ADDIS ABABA UNIVERSITY

ADDIS ABABA INSTITUTE OF TECHNOLOGY

SCHOOL OF ELECTRICAL AND COMPUTER ENGINEERING

Pulmonary Nodule Segmentation in Lung CT Images by Post Processing U-Nets Using Average Ensemble Learning Technique

by

Habtamu Yetwale Dememew

A thesis submitted to the School of Electrical and Computer Engineering in partial fulfillment of the requirements for the Degree of Master of Science in Computer Engineering

June 24, 2020

Addis Ababa, Ethiopia

ADDIS ABABA UNIVERSITY

ADDIS ABABA INSTITUTE OF TECHNOLOGY

SCHOOL OF ELECTRICAL AND COMPUTER ENGINEERING

By: Habtamu Yetwale Dememew

Advisor: Mr. Menore Tekeba

This is to certify that the thesis prepared by *Habtamu Yetwale*, titled *Pulmonary Nodule Segmentation in Lung CT Images by Post Processing U-Nets Using Average Ensemble Learning Technique* and Submitted in partial fulfillment of the requirements for the Degree of Master of Science in Computer Engineering compiles with the regulations of the University and meets the accepted standards with respect to originality and quality.

Approved and signed by board of Examining Committee:

	Name	Signature	Date
Dean, School of Electrical and Computer Engineering:	Dr. Yalemzewd Negash		09/07/2020
Advisor:	Mr. Menore Tekeba		14/07/2020
Internal Examiner:	Mr. Bisrat Derebssa		09/07/2020
External Examiner:	Dr. Surafel Lemma		14/07/2020

Declaration

I, the undersigned, declare that this thesis is my original work and has not been presented for a degree in any other university, and that all source of materials used for the thesis have been properly acknowledged.

Declared by:

Name	Signature	Date
Habtamu Yetwale Dememew		07/07/2020

Approved by my advisor:

Name	Signature	Date
Mr. Menore Tekeba		14/07/2020

Acknowledgments

I am here and be what I am by the help of the son of **St. Mariam** almighty **God**. They are always at the back and front on all my ways. የድንግል ማርያም ልጅ ልዑል እግዚአብሄር ፊጽሞ የተመሰገነ ይሁን::

A special and deep thanks goes to cooperative, positive, kind, and an open person, **Mr. Menore Tekeba**, who was my advisor. He was always in place whenever I need him.

I would also like to thank LUNA16 challenge and Kaggle competition organizers. For me, they were a place where I can found any issues solved.

Lastly, I want to convey my appreciation and love to all my family members (my mother, my father, my brothers, and my sister), they are mean a lot to me. Especially, I thank my mother **Berie** and my brother **Muler**; I know all your efforts spent to put me here.

Abstract

Pulmonary nodules are potential manifestation of lung cancer and accurate segmentation of different pulmonary nodules from lung Computed Tomography (CT) scan images is important clinical relevance in diagnosis, prognosis and treatment of lung cancer. However, due to the highly heterogeneous type, size, location and shape of nodules, segmentation of pulmonary nodules is very challenging. In this study we propose and present an improved pulmonary nodule segmentation method based on thresholding and morphological operators, lungs region segmentation algorithm, modified U-Net model for pulmonary nodules detection and segmentation, and average ensemble learning as a post processing technique. First, as a pre-processing stage that includes lungs region segmentation, normalization, cropping and resizing has done on the raw input CT scan images. Thresholding and morphological operators algorithm is simple and yields good result in accurately segmenting the lung parenchyma to reduce the search space. Then modified U-Net model is loaded to detect and segment pulmonary nodules. U-Net is a widely used Convolutional Neural Network (CNN) structure for end-to-end segmentation and can be used on the entire image classification labels over the entire input pixels so that it is more efficient and expected to yield better performance. Furthermore, instead of picking the best U-Net network structure, we applied average ensemble learning method as a post processing technique. An ensemble of three U-Net models having different network structure, trained on the same dataset with different hyper-parameters, can generally improve the overall segmentation performance. The performance of our proposed method is trained, tested, and evaluated using 858 lung CT images and their corresponding ground truth nodule masks obtained from Lung Nodule Analysis 2016 (LUNA16) dataset and achieved evaluation results of 0.848 mean Dice Similarity Coefficient (DSC), 0.965 mean accuracy, 0.826 mean sensitivity, and 0.983 mean specificity. Moreover, we compared our test results with other methods results to show our approach's performance. Experiments and these preliminary results showed that our proposed method can effectively improve the segmentation accuracy of pulmonary nodules and the effectiveness of our approach. Generally, the algorithms used here are simple, effective and practical.

Keywords: ***CT, LUNA16, thresholding and morphological operators, pulmonary nodules, U-Net, average ensemble learning, DSC***

Table of Contents

LIST OF TABLES

LIST OF FIGURES

List of Abbreviations

ACM: Active Contour Model

Adam: Adaptive Moment Estimation

ANN: Artificial Neural Network

BN: Batch Normalization

CAD: Computer Aided Diagnosis

CNN: Convolutional Neural Network

CT: Computed Tomography

DSC: Dice Similarity Coefficient

EM: Electron Microscopy

FCN: Fully Convolutional Network

fCNN: fully Convolutional Neural Network

GPU: Graphical Processing Unit

GT: Ground Truth

HU: Hounsfield Units

IDRI: Image Database Resource Initiative

IOU: Intersection Over Union

ISBI: International Symposium on Biomedical Imaging

LIDC: Lung Image Database Consortium

LUNA16: Lung Nodule Analysis 2016

MBM: Mean Binary Mask

MLP: Multi Layer Perceptron

ReLU: Rectified Linear Unit

ROI: Region Of Interest

SGD: Stochastic Gradient Descent

SVM: Support Vector Machine

TanH: Hyperbolic Tangent

VOI: Volume Of Interest

2D: 2 Dimensional

3D: 3 Dimensional

Chapter One

1. Introduction

1.1. Background

Globally, lung cancer is the major source of cancer based deaths [1]. One of the main causes for such a high rate of death is the fact that it is very difficult to detect malignant lung nodules. Usually pulmonary nodules are detected lately when they become too large, and/or too advanced and it is very hard to treat and cure them effectively [2]. Therefore, it is very essential to detect and segment pulmonary lung nodules early for lung screening so that lung cancer can be treated effectively in the stage where it can be cured.

The diagnosis of lung cancer at early stage, which is typically manifested in the form of pulmonary nodules, is an efficient way of improving the survival rate and it is critical for treatment planning [3]. Nowadays, even though modern imaging techniques such as CT scanning; that can accurately capture the images of lungs are available, identifying, finding and locating the cancerous pulmonary nodules is still a difficult task for medical persons.

CT is the newest type of medical imaging equipment with high density resolution, and adequate image information that can make a variety of processing on images [4]. CT has a significant value for early detection of pulmonary lung nodules and it has been widely applied in diagnosing and screening of lung cancer all over the world.

Physicians use different metrics such as size, intensity, shape, and texture as a way of differentiation between healthy and abnormal lung tissue anatomical structures. But, it is a challenging task due to a huge amount of data generated by CT scan. Therefore, the development of automatic CAD tools is needed to help physicians to analyze and evaluate the CT scan images more accurately.

Pulmonary nodules that appear as round or oval opacity in lung CT images are early indications and common causes of lung cancer [3, 5]. In general, segmenting pulmonary nodules that differs

in type, size, shape and location is the first necessary step in the development of an automatic CAD system.

Currently, deep learning algorithms specifically, convolutional networks, are rapidly becoming a practice of choice for evaluating medical images. CNNs are a type of deep artificial neural networks inspired by the biological functions of neurons and widely used in the areas of computer vision. CNNs have been shown their ability in learning hierarchically organized low to high-level features directly from raw images, and yield state-of-the-art performance in image classification [6], super resolution, object detection [7] and semantic segmentation tasks [8]. However, very small has been done in using CNN approaches to segment pulmonary nodules from lung CT images, primarily due to the limited size of labeled datasets.

This work demonstrates a CNN-based medical image segmentation method on lung CT images, with the task of pulmonary nodule segmentation. For our task, a network architecture which is same as a U-Net model architecture developed by Ronneberger et al. [9] is used. The two major architectural qualities in U-Net are the arrangement of an equal amount of up-sampling and down-sampling layers and it integrates them with the so called skip connections between opposing convolution and deconvolution layers which associate features from the contracting and expanding paths. From a training perspective of U-Net model architecture, this means that, the whole image can be processed by the model within one single forward pass to produce a segmentation map directly [9]. This new feature of the U-Net enables us to consider the entire context of the image and this can be taken as a benefit in comparison with to traditional CNNs.

1.2. Motivations and Problem Formulation of the Study

Lung cancer is more life-threatening than other types of cancers. Because it starts to spread very early after its formation. Accurately and effectively identifying lung cancer in the early stage is beneficial. Regardless of the advantages in helping early diagnosis, physicians do not always make the best use of the data acquired from the imaging devices. With the increasing popularity of regular CT screening, the amount of CT scans grows rapidly and it takes a great amount of time for physicians to manually read through all the scans. In addition, physicians as humans may have limitations of visual system, insufficient training and experience, and the diagnosis may also be affected by factors such as fatigue and distraction that contribute to inefficient use of

available information [10]. One of the main difficulties in medical diagnosis is the subjectivity of the physician's decisions. When different experts are asked to trace the boundary of a pulmonary nodule, their response can show variations that are sometimes very significant. These variations depend on the training and experiences that each individual received his or her particular area of speciality. Generally, using manual methods to segment pulmonary nodules is tedious, time consuming and subjective in identifying the correct locations of the nodules. In such scenarios, automated techniques of image processing and analysis can be used as a medical aid tool to minimize these difficulties. The system can also significantly reduce the work for labeling large dataset that is useful in training more sophisticated models.

Now days, CAD becomes one of those the most active research areas in the field of medical image analysis. Automatically segmenting pulmonary nodules in CT scans is one of those mostly studied CAD applications. CAD can assist physicians to efficiently identify lung cancer through CT scans and also significantly reduces the work for labeling larger dataset [11]. Several written papers give overall overview about current developments in the field and describe the major characteristics and shortcomings of the different methods. Some of the drawbacks of the existing pulmonary nodule segmentation methods and algorithms include; reduced accuracy in detecting and segmenting true nodule pixels, reduced segmentation overlap accuracy between true nodule and true non-nodule pixels, less accurate in predicting the exact location of pulmonary nodules in the given scanned images, need of human intervention to locate the ROI, limited to a certain type, size, or location of a pulmonary nodules, and being poor to accurately segment a pulmonary nodules that varies in type, size, shape, and location and these makes the physicians to lose their confidence in using the CAD system as a useful tool.

Therefore, we should propose an automatic approach that helps perfect segmentation of pulmonary nodules from lung images. The development of deep learning models that works over neural networks monitored by the current advances in computational power enables new computer vision based intelligent diagnostics to be developed. These diagnostics are capable to analyze lung CT images, performing accurate segmentations in order to detect and segment pulmonary nodules and to make final decisions on the patient's health with respect to the best clinical views. However, only very deep convolutional neural networks with large amounts of

trainable parameters and large size valuable and annotated data are able to approach this kind of segmentations and to make the system converge properly.

In this thesis work, we combined thresholding and morphological operators which is a good choice for lungs region segmentation algorithm to minimize the search space together with a well-known convolutional network architecture called U-Net model. Furthermore, ensemble learning method is applied on three different network architectures of a U-Net model as a post processing technique.

At the end of this work, we are expected to evaluate the achievements of our proposed method towards the following scientific research questions:

Research Questions:

1. Can we improve the accuracy of pulmonary nodules segmentation in CT images using post processed U-Nets?
2. Can we use the same method to segment different pulmonary nodules that varies in type and size?

1.3. Objectives of the Study

1.3.1. General Objective

- To implement an automatic system that can accurately segments pulmonary nodules from lung CT images using CNN.

1.3.2. Specific Objectives

- To improve the sensitivity of a segmentation method.
- To improve the dice coefficient of a segmentation method.
- To improve the segmentation accuracy of different pulmonary nodules that varies in type, size, shape, and location.

1.4. Significance of the Study

- Since there are many lung cancer victims in Ethiopia, and the diagnosis made and treatment given by a physician is not yet supported by accurate automatic CAD system, this work motivates, satisfies, and reduces the work load of physicians, and increases the patients' satisfaction by assisting physicians to deliver better diagnosis and treatment.
- This is an automatic CAD system that helps physician to get pulmonary lung nodules locations information with a reduced error rate and time.
- This system increases the existence of lung cancer patients by enabling physicians to provide appropriate medication for early treatment of the disease.
- Our thesis improves the objectivity of lung cancer diagnosis and greatly reduces the subjectivity of the specialist decision. Therefore, it helps physician to develop their confidence in using automatic CAD systems as a useful tool.
- This study can be used as a reference for those researchers who want to develop an automated system for pulmonary nodules segmentation from lung images in a different approach.

1.5. Contributions of the Study

Our thesis has the following contributions:

1. For accurate and exact pulmonary nodules segmentation, it is best to apply a mechanism that can minimize the search space by segmenting the lung structure. For this purpose, we applied an algorithm called thresholding and morphological operators that can segment the lungs region well enough.
2. We modified the original U-Net architecture of Ronneberger et al. [9] to reduce the memory requirement and to get the best U-Net architecture for our task. We also applied batch normalization and dropout at each layer to improve learning that results better performance in segmenting pulmonary nodules from lung CT images.
3. We combined thresholding and morphological operator lungs region segmentation algorithm with the modified U-Net model to develop an automatic CAD system that can outperform the existing related systems.

4. We further improved the performance of our single U-Net model by using three U-Nets having different network architectures and hyper parameter values to be post processed by using average ensemble learning method (MBM).

1.6. Scope and Limitations of the Study

1.6.1. Scope of the Study

- This thesis work attempts to design, implement and test pulmonary nodule segmentation from lung CT images using a modified convolutional network architecture called U-Net model together with thresholding and morphological operator lung segmentation algorithm.
- For better and robust prediction, we used three U-Nets to apply average ensemble learning method (MBM) as a post processing technique.
- Our system can segment those pulmonary lung nodules having sizes from 3mm to 30mm in diameter, thicknesses less than 2.5 mm, and found on lung parenchyma, attached with the lung wall, or attached with blood vessels of lungs.
- Since we do not have a dataset that contains pulmonary lung nodules based on their type, we evaluated the performance of our proposed system in segmenting different pulmonary nodule types in general.
- We test and measure our model with different hyper parameters, and with a post processing technique.
- We compared and contrasted our proposed system performance with previously implemented systems.

1.6.2. Limitations of the Study

- Since the malignancy rate of pulmonary nodules having sizes lower than 3mm and greater than 30mm in diameter is less than 5%, our system will not consider them to segment from the given lung CT images.
- 3 dimensional lung CT scan images can’t be used to train and test the model.
- We did not test the effectiveness of our proposed system in segmenting pulmonary nodules by considering their type and location, separately.
- Our system can’t classify the segmented pulmonary nodules as malignant or benign.

- We did not test our proposed system for the analysis and diagnosis of CT scan images other than lung CT scans.

1.7. Research Methodology

This thesis work presents an automatic system that incorporates pre-processing and lung segmentation algorithm to reduce the search space, modified U-Net model for nodule detection and segmentation and average ensemble learning method as a post processing mechanism to improve pulmonary nodules segmentation performance. We proposed this method due to the lower effectiveness of the existing pulmonary nodules segmentation systems for early detection of lung cancer. This proposed system is evaluated using metrics DSC, accuracy, sensitivity, and specificity. Therefore, in order to conduct and accomplish this research work successfully, a number of steps have been followed.

1.7.1. Review of Related Literatures

This section is the major portion of our work which should be done continuously throughout the task until final documentation preparation. Before starting the actual work, deep studies have been made in the literatures written on the field of pulmonary nodules segmentation, to have a clear picture and understanding about the research idea. Researches previously written on pulmonary lung nodules segmentation have to be reviewed in order to get an understanding of the various techniques and methods of automatic pulmonary lung nodules segmentation systems.

1.7.2. Dataset Collection

To implement our system, which segments pulmonary lung nodules; collections of CT scan images are needed. These lung CT scan images are obtained from LUNA16 [12] database, which is a subset of LIDC/IDRI [13] database. LIDC/IDRI database is standard CT scan library and is mainly to build a system for pulmonary lung nodules segmentation. In this thesis, we used the dataset from LUNA16 database.

1.7.3. System Design, Implementation, and Testing

System design involves a series of very essential steps. Lungs region segmentation and pre-processing is the primary stage that should be accomplished. This is done to reduce our ROI and

to make our system more accurate. If we are directly inputting the raw CT scan images to our neural network, it is becoming a difficult task to identify (detect and segment) the exact location of a pulmonary nodules, especially juxta-pleural and juxta-vascular. For this purpose we used a good algorithm called thresholding and morphological operators. Then normalization, cropping and resizing is followed to remove those artifacts, noises, and the region of an image that contains less useful information but still available, from the segmented lung CT scan images.

Once our ROI is identified, it can be given directly to the proposed neural network. For better accuracy and memory efficiency, we proposed a model which is the adapted form of the original U-Net model designed by Ronneberger et al. [9]. We further improved a single U-Net performance by using average ensemble learning method as a post processing technique, which can combine the outputs from three U-Net models having different architectures and hyper parameter values.

We implement and test the proposed system using python language. Python is one of modern and currently developed object oriented programming language. To write the source code, to store and read the files, and to train the model, we need to install a compiler. First and for most we have to install anaconda. Anaconda is a powerful tool that contains more than 1000 libraries. Then we have install JetBrains Pycharm Community Editions 2018 as a compiler, which has the ability to work with python language and anaconda software tool. JetBrains Pycharm, in general, is mostly used in researches.

1.8. Organization of the Thesis

The remaining of this thesis report is organized as follows. Chapter two elaborates the theoretical background of the research. Chapter three discusses the researches that have been made on the segmentation of pulmonary lung nodules from lung CT images using different approaches. Design and implementation of automatic pulmonary lung nodules segmentation system from lung CT images is presented in Chapter four. The experimental test results, results discussion, comparisons, and the evaluation to the response of research questions are elaborated in Chapter five. Lastly, in Chapter six conclusions and future works are described.

Chapter Two

2. The Theoretical Background of the Research

2.1. Introduction

Having a clear picture about the theoretical concepts that are foundations for our thesis work is a key phase in the research development process. Understanding the basic notions behind the research enable us to gather information about current theoretical and scientific knowledge and to develop our own know how and realization regarding concepts such as pulmonary nodules, deep learning algorithms, optimization of deep learning algorithms, CAD, CT images and image segmentation, and thresholding and morphological operations. Generally, it is a critical summary of the background knowledge on the topics that are the base for our thesis work.

2.2. Lung Cancer

Lungs are soft organs that are located at the upper part of the body. Lungs are two in number that are slightly different in size and shape. Left lung is a bit smaller in size than the right lung since heart is located near it. Oxygen and other different gases come from air and moves into the body through the lungs. The oxygen is then transported from the lungs into the blood stream and hold through the body. Red blood cells store the carbon dioxide and send it back to the lungs, and then the lungs exhale it from the body.

Among many other cancers, lung cancer is the type of cancer currently taken as one of the most common and dangerous type of cancer which unchecks the growth of unusual cells either in one or in both of the lungs [14]. These irregular cells cannot do the functions of healthy human cells and cannot mature into normal cells. This irregularity disturbs the proper regular functioning of the lungs in supplying of oxygen to the human body through the blood. Although there are many advances in treatment procedures, the lung cancer which is at an advanced stage or late stage is not often easily curable [2]. Based on the report of World Health Organization, in 2012, lung cancer resulted in 1.59 million deaths across the world & caused approximately two times more deaths than liver cancer, the 2nd primer reason of death by cancer. In 2017, 222,500 new lung cancer cases occurred in USA with estimated 155,800 deaths. The 5-year relative existence rate

for lung cancer is 15% for men and 21% for women. However, the existence rate of lung cancer patients can increase to 55% if it is detected early [15].

2.3. CT Scan Images, CT, and CAD

Images are considered as one of the major important medium of conveying information in the area of computer vision and understanding the information obtained from images can be used for other tasks [16]. An image can be a 2D or 3D picture and a digital image is a numerical representation of a 2D image. A digital image is consisted of a limited number of elements having a specific location and value, are called picture elements, image elements, and pixels [17]. Pixels are the tiniest individual element in an image, holding finite, discrete, quantized values that represent the brightness, intensity or gray level at any specific point. Almost all CT images are now digital, thus allowing increasingly complex image reconstruction methods as well as image analysis methods within or as an additional to picture archiving and communication systems. Now a day, using high resolution CT image dataset permits the quantitative measurements such as size, shape and density, for each nodule to be made [18].

For around 20 years, CT examinations have been broadly used for early detection of cancer. CT as a new type of medical imaging equipment, has high density resolution, adequate image information, and makes a variety of post-processing of images [9]. CT outperforms conventional radiography in the screening of lungs. This is because of CT scan can produce very detailed high-resolution images, it can indicate early-stage lesions that are too small to be detected by conventional X-ray, it has the ability to screen nodules which can be malignant, and it can also detect nodules which are as small as 3 mm in diameter. In general, CT has a significant value for earlier detection of pulmonary lung nodules. CT scan of lung generates continuous cross-sectional images, and to approve the cancerous nature of lung, it is essential to analyze every cross-section [1]. Physicians use metrics such as size, intensities, observable shapes, and textures as a way of visual differentiation between healthy and abnormal anatomical structures of the lung tissues. However, it is one of among the most difficult tasks performed by physicians due to the vast amount of data generated by CT scan. The physician needs to apply additional efforts to analyze each cross-sectional image of the lung and therefore there is high probability of an error.

