convolutional neural networks-based lung nodule classification: A surrogate-assisted evolutionary algorithm for hyperparameter optimization," *IEEE Transactions on Evolutionary Computation,* vol. 25, pp. 869--882, 2021.

[26] . X.-p. Wang, W. Zhang and Y. Cui, "Tumor segmentation in lung CT images based on support vector machine and improved level set. Optoelectron.," *Optoelectronics Letters,* vol. 11, pp. 395-400, 2015.

[27] M. H. and I. S. , "A Closer Look at Prototype Classifier for Few-shot Image Classification," in *ICLR 2022 Conference,* 2022.

[28] D. Liu, X. Gao and Q. Shen, "Prototypical Network for Radar Image Recognition with Few Samples," *Journal of Physics: Conference Series,* vol. 1634, p. 012116, Sep 2020.

[29] K. Dang, N. Tajbakhsh and X. Ding, "A Location-Sensitive Local Prototype Network For Few-Shot Medical Image Segmentation," 2021.

[30] J. Kotia, A. Kotwal, B. Rishika and R. Mangrulkar, "Few Shot Learning for Medical Imaging," *Studies in Computational Intelligence,* pp. 107-132, 01 2021.

[31] J. Bromley, J. Bentz, L. Bottou, I. Guyon and Y. Lecun, "Signature Verification using a "Siamese" Time Delay Neural Network," *International Journal of Pattern Recognition and Artificial Intelligence,* vol. 7, p. 25, 08 1993.

[32] Y. Yi, C. Chen and T. Zhang, "A Survey on Siamese Network: Methodologies, Applications and Opportunities," *IEEE Transactions on Artificial Intelligence,* vol. PP, pp. 1-21, 12 2022.

[33] G. R. Koch, "Siamese Neural Networks for One-Shot Image Recognition," 2015.

[34] K. Clark, B. Vendt, K. Smith, J. Freyman, J. Kirby, P. Koopel, S. Moore and P. Stanely, "The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository," *Journal of digital imaging,* vol. 26, 07 2013.

[35] J. Hofamanninger, F. Prayer, J. Pan, S. Röhrich and H. Prosch, "Automatic lung segmentation in routine imaging is primarily a data diversity problem, not a methodology problem," *European Radiology Experimental,* vol. 4, p. 50, 08 2020.

[36] O. Ronneberger, P. Fischer and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *LNCS,* vol. 9351, pp. 234-241, 10 2015.

[37] R. Haralick, K. Shanmugam and I. Dinstein, "Textural Features for Image Classification," *IEEE Trans Syst Man Cybern,* Vols. SMC-3, pp. 610-621, 01 1973.

[38] J. Yosinski, J. Clune, Y. Bengio and H. Lipson, "How transferable are features in deep neural networks?," in *Advances in Neural Information Processing Systems*, vol. 27, Curran Associates, Inc, 2014.

[39] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *ICLR 2015,* 09 2014.

[40] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," pp. 770-778, 2016.

[41] G. Huang, Z. Liu and K. Q. Weinberger, "Densely Connected Convolutional Networks," *2017 IEEE Conference on Computer Vision and Pattern Recognition(CVPR),* pp. 2261.-2269, 2016.

[42] J. Snell, K. Swersky and R. Zemel, "Prototypical Networks for Few-shot Learning," 2017.

[43] R. Siegel, K. Miller and A. Jemal, "Cancer statistics," *CA: a cancer journal for clinicians,* vol. 69 (1), pp. 7-34, 2019.

[44] M. o. Defence, Machine Learning with Limited Data, DSTL/PUB126061.

[45] N. Khehrah, M. S. Farid, . S. Bilal and M. K. Hassan, "Lung Nodule Detection in CT Images Using Statistical and Shape-Based Features," *Journal of Imaging,* vol. 6, 2020.

[46] C. S. Feng and D. Pan, "An improved approach of lungs image segmentations based on watershed algorithms," in *7th International Conference on Internet Multimedia Computing & Service,* 2015.

[47] M. Lavanaya and M. Kanan, "Lung Lesion Detection in CT Scan Images Using the Fuzzy Local Information Cluster Means (FLICM) Automatic Segmentation Algorithm and Back Propagation Network Classification," *Asian Pacific Journal of Cancer Prevention,* p. 3395–3399., 2017.

[48] T. S. Roy, N. Sirohi and A. Patle, "Classification of lung image and nodule detection using fuzzy inference system," in *International Conference on Computing, Communication Automation*, 2015.

[49] M. Fink, "Object Classification from a Single Example Utilizing Class Relevance Metrics," in *Advances in Neural Information Processing Systems*, MIT Press, 2004, p. 449–456.

[50] A. Parnami and M. Lee, "Learning from Few Examples: A Summary of Approaches to Few-Shot Learning," 2022.

[51] M. Drozdzal, E. Vorontsov, G. Chartrand and S. Kadoury, "The Importance of Skip Connections in Biomedical Image Segmentation," 08 2016.

[52] Z. Xiao, B. Liu, L. Geng, F. Zhang and Y. Liu, "Segmentation of lung nodules using improved 3D-UNet neural network," *Symmetry*, vol. 12, p. 1787, 2020.

[53] "The Cancer Imaging Archive," [Online]. Available: https://www.cancerimagingarchive.net.