

LUNG NODULE DETECTION AND SEGMENTATION USING A PATCH-BASED MULTI-
ATLAS METHOD

by

MUSTAFA NOOR-E-ALAM

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy at
The University of Texas at Arlington
May 2017

Arlington, Texas

Supervising Committee:

Venkat Devarajan, Supervising Professor
Jonathan W. Bredow
Michael T. Manry
Ioannis D. Schizas
Ganesh Sankaranarayanan

Copyright © by Mustafa Noor E Alam 2017

All Rights Reserved



*I dedicate this thesis to
my father MD Khorshed Alam, and my mother Dilruba Yeasmin*

ACKNOWLEDGEMENTS

First of all, I would like to express my gratitude to Professor Venkat Devarajan for his overall support throughout my doctoral studies and in particular during my research.

I would also like to thank Dr. Michael Manry, Dr. Jonathan Bredow, and Dr. Ioannis D. Schizas and Dr. Ganesh Sankaranarayanan for their support and for agreeing to become my doctoral committee members.

I also wanted to thank Dr. Ganesh Sankaranaryana of Baylor Health Care System for his continuous support. Technical discussions with Dr. Sankaranaryana were helpful in providing insight into the practical importance of my research problem.

I would like to acknowledge all the faculty members and staff of the Electrical Engineering Department of the University of Texas at Arlington for their friendly and helpful attitude, which gave me a very pleasant and wonderful experience during my doctoral studies. I would like to thank all my lab mates especially B. Namazi, P. Acharjee, Dr. M. Lakshman, A. Habib, and Dr. G. Toscano for all their thoughtful discussions and ideas.

I would like to take the opportunity to recognize the sacrifices my parents have made for my sister and me. No words can express my gratitude for my father MD. Khorshed Alam and my mother Dilruba Yeasmin, for their endless and unconditional support. I would also like to express my gratitude for my sister Tahsina Alam and her husband, brother Mahbub Zaman for all their advice and support.

Finally, I would like to conclude by appreciating the encouragement and a great deal of patience shown by my wife Nazifah Islam during the entire period of my studies.

Without all my family members support, I wouldn't have been in a position to write a single word of my dissertation.

April 24th, 2017

ABSTRACT

LUNG NODULE DETECTION AND SEGMENTATION USING A PATCH-BASED MULTI-ATLAS METHOD

Mustafa Noor E Alam

The University of Texas at Arlington, 2017

Supervising Professor: Venkat Devarajan

CT image based lung nodule detection is the most widely used and accepted method for detecting lung cancer. Most CT image based methods rely on supervised learning, which has a high number of false positives and need a large amount of pre-segmented training samples. These problems can be solved if an optimally small number of training sample images can be created, where each sample has lung nodules of similar size and shape as the target image of the actual patient. Based on this hypothesis, two algorithms are proposed for 2D CT images and 3D CT images respectively. Both algorithms use size and shape characteristics of nodule candidates and, patch-based image segmentation.

The proposed algorithm for 2D images has three steps: a) a small set of atlases is selected by comparing the target image with a larger set of atlas images using a size-shape based feature vector, b) lung nodules are selected using a patch-based method, where each pixel of a target image is labeled by comparing the image patch, centered by the pixel, with patches from an atlas library. The most probable labels are then chosen according to a defined closest match criterion. c) a Laplacian of Gaussian blob detection method is then developed to find the segmented area of

the lung nodule. The method is tested for more than 25 test slices, where each test image is applied to more than 200 atlas images. For non-attached nodules in the size between 3 mm to 30 mm, the sensitivity of the proposed algorithm is 100% and no false positive was found.

For 3D images, the algorithm is significantly changed. This algorithm has three steps: a) In the first step, nodule candidates of the current patient are detected and different features are extracted from each nodule b) The second step is the ‘atlas selection step’, in which two or three very similar lung images (atlas image) are selected from a group of atlases by a nodule-based atlas search process c) In the third and final step, the *nodule based patch comparison* process is developed to determine the accurate size and shape of the lung nodules. The proposed method has been proven accurate in recognizing all the non-attached nodules, which are bigger than 3 mm of radius, when applied on a population of twelve patient’s image datasets.

CONTENTS

ACKNOWLEDGEMENTS	iv
ABSTRACT	vi
LIST OF ILLUSTRATIONS	xi
LIST OF TABLES	xv
CHAPTER 1 Introduction	1
1.1 Computed Tomography.....	2
1.2 Dissertation Problem Statement and Goal	4
CHAPTER 2 Literature Review	5
2.1 Cancer detection:.....	5
2.2 Multi-atlas based segmentation method:.....	9
2.2.1 Generation of Atlases:.....	10
2.2.2 Offline learning:	10
2.2.3 Registration:	10
2.2.4 Atlas selection:.....	11
2.2.5 Label propagation:	11
2.2.6 Online learning:	11
2.2.7 Label fusion:	11
2.2.8 Post-processing:	11
CHAPTER 3 Proposed algorithm for 2D implementation.....	12
3.1 Step1: Atlas selection.....	15
3.1.1 Morphological Operation:.....	16
3.1.2 Thresholding:	16

3.1.3	Size and shape based Feature vector:.....	17
3.1.4	Removing big connected components:	19
3.1.5	K-NN algorithm:.....	20
3.2	Step2: Image segmentation using image patch comparison.....	20
3.2.1	Feature vector develop:.....	21
3.2.2	Label Selection:	22
3.3	Post Processing.....	23
CHAPTER 4	Experimental Results for 2D implementation.....	25
4.1.1	Final result	37
CHAPTER 5	Proposed algorithm for 3D implementation.....	38
5.1	Step 1: Nodule candidate detection in both atlas and target images	38
5.1.1	Thresholding:	40
5.1.2	Lung region extraction:.....	40
5.1.3	Morphological operation:.....	40
5.1.4	Contour correction:	41
5.1.5	Nodule candidate detection using connected component analysis	45
5.2	Step 2: Atlas selection by nodule-based atlas search	47
5.2.1	Euclidian Distance-Weight table:	48
5.2.2	Nodule similarity table:.....	49
5.3	Step 3: Image segmentation	50
CHAPTER 6	Experimental results for 3D implementation	53
CHAPTER 7	Conclusions and Future work.....	70
7.1	Comparisons of the 2D and the 3D method:.....	70
7.2	Contribution:	72
7.3	Future work:.....	73

Appendix A.....	74
Appendix B.....	77
REFERENCES	83

LIST OF ILLUSTRATIONS

Figure 1-1 A typical CT scan setup	3
Figure 1-2 Example slices of chest scan by MDCT (a) Axial slice (b) Coronal slice (c) Sagittal slice (d) Body planes.....	3
Figure 2-1 Typical algorithm for generic lung nodule detection of lesions in medical images	6
Figure 2-2 A sample lung segmentation step	7
Figure 2-3 Basic Building blocks for a typical MAS algorithm	9
Figure 3-1 Relationship between TP (True positives) and FP (False positives)	12
Figure 3-2 Block diagram of the proposed algorithm.....	14
Figure 3-3 Block diagram of the proposed atlas selection step	15
Figure 3-4 Structuring element used for morphological operation.....	16
Figure 3-5 Convex hull of the object A.....	17
Figure 3-6 Block diagram of the proposed image segmentation step using patch comparison....	21
Figure 3-7 Image patch and their comparison	22
Figure 3-8 The post processing step	23
Figure 4-1 Target image from one patient	25
Figure 4-2 Target image after morphological operation.....	25

Figure 4-3: Output of conversion to binary images image	26
Figure 4-4 Likely candidates for nodules based on eliminating areas larger than typical nodules	26
Figure 4-5 Four steps of the atlas image taken very close to the neck for patient 2 (a) Original CT image (b) After morphological operation (c) After converting to a binary image and (d) After removing connected components due to a bigger area. It can be manually verified that this is not a good atlas image.....	29
Figure 4-6 Four steps of another atlas image taken very close to the diaphragm of patient 2 (a) original CT image (b) After morphological operation (c) After converting to binary (d) After removing connected components due to a bigger area. Clearly, this is not a good atlas image either.	30
Figure 4-7 Four steps of the third atlas image taken at mid-level between neck and diaphragm for patient 2 (a) original CT image (b) After morphological operation (c) After converting to binary (d) After removing connected components due to a bigger area. This is the atlas closest to the target image.....	31
Figure 4-8 Normalized Euclidian distances from the target image	32
Figure 4-9 (a) Output of the second stage (b) The final output with the segmented lung nodule	33
Figure 4-10 Two atlases manually selected from two additional patients as input to the step-2 of the program (a) Atlas – 2 lung image (b) Segmented lung nodule from atlas – 2 (c) Atlas – 3 lung image (d) Segmented lung nodule from Atlas - 3.....	35
Figure 4-11 The final output with the segmented lung nodule.....	36
Figure 5-1 Block diagram of the nodule candidate detection step.....	38
Figure 5-2 Block diagram of 3D implementation.....	39

Figure 5-3 8-connected chain coding.....	42
Figure 5-4 Critical points of different angles a) two consecutive points of 135° b) One concave point of 90° c) one point of 90° and one point of 135° d) one point of 135° and one point of 90° e) two points of 90°	42
Figure 5-5 Contour correction process (overall block diagram).....	43
Figure 5-6 Contour correction process for a 2D slice.....	44
Figure 5-7 3D connected component method using 26 connectivity	45
Figure 5-8 A nodule enclosed by a sphere.....	46
Figure 5-9 Block diagram of the atlas selection step.....	47
Figure 5-10 Image segmentation using patch-based segmentation	50
Figure 5-11 Image patches and their comparison.....	51
Figure 6-1 3D view of lungs (rendered by 3Dimviewer)	53
Figure 6-2 2D slice of an image a) before thresholding b) after thresholding.....	54
Figure 6-3 Three different initial thresholds and the final threshold for a 3D image.....	54
Figure 6-4 Histogram of the 3D image consisting of 212 2D slices.....	55
Figure 6-5 Seven different iterations of the thresholding process with the initial threshold of 1800	56
Figure 6-6 (a) before lung region extraction (b) after lung region extraction	57

Figure 6-7 A CT image where the critical points are circled and adjacent points are joined.....	58
Figure 6-8 After filling out the polygons.....	59
Figure 6-9: CT image slice after final iteration of the contour correction.....	60
Figure 6-10 Contour correction in coronal direction a) before contour correction b) after contour correction	61
Figure 6-11 Two lung nodule candidates and their sphericity.....	62
Figure 6-12 Different lung nodule candidates with their sphericity	63
Figure 6-13 Two lung nodule candidates with sphericity less than 0.2.....	64
Figure 6-14 3D image of the segmented nodule by patch comparison.....	67
Figure 6-15 (a) Original image b) Final segmented nodule by the proposed algorithm (c) Manually segmented nodule from the same image.....	68
Figure 6-16 (a) Original image b) Final segmentation by the proposed algorithm	69
Figure 6-17 (a) Original image b) Final segmentation by the proposed algorithm	69

LIST OF TABLES

Table 4-1 Size and shape based feature vector values of the target image.....	27
Table 4-2 Figure number and Euclidian distance from the target image.....	33
Table 6-1 Euclidian Distance-Weight table of target nodule 1.....	65
Table 6-2 Nodule similarity Table of a sample target image.....	65
Table 6-3 Nodule similarity index for the all the atlases from a Nodule Similarity table	67

CHAPTER 1

INTRODUCTION

Lung cancer is one of the leading causes of deaths in the world [1] and the second most common cancer in the United States. Approximately 222,500 new patients will be diagnosed with cancer in the USA in the year 2017 [2]. Lung cancer is caused by uncontrolled growth of cells in the lung and early detection of uncontrolled cell growth can increase the survival rate of the cancer patients to a great extent. Advancement in computed tomography (CT) technology has offered a great opportunity in such early detection and therefore helping lung cancer patients to receive early treatment.

Focal opacity of size range between 3 mm and 30 mm is defined as a nodule. There are different sized lung nodules. Anything less than 3 mm is called a micro nodule, and anything greater than 30 mm is called a mass [3]. CT technology has enabled visualization of very small or low-contrast nodules that can be barely seen in conventional radiograms.

Careful observation of the CT images by an expert radiologist is the common practice for detecting lung nodules. Since 1980, multiple attempts have been made to create an automated system to detect lung nodules in order save time and cost. None of these systems is completely successful in finding lung nodules automatically, but some of them have gone a long way to help the radiologists to reduce their workload. These systems are typically grouped under Computer Aided Diagnostic (CAD) systems. Some of the parameters used to define the efficiency of a CAD system are [4]

1. Sensitivity: Sensitivity is defined as

$$Sensitivity = \frac{TP}{TP + FN} \quad (1.1)$$

Where, TP = True positive, i.e., a positive output for a sample that has the disease and

FN= False negative, i.e., a negative output for a sample that has the disease.

2. False positives (FP), a positive output for a sample that has no disease.
3. Processing time.
4. Level of automation.
5. Implementation costs.
6. Detection capability of different types, sizes and shapes of nodules, such as isolated nodules, micronodules, partially solid, nodules attached to lung edges or lung cavities.

1.1 Computed Tomography

Computed Tomography (CT) is a widely-used method to create an image of the interior of a living human body. Figure 1-1 shows a typical setup for CT scan. In the CT technique, X-ray beams are passed through a section of a human body from different directions. X-ray sensors are positioned on the opposite side of the body. Contemporary CT scanners can obtain up to 320 CT slices at a time. A 2.5 mm slice (2D) is the standard for a routine chest protocol. A typical chest CT scan may contain 400 slices, where each slice is a 512×512 pixels frame. In effect, MDCT produces high-quality volumetric images (3D). Figure 1-2 shows three slices of MDCT chest scan from three different planes.

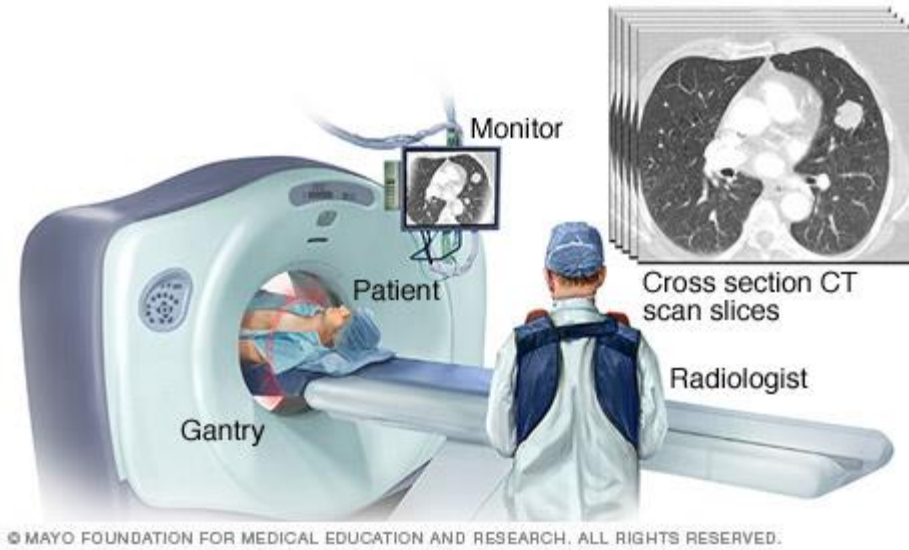
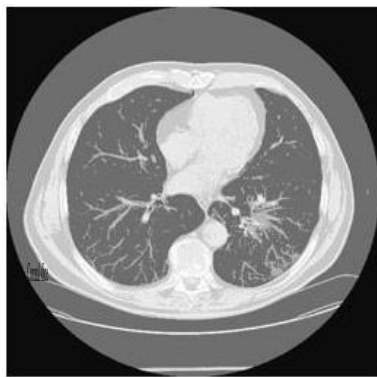


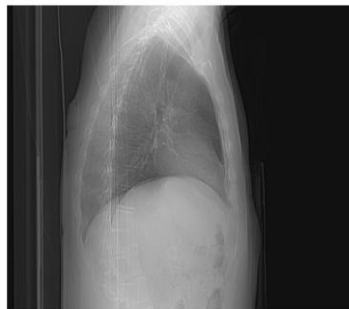
Figure 1-1 A typical CT scan setup



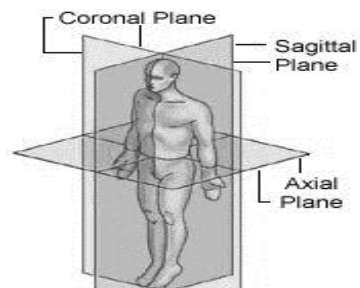
(a) Axial slice



(b) Coronal slice



(c) Sagittal slice



(d) Body planes

Figure 1-2 Example slices of chest scan by MDCT (a) Axial slice (b) Coronal slice (c) Sagittal slice (d) Body planes

1.2 Dissertation Problem Statement and Goal

Automated lung nodule detection and segmentation is a challenging problem and much work has been accomplished in this field. And yet, in practice, lung nodule detection is performed manually by expert radiologists, which is expensive and time-consuming. This is because, most of the previous work related to ‘automation’ were in fact implemented using supervised/unsupervised learning methods, which have a high number of false positives and need a large amount of already segmented training samples. Also, most of the methods were applied only to a limited number of data sets.

Thus, in this research, a novel approach has been taken to solve those problems. Two algorithms have been developed for 2D and 3D images respectively. In both algorithms, size and shape characteristics of lung nodules and patch based image segmentation are used. In the 2D algorithm a new size and shape based feature vector is developed and, for 3D algorithm a nodule based atlas search method is developed.

Thus, these two algorithms should have high nodule detection sensitivity so that they can detect different types of nodules.

The algorithms should have a small number of false positives and should be computationally efficient.

CHAPTER 2

LITERATURE REVIEW

2.1 Cancer detection:

A good amount of literature exists for lung nodule detection and segmentation from 2D and 3D CT images. Each paper follows an algorithmic structure with different steps. Some of the steps are common in most of the papers. These common steps are: Image acquisition, preprocessing, lung segmentation, nodule detection and false positive reduction [5].

The first step of the process is the lung image acquisition. There are different public databases available for research purposes. Examples are The Reference Image Database to Evaluate Therapy Response(RIDER) [6], Lung Image Database Consortium (LIDC) [7] , Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI) [6], Early Lung Cancer Action Program (ELCAP) [8] etc. Many researchers have also used different private databases obtained from their partner hospitals.

The second step of the lung segmentation is the preprocessing. Preprocessing helps to reduce noise and artifacts from CT images. In addition, preprocessing steps may be needed because different CT scanners produce different images of different sizes and intensity. These are not essential steps and therefore, not all papers report preprocessing. Different methods are used for preprocessing. Cascio, Donato, et al. [9] employed linear isotropic interpolation for data reconstruction to make uniform 3D spatial reconstruction. Soltanized et al. [10] and Kim et al. [11] performed median filtering for smoothing and then morphological operation for reducing noise. Pu et al.[12],Gori et al. [13], Wei et al. [14] and Retico et al. [15] selected Gaussian filtering to eliminate artifacts.