Therefore, developing of CAD tools are needed to assist physicians in analyzing and evaluating the CT scan images more accurately by reducing the image reading time. A well-built CAD system assists physicians in processing images for detection, categorization and extraction of irregularities and also helps in classification of image features between healthy and abnormal. CAD system is a very significant field of biomedical image analysis in which computer vision, image processing, visualization, and machine learning approaches have played main roles in the development and validation of these systems [19]. In many of the cases, CAD provides a very important 2nd choice when physicians diagnosis patient at cancer CT screenings. A CAD system is also influential in minimizing the amount of false negative diagnosis. The successfulness of a CAD system is evaluated in terms of accuracy in diagnosis, speed and its degree of automation [11].

2.4. Lungs Region Segmentation

Image segmentation is the way of dividing a digital image into several regions or objects that contains sets of pixels sharing similar properties or characteristics which are given different labels for representing different regions or objects [20]. The purpose of segmentation is to shorten and/or alter the illustration of an image into something that is more expressive and simple to analyze. Specifically, segmentation in medical images is used to extract and produce quantitative information for the interested organs or for those organs that have lesions within it. Generally, a segmentation problem comprises of two tasks; object recognition and object delineation. Object recognition is a way of determining the locations of the target objects on the image whereas object delineation is drawing of the object's spatial extent and composition [20].

For automatic image analysis in pulmonary nodules segmentation, lung segmentation is usually a significant first step. It plays a key role in pulmonary nodule segmentation by enhancing reliability, accuracy, and precision, and by reducing the computational cost of segmentation. Accurate lungs region segmentation improves the detection, segmentation, and quantification of defects within the lungs. When lung segmentation fails to correctly outline the borders of the lungs, findings might be missed or findings outside the lungs might be included in the analysis. Accurate lungs region segmentation is important to measure the interstitial lung diseases, and also for accurate detection, segmentation, and characterization of pulmonary nodules. The

significance of exact lung segmentation for a CAD scheme was illustrated by [21]. They showed that, computer-aided detection systems lost 17% of all true pulmonary nodules due to wrong lungs region segmentation. They also proved that when the lungs region segmentation done correctly, the percentage of missed nodules is reduced to 5%.

2.4.1. Thresholding and Morphological Operations Segmentation Technique

Segmentation by thresholding is an easy but influential method for segmenting images that has bright objects on dark background. Thresholding operation changes an image with several levels into a binary image i.e., it selects a good threshold T, to divide image pixels into several regions and to isolate objects from background. The gray-level histogram of an image is usually taken as an efficient tool for development of image thresholding algorithms [22]. Thresholding creates a binary image from grey-level images by changing all pixels below some threshold level to zero and all pixels above that threshold level to one.

Morphology operations can be described as a collecting of image processing techniques that processes images by taking shapes into consideration. These morphological operations considered applying of a structuring element on top of an input image to create an image of same size. In such operation, comparing pixels in the information picture with its neighbours is taken into account to approximate every pixel in the return image. This is done by picking the size and state of the area. At that point, we can develop a morphological operation that is subtle to specific shapes in the input image [23]. The structure element is a matrix consists of 0’s and 1’s, where the 1’s are called the neighbours. The value of each pixel in the output image is set with respect to in comparison with the corresponding neighbour pixels in the input image. The structuring element has many shapes; in relation to its applications. The most basic and commonly used morphological operations are dilation, and erosion. Dilation is to include pixels to the borders of objects in an image, while erosion is to remove pixels on object borders. Closing and opening are other morphological operations obtained from the basic operations of dilation and erosion. Closing is the application of dilation then erosion while opening is erosion followed by dilation. The amount of pixels that are added or even removed from the structure in an image highly depends on the size and shape of the structuring element that is applied to process that image. In morphological operations, the condition of any given pixel in the output

image can be decided by applying a rule to the studied pixel and its neighbours in the input image [17].

2.5. Pulmonary Nodules

A pulmonary nodule in radiology is described as a mass in the lung usually spherical in shape, with moderately identifiable margins and diameters smaller than 3cm [24]; however, it can be distorted by surrounding anatomical structures such as the pleural surface. Pulmonary nodules can also be defined as clusters of tissue in the parenchyma of lung. Based on their position in the parenchymal region, they are classified under four categories [25], as shown on Figure 2.1: (A) well-circumscribed nodules are discrete, well-marginated, with oval opacity, and entirely enclosed by lung parenchyma; (B) pleural-tail nodules are placed near the pleural surface and attached by thin line–like structure; (C) vascularized nodules are significantly connected with neighbouring vessels and located centrally in the lung; and (D) juxta-pleural nodules are significantly connected with pleural surface.

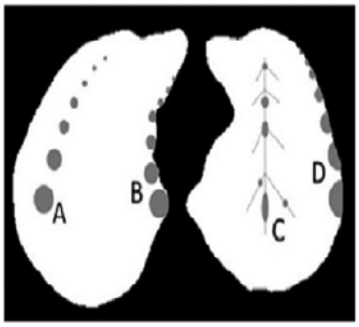


Figure 2.1 Template model of lung nodule [1].

The internal texture of pulmonary nodules can be solid, part-solid, and non-solid as shown on Figure 2.2. The solid nodules completely obscure the whole lung parenchyma within it. The non-solid nodules denote to focal nodular opacity with dim increase in attenuation and do not obscure the underlying parenchymal structure including vessels and airways. The part-solid nodules combined solid and non-solid nodules and indicate huge deviations of intensity within it [26].

Many related reported works specifies that nodules of non-solid and part-solid types are frequent and have higher risks of being malignant than solid ones.

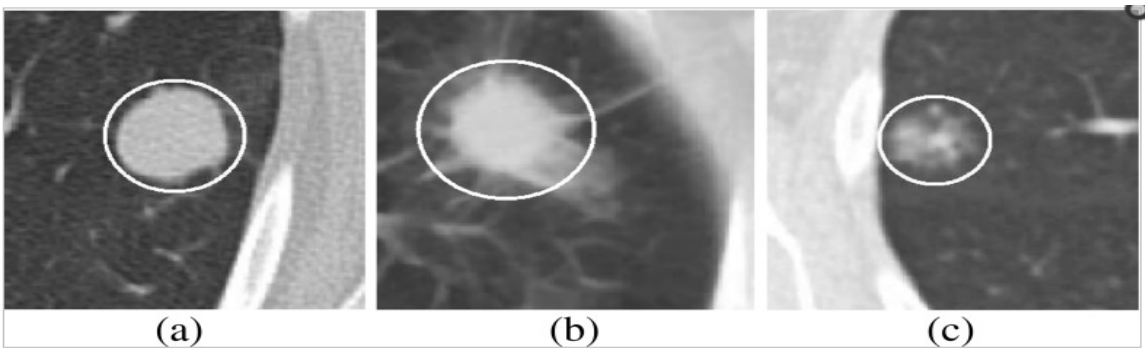


Figure 2.2 Pulmonary nodules based on internal texture: a) solid, b) part-solid and c) non-solid [26]

The segmentation of lung nodules is an essential part of a system that is used in the prevention and diagnosis of lesions. Automatic Segmentation of pulmonary nodules is required for robust, capable, and effective CAD system design and the performance of a segmentation technique strongly dependent on the extraction of foreground region, and the exactness of removal of pleural surface and blood vessels connected with it [1]. CAD system aims to improve the ability to detect and segment the nodules and can also help in classifying the nodules as malignant or benign. While automatic detection of the lung pulmonary nodules is a very significant task, segmenting the nodules after detecting them is also an equally difficult task due to different pixel intensities, non-uniform nodules shape, and the nodules positions.

2.6. Deep Learning

Deep learning is the newest area in the machine learning field. It is based on ANN, which are biologically-inspired computing systems. Today, deep learning algorithms are widely used in image processing, speech recognition, natural language processing, audio recognition, and bio-informatics. Deep learning methods have the main advantage over old-fashioned machine learning algorithms like SVM or shallow neural networks. The main advantage is that deep learning algorithms extract the features in the data by themselves. Therefore, there is no need for human intervention during the training process. Besides, this feature extraction mechanism generates features that are hard for a human to think and implement. Compared to other techniques such as standard machine learning algorithms, deep learning is a preferred option; its performance increases as the data scales. Machine Learning algorithms such as SVM, random

forest and etc, performance plateaus when data increases in these algorithms [27] as shown on Figure 2.3.

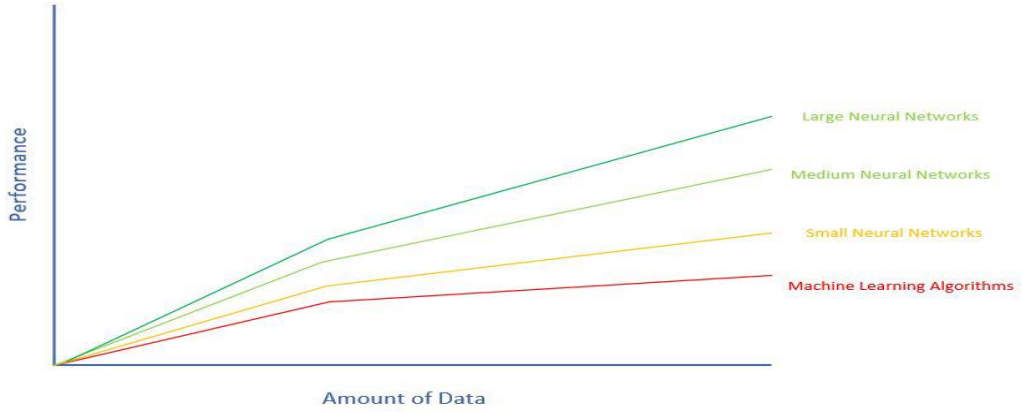


Figure 2.3 Performance of Neural Networks compared to Traditional Learning Algorithms by [27].

2.6.1. Artificial Neural Network

An ANN is an interconnected architecture where there exists an input layer where input data is placed, a hidden layer(s) where artificial neurons are stacked on top of each other and an output layer where the prediction or classification is made [28]. An ANN is usually a combination of forward and backward propagation techniques. Forward propagation is a technique in which data moves through from the corresponding input layer, hidden layers and output layer sequentially. But back propagation is the opposite of forward propagation, it feeds the network backwards. Back propagation is used to adjust the weight of the neural network after the errors have been computed by forward propagation algorithms. A hidden layer usually consists of a weight which is used in an optimization algorithm such as gradient descent, an activation function such as sigmoid, tanh, relu etc. and a loss function to calculate the loss or error of the function which we use to back propagate to adjust the weights.

Activation functions are an important part of a neural network. Since most real world data that we want our neural network to learn is non-linear, activation functions allow neural networks to create non-linear functions to solve problems. The three major currently used activation functions are Sigmoid, TanH and ReLU, but ReLU has been found to be the best in most

situations. Activation functions allow or stop neurons from firing into the next layer by comparing it into the activation function thresholds [29]. Activation functions are used both in the forward and backward propagation where in the forward propagation an activation function is used to calculate the loss where the output of a function is compared to the a real number and in backward propagation they update the parameters of the neural network. Figure 2.4 shows commonly used activation functions in a neural network.

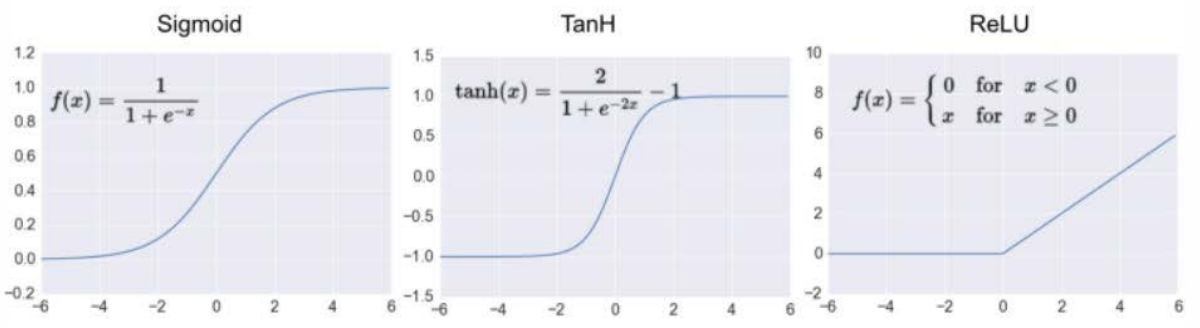


Figure 2.4 Common activation functions [29]

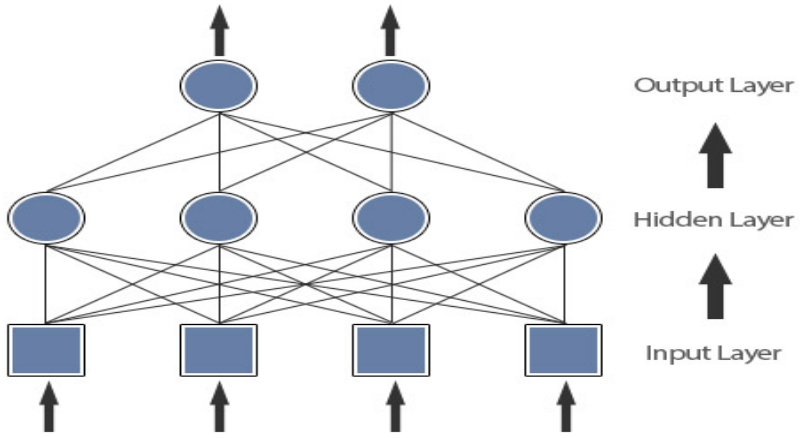


Figure 2.5 Artificial Neural Network also known as Multi-Layer Perceptron [28]

2.6.2. Convolutional Neural Network

CNN or ConvNets is a type of feed-forward ANN. CNNs use very little pre-processing compared to regular neural network models. Network model in CNN learns the features itself by creating a relationship between adjacent nodes. CNNs are generally composed of convolution layers, pooling layers, and fully connected layers. CNNs use filters to extract the specific

patterns, use pooling to help the model to ignore redundant data, and use MLP as a final step to vectorize the resulting data.

2.6.2.1. CNN Architecture

Like a Neural Network, a typical CNN consists of a multiple hidden layers called a convolutional layer where the linear function computes the strided convolutions over an image to extract features. CNN also consists of a pooling layer function such as Max Pool to reduce the size of the image in the neuron and to speed up the computation. Pooling layer does by extracting the features of the neuron image and ignoring the rest, this makes the network more robust. Fully connected layer in CNN, which is like a hidden layer in a neural network where the sum of the outputs of each layer are attend and where each value is an input to the next layer followed by an activation function and an output [30].

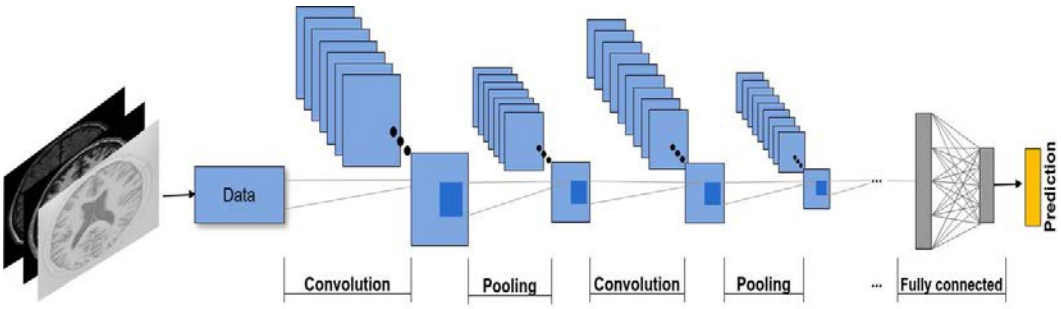


Figure 2.6 Convolutional Neural Network Architecture Example [31]

In a CNN, the linear function that is used is called a convolutional layer. Each neuron in the hidden layer extracts different features by using image processing feature detectors. These features are extracted using a kernel or filter [32]. Figure 2.7 shows an example of how strided convolutions works on an image. The lower is the original image and the upper is the output of the convolutions. Thus, the outcome of the convolutions decreases the dimension of the original image.

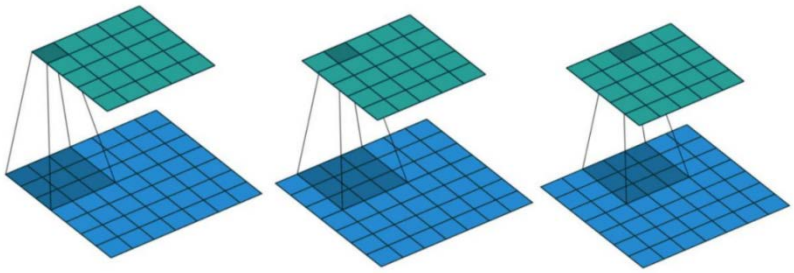


Figure 2.7 Strided convolution operation example [32]

The pooling layer is work out after the convolutional layer. The reason why pooling included in the architecture is to further reduce the dimensions of the convolutional layer and just extract out the features to make the model more robust. Pooling can be max pooling or average pooling. Max pooling extracts out the highest pixel value out of a feature while average pooling calculates the average pixel value that has to be extracted.

Convolution and pooling layers are used to generate high level features. With the help of a fully connected layer, these high-level features can be multiplied by the weights of the hidden layers [31]. A convolutional layer followed by a ReLU and pooling layer can be used multiple times before the fully connected layer. In CNN architecture, earlier convolution operations look for low level features. Instead, later convolution operations try to look for high level features, which are specific to the training data. Therefore, the model will extract more high level features if there are more convolutional layers in the network.

2.6.3. U-Net Model

A U-Net model is a different variation of CNN. It has been mainly used for biomedical image segmentation, but it is being used in a variety of ways to segment regions of interest [9].

2.6.3.1. U-Net Architecture

Ronneberger et al 2015 [9], indicates that a U-Net model comprises of a contracting path and an expanding path. This enables the U-Net network to create features that are necessary to detect and predict the patterns that are being looked. The problem associated with the traditional CNN is that pooling layers down-sample the size of the image and hence it makes some data to be lost as it feed forwards. The unique design of the U-Net model lies in its expanding path, which

consists of an up-convolution that merges or concatenates layers. The up-convolutions take a down-sampled sized image and expand the borders to escalate the size of the image. The concatenation layers then takes output from an earlier convolution layer and merge it with the previous up-sampled image. This concatenated data is then given to another convolution layer. The network is capable to send data from an earlier network to a later one through skip connections. This is how the network is able to segment the images after it has found the features.

2.6.4. Enhancing Deep Neural Networks

Deep learning is a highly iterative process and training is extremely slow and can get stuck on plateaus i.e. the model may not get better. To avoid this, different techniques have been applied on the model during training for a better. There are many new techniques evolved constantly to improve a neural network.

2.6.4.1. Batch Normalization

In old-fashioned deep networks, very big learning rate may leads to vanishing or exploding gradients as well as get stacking in poor local minima. Batch normalization is a currently popularized approach that can addresses these issues for fastening deep network training by making data standardization an integral part of the network architecture. It offers a definition for feed-forwarding the input and calculating the gradients with regard to the parameters and its own input via a backward pass. Batch normalization can be seen as yet another layer that can be inserted into the model architecture. In practice, it is inserted after a convolutional or fully connected layer but before the output is given to an activation function. In convolutional layers, the various components of similar feature maps .i.e. the activations at various positions are normalized in a similar way so as to respect the convolutional behavior. Hence, the entire activations in a mini-batch are standardized over whole positions, instead of per activation. By standardizing activations all over the network, it protects a slight adjustment of parameters from being amplified into greater and suboptimal alterations in activations of gradients [33].

2.6.4.2. Regularization using Dropout

Neurons in a neural network have a high tendency to adapt co-dependency among each other during training time. This means that certain activations become more heavily weighted and results in a network becoming overfit to the training set [34]. To stop the network from being overfitting, dropout is used.

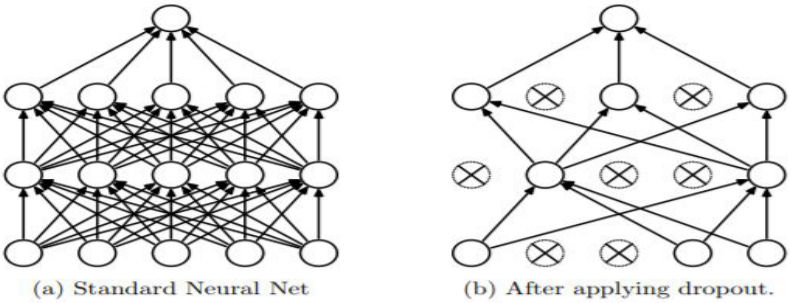


Figure 2.8 Dropout on a Neural Network [34]

Dropout is a regularization method that decreases neuron's ability to be interdependent among one another during training time. This is done by shutting off random neurons in the network during training time. This encourages and forces the network to learn better features that are more robust relative to random subsets of other neurons [34]. Dropout introduces an extra hyper-parameter .i.e. the probability of retaining a unit p. This hyper-parameter controls the intensity of dropout.

