FIRAT UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
TÜRKİYE

DETECTION OF CANCER AREA IN LUNG IMAGES WITH THE HELP OF DEEP LEARNING ALGORITHM

Shivan Hasan Mohammed

Master's Thesis

DEPARTMENT OF COMPUTER ENGINEERING
Program of Graduate School of Natural and Applied sciences

JANUARY 2021

FIRAT UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
TÜRKİYE

Department of Computer Engineering
Graduate School of Natural and Applied sciences

Master's Thesis

DETECTION OF CANCER AREA IN LUNG IMAGES WITH THE HELP OF DEEP LEARNING ALGORITHM

Author
Shivan Hasan Mohammed

Supervisor
Assoc. Prof. Ahmet ÇINAR

JANUARY 2021
ELAZIG

FIRAT UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
TÜRKİYE

Department of Computer Engineering
Graduate School of Natural and Applied sciences

Master's Thesis

Title:	Detection Of Cancer Area In Lung Images With The Help Of Deep Learning Algorithm
Author:	Shivan Hasan Mohammed
Submission Date:	03 December 2020
Defense Date:	25 January 2021

THESIS APPROVAL

This thesis, which was prepared according to the thesis writing rules of the Graduate School of Natural and Applied Sciences, Fırat University, was evaluated by the committee members who have signed the following signatures and was unanimously approved after the defense exam made open to the academic audience.

		Signature
Supervisor:	Assoc. Prof. Ahmet ÇINAR Firat University, Faculty of Science	Approved
Chair:	Assoc. Prof. Taner TUNCER Firat University, Faculty of Science	Approved
Member:	Prof. Dr. Ali KARCI Inonu University, Faculty of Science	Approved

This thesis was approved by the Administrative Board of the Graduate School on

.... / .... / 20 ....

Signature

Assoc. Prof. Dr. Kürşat Esat ALYAMAÇ
Director of the Graduate School

DECLARATION

I hereby declare that I wrote this Master's Thesis titled " Detection Of Cancer Area In Lung Images With The Help Of Deep Learning Algorithm" in consistent with the thesis writing guide of the Graduate School of Natural and Applied Sciences, Fırat University. I also declare that all information in it is correct, that I acted according to scientific ethics in producing and presenting the findings, cited all the references I used, express all institutions or organizations or persons who supported the thesis financially. I have never used the data and information I provide here in order to get a degree in any way.

25 January 2021

Shivan Hasan Mohammed

PREFACE

I would like to dedicate this dissertation to my family. A special dedication to my mother and my father, who encouraged me to further my study and career, and who also, taught me to never stop setting and reaching newer and bigger goals. I dedicated to my dear supervisor Assoc. Prof. Ahmet ÇINAR

Shivan Hasan Mohammed
ELAZIG, 2021

TABLE OF CONTENTS

ABSTRACT

Detection Of Cancer Area In Lung Images With The Help Of Deep Learning Algorithm

Shivan Hasan Mohammed

Master's Thesis

FIRAT UNIVERSITY
Graduate School of Natural and Applied Sciences

Department of Computer Engineering
Graduate School of Natural and Applied sciences
January 2021, Page: xi + 39

With the acceleration of the process of industrialization, and the increasing number of smokers, the number of deaths due to lung cancer remains high. A large amount of clinical data found that lung cancer appeared in the form of lung nodules in the early stage. CT imaging is one of the non-invasive detection methods to find early lung nodules. The use of computer-assisted CT lung nodule detection system can help doctors in early detection and diagnosis. However, due to the difference in radiation doses and machine parameters, the brightness, contrast, or irradiation angle of the image database of different hospitals are different. The training of the detection system cannot improve the performance of lung nodule classification algorithm by adding data. This causes problems such as low data usage, high label training costs, and easy expiration of classification models. This research provides the Convolutional Neural Network (CNN) and its architectures, such as AlexNet, GoogleNet, Resnet18 and Resnet50. We have included transfer learning by using CNN's pre-trained architectures. These architectures are tested with large data sets which are SPIE AAPM lung image dataset, which is freely available on the web. Due to the lack of training data, data augmentation technique has proposed to increase the number of training data artificially. The Convolution Neural Network architectures' performances were evaluated by some matrixes, which are confusion matrix, recall, precision and f1-score. Based on the above work, Matlab R2019b is used to program the pre-trained networks and obtaining the results based on transfer learning.

Keywords: Lung cancer; Transfer Learning; pre-trained model; CNN.

ÖZET

Derin Öğrenme Algoritması İle Akciğer Görüntülerinde Kanser Alanının Tespiti

Shivan Hasan Mohammed

Yüksek Lisans Tezi

FIRAT ÜNİVERSİTESİ
Fen Bilimleri Enstitüsü

Bilgisayar Mühendisliği Anabilim Dalı
Fen Bilimleri Enstitüsü
Ocak 2021, Sayfa: xi + 39

Sanayileşme sürecinin hızlanması ve sigara içenlerin sayısının artması ile akciğer kanserine bağlı ölümlerin sayısı yüksek olmaya devam ediyor. Büyük miktarda klinik veri, akciğer kanserinin erken aşamada akciğer nodülleri şeklinde ortaya çıktığını buldu. CT görüntüleme, erken akciğer nodüllerini bulmak için invazif olmayan tespit yöntemlerinden biridir. Bilgisayar destekli CT akciğer nodülü tespit sisteminin kullanılması, doktorlara erken teşhis ve teşhis konusunda yardımcı olabilir. Bununla birlikte, radyasyon dozları ve makine parametrelerindeki farklılık nedeniyle, farklı hastanelerin görüntü veritabanlarının parlaklığı, kontrastı veya ışınlama açısı farklıdır. Saptama sisteminin eğitimi, veri ekleyerek akciğer nodülü sınıflandırma algoritmasının performansını iyileştiremez. Bu, düşük veri kullanımı, yüksek etiket eğitim maliyetleri ve sınıflandırma modellerinin kolay sona ermesi gibi sorunlara neden olur. Bu araştırma, Evrişimli Sinir Ağı'nı (CNN) ve AlexNet, GoogleNet, Resnet18 ve Resnet50 gibi mimarilerini sağlar. CNN'nin önceden eğitilmiş mimarilerini kullanarak aktarım öğrenimini dahil ettik. Bu mimariler, web'de ücretsiz olarak bulunabilen SPIE AAPM akciğer görüntüsü veri seti olan büyük veri setleriyle test edilir. Eğitim verilerinin eksikliğinden dolayı, veri artırma tekniği eğitim verisi sayısını yapay olarak artırmayı önerdi. Evrişim Sinir Ağı mimarilerinin performansları, karışıklık matrisi, geri çağırma, kesinlik ve f1-score olan bazı matrislerle değerlendirildi. Yukarıdaki çalışmaya dayanarak, Matlab R2019b, önceden eğitilmiş ağları programlamak ve aktarım öğrenmeye dayalı sonuçları elde etmek için kullanılır.

Anahtar Kelimeler: Akciğer kanseri; Transfer Öğrenimi; önceden eğitilmiş model; CNN.

LIST OF FIGURES

Page

LIST OF TABLES

Page

ABBREVIATIONS

DL	: Deep Learning
TL	: Transfer Learning
ML	: Machine Learning
CT	: Computed Tomography
CNN	: Convolutional Neural Network
DA	: Data Augmentation
AAPM	: American Association of Physicists in Medicine
NCI	: National Cancer Institute

1. INTRODUCTION

In recent years, environmental pollution and food safety issues have been a serious threat to people's health, most urban PM2.5 high, triggered a large number of pneumonia and lung cancer and other lung diseases [1]. According to incomplete statistics, fine airborne particles are the leading cause of lung cancer. In addition to improving the living environment, for lung cancer, early detection, early diagnosis, and early treatment are also necessary means to improve patient survival. Lung cancer has extremely complex characteristics in clinical medicine. Patients have a shorter onset time and have a higher probability of being malignant. Once discovered, patients are mostly in the late stage of the disease, and the cure rate is extremely low.

The main feature of the medical and clinical tissue anatomy diagnosis method for lung cancer is to make a clinical diagnosis of lung cancer based on histology and lung cancer imaging anatomy. The main features of the histoanatomical diagnosis method of lung cancer include bronchoscopy anatomical examination, biopsy tissue anatomy examination and sputum cancer cell anatomy examination of lung cancer. However, such diagnostic methods can be used as a diagnosis in sputum and the basis and standards for detecting early lung cancer. Still, for those lung cancer patients at different early stages of the disease and different life types, cancer cells may not always appear in their sputum. The main feature of the diagnostic Method of lung cancer imaging anatomy is to intuitively understand and observe the size, morphology and biopsy of cancer cells and human lung lesions by using computerized tomography and x-ray anatomical diagnosis methods of lung cancer. Due to its location, this diagnostic Method effectively avoids the interference and damage caused by cancer cell biopsy tissue to the body of lung cancer patients. Nowadays, the Method of imaging anatomy for the diagnosis of lung cancer is the most widely used diagnostic Method in the current diagnosis of lung cancer.

In recent years, with the continuous development of tomography technology, this technology has become the most effective lung cancer test method. The lung Computed Tomography (CT) image is a single-channel high-level image. The doctor can find the area of cancer and it is benign or malignant by observing the CT image of the chest scan and combining the patient's clinical symptoms. It is precise because chest CT scans have become more and more common that it also brings tremendous pressure to doctors to read the film. In the process of reading the film based on personal work experience and relevant knowledge, doctors will inevitably have missed and misjudged situations; this is extremely unfavorable for the diagnosis and treatment of patients. Therefore, the use of computer technology to process the lung CT image model can significantly improve the speed and accuracy of reading the picture, which is also a popular research direction in lung cancer detection and diagnosis in recent years.

Deep learning is a rapidly developing field and has many influences on the field of medical imaging. Currently, medical images are diagnosed by radiologists, physicians, etc. But this diagnosis becomes very subjective. Radiologists usually have to examine a large number of these images carefully, and long-term inspection of these images may cause fatigue and lead to errors. Therefore, this needs to be automated. Deep learning algorithms such as Convolutional Neural Network (CNN) are commonly used to detect and classify tumors. But they are usually limited by assumptions made when defining the elements. This leads to a decrease in sensitivity. However, deep learning may be an ideal solution because these algorithms can learn features from raw image data. One of the challenges in implementing these algorithms is the lack of labeled medical image data [2]. Although this is a limitation on all applications of deep learning, the restriction on medical image data is even more significant due to the confidentiality of patients.

In this research, to identify the pulmonary nodules as benign or malignant, the use of a standard dataset has proposed to test the competitors more precisely. SPIE-AAPM [3] dataset is utilized in our research.

The main feature of the early detection and analysis method of lung cancer based on this deep learning is to extract the features of lung nodules in the included images by analyzing the images of lung CT through the computer system and its auxiliary diagnosis system. The detection, segmentation, and classification of early lung nodules that may be contained in the obtained images are studied, and finally, it can be diagnosed and judged whether the patient's lung nodules may have lung cancer. Among them, the performance of the system is mainly reflected in the ability to correctly classify the early lung nodules and their severity of danger.

The early detection data of lung cancer experiment in this article mainly comes from SPIE-AAPM, an open source CT image database for lung detection. The image database already contains a large number of early lung nodule samples and their corresponding lung nodule labels. Therefore, the SPIE-AAPM database is often used by researchers as a model to verify the correct classification of benign and malignant lung nodules, including the optimized model of the lung nodule classifier and the image feature extraction of the classifier. The main purpose of classifying images of benign and malignant lung nodules is to provide doctors and patients with lung nodules with a more scientific and reliable auxiliary classification result, so that his diagnosis and treatment process can be more accurate, and in this way effectively help patients reduce the clinical examination and reading and workload of pulmonary nodules doctors.

In this work, Transfer learning [4] is one of the most common strategies of deep learning has been proposed. Transfer learning, as the name implies, is to transfer the learned and trained model parameters to a new model to help new model training. Considering that most data or tasks are related, we can use transfer learning to share the learned model parameters (also known as the knowledge learned by the model) to the new model in some way to speed up and optimize The learning efficiency of the model does not need to learn from scratch (starting from scratch, tabula rasa) like most networks.

Transfer learning is a useful technique where we might use a pre-trained model like (AlexNet [5], ResNet18 [6], Googlenet [7], and ResNet50) and adjust the network for the next application by making specific improvements in the architecture of the network [8]. AlexNet and the other pre-trained networks have trained for 1000 classes of real-world images. We can use this network to identify some other classes with some adjustment to the network (i.e.in our research, we used two classes Benign and Malignant, instead of 1000 classes; furthermore, we modified the last three layers and fully connected layer to adjust for our work.