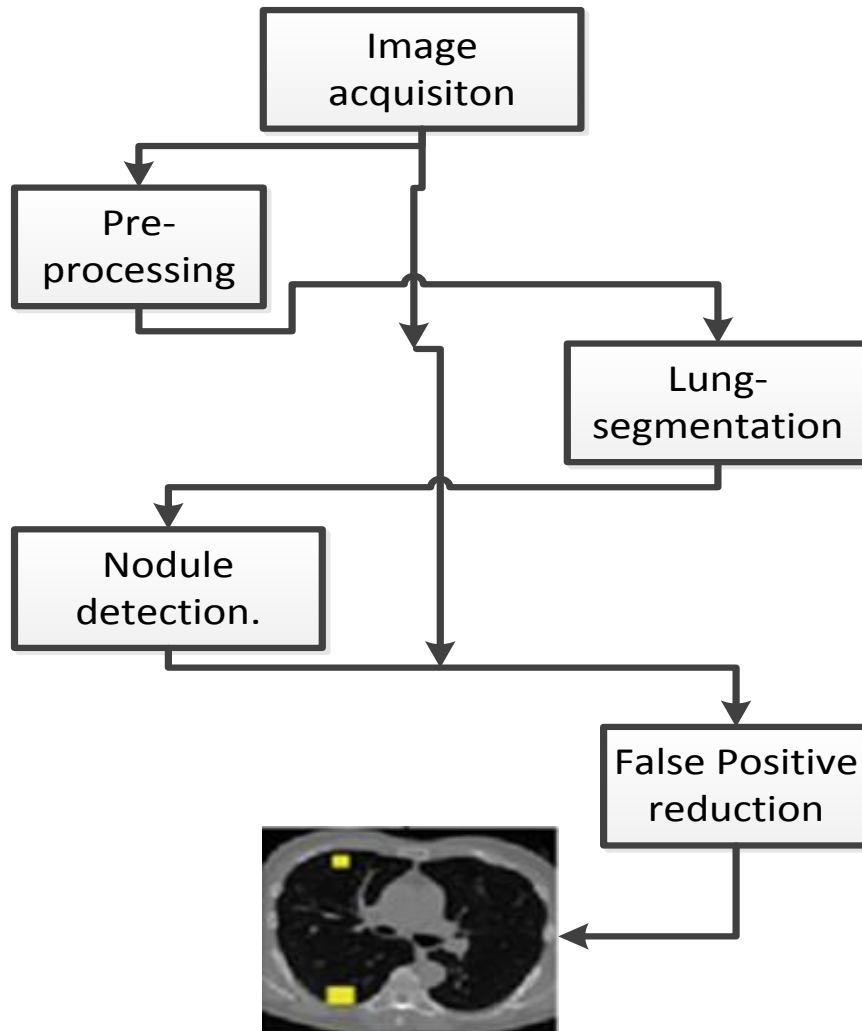


Figure 2-1 Typical algorithm for generic lung nodule detection of lesions in medical images [5]

The algorithm by Namin et al. [16] had several steps for preprocessing: first was voxel size unification and the second was normalization of intensity range. They also implemented a Gaussian filter for image de-noising. Liu et al. [17] used a 2D dot enhancement filter for obtaining a nodule with high sensitivity. Garnavi et al. [18] utilized a low pass filtering with disk

and Gaussian parameters. Diciotti et al. [19], Sluimer et al. [20] employed Laplacian of Gaussian filter for image enhancement, Bae et al. [21] a morphological filter and Ochs et al. [22] a sphericity structure enhancement filter.

The third step is the lung segmentation. Lung segmentation refers to segmentation of the lung region from muscle and fat tissue. In general, lung segmentation methods can be divided into two groups: methods based on thresholding and shape or edge based methods [4]. Figure 2-2 shows a sample lung segmentation step. Choi and Choi [23] implemented a three-step method for lung volume segmentation: first step is thresholding, second step is to find the lung volume using 3D connected component labeling and the third step is segmentation refinement.

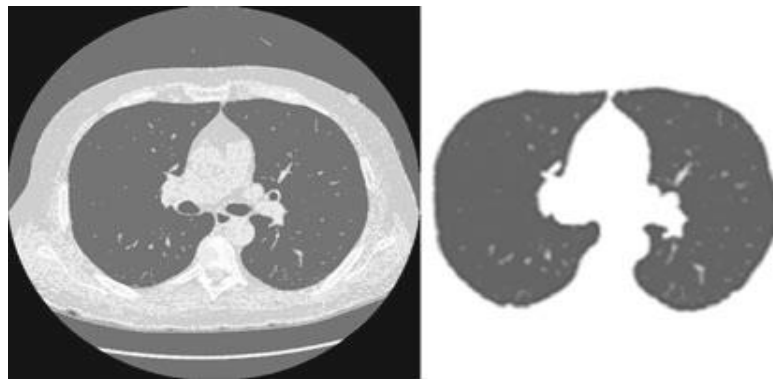


Figure 2-2 A sample lung segmentation step [5]

Keshani et al. [24] employed several steps for lung segmentation. In the first step, binary images were collected by an adaptive, fuzzy thresholding process. Then, two windows were applied to get a hole-free mask. Thereafter, active contour modeling was applied to find the lung area. Kim et al. [25] implemented a deformable model to segment the lung boundaries. Belloti et al. [26] proposed a region growing and active contour modeling to segment the lung region.

The fourth step is the nodule candidate detection. Choi and Choi [23] developed Hessian based matrix for detection of nodule candidates. El-baz et al. [27] implemented a genetic algorithm based template matching for nodule candidate detection. In their method, in the first step, a genetic algorithm was developed to determine the target position and a proper radius in an observed image. In the second step, they used template matching. Cascio, Donato, et al. [9] implemented a 3D mass spring model for modeling the nodule. Suiyuan and Junfeng [28] used thresholding to find the region of interest and then they applied connected component analysis for finding nodule candidates. Xu et al. [29], Aoyama [30] and Wang et al. [31] applied dynamic programming to isolate nodule contour. Fan et al. [32] developed a 3D template matching algorithm to segment the nodule. Kostic et al. [33] implemented a 3D morphological processing to segment the lung nodules. Enquobahrie et al. [34] applied surface morphology analysis and volume occupancy analysis to detect the nodule candidate. Li et al. [35] developed a rule based classifier to detect nodules. Kawata et al. [36] implemented k-means clustering procedure and the linear discriminate (LD) classifier for nodule detection. Matsumoto et al. [37] developed new filter named quantized convergence index filter for nodule detection. Jia et al. [38] used hessian based detection method and Fukano et al. [39] developed a morphology filter to detect nodule candidates. Zhao et al. [40], Dehmeshki et al. [41] applied support vector machine to classify nodules.

The fifth step is the false positive reduction using feature classifiers. Choi and Choi [23] and Santos et al. [42] developed support vector machine classifiers for nodule classification. El-baz et al. [27] developed a Bayesian supervised classifier for false positive reduction. Matsumoto et al. [43] implemented rule based classifiers. In their method, seven features were calculated for

each nodule candidate. For each feature, a cutoff value was chosen above which the nodule candidate was labeled as a true nodule.

The implementation of our proposed 2D and 3D methods has the conventional image acquisition and preprocessing steps. CT Images of RIDER database [6], which is acquired from *The Cancer Imaging Archive* (TCIA) [44] are also our source images for implementation of both the algorithms. Two preprocessing steps are developed in both implementations, which are: *thresholding, morphological operation.*

2.2 Multi-atlas based segmentation method:

In multi-atlas segmentation (MAS), a target image is segmented by manipulating a set of atlases (already pre-segmented images). The idea was first introduced and popularized by Rohlfing, et al. [45], Klein, et al. [46], and Heckemann, et al. [47]. In their method, atlases are used and manipulated for segmentation of a new image instead of using a model based segmentation. A typical block diagram for an MAS algorithm is shown in Figure 2-3.

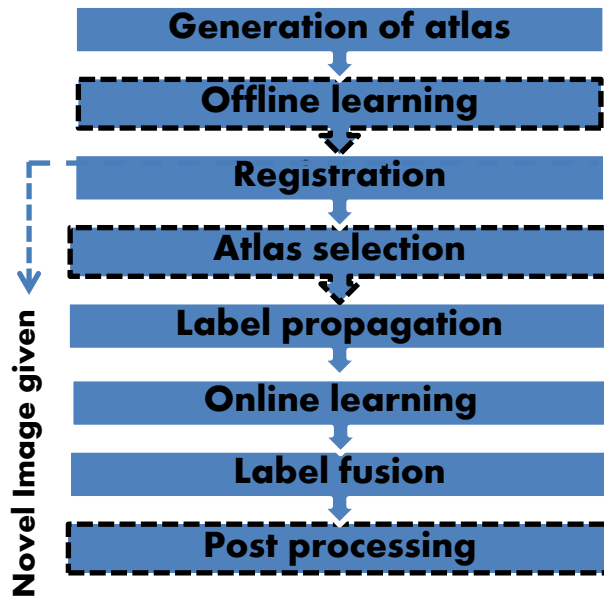


Figure 2-3 Basic Building blocks for a typical MAS algorithm [48]

The basic steps are

2.2.1 Generation of Atlases: Atlases are already segmented images. Typically, domain specific experts, who rely on their experience and textbook anatomical knowledge, create atlases. They also use commercial software and multiple image modalities as needed [49].

2.2.2 Offline learning: After the generation of atlases, often they are reviewed offline and additional information may be gathered for later processing in the segmentation step. For example, Ramus et al. [50] evaluated the atlases independently and ranked them based on intensity, deformation, and reference. Thereafter, they used the Spearman rank correction method to cluster the atlases.

2.2.3 Registration: Registration is the process of establishing spatial correspondence between images. In the image registration process, a similarity parameter is first defined. One image (atlas) is deformed by employing a deformation model to maximize this similarity parameter. Different types of registration methods are available. Rohlfing et al. [45] invoked a spline based deformation model and normalized mutual information as the similarity measure. Candemir et al. [51] implemented a sift-flow algorithm for deformation of the atlas and a SIFT descriptor based similarity parameter. Lee et al. [52] developed a particle guided group wise image registration method. In their method, the same number of particles/pixels are chosen from different subjects and a correspondence was sought by minimum description length (MDL). Some recent papers described non-local means methods and patch-based search strategies [53] [54] [55] [56].

2.2.4 Atlas selection: After registration, not all atlases are selected for the next step of label propagation. By excluding irrelevant atlases, the segmentation can be improved because irrelevant atlases misguide the segmentation process and increase computation time. Aljabar et al. [57] showed that the atlas selection process, using a majority voting approach improves the accuracy of label propagation process.

2.2.5 Label propagation: After the selection of atlases, the labels of the target image are decided. Early MAS (Multi-Atlas segmentation) algorithms [47] used the nearest neighborhood algorithm for choosing the target image pixel label.

2.2.6 Online learning: Some MAS algorithms have an online learning step, where the performance of the algorithm is improved by exploiting the relationship between atlases and the target image. For example, Van Rixooort et al. [58] divided the target image into blocks, where each block was used to improve the local registration process.

2.2.7 Label fusion: It is one of the most important steps of the MAS algorithm. In this step, the propagated labels are combined to get the final labels of the target image. Early MAS algorithms [45] [46] [47] employed a majority voting method, where the label of a pixel of a target image was decided by the majority of the pixel labels of the selected atlases. In recent papers, weighted voting is applied. For example, Artaechevarria et al. [59] described a weighted voting method, where weights are chosen based on local cross-correlation.

2.2.8 Post-processing: Label-fusion may not create the final segmented image. In that case, the segmented image is fed to another algorithm to estimate the labels of the final image. For example, Fritcher et al. [60] applied an edge-based geodesic active contour segmentation algorithm as the post processing step.

CHAPTER 3

PROPOSED ALGORITHM FOR 2D IMPLEMENTATION

In a supervised/unsupervised learning method, the system has to detect all nodules to increase sensitivity. Thus, the system must be trained with different types of lung nodules (different shapes, sizes etc.). Sometimes, it is challenging to differentiate a lung nodule from an artery or vein. Thus, a large set of training samples is needed so that the system doesn't "miss" any nodule. However, if a large set of training samples is used, the system will also create a lot of 'false positives'. On the other hand, attempts made to decrease the 'false positives' will also decrease the 'true positives' [4]. Figure 3-1 shows the relationship between TP (True positives) and FP (False positives).

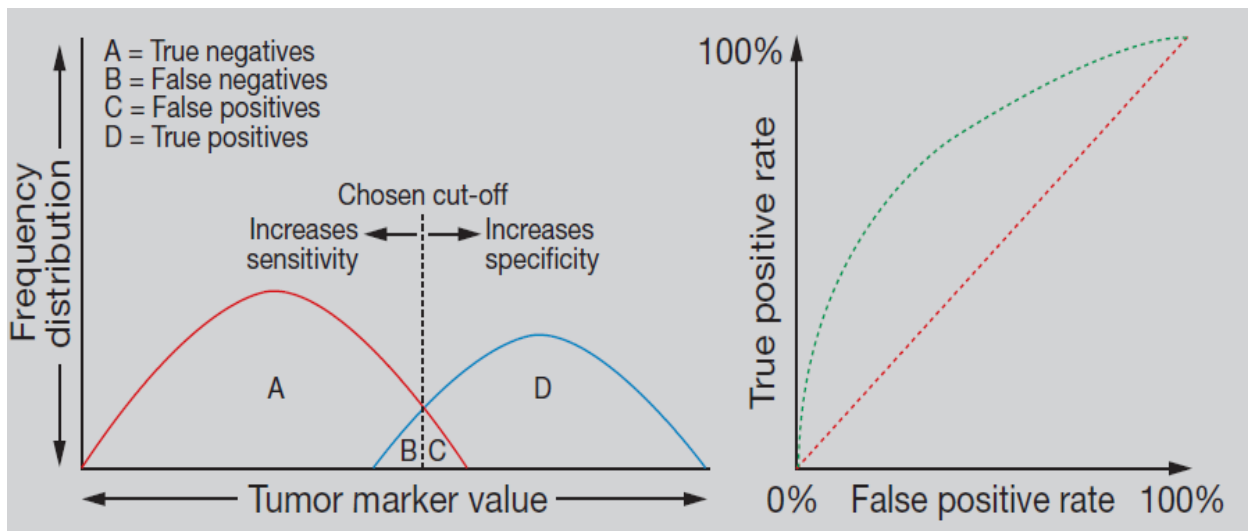


Figure 3-1 Relationship between TP (True positives) and FP (False positives) [61]

This problem can be solved in theory, if a set of optimally small number of training samples can be created, where each sample has lung nodules of similar size and shape as the

target image of the actual patient. A multi-atlas segmentation based method for 2D images is proposed to aid in this approach.

The details of the proposed method are shown in Figure 3-2. The proposed method has three steps. In the first step an atlas (already segmented image) selection process is developed, whereby a set of very similar atlases is selected. In this process, one, two or several atlases can be selected [48] by comparing the nodules in the nominal atlas images with the nodules in the target image. One of the novelties of the proposed approach is that a size and shape based feature detector is developed for the selection the atlases.

“Patch” refers to a subimage. In the second step, a patch-based method is chosen for lung nodule detection. In this process, each pixel of a target image is labeled by comparing an image patch, centered by the pixel, with patches from an atlas library and, choosing the most probable label [53] according to the closest matches. Note that in the second step, labeling is accomplished at the pixel level. While this approach makes it easier to get the nodule’s centroid position, it may be hard to find the properly segmented *shape*. The patch-based method is chosen, because it reduces the processing time without losing accuracy [54] . The nominal patch size chosen for 2D implementation is 9 x 9.

The third step is the post-processing step. In this step, the shortcoming with respect to the segmented shape determination mentioned above is mitigated by Laplacian of Gaussian (*LoG*) blob detection to find the accurately segmented area of the lung nodule [62]. The proposed algorithm is described in greater detail below.

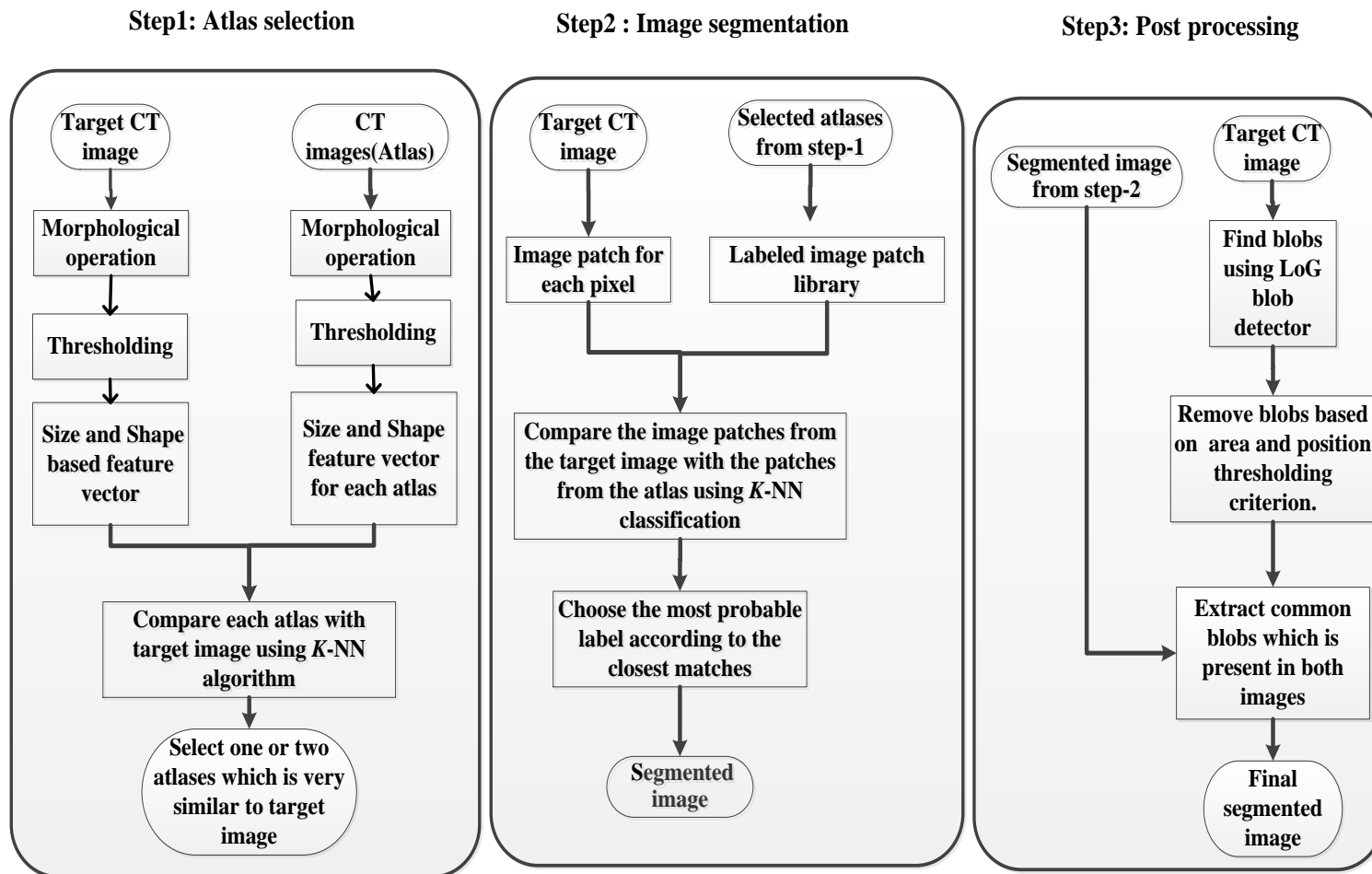


Figure 3-2 Block diagram of the proposed algorithm

3.1 Step1: Atlas selection

Atlas selection is the first step of the proposed algorithm. In this step, an atlas (already segmented image) is selected which is very ‘similar’ to the target image. The block diagram of this step is shown in Figure 3-3. In this step, as a preprocessing step, a morphological operation is first performed to reduce noise in the acquired target patient image. After that a thresholding process is performed to binarize the image. Thereafter, a size and shape based feature descriptor and the *K*-NN classification method are used for selecting the atlas.