2.6.4.3. Weight Initializations

One of the major problems faced in deep learning is vanishing or exploding gradients. This is when the activations become smaller or bigger at each layer. If the weights in a network are too small, then the signals contract as it passes through each layer until it’s too small to be useful, conversely, if the weights in the network start too large, then the signal becomes too large to be useful. This has a terrible consequences to update the parameters on the weights as the network back propagates, giving either very small or very large updates into the network. It means that the larger the network is, the more likely the network to be slow to train and inaccurate. Weights can be initialized either simply randomly, using Xavier, or He. Xavier, and He initialization techniques put neurons into a Gaussian or uniform distribution and draws weights from the

distribution at a certain mean and variance. These approaches used as a good starting point for initializing and mitigating the probabilities of exploding or vanishing gradients. They set the weights neither too much bigger that 1, nor too much less than 1. So, the gradients do not vanish or explode too quickly. The weights initialized by He is random still but vary in range based on the size of the preceding layer of neurons and this gives an organized initialization; hence faster and efficient gradient descent [35].

2.6.4.4. Gradient Descent

Gradient descent is the learning algorithm used in deep learning. It is mostly used in the back propagation step in the training of the neural network. In a neural network, the loss function measures the variation between a predicted value and the actual value. The idea behind gradient descent is to increase the quality of the weights and to minimize a loss function in the neural network using these weights. Gradient descent uses the cost function to calculate the derivative and this is also known as calculating the gradient of a loss function. Calculating the gradient computes the best direction in which the weights must be updated to guarantee a best outcome for minimizing the loss function. The learning rate parameter then is used to determine the size of the steps we take given the direction [36]. The closer the network gets to convergence i.e. lowest point in the loss function, the more accurate the neural network becomes.

There are many different types of gradient descent that can improve the algorithm's capability. From these variations, batch gradient descent, stochastic gradient descent and mini-batch gradient descent algorithms are some of them. Stochastic gradient descent is a type of gradient descent algorithm that performs a parameter update on each training instance and its corresponding label. SGD removes the redundancy problem which means that it could be faster during training. It also means that stochastic gradient descent could jump out of a local minima during training and go towards a better local minima [36].

2.6.4.5. Gradient Descent Optimization Algorithms

Although there are different gradient descent algorithms with their own advantages and disadvantages, there are still some issues regarding training the network that has to be solved in a different way. Choosing a learning rate can be difficult, as statically predefining the learning

rate, can either mean slow convergence when it's too small or swing about the least when the learning rate is very large. Choosing the same learning rate means that an update applies to each parameter and this can be a problem if we don't want to update the weights that find less frequent features in our data [36] [37]. Therefore, to solve this problem, there needs to be a way to choose the learning rate dynamically.

Adam optimization algorithm is a new version of SGD that can compute adaptive learning rates. This technique was invented by Diederik Kingma and Jimmy Ba. Adam maintains an exponentially decaying mean of the squared gradients and exponentially decaying mean gradients. Since these 2 values are 0 when initialized, Adam uses bias correction for the first and second moment estimation for correcting the algorithms initial bias towards 0. These 2 values are kept to later update the parameters [37]. Diederik Kingma and Jimmy Ba also proofs Adam outperform the other optimization algorithms when used with dropout.

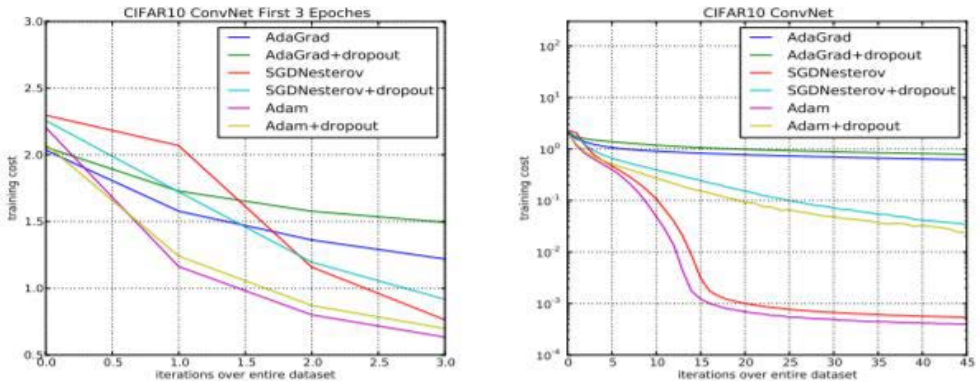


Figure 2.9 Adam performance on CIFAR-10 Benchmark [37]

2.7. Summary

Among many other cancers, lung cancer is the type of cancer currently considered as one of mostly occurred dangerous types of cancer which unchecks the growth of unusual cells either in one or in both the lungs. Although there are many improvements in treatment measures, the lung cancer which is at a progressive or late stage is often not curable easily [2]. Therefore, early stage cancer detection is needed. For around 20 years, CT examinations have been broadly used for earlier detection of cancer [5]. However, it is considered as the most challenging task performed by physicians due to the vast quantity of data generated by CT scan. Physicians usually used metrics such as size, texture, observable shape, and intensity, as a way of visual

discrimination between healthy and abnormal anatomical structures of the lung tissues. Therefore, there is a need to develop CAD tools that can assist physicians in analyzing and evaluating the CT scan images more accurately by reducing the image reading time.

For automatic lung CT scan image analysis, lung segmentation is usually an important first step. It plays a vital role in pulmonary nodule segmentation by enhancing reliability, accuracy, and precision, and by reducing the computational cost of segmentation. Accurate lungs region segmentation improves the detection, segmentation, and quantification of abnormalities within the lungs. Thresholding and morphological operators is an easy but operational type of algorithm that can be used in segmentation tasks.

A pulmonary nodule in radiology is demarcated as a mass in the lung usually spherical in shape, with moderately identifiable margins and diameters smaller than 3cm [24]. Pulmonary nodules segmentation is a pre-requisite step in developing a CAD system. Even though automatic detection of pulmonary lung nodules is a very essential task, segmenting the nodules after detecting them retains an equal critical and difficult task. This is because of, overlapped pixel intensities at different regions of the lung CT scans, non-uniform shape of nodules present inside parenchyma region, the position of the nodules inside the lung, and segmentation needs to be carried out continuously on consecutive slices to monitor the behaviour of nodule candidates from one slice to the other [1].

Deep learning models have been showing their dominance in classification, object detection and (medical) image segmentation. Deep learning models, specifically, convolutional networks, are rapidly becoming a preferred methodology for analyzing medical images. CNNs are generally composed of convolution layers, pooling layers, and fully connected layers. CNNs extract the specific patterns by using filters. A U-Net model is a different variation of a CNN. U-Net comprises of contracting and expanding paths and instead of fully connected layer; it uses 1x1 convolutional layer as a final layer. It has been mainly used for biomedical image segmentation problems. Training deep neural network can get better using different systems such as batch normalization, dropout, weight initialization techniques, gradient descent learning and optimization algorithms.

Chapter Three

3. Literatures Review

3.1. Introduction

Reviewing reported works that are related to our task is another key stage in the research development process. We reviewed various studies that were conducted globally to understand the methodologies used and the algorithms implemented in connection to segmenting pulmonary lung nodules. This helped us to have a wider perspective on pulmonary nodules segmentation from CT images and to identify the shortcomings of existing CAD systems. In addition, in this section, various lung CT image segmentation algorithms, methods, and approaches have been examined.

Approaches in segmenting pulmonary lung nodules involve the detection of the ROI bounding the nodule area and perform segmentation inside this ROI. Some of these methods can be region growing methods, morphology methods, active contour methods, or deep-learning methods. Before deep networks have been introduced, the well performing approaches mostly depend on manually crafted features which categorize pixels individually.

3.2. Morphological, Active Contour, and Region Growing Methods

In morphology methods, morphological operations such as logic opening operation were applied to remove the nodules attached with vessels and pleural surface, and then the connected component selection can separate lung nodules while in region growing approaches, segmentation starts with a user-specified seed point, and voxels are included into nodule set iteratively until the pre-defined converge criterion is satisfied. For example:

Kubota et al. [38] proposed a new algorithm that can be applied to solid, non-solid and part-solid types and solitary, vascularized, and juxtapleaural types. Their algorithm consists of six stages. The first stage called voxel transformation assigns a number between 0 and 1 that is comparative to their belief of the voxel, depending on the tick point, being a part of a forefront given its CT value. Then they isolate any body structures and opacity exists in the input volume of normal lung parenchyma using figure ground separation. Thirdly, they locate the center of the nodule by

computing the Euclidean distance map. In their fourth stage, region growing is applied to extract all voxels of the nodule and to isolate any body structures attached with the nodule. Surface extraction is their fifth stage which is responsible to remove the false positives. And finally they perform convex hull operation to segment the nodule by taking the convex hull of the result obtained from surface extraction stage and taking the intersection between the convex hull and the foreground. The algorithm proposed by [38] is semi-automated and needs a click point as an input. Their algorithm was tested on 23 nodules provided by LIDC and achieved a mean segmentation overlap of 0.69.

Moltz et al. [39] developed a new algorithm which is able to segment most juxta-pleural lung nodules. Initially, they apply region growing using the ROI center as a seed point to identify the nodule from other lung structures and then subsequent ray casting approach is followed to identify the boundary points between the nodule and the parenchyma. Secondly, they use ellipsoid approximation to reconstruct the shape by fitting an ellipsoid to the boundary points, assuming that the nodule has approximately an ellipsoid shape. Then convex hull operation is applied to use a slightly dilated version of the ellipsoid as a new minimal ROI and to reconstruct the original lung shape within this ROI. Finally, the adjacent structures are cut off by morphological processing within a dilated version of this ellipsoid. For their evaluation, they used a database of 333 ROIs juxta-pleural lung nodules from various patients, clinics and CT scanners with seed points set manually by radiologists. All of the nodules in their study had direct contact to the pleura and achieved a segmentation accuracy of 94%.

In addition, Sudipta [26] designed a segmentation outline for the whole types of nodules by considering their internal texture (solid/part-solid and non-solid) and external attachment (juxta-pleural and juxta-vascular). In their approach, after VOI selection and isotropic resampling is done, primarily pulmonary nodules are classified into solid/part-solid and non-solid class by examining concentration spread in the center of nodule. Then two different pre-processing techniques are designed for solid/part-solid and non-solid pulmonary nodules, correspondingly. To pre-process solid/part-solid nodules they used ellipsoid enclosing method while for those non-solid nodules they used anisotrophic diffusion together with ellipsoid enclosing. Thirdly, they remove pleural surfaces from the nodules. The pleural surface removal of solid, part-solid and non-solid nodules is done according to the method of [39]. Finally, they proposed

vasculature pruning technique to remove vessels from solid/part-solid nodules and selective enhancement filtering to remove vessels from non-solid nodules. The method proposed by [26] needs human intervention to select a seed point and click on the pulmonary nodule. They evaluate their method performance using 891 LIDC/IDRI pulmonary nodules and achieved a mean dice similarity coefficient of 0.46, mean accuracy of 0.99, mean sensitivity of 0.58, and a mean specificity of 1.0.

Among various morphological segmentation approaches, active contour was one of the very widespread and effective one that many researchers used to study the segmentation of pulmonary nodules connected with vessels and the pulmonary wall. For example:

Keshani et al. [40] proposed a technique for lung nodule detection, segmentation and recognition appearing in CT images. Their first task is to segment the lungs region. In their lung segmentation approach, the first step is binarizing lung CT images using adaptive fuzzy thresholding. Second, two windows of different sizes are applied to obtain a hole- free mask. Third, small and large non-isolated nodules connected to the lung wall will be transferred to isolated ones by applying two 45^{o} rotated window. Finally, this starting mask is used to apply ACM automatically and to segment the lung area exactly. After they did lung segmentation, they extract the features of the lungs area to classify it into two classes; nodules and non- nodules. For this binary classification setting, they utilize SVMs. During classification, some small deformations in the edges of nodules may be detected. Therefore, as a solution, they propose to use active contour based nodule edge taking out as the next step to resolve this deformation problem and to reduce the number of false positives. At this stage, all solid and cavitary nodules are accurately segmented. Finally, they categorize the lung tissues into four classes; namely lung wall, parenchyma, bronchioles and nodules. They evaluate and verified their proposed method on four groups of datasets and achieved an overall detection rate of 89% and False Positive value of 7.3/scan. Their system also attained a dice coefficient of 82% when tested on 19 nodules.

Most of the methods and approaches listed above, in addition to the human involvement, they produce segmentation results which are not satisfactory, that needs to be improved and may be undesirable for clinical usage. Therefore, it is essential to develop suitable techniques which can segment lung nodules with higher performance and less segmentation error rate.

3.3. Deep Learning Based Approaches

Old-fashioned architectures for solving medical image segmentation problems and the degree of success they enjoyed have been greatly dependent on hand-crafted features. However, lately, deep learning techniques have offered a fascinating alternative by automatically learning the problem of specific features. With this new model, the problem of a specific feature is now being re-viewed from a deep learning perception. The achievements of deep convolutional neural networks for object segmentation and classification has more recently directed researchers to get a plus from their feature learning abilities for well-thought-out prediction for tasks such as segmentation.

An earlier application of Deep Neural Networks to medical image segmentation is reported by Ciresan et al. [41]. They proposed a system to address a central problem of neuro-anatomy, namely, the automatic segmentation of neuronal structures depicted in stacks of EM images. To segment biological neuron membranes, they used a special type of deep artificial neural network which consists of a succession of convolutional, max-pooling and fully connected layers, as a pixel classifier. Their classifier is trained by plain gradient descent on a 512 x 512 x 30 stack with known ground truth, and tested on a stack of the same size given by the organizers of the ISBI 2012 EM Segmentation Challenge. They validate their approach on the publicly-available dataset provided by the organizers of the ISBI 2012 EM Segmentation Challenge, which represents two portions of the ventral nerve cord of a Drosophila larva and achieved average rand error of 48, average warping of 434, and average pixel error of 60.

Bronmans et al. [42], apply a 3-dimensional CNN model to perform a pixel-wise nodule segmentation task. Diagnosis of lung cancer triggered by malignant nodules in CT scans done by pulmonary radiologists commonly follow three stages; nodule localization, nodule detection, and nodule segmentation. They did the nodule localization stage using a relatively simple fully-connected CNN model implemented in Tensor-Flow, consisting of four convolutional layers followed by a single fully-connected layer while they are implementing nodule detection and segmentation using 3D CNN. Their model is trained and evaluated by a total of 1132 samples obtained from LIDC. From these available dataset, their model used 75% which is 849 samples for training, 11.9% which is 135 samples for testing, and 13.1% which is 148 for validation. The

evaluation results indicate that their model achieved 87.4 accuracy, 68.0 sensitivity, and 87.7 specificity.

CNNs systems like the methods proposed by [41] and [42], have shown successful application to the segmentation problems. However, traditional standard CNNs simply focused on local textual features. As a result, some important global features are lost certainly. In addition, as the depth of a CNN network increase, the amount of feature maps often increases and this results in a dramatic increase in the amount of parameters while the network complexity increased and this may lead to over-fitting.

To address this problem, recent semantic segmentation algorithms, convert an existing CNN architecture constructed for classification to a FCN. Long et al. [8] proposed the replacement of fully connected layers with convolutions using 1 x 1 sized filters, which can effectively solve the over storage problem. According to [8], FCNs can obtain a grainy label map from the network by classifying every local region in image, and do a simple de-convolution, which is executed as bilinear interpolation, for pixel-level labeling. Long restore the down-sampled feature maps to the original size of the input using de-convolution operations. The grainy segmentation map obtained through de-convolution is made finer by integrating it with segmentation maps calculated from earlier steps in the network, where the input has been down-sampled fewer times.

An alternative way of up-sampling feature maps is used by Badrinarayanan et al. [43]. The author stores the indices of activations that were preserved by max-pooling layers. The indices are later used for up-sampling. Un-pooling feature maps in this fashion produce a sparse feature maps in which all locations but the ones previously selected by max-pooling contain zero-values.

Wang et al. [44], proposed Joint learning for Pulmonary Nodule Segmentation, Attributes and Malignancy Prediction. Their proposed method uses an end-to-end multi-task and interpretable 3D CNN to concurrently forecast the malignancy of lung nodules, segment the nodule areas and study nodule characteristics. They believe that the suggested technique is capable to deliver semantic advanced characteristics as well as the region of pulmonary nodules, which enables radiologists to understand how to make decisions easily. To verify the effectiveness of their proposed method, they implemented and evaluated it on public LIDC/IDRI dataset, which

includes 1010 patients (1018 scans) and 2660 nodules with slice thickness varying from 0.45 mm to 5.0 mm and their model achieved a nodule malignancy prediction of 97.58%, an attributes prediction of 89.33% and a nodule segmentation accuracy of 73.89%.

In patch based systems, like the method proposed by [44], segmenting and classifying each pixel can be met by applying a classifier to patches of the input in a sliding window fashion. In these systems, each pixel is classified by extracting a patch around it and having it classified by the network. But this approach has drawbacks of time-inefficiency caused by duplicate computation, lower accuracy when the patch size gets smaller, and storage overhead and ineffective when high resolution images are processed.

The methods proposed by [8], [43], and [44] uses a fixed up-sampling techniques. Fixed up-sampling techniques are not able to exactly learning how to do up-sampling. Instead, they store the indexes of the pixels to be up-sampled and then either duplicate the pixel values or take one value and fill the rest of the neighbor pixels by zeros. Using fixed up-sampling simply transforms a small input into a large image output. It is not able to fill in useful details in the up-sampling operation to translate the input feature maps into meaningful details.

Therefore, there should be a mechanism that can overcome the problems of patch based systems and fixed up-sampling operations. The method suggested by Ronneberger et al. [9], has a network structure that can take the whole image without patch extraction with transpose convolution which is a learnable up-sampling technique.

Ronneberger et al. [9], build a system upon a more elegant architecture, the so-called “fully convolutional network” [8]. Their architecture involves a contracting path to earn context and an equal expanding path that allows exact localization. They modify and extend the architecture of [8] by making it to work with very few training images and yields more precise segmentations. The other important modification incorporated in the architecture of [9] is in the up-sampling part. They had also a huge amount of feature channels, which permit the network to transmit context data to the next level resolution layers. Their network is trained using input images and their respective segmentation maps with the stochastic gradient descent implementation of Caffe. They demonstrate the application of their u-net model on three different segmentation tasks. The first job is done on the segmentation of neuronal structures in electron microscopic recordings.

The dataset is provided by the EM segmentation challenge. The training data is a collection of 30 images with 512x512 pixels from serial section transmission electron microscopy of the Drosophila first instar larva ventral nerve cord. Each and every image of cells and membranes are given with their respective fully annotated ground truth segmentation map. The test set is publicly available. Without any further pre or post-processing, their u-net achieves a warping error of 0.0003529 and a rand-error of 0.0382. They also applied the proposed u-net model to a cell segmentation task in light microscopic images. The first data set “PhC-U373”2 comprises of Glioblastoma-astrocytoma U373 cells on a polyacrylimide substrate logged by phase contrast microscopy. On the first dataset which consists of 35 partly annotated training images, they succeeded a mean IOU of 92%. The second data set “DIC-HeLa”3 are HeLa cells on a flat glass logged by differential interference contrast microscopy. On the second dataset which consists of 20 partly annotated training images, they succeeded a mean IOU of 77.5%.

From the related works that we reviewed, we understood that a new approach is expected to get an improved performance in segmenting pulmonary nodules. We tried to increase pulmonary nodules segmentation accuracy by using deep learning model in combination with pre and post processing stages. We had designed a modified deep convolutional neural network architecture similar to U-Net model of Ronneberger. This U-Net model is used together with thresholding and morphological operators lungs region segmentation algorithm. We then tried to further enhance the performance of our modified U-Net model by using average ensemble learning method as a post processing technique. This ensemble learning method combines three U-Nets having different network architectures and hyper parameter values.

3.4. Summary

Approaches for lung nodule segmentation involved the detection of a ROI covering the nodule area and segmentation inside this ROI. Before the introduction of deep networks, the highest achieving approaches commonly depend on manually crafted features which classify pixels individually. Researchers conducted pulmonary nodule segmentation based on the conventional approaches depends on the artificial extraction of lung nodules, which are low-level features. Conventional methods need a separate feature extraction procedure and selecting best features that represent nodules depends on the knowledge and experience of medical experts and this

affects the overall performance of the segmentation system. The performance of these conventional approaches can be improved using state of the art learning algorithm called deep learning algorithms.

Many literatures reviewed were implementing traditional CCN architecture for pulmonary nodules segmentation. Some of the research works tried to modify the original CNN architecture to reduce the complexity of the network and this make training easier. Network architectures of some of the papers consist of convolutional and de-convolutional layers that help us to have a full context of an image and make the training better. Many research works reported in the literature were mainly focused on the training part of deep neural networks with large amount of training data. And only performing deep training may not be a guarantee for accurate pulmonary nodules segmentation.

We designed a system by combining a technique that can minimize the search space together with deep convolutional neural network. Search space reduction is done by segmenting the lung structure via a technique called thresholding and morphological operators. We combined the proposed lungs region segmentation algorithm with modified U-Net model. We then further used average ensemble learning method to integrate the predictions obtained from three different network architectures with various hyper parameter values of a U-Net model, as a post processing technique. Generally, combining different architectures, like us, suggests that a best result will be obtained.