1.1. Literature Review

Medical imaging technology plays an essential role in the detection, diagnosis and treatment of diseases. Due to the instability of human expert experience, deep learning technology is expected to assist researchers and physicians in improving the accuracy of imaging diagnosis and treatment and reducing the imbalance of medical resources [9]. Many researchers have employed deep learning in the field of medical image.

E. Cengil et al. [10] proposed deep learning approach in order to classify lung cancer detection and classification on binary dataset. This study has used TensorFlow libraries and 3D-CNN architecture from deep learning. The data have used is from a SPIE-AAPM-LungX dataset of CT scan. Furthermore, after evaluating the model by confusion matrix, the model accuracy got 70 %.

Z. Shi et al.[11] The study have implemented deep Convolutional Neural Network rely on Transfer learning (TL) to detect a pulmonary nodule on CT images, it employed a pre-trained model in Convolutional Neural Network architectures well-known VGG16 that is used for feature extraction and also the study is used Support Vector Machine (SVM) to classify a nodule from lung CT images.

B. Almas et al.[12] This study, the performance of Convolutional Neural Networks such as GoogleNet and AlexNet are examined for the detection of lung cancer. This research used LIDC-IDRI cancer image dataset. Using GoogleNet and AlexNet, the preprocessed cancer images are trained to determine the portion of the lungs suffering from cancer. The detection of lung cancer by the use of GoogLeNet and AlexNet is used for network training and identification of images.

T. K. Sajja et al. [13] The implementation of convolutional neural network (CNN) has demonstrated to classify between benign and malignant tissues on CT images. To decrease the computation time, the authors recommended a deep neural network based on GoogleNet with a maximum dropout ratio. At the time of learning, this network decreases the over-fitting by using the dropout layer. The densely connected architecture of the proposed network was sparsified, with 60 percent of all neurons installed on dropout layers, to reduce the computational expense and prevent over-fitting in network learning. This study used LIDC-IDRI cancer image dataset and compared with several pre-trained Convolution Neural Networks, such as, AlexNet, GoogleNet and ResNet50.

H. Jiang et al. [14] This research proposed an efficient pulmonary nodule identification method based on multiple group patches taken from the lung images, which are improved by the Frangi filter. In this study, convolution neural networks model is constructed to learn the knowledge of radiologists for detecting nodules on CT images. The scheme was significantly reduced the false positive depending on the results.

R. V. M. Da Nóbrega et al. [15] Inspired by the performance of deep learning in the classification of natural and medical images, the approach proposed investigates the efficiency of deep transfer learning for malignancy classification of lung nodules. This research used pre-trained networks like (VGG16, VGG19, MobileNet, Xception, InceptionV3, ResNet50, Inception-ResNet-V2, DenseNet169, DenseNet201, NASNetMobile and NASNetLarge) as feature extractors to process on CT scans. The Lung Image Database Consortium and Image Database Resource Initiative (LIDC/IDRI) was used for training the networks. For classifying these feature extraction, they used Naive Bayes, MultiLayer Perceptron (MLP), Support Vector Machine (SVM), Near Neighbors (KNN) and Random Forest (RF) classifiers. In addition, the comparison for evaluation of performances' classifiers, some matrix have computed such as, Accuracy, Area Under the Curve, True Positive Rate, Precision and F1-Score.

R. Anirudh et al. [16] the authors have implemented a three-dimensional convolutional neural network CNN architecture to detect lung nodule on CT images. This study, for training the Convolution Neural Network, they utilized unsupervised segmentation to grow out the 3D region on images. The SPIE-LUNGx dataset has experimented through the training the models.

R. M. Devarapalli et al. [17] in this research, there are different types image processing approach have innovated to detect cancer on images which is Implemented such as median-wiener filter in the preprocessing step. To detect whether the nodule is cancerous or non-cancerous, the classification algorithm have utilized , such as Support Vector Machines (SVM), Forward Neural Networks, Back Propagation model, and Convolution Neural Networks (CNN).

Throughout the literature studies, deep learning have conducted with lots of different algorithms; owing to its approval, academic and science domain have recently acquired deep learning. In this research, the lung cancer classification is given out, demonstrating to the early detection of lung cancer. For processing the models on CT scans images a Convolutional neural networks have advocated. Some of the popular pre-trained networks namely (Alexnet, Googlenet, Resnet18 and Resnet50) are used, and Transfer Learning [4] for implementing our thesis.

1.2. Purpose of the study

The purpose of this thesis is to use deep learning algorithms in the process of determining whether there are cancers on lung images within biomedical images or not. Especially by using deep learning models to ensure a more accurate and more robust structure is revealed. Classification and clustering methods for this process will be examined, and the most suitable of these deep learning architectures will be revealed. Results will be given comparatively.

1.3. Outline of the thesis

The current research is presented as follow:

Chapter Two is a theoretical background of the thesis. In addition, CNN architectures are described and provide definitions of deep learning and classification algorithms. Moreover, it gives the descriptions of the evaluation matrix, and models adopted in this thesis.

Chapter Three present with the methodology used in this study. Furthermore, this chapter, the explanation of deep learning techniques was provides. Also, Chapter Three will explain the dataset and several evaluation methods which are used in the study.

In Chapter Four, the results of the thesis will be presented, and also the comparison with other existing methods will be demonstrated. Moreover, this chapter is focused on how model have experimented the results.

Chapter Five illustrate the conclusion of the research and future works.

2. THEORETICAL BACKGROUND

2.1. Lung cancer

Lung cancer refers to the uncontrolled growth of abnormal cells in one or both lung tissues, which have the property of developing into malignant tumors [18]. Lung cancer can exist in the trachea (bronchi) or cavernous lung tissue (alveoli). Lung cancer may also be caused by cancers that have spread from another part of the body, but these cancers are not considered lung cancers themselves.

Primary lung cancer can be divided into two categories:

- **Non-small cell lung cancer**

Non-small cell lung cancer is the more common of the two types of lung cancer, accounting for about 80% of lung cancers, and is less malignant than small cell lung cancer. If caught early, non-small cell lung cancer may be cured by surgery, radiation therapy or chemotherapy.

- **Small Cell Lung Cancer**

Small cell lung cancer grows fast and spreads quickly to other parts of the body through blood vessels and lymph. Generally, patients are already at an advanced stage when they are diagnosed with this cancer. Treatment is usually chemotherapy, or a combination of chemotherapy and radiation therapy.

2.2. Causes of lung cancer

Doctors cannot fully explain why some people get lung cancer, but others do not. But what we do know is that some risk factors may make people more susceptible to lung cancer. Smoking is the most critical risk factor for lung cancer, and more than 80% of lung cancer cases worldwide are caused by smoking. Harmful substances in tobacco can cause damage to lung cells. Over time, these cells may become cancerous. This is why cigarettes, pipes or cigars can cause lung cancer. Non-smokers may also develop lung cancer from inhaling second-hand smoke. The more smoke you inhale, the higher your risk of cancer [19].

Similarly, the earlier a person starts smoking and the more years of smoking, the greater the risk of lung cancer. Other risk factors for lung cancer include exposure to radon (a radioactive gas) radiation, asbestos, arsenic, chromium, nickel, soot, tar, etc., air pollution, and ageing in the workplace. People with a family history of lung cancer have a slightly higher risk. People who have had lung cancer have a higher chance of recurring tumors in the lungs. The best way to prevent lung cancer is martial law or not to start smoking.

2.3. Cancer screening

Screening refers to testing to find potential cancer before symptoms appear. The goal is to detect cancer as early as possible so that it can be treated at an early stage. Screening tests can help doctors detect and treat cancer as early as possible. Several possible tests for lung cancer screening have been studied, such as low-dose lung CT scans. However, studies have found that this scan is only effective for high-risk groups.

2.4. Deep Learning

Deep learning (DL) is a research direction in the field of machine learning. It is a representation learning algorithm composed of multiple complex non-linear transformation structures, including multiple intermediate processing layers, which can automatically perform high-level abstraction on data [20]. Generally, deep neural networks are feed-forward networks with one or more hidden layers, similar to shallow neural networks, and can provide modeling for extremely complex non-linear problems. However, the more layers of the shallower network can abstract the model at a higher level, thereby improving the ability of the model. There are many types of deep learning models, such as deep confidence networks. Each layer in the network is a typical restricted Boltzmann machine, which can be trained layer by layer without supervision; convolutional neural networks greatly reduce the number of parameters that need to be optimized. Accepts two-dimensional or three-dimensional image block input, suitable for capturing local information of vision; cyclic neural network is a type of sequence data input, information flow is one-way transmission, with feedback connection and memory neural network, suitable for extracting sequence nonlinearity the timing characteristics of the data.

In today's medical field [9], the diagnosis method of medical images is still manual reading. Manual reading depends on the clinical experience of doctors, the workload is enormous, subjective factors have a significant influence on the diagnosis results, and it is easy to cause misdiagnosis. These problems indicate that it is imperative to seek more efficient medical image analysis methods. DL learns knowledge from a large number of specific sample data (internal laws and representation relationships in the data, etc.), and then applies the learned knowledge to real-world scenarios to solve practical problems, which can reduce the interference of subjective factors; DL can learn data Multiple abstract levels of features can construct more complex models; DL has demonstrated excellent performance in speech processing, image processing, machine translation, natural language processing and other fields, especially in the field of computer vision where natural images are the processing object A breakthrough has been made. These characteristics make the application of deep learning methods and medical image analysis possible.

2.5. Classification

Image classification is one of the earliest applications of DL into the medical imaging field, which can be divided into whole image classification or target classification. Entire image classification usually takes the entire image and its category labels as input and performs end-to-end training on the DL model to learn the data characteristics of its specific category.

Image classification and recognition based on deep learning has always been a very important research direction in the field of medical image analysis. The essence of the problem of image classification and recognition is the process of data fitting. It is a network model that can simulate human brain analysis to understand the data through a trained network model [7], predict the input inspection image data, and output a description of whether there is a certain disease or a quantified and graded diagnostic variable of the severity of a certain disease to assist doctors in diagnosing the condition and subsequent image research. At present, a considerable number of studies have confirmed that the image classification and recognition method based on deep learning has apparent advantages in processing sensitivity, true negatives, and accuracy.

Today, convolutional neural networks are widely used and have become one of the mainstream technologies in image classification and recognition.

2.6. Deep Learning in medical images field

Deep learning indicates a group of algorithms which is used numerous machine learning algorithms to work out many problems such as images, videos and texts on a multi-layer neural network [20]. Deep learning can be divided into neural networks from a broad category, but there are several changes in a particular performance. The fundamental of deep learning is feature learning, which goals to get hierarchical feature information through a hierarchical network, so regards to solve the significant problem of manually designing features in the recent. Deep learning is a substructure that includes many powerful algorithms, such as Convolutional Neural Networks (CNN).

In the past few decades, with the speedy expansion and popularization of medical imaging tools, medical image data has exploded. How to efficiently and precisely analyze medical images has become a significant challenge. On the other hand, driven by electronic information technology, the use of computer technology to assist medical image analysis can not only improve the efficiency of radiologists' analysis and diagnosis but also improve their accuracy.

Therefore, computer-aided detection and diagnosis are becoming an intersecting research field that has received more and more attention. In the last few years, deep learning technology has obtained great success in image recognition, target detection and other fields, and has begun to be applied to medical image analysis. The research content of this thesis is to use deep learning technology to automatically detect diseased tissues in medical images, especially in Lung images. One of the primary applications of deep leaning that used in several types of research is image classification methods.

In medical image classification, one or more images are usually used as input, and a single diagnostic variable is used as output (for example, whether there is a disease). CNN is the current standard in the field of image classification [9].

Image classification is one of the earliest applications of deep learning into the medical imaging field, which can be divided into whole image classification or target classification. Entire image classification usually takes the entire image and its category labels as input. It performs end-to-end training on the deep learning model to learn the data characteristics of its specific category. Typical work is the classification of lung cancer published in IEE by Firat University in 2017 [10]. Target classification is commonly used in medicine to classify or classify specific lesions, such as the classification of benign and malignant lung nodules.