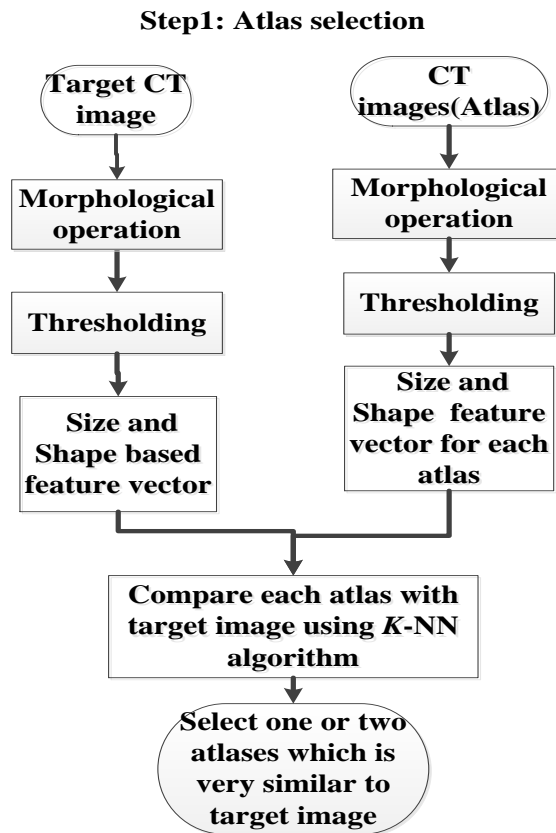


Figure 3-3 Block diagram of the proposed atlas selection step

3.1.1 Morphological Operation: A morphological operation is performed to reduce noise and to remove irrelevant objects (like fissure). In the proposed algorithm, the well-known morphological ‘erosion’ is applied to the image, in which the value of the output pixel is the minimum value of all the pixels in its neighborhood. This algorithm is geared toward detecting nodules of the size ranging from 3 mm to 30 mm radius [4]. Each pixel typically encompasses approximately 1 mm x 1 mm. Therefore, in the proposed algorithm, a disk-shaped structuring element with a radius of 3 pixels is chosen. Figure 3-4 shows the structuring element used in the algorithm.

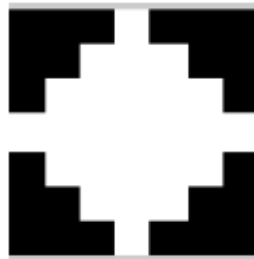


Figure 3-4 Structuring element used for morphological operation

3.1.2 Thresholding: By a thresholding process, the lung region is separated from the muscle, fat and parenchyma tissue. Thresholding works well where the intensity difference between different portions of the image is high. The intensity of lung tissue is mostly between -950 to -500 Hounsfield Unit (HU), and those of muscle fat and bone tissue are above -500 HU [63]. Thus, there is a significant intensity difference. In our method, the threshold is chosen to be the middle point between the highest intensity and the lowest intensity of a CT image slice.

3.1.3 Size and shape based Feature vector: As a part of detecting shapes of interest in each atlas image, a standard connected component analysis is performed using an 8-connected neighborhood. Given that the performance of the K -NN algorithm is sensitive to the *relevance* of the chosen feature vector, the feature vector was carefully chosen to reflect the characteristics of a typical nodule. Also, each connected component is associated with a significant number of features - nine were chosen in this step.

1) **Area:** Total number of pixels inside the connected component.

2) **Convex area:** The convex hull of a region is defined as the smallest convex envelope, which contains the connected component. In Figure 3-5 black lines show the convex hull of the object A.

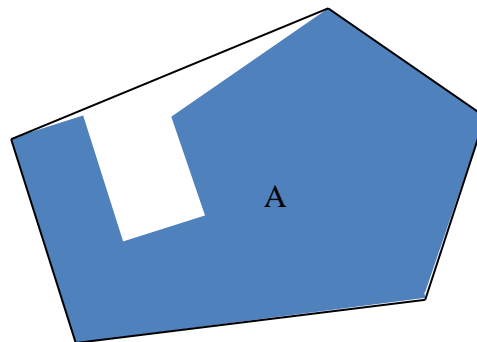


Figure 3-5 Convex hull of the object A

3) **Perimeter of the actual shape:** A perimeter is a path that surrounds the connected component.

4) **Convex perimeter:** Perimeter of the convex area.

5) **Equivalence diameter:** The diameter of a circle with the same area as the connected component. The calculation is shown in equation (3.1).

$$\text{Equivalent diameter} = \sqrt{\frac{4 \times \text{area}}{\pi}} \quad (3.1)$$

6) **Major axis length:** The length (in pixels) of the major axis of the ellipse that has the same normalized second central moment as the connected component.

7) **Minor axis length:** The length (in pixels) of the minor axis of the ellipse that has the same normalized second central moment as the connected component.

For 2D data, the *covariance matrix* is calculated from the x, y position of a connected component. For any *multivariate normal distribution*, the covariance matrix shows the shape of the distribution. The typical contour lines of a multivariate normal distribution results in an *ellipse*. To find the major axis length of that ellipse the eigenvalues of the covariance matrix is calculated. For a 2×2 covariance matrix, there will be two eigenvalues. One of the eigenvalues is the *major axis* and the other is the *minor axis*.

8) **Circularity:** A value that shows how circular a connected component is. It is calculated from

$$\text{Circularity} = \frac{4\pi \times \text{area}}{(\text{Perimeter})^2} \quad (3.2)$$

The value of circularity is in the range [0 1]. 0 corresponds not circular at all, and 1 suggests fully circular.

9) **Centroid:** The center of mass of the region. It is calculated by averaging the x and y positions of all the points inside the connected component.

3.1.4 Removing big connected components: After finding all the connected components in an image, the connected components, which have greater than 50 mm x 50 mm (64 pixel \times 64 pixels) are removed, since most of the lung nodules will likely have the size less than 50 mm square [4].

For each image and for each of the first eight features, three parameters are calculated (max, average and minimum). For example, for the area feature, three feature points will be:

1. Maximum area of all the connected components in an image.
2. Average area of all the connected components in an image
3. Minimum area of all the connected components in an image.

Similarly, the centroids are calculated (x , y position of leftmost, rightmost, topmost and bottommost connected component). The number of connected components is also added as a feature point. Thus, the total number of feature points for each image (both nominal atlases and target patient image) is $8 \times 3 + 8 + 1 = 33$. Using these feature points, a 33-element feature vector is created for each nominal atlas image (also called training images in the context of the selection process) and the target image. Feature vectors are created for all the images (target image and all the atlas images).

3.1.5 K-NN algorithm: In the next step, a K -NN (K nearest neighborhood) algorithm [64] is applied to find the feature vector, which is closest to that of the target image. The K -NN algorithm is a well-known and simple class-based algorithm, in which its K nearest neighbors classifies the output sample. The K nearest neighbors are considered to be in the same class. If the user decides that $K=1$, only one nearest neighbor will be chosen. In that case, the class of the output sample will be the class of the closest sample. For each feature point, L^2 - normalization is performed for all the images, which in effect makes the weight of each of the nine features equal.

3.2 Step2: Image segmentation using image patch comparison

In this step, a patch-based image segmentation method is applied for lung nodule detection. In this method, each pixel of a target image is labeled by comparing the image patch centered by that pixel with patches from an atlas library. The most probable label [53] is chosen based on a Euclidean distance measure applied to the K -NN classifier.

In classification literature, a patch, in general, is defined as a small sub-volume of the image. In our 2D patch-based method, a patch refers to pixels instead of voxels with similar surrounding neighbors. Figure 3-7 describes the workings of a patch-based comparison.

Step2 : Image segmentation using patch comparison

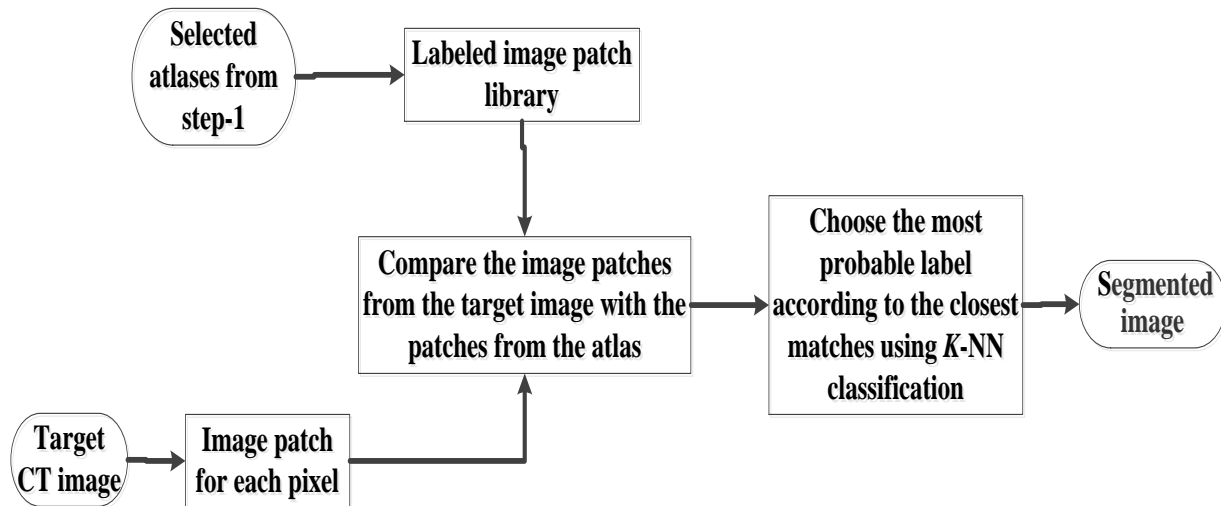


Figure 3-6 Block diagram of the proposed image segmentation step using patch comparison

3.2.1 Feature vector develop: For each image patch, a feature vector is created. The following logic is used to determine the five features: 1) Gray value: Since the gray value associated with a lung nodule is typically distinctly different from the gray values associated with the rest of the lung area and bronchiole, it is a good distinguishing characteristic and therefore a candidate for a feature [27] . Thus, mean value, area, and maximum intensity of the image patch area are selected as feature points. 2) Spatial information, i.e. x and y coordinates of the pixel under consideration for segmentation is clearly a useful feature point [24] . Thus, a five-dimensional feature vector is created for each image patch.

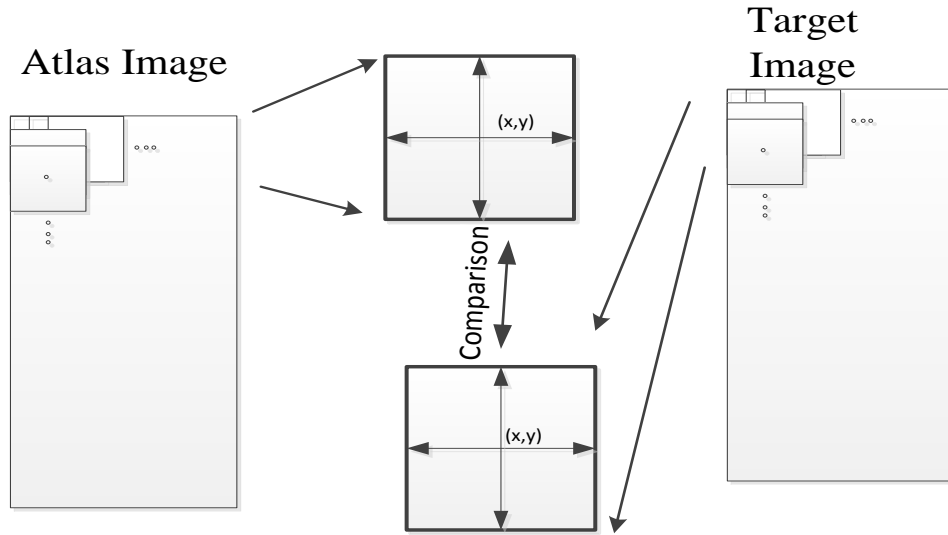


Figure 3-7 Image patch and their comparison

3.2.2 Label Selection: In this step, an image patch library is created from the selected atlases. An image patch can be automatically selected from any area of the atlas (A significant number of patches can be optionally selected from the nodule area using manual intervention). From each image patch, a feature vector is created. Similarly, for the target image, a feature vector needs to be created for all the pixels of the image. Thereafter, a K -NN classification step is used to choose the most probable label.

Let us assume,

For a selected atlas a , $P(a_j)$ is the feature vector for pixel j of the atlas image.

For the target image t , $P(t_i)$ is the feature vector for pixel i of the image.

Thus, the Euclidian distance between two patches is

$$d_{ij} = \left\{ \left\{ P(a_j) - P(t_i) \right\}^2 + \alpha \{i - j\}^2 \right\} \quad (3.3)$$

Here α = optional weighting applied to control the influence of spatial correspondence. In our implementation, the value of α is taken as 10 [54]. If the Euclidian distance d_{ij} between patch $P(t_i)$ and $P(a_j)$ is a minimum, the label of the pixel i of the target image will be the label of the pixel j of the atlas image.

$$Label_{pixel\ i\ of\ the\ target\ image} = Label_{pixel\ j\ of\ the\ atlas\ image}$$

3.3 Post Processing

Using patch-based image segmentation described above, it is possible to get the nodule position, but often it may be difficult to get the properly segmented area. In this step, the shortcoming with respect to the segmented shape determination mentioned above is mitigated by *LoG* (Laplacian of Gaussian) *blob* detection to find the accurately segmented area of the lung nodule [62]. Figure 3-8 shows the post-processing step.

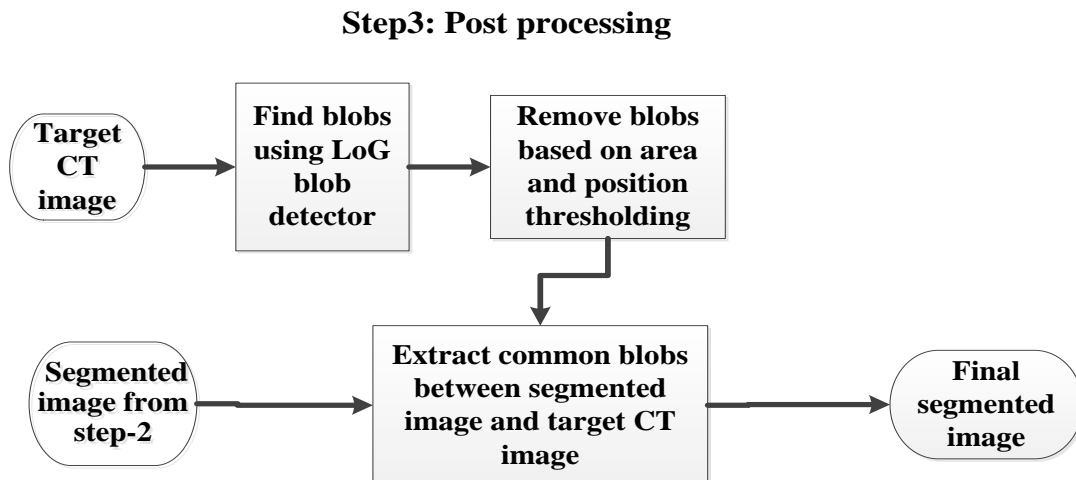


Figure 3-8 The post processing step

The final segmented image is obtained using the equation

$$I_{segmented} = I_{LoG\ convolved} \cap I_{patch\ based\ segmentation}$$

CHAPTER 4

EXPERIMENTAL RESULTS FOR 2D IMPLEMENTATION

Some of the experimental results are shown in this chapter. One image is taken from one patient and treated as a target image.



Figure 4-1 Target image from one patient

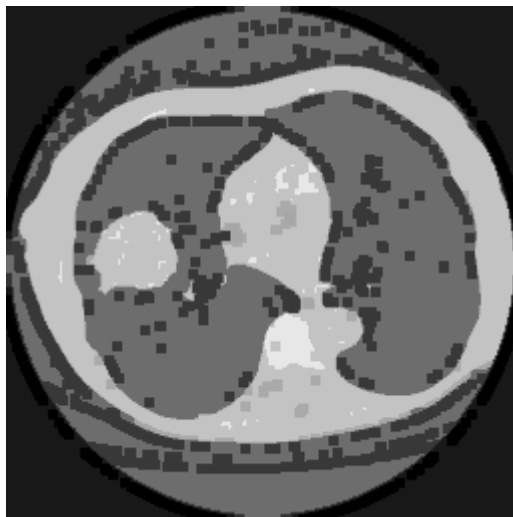


Figure 4-2 Target image after morphological operation



Figure 4-3: Output of thresholding of the image

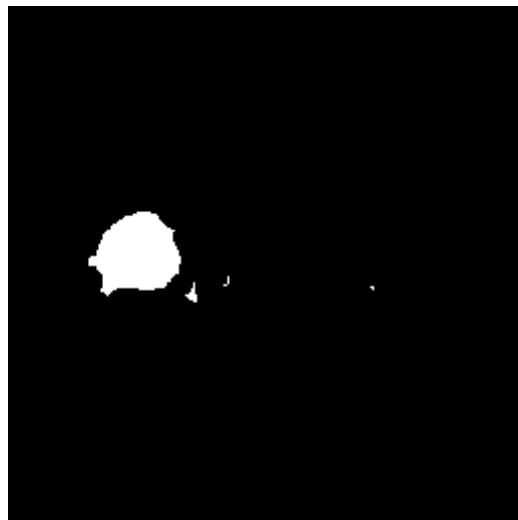


Figure 4-4 Likely candidates for nodules based on eliminating areas larger than typical nodules

The tumor was manually observed and understood. It was verified that Figure 4-4 indeed detected the tumor correctly in the target image. Now, this target image is ready to be tested with the proposed algorithm against the atlas image(s) from *a second patient*. Table 4-1 illustrates the typical value of the individual feature vector elements. At this time, no significance has been assigned and all the parameters are equally weighted.

Table 4-1 Size and shape based feature vector values of the target image

	<i>Feature</i>	<i>Value</i>
1	No of connected components-	4.00
2	Left most Centroid x-	123.70
3	Left most Centroid y-	64.68
4	Right most Centroid x-	143.67
5	Right most Centroid y-	92.26
6	Bottom most_ Centroid x-	64.68
7	Bottom most_ Centroid y-	123.73
8	Top most_ Centroid x-	181.67
9	Top most Centroid y-	140.33
10	Maximum area-	1316.00
11	Average area-	338.00
12	Minimum area-	3.00
13	Maximum Convex Area-	1452.00
14	Average Convex Area-	374.75

	<i>Feature</i>	<i>Value</i>
15	Minimum Convex Area-	3.00
16	Maximum Perimeter-	150.47
17	Average Perimeter-	47.15
18	Minimum Perimeter	3.41
19	Maximum Convex Perimeter-	43.00
20	Average Convex Perimeter-	20.00
21	Minimum Convex Perimeter	9.00
22	Maximum Equivalent Diameter	40.93
23	Average Equivalent Diameter	12.88
24	Minimum Equivalent Diameter	1.95
25	Maximum Major Axis Length	44.03
26	Average Major Axis Length	15.60
27	Minimum Major Axis Length	2.58
28	Maximum Minor Axis Length	39.06
29	Average Minor Axis Length	12.09
30	Minimum Minor Axis Length	1.76
31	Maximum Circularity	0.35
32	Average Circularity	0.16
33	Minimum Circularity	0.03

Size and shape based feature vector is applied on the second patient image with 212 atlases to find the image, which may have a similar lung nodule to the test image. These 212 images are now treated as the nominal atlases.