Chapter Four

4. Design and Implementation of Pulmonary Nodules Segmentation System

4.1. Introduction

In this paper, we propose an automatic pulmonary nodule segmentation method based on CNNs. Old-fashioned CNNs focus only on local region features and disregard global region features even though both are important for pixel detection and segmentation. Moreover, pulmonary nodule can appear in any place of the lung and be any size and shape in patients. Since both local and global features play important role in segmentation tasks, we proposed to use U-Nets having an adapted form of the original network architecture of Ronneberger et al. [9] and post processed with average ensemble learning method (MBM). For better segmentation performance, we combined the proposed model with thresholding and morphological operators lung segmentation algorithm; which could automatically detect and segment pulmonary nodules having a size between 3mm to 30mm, that appear in any place of the lung and having any shape in patients from lung CT images.

Before the CT image is taken by the model, our initial approach was to simply input the lung CT scan images into morphological and thresholding lung segmentation algorithm. This lung segmentation algorithm will reduce our search space by segmenting only the two lungs region. We then pre-process the lung CT scan images using normalization, and cropping and resizing. Normalization is done to avoid some contrast issues while cropping is done to avoid the area of the segmented image that contains less useful information. The lung segmentation algorithm hides the bone, external air and other elements that could make our data noisy, and retain only the lungs information for the network. Then we are going to input the segmented and pre-processed lungs region into the modified U-Net model to detect and segment the pulmonary nodules from the given lung CT images. And finally our post processing stage uses average ensemble learning technique (MBM) that increases the accuracy of segmenting pulmonary nodules and improve the performance of our system as a whole.

Generally, our proposed methodology has five pipelines:

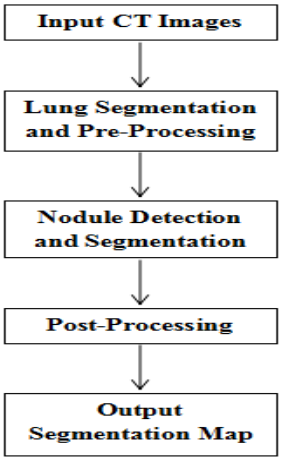


Figure 4.1 Work flow of our proposed methodology

4.2. CT Images

This step is responsible for finding a set of images used by automatic CAD systems. There are some public databases used for training and testing. The LIDC/IDRI is a public database which is the result of collections from seven academic institutions and eight medical imaging enterprises assisting the development, training, and evaluation of CAD approaches for the lung cancer diagnosis. As of February 2015, the database contained 1018 CT scans acquired using single and multi-slice scanners with slice thickness ranging from 0.6 mm to 5 mm. Each scan is provided along with annotations independently generated by four experienced radiologists associated with different institutions. Whole frameworks for all nodules between 3mm and 30mm in diameter were obtained together with their radiological features. Each LIDC/IDRI scan was annotated by experienced thoracic radiologists in a two phase reading process. In the first phase, four radiologists annotated the scans independently. All lesions were marked as nodule ≥ 3 mm; nodule < 3 mm; non-nodule. Secondly, independently annotated nodules were exposed to each radiologist, who reviewed all of the marks again [13].

For this study, to train and test our multi-stage framework, we utilize the LUNA16 dataset, which provides nodule annotations.

The LUNA16 dataset is a collection of 888 axial CT scans of patient chest cavities taken from the LIDC/IDRI database, in which only scans with a slice thickness less than 2.5 mm and nodules with diameter >= 3 mm are included. This resulted in a set of 1,186 nodules annotated by at least 3 radiologists. For each patient, the CT scan data comprises of a different amount of images (normally around 100-400) of an axial slice of 512 x 512 pixels. In each scan, the location and size of the pulmonary nodules are agreed by at least three radiologists [12].

The images are structures in 10 subsets and in each subset the images are stored in MetaImage (mhd/raw) format. Each .mhd file is stored with a separate .raw binary file for the pixel data.

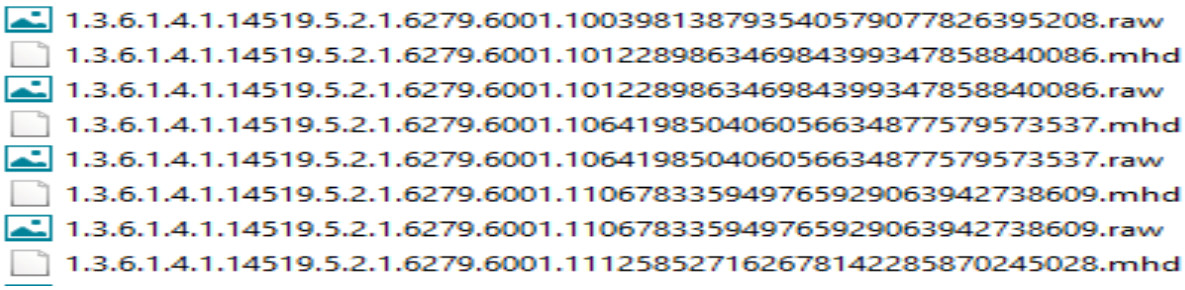


Figure 4.2 Sample raw lung CT scan images in mhd format from LUNA16

The annotation file is in a csv file that contains of a single finding per line. Each line holds the SeriesInstanceUID of the scan, the x, y, and z position of each finding in world coordinates and the corresponding diameter in mm. Generally, in LUNA16 dataset, 1186 nodules annotated cross 601 patients by the majority of the radiologists (at least 3 out of 4 radiologists) as positive examples in reference standard [12].

seriesuid	coordX	coordY	coordZ	diameter_mm
1.3.6.1.4.1.14519.5.2.1.6279.6001.100225287222365663678666836860	-128.699	-175.3192718	-298.3875064	5.651470635
1.3.6.1.4.1.14519.5.2.1.6279.6001.100225287222365663678666836860	103.7837	-211.9251487	-227.12125	4.224708481
1.3.6.1.4.1.14519.5.2.1.6279.6001.100398138793540579077826395208	69.63902	-140.9445859	876.3744957	5.786347814
1.3.6.1.4.1.14519.5.2.1.6279.6001.100621383016233746780170740405	-24.0138	192.1024053	-391.0812764	8.143261683
1.3.6.1.4.1.14519.5.2.1.6279.6001.100621383016233746780170740405	2.441547	172.4648812	-405.4937318	18.54514997
1.3.6.1.4.1.14519.5.2.1.6279.6001.100621383016233746780170740405	90.93171	149.0272657	-426.5447146	18.20857028
1.3.6.1.4.1.14519.5.2.1.6279.6001.100621383016233746780170740405	89.54077	196.4051593	-515.0733216	16.38127631
1.3.6.1.4.1.14519.5.2.1.6279.6001.100953483028192176989979435275	81.50965	54.9572186	-150.3464233	10.36232088
1.3.6.1.4.1.14519.5.2.1.6279.6001.102681962408431413578140925249	105.0558	19.82526014	-91.24725078	21.08961863
1.3.6.1.4.1.14519.5.2.1.6279.6001.104562737760173137525888934217	-124.834	127.2471546	-473.0644785	10.46585391

Figure 4.3 Sample nodules annotation information.

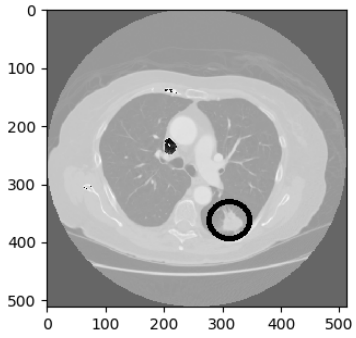


Figure 4.4 Sample slice that indicates a cross section of a patient chest cavity with annotated nodule

4.3. Lung Segmentation and Pre-Processing

Since the acquired data can bring inconsistencies in dataset, it's necessary to carry pre-processing steps to ensure that the model receives uniform inputs. A CT scan image of lung is filled with noise from surrounding tissues, bone, and air. Therefore, the first thing to do is image pre-processing. Pre-processing of an image is the procedure to enhance the quality and interpretability of the input image by reducing the noise, by removing unwanted objects, and by normalizing the concentration of the individual particles in images prior to computational processing to make the search efficient for the proposed CAD system. This stage is also important since the lungs also contain many other structures that can be confused with nodules. The Pre-processing treatments that are performed on the lung CT slices (images) gives better results in nodule detection and segmentation. In LUNA16 dataset, the CT Scans are saved in '.mhd' and the pixel values are stored in '.raw' file format. The first thing that we are going to do is converting the acquired image from .mhd file format to .npy file format, since we are going to evaluate each pixel separately for easier processing. Then we preprocess the CT scans using segmentation, normalization, and cropping and resizing. Accurate lungs region segmentation is appropriate to speed up pulmonary nodules detection and segmentation. For our task, for accurate lungs region segmentation, we used an algorithm called *thresholding and morphological Operators*.

4.3.1. Lung Segmentation

The image data is in MetaImage format and measured in HU. Table 4-1 shows the different attenuation values for lung and other substances (tissues) and each substance has different attenuation value which is measured in HU.

Substance	**HU Range**
Air	-1000
Lung	-400 -600
Nodule	-150
Fat	-60 -100
Water	0
Soft Tissue	+40 +80
Bone	+400 +1000

Table 4-1 Attenuation values of the different tissues [45].

Segmentation of the lungs region is applied on the images of CT which were taken from LUNA2106 dataset. The lungs region segmentation is an essential and crucial pre-processing step. It basically extracts the portions of the lungs from CT chest image to minimize the search space, to find the accurate ROIs, and to decreases the overhead for further processing. Lung segmentation is the procedure of extracting of the lungs in CT chest images and separates the ROIs from muscles, fats, and other attached pulmonary. The presence of noise, low contrast and intensity inhomogeneity in CT images makes the lung segmentation a difficult task. But the most challenging task in many literatures is to recognize the nodules attached to pleural surface (i.e. juxta-pleural) and nodules attached with the blood vessels (i.e. juxta-vascular) and to include these regions in the segmented lungs. To segment the lungs region, different approaches such as thresholding, watershed, region growing, and active contour model can be used but each has their own limitations. For example, thresholding is simple and computationally efficient lungs region segmentation but it is unable to segment two closely connected objects accurately if their boundary is not defined clearly. Similarly, watershed lungs region segmentation approach provides the segmentation result with relatively low computational cost but has a limitation of over-segmentation. Particularly, region growing lungs region segmentation approach is

computationally intensive and the performance depends on its initial seed point. In addition, active contour model is another type of lung segmentation approach which has the benefit to accurately segment the lung border and excludes tiny white parts exist on the lung area but has a disadvantage that the construction of initial contours may require human interaction.

Therefore, we propose a method that uses thresholding and morphological operators that can segment lungs region with those nodules attached with the lung wall and vessels of the parenchyma. In the proposed method, the boundary of the lungs region becomes smooth to remove objects attached with the lung walls.

In our proposed lungs region segmentation algorithm, morphological operations is applied to remove vascular and pleural surfaces and/or to fill holes in the lung parenchyma. A structuring element is a shape mask used in basic morphological operations. They can be any shape and size that is digitally representable (i.e. 0's and 1's) and each has an origin. For our task, we used a differently sized disk shape structuring element according to the position of an object going to be filled or removed. We applied dilation, erosion, closing and opening morphological operations. Dilation operation takes in a binary image, places the origin of structuring element over each 1-pixel, and ORs the structuring element with the input image at the corresponding. While erosion operation takes in a binary image, places the origin of structuring element on each pixel position, and ORs the structuring element with the input image only if each 1-pixel position of the structuring element overlaps with 1 of the input binary image. Closing and opening is the compound operations of dilation then erosion and erosion then dilation using the same structuring element, respectively. Closing is similar in some ways to dilation in that it leads to extend the borders of foreground (bright) regions in an image while shrinking the background color holes in such regions. But closing is less destructive of the original borderline shape. Similarly, the basic outcome of an opening operation is just like erosion in which it leads to eliminate some of the bright pixels from the boundaries of regions of foreground pixels. However, opening is less destructive than erosion in general.

Steps in thresholding and morphological operators lungs region segmentation algorithm:

Step 1: Create a binary image by applying KMeans on the center of the given CT image using an average threshold value of -420 HU [45].

$$\text{binary}(i,j) = \begin{cases} 1, & \text{if}(i,j) < T \\ 0, & \text{otherwise} \end{cases} \quad \ldots(1)$$

Here T is -420 HU and binary (i, j) is the binary image after applying thresholding. To convert HU to gray-level value, we used equation (2).

$$\text{Gray level value} = 1024 + T \quad \ldots(2)$$

Step 2: Find and label various connected components of all areas in the image.

Step 3: Apply closure operation with a disk of radius 3 to fill smaller areas which appears at the boarder of the lung and to keep Juxta-pleural nodules [45].

$$I \bullet S = (I \oplus S) \ominus S \quad \ldots(3)$$

Note that I is the input image and S is the structuring element.

Step 4: Extract the two largest lung areas of the connected components.

Step 5: Clear image border to remove the regions which are connected to the boundary of the image [46].

$$R = \begin{cases} 0, & \text{border}(R) = B \text{ pixels} \\ 1, & \text{border}(R) \neq B \text{ pixels} \end{cases} \quad \ldots(4)$$

Note that R is the regions with different labels and B pixels are the border pixels of the image.

Step 6: Apply morphological closing operation with disk of radius 12 inside the two largest lung areas to fill image holes in these largest connected areas.

$$I \bullet S = (I \oplus S) \ominus S \quad \ldots(5)$$

Step 7: To engulf the nodules attached to vessels, an erosion operation with a disk structuring element size of 4 is applied [45].

$$I = I \ominus S \quad \ldots(6)$$

Step 8: Apply morphological dilation operation with a disk structuring element size of 10 to fill holes created by removed blood vessels.

$$I = I \oplus S \qquad \ldots (7)$$

Step 9: Superimpose the resultant binary image with the given image to extract lungs region.

Step 10: Smooth borders by applying a morphological opening operation on the image from step 8 with a disk shape structuring element of radius 1 [45].

$$I \circ S = \{z|(S)z \subseteq I\} \qquad \ldots (8)$$

Note that I is the reference binary image, S is the structuring element of shape disk with the size 1 and z is the outcome of equation 8. Then subtract the output image of step 10 from the image of step 8.

$$B = I - z \qquad \ldots (9)$$

Note that B contains an image which has smooth borders.

4.3.2. Normalization

Normalization is a pre-processing stage where we can find new range from an existing one. It is helpful in prediction a lot. To maintain the large variation of the prediction, normalization technique is required to make them closer by structuring the given dataset in well-organized manner. As the next pre-processing step, our proposed model's input data is normalized. According to Andrew .et.al [47], normalization of input data can speed up the training process. Normalization is done on the image by applying the linear scaling to squeeze all pixels of the binarized image to values between 0 and 1. Dividing each input image by its standard deviation squeezes the pixel values from pixel values between 0 and 255 to between 0 and 1 [48]. Generally, applying normalization on an image makes optimization a little bit easier, it makes all features to be in the same range and contribute equally, and it also avoids some contrast issues. To do normalization, we go through the following steps, as stated by [49]:

i. Select the range of data of any size

ii. Write the code to read the dataset

iii. scale down the range of data in between 0 and 1

iv. use the newly generated scaled down data for further processing

v. finish

4.3.3. Cropping and Resizing

Image cropping and resizing are the most popular image processing operations. After we have segmented the lungs part, the black background occupies more space and these should be ignored since they are not our interest. Cropping an image generates a rectangular region of interest from the segmented lung mask. Cropping makes us to focus a particular region of the image and discard areas of an image that contains less useful information. To crop the segmented image, we need to specify a pair of (*min_row, min_col*) and (*max_row, max_col*) pixel coordinates that defines the corners of the new, cropped image. Cropping reduces the image size but retains the pixels information. So to turn the cropped image size back to its original dimension, resizing is followed. Resizing is altering the image size without removing anything out. Resizing neither create details that did not exist nor loses information to/from the cropped image. To resize, we need to mention the width and height of the resulting new image. Our resizing rescales the cropped lung mask image back to 512x512 dimensions. Cropping and resizing should also be done on the ground truth nodule mask with similar pixel coordinates and dimension that are used in lung mask, to keep changes uniform between them.

4.4. Pulmonary Lung Nodule Detection and Segmentation

4.4.1. U-Net

The U-Net is a convolutional neural network architecture specifically designed by O.Ronneberger et al. [9] from the University of Freiburg to solve Biomedical Image Segmentation problems. U-Net is a 2D FCN architecture that is popular for biomedical image segmentation, which has consistent output, needs less training data and memory, computationally cheap, easy to train, and it can receive 1 or N number of images slice by slice as an input during training.

In this work, we developed a fully automated 2D model designed to detect and segment pulmonary nodules from lung CT images. We develop a modified version of U-Net model

adapted from the original architecture and train a segmentation model as described in [9] on LUNA16 dataset.

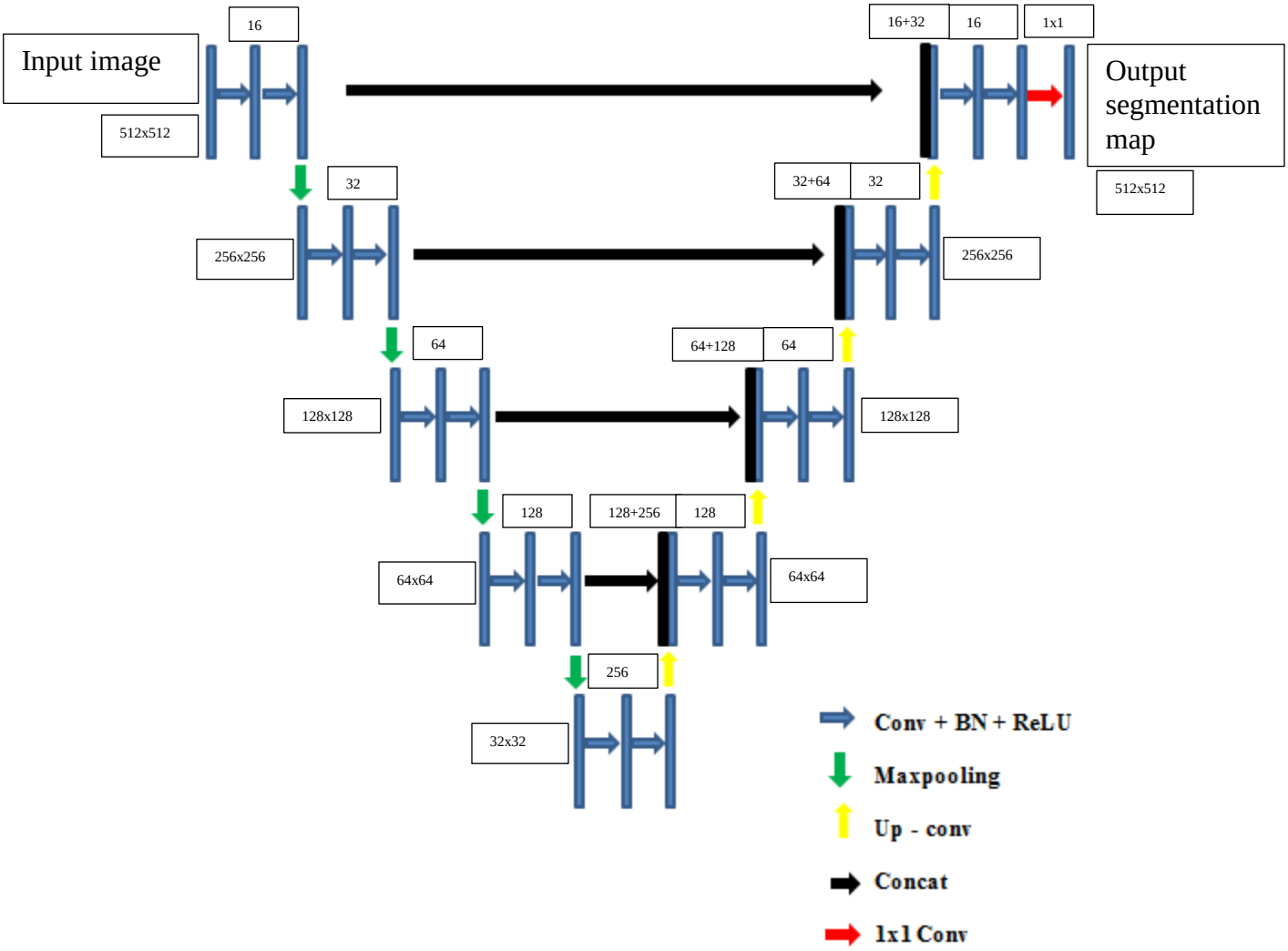


Figure 4.5 Modified U-Net model with a depth of 4 for pulmonary nodule detection and segmentation.

From Figure 4.5, the numbers at the top of each network block is the amount of feature maps and the numbers at the left and right of each network block is the size of feature maps in pixels.