2.7. Convolution Neural Network

CNN is a discriminative depth structure, including some modules; each module is generally composed of a convolutional layer and a pooling layer. If necessary, there are other layers behind, namely corrected linear unit (ReLu) and batch normalization (BN). The last part of the network is generally a fully connected layer, which constitutes a standard multi-layer neural network. Concerning structure, these modules are commonly assembled one by one to form a depth model, so that space and configuration information of the input 2D or 3D image can be fully utilized [20]. By performing the convolution operation on the input image, the convolution layer shares all the weights. In actual fact, the function of the convolutional layer is to detect the local features of different positions in the input image/feature map such as medical CT scans. In medical image analysis, CNN can be said to be one of the most successful applications of deep learning in the field of medical diagnosis. In 2015, scholars from the University of South Florida used one of the variants of CNN-Multi-scale Convolution Neural Network (Multi-scale Convolution Neural Network) to enable computers to identify lung nodules (pulmonary nodules) from chest CT scan images. It is one of the bases for the diagnosis of lung cancer), and the accuracy of the proposed model is as high as 86.84%. However, using images of lung cancer, Teramoto-trained CNN successfully detects lung cancer with only 71% accuracy.

2.7.1. Convolutional layer

The convolutional layer has two fundamental properties: the first property is the local connection. Each neuron in the convolutional layer (assuming it is the i-th layer) is only connected to the neuron in a particular internal window in the next layer (the i-1th layer) to form a local connection network. The convolution kernel as a parameter is the same for all neurons in the i th layer. The second property is weight sharing. Weight sharing can be understood as a convolution kernel that only captures a specific local feature in the input data, so if you want to extract multiple features, you need to use multiple different convolution kernels. As shown in the Figure 2.1 (a) is the fully connected layer in the fully connected neural network. The weight matrix of the fully connected layer has many parameters, and the training efficiency will be relatively low. Furthermore, Figure 2.1 (b) is the convolutional layer in the convolutional neural network. The weights of the connections of the same color in the convolutional layer are the same so that the number of parameters of the weight matrix obtained is greatly reduced.

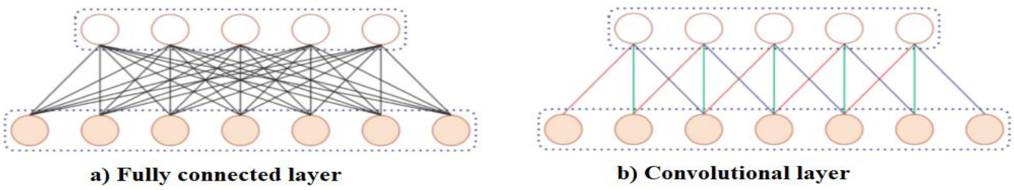

Figure 2.1. CNN-Convolutional Layers

2.7.2. Pooling layer

The position of pooling layer is to perform feature selection and decrease the number of features, thereby reducing the number of arguments. However, the convolutional layer can significantly decrease the number of connections in the network; the number of neurons in the feature map set does not significantly decrease. If followed by a classifier, the input dimension of the algorithm is still very high, and it is facing downwards to over-fitting. To address this issue, a pooling layer can be added after the convolutional layer to decrease the feature dimension and get rid over-fitting. Pooling refers to down-sampling each area to obtain a value as a generalization of this area. Commonly used pooling functions are:

(1) Max Pooling: The most common pool type is max pooling, which occupies the maximum value in each window. These window sizes need to be specified in advance. This will reduce the size of the feature map while retaining important information.

(2) Average /Mean Pooling: generally take the average value of all neuron activity values in the area. The following is an example of the maximum pooling process in the pooling layer, as shown in Figure 2.2:

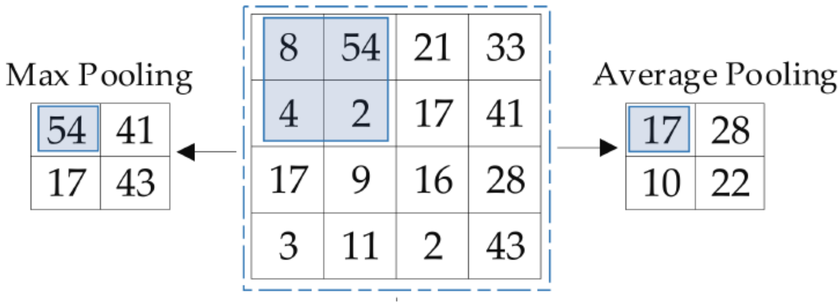

Figure 2.2. Operation of pooling.

2.7.3. Fully connected layer

Fully Connected Layer is a process of linear feature mapping, which maps multi-dimensional feature input to two-dimensional feature output. The high-dimensional represents the sample batch, and the low-level often corresponds to the task goal (for example, classification corresponds to each category Probability). The fully connected layer chiefly repairs the features to decrease the loss of feature information; the output layer is mainly prepared to output the final target result.

2.7.4. Dropout

Dropout [21] is a regularization technique to prevent over-fitting. The specific method is to make a probability decision on the input of each hidden layer. For example, we set the probability to 0.5 (usually named keep_prob), and according to 0.5, randomly generate one and hide A vector with the same number of neurons in the layer, the ratio of true:false is 1:1 (because keep_prob=0.5), multiplying with the neurons in the hidden layer, then half of the neurons in the hidden layer will be discarded and will not participate in training . Repeat the iterative appeal operation. Figure 2.3 demonstrate dropout technique.

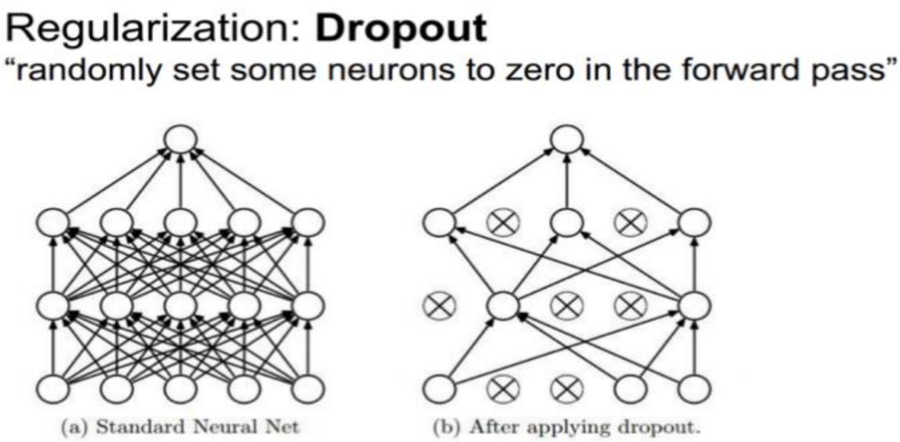

Figure 2.3. dropout technique [22]

If we have multiple different models to make predictions together, the generalization error will be effectively reduced. The problem is that training multiple models is computationally expensive. Dropout provides us with a new idea: let these models share the same weight coefficients, but the output results of neurons are not the same. Expressly, the output of the hidden neuron is set to zero with a 50% probability. In this way, each time backpropagation, the reference loss is calculated by a different model. In general, Dropout technology breaks the dependence between neurons and forces the network to learn more robust neuron connections. We only use it in the fully connected layer because there are so many connections in the fully connected layer. Dropout is not used in the testing phase. Dropout will prolong the convergence time, but can effectively suppress over-fitting.

2.7.5. Batch Normalization

Since there are cascaded non-linear operations in the network, the multi-layer architecture is highly non-linear. In addition to rectification nonlinearity, normalization is another non-linear processing module that plays an essential role in the CNN architecture. Batch Normalization [23] (batch normalization) realizes the preprocessing operation in the middle of the neural network layer, that is, the input of the previous layer is normalized and then enters the next layer of the network, which can effectively prevent "gradient dispersion ", to accelerate network training. The specific algorithm of Batch Normalization is shown in the Figure 2.4.

Input: Values of x over a mini-batch: $\mathcal{B} = \{x_{1\ldots m}\}$;
Parameters to be learned: γ, β

Output: $\{y_i = \mathrm{BN}_{\gamma,\beta}(x_i)\}$

$$\mu_{\mathcal{B}} \leftarrow \frac{1}{m}\sum_{i=1}^{m} x_i \qquad \text{// mini-batch mean}$$

$$\sigma_{\mathcal{B}}^2 \leftarrow \frac{1}{m}\sum_{i=1}^{m} (x_i - \mu_{\mathcal{B}})^2 \qquad \text{// mini-batch variance}$$

$$\widehat{x}_i \leftarrow \frac{x_i - \mu_{\mathcal{B}}}{\sqrt{\sigma_{\mathcal{B}}^2 + \epsilon}} \qquad \text{// normalize}$$

$$y_i \leftarrow \gamma \widehat{x}_i + \beta \equiv \mathrm{BN}_{\gamma,\beta}(x_i) \qquad \text{// scale and shift}$$

Figure 2.4. Batch Normalization

2.7.6. Activation Layer

The Activation Layer is responsible for activating the special diagnosis extracted by the convolutional layer. Since the convolution operation is to perform the corresponding linear transformation on the input image and the convolution kernel, an activation layer (non-linear function) needs to be introduced to make it non-linear Mapping. The activation layer is composed of non-linear functions, as shown in Figure 2.5, such as sigmoid, tanh, and relu. The most commonly used activation function is Relu, also called a linear rectifier. The formula is expressed as: $f(x) = max(x, 0)$.

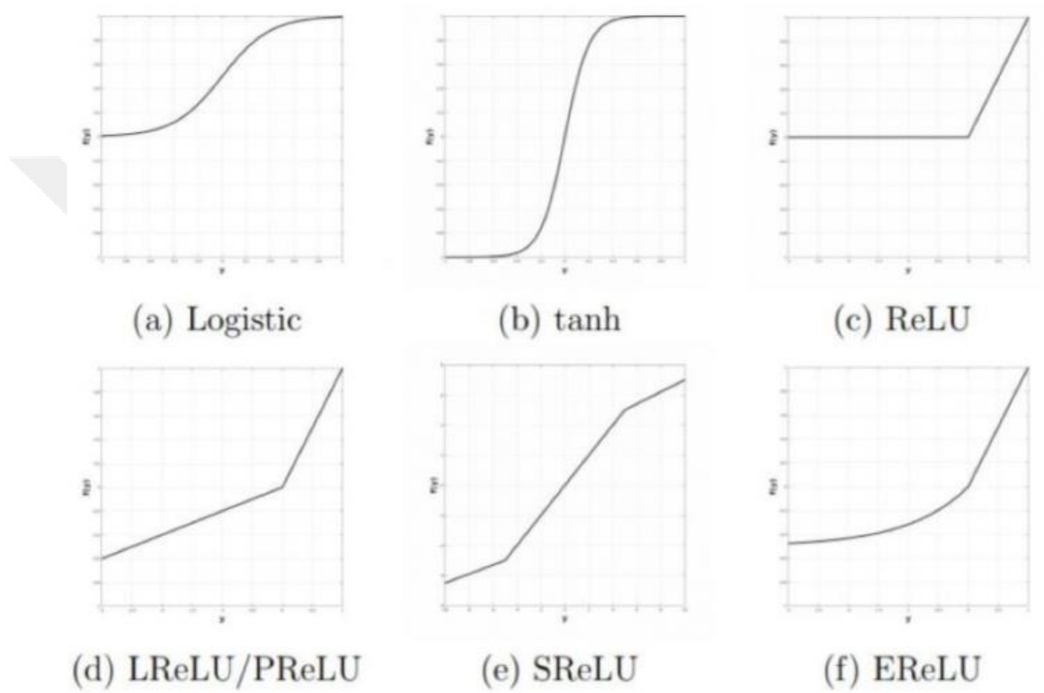

Figure 2.5. Activation functions

2.8. CNN architectures

2.8.1. AlexNet

In 2012, AlexNet proposed by Alex Krizhevsky and others won the ImageNet competition with a significant advantage [24]. This achievement has dramatically stimulated the industry's interest in neural networks and created a way to use deep neural networks to solve image problems. Later, more and more excellent results have emerged in this field. Compared with LeNet, AlexNet has a deeper network structure, including 5-layer convolution and 3-layer full connection. At the same time, the following three methods are used to improve the training process of the model:

Data augmentation: is a processing approach commonly used in deep learning. Indiscriminately add some changes to training to data, such as translation, zoom, crop, rotate, flip or increase or decrease brightness, etc., to generate a sequence of images which are similar but not identical to the original Sample, thereby enlarging the training dataset. Thus, the training data can be randomly changed to avoid the model from relying too much on certain features, and over-fitting can be suppressed to a certain extent.

- It uses dropout to avoid over-fitting in the model.
- It uses ReLU activation function to decrease gradient vanishing.