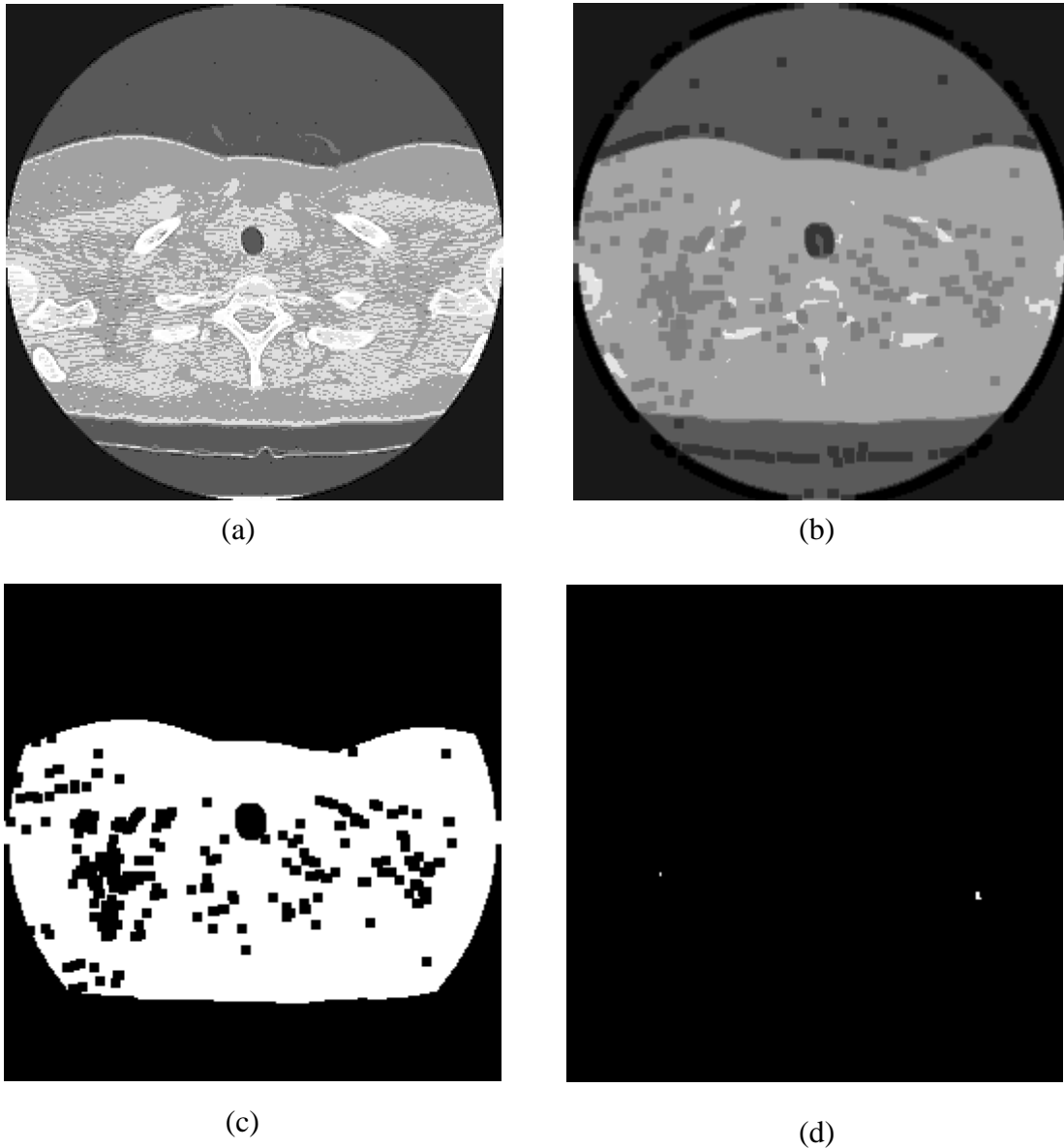
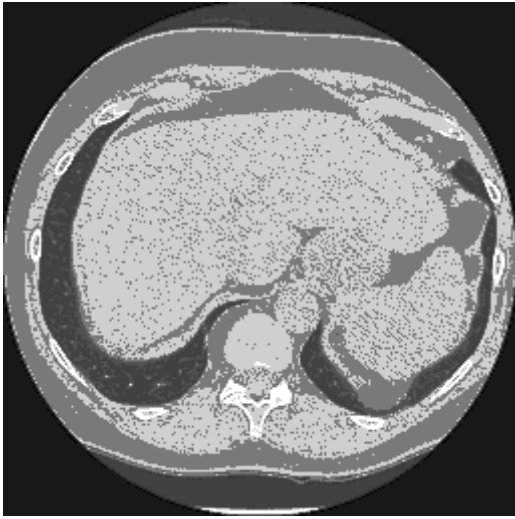
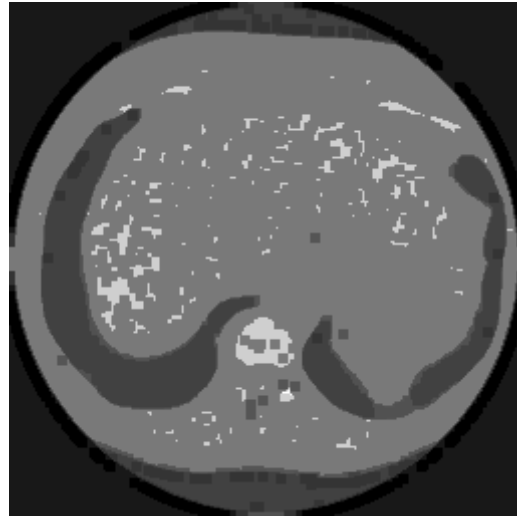


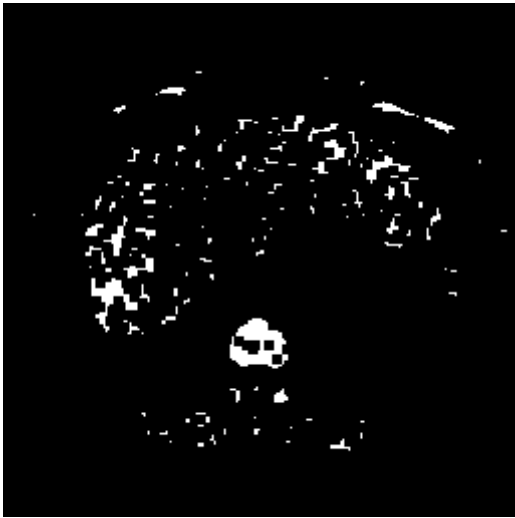
Figure 4-5 Four steps of the atlas image taken very close to the neck for patient 2 (a) Original CT image (b) After morphological operation (c) After converting to a binary image and (d) After removing connected components due to a bigger area. It can be manually verified that this is not a good atlas image.



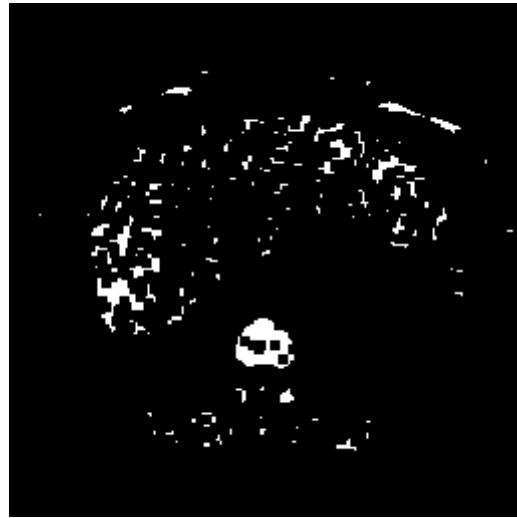
(a)



(b)



(c)



(d)

Figure 4-6 Four steps of another atlas image taken very close to the diaphragm of patient 2 (a) Original CT image (b) After morphological operation (c) After converting to binary (d) After removing connected components due to a bigger area. Clearly, this is not a good atlas image either.

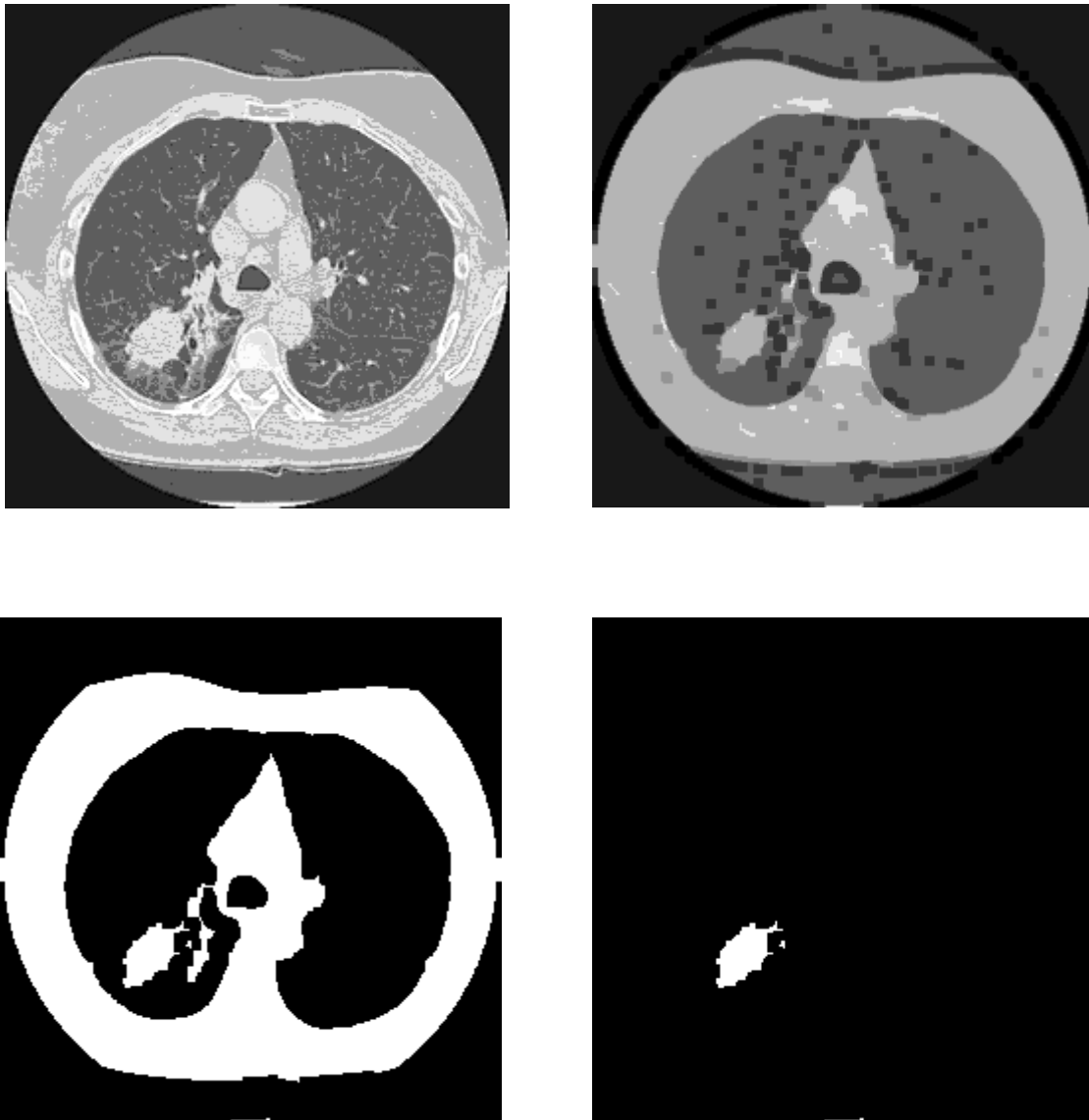


Figure 4-7 Four steps of the third atlas image taken at mid-level between neck and diaphragm for patient 2 (a) Original CT image (b) After morphological operation (c) After converting to binary (d) After removing connected components due to a bigger area. This is the atlas closest to the target image.

Figure 4-8 shows the normalized squared Euclidian distances between the target image and the different atlas images in increasing order.

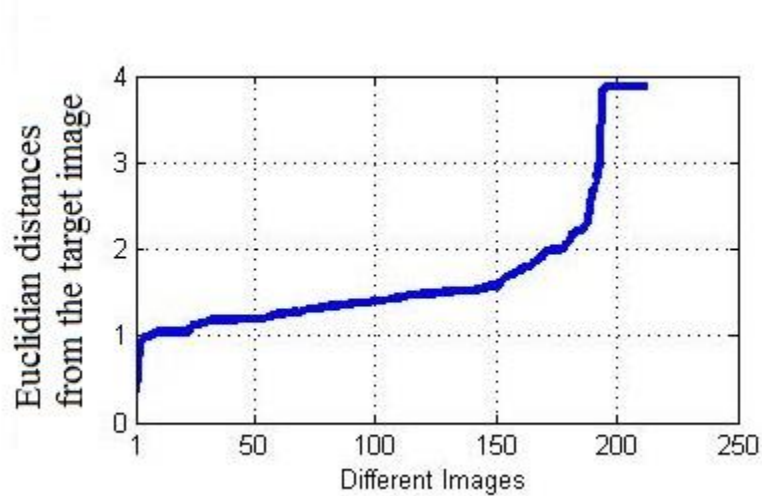


Figure 4-8 Normalized Euclidian distances from the target image

After the atlas selection step, the most ‘similar’ image (atlas) found is shown in Figure 4-7. The automated atlas selection process successfully selected this atlas image as can be manually verified.

Table 4-2 shows the Euclidian distances calculated between the target image feature vector and the atlas image feature vector for all the three images (atlases) shown above. The Euclidian distance for Figure 4-7 was found to be 0.47, which proves that this target patient essentially has the same nodule or tumor as the well understood atlas from the patient.

Table 4-2 Figure number and Euclidian distance from the target image

Figure number	Euclidian distance from target image
Figure 4-7	0.47
Figure 4-5	2.65
Figure 4-6	1.35

This can also be manually verified by examining Figure 4-4 and Figure 4-7. Once the atlas was selected, the step 2 was run with the target image. The final output is shown in Figure 4-9. It can be seen that, the output is logically the correct one. Figure 4-9 shows the output where a single atlas is used in the second step.

Although it might seem inappropriate to run the step 2 with just one atlas image, the exercise proved that when a correct answer is presented to the detection scheme, it finds it. Note the atlas selection process is also automated. Also, this effort helped validate the code.

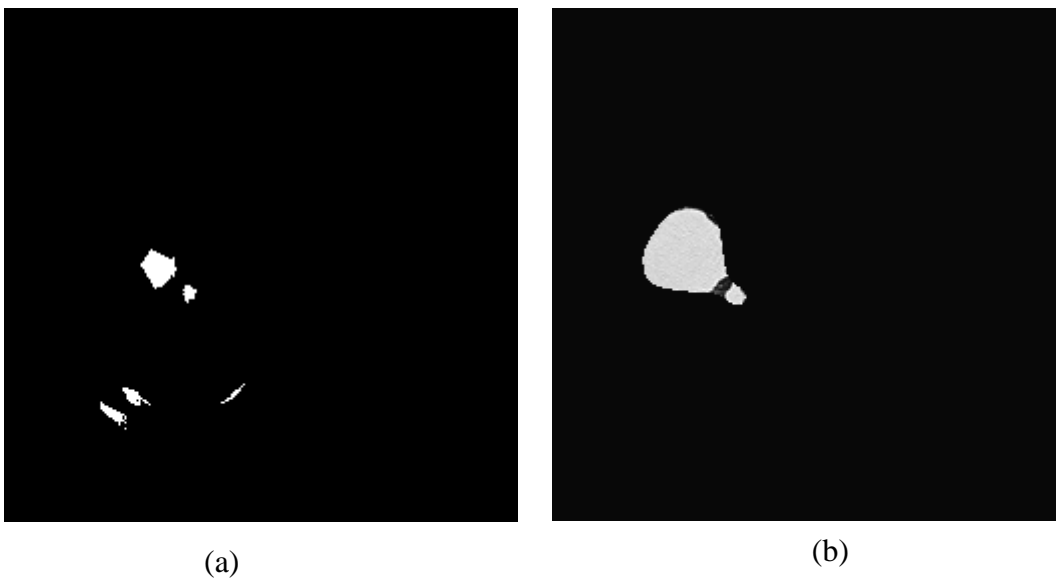


Figure 4-9 (a) Output of the second stage (b) The final output with the segmented lung nodule

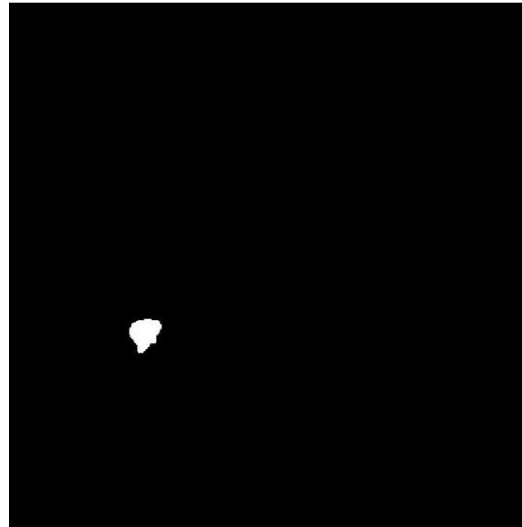
An added benefit, if testing with a single patient is that the atlas selection process can be efficiently used to cull all the slices corresponding one patient to select the most likely atlas candidate(s) from that one patient. This could be a major labor and time saver for the radiologist.

While the atlas selection process might be useful with a single patient's slices, recall that the objective of this overall dissertation is to compare the current patient's tumor (nodule) to that a variety of well-marked and understood nodules from multiple patients so that the current patient's tumor can be identified. In order to accomplish this, multiple atlases may be selected from multiple patients during step-2 of the algorithm.

In the following, the step -2 of the experiment was rerun with the atlases from three patients. Thus, in this experiment, three atlases of three patients pre-selected. The atlases are shown in Figure 4-10.



(a)



(b)



(c)



(d)

Figure 4-10 Two atlases manually selected from two additional patients as input to the step-2 of the program (a) Atlas – 2 lung image (b) Segmented lung nodule from atlas – 2 (c) Atlas – 3 lung image (d) Segmented lung nodule from Atlas - 3

Figure 4-11 shows the final segmented image after post-processing.



Figure 4-11 The final output with the segmented lung nodule

It can be seen that the experiment with three atlases corresponding to three different patients continues to validate the algorithm in that it found a nodule in patient 1 that is closest in size and shape to the target patient (manual comparison to Figure 4-4). Thus, if patient 1 had a specific sarcoma x, then the system would provide a nominal automated diagnosis for the target patient.

Pixel by pixel patch comparison is a very accurate but a time-consuming segmentation process, because to compare two whole images of $512 * 512$ pixels we have to do the patch comparison of $512 * 512$ times for each pixel. Thus, in our proposed algorithm, size and shape based feature vector is used to reduce the number of atlas images. Size and shape based feature vector is much faster because it only calculates different feature values of probable nodules in each image. In our process, three atlas images have been selected in the atlas selection stage from 212 atlases. Thus, the atlas selection process reduces processing time by orders of magnitude

4.1.1 Final result: The 2D method is tested for more than 25 test images, where each test image is applied to more than 200 atlas images.

Five of the test images are known to have lung nodule and they are chosen to determine the sensitivity of the algorithm. The algorithm accurately detected each of the non-attached nodules. Thus, for nodules in the size between 3 mm to 30 mm, the sensitivity of the proposed algorithm is 100% for the given test data. Appendix A presents these results.

The twenty non-noduled images were evaluated for the occurrence of false positives. None of them showed false positive. Appendix B presents these results.