4.4.2. Network Architecture

Our network architecture is illustrated on Figure 4.5. We feed the whole lung CT scan images and ground truth nodule masks as a single array of image slices that can be passed as input

channels to the model. The model a combination of contracting path (left side) and an expansive path (right side). The left side aims at extracting the features of an input image for classifying each pixel into pulmonary nodule and non-pulmonary nodule while the right side aims at locating pulmonary nodules more precisely, by combining with the contextual information from the left side. The contracting part is composed of four pairs of convolutional layers, each pair followed by a 2x2 max-pooling with stride of 2. The first layer of the network has 16 feature maps with an input resolution of 512x512. Every stage on the contracting path consists of a series of a 3x3 convolutions (conv 3x3), batch normalization (bn), rectified linear unit (ReLU) and forms a conv + bn + relu block. Two of conv + bn + relu blocks in series form a Conv block that doubles the amount of feature channels while the resolution of each consecutive convolutional layer pair reduced. The final max-pooling layer is followed by two more convolutional layers, and then by the symmetrically constructed up-sampling part, composed of four blocks of 3x3 up-sampling and pairs of convolutional layers. Similar to contracting stage, every step in expanding stage has a sequence of conv 3x3, bn, ReLU and a series of two consecutive blocks halves the amount of feature channels while the resolution of each consecutive convolutional layer pair increased. The images are up-sampled with a 3x3 up-convolution (transpose convolution) with stride 2. For up-sampling of images, a sequence of 3x3 up-convolution with stride 2, concatenations and a Conv block forms a de-conv block. The skip-connections, recognized as simple concatenations of feature maps, combine the output of an up-sampling layer with the corresponding features from the contracting part. The skip connections enables successive convolution layers in the up-sampling part to learn and produce better localized output, allowing the network to perform more accurate segmentations. All layers use ReLU non-linearity, except for the last layer where sigmoid is used to select the best scoring class. A final 1x1 convolution layer maps the 16 component feature vector to pulmonary nodule and non-nodule classes per pixel.

We follow the basic architecture proposed by [9], but we made some modifications in order to adapt it to our problem of pulmonary lung nodules segmentation:

i. Batch Normalization [33], dropout [34], and He weight initialization [35] is used in all layers to improve learning.
ii. We used 3x3 up-conv at the expanding path.

iii. The input size and the size of further layers are modified to match the size of our input images.

iv. The amount of feature maps in all layers is decreased by a factor of 4 compared to the original architecture, in order to speed-up the learning process and enables the model to fit in the memory of the GPU we are currently using.

4.5. Post-Processing

To enhance the performance of our network and/or to enhance the results of our initial segmentation, further processing may be essential which will be done in the post processing stage. We propose a post-processing technique using average ensemble learning method that can create durable segmentation nodule locations, removes false positives and false negatives from the segmentation results. Average ensemble learning method integrates the predictions from multiple models or network architectures and then take their mean as a final prediction to get better predictive performance than that could be obtained by any single model [50] [51]. Our average ensemble learning method combines three different network architectures with various hyper parameter values of a U-Net model. Ensemble learning methods, in general can basically improve the prediction performance and the robustness of a segmentation method [52]. Hence our average ensemble learning method can enhance the robustness and the stability of pulmonary lung nodules segmentation from lung CT scan images. Table 4-2 contains detailed hyper parameter values of the three different network architectures of a U-Net model.

Model	Network depth	Convolutional Layers	Input image dimension	Input feature maps at first layer	Learning rate	Dropout rate	Weight Initialization
U-Net1	4	23	512x512	8	0.01	0.2	He_normal
U-Net2	3	18	512x512	6	0.001	0.5	He_normal
U-Net3	2	13	512x512	6	0.00001	0.5	He_normal

Table 4-2 Detailed hyper parameters values of the three U-Nets used at the post processing stage

4.6. Implementation

The proposed system is implemented in Python language, Anaconda platform, Keras framework with plaidml backend, and JetBrains PyCharm IDE.

- **Python** is a high level interpreted language used for general purpose programming. It is widely used for scientific computing and can be used for a wide variety of general tasks from data mining to software development. Python is the main language used in this thesis.
- **Anaconda** is a popular data science platform where we can create data science projects and machine learning. Libraries such as NumPy, Pandas, Matplotlib, PlaidML, etc and IDE's such as PyCharm, Jupyter Notebook, Spyder, etc come with Anaconda.
- **PlaidML** is a deep learning software framework that supports GPUs from diverse hardware dealers. PlaidML uses OpenCL to run on GPUs manufactured by NVIDIA, AMD, or Intel, and acts as the backend for Keras to support deep learning programs.
- **Keras** is a deep learning framework/library written in Python used for easy and fast prototyping and has the ability of running on top of TensorFlow, Theano, or PlaidML. It was made with a target for the capability of fast experimentation and supports convolutional networks and recurrent networks; as well as combinations of the two.
- **PyCharm** is open source software IDE for python and scientific development that allows developers to create and share documents. The Anaconda distribution comes with a variety of software that includes PyCharm for scientific computing.

We run all the experiments on a personal laptop windows machine running windows 8.1 equipped with 6 GB of memory, Intel(R) Core(TM) i5-4210U CPU clocking at 2.40GHz. The U-Net training and testing was carried out on a 2 GB AMD Radeon HD 8500M GPU.

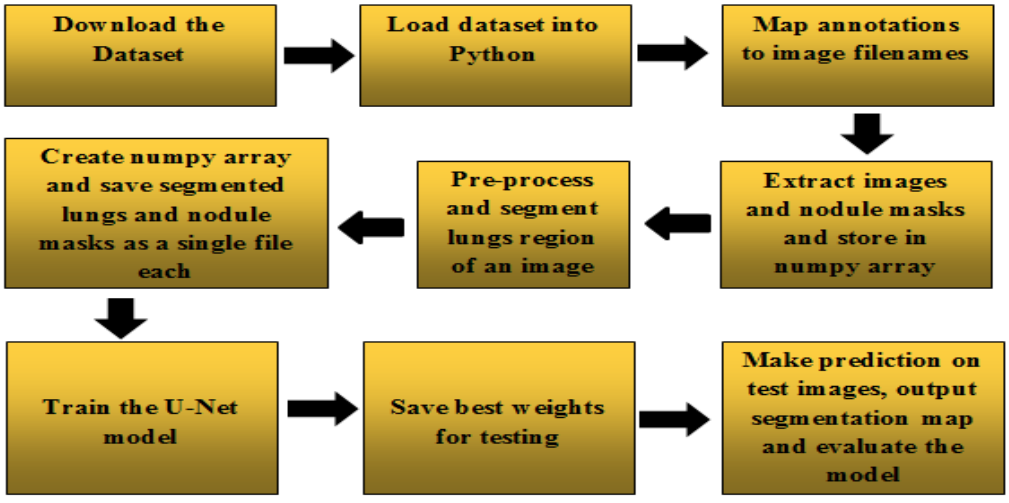


Figure 4.6 Block diagram of our implementation details

4.6.1. Mapping Annotations and Extracting Lung CT Images and Nodule Masks

First we installed anaconda and then import the necessary libraries for reading, processing and visualization of data. We started our work by loading the lung CT image and the ground truth nodules information. The lung CT images are given in 10 different subset folders and SimpleITK is one of those useful anaconda libraries used for reading .mhd files stored in LUNA16 database. The lung CT image data is given as a flexible stack of 512 x 512 arrays. The nodule location information is an (X, Y, Z) coordinate in millimeter using a coordinate system defined in the .mhd file and these nodule locations are given relative to the coordinate system defined and generated by the CT scanner. While reading lung CT images, only those files that have nodule information will be loaded and we mapped each nodule to each corresponding patient scan slices. Nodule masks will be created for those slices based on the nodule information given in *annotations.csv*. The nodule information i.e. the nodules locations given in (X, Y, Z) coordinate in millimeter and the size given in diameter_mm is stored in .csv file. We used *cs* library to read annotations file and a *pandas DataFrame* to keep track of the case numbers with their corresponding nodule information. Since the nodules locations are given in real world coordinate system, to find the corresponding voxel coordinates of a nodule, we used getOrigin() and getSpacing() method of the itk image object.

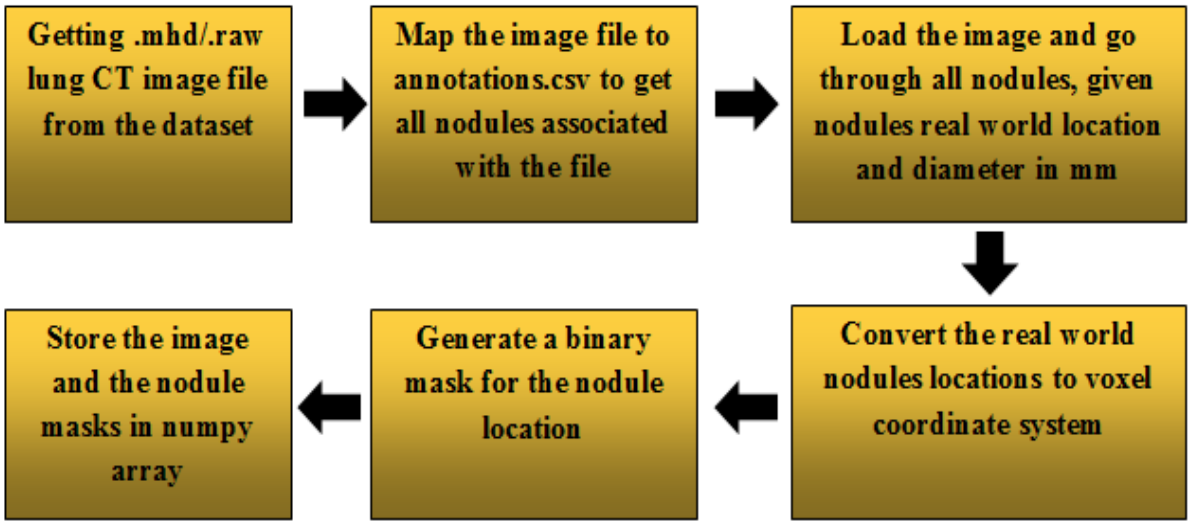


Figure 4.7 Block diagram to map and extract lung CT images and nodule masks

The output is therefore 2 files of each patient scan i.e. a set of lung CT images and a set of corresponding ground truth nodule masks. *Numpy* is a library required to be imported to store those files in numpy array.

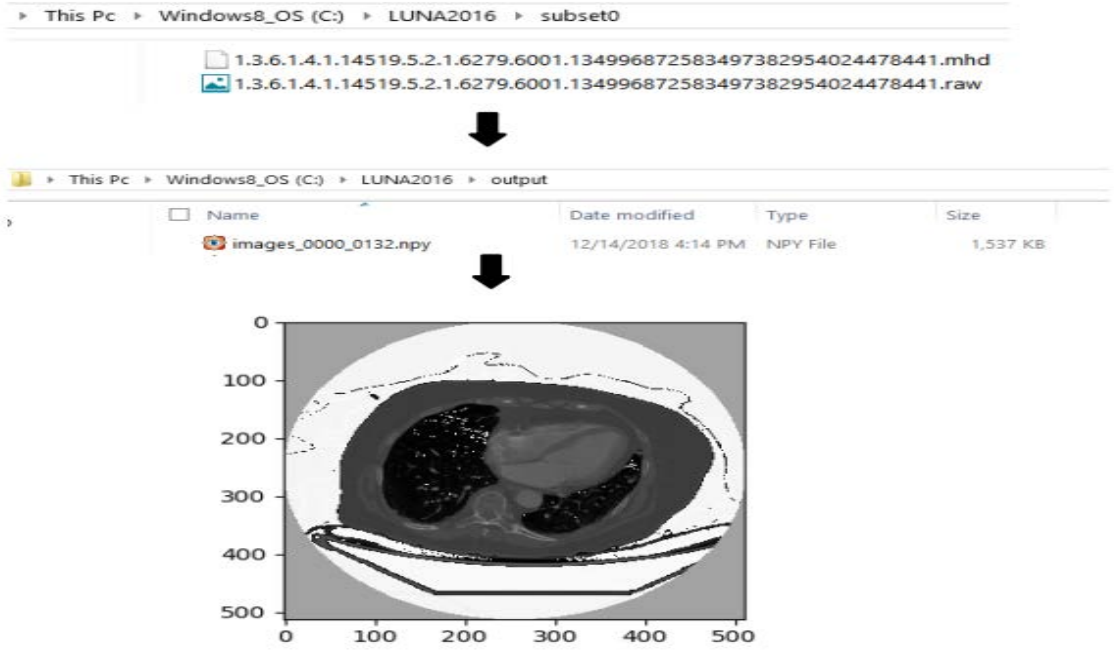


Figure 4.8 Sample raw lung CT image extracted from LUNA16 dataset

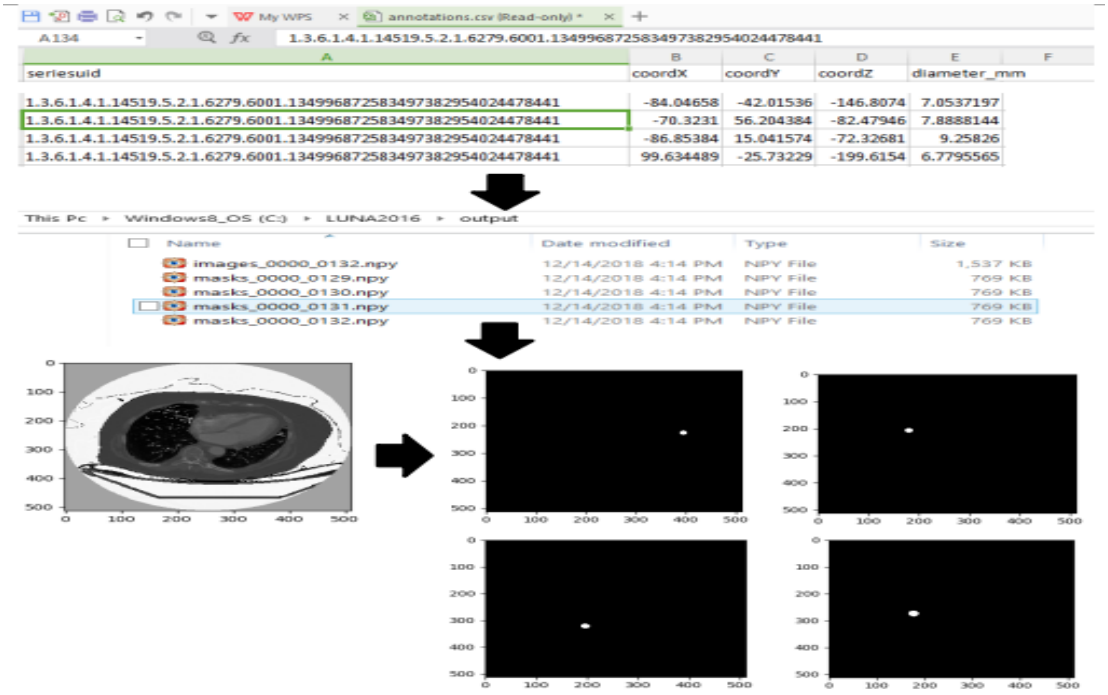


Figure 4.9 Sample raw lung CT image with its annotated binary nodule masks

4.6.2. Lung Segmentation and Pre-processing

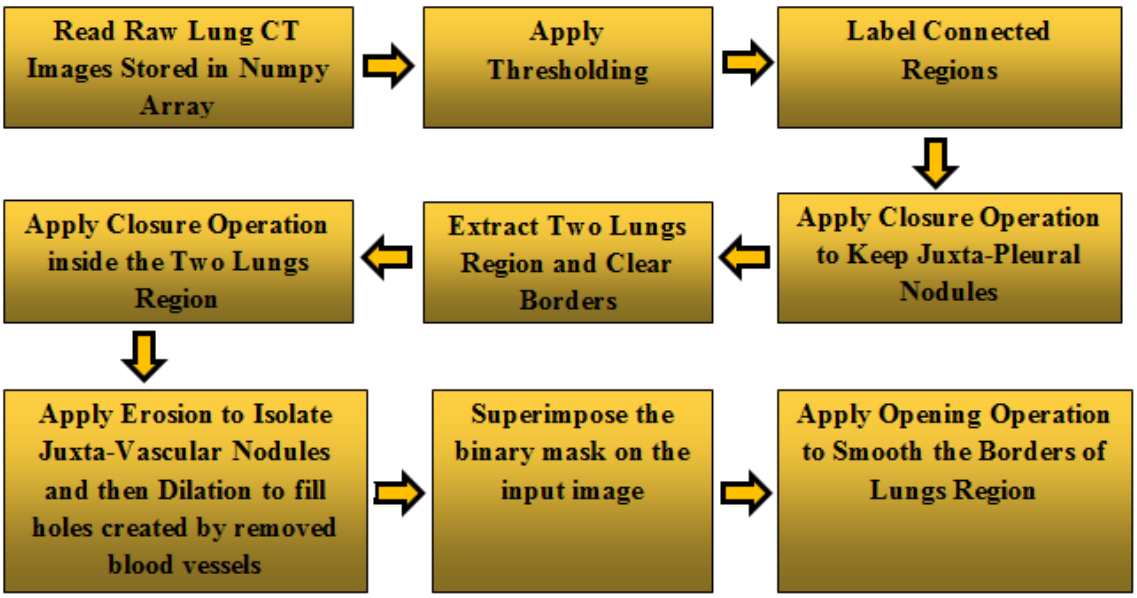


Figure 4.10 Block diagram of the proposed lung segmentation algorithm.

After extracting lung CT images and binary nodule masks, the first step in pre-processing is the segmentation of lung structures. This is because it is obvious that the region of interest lies within the lungs region. It is visible that the lungs are the darker regions in the CT Scans. The bright regions within the lungs are either the blood vessels or air. For this task, in addition to previous imports, we'll need to import some skimage image processing modules such as morphology, measure, KMeans, resize, and disk.

Thresholding is done first in lungs region segmentation to classify the regions of the image as foreground and background. Thresholding converts a given lung CT image into binary. To perform thresholding, we should first find the average HU (pixel) values near the lungs to take a look at the intensity distribution. We found an average threshold of value of -420HU (604) to be used at all places, because it was found in experiments that it works just fine [45]. But different scanners may have datasets with different intensity distribution. So to improve the thresholding finding, we moved our underflow and overflow over the pixel spectrum. Now by using this threshold value we applied KMeans with k=2 on the center of the image to separate non lung (radio opaque tissues) from lung (radio transparent tissues). Then we are going to label each region and obtain the region properties. The connected component operation takes in a binary

image to produce a labeled image and each region's pixel has an integer label of either the background or a component. Finding the lung nodule regions is a very hard problem. This is because there are nodules that are directly attached with blood vessels and some of them may present at the boundary of the lungs region. During thresholding we removed the non-lung parts of the image by separating the radio-opaque tissues from the radio-transparent tissues .i.e. the lungs. When we remove the radio-opaque tissue pixels, some pixels of nodules and bronchioles found at the lungs boundary which have similar intensity with them may be removed mistakenly and holes may be created. Since juxta_pleural nodules found at the lungs boundary, these removed pixels may represent them. Therefore, we should come up a solution that can preserve these pixels. For this purpose we applied smaller size morphological closing operation on the labeled image. The next step in lungs region segmentation is extracting the left and right lung regions from the labeled regions. The two largest regions can be extracted by applying a bounding box that is too large in either dimension. This bounding box should be applied on the center of the image since the lungs are far away from the upper and the lower end of the image. This also enables us to clear the borders by removing any regions that are too close to the top and bottom of the image. The bounding box should also be used together with an erosion operation to remove those regions found in between or around the two lungs region and got extracted along with the two largest areas. Similarly, after extracting the two largest regions, there may exists holes created due to mistakenly classified lung parenchyma pixels of nodules and bronchioles having similar intensity with the radio-opaque tissues as non-lung pixels. We proposed larger size morphological closing operation to be done on the two largest regions to fill holes appeared on the lung parenchyma. This way we can preserve mistakenly classified pixels to be part of the lung structure. Still there are lots of noises because of blood vessels. Juxta-vascular nodules are the nodules attached with blood vessels. So we have to isolate these nodules from the attached blood vessels. As a solution we proposed a morphological erosion operation to make the lungs region engulf vessels attached to nodules. Erosion also helps us to remove graininess from the lungs region. Then morphological dilation operation should be followed to fill the holes created during the removal of blood vessels. Finally, we superimposed the binary mask on the original image followed by morphological opening operation to smooth the borders of the lung mask. Smoothing is required to remove those objects got extracted along with the lung wall at the boundary of the segmented lungs region which are mistakenly considered as part of the lung

mask. We segment lung structures and tried not to miss the possible regions of interest from each slice of the CT scan image of all patients.