The specific structure of AlexNet is shown in Figure 2.6:

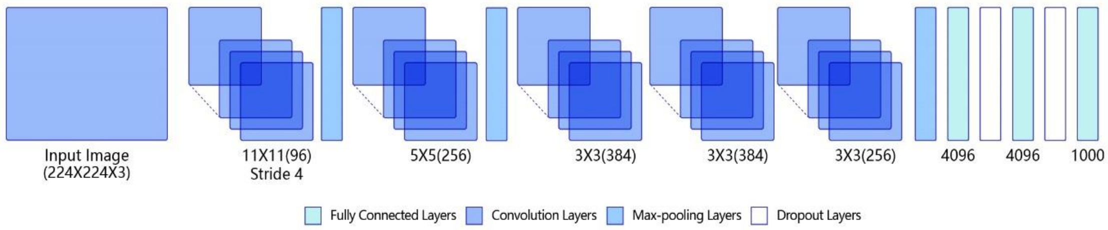

Figure 2.6. Alexnet architecture

There are four modules in Alexnet architecture:

- The first module contains a 96-channel convolution with 11 x 11 steps of 4 and a maximum pooling.
- The second module contains 5 x 5 steps of 256-channel convolution and a maximum pooling.
- The third module: contains two 3 x 3 384 channels and a 3 x 3 256 channel convolution, followed by a maximum pooling
- The fourth module: Contains two fully connected layers with 4096-channel input, each fully connected layer is followed by a Dropout layer to suppress over-fitting and the last 1000-channel fully connected layer.

2.8.2. GoogLeNet

GoogLeNet is the champion of the 2014 ImageNet competition. Its main feature is that the network not only has depth but also has "width" horizontally. Due to the massive difference in the spatial size of image information, it is more challenging to select the appropriate convolution kernel size to extract features. Image information with a more comprehensive spatial distribution range is suitable for extracting its features with a larger convolution kernel, and image information with a smaller spatial distribution range is suitable for extracting its features with a smaller convolution kernel. In order to solve this issue, GoogLeNet proposed a solution called Inception module. As shown in Figure 2.7:

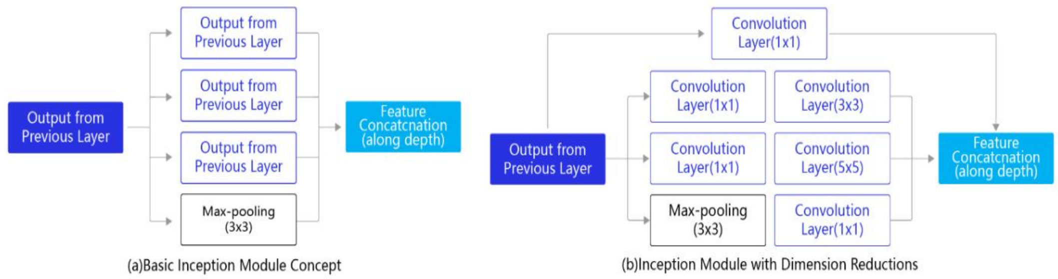

Figure 2.7. Inception module

Among them Figure 2.7(a) is the design idea of the Inception module, which uses 3 convolution kernels of different sizes to perform convolution operations on the input image, and adds maximum pooling. The outputs of these 4 operations are spliced along the channel dimension to form the output feature map of will contain the features extracted by convolution kernels of different sizes. The Inception module adopts a multi-path design. Each branch uses a different size convolution kernel. The number of channels in the final output feature map is the sum of the output channels of each branch, which will result in output channels the number becomes very large, especially when multiple Inception modules are used in series operation, the amount of model parameters will become very large. In order to reduce the amount of parameters, the Inception module uses the design method in Figure 2.7 (b). Before each 3x3 and 5x5 convolutional layer, add a 1x1 convolutional layer to control the number of output channels; add it after the maximum pooling layer The 1x1 convolutional layer reduces the number of output channels. Based on this design idea, the structure shown in the above Figure 2.7(b) is formed. The architecture of GoogLeNet is shown in the Figure 2.8 below. 5 blocks are used in the main convolution part, and a 3 × 3 maximum pooling layer with a stride of 2 is used between each block to reduce the output height and width.

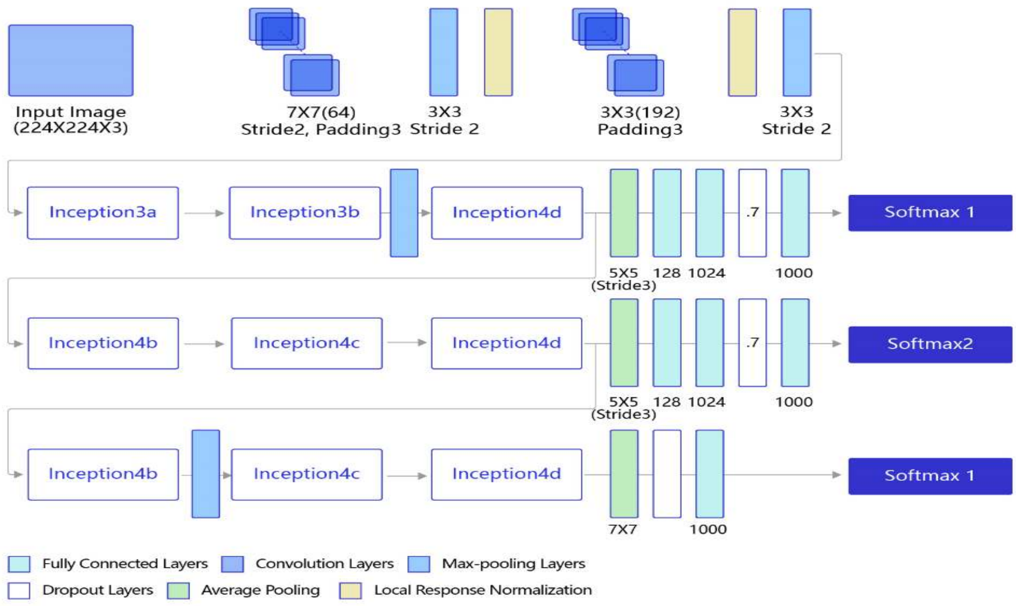

Figure 2.8. GoogLeNet architecture

GoogLeNet modules are listed as below:

- The first module uses a 64-channel 7 × 7 convolutional layer
- The second module uses 2 convolutional layers: first a 64-channel 1 × 1 convolutional layer, and then a 3 × 3 convolutional layer that increases the channel by 3 times.
- The third module connects 2 complete Inception blocks in series.
- The fourth module connects 5 Inception blocks in series.
- The fifth module connects 2 Inception blocks in series.
- The fifth module is followed by the output layer, which uses the global average pooling layer to change the height and width of each channel to 1. Finally, it connects to a fully connected layer whose output number is the number of label categories.

Moreover, the two auxiliary classifiers softmax1 and softmax2 shown in Figure 2.8 are added to the original author's study. As shown in Figure 2.8, the loss functions of the three classifiers are weighted and summed during training to alleviate the phenomenon of gradient disappearance. The procedure here is simplified, and no auxiliary classifier is added.

2.8.3. ResNet

ResNet was the champion of the ImageNet competition in 2015 and reduced the error rate of image classification recognition to 3.6%. This result even exceeded the accuracy of standard human eye recognition. Through the learning of the previous classic models, we can find that with the continuous expansion of deep learning, the number of layers of the model is increasing, and the network structure is becoming more and more complex. So if you deepen the network structure, you will get better results? In theory, assuming that the newly added layers are all identity mappings, as long as the original layer learns the same parameters as the original model, the deep model structure can achieve the effect of the original model structure. In other words, the solution of the original model is only a subspace of the solution of the new model, and a better result should be found in the solution space of the new model than the subspace corresponding to the solution of the original model. However, practice shows that after increasing the number of layers of the network, the training error tends to increase instead of decreasing.

Kaiming He et al [25]. proposed the residual network ResNet to solve the above problems. The basic idea is shown in Figure 2.9 below.

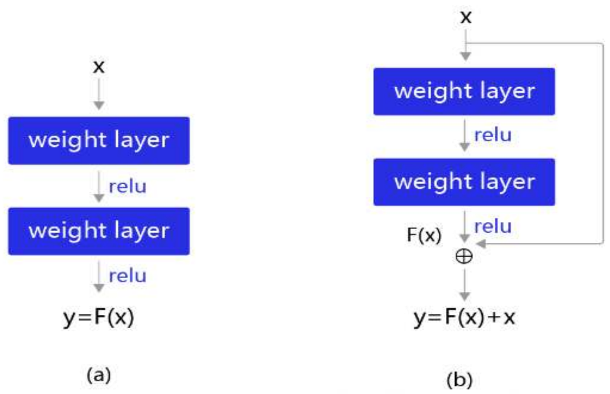

Figure 2.9. Residual block of Resnet (a) connection module (b) Bottleneck module

- Figure 2.9 (a): When adding network, map x to y=F(x) output.
- Figure 2.9 (b): (a) is improved, the output y=F(x) + x is not directly learning the representation of the output feature y, but learning y-x.

If you want to learn the representation of the original model, you only need to set all the parameters of F(x) to 0, then y=x is an individuality mapping. F(x) = y − x is also called the residual term. If the mapping of x→y is close to the identity mapping, it is easier to learn the residual term in (b) than (a) to learn the complete mapping form.

The structure of (b) is the basis of the residual network, and this structure is also called a residual block (residual block). The input x can propagate the data forward faster or propagate the gradient backward through the cross-layer connection.

The Figure 2.10 shows the structure of ResNet-50, which contains 49 layers of convolution and 1 layer of full connection, so it is called ResNet-50.

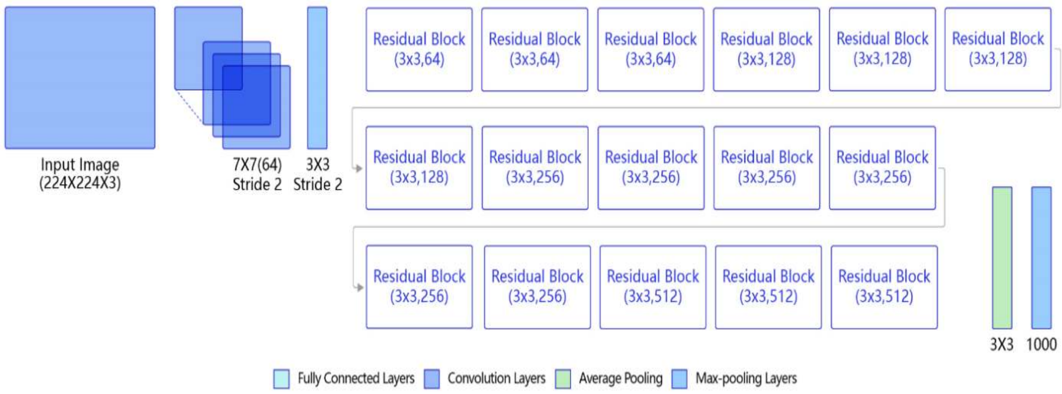

Figure 2.10. Resnet architecture

3. MATERIAL AND METHOD

3.1. Dataset

The data set [3] is produced by SPIE, American Association of Physicists in Medicine (AAPM), National Cancer Institute (NCI) and from Chicago University, University of Michigan, and Oak Ridge National Laboratory, launched in 2015 A challenge provided by LUNGx. In this research, data used were obtained from The Cancer Imaging Archive (TCIA) [26].the objective of this dataset is to identify the pulmonary nodules as benign or malignant. A dataset include of 83 CT scans from 70 different patients; there are 22,489 unique numbers of images in the whole dataset. From this 70 patients CT images, 10 CT images (subjects) are used for training and the other 60 CT images (subjects) are used for testing. All CT scans have 512 x 512 pixels size, and they are stored in Digital Imaging and Communication in Medicine (DICOM) format. Figure 3.1 shows two different classes of the dataset.