CHAPTER 5

PROPOSED ALGORITHM FOR 3D IMPLEMENTATION

The 3D implementation of our proposed algorithm has three steps, the block diagram of which is shown in Figure 5-2. In the first step, which is called the *nodule candidate detection*, nodule candidates of the current patient are detected. The second step is the *atlas selection step*, in which two or three very similar lung images (atlas image) are selected from a group of atlases (patient images) by a *nodule based atlas search* process. In the third and final step, the patch comparison process is developed to determine the accurate size and shape of the lung nodules. The details of the method are explained in following paragraphs.

5.1 Step 1: Nodule candidate detection in both atlas and target images

In the nodule candidate detection step, all the possible nodule candidates are separated and different features of the nodules are extracted. Figure 5-1 shows the *nodule candidate detection* step. In this step, CT voxel images are acquired and three different preprocessing steps are applied to all the CT images to remove noise. They are a) thresholding b) lung-region extraction and c) morphology. After that ‘Contour Correction’ process is applied to find out the ROI.

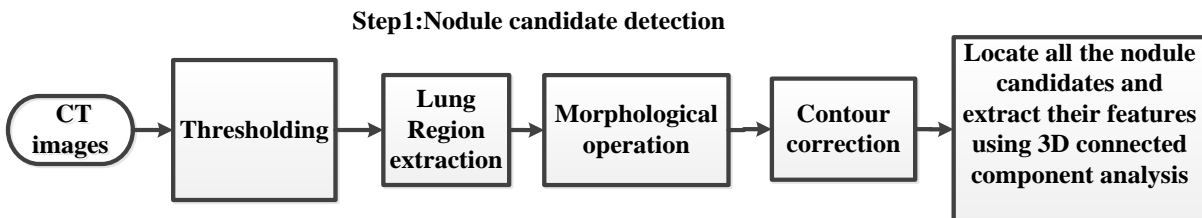


Figure 5-1 Block diagram of the nodule candidate detection step

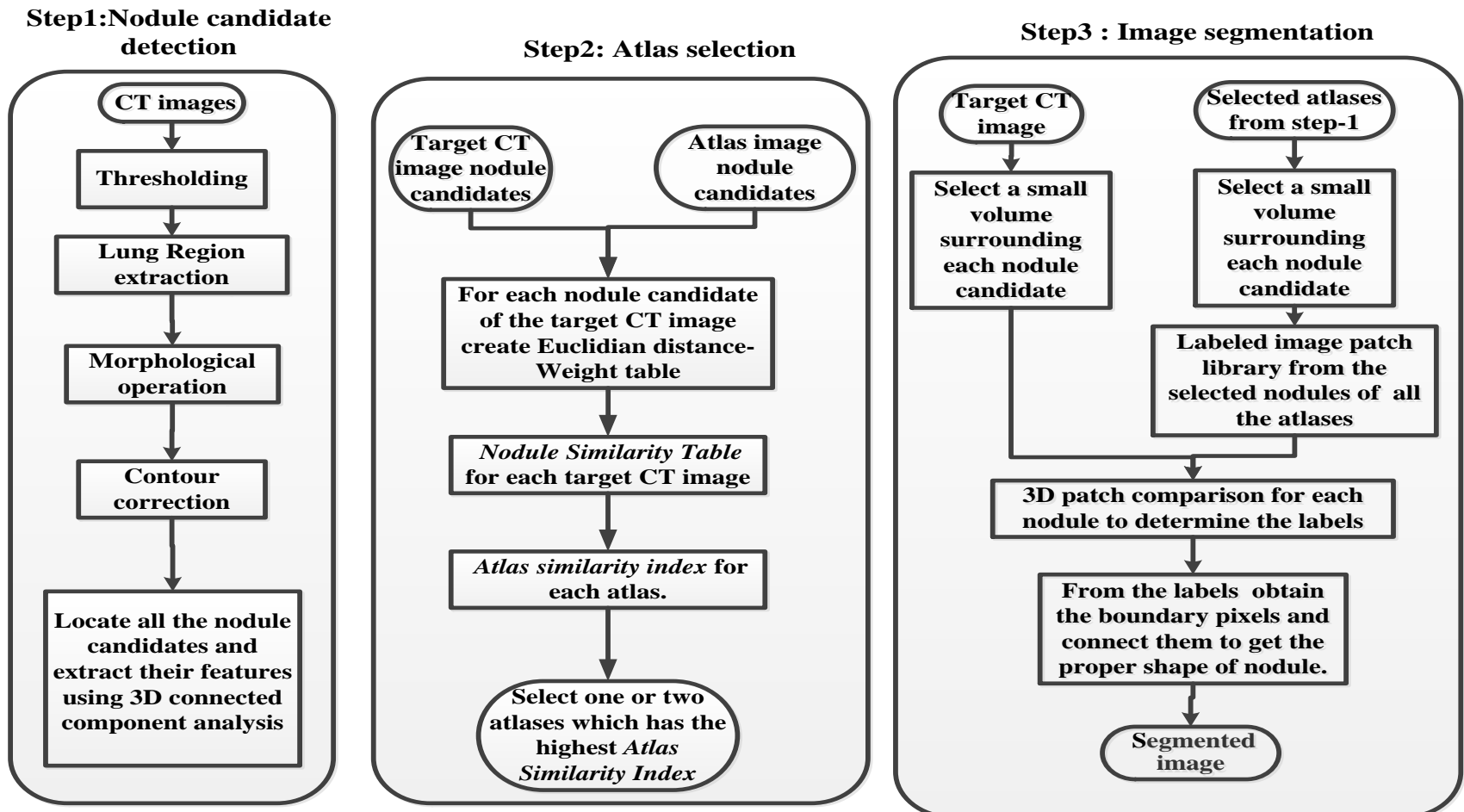


Figure 5-2 Block diagram of 3D implementation

5.1.1 Thresholding: By thresholding process, an object is separated from the background in an image. Iterative thresholding method is used here to separate the lung region from the background [23]. In this method, an initial threshold value of $T^{(0)}$ is assumed, using which, the mean intensity of the background (μ_b) and the object (μ_o) are calculated. Using these two mean intensities a new intensity $T^{(1)}$ is calculated. The threshold of the i -th iteration is calculated from equation (5.1).

$$T^{(i)} = \frac{(\mu_o)^{i-1} + (\mu_B)^{i-1}}{2} \quad (5.1)$$

The iterative process of the thresholding continues until the value of the threshold converges. The converged threshold value is a constant value for a particular image regardless of the initial threshold value ($T^{(0)}$).

5.1.2 Lung region extraction: In this step, a 3D connected component labeling method is used to remove the area outside of the lung region.

5.1.3 Morphological operation: Morphological erosion is applied to remove noise from 3D image. A spherical structuring element is used to implement erosion.

5.1.4 Contour correction: A region of interest is defined as the part of the image where a nodule is being searched. If any nodule is attached to the wall of the lung, it is challenging to find it using only the thresholding process. A healthy lung has a well-contoured shape which typically doesn't have any sharp discontinuity on its wall. If there is any sharp discontinuity does exist it is likely due to a nodule attached to the wall of the lung. Based on this hypothesis, the contour of the lung wall is extracted and, a contour correction method is applied in several steps in our algorithm. Figure 5-5 shows the block diagram of the overall contour correction process. Each 2D slice is split in half (the rationale will be explained later) in the axial direction and the contour correction is applied to each half. The two halves are then re-attached.

Thereafter, the coronal view derived from the processed axial view is considered for further processing. The logic is that, some nodules, that might not have been visible in the processed axial view, can be captured on further examination of the derived coronal view. Then, as before, a 2D morphological erosion is performed in the coronal direction to remove unnecessary edges created during the contour correction process in the axial direction. The images are split into halves in the coronal direction and the contour correction process is applied in each half. The resulting two slices are finally attached to make a complete 2D slice.

The reason for splitting the image is to separate the left and right lungs and individually apply the contour correction process to each lung. There are many critical points in the axial images at the posterior section of the lung regions. If the left and right lungs are not separated, these critical points may connect the two lungs. This could result in the region between the lungs being filled out and the heart being included in the ROI.

Figure 5-6 shows the details of the contour correction process for each 2D slice. The process starts with the initialization of the number of critical points N to 0. For each slice, an 8-connected chain coding method is applied to the lung boundary to determine the local lung contour [65] [66]. Figure 5-3 shows the 8-connected chain coding method.

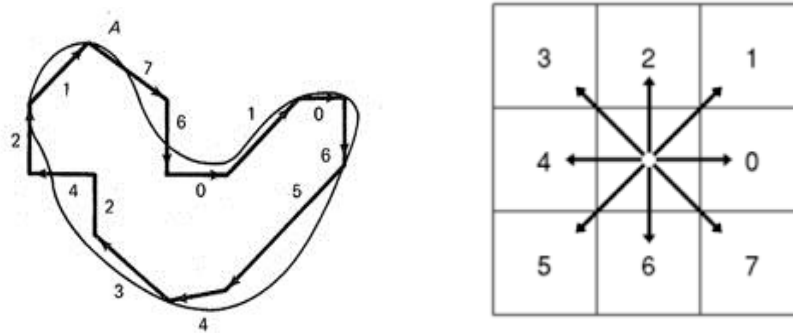


Figure 5-3 8-connected chain coding

A critical point is defined, if two consecutive concave points have the angle 135° or, one concave point has an angle of 90° [67]. Figure 5-4 shows different combinations of concave points of angle 135° and angle 90° to form a critical point.

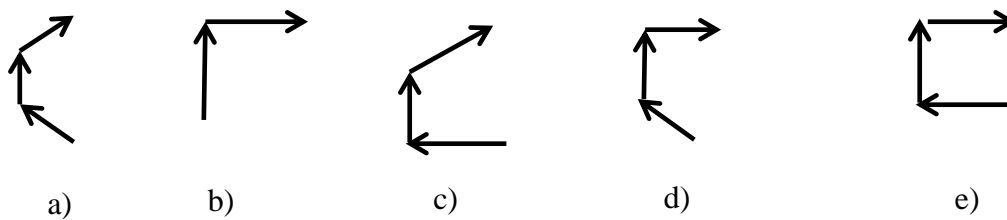


Figure 5-4 Critical points of different angles a) two consecutive points of 135° b) One concave point of 90° c) One point of 90° and one point of 135° d) One point of 135° and one point of 90° e) Two points of 90°

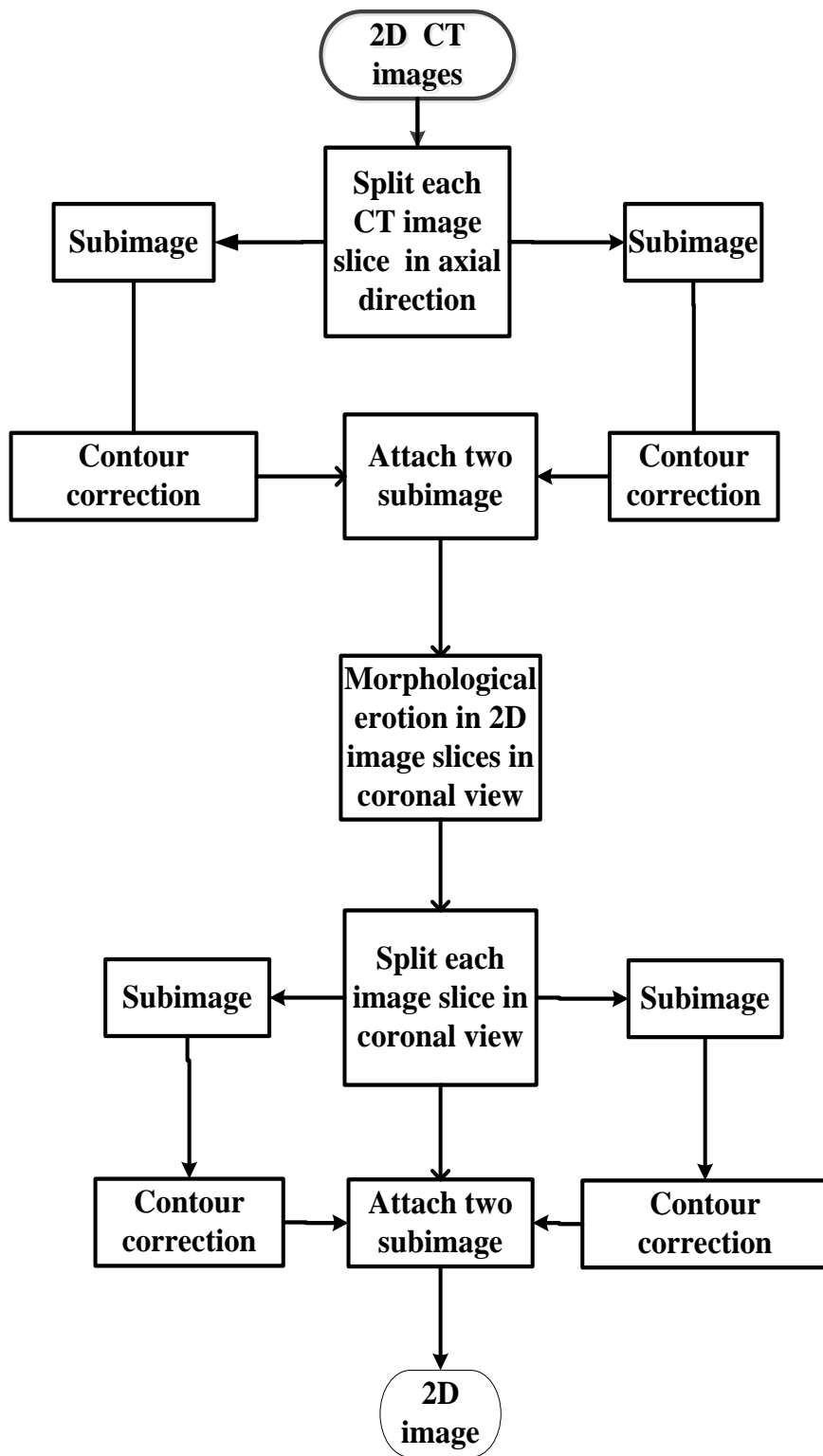


Figure 5-5 Contour correction process (overall block diagram)

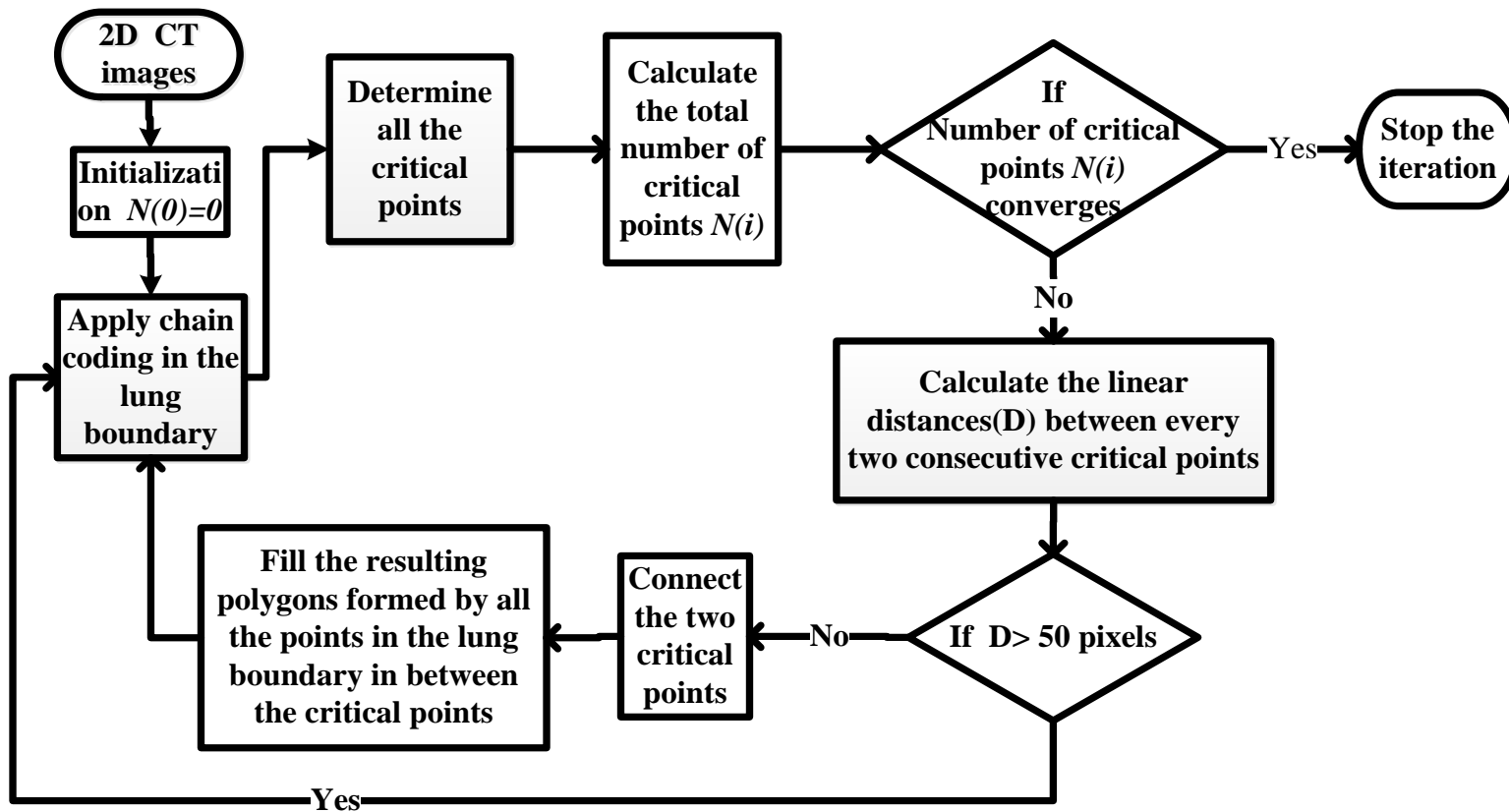


Figure 5-6 Contour correction process for a 2D slice

Thereafter, all the critical points are detected, and distances in number of pixels between every two consecutive critical points are measured. The presence of a critical point indicates the presence of a sharp curvature inside the lung wall. If the distances (D) between two consecutive critical points is less than, say 50 pixels, there is a high possibility that a small nodule is attached to the lung wall [65]. Then, the critical points are connected, and any gap is filled out. If the distance is too large (say, > 50 pixels) the big curvature could be due to trachea or bronchus.

5.1.5 Nodule candidate detection using connected component analysis: After detecting the ROI by the contour correction process, a 3D connected component method is applied to extract possible nodule candidates. The 3D connected component analysis consists of 26-connected neighborhood voxels as shown in Figure 5-7. Each of these candidates is examined further to determine the various useful features for possible nodule labeling. Six different features are extracted from each nodule candidate.

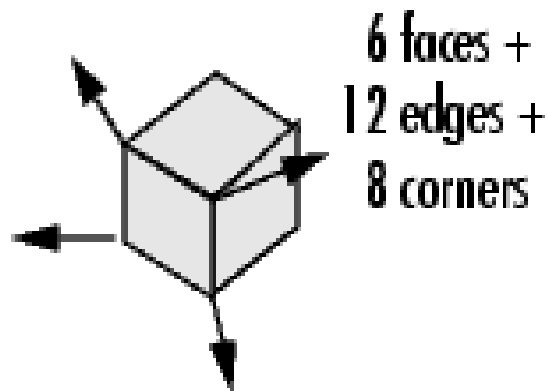


Figure 5-7 3D connected component method using 26 connectivity [68]

The features are

- 1) **Volume:** Total number of pixels inside the connected component.
- 2) **Surface area:** Total surface area of the connected component.