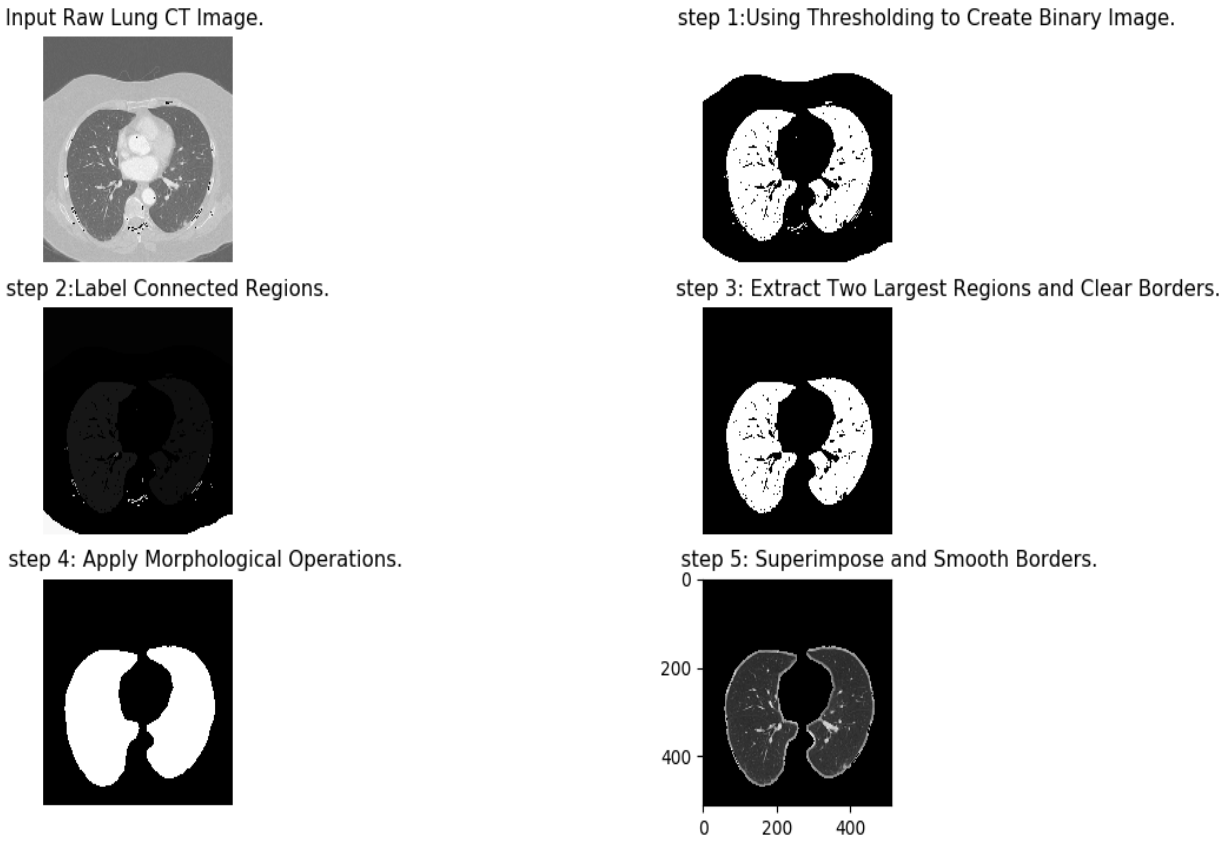


Figure 4.11 Step by step lungs region segmentation algorithm applied on sample image

Once we have the segmented lung mask, then we perform some final pixel normalization on the segmented lungs region. Normalization finds the mean and standard deviation of the masked region and then subtracts the mean from each masked image followed by dividing each resultant masked image by its standard deviation to squeeze pixels of the segmented image to be between 0 and 1. In addition, normalization sends the background pixels to the lower end of the pixel distribution. This is due to the fact that the mask sends the non ROI in the image to 0, and this operation is not sensitive to the pixel value distribution.

As a final pre-processing step, we performed cropping and resizing on the segmented and normalized lung mask. We cropped the segmented lung mask using the specified coordinates of (*512, 512*) and (*0, 0*) as a bounding square of the lungs ROI. Then we resized the cropped image back to *512* X *512* dimensional images.

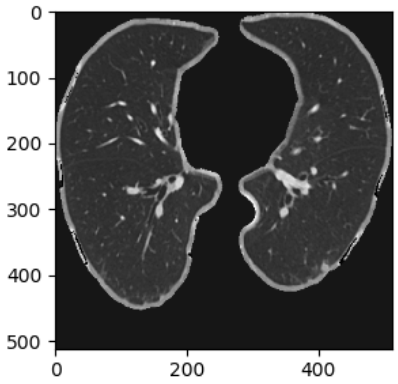


Figure 4.12 Segmented, Normalized, Cropped, and Resized sample lung CT image

The final product is a set of lungs that is ready to be compiled into our training set. Totally, we do have 601 patients' lung CT scan images and 1186 pulmonary nodules are labeled across them. The subset of patients' segmented and pre-processed lung CT scan images and their respective trimmed and rescaled annotated nodule masks are randomized and sent to a single file as a one channel numpy array to be taken as an input for our network model.

4.6.3. Training, Testing and Evaluating

After pre-processing the input images, the next task is to train the model for segmentation. The model is written in keras. It takes a 2D slice as input and returns a 2D slice of the same size as an output. Figure 4.13 shows how training, testing and evaluating is done on a U-Net model.

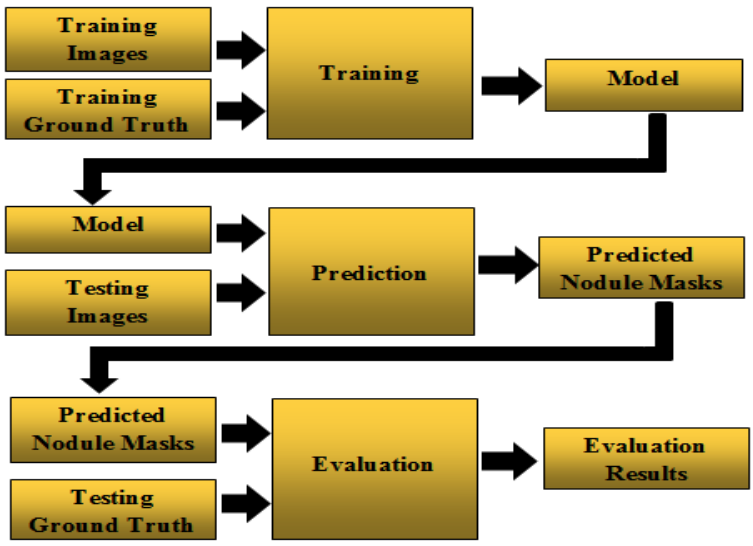


Figure 4.13 Training, testing, and evaluating our U-Net model

During training, our U-Net model takes an array of 512x512 lung CT images and their respective ground truth nodule masks to train slice by slice. The nodule mask is provided with size of 512 x 512 where nodule pixels are 1 and rest are 0. The input images and their respective ground truth nodule masks are used to train the network with Adam optimization algorithm and dice loss function.

Our network models have 2, 3, and 4 layers deep with 13, 18, and 23 convolutional layers and trained for many numbers of epochs using Adam optimization algorithm with a batch size of 1. During training, the model will take training set images with their corresponding nodule masks. The model is trained to output the same shape like an input images of 512 x 512 where each pixels of the output has a value between 0 and 1 indicating the probability the pixel belongs to a nodule or not. This is done by taking the slice with the corresponding ground truth nodule masks to the 1x1 conv layer of the final U-Net layer. During training, best models are checkpointed and stored for testing. Once we completed the training, we will load the best model and the testing set images. During testing, the final output of the network is the same shape as the input, but each pixel of the output, rather than containing visual information, contains the predicted probability that the corresponding pixel in the input belongs to the nodule. We generate our predicted segmentation by binarizing the predicted probabilities. Pixels having a score map of greater than 0.5 rounded up to 1 and are labeled as pulmonary nodule and pixels having a score map of less than or equal to 0.5 rounded down to 0 and are labeled as non-pulmonary nodule [53]. Finally, we will compare each pixels of automatically generated nodule mask by our model with the corresponding ground truth nodule mask to evaluate the effectiveness and the achievements of the proposed system. During training and testing, to check the success of the proposed model in detecting and segmenting the exact pulmonary nodules and their locations, we incorporated 8 slices without nodule region in addition to those slices having nodules. Ideally the output of U-Net would give us the exact nodules and their location.

4.6.4. Post processing

Pulmonary nodules segmentation performance from lung CT images using a single U-Net network can be further enhanced by applying a post processing technique. Therefore, we proposed a new post processing mechanism using average ensemble learning method (MBM).

For this task, we used three U-Nets having different network architectures with various hyper parameter values. With MBM method, we first binarized the score maps of the predicted outputs from each network and then averaged. Both binarization and averaging were performed pixel-wisely. The output of the averaging operation is the mean value of the three input pixel values, and the binarization is conducted like equation (10).

$$\text{binary mask}(x,y) = \begin{cases} 1, & \text{score map}(x,y) > 0.5 \\ 0, & \text{score map}(x,y) \leq 0.5 \end{cases} \quad \ldots (10)$$

Where score map(x, y) represents the pixel value at the specific location (x, y) in score map, which is the result of a U-Net in the testing stage. And binary mask(x, y) represents the pixel value of the binary mask at the respective position that can be taken as the segmentation results [53].

Table 4-3 shows how average ensemble learning (MBM) works. We have used three different network architectures with various hyper parameter values of a U-Net model, for MBM technique. For example, let us assume the score values at a specific pixel obtained from the three U-Nets U-Net1, U-Net2, and U-Net3 are 0.8, 0.6, and 0.4 respectively. A possible situation is that both U-Net1 and U-Net2 have converged to a normal optimum and have lower error rate whereas U-Net3 suffered an abnormal local optimum and has higher error rate. In this case, average ensemble learning method (MBM) combines the decisions from these three different network architectures having various hyper parameter values of a U-Net model to reduce the false negative rate caused by the abnormal model and improves the overall segmentation performance. Average ensemble learning method takes advantage from these diverse network architectures and hyper parameter values of a U-Net model to produce segmentation map which is better than a single best performing network architecture of the model.

U-Net	**1**	**2**	**3**
Score	0.8	0.6	0.4
Threshold	1	1	0
Average	0.6667		
MBM	1		

Table 4-3 An example demonstrating how average ensemble learning (MBM) works

4.7. Evaluation Metrics

4.7.1. Similarity Metric

As a similarity evaluation metric, we used DSC, to measure an overlap between the ground truth (GT) and the predicted segmentation mask (S).

At training, we apply a pixel-wise sigmoid activation in the final layer of the model to get the predicted probabilities p((x,y); i) for each class i at each pixel (x,y). The targets at location (x,y) is denoted by t(x,y) [54]. To evaluate the accuracy of an automatically segmented pulmonary nodule region, we employed the DSC score, which is also known as the similarity index. DSC is a metric used for measuring the similarity of two sets, which is defined in equation (11).

$$\mathrm{DSC}(\mathrm{GT}, \mathrm{S}) = \frac{2|\mathrm{GT} \cap \mathrm{S}|}{|\mathrm{GT}| + |\mathrm{S}|} \quad \ldots (11)$$

Where GT represents the nodule region of the ground truth segmentation of a pulmonary nodule, S represents the nodule region covered by the automatic segmentation, and $|\mathrm{GT} \cap \mathrm{S}|$ refers to the overlap nodule region between ground truth and automatic segmentation [55] [56].

Loss functions measures the difference by comparing ground truth with the corresponding predictions that our network produced using the current parameters. Loss functions sometimes also known as cost function, objective function, or criterion.

Extending DSC score as a loss function, improves the performance when dealing with situations where background pixels are higher than the ground truth pixels [54]. Generally, in order to maximize the dice score, the negative loss should be minimized using equation (12).

$$\text{dice loss function} = -\mathrm{DSC} \quad \ldots (12)$$

4.7.2. Classic Measurements

As a classic measurement metric, we used Sensitivity, Specificity, and Accuracy.

Sensitivity - we used sensitivity to measure the portion of positive pixels in the ground truth that are also identified as positive by our system. The sensitivity is the percentage of nodule region

recognized by our segmentation method. It is used to check the algorithm's sensitivity for segmenting proper pulmonary lung nodule pixels. The Sensitivity of our model can be calculated by equation (13).

$$\text{Sensitivity}(GT, S) = \frac{|GT \cap S|}{|GT|} \quad \ldots (13)$$

Where GT represents the nodule region of the ground truth segmentation, S represents the nodule region covered by the automatic segmentation, and $|GT \cap S|$ refers to the overlap nodule region between ground truth and automatic segmentation [55].

Specificity - we used specificity to measure the portion of negative pixels in the ground truth that are also recognized as negative by our system. The specificity is the percentage of non- nodule region recognized by our segmentation method. It tests how sensitive the algorithm is to the segmentation of correct non-nodule pixels. The specificity of our model can be calculated by equation (14).

$$\text{Specificity}(GT, S) = \frac{\left|\overline{GT} \cap \overline{S}\right|}{\left|\overline{GT}\right|} \quad \ldots (14)$$

Where $\overline{GT}$ represents the non-nodule region covered by the ground truth segmentation, $\overline{S}$ represents the non-nodule region covered by the automatic segmentation, and $\left|\overline{GT} \cap \overline{S}\right|$ refers to the non-nodule region covered by both the ground truth and automatic segmentation maps [57].

Accuracy - we used accuracy to measure the portions of the positive and negative pixels in the given image that are also recognized as positive and negative by our system, accordingly. The accuracy is the percentage of nodule and non- nodule region recognized by our segmentation method. It tests how sensitive the algorithm is to the segmentation of correct nodule and non-nodule pixels.

$$\text{Accuracy}(GT, S) = \frac{|GT \cap S| + \left|\overline{GT} \cap \overline{S}\right|}{|GT| + \left|\overline{GT}\right|} \quad \ldots (15)$$

Where GT represents the nodule region of ground truth segmentation, S represents the nodule region covered by the automatic segmentation, $|GT \cap S|$ refers to the overlap nodule region between ground truth and automatic segmentation. $\overline{\mathrm{GT}}$ refers the non-nodule region of the ground truth segmentation, $\overline{\mathrm{S}}$ refers the non-nodule region of the automatic segmentation, $\left|\overline{\mathrm{GT}} \cap \overline{\mathrm{S}}\right|$ refers overlap non-nodule region between the ground truth and automatic segmentation. The denominator refers all pixels from the ground truth and automatic segmentation [56].

4.8. Summary

In this chapter of the thesis, we tried to show the design and implementation of a system for automatic detection and segmentation of pulmonary nodules. To do that first the images should be downloaded from the publicly available database; in our case we used the data from LUNA16 dataset. These images have a file format of .mhd and we have converted to .npy file format for further processing. Then we reduced our ROI by removing unwanted sections of the lung CT scan images. For this task, we proposed a lung segmentation algorithm called thresholding and morphological operators. This lungs region segmentation algorithm performed good in accurately segmenting the lungs boundary. The algorithm tried to effectively isolate those nodules connected with blood vessels and the lung wall. The lungs region segmentation algorithm is then followed by normalization, cropping, and resizing. Normalization squeezes the pixels of the segmented lung mask to be between 0 and 1 while cropping is used to remove portion of a segmented lung mask that contains less useful information. And finally, we combined the proposed lungs region segmentation algorithm with U-Net model. Our U-Net model is adopted from the model of [9] with some modifications. Our proposed network architectures have depths of two, three and four with 13, 18 and 23 convolutional layers, respectively. For further performance improvement, we combined the outputs from these three U-Net models having different network architectures and hyper parameter values. This is done by using average ensemble learning method (MBM) as a post processing technique. MBM binarizes the outputs from these networks based on a threshold value of 0.5, taking their mean followed by binarization. At the post processing stage, MBM combines the outputs from three different network architectures with various hyper parameter values of a U-Net model to gain a plus from their variation in prediction and to get the overall performance improved.

Chapter Five

5. Results and Discussion

5.1. Introduction

This chapter comprises the details of experimental procedures to setup hyper parameter values and the experimental test results are discussed in detail. A large set of experimentations were done to proof the performance and the effectiveness of the proposed approach. In the results section, the dataset properties and experimental setups used in training and testing of the models are elaborated clearly. Under discussion section, we compared our final test results with the results of other techniques using similarity and classical evaluation metrics. The response towards the achievement of the research questions are also assessed under discussion section.

5.2. Results

5.2.1. Dataset properties

For our training and testing, we used LUNA16 dataset. The initial dataset properties are described on Table 5-1.

File Format	Patient CT scans with nodules	Annotated nodules by at least three radiologists	Slice thickness and Nodule size	Pixel Dimension
MHD /RAW	601 out of 888	1186	Less than 2.5mm thickness, 3-30 mm in diameter,	512 X 512

Table 5-1 LUNA16 dataset properties

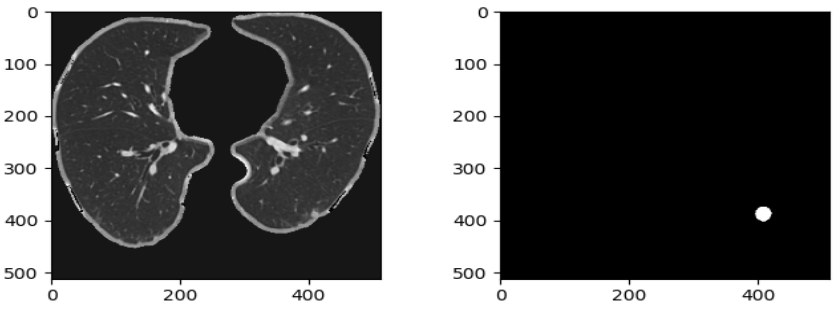


Figure 5.1 Sample processed image and nodule mask (Ground Truth) for input to the model.

5.2.2. Train and Test set

Due to our machines memory constraints, we used 858 images and ground truth nodule masks from the available dataset, as described on Table 5-1. And from these selected dataset, 80% was used as training set and 20% was kept as a testing set. The training set and the testing set images and their respective nodule masks are selected randomly. During training, from these 80% of the training set, we divided it to use 80% of them for training and 20% for validation. The test set is completely kept out of the training process. The choice of this ratio is based on the amount of data available for this research. When we analyze the available data, it shows us that there is much variance in the size and location of the nodules. Therefore the model gets enough training examples and testing data to generalize well and to evaluate the model after training. Therefore, the proposed ratio is a reasonable choice. Figure 5.2 shows the dataset classification graphically.

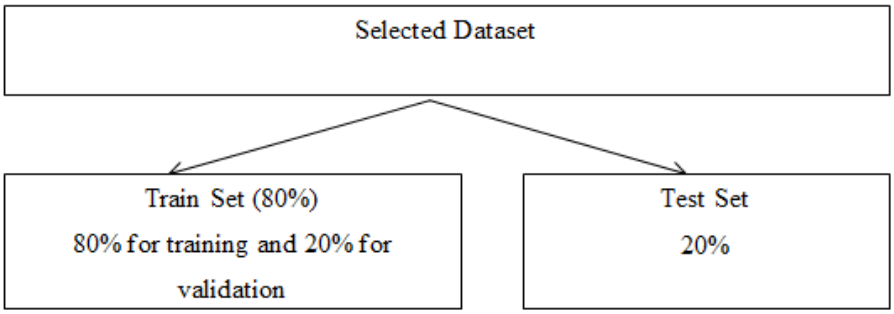


Figure 5.2 splitting the selected dataset into train, validation and test set

The train set is used to train and validate the proposed model whereas the test set is used to test our segmenter. Generally, we used 549 slices with their corresponding ground truth nodule masks for training, 138 slices with their corresponding ground truth nodule masks for validation, and 171 slices with their corresponding ground truth nodule masks for testing.

5.2.3. Experimental Setups

In this section, we present the results of the proposed method and the values of different hyper parameters are determined experimentally.

To train and test the proposed automatic CAD system scientifically, an experiment has been carried out. We have specified different system's hyper parameters and their values are selected

scientifically based on experimental procedure and this is a critical issue to design and develop a system model that has great performance. System performance is strongly dependent and has a great relationship with network parameter selection. Since there are many hyper parameters and infinite possible situations to test in each phase, we chose best values for each hyper parameters of the network architecture of the model. Recommended Values of the hyper parameters can be obtained from the papers and/or their values can also be decided from routine experiments. In this proposed system different parameters were tested, and the model with smallest loss or error rate is selected. The goal of testing different hyper parameters is to decrease the percentage of loss function while improving the performance values for our evaluation metrics.

Dataset size: The first experiment was related to the size of the labeled dataset. Naturally, the more data we have the better result we will get but small initial data will produce wrong predictions. But, due to our memory constraint, we used 858 images out of the available dataset.

Batch size: During training and testing of the model, we do not change the value of the batch size. This is because of our machine can support only a single batch size. The batch size selection is strongly dependent on the size of the physical memory of GPU installed on our machine. Even though we used a batch size of 1 due to our smaller size of GPU, it has an advantage for our model to have smaller fluctuations.

Evaluation metrics: In statistical analysis, the DSC, the accuracy (Acc), the sensitivity (Sen), and the specificity (Spe) were measured. The values of these metrics varied from 0 to 1 where 0 indicate whole disagreement and 1 indicates whole agreement between the automatic prediction and the ground truth. Therefore, the higher the values of these metrics indicate that the system is performed significantly well.

Optimizer, learning rate and regularizer: Initially, we configured our U-Net model to have depth of three. This is a different architecture from the generalized version of U-Net since it is possible to have a U-Net of greater or smaller depth. The reason why we want to test a U-Net of smaller depth is to see if we can find best performance with limited memory expense. Our network was trained with Adam optimizer. An Adam optimizer was used with a learning rate of 0.01. We decreased the learning rate value by a factor of 0.1 if the training loss gets plateau for the number of 2 epochs. We integrated this learning rate with a regularizer called dropout to

incorporate into our model with a probability of 0.2. In addition, at each convolutional layer, batch normalization was added to improve the accuracy and enable faster convergence for our system.

Weight Initialization	Regularizer	Optimizer	Batch size	Learning Rate	Epochs	Mean DSC
Random	dropout(0.2)	Adam	1	0.01	5	0.35

Table 5-2 Measured DSC value from U-Net model of depth three

Weights initialization: Given the network architecture and optimizer, it is interesting to see whether it is possible to optimize the model. Instead of using random weight initialization for our model, we tested the network with different weight initialization. The weights of the network can be initialized using He initialization technique which is another form of random initialization.