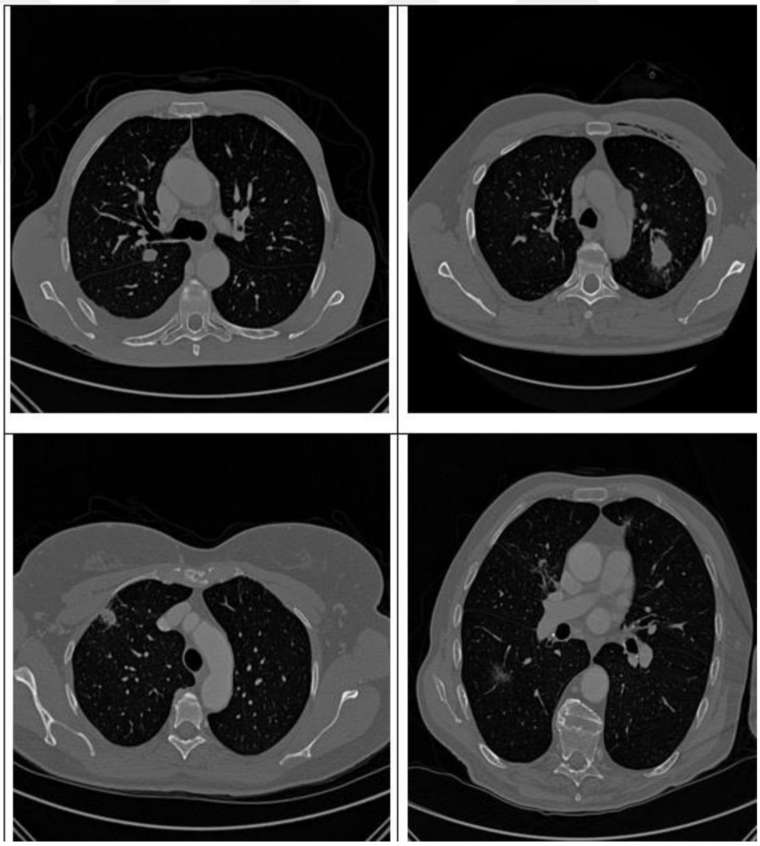

Figure 3.1. Lung cancer and non-cancer samples

3.2. Data Augmentation

Data enhancement [27]. Allows limited data to generate more data, increases the number and diversity of training samples (noisy data), and improves model robustness. Deep learning require a huge number of parameters, many of them have millions of parameters, and making these parameters work correctly requires plenty of samples for training. However, in many practical projects, it is difficult to find sufficient data to complete the task. Randomly changing the training samples can reduce the model's dependence on certain attributes, thereby improving the generalization ability of the model. For instance, we can crop the image in different methods so that the object appears in different positions of the image in different samples, which is also reducing the sensitivity of the model to the target position. For instance, we can also adjust factors such as brightness, contrast, saturation, and hue to decrease the model's sensitivity to color.

Data enhancement is generally used for images, which are some commonly used methods, such as random crop, random inversion, random contrast enhancement, color change. Generally speaking, random inversion and a small ratio of random resize, followed by random Crop are more commonly used. Image data enhancement is usually only for training data and less used for test data.

Data enhancement is a regularization technique that mainly operates on training data. As the name implies, data augmentation uses a series of methods to randomly change the training data, such as translation, rotation, shearing, and flipping. The detailed transformation amplitude of the data enhancement needs to be designed according to the specific application data. Just pay attention to one point: applying these simple transformations cannot change the label of the input image. Each image obtained through enhancement can be considered a new image. In this way, we can continuously provide new training samples to the model so that the model can learn more discriminative and generalized features.

The application of data enhancement technology can improve the accuracy of the model and help reduce over-fitting [27]. Besides, data enhancement can also increase the amount of data and reduce the large number of manually labeled data sets required for deep learning. Although the more "natural" training samples are collected, the better, but when it is impossible to increase the real training samples, data augmentation can be used to overcome the limitations of small data sets.

In this study, we have used some image processing operations such as the rotation technique to enhance the training data. Rotating has implemented by applying angles of 45, 90, 135, 180, 225, 270, 315 degrees, as shown in Figure 3.1 for two reasons[28]; we have chosen with rotation over other techniques. Firstly, doctors can examine histological images of Lung cancer from different angles without impacting the process of diagnosis. Therefore, utilizing data augmentation using rotation moreover improves the dataset. Secondly, the rotation technique has enlarged the dataset's size without impacting the quality of the input images [29].

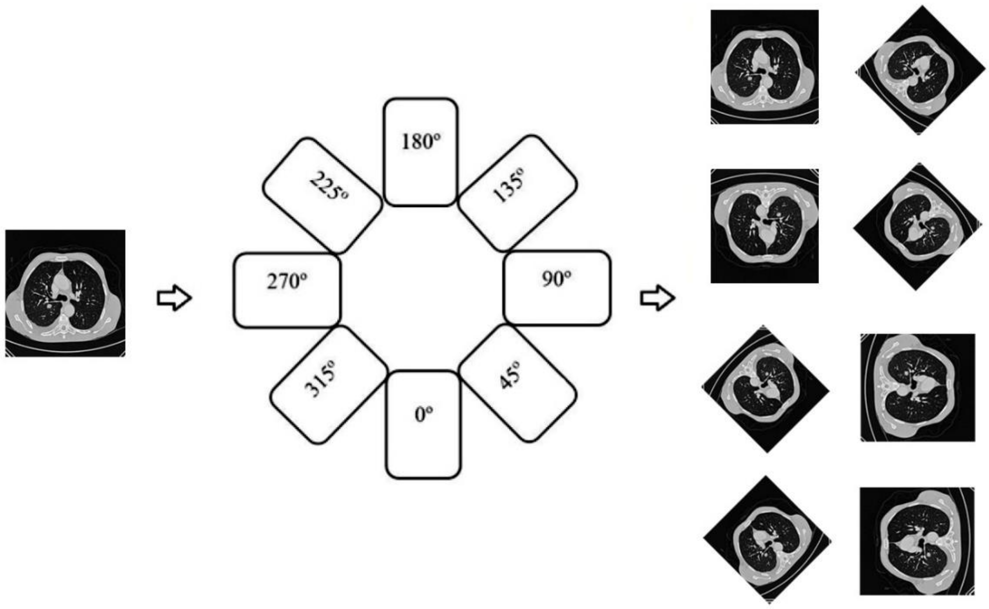

Figure 3.2. Rotation Process in different angels

3.3. Transfer learning

The massive victory of deep learning stems from the certainty that it requires a large number of labeled training data to obtain excellent learning performance. However, in current medical lung image analysis, this recommendation is hard to meet, because expert labeling is high-priced, and data on some diseases such as nodules or lesions are rare [30]. Therefore, in medical lung image analysis, how to use limited training samples to train deep models has become an open challenge. One of the most ordinary issues when using limited training data is that the model is prone to over-fitting. In order to solve the problem of model over-fitting, there is the primary way to choose is known as transfer learning.

Transfer learning, has been extensively adopted and has shown outstanding performance improvement capabilities, and it does not need extensive training data. This Method keeps away from expensive data annotation work in specific application areas. According to the research of S. J. Pan et al.[4] , transfer learning has classified into three types: inductive transfer learning, that is, irrespective of whether the target domain is the same as the destination domain, the target and destination tasks are different; direct transfer learning, that is, the target task it is the same as the destination task, but the target domain is different from the destination domain; and unsupervised transfer learning, that is, inductive transfer learning is similar, except that the target task is different from the destination task, but is related to the destination task.

The transfer learning approach requires two chief strategies: first, use the pre-trained network as a feature extractor that is, learn features from scratch; and second, Fine-tune the pre-trained network on medical images. This method is extensively used in current Medical lung image analysis. In some specific tasks, these two strategies have achieved good performance [11,15].

Transfer learning is a useful technique where we might use a pre-trained model like (AlexNet, ResNet18, Googlenet, and ResNet50) and adjust the network for the next application by making specific improvements in the architecture of the network [8]. AlexNet and the other pre-trained networks have trained for 1000 classes of real-world images. We can use this network to identify some other classes with some adjustment to the network (i.e.in our research, we used two classes Benign and Malignant, instead of 1000 classes; furthermore, we modified the last layer and fully connected layer to adjust for our work.

The mechanism of using any pre-trained models will be described in Figure 3.3. The standard pre-trained models like (AlexNet, ResNet18, Googlenet, and ResNet5) have proposed for well-organized classification. In this study, all models have altered the last three layers to adjust the new image classification.

Some other strategies also need attention in the next section, such as data augmentation/ data enhancement.

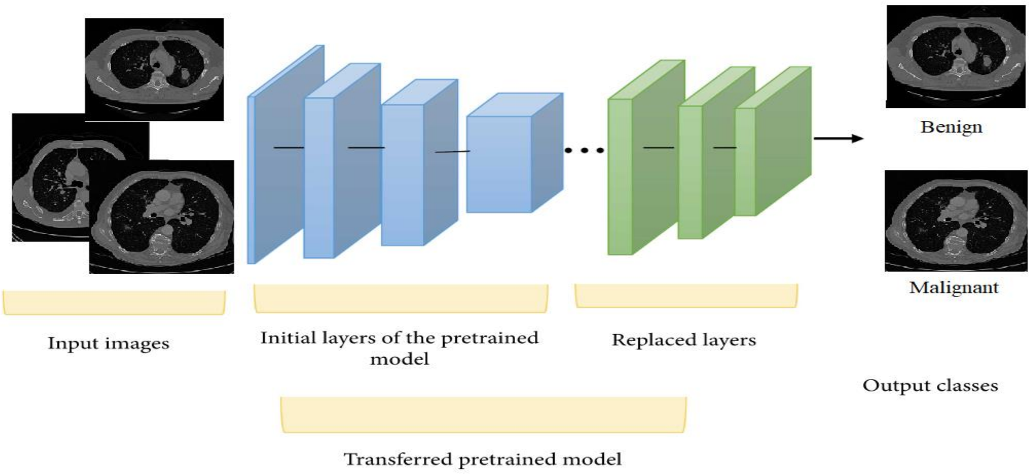

Figure 3.3. Transfer Learning Process using pre-trained models.

3.4. Evaluation metrics

When the training of a network is completed it needs to be analyzed with respect to performance. Several evaluation metrics can be utilized to analyze the performance, such as accuracy, sensitivity (recall), specificity, confusion matrix and F1-score.

The accuracy [31] is a measure of the ability of a network to predict the inputs. It is defined by Equation (3.1), where all the correct predictions are summed and divided by the total number of predictions. Depending on the value of the prediction, it is rounded to zero or one and compared to the true label. If the prediction and the label are equal, then it is a correct prediction. Typically, the average accuracy is plotted as a function of epochs, and the plot provides an overview of the average accuracy during the training.

$$\text{Accuracy} = \text{TP} + \text{TN} / \text{TP} + \text{FP} + \text{FN} + \text{TN} \tag{3.1}$$

The sensitivity (recall) is a measure of how good the network is at predicting a tumor, and is the probability that a prediction will be positive when a tumor is present. It is also called the true positive rate (TPR) it is defined by Equation (3.2).

$$\text{Sensitivity} = \text{TP} / \text{TP} + \text{FN} \tag{3.2}$$

Precision is also known as positive predictive value (PPV), which is the ratio of true positives among individuals identified as positive. Negative predictive value (NPV) is the ratio of true negatives among individuals who are identified as negative. It is defined by Equation (3.3).

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP}) \tag{3.3}$$

The specificity is a measure of how good the network is at classifying non-tumors, and is the probability that a prediction will be negative when a tumor is not present. The specificity is also called the true negative rate. The specificity is defined by Equation (3.4).

$$\text{Specificity} = \text{TN} / \text{TN} + \text{FP} \tag{3.4}$$

Where TP is true positive, the number of correct classifications of the tumors. The TN is true negative, the number of correct classification of the non-tumor. The FP is false positive, the number of the tumors that have been classified as non-tumors. The FN is false negative, the number of the non-tumors that have been classified as tumors [5]. Normally, TP, TN, FP and FN are used to obtain a confusion matrix, defined by Equation (3.5). The higher values on the diagonal are the better classifier.

$$\text{Confusion matrix} = \begin{pmatrix} \mathrm{TN} & \mathrm{FP} \\ \mathrm{FN} & \mathrm{TP} \end{pmatrix} \tag{3.5}$$

The F1-Score indicator combines the output results of Precision and Recall. The value of F1-Score ranges from 0 to 1. 1 constitutes the best output of the model, and 0 represents the worst output of the model. It is defined by equation (3.6).

$$\mathrm{F1-score} = 2 * (\mathrm{Precision} * \mathrm{Recall}) / (\mathrm{Precision} + \mathrm{Recall}) \tag{3.6}$$

4. RESULT AND DISCUSSION

4.1. Experimental result

4.1.1. Convolutional Neural Network settings

CNN is a feed-forward network; that is, the flow of information from input to output is one-way. Just as the Artificial Neural Network (ANN) is inspired by biology, so is CNN. There are many variants of CNN architecture. Usually, they are composed of convolutional layers and subsampling layers, which are grouped into modules. In a feed-forward network, one or more fully connected layers are finally connected. These modules are usually stacked together to form a deep model. The illustration of CNN architecture is in Figure 4.1.