3) **Sphericity:** A value that shows how spherical a connected component is. Equation (5.1) shows how to calculate the sphericity.

$$Sphericity = \frac{\pi^{\frac{1}{3}} (6 \times Volume)^{\frac{2}{3}}}{Surface\ area} \quad (5.1)$$

The value of sphericity is in the range [0, 1], where 0 corresponds not spherical at all and 1 suggests a perfect sphere.

4) **Equivalent diameter:** The diameter of a sphere with the same volume as the connected component.

$$Equivalent\ diameter = 2 \times \left(\frac{3 \times Volume}{4 \times \pi} \right)^{\frac{1}{3}} \quad (5.2)$$

5) **Meridional eccentricity:** It is the eccentricity of an ellipsoid formed by a plane through the longest and shortest axes of the connected component.

6) **Sphere volume ratio:** It is the ratio of the volume of the sphere which totally encloses the connected component to the volume of the connected component. The radius of the sphere is the distance between the centroid and the farthest point of the connected component Figure 5-8 shows a nodule enclosed by a sphere.

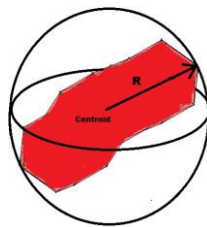


Figure 5-8 A nodule enclosed by a sphere

7) **Major axis length:** The length (in pixels) of the major axis of the ellipsoid that has the same normalized second central moments as the region.

5.2 Step 2: Atlas selection by nodule-based atlas search

In this step, a few atlases are selected based on size and shape features of nodule candidates. Figure 5-9 shows the block diagram of the proposed atlas selection step.

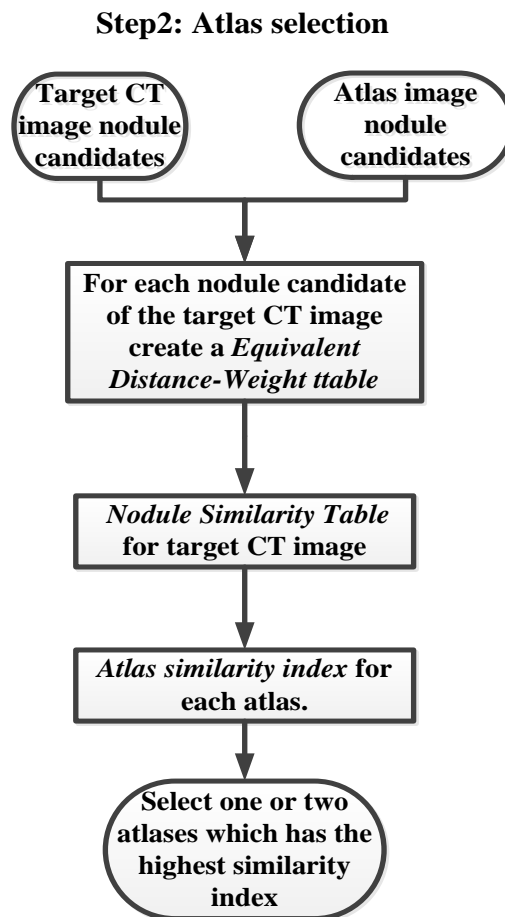


Figure 5-9 Block diagram of the atlas selection step

For all the images (target image and atlas image) extremely aspherical nodule candidates are removed (sphericity less than 0.2) because most of the cancerous lung nodules are known to be spherical [69]. For each nodule candidate of the target patient, most similar nodules are searched in all the atlas images in two steps.

First: The nodule candidates of the atlas image which are in the volume range of $10 \cdot A > V > A/10$ are searched, where A = Volume of the target nodule candidate and V = Volume of the atlas nodule candidate.

Second: A K nearest neighborhood (K -NN) algorithm is applied to find similar nodule candidates. The Euclidian Distance-Weight table is created for each nodule.

5.2.1 Euclidian Distance-Weight table: For each target nodule candidate, a Euclidian Distance-Weight matrix is created in which, a column has been added where the weights of atlas nodule candidates are calculated. The weights of each atlas nodule are calculated by equation (5.3)

$$weights_{atlas\ nodule} = \frac{ED_{smallest\ distance} \times ER_{target\ nodule}}{ED_{atlas\ nodule}} \quad (5.3)$$

Where, ED = Euclidian Distance.

ER = Equivalent Radius.

For each target nodule one *Euclidian Distance-Weight table* is created. These weights are used in the *Nodule Similarity Table* to determine the *Nodule Similarity Index* for each atlas.

5.2.2 Nodule similarity table: For the nodule candidates of the target image, which has the equivalent radius greater than 3mm, a table is created to show the most similar nodules (and their corresponding atlas ID and *weight* from Euclidian Distance-Weight table). The table is named as *Nodule Similarity Table*. From the nodule similarity table, weights of all the atlas nodules of a specific atlas are added to calculate the *Nodule Similarity Index* of an atlas. The logic for this operation is as follows: the overall objective of atlas selection is to find one or more atlases that are “similar” to the target patient CT image. “Similarity” can be logically considered high if a) the number of similar nodule candidates in both target patient CT and an atlas are high. This can be accomplished using the *K*-NN approach. However, if the number of nodules is considered the only parameter without considering the nodules’ sizes, then a very big nodule and a very small nodule in the same target image may end up being equally important. b) The size of the nodule can also be taken into account by modifying the Euclidian distance criterion to a weight criterion. To estimate the volume of the nodule candidates (as a measure of their sizes), their equivalent radii are therefore used as a part of the weight criterion. Now, if equivalent radius *and* the number of nodules are both considered for similarity estimation, the selected atlas will be the most “similar”. The number of nodules criterion is covered by the Euclidian distance within the *K*-NN approach. The size of the nodule criterion is covered by the weights. The nodule similarity table combines both of these criteria. The atlas with the highest *Nodule Similarity Index* is the closest atlas of our target image.

5.3 Step 3: Image segmentation

In this step, patch-based nodule segmentation method is applied to determine the proper size and shape of the nodule. Figure 5-10 shows the nodule segmentation using patch comparison process.

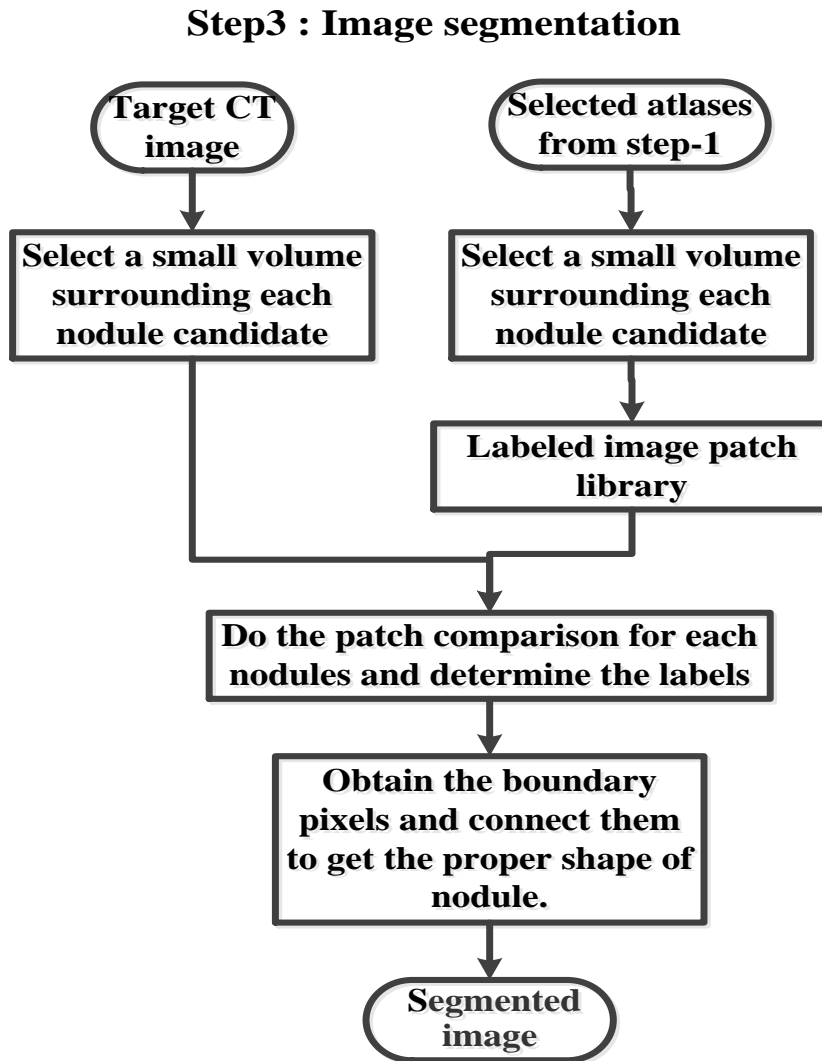


Figure 5-10 Image segmentation using patch-based segmentation

In image segmentation literature, a “patch” refers to a sub-volume of the 3D image. In patch based segmentation, for each pixel of a target image, a patch centered by the pixel is selected and is compared with other patches from a patch library. In the 3D algorithm, patches are created only from a cube surrounded by each lung nodule, instead of using the whole image. In other words, the segmented nodule drives the selection of the voxel patch for comparison. This process avoids unnecessarily looking for corresponding nodules over the entire CT scan. This also avoids the need to overtly register atlas image to the target image. The height, width and length are chosen as $2R$, where R = distance between the centroid and the farthest point of a nodule candidate. Figure 5-11 shows the image patches and their comparison process in our method.

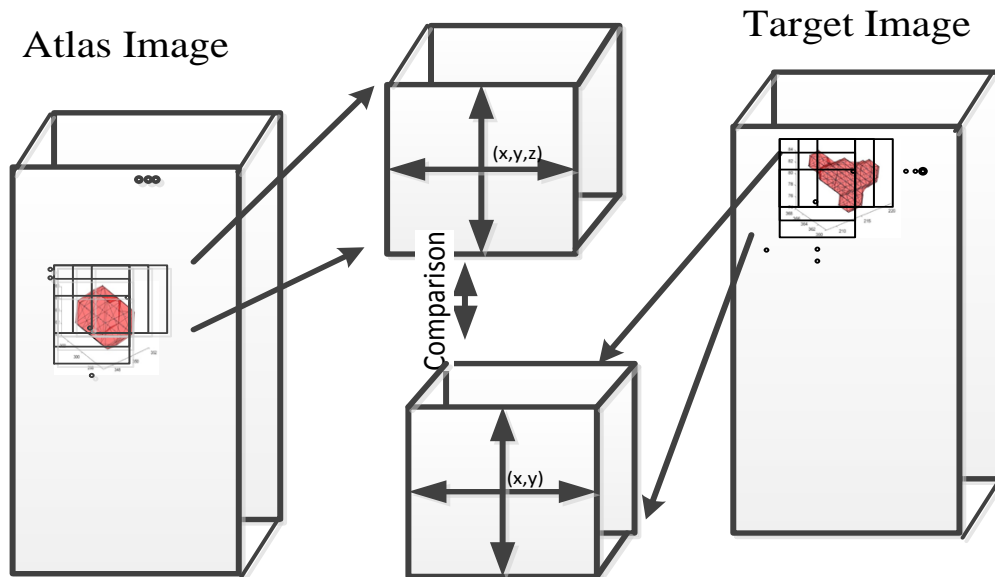


Figure 5-11 Image patches and their comparison

For the patch comparison, a feature vector is created [70] for each patch, and a K -NN classification is used for finding the closest patch. After finding the closest patch the label of the 'closest patch' is taken as the label of our target patch. The details of the process are same as in the 2D operation. Once the patch comparison process is completed, a boundary selection algorithm is used to determine the proper boundary of the nodule.

CHAPTER 6

EXPERIMENTAL RESULTS FOR 3D IMPLEMENTATION

In this chapter, representative results of the proposed 3D algorithm are presented. The incremental results and figures logically follow the block diagram in Figure 5-2. The implementation starts with acquiring CT images from RIDER databases [6]. Thereafter, each step of the proposed algorithm is performed through MATLAB coding. Figure 6-1 shows the 3D view of a deceased lungs rendered by 3Dimviewer [71].

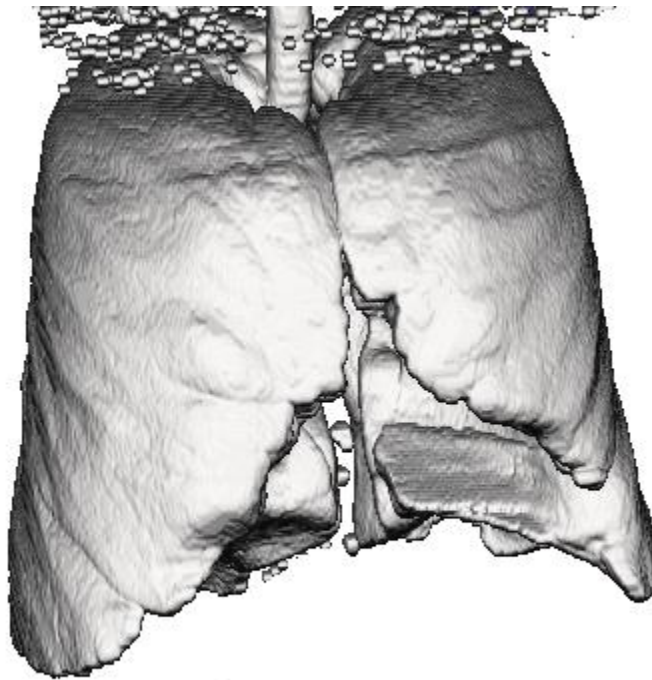


Figure 6-1 3D view of lungs (rendered by 3Dimviewer)

Figure 6-2 shows a 2D slice of an image before and after thresholding. Figure 6-3 shows that the threshold of the 3D image converges even though the initial intensities are different. The X-axis has the iteration number and Y-axis the intensity gray value. For instance, if 1800 was chosen as

the initial intensity, following the graph clarifies that the thresholding process eventually calculates the threshold to be approximately 411.

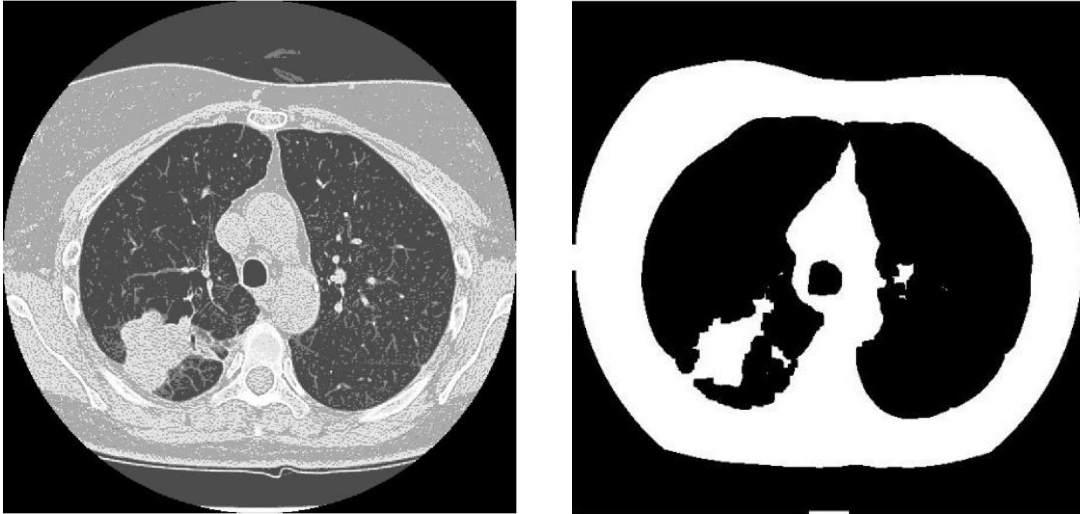


Figure 6-2 2D slice of an image a) before thresholding b) after thresholding

Also note that the choices of 10 or 800 as initial intensities also converge to 411.

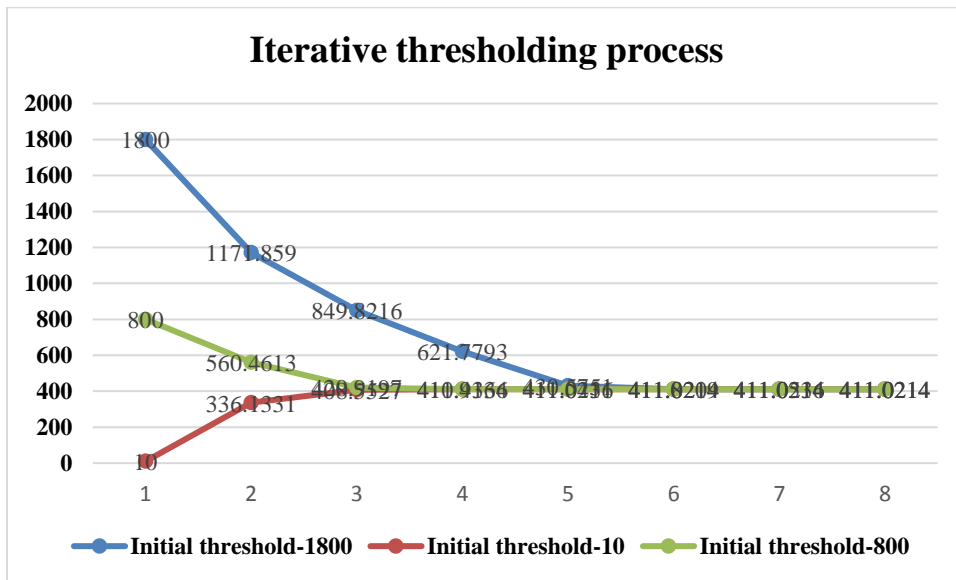


Figure 6-3 Three different initial thresholds and the final threshold for a 3D image

Figure 6-4 illustrates the histogram of the 3D image corresponding to the representative 2D slice shown in Figure 6-2. The 3D image consisted of 212 slices. From the histogram, it can be seen that high and low intensities are very clearly differentiable by the threshold.

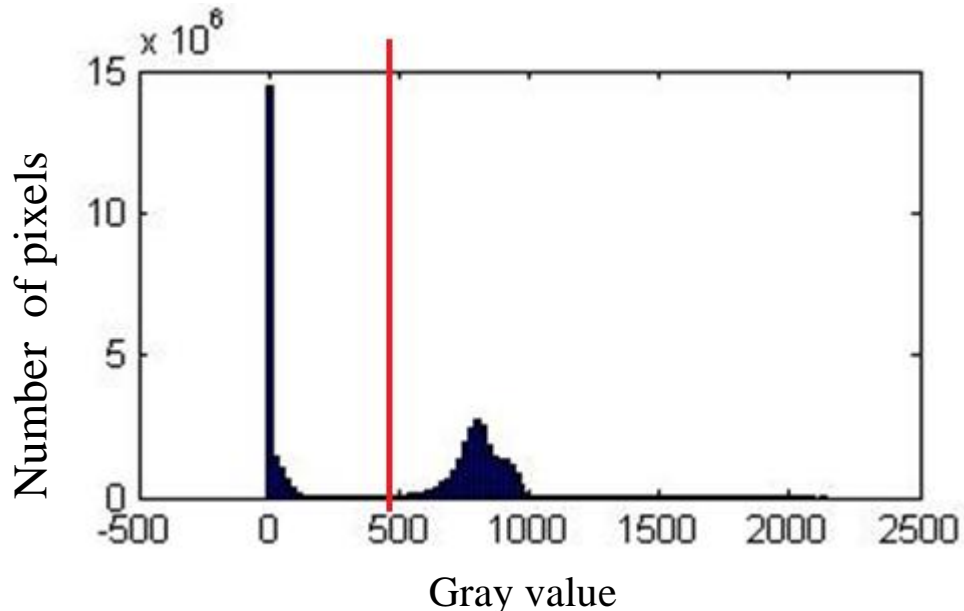


Figure 6-4 Histogram of the 3D image consisting of 212 2D slices

Figure 6-5 shows the output images of different iterations of the thresholding process, where the initial threshold is chosen as the gray level value of 1800. Figure 6-5(a) corresponds to the output with the initial intensity of 1800. In the next iteration, threshold intensity reduces to 1171.859. The image corresponding to this threshold is in Figure 6-5(b). Finally, the threshold value of 411 does not change over Figure 6-5(e), Figure 6-5(f) and Figure 6-5(g). The reason for this can be clearly understood by looking at the figure. It can be seen from Figure 6-3 that, in iteration 5, iteration 6 and iteration 7 thresholds are almost same and that is 411. Thus, the threshold is converged and the final threshold is 411. Similar convergence can be observed for initial threshold values of 10 and 800.