As we can see from Table 5-2 above, the overlap accuracy measured by DSC is not that much satisfactory and we need to make a change on the values of different hyper parameters to check their impact on the model's performance.

Therefore, to see the effect of changing the values of different hyper parameters, we did our next experiment by increasing the amount of epochs, by changing the optimization method, by changing the dropout probability, and by changing the learning rates but still did not have an effect on the overall performance of the proposed system. However, as we can see later on the last row of the table presented in the discussion section, the DSC value got an improvement. This means that the way we are initializing our network matters. Therefore, the drawbacks of normal random weight initialization can be reduced by another initialization mechanism called He initialization. Typically, He is a type of weight initialization technique used mostly with ReLU activation function [35]. With proper hyper parameter values setting, let us train our model and plot the learning curve for multiple numbers of epochs.

Figure 5.3 U-Net learning curve with a network depth of three

As shown on Figure 5.3, the amount of epochs now has an effect on the DSC value when proper hyper parameter values are set. With best model setting, DSC increases while decreasing the loss until the model convergence.

Depth: In order to confirm whether the deeper version of U-Net actually helps for this segmentation problem, the results found from the U-Net models with depth of three and four are compared with each other. A depth of more than four was not possible due to the smaller size of our GPU memory.

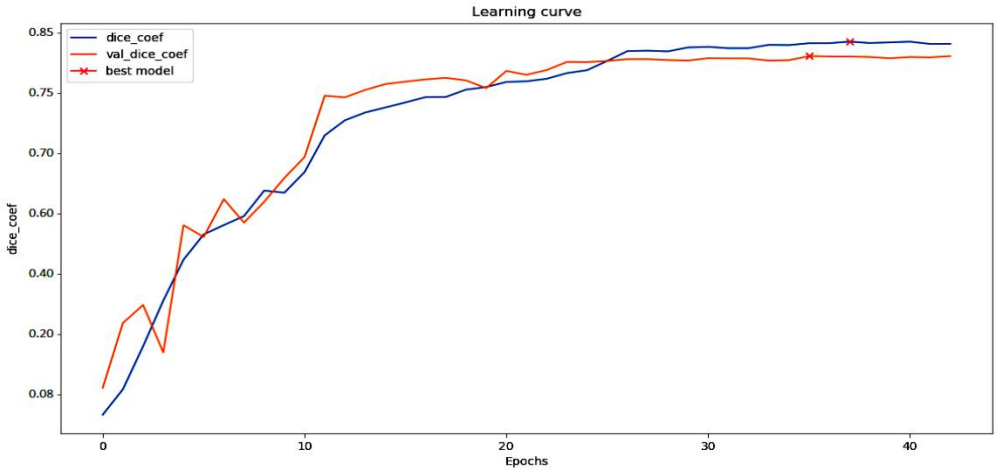
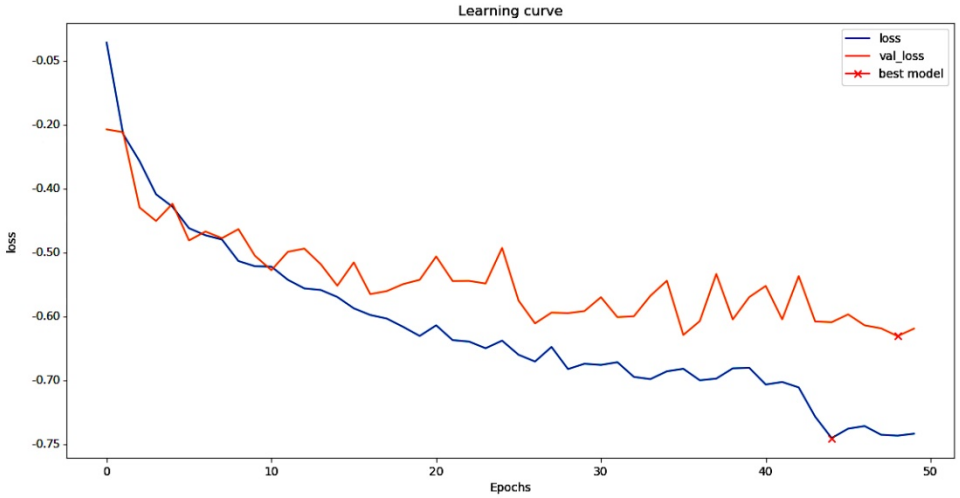
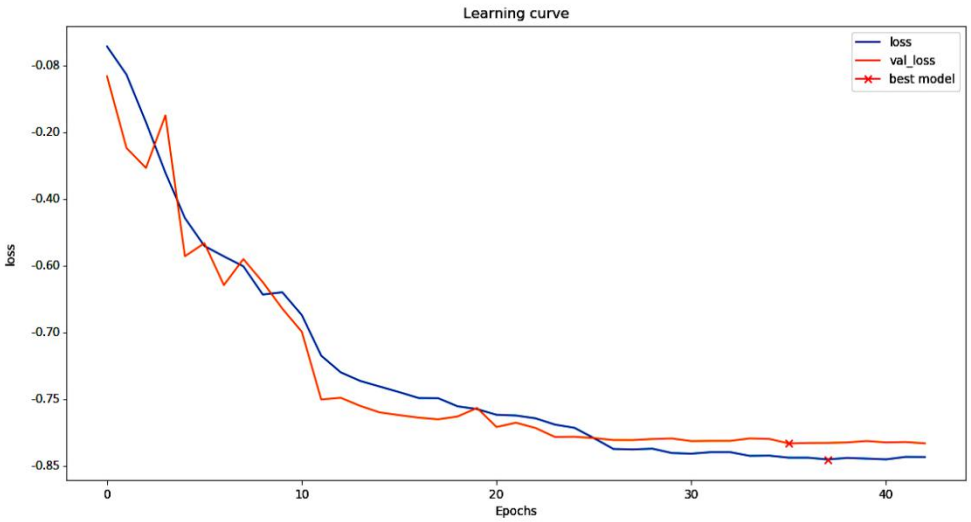


Figure 5.4 U-Net learning curve with a network depth of four

Comparing Figures 5.3 and 5.4, increasing the depth of the model from three to four improved the overall segmentation performance. But this may not always be true. The depth of the network should be compatible with the size of the dataset. When we increase the depth of the network, the size of the dataset should be considered. Deepening the network with small size of dataset can't bring improvement on performance.



a) A network with depth of three



b) A network with depth of four

Figure 5.5 dice loss per epoch for our U-Net models a) with depth of three and b) with depth of four

To further improve our models performance, we applied a post processing mechanism using average ensemble learning method (MBM). For this task we used three U-Nets which have different network architectures and hyper parameter values, as specified on Table 4-2. We trained and tested each model sequentially and then combine their predicted outputs using a post processing technique MBM to produce the final prediction.

To show the contribution of the post processing stage, we compared the DSC values obtained from the best performing single U-Net model with MBM using box plots. We recorded the mean DSC values at 5^{th}, 10^{th}, 15^{th}, 20^{th}, 25^{th}, 30^{th}, 35^{th}, and 40^{th} epochs obtained from best single U-Net model and MBM. Note that training and testing is done separately at each epoch on both of them to compare the DSC values obtained. In addition, we compared the DSC values of individual U-Nets with MBM at epoch 40.

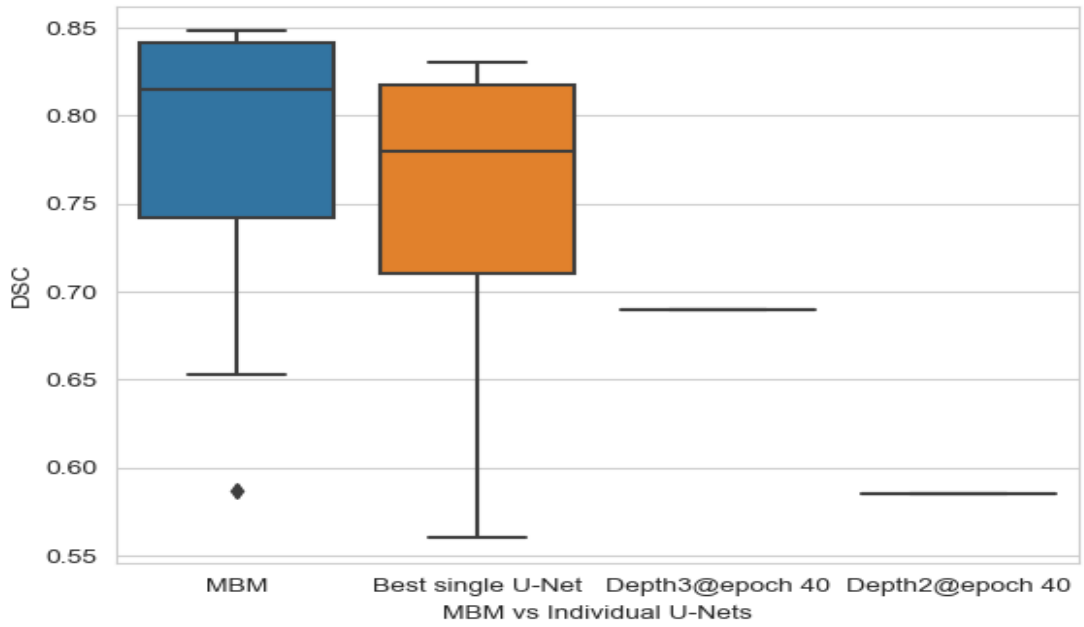


Figure 5.6 Comparing DSC values of individual U-Nets with MBM of three U-Nets

5.3. Discussion

Weight Initialization	Regularizer	Optimizer	Batch size	Learning Rate	Epochs	Mean DSC
Random	dropout(0.2)	Adam	1	0.01	5	0.35
Random	dropout(0.2)	Adam	1	0.01	10	0.35
Random	dropout(0.2)	SGD	1	0.001	5	0.35
Random	dropout(0.5)	SGD	1	0.00001	10	0.35
He-normal	dropout(0.2)	Adam	1	0.01	5	0.43

Table 5-3 The effects of changing optimizers, learning rates, dropout probability, number of epochs, and initialization method.

When training our network, the architecture was tested by varying certain hyper parameters. This is because as it can be seen on Table 5-2, by our initial network setting, the segmentation performance of our model was not satisfactory and a change on different hyper parameter values was needed. Therefore, a very thorough search has been done in order to get the best hyper parameter values. Initializations of parameters in deep neural networks have a big effect on the performance of the networks, as shown on Table 5-3. Recent deep CNNs are mostly initialized by normal random weights drawn from Gaussian distributions with fixed standard deviations [58]. But from Table 5-3, using normal random initialization of weights in neural networks is critical to learn good mappings from input to output. This is because of the search space involving many weights during training are very large and there are also multiple local minimums within which the back-propagation may be trapped. Normal random weight initialization makes the model to diverge or to converge slowly [59]. Therefore, a robust initialization method that removes an obstacle for training deep networks should be found. The initialization scheme developed by [35], enabled convolution activations to hold a constrained mean and this makes training of deep networks effective. As tabulated on the last raw of Table 5-3, a significant change occurred when we change the weight initialization method to He-normal. Using He weight initializations made the segmentation overlap to be increased. In this initialization method the weights are still random but vary in range based on the size of the preceding layer of neurons. The weights are initialized by considering the size of the preceding

layer, which helps in getting a global minimum of the loss function faster and more effective and also provides faster and more effective gradient descent.

On Figures 5.3 and 5.4, we tried to show the impact of increasing the depth of our U-Net. Figure 5.3 show us that a U-Net model of depth three gets its best value at epoch number 44 with a DSC value of 0.74 while Figure 5.4 shows a U-Net model of depth 4 gets its best value at epoch number 38 with a DSC value of 0.83. On both U-Net models, as can be seen on Figures 5.3 and 5.4, we controlled the training by configuring early stopping. The training will be stopped if the training loss never decreased for about 5 epochs. From those figures we can conclude that increasing the depth from three to four increases the performance of a segmentation overlap. But increasing the depth of the network needs more memory requirements that could result in surprising cost of training time and GPU resources. In order to tackle this problem we tried to use appropriate number of feature maps. When the network depth increases, the number of parameters to be trained will be increased and increasing the number of trainable parameters makes the prediction more accurate by generating better segmentation accuracy. In addition, increasing the depth of the U-Net model allows us to have the full context of the given lung CT scan image to be taken into account, and by having more convolutional layers, the deepen U-Net model allow us to recognize more complex patterns from input images. Figure 5.5 showed the comparison between a U-Net model of depth three and depth four based on their loss. As it can be seen, the loss value and the gap between the training loss value and the validation loss value are very small in case of a network having depth of four than three. In addition, the network converges at earlier epochs in a network having a depth of four than three. In both network architectures, the validation loss is greater than the training loss. But in the network having a depth of four, its validation loss is the lowest but slightly higher than its training loss, which indicates us that our training is free from over fitting and under fitting; it is in a good fit condition. Generally, from Figures 5.3, 5.4, and 5.5, the model with larger depth has got some advantages over the model with a smaller depth and provides better results.

When we tried to train and test our best network architecture of the U-Net model at different times, we found closer but different values (i.e. there are some fluctuations). So we should build a system that is more stable and we proposed an ensemble learning method at the post processing stage to further improve the overall performance. According to many papers, the winners of most

recent machine learning challenges are those who employed ensemble learning methods. Ensembling is a combination of multiple models whose separate predictions are integrated in some manner to create final prediction [50]. Ensembling can be done using homogenous or heterogeneous models [50] [51]. Homogeneous ensemble consists of members having a single-type base learning algorithm whereas; heterogeneous ensemble contains members having different base learning algorithms [51] [60] [61]. Experimental analyses demonstrate that both ensemble methods outperform single model predictions and heterogeneous ensemble is better [50] [51]. There are two main approaches for model combination: averaging and voting. Averaging is a type of ensembling mostly used for numeric output combinations [50]. We developed an average ensemble modeling (MBM) using three U-Nets. In order to introduce high variance between them, we configured different network architectures with various hyper-parameter values to be trained in diverse ways. Different networks execute differently, with behavior mostly affected by architectural choices and training settings. About the architectures, they differ in depth and number of filters. Models with large filter sizes may show better localization abilities but have lower sensitivity to fine texture than models focusing local information. The setting of hyper-parameters for the optimization can strongly affect performance. It is often observed by practitioners that the choice of the optimizer and its configuration, for instance the learning rate and the regularizations, can make differences in segmentation results [52]. For the same machine learning algorithm, we can enhance diversity by using different hyper parameter settings and subsets of variables [62]. Our average ensemble learning method combines these three different network architectures with various hyper parameter values of a U-Net model and it helps us to take advantage from their differences. These three networks are trained and tested sequentially with their variant hyper parameter values and their predicted pixel values are stored. With different network architecture and hyper parameter values, there is a possibility of different pixel values to be predicted on the same position. Then, our average ensemble learning method (MBM) combines the predicted pixel values of the three networks to generate a new and best segmentation map. MBM binarizes the predicted sore values from those three U-Nets using a threshold value of 0.5 and then taking their average followed by binarization. Figure 5.6 contains the recorded statistics of the results from individual U-Nets and MBM. As shown on the figure, the median of our best single U-Net model is lower than the median of MBM and the 50% values distribution of the measured performance

data of best single U-Net model is between 0.712 and 0.82 whereas, the 50% values distribution of MBM is between 0.74 and 0.842. In addition, at epoch 40 MBM achieved a mean DSC of 0.848, best single U-Net achieved a mean DSC of 0.83, depth 3 achieved a DSC of 0.69, and depth 2 achieved a mean DSC of 0.585. According to the statistics from Figure 5.6, we observed that the DSC values of MBM are better than the best single U-Net model. With one U-Net delivering lower segmentation results, MBM can weaken the influence of that lower-valued U-Net and deliver segmentation results that are superior to those obtained from the best single U-Net network architecture. This means that MBM improved the local optimum and instability problems of the model and it made our predicted outputs improved, more stable and robust by lowering the loss. Generally, from Figure 5.6 we can say that applying average ensemble learning method as a post processing technique using three different network architectures with various hyper parameter values of a U-Net model improved the overall performance of the segmentation system obtained from single best network architecture of a U-Net model by approximately 2%.

5.3.1. Results Comparison

In this section we attempt a comparison between our experimental results with the results reported by other research groups for segmenting pulmonary nodules. The performances of segmenting all types of pulmonary nodules were measured by mean values of DSC, accuracy, sensitivity, and specificity. Now let us have a report of different methods for the segmentation of pulmonary nodules from lung CT scan images, in addition to ours.

Authors	Method	Dataset	Number of images and ground truths	Nodule type	Mean Acc	Mean Sen	Mean Spe
Our	U-Nets with MBM	LUNA16	858	All	**0.965**	**0.826**	**0.983**
Bronmans	3-dimensional CNN	LIDC	1132	All	0.864	0.680	0.877

Table 5-4 Comparing the classic measurement results of different systems

Note that: the statistical values of the methods used for comparison on Table 5-4 and Table 5-5 are directly taken from the papers. i.e. we didn't implement them.

To compare the performance of our method with other techniques, we used the model proposed by Bronmans et al. [42]. Table 5-4 shows the performance comparison results in terms of accuracy, sensitivity, and specificity statistical measures. As shown on Table 5-4, our method achieved higher performance rate in terms of accuracy, sensitivity, and specificity compared to the other previous method. Better segmentation accuracy and specificity values scored by our proposed approach due to the larger number of background pixels as compared to the nodule pixels and the pre-processing method employed to minimize the region of interest. The value of specificity lies within a large band. Since accuracy calculates the segmentation performance of nodule and background pixels, the value of accuracy is also affected by specificity, therefore, accuracy also lies within a large band. The increased percentage of specificity greatly increases the percentage of accuracy, since accuracy is strongly dependent on specificity. But the accuracy and specificity values obtained from [42] is the lowest. This may be due to the fact that in the work of [42], there is no use of extensive pre-processing such as segmenting only the lungs from within the given raw CT input image.

Sensitivity is related to measuring the percentage of segmenting true nodule pixels. Approximately, 83% of the true nodule pixels of the lung CT image are correctly identified by our proposed system, which is relatively best compared to the other method presented on Table 5-4. The percentage of accuracy, sensitivity, and specificity is directly related to segmentation overlap. More overlapping between the predicted segmentation map and the ground truth segmentation map enhances the percentage of correctly identifying true nodule and true non-nodule pixels. The approach proposed by [42] is a type of deep learning algorithm which has a relatively good ability to identify true nodule pixels than the conventional shallow approach. Our proposed approach measures a mean sensitivity value of 0.826 which is the best and outperforms the previous method. Generally, Table 5-4 shows the desirable performances our proposed segmentation method scored using the given dataset in terms of accuracy, sensitivity, and specificity.

DSC is used to evaluate the quality of the model segmentation results. The DSC primarily calculates the similarity between the automatic segmentation map and the ground truth, and it is an evaluation function of the similarity between the sets. If the two sets are similar, the DSC will

be larger. Therefore, the larger the DSC, the more similar two images are, and the equivalent segmentation result is more accurate.

Authors	Method	Dataset	Number of images and ground truths	Nodule type	Mean DSC
Our	U-Nets with MBM	LUNA16	858	All	**0.848**
Wang	3D CNN-Encoding/Decoding	LIDC/IDRI	1018	All	0.739
Kubota	Morphological approaches and convexity models	LIDC	23	All	0.69

Table 5-5 Comparing the DSC value of different systems

From Table 5-5, the DSC value of [38] is the smallest of all other methods. Similar to our approach, the methods of [38], and [44] are applicable to segment different types, shapes, sizes, and locations of pulmonary nodules. The methods are differentiated by their segmentation overlap performance. Comparing the values presented from [38], [44], and our proposed method, the DSC value of [38] is the lowest and the value obtained from our proposed method is the highest. This is to mean that the works of [38] and [44] correctly identifies less number of true nodule and true non-nodule pixels and there exists many pixels that are incorrectly classified as non-nodule but actually nodule and vice versa. Our approach is able to identify many true nodule and true non-nodule pixels correctly than the works of [38] and [44]. The approach proposed by us and [44] is a type of deep learning algorithm. From the results of [38], and [44], we can conclude, the deep learning algorithm proposed by [44] outperforms the conventional method proposed by [38] with a DSC value of 0.739. Comparing the result obtained from our approach with the value obtained from [44], ours is better. Some of the reasons that make the approach proposed by [44] perform less efficiently includes; first the method proposed by [44] is patch based system which may have a chance of duplication and/or missing of some important global features, second unlike our system, the decoder part of the network developed by [44] used a fixed up-sampling technique and it is not efficient to extract and fill in useful features in the up-sampling operation when translating the input feature maps to output images, third the depth of the network developed by [44] is smaller than ours and by consensus the deeper the network the

better the result is. To conclude, there is a great improvement in segmentation overlap of our proposed approach. The DSC value of our method outperforms the methods of [38], and [44] with a mean DSC value of 0.848. This is to mean that our proposed automatic system can better identify true nodule pixels of an image. The higher the value of DSC, the lower the loss and it shows us the robustness of our model.

In our work we used the advanced and state of the art concept of deep learning algorithms called U-Net. Our average ensemble learning technique combines the predicted outputs from three different network architectures with various hyper parameter values of a U-Net model to automatically and effectively extract both local and global features for lung nodules using convolutional layers of contracting and expanding paths and this makes it to have a better performance in the pulmonary nodules segmentation problem. The features generated by our U-Net deep learning model are relatively good enough to be used in clinical usage. From Tables 5.4 and 5.5, our experimental results indicate us that combined and post processed U-Net models using average ensemble learning method have better performance in terms of accuracy, sensitivity, specificity, and DSC.