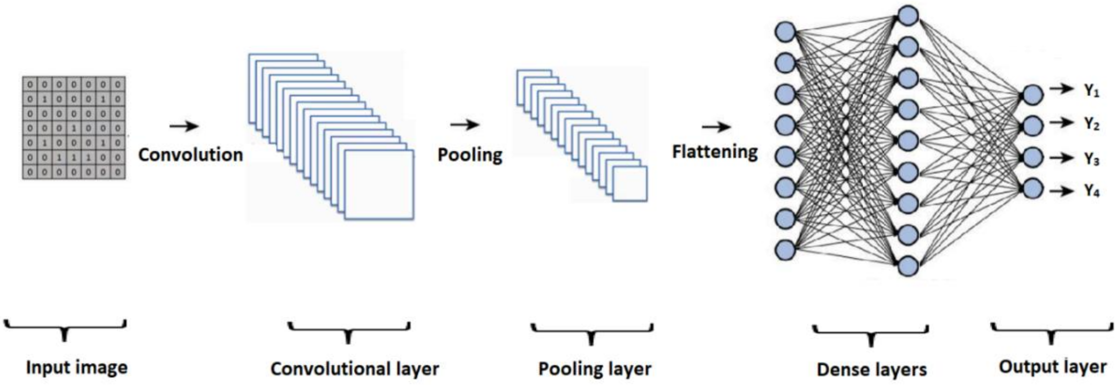

Figure 4.1. Representation of CNN architecture

The image is directly input into the CNN network, followed by several stages of convolution and pooling, and then through the feature representation of one or more fully connected layers, and finally, through a fully connected layer to output class labels. In this thesis, we utilized CNN models like (AlexNet, ResNet18, Googlenet, and ResNet50), and the modified models have altered the last three layers to adjust the new image classification:

- **AlexNet**: the last three layers of the constructed network with a group of layers are changed fully connected layer (FC), softmax layer, and output layer (classification) to classify images into relevant classes.
- **ResNet18:** The network layers (fc1000, prob, and Classification Layer_predictions) are replaced with fully connected layer, softmax layer, and classification output layer. Afterward, the last remaining transferred layer on the network (pool5) is linked to the novel layers.
- **GoogleNet**: also, the last three layers of the network are modified. The layers loss3-classifier, prob, and classification output layer are adjusted with a fully connected layer,

softmax layer, and an output layer. Later, the last transferred layer still existing on the network (pool5_drop7x7_s1) is connected to the new layers.

- **ResNet50:** The network's (fc1000, fc1000_softmax, and Classification Layer_fc1000) layers are replaced with fully connected layer, softmax layer, and classification output layer. Afterwards, the last remaining transferred layer on the network (avg_pool) is linked to the novel layers.

Therefore, each pre-trained networks have their characteristics. Besides, networks have depth, parameters in a million and input image size. Table 4.1 demonstrates the main properties of the models.

Table 4.1. Main characteristics of each network

Network	Depth	Parameters (Millions)	Image Input Size
AlexNet	8	60	227 × 227
ResNet 18	18	11.7	224 × 224
GoogleNet	22	7	224 × 224
ResNet 50	50	25.6	224 × 224

4.1.2. Hardware and software configuration

Hardware configuration

In this study, all experiments result was using a PC with i7 processor, 8 GB memory, and Windows 10 64-bits. In addition, MATLAB R2019b software was used for performance evaluation of the network's architecture. Moreover, deep learning Toolbox has utilized, which allows a framework to design and implement CNNs, where applications and graphics help to visualize network activations and monitor the progress of network training. All details are shown in Table 4.2.

Table 4.2. Hardware identifications

Hardware	Characteristics
Memory	8 GB
Processor	Intel Core i7-7600 CPU @ 2.80 GHz
Graphics	Intel® HD Graphics 620, 4 GB
Operating system	Windows 10, 64 bits

Software configuration

The study was experienced in SPIE-AAPM dataset, this dataset consist of 22,489 CT scans. The dataset was divided in to two classes, malignant and benign. 70% of this data was used for training and 30% was used for testing. There are several hyper-parameters used in training option in Matlab environment. In the field of deep learning, the choice of optimization algorithm is also the top priority of a model. One of the simple, effective algorithms that used in deep leaning is Stochastic Gradient Descent (SGD).

In the SGD algorithm [32] in each iteration, gradient descent updates the independent variable along the gradient of the current position according to the current position of the independent variable. Since the number of samples selected each time is relatively small, the loss will decrease in an oscillating manner.

In order to train the deep neural network more effectively, based on the standard small batch gradient descent method, some improved methods are often used to speed up the optimization. Common improvement methods are mainly improved from the following two aspects: learning rate attenuation and gradient direction optimization, where learning rate attenuation can be understood as adjusting the step size of parameter update, and gradient direction optimization can be understood as the direction of update make adjustments. These improved optimization methods can also be applied to batch or stochastic gradient descent methods [33]. In addition, learning rate has used is 0.0001, maximum number of epochs was 20 epochs, ValidationFrequency was 3 and we make a shuffle on dataset in every epoch. Finally, all the CNN's networks are trained by the mini batch size of 64. The Table 4.3 is showed the details.

Table 4.3. CNN network's training option hyper-parameters and values

Training option Parameters	**Value**
Optimization algorithm	SGDM
Learning rate	0.0001
Epochs	20
Shuffle	Every epoch
Batch size	64
ValidationFrequency	3

4.2. Results

In this thesis, an evaluation of state-of-the-art pre-trained networks for the task of classification of lung cancer infection using CT images was done. The objective of this research was to compare the CNN models evaluating the accuracy, precision, sensitivity, specificity and F1-Score by fine-tuning. The results are shown in Table 4.4.

Table 4.4 Pre-trained networks' results

Performance Measures	**Alexnet**	**Resnet18**	**Googlenet**	**Resnet50**
Accuracy	**98.52**	97.97	94.10	97.05
Recall	99.63	98.89	94.10	98.52
Precision	97.47	97.10	94.10	95.70
Specificity	97.41	97.04	94.09	95.57
F1-score	98.54	97.99	94.10	97.09
Time (min.)	67	192	209	525

All the models demonstrated similar and statistically significant performance. Starting with the accuracy matrix, Alexnet obtained a higher result with 98.52 %, followed closely by Resnet18, Goolglenet, and Resnet50 with 97.97 %, 94.10 %, and 97.05 %, respectively. On the other hand, the precision matrix, the higher percentage, is obtained by Alexnet with 97.47 %, and the lowest one is by Googlenet with 94.10 %, while the result of Resnet18 is 97.10 % and the result of Resnet50 is 95.70 %. Finally, in the f1-score, specificity, and sensitivity measurements, the performance of Googlenet was low, with 94.10 %, 94.09 %, and 94.10 %, respectively. In contrast to Alexnet that obtained the most acceptable percentage in the evaluation matrix with 98.54 %, 97.41 % and 99.63 %, respectively. Moreover, according to the processing time for every CNN architectures to perform the classification methods, Alexnet presented the best performance in taking the shortest time with 67 minutes, followed closely by Resnet18 with 192 minutes, while the extended classification of processing time was for Googlenet and Resnet50 with 209 and 525 minutes respectively. All the above discussions are mentioned in Table 4.4.

Additionally, the confusion matrix of the pre-trained models with the best consequence based on the performance measures, which are Alexnet, Resnet18, Googlenet and Resnet50. Depending on the results, it is possible to visually evaluate the classifier's performance and determine which classes are spotlighted by the neurons of the pre-trained models. On the confusion matrix, the rows are present with the output class, while the columns are present to the actual class.

As identified in the confusion matrix, considering in Figure 4.2, the Alexnet network correctly classified 264 of 271 standard images. However, it approximated 7 standard images as lung cancer furthermore, while predicting 270 of 271 lung cancer images correctly and predicted one lung cancer images incorrectly.

Figure 4.2. Confusion matrix of Alexnet network

As seen in Figure 4.3 the Alexnet network trained, it obtained the accuracy, learning curve and loss function. While the trained network with Alexnet, it obtained 98.52% of accuracy rate.

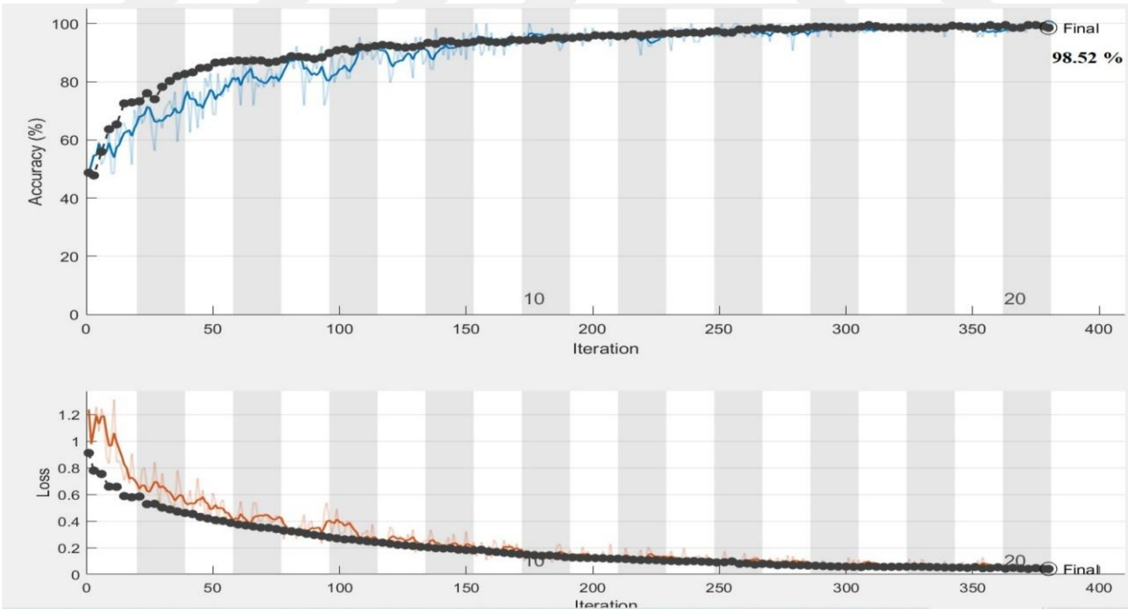

Figure 4.3. Learning curve and loss function of Alexnet network

As identified in the confusion matrix, considering in Figure 4.4, the Resnet18 network correctly classified 263 of 271 standard images. However, it approximated 8 standard images as lung cancer furthermore, while predicting 268 of 271 lung cancer images correctly and predicted 3 lung cancer images incorrectly.

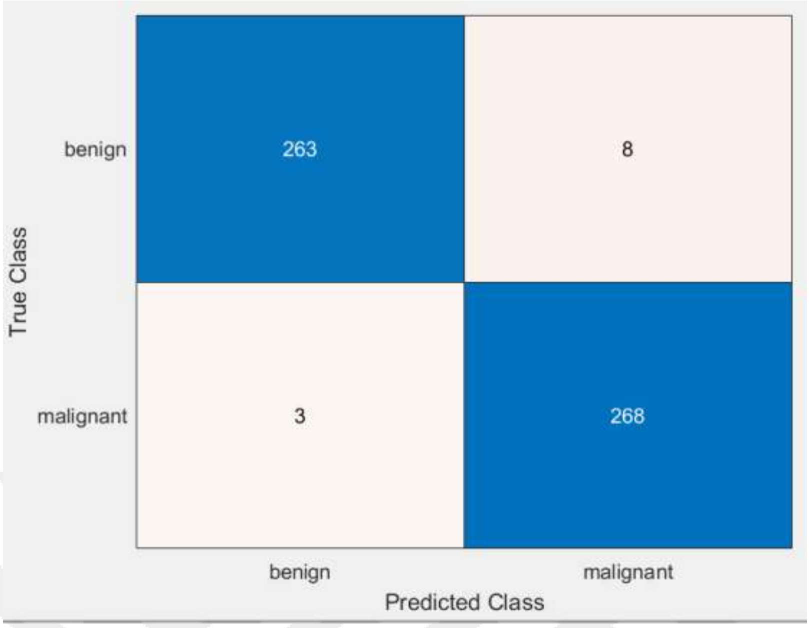

Figure 4.4. Confusion matrix of Resnet18 network

As seen in Figure 4.5 the Resnet18 network trained, it obtained the accuracy, learning curve and loss function. While the trained network with Resnet18, it obtained 97.97% of the accuracy rate.