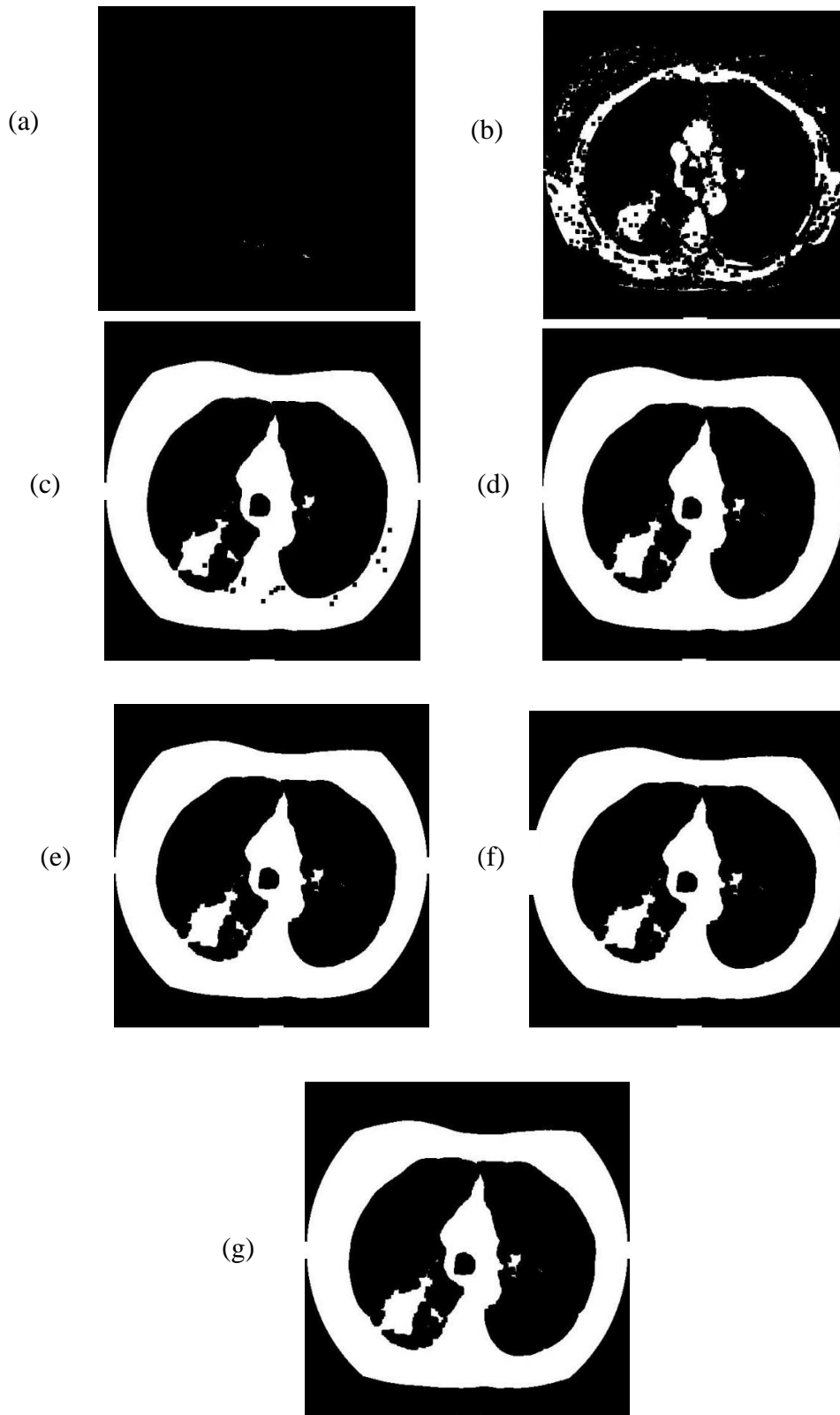


Figure 6-5 Seven different iterations of the thresholding process with the initial threshold of 1800

In the next step, the lung region extraction is completed. Figure 6-6 shows a 2D slice before and after the lung region extraction using 3D connected component method.

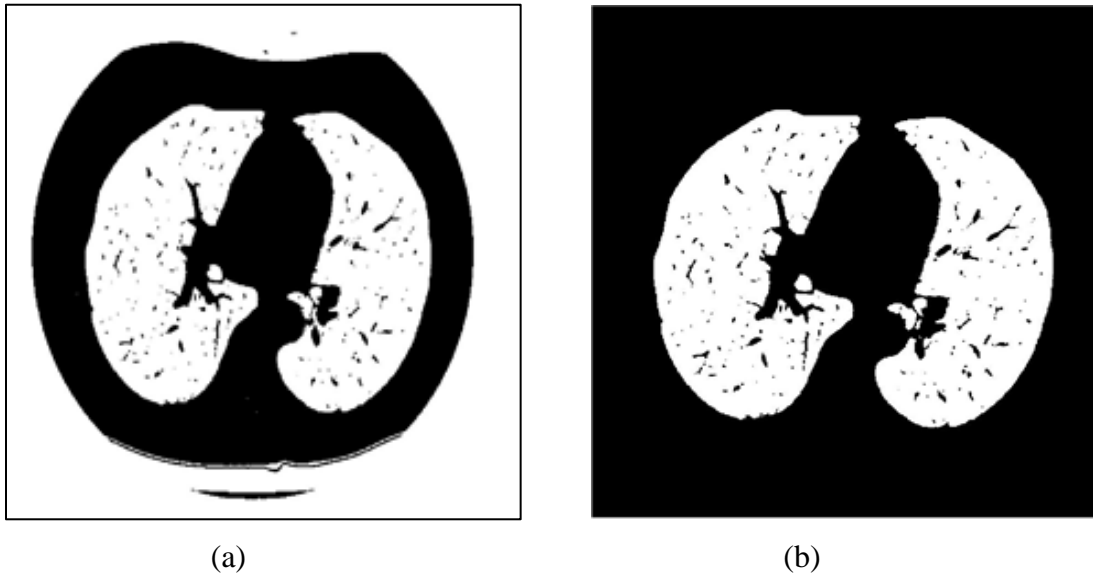


Figure 6-6 (a) before lung region extraction (b) after lung region extraction

Thereafter, the contour correction process is performed which is detailed in the following paragraphs.

Figure 6-7 shows a CT image, where the critical point selection is performed. In this image, the critical points are circled and the adjacent points are joined by blue lines. Thereafter, the polygons created by the connecting lines and the points in the lung contour in between the pair of critical points, are filled. Figure 6-8 shows the CT image where the polygons created by critical points are filled. This is the first iteration. In the next iteration, again critical points are selected and adjacent critical points are connected. In this iteration, the total number of critical points would be lower than the previous iteration due to filling out polygons. The process continues until the number of critical points is converged.

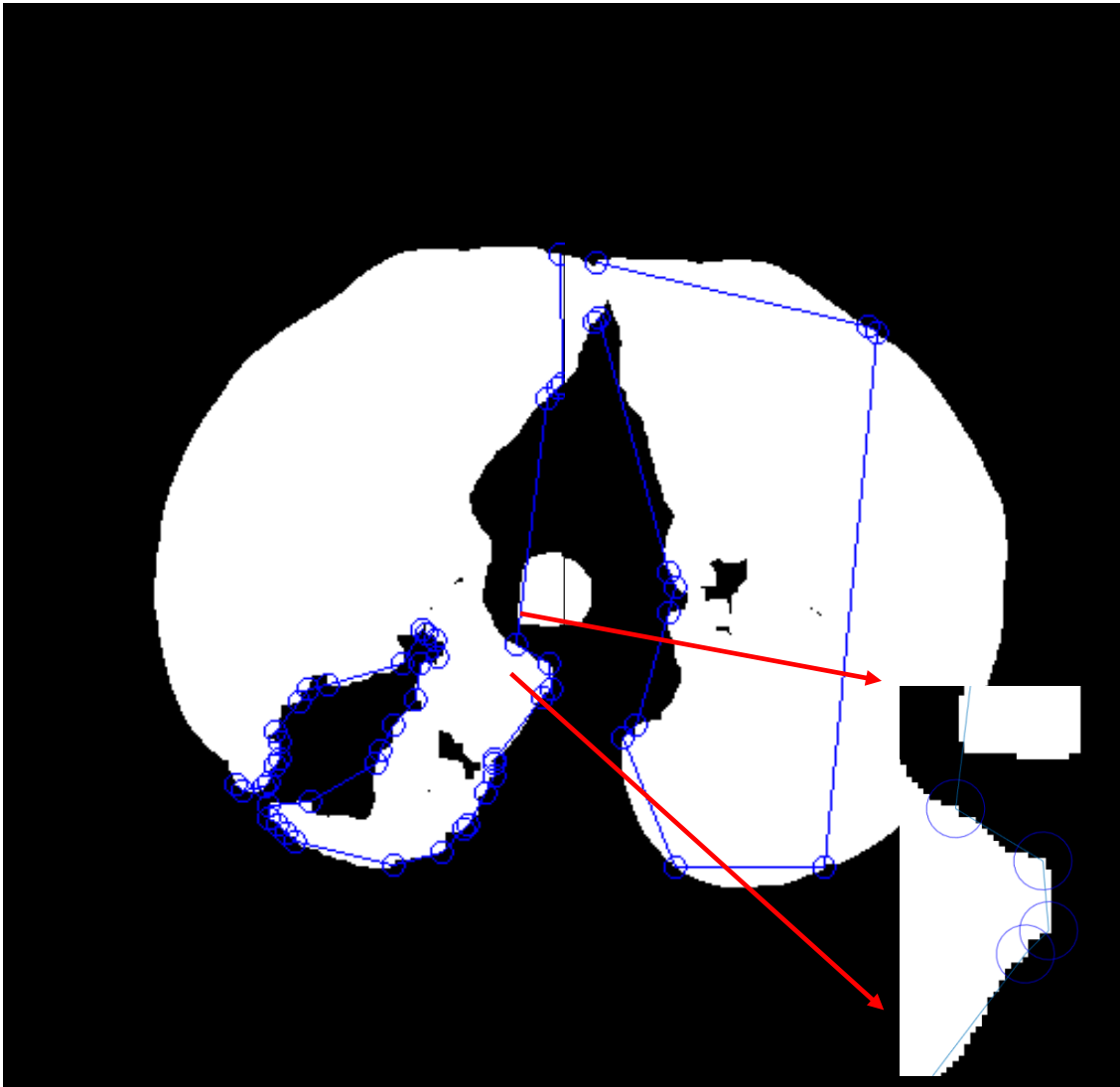


Figure 6-7 A CT image where the critical points are circled and adjacent points are joined.

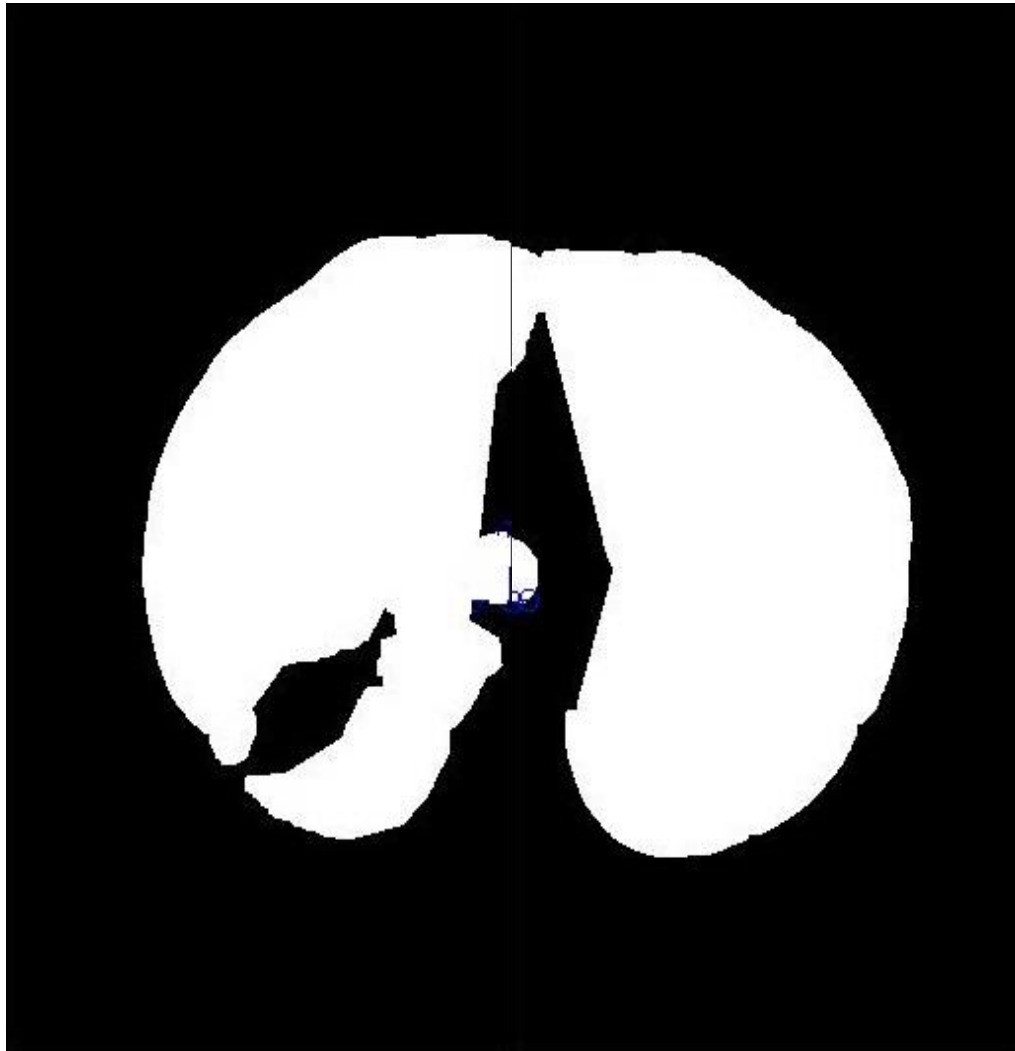


Figure 6-8 After filling out the polygons

Figure 6-9 shows the CT image after final iteration of the contour correction process for each slice.

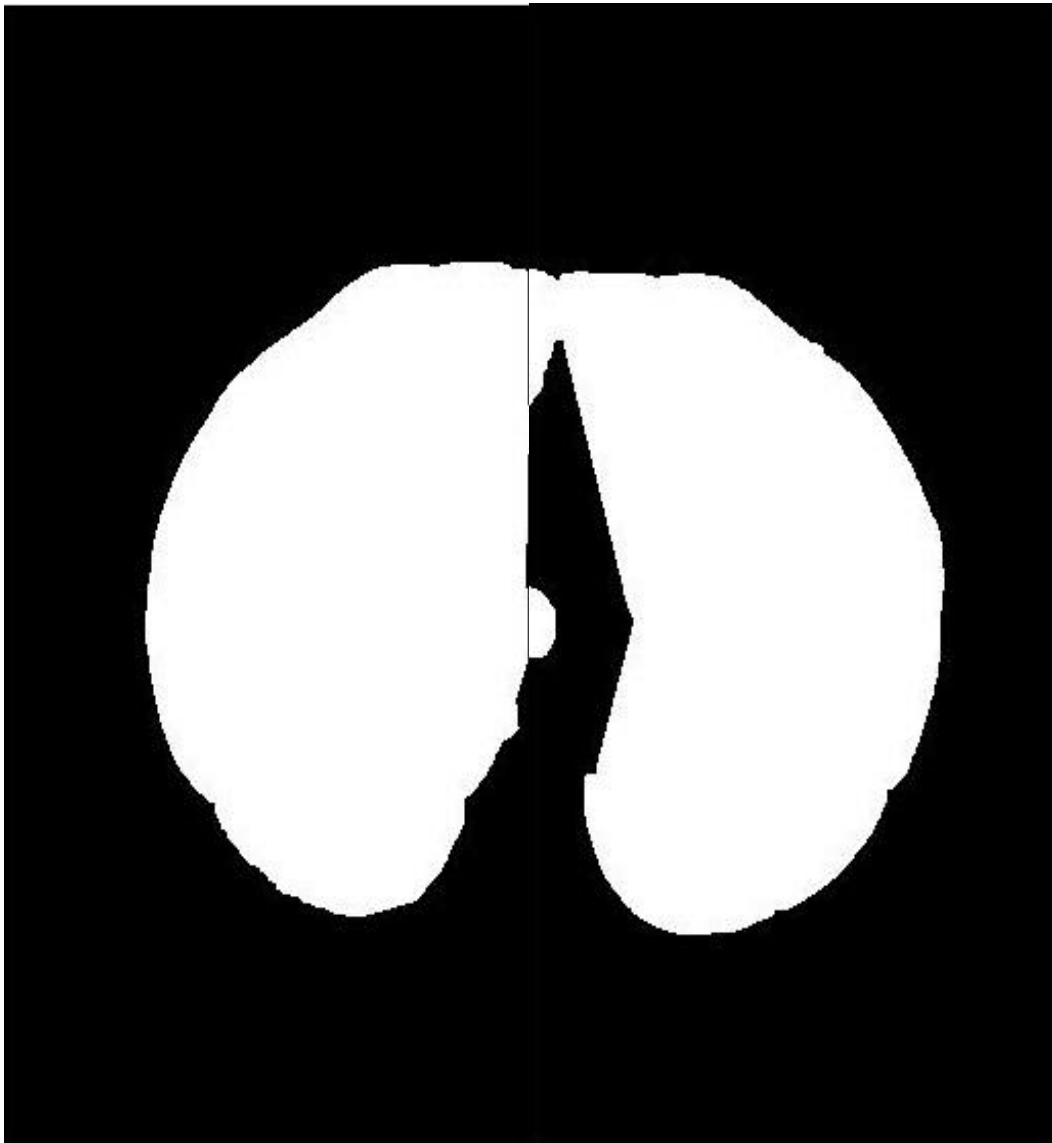
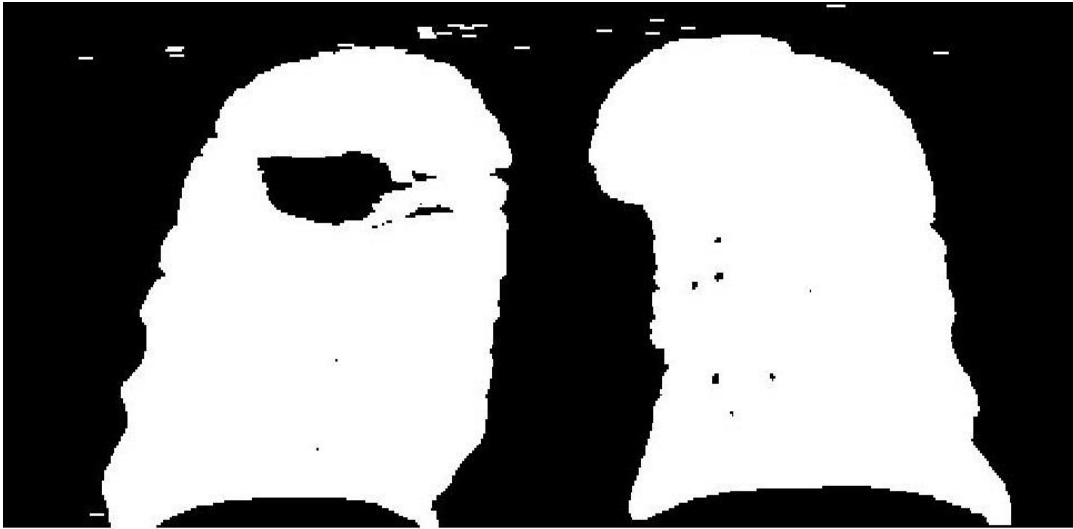


Figure 6-9: CT image slice after final iteration of the contour correction

After performing the contour correction process in the axial direction, the view of the 3D image is changed to the coronal direction and the contour correction process is performed again. Figure 6-10 shows an image before and after the contour correction in the coronal direction.