5.3.2. Research Questions Evaluation

At the beginning of this work, we raised two questions that need to be responded to assess the achievements and the overall performances of our proposed approach.

Research question 1: Can we improve the accuracy of pulmonary nodules segmentation in CT images using post processed U-Nets? **Yes**

U-Net is a class of convolutional neural network that have the ability to extract both local and global features to get the full context of an image. From the evaluation results obtained, the values of accuracy, sensitivity, specificity and DSC are better than those obtained from previous systems. Sensitivity is the measure of accuracy in segmenting true nodule pixels. Similarly specificity is the measure of accuracy in segmenting true non-nodule pixels. And accuracy measures the percentage of accurately segmenting true nodule and true non-nodule pixels while DSC is the measure of segmentation overlap accuracy between predicted segmentation mask and ground truth mask. Our proposed approach outperforms previously implemented systems with

accuracy of 0.965, sensitivity of 0.826, specificity of 0.983, and a DSC value of 0.848. Therefore, our approach improves the accuracy of segmenting true nodule pixels, true non-nodules pixels and the segmentation overlap accuracy of pulmonary nodules from lung CT images. In general, combining three different network architectures with various hyper parameter values of a U-Net model using average ensemble learning at the post processing stage improved the percentage of accuracy in segmenting true nodule pixels, true non-nodule pixels and the segmentation overlap accuracy.

Research question 2: Can we use the same method to segment different pulmonary nodules that varies in type and size? **Yes**

Our lung segmentation algorithm called thresholding and morphological operators incorporates steps and techniques to keep pulmonary nodules located at different parts of the lung structure. Well-circumscribed pulmonary nodules completely reside on the lungs parenchyma whereas juxta-pleural and juxta-vascular pulmonary nodules found attached with the lung wall and the blood vessels, respectively. The dataset that we used contains nodules having a nodule size of 3-30 mm in diameter and those annotated nodules are enclosed with a complete circle whatever the shape of the nodule is. In the algorithm, an average HU value of -420 [45] used to consider the entire lung tissue and the connected components with vessels and lungs wall. We applied a morphological closing operation to keep juxta_pleural nodules that are attached with the lung wall and a morphological erosion operation is applied to isolate juxta_vascular nodules that are attached with blood vessels of the lungs parenchyma. Based on their internal texture, pulmonary nodules can also be classified as solid part-solid, and non-solid. Our proposed system consists of post processed U-Nets. U-Net is composed of contracting and expanding paths of convolutional layers which has the ability to extract both local and global features of an image. Local features obtained from the contracting path are used to extract the features, textures, and appearances of the nodules and the global features obtained from the expanding path are used to extract the structure and location of the nodules.

To justify the ability of our proposed approach in segmenting different types of pulmonary nodules, we collected sample slices of each category from LUNA16 which is a subset of LIDC/IDRI dataset. In LIDC/IDRI, the pulmonary nodules with texture index 1 are considered

as non-solid, pulmonary nodules with texture index 2 or 3 are considered as part-solid, and pulmonary nodules with texture index 4 or 5 are considered as solid [26]. This data is provided in a separate XML file.

Nodule type	No. of slices & pulmonary nodules	Sizes in diameter	Mean DSC
Solid	60	3-30mm	0.601
Part-Solid	60		0.579
Non-Solid	60		0.523

Table 5-6 Separate segmentation performance for Solid, Part-Solid, and Non-Solid nodules

As we can see on Table 5-6, our proposed system can segment different pulmonary nodules that varies in type even though the segmentation performance can even be further improved by using many slices for each category during training and testing of the proposed model. Totally we have given 1,318 XML files and under each xml file there are multiple numbers of patients' lung CT scan images. Filtering out and categorizing all pulmonary nodules from such huge amount of data takes too much time since the mapping is going to be done manually using the scans texture index. In addition, based on the given XML file data many of the pulmonary nodules are solid and we are unable to extract all lung CT images with their corresponding nodules at a time due to shortage of our disk space. We just did this only to show our system's ability in segmenting different pulmonary nodules that varies in type. Comparing the performances obtained from each type, the DSC value of solid pulmonary nodules is better than the other two. This is because of solid pulmonary nodules mostly residing on the lung parenchyma whereas non-solid pulmonary nodules may incorporate the underlying parenchymal structure including vessels and airways and training on more number of slices is needed to get better performance. Part-solid nodule is the combination of solid and non-solid nodule. In general, we can say that integrating post processed U-Nets with thresholding and morphological lungs region segmentation algorithm enables us to extract the features and locations of different nodules found on different regions of the lung accurately.

Chapter Six

6. Conclusion and Future Works

6.1. Conclusion

The main goal of this research work is to implement a robust pulmonary lung nodules segmentation system using convolutional neural network algorithm. The technique had been applied in lung CT scan images for segmentation of pulmonary lung nodules, and a comparison were made with existing segmentation methods such as 3-dimentional CNN, 3D CNN - Encoding/Decoding, and morphological approaches and convexity models.

In this paper, to implement an automatic CAD system for the segmentation of pulmonary lung nodules from lung CT scan images, the type of CNN architecture called U-Nets network is integrated with thresholding and morphological operators lungs region segmentation algorithm. We further improved the performance of a single U-Net network at the post processing stage. Our post processing stage combined three U-Nets having different network architectures using average ensemble learning method (MBM). MBM made the results improved and more stable by producing a better local optimum value with decreased error rate. MBM enable us to take advantage of the different network architectures and hyper parameter values from their variation. MBM combines the predicted values from three different network architectures with various hyper parameter values of a U-Net model to produce one best segmentation map. MBM can weaken the influence of the lower-valued U-Net network architecture, pixel wisely and deliver a segmentation result that is superior to the best individual U-Net model network architecture. MBM binarizes each pixel values obtained from these three individual U-Nets using a threshold value of 0.5, then take the average of the three pixel values obtained from each U-Net followed by binarization. The proposed approach is trained, validated, and tested on 858 lung CT scan images with their corresponding ground truth nodule masks got from LUNA16 database [12]. Our proposed approach gave us evaluation results of 0.848 mean DSC, 0.965 mean accuracy, 0.826 mean sensitivity, and 0.983 mean specificity. The results clearly showed that our approach is much ahead in terms of accuracy, sensitivity, specificity, and dice similarity coefficient and suited well for the segmentation of pulmonary lung nodules when compared to other previously

implemented algorithms. The automatically segmented pulmonary nodules from our proposed system can be used to diagnosis and treat pulmonary nodules at early stage; since they are the manifestations of lung cancer. The output obtained from our proposed method might be used by health-care professionals as a second opinion and would bring high level of accuracy in diagnosis.

Generally, we showed that the deep learning algorithm specifically U-Net model together with thresholding and morphological operators lungs region segmentation algorithm can deliver good performance. Further performance improvement can be achieved by using average ensemble learning method as a post processing technique. Our average ensemble learning method (MBM) combines the predicted outputs from three different network architectures with various hyper parameter values of a U-Net model to provide a segmentation result which is better than the best individual U-Net model network architectures. U-Nets with MBM can learn deep features and able to capture the differentiation between the nodule and background pixels with highest accuracy.

6.2. Future Works

Some of the proposals that can be the extension of this work include the following:

1. We used 2-dimensional images as training and testing set for our model. In the future we will try to use 3-dimensional lung CT scan images with the proposed approach for the segmentation of pulmonary lung nodules.
2. We mainly focused only on the segmentation of pulmonary nodules. But we can further segment and classify pulmonary nodules as malignant or benign also. This could be a better way to diagnosis and treat lung cancer patients as early as possible.
3. The post processing method of this work was not tested in other deep learning models and applications. Implementing average ensemble learning method (MBM) together with other deep learning algorithms can be one of our future activities.
4. Thresholding and morphological operators lungs region segmentation algorithm is used together with three different network architectures with various hyper parameter values of a U-Net model which are post processed by MBM. Using our proposed approach for other medical image segmentation problems can be a future task.

References

[1] S. Krishnamurthy, G. Narasimhan, and U. Rengasamy," Three-dimensional lung nodule segmentation and shape variance analysis to detect lung cancer with reduced false positives," *J Engineering in Medicine*, vol.230, 1, pp.58-70, Jan 2016.

[2] Pataer A., Kalhor N., Correa A. M., Raso M. G. Erasmus, J. J. Kim, E. S. & Vaporciyan, A., 2012, "A Histopathologic response criteria predict survival of patients with resected lung cancer after neoadjuvant chemotherapy," *J Thorac Oncology*, vol.7, pp.825-832, 2012.

[3] Gould MK, Donington J, Lynch WR, et al., "Evaluation of individuals with pulmonary nodules: when is it lung cancer? Diagnosis and management of lung cancer: American college of chest physicians evidence-based clinical practice guidelines," *Chest*, vol.143, 5, pp.e93S–e120S, May 2013.

[4] Sluimer I, Schilham A, Prokop M, van Ginneken B, " Computer analysis of computed tomography scans of the lung: a survey," *IEEE Trans Med Imaging*, vol.25, 4, pp.385-405, May 2006.

[5] Ko JP, Naidich DP, "Computer-aided diagnosis and the evaluation of lung disease," *J Thorac Imaging*, vol.19, 3, pp.136–155, Jul 19 2004.

[6] Qing Li, Weidong Cai, Xiaogang Wang, Yun Zhou, David Dagan Feng and Mei Chen. "Medical Image Classification with Convolutional Neural Network," *13th International Conference on Control, Automation, Robotics & Vision*, Marina Bay Sands, Singapore, 10-12th Dec 2014, pp.844-848.

[7] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *arXiv:1506.01497v3 [cs.CV]*, 6 Jan 2016.

[8] E. Shelhamer , J. Long , T. Darrell, "Fully convolutional networks for semantic segmentation," *arXiv: 1605.06211v1 [cs.CV]*, 20 May 2016.

[9] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” *MICCAI 2015*, pp.234-241, 2015.

[10] W. J. Tuddenham, “Visual search, image organization, and reader error in roentgen diagnosis,” *Studies of the psycho-physiology of roentgen image perception, Radiology 78*, pp.694-704, 1962.

[11] M.V. Sprindzuk et al, “Lung cancer differential diagnosis based on the computer assisted radiology: The state of the art,” *Polish journal of Radiology*, vol.75, 1: pp.67-80, Mar 2010.

[12] Grand Challenge (2016). Lung Nodule Analysis 2016 [online]. Available: https://luna16.grand-challenge.org/Download/#.

[13]. Samuel G Armato, Geo.rey McLennan, Luc Bidaut, Michael F McNi.-Gray, Charles R Meyer, Anthony P Reeves, Binsheng Zhao, Denise R Aberle, Claudia I Henschke, Eric A Ho.man, and others, “The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A Completed Reference Database of Lung Nodules on CT Scans,” *Medical physics*, Vol.38, pp. 915–931, , Feb 2011.

[14] National Cancer Institute. Non-small Cell Lung Cancer Treatment - Patient Version [Online]. Available: https://www.cancer.gov/types/lung/patient/non-small-cell-lung-treatment-pdq.

[15] World Health Organization. Cancer fact sheet No.297 [online]. Available: https://www.who.int/en/news-room/fact-sheets/detail/cancer.

[16] Mehmet Sezgin and Bulent Sankur, “Survey over image thresholding techniques and quantitative performance evaluation,” *Journal of Electronic Imaging*, vol.13, 1, pp.146–165, Jan 2004.

[17] Rafael C. Gonzalez and Richard E. Woods, *Digital Image Processing*, Third Edition. Asia: Pearson Education, 2008.

[18] Twair AA, Torreggiani WC, Mahmud SM, Ramesh N, Hogan B, “Significant savings in radiologic report turnaround time after implementation of a complete picture archiving and communication system (PACS),” *Journal of Digit Imaging*, vol.13, 4, pp.175–177, Nov 2000.

[19] Anitha. S and Sridhar.S, “Segmentation Of Lung Lobes And Nodules In CT Images,” *An international Journal of Signal and Image Processing*, Vol.1, 1, Jul 2010.

[20] G.Ramya and A.S Shanthi, “Segmentation Techniques for Image Analysis: A Survey,” *International Journal of Trend in Research and Development*, Vol.2, 6, pp.32-38, 2015.

[21] S. G. Armato and W. F. Sensakovic, “Automated lung segmentation for thoracic CT: Impact on computer-aided diagnosis,” *Academic Radiology*, vol.11, 9, pp.1011–1021, Sep 2004.

[22] S. Uzelaltinbulat, B. Ugur. “Lung tumor segmentation algorithm,” *9th International Conference on Theory and Application of Soft Computing, Computing with Words and Perception, ICSCCW 2017*, Budapest, Hungary, 24-25 Aug 2017, pp.140–147.

[23] Assefa, M., Faye, I., Malik, A. S., & Shoaib. “Lung nodule detection using multi-resolution analysis,” *Proceedings of 2013 ICME International Conference on Complex Medical Engineering*, Beijing, Chaina, 25-28 May 2013, pp.457-461.

[24] JH Austin et al., “Glossary of terms for CT of the lungs: recommendations of the Nomenclature Committee of the Fleischner Society,” *Radiology*, vol.200, pp.327–331, 1996.

[25] Kostis, W., Reeves, A., Yankelevitz, D., Henschke, C., 2003, “Three-dimensional segmentation and growth-rate estimation of small pulmonary nodules in helical CT images,” *IEEE Transactions on Medical Imaging*, vol. 22, 10, pp.1259–1274, Oct 2003.

[26] Sudipta Mukhopadhyay, “A Segmentation Framework of Pulmonary Nodules in Lung CT Images,” *J Digit Imaging*, 2015.

[27] Coursera. Why is Deep Learning taking off? deeplearning.ai [online]. Available: https://www.coursera.org/learn/neural-networks-deep-learning/lecture/praGm/why-is-deep-learning-taking-off.

[28] Lee Jacobson. Introduction to Artificial Neural Networks – Part 1 [online]. Available: http: // www.theprojectspot.com/ tutorial-post/ introduction-to-artificial-neural-networks-part- 1/7.

[29] Matthew Mayo, KDnuggets. Neural Network Foundations, Explained: Activation Function [online]. Available: https://www.kdnuggets.com/2017/09/neural-network-foundations-explained-activation-function.html.

[30] Coursera. Computer Vision-deeplearning.ai [online]. Available: https://www.coursera.org/learn/ convolutional-neural-networks/lecture/Ob1nR/computer-vision.

[31] Jose Bernal, Kaisar Kushibar, Daniel S. Asfaw, Sergi Valverde, Arnau Oliver, Robert Marti, Xavier Llado, "Deep convolutional neural networks for brain image analysis on magnetic resonance imaging: a review," 11 Jun 2018.

[32] vdumoulin/conv_arithmetic. A technical report on convolution arithmetic in the context of deep learning [online]. Available: https://github.com/vdumoulin/conv_arithmetic.

[33] S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," *arXiv:1502.03167v3 [cs.LG]*, 2 Mar 2015.

[34] N. Srivastava, G.Hinton, A.Krizhevsky, I.Sutskever, and R.Salakhutdinov, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," *Journal of Machine Learning Research*, vol.15, pp.1929-1958, 2014.

[35] K. He, X.Zhang, S.Ren, and J.Sun, "Delving Deep into Rectifiers: Surpassing Human-Level Performance on Image Net Classification," *arXiv:1502.01852v1 [cs.CV]*, 6 Feb 2015.

[36] Sebastian Ruder, "An overview of gradient descent optimization algorithms," *arXiv:1609.04747v2 [cs.LG]*, 15 June 2017.

[37] Diederik P. Kingma, Jimmy Lei Ba, "Adam: A Method for Stochastic Optimization," *arXiv:1412.6980v9 [cs.LG]*, 30 Jan 2017.

[38]. Toshiro Kubota, Anna K. Jerebko, Maneesh Dewan, Marcos Salganicoff, Arun Krishnan, “Segmentation of pulmonary nodules of various densities with morphological approaches and convexity models,” *Medical Image Analysis,* vol.15, pp.133–154, 2011.

[39] Jan Hendrik Moltz, Jan-Martin Kuhnigk, Lars Bornemann, Heiz-Otto Peitgen,. “Segmentation of juxtapleural lung nodules in CT scans based on ellipsoid approximation,” *MeVis Research GmbH - Center for Medical Image Computing*, Bremen, Germany. Pp.25-32.

[40] Zohreh Azimifar, Mohsen Keshani, Farshad Tajeripour, Reza 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.

[41] D. C. Ciresan, A. Giusti, L. M. Gambardella, and J. Schmidhuber, “Deep neural networks segment neuronal membranes in electron microscopy images,” 2012.

[42] Bernard Bronmans, Evert Haasdijk, “Lung Nodule Segmentation Using 3D Convolutional Neural Networks,” *Research paper Business Analytics*, 2 Feb 2018.

[43] V. Badrinarayanan, A. Kendall, and R. Cipolla. “Segnet: A deep convolutional encoder-decoder architecture for image segmentation,” *arXiv:1511.00561v3 [cs.CV]*, 10 Oct 2016.

[44] Botong Wu, Zhen Zhou, Jianwei Wang, Yizhou Wang, “JOINT LEARNING FOR PULMONARY NODULE SEGMENTATION, ATTRIBUTES AND MALIGNANCY PREDICTION,” *IEEE International Symposium on Biomedical Imaging (ISBI),* 10 Feb 2018.

[45] Eng. Michael Samir Labib Habib,”A computer aided diagnosis system (CAD) for the detection of pulmonary nodules on CT scans,” Msc thesis, Dept.Systems and Biomedical Eng., Cairo univ.,Giza, Egypt,2009.

[46] Aamir Saeed Malik and Tae-Sun Choi, ”A novel algorithm for segmentation of lung images,” *ISBMDA 2006, LNBI 4345,* pp.346-357, 2006.

[47] Coursera. Normalizing inputs-deeplearning.ai [online]. Available: https://www.coursera.org/lecture/deep-neural-network/normalizinginputs.-lXv6U.

[48] Nikhil B. Image Data Pre-Processing for Neural Networks [online]. Available: https://becominghuman.ai/image-data-pre-processing-for-neural-networks-498289068258.

[49] S.Gopal Krishna Patro, Kishore Kumar sahu, “Normalization: A Preprocessing Stage”.

[50] Aleksandra Petrakova, Michael Affenzeller, Galina Merkurjeva, “Heterogeneous versus Homogeneous Machine Learning Ensembles,” *Information Technology and Management Science*, pp. 135-140, Dec 2015.

[51] Yllias Chali, Hassan Sadid, Mustapha Mojahid. “Complex Question Answering: Homogeneous or Heterogeneous, Which Ensemble Is Better?,” *19th International Conference on Application of Natural Language to Information Systems (NLDB 2014)*, Montpellier, France, 18-20 Jun 2014, pp. 160-163.

[52] Kamnitsas, Bai, Ferrante, McDonagh, Sinclair, et al., “Ensembles of Multiple Models and Architectures for Robust Brain Tumour Segmentation,” *arXiv:1711.01468v1 [cs.CV]*, 4 Nov 2017.

[53] Milletari, F., Navab, N., Ahmadi, S.A., “V-Net: Fully convolutional neural networks for volumetric medical image segmentation,” *arXiv:1606.04797v1 [cs.CV]*, 15 Jun 2016.

[54] J. Patravali, S. Jain, and S. Chilamkurthy, “2D-3D Fully Convolutional Neural Networks for Cardiac MR Segmentation,” *arXiv:1707.09813v1 [cs.CV]*, 31 Jul 2017.

[55] V. Yeghiazaryan and I. Voiculescu, “An Overview of Current Evaluation Methods Used in Medical Image Segmentation,” 3 Nov 2015.

[56] StackExchange [2017]. Cross Validated – Is the Dice coefficient the same as accuracy [online]. Available: https://stats.stackexchange.com/questions/195006/is-the-dice-coefficient-the-same-as-accuracy.

[57] M. Havaeia, A. Davyb, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, C. Pal, P. Jodoin, H. Larochelle, “Brain Tumor Segmentation with Deep Neural Networks,” *arXiv:1505.03540v3 [cs.CV]*, 20 May 2016.

[58] Mishkin, Dmytro and Matas, Jiri, “All you need is a good init,” *CoRR*, 2015.

[59] A. Krizhevsky, I. Sutskever, and G. Hinton, “Image net classification with deep convolutional neural networks,“ *In NIPS*, 2012.

[60] Stack overflow. Homogeneous vs heterogeneous ensembles [online]. Available: https://stackoverflow.com/questions/49445446/homogeneous-vs-heterogeneous-ensembles.

[61] Maryam Sabzevari, Gonzalo Mart´ınez-Mu˜noz, Alberto Su´arez, “Building heterogeneous ensembles by pooling homogeneous ones,” *arXiv:1802.07877v2 [cs.LG]*, 9 Jun 2019.

[62] SaS BLOGS. Why do stacked ensemble models win data science competitions? [online]. Available: https://blogs.sas.com/content/subconsciousmusings/2017/05/18/stacked-ensemble-models-win-data-science-competitions/.

Appendix

Sample Pulmonary Nodules Mask predicted by our system versus Ground truth

Raw Lung CT image	Processed Lung CT Image	Ground Truth	Predicted Mask