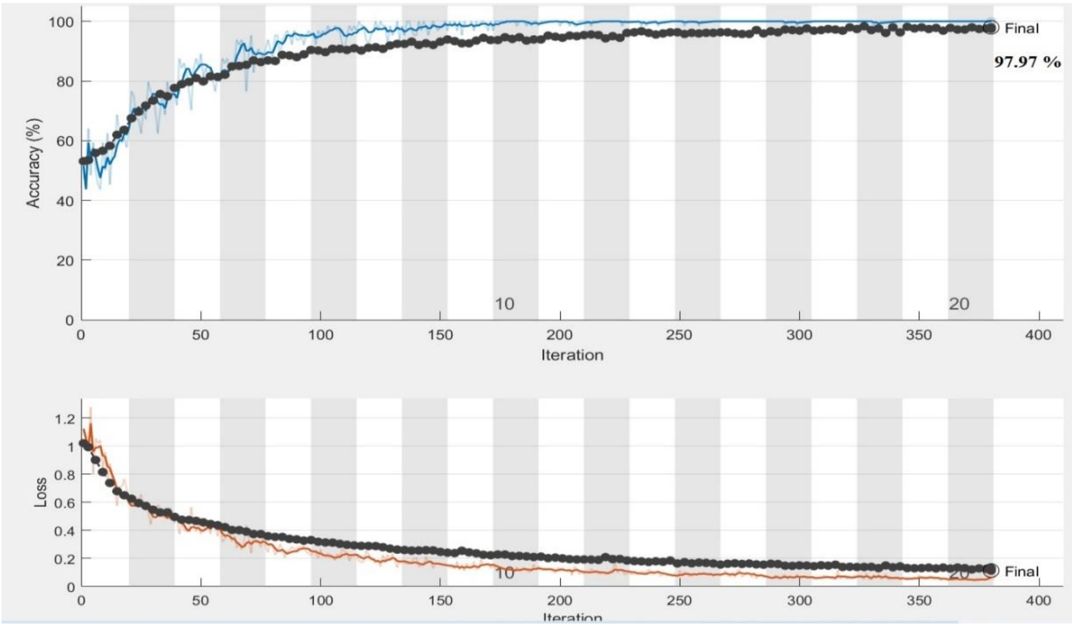

Figure 4.5. Learning curve and loss function of Resnet18 network

As identified in the confusion matrix, considering in Figure 4.6, the Googlenet network correctly classified 255 of 271 standard images. However, it approximated 16 standard images as lung cancer furthermore, while predicting 255 of 271 lung cancer images correctly and mispredicted 16 lung cancer images.

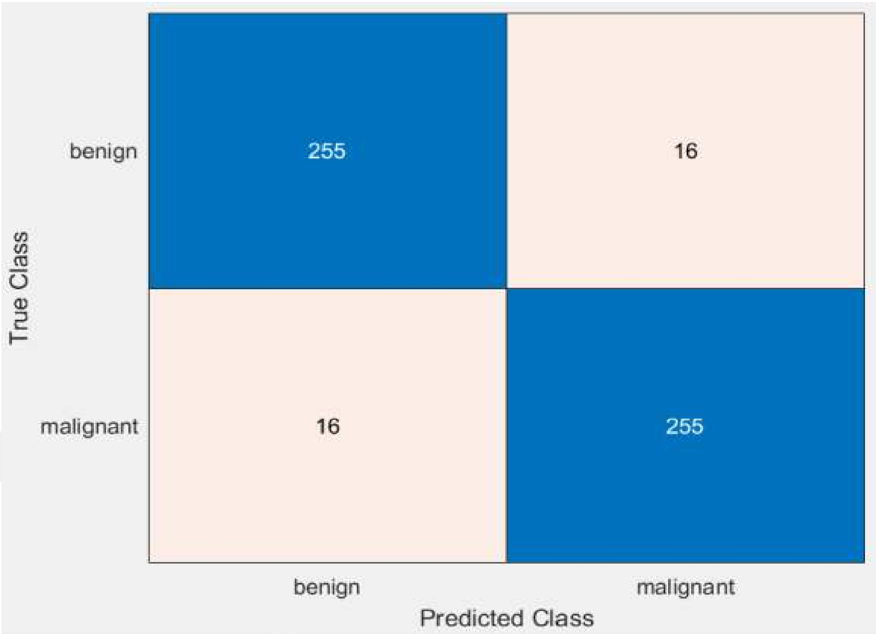

Figure 4.6. Confusion matrix of Googlenet network

As seen in Figure 4.7 the Googlenet network trained, it obtained the accuracy, learning curve and loss function. While the trained network with Googlenet, it obtained 94.10% of accuracy rate.

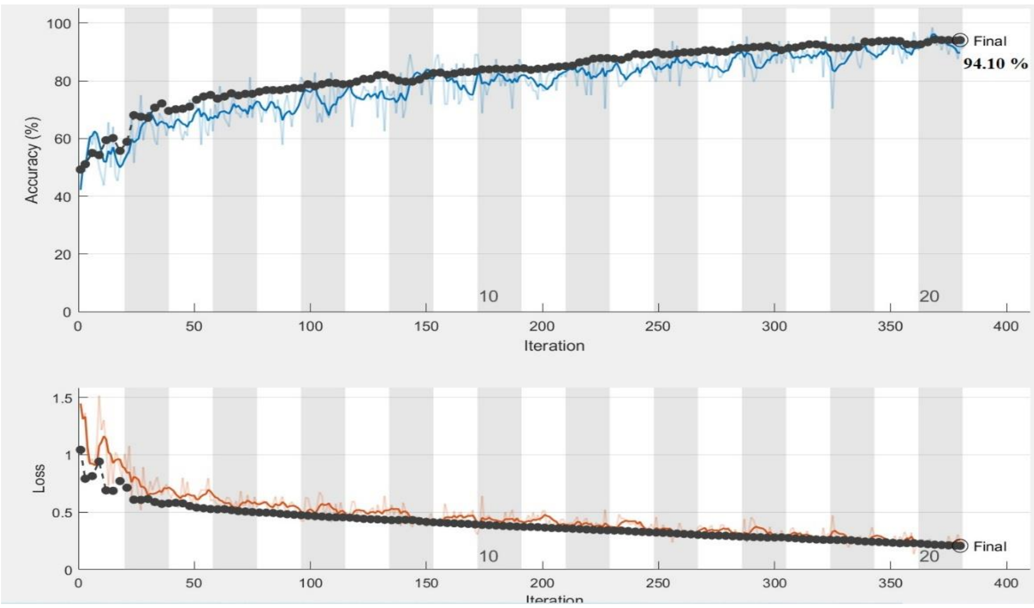

Figure 4.7. Learning curve and loss function of Googlenet network

As identified in the confusion matrix, considering in Figure 4.8, the Resnet50 network correctly classified 259 of 271 standard images. However, it approximated 12 standard images as lung cancer furthermore, while predicting 267 of 271 lung cancer images correctly and mispredicted 4 lung cancer images.

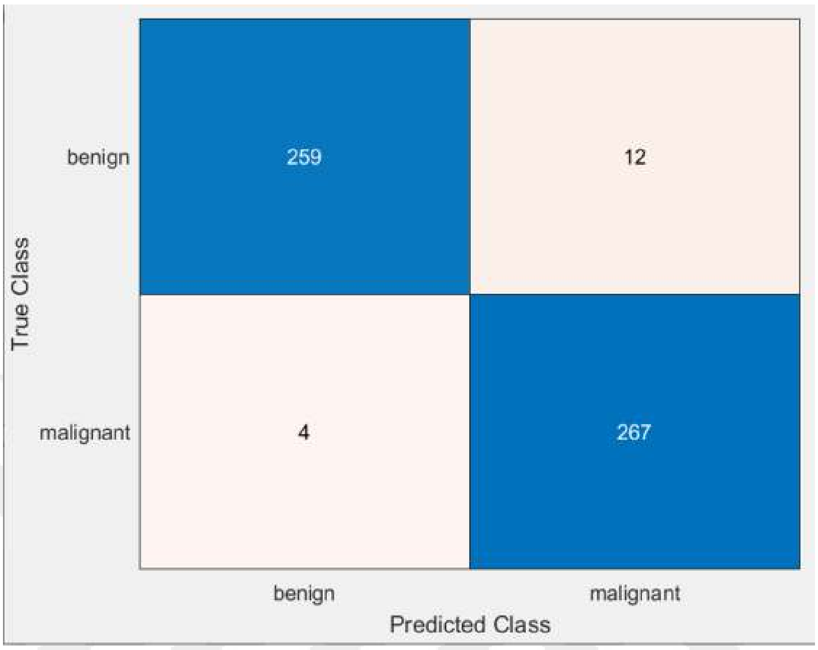

Figure 4.8. Confusion matrix of Resnet50 network

As seen in Figure 4.9, the Resnet50 network trained, it obtained the accuracy, learning curve and loss function. While the trained network with Resnet50, it obtained 97.05% of accuracy rate.

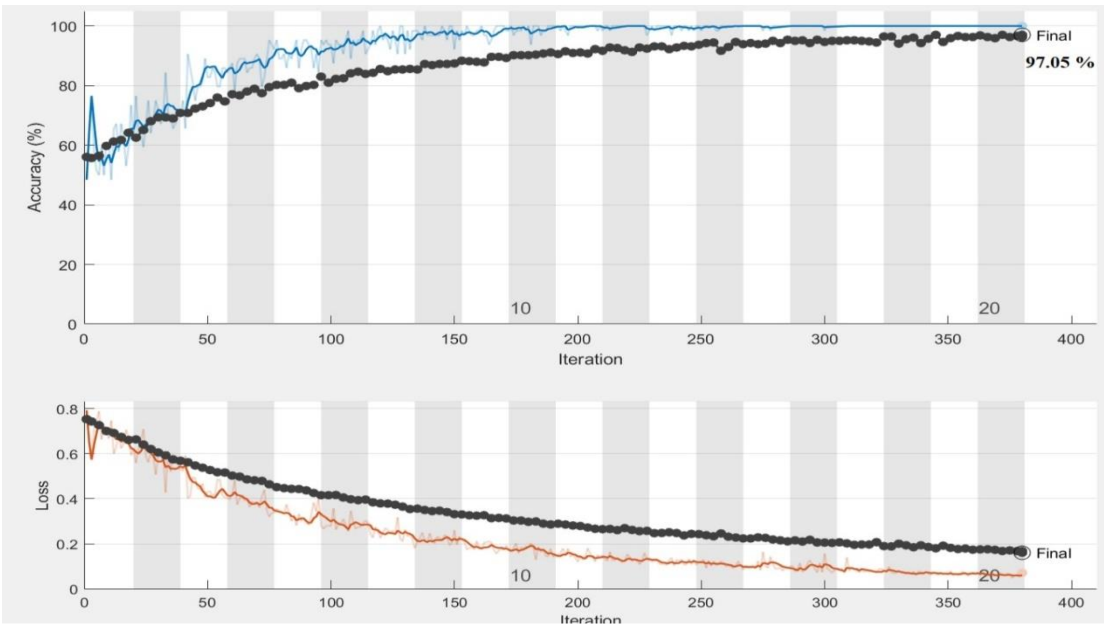

Figure 4.9. Learning curve and loss function of Resnet50 network

4.3. Comparison of the results with existing methods

In this section, we compare our pre-trained models' performance with some other existing works, which are mentioned in the literature review. Table 4.5 is demonstrating the details comparisons.

Table 4.5. Pre-trained networks compared with existing approaches

Dataset information				Existing Method (other methods)			Our Proposed Method			
S. No	No. of samples	Training	Testing	Authors	Methods Used	Results	AlexNet	ResNet18	GoogleNet	ResNet50
1	1006	90%	10%	H. Jiang et al. [14]	Filter, Convolution Neural Networks	Sensitivity 94%	Sensitivity 99.63 %	Sensitivity 98.89 %	Sensitivity 94.10 %	Sensitivity 98.52 %
2	1536	80%	20%	Raul V. et al. [15]	CNN-ResNet50 with SVM-RBF	accuracy 88.41 %	Accuracy 98.52 %	Accuracy 97.97 %	Accuracy 94.10 %	Accuracy 97.05 %
3	1006	80%	20%	T. K. Sajja et al. [13]	AlexNet GoogleNet ResNet50 and densely connected architecture	Accuracy 100 %, 98.84 %, 100 %, and 100%	Accuracy 98.52 %	Accuracy 97.97 %	Accuracy 94.10 %	Accuracy 97.05 %

According to Table 4.5, the existing methods and our pre-trained networks are compared. H. Jiang et al. [14] nominated Frangi filter and Convolution Neural Networks approach which are tested on 1006 CT scan of LIDC dataset with 10% testing and 90% training they obtained 94% sensitivity. The pre-trained networks (Alexnet, Resnet18, Googlenet and Resnet50) are tested on 1804 CT images with 70% training and 30% testing they got (99.63%, 98.89%, 94.10% and 98.52%)sensitivity, respectively.

Raul V. et al.[15] Proposed CNN-ResNet50 with SVM-RBF methods; used 1536 CT samples with 20% testing and 80% training obtained 88.41% accuracy. Our pre-trained networks (Alexnet, Resnet18, Googlenet and Resnet50) obtained (98.52%, 97.97%, 94.10% and 97.05%) accuracy correspondingly.

T. K. Sajja et al.[13] Used AlexNet GoogleNet ResNet50 and densely connected architecture they tested on 80% training and 20% testing dataset samples they got 100 %, 98.84 %, 100 %, and 100 % accuracy respectively.

4.4. Discussion

Our research SPIE-AAPM dataset has been used; the dataset is a part of the medical imaging conference at SPIE in 2015. It Includes 22,489 pulmonary CT images with the real data that is the position of the tumor and other details of which patient has a cancer tumor and which patient has a non-cancer tumor. Deep learning needs a massive amount of data to correctly work on the training dataset so that Data augmentation has been used to increase the number of labeled data artificially. Transfer learning can decrease the required training data and reduce the training time for deep learning.

In this phase, the dataset used are presents two classes of Lung infection; the dataset was divided into 70% for training and 30% for testing. Therefore, four pre-trained models, namely AlexNet, ResNet18, GoogleNet and ResNet50 were trained by the concept of fine-tuning, which consists of replacing the last three layers, where the final output layer has to be well-matched with the number of classes.

These pre-trained models are evaluated by different metrics such as accuracy, precision, recall, and F-Score. As shown in Table 4.4, which offers a general result performance, AlexNet achieved better results than the other architectures. Besides, Googlenet, shows low performance.

An overview of pre-trained networks (AlexNet, ResNet18, Googlenet and ResNet50) for the role of classifying Lung cancer infections using CT images was done. The purpose of this study is to compare the Convolutional Neural Network evaluating the accuracy, recall, precision, and f1-score by fine-tuning.

CONCLUSION

As we all know, Deep learning the training of the model relies on a large amount of data. Therefore, the two techniques used in our study first technique is Transfer Learning. The second is a data augmentation technique to tackle the lack of training data and over-fitting issues and not build a network from scratch. Through using a dataset with fine-tuning models (AlexNet, Resnet18, Googlenet, and Resnet50), it could classify lung cancer data, i.e., Classification of Benign and Malignant. Our model's learning in 20 Epoch has obtained a high accuracy and very fit learning curve with suitable training time. By applying these techniques and obtaining a high performance of the model evaluation, we early confirmed lung cancer detection to protect patients from death. In the future, we would compare our model to some other models and improved the data with the proposed model.

REFERENCES

[1] World Health Organization, WHO's Urban Ambient Air Pollution database - Update 2016, Who. (2016) 1–7. http://www.who.int/phe/health_ topics/outdoorair/ databases/AAP_ database_ summary_ results_2016_v02.pdf?ua=1.

[2] M.J. Willemink, W.A. Koszek, C. Hardell, J. Wu, D. Fleischmann, H. Harvey, L.R. Folio, R.M. Summers, D.L. Rubin, M.P. Lungren, Preparing medical imaging data for machine learning, Radiology. 295 (2020) 4–15. https://doi.org/10.1148/radiol.2020192224.

[3] L.P. Armato III, Samuel G.; Hadjiiski, Lubomir; Tourassi, Georgia D.; Drukker, Karen; Giger, Maryellen L.; Li, Feng; Redmond, George; Farahani, Keyvan; Kirby, Justin S.; Clarke, SPIE-AAPM-NCI Lung Nodule Classification Challenge Dataset, (2015). https://doi.org/10.7937/K9/ TCIA.2015.UZLSU3FL.

[4] S.J. Pan, Q. Yang, A survey on transfer learning, IEEE Trans. Knowl. Data Eng. 22 (2010) 1345–1359. https://doi.org/10.1109/TKDE.2009.191.

[5] G.E.H. Alex Krizhevsky, Ilya Sutskever, Handbook of approximation algorithms and metaheuristics, Handb. Approx. Algorithms Metaheuristics. (2007) 1–1432. https://doi.org/ 10.1201/ 9781420010749.

[6] V. Sangeetha, K.J.R. Prasad, Syntheses of novel derivatives of 2-acetylfuro[2,3-a]carbazoles, benzo[1,2-b]-1,4-thiazepino[2,3-a]carbazoles and 1-acetyloxycarbazole-2- carbaldehydes, Indian J. Chem. - Sect. B Org. Med. Chem. 45 (2006) 1951–1954. https://doi.org/10.1002/chin.200650130.

[7] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. 07-12-June (2015) 1–9. https://doi.org/10.1109/CVPR.2015.7298594.

[8] H.C. Shin, H.R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, R.M. Summers, Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning, IEEE Trans. Med. Imaging. 35 (2016) 1285–1298. https://doi.org/10.1109/TMI.2016.2528162.

[9] G. Litjens, T. Kooi, B.E. Bejnordi, A.A.A. Setio, F. Ciompi, M. Ghafoorian, J.A.W.M. van der Laak, B. van Ginneken, C.I. Sánchez, A survey on deep learning in medical image analysis, Med. Image Anal. 42 (2017) 60–88. https://doi.org/10.1016/j.media.2017.07.005.

[10] E. Cengil, A. Çinar, A Deep Learning Based Approach to Lung Cancer Identification, 2018 Int. Conf. Artif. Intell. Data Process. IDAP 2018. (2019). https://doi.org/10.1109/IDAP.2018.8620723.

[11] Z. Shi, H. Hao, M. Zhao, Y. Feng, L. He, Y. Wang, K. Suzuki, A deep CNN based transfer learning method for false positive reduction, Multimed. Tools Appl. 78 (2019) 1017–1033. https://doi.org/10.1007/s11042-018-6082-6.

[12] B. Almas, K. Sathesh, S. Rajasekaran, A Deep Analysis of Google Net and Alex Net for Lung Cancer Detection, Int. J. Eng. Adv. Technol. 9 (2019) 395–399. https://doi.org/10.35940 /ijeat.b3226.129219.

[13] T.K. Sajja, R.M. Devarapalli, H.K. Kalluri, Lung cancer detection based on CT scan images by using deep transfer learning, Trait. Du Signal. 36 (2019) 339–344. https://doi.org/ 10.18280/ts. 360406.

[14] H. Jiang, H. Ma, W. Qian, M. Gao, Y. Li, An Automatic Detection System of Lung Nodule Based on Multigroup Patch-Based Deep Learning Network, IEEE J. Biomed. Heal. Informatics. 22 (2018) 1227–1237. https://doi.org/10.1109/JBHI.2017.2725903.

[15] R.V.M. Da Nóbrega, S.A. Peixoto, S.P.P. Da Silva, P.P.R. Filho, Lung Nodule Classification via Deep Transfer Learning in CT Lung Images, Proc. - IEEE Symp. Comput. Med. Syst. 2018-June (2018) 244–249. https://doi.org/10.1109/CBMS.2018.00050.

[16] R. Anirudh, J.J. Thiagarajan, T. Bremer, H. Kim, Lung nodule detection using 3D convolutional neural networks trained on weakly labeled data, Med. Imaging 2016 Comput. Diagnosis. 9785 (2016) 978532. https://doi.org/10.1117/12.2214876.

[17] R.M. Devarapalli, H.K. Kalluri, V. Dondeti, Lung cancer detection of ct lung images, Int. J. Recent Technol. Eng. 7 (2019) 413–416.

[18] Cancer Council, Understanding Lung Cancer, Cancer Counc. Aust. (2016) 9–11. http://www. cancervic.org.au/downloads/resources/booklets/Understanding-Lung-Cancer.pdf.

[19] M.J. Thun, When truth is unwelcome: the first reports on smoking and lung cancer., Bull. World Health Organ. 83 (2005) 144–145. https://doi.org/10.1590/S0042-96862005000200015.

[20] Y. Lecun, Y. Bengio, G. Hinton, Deep learning, Nature. 521 (2015) 436–444. https://doi.org /10.1038/nature14539.

[21] G. Larsson, M. Maire, G. Shakhnarovich, FractalNet: Ultra-Deep Neural Networks without Residuals, (2016) 1–11. http://arxiv.org/abs/1605.07648.

[22] N.S. Geoffrey, A.K.I. Sutskever, R. Salakhutdinov, Department, Dropout: A Simple Way to Prevent Neural Networks from Overfittin, 2016 Int. Conf. Adv. Electr. Electron. Syst. Eng. ICAEES 2016. 15 (2016) 520–525. https://doi.org/10.1109/ICAEES.2016.7888100.

[23] L. Zhao, M. Li, D. Meng, X. Li, Z. Zhang, Y. Zhuang, Z. Tu, J. Wang, On the Connection of Deep Fusion to Ensembling Liming, IJCAI Int. Jt. Conf. Artif. Intell. 2018-July (2018) 3170–3176. https://doi.org/10.24963/ijcai.2018/440.

[24] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A.C. Berg, L. Fei-Fei, ImageNet Large Scale Visual Recognition Challenge, Int. J. Comput. Vis. 115 (2015) 211–252. https://doi.org/10.1007/s11263-015-0816-y.

[25] S.X.Z.S.R.J. Kaiming He, Deep residual learning for image steganalysis, Multimed. Tools Appl. (2017) 1–17. https://doi.org/10.1007/s11042-017-4440-4.

[26] K. Clark, B. Vendt, K. Smith, J. Freymann, J. Kirby, P. Koppel, S. Moore, S. Phillips, D. Maffitt, M. Pringle, L. Tarbox, F. Prior, The cancer imaging archive (TCIA): Maintaining and operating a public information repository, J. Digit. Imaging. 26 (2013) 1045–1057. https://doi.org /10.1007/ s10278-013-9622-7.

[27] C. Shorten, T.M. Khoshgoftaar, A survey on Image Data Augmentation for Deep Learning, J. Big Data. 6 (2019). https://doi.org/10.1186/s40537-019-0197-0.

[28] L. Alzubaidi, O. Al-Shamma, M.A. Fadhel, L. Farhan, J. Zhang, Y. Duan, Optimizing the performance of breast cancer classification by employing the same domain transfer learning from hybrid deep convolutional neural network model, Electron. 9 (2020). https:// doi.org/ 10.3390/ electronics9030445.

[29] L. Perez, J. Wang, The Effectiveness of Data Augmentation in Image Classification using Deep Learning, (2017). http://arxiv.org/abs/1712.04621.

[30] N. Tajbakhsh, J.Y. Shin, S.R. Gurudu, R.T. Hurst, C.B. Kendall, M.B. Gotway, J. Liang, Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?, IEEE Trans. Med. Imaging. 35 (2016) 1299–1312. https://doi.org/10.1109/TMI.2016.2535302.

[31] H. M, S. M.N, A Review on Evaluation Metrics for Data Classification Evaluations, Int. J. Data Min. Knowl. Manag. Process. 5 (2015) 01–11. https://doi.org/10.5121/ijdkp.2015.5201.

[32] S. Ruder, An overview of gradient descent optimization algorithms, (2016) 1–14. http://arxiv. org/ abs/1609.04747.

[33] X. Cui, W. Zhang, Z. Tüske, M. Picheny, Evolutionary stochastic gradient descent for optimization of deep neural networks, Adv. Neural Inf. Process. Syst. 2018-Decem (2018) 6048–6058.

CURRICULUM VITAE

[redacted]

PERSONAL INFORMATION

RESEARCHER INFORMATION

[redacted]

EDUCATION

Master Degree **:** "Detection of Cancer Area in Lung Images With the Help of Deep Learning Algorithms"

Firat University, Graduate School of Natural and applied Sciences, Department of Computer Engineering, 2020

Supervisor: Assoc.Prof.Dr. Ahmet ÇINAR

Bachelor **:** Duhok University, Science Faculty, Department of Computer Science, 2012

High School **:** Rubad / Brayati, Duhok, 2007

RESEARCH EXPERIENCES

Laboratory Instruments you use, such as test systems etc. you know

Computer Programming languages available (Python, C++, MATLAB, vb .Net.)

Computer programs you can use (CAT/CAM, PROTEUS, CALCULUS, vb.)

WORK EXPERIENCE

2013 – till now: I have about 10 years' experience working with academic organization with more than 7 years working with the technology company

ACADEMIC ACTIVITIES

Paper: Under Publishing