(a)



(b)

Figure 6-10 Contour correction in coronal direction a) before contour correction b) after contour correction

Through the process of contour correction, the final ROI is selected, where the possible lung nodules are searched. In the next step, all the nodule candidates are located and their features are extracted by a 3D connected component method. Figure 6-11 shows four lung

nodule candidates. 3D plot was created using triangular surface plot method [72]. Figure 6-11(a) shows a lung nodule in x,y and z axis. In the x direction the nodule position is between pixel 348 to 352 and in the y direction between 298 to 302 pixel in a 512 by 512 pixel image. For the z direction, the position is between slice number 61 to 65. Therefore the the nodule candidate is enclosed in a volume of 5x5x5 voxel cluster. Figure 6-11(b) illustrates another lung nodule candidate enclosed within 5x5x15. Figure 6-12 shows some additional lung nodule candidates with different sphericiity and sizes. Different colors don't have any significance.

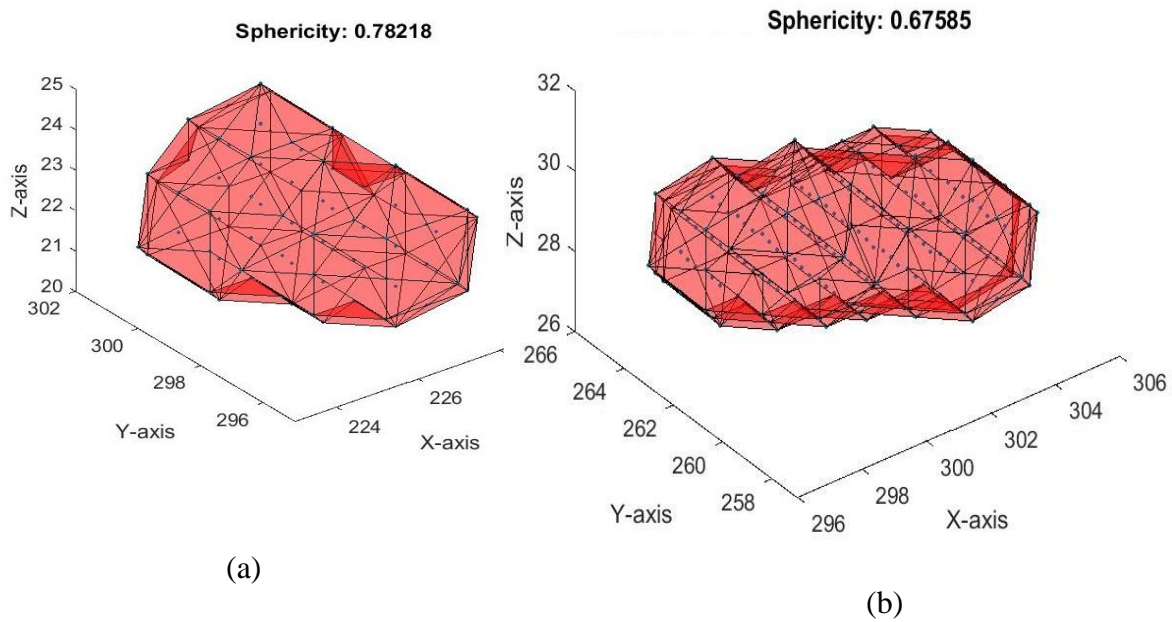


Figure 6-11 Two lung nodule candidates and their sphericity

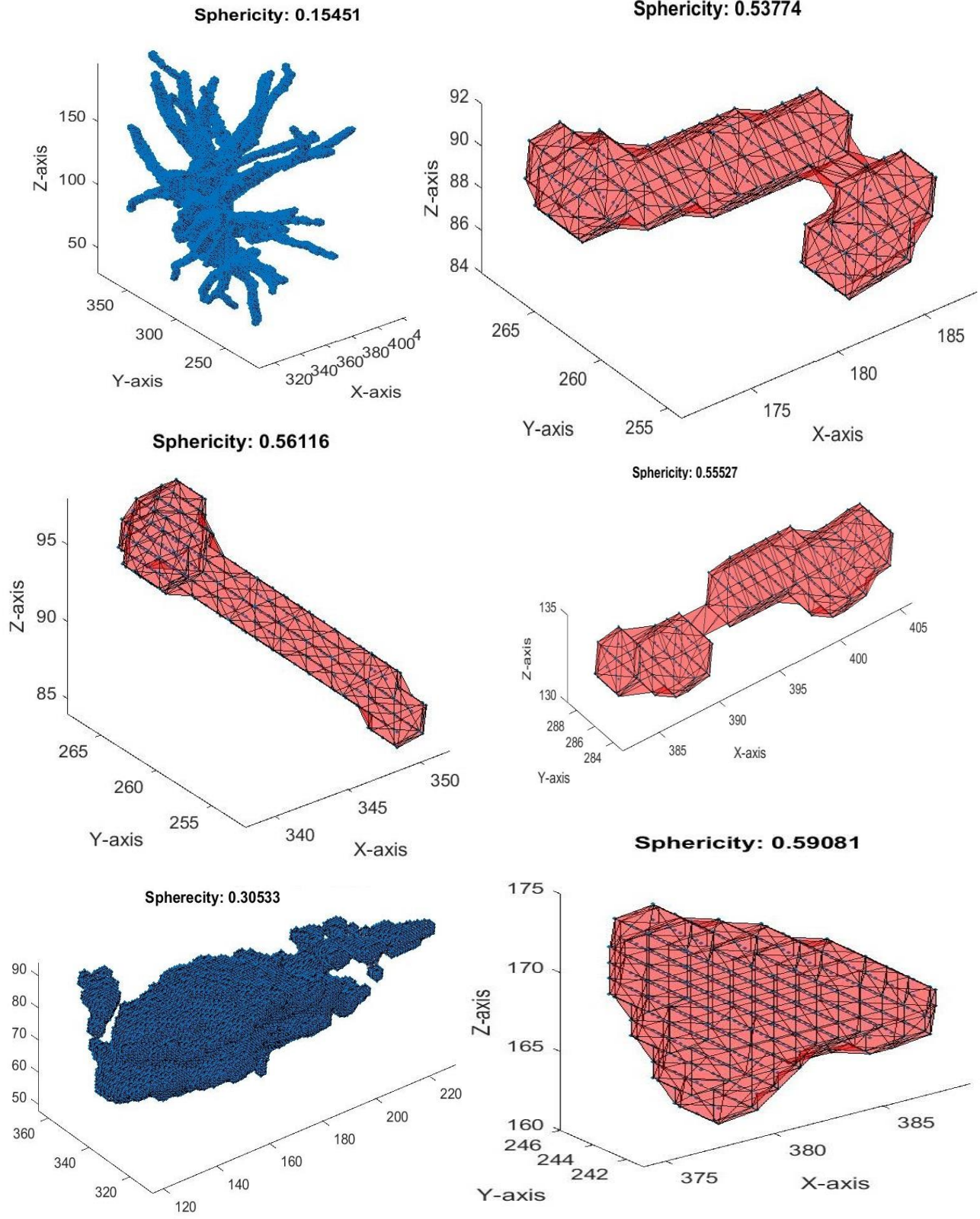


Figure 6-12 Different lung nodule candidates with their sphericity

From all the nodule candidates, nodules with very low sphericity (0.2) can be removed from the list as explained in chapter 5. Figure 6-13 shows two lung nodule candidates with very low sphericity.

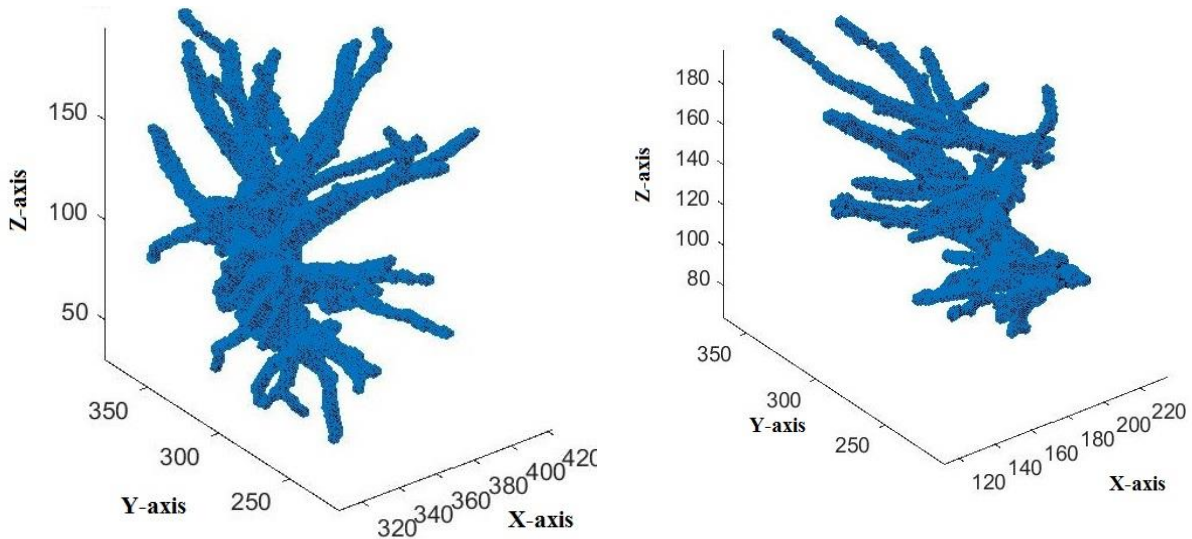


Figure 6-13 Two lung nodule candidates with sphericity less than 0.2

As an example, assume that the sphericity test leaves us with 15 nodules for the patient under consideration. Assume there are twelve atlases each with 15 nodules for a total of 300 atlas nodule candidates. After extracting all six features from each of the 15 target lung nodule candidates, the *K*-NN method is applied to find the closest atlas nodules for each target nodules. To apply the *K*-NN method, the Euclidian distance is calculated between each target and atlas nodule. Table 6-1 illustrates the Euclidian Distance-Weight table for a target nodule candidate, where the Euclidian distance, Nodule ID, Patient ID and the Weight are shown. Weights are calculated using equation (5.3). For instance, the first row of the table shows an Euclidian distance of 0.015766 between the first nodule of target patient and nodule number 19 of Atlas 11.

Notice the first row of the Table also has this smallest Euclidian distance. Therefore, according to equation (5.3), for this nodule ID and atlas ID, the weight is the same as the radius of target nodule. Further, for the remaining rows, the weight would be less by the ratio of the smallest Euclidian distance to the actual Euclidian distances

Table 6-1 Euclidian Distance-Weight table of target nodule 1

Euclidian distance	Atlas Nodule ID	Atlas ID	Weight
0.55	1	8	21.79
0.58	1	5	20.74
0.59	1	2	20.10
0.62	1	11	19.10
0.69	2	5	17.41
0.89	2	8	13.41
1.01	1	9	11.83
1.10	2	12	10.89

From Euclidian Distance-Weight table of each nodule candidate of the target image, a *Nodule Similarity Table* is created. Table 6-2 shows a sample *Nodule Similarity Table*.

Table 6-2 Nodule similarity Table of a sample target image

Target image Nodule ID (From biggest volume to smallest volume)	Closest nodule atlas ID	Second Closest nodule atlas ID	Third Closest nodule atlas ID	Weight of the closest nodule candidate	Weight for second closest nodule	Weight for third closest nodule
1	<u>8</u>	5	2	<u>21.79</u>	20.74	20.10
2	<u>8</u>	6	3	<u>18.51</u>	8.77	8.18
3	5	11	3	9.46	8.57	7.19
4	5	7	9	9.29	4.87	3.91

Target image Nodule ID (From biggest volume to smallest volume)	Closest nodule atlas ID	Second Closest nodule atlas ID	Third Closest nodule atlas ID	Weight of the closest nodule candidate	Weight for second closest nodule	Weight for third closest nodule
5	2	5	11	8.23	5.86	4.95
6	7	12	10	7.81	4.15	3.11
7	2	11	3	7.81	0.64	0.18
8	10	9	12	6.14	5.08	3.07
9	9	9	10	5.70	5.38	4.66
10	3	4	4	5.60	3.99	3.93
11	11	11	7	5.47	3.09	2.39
12	3	4	4	5.46	4.07	2.74
13	4	5	<u>8</u>	5.41	3.91	<u>1.41</u>
14	<u>8</u>	9	5	<u>5.25</u>	3.54	3.33
15	4	11	5	5.19	4.90	4.09

From the Nodule Similarity Table, weights of each atlas are added to find the *nodule similarity index*. For example, in Table 6-2 the weights corresponding to Atlas-8 are 21.79, 18.51, 5.25, and 3.91 (Underlined). If all these weights are added, the total weight for Atlas-8 is 46.97. This total weight is called *nodule similarity index*. Through this process, *nodule similarity index* is calculated for the atlases present in the *Nodule Similarity table*. For example, Table 6-3 shows all the *nodule similarity indices* for the all the atlases present in the Nodule Similarity table. The top row indicates the Atlas-ID s and the bottom row shows the *Nodule similarity index* of each ID.

Table 6-3 Nodule similarity index for the all the atlases from a Nodule Similarity table

Atlas-ID	5	8	2	11	3	4	9	7	10	6	12
<i>Nodule similarity index</i>	56.68	46.97	36.14	27.62	26.61	25.32	23.61	15.08	13.91	8.77	7.22

The atlas with the highest *nodule similarity index* is the ‘closest’ atlas for potentially all the nodules of the target image. Thus, Atlas-5, Atlas-8, and Atlas-2 are the closest and the second closest and the third closest atlases, respectively.

Once the closest, second closest and third closest atlases are selected, patch comparison process is applied to segment the lung nodules. For example, segmented 3D nodule by patch comparison between target image Nodule-1 and Atlas-8 nodule 1 is shown in Figure 6-14. From the figure, it can be seen that the nodule position is between 140 pixel to 200 pixel in x direction, and in between 300 to 360 pixel in y direction and between 40 slice to 90 pixel in z direction resulting in a nodule enclosed in 60x60x50 voxel cluster.

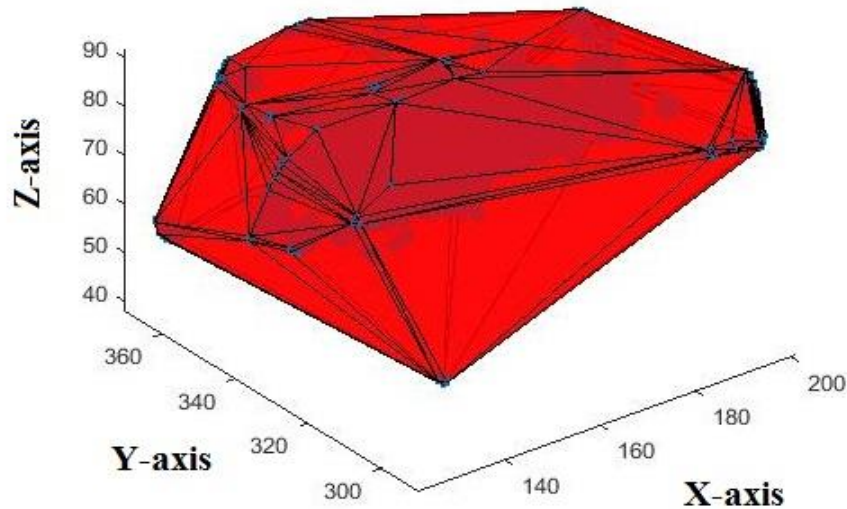


Figure 6-14 3D image of the segmented nodule by patch comparison

Figure 6-15 illustrates the 2D view of the final segmented nodule. Figure 6-15(a) shows the target CT image. Figure 6-15(b) shows the 2D slice of a segmented nodule resulting from the automated algorithm of this dissertation. Figure 6-15(c) shows a 2D slice of a manually segmented nodule (the correct answer). As can be seen, that the automated algorithm successfully determined the correct nodule location *and* the approximate shape.

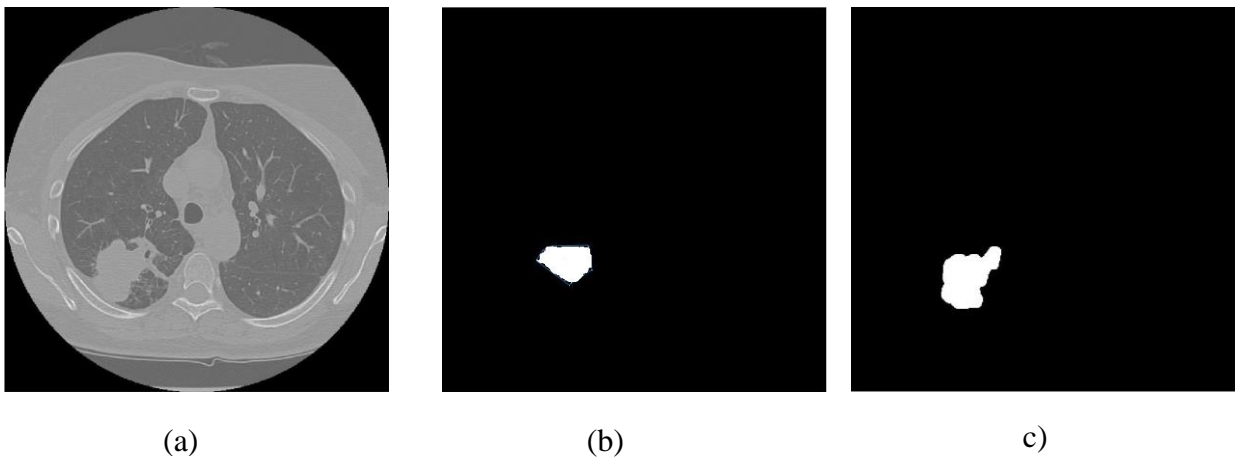


Figure 6-15 (a) Original image b) Final segmented nodule by the proposed algorithm (c) Manually segmented nodule from the same image.

The patch segmentation process was applied to two other nodules of the target image to test the validity of the process. Figure 6-16 shows the 2D output of the patch comparison performed between Nodule-2 and (Atlas-8 nodule 2) shown in Table 6-2. The segmented image turned out to be blank, which proves that there is no lung nodule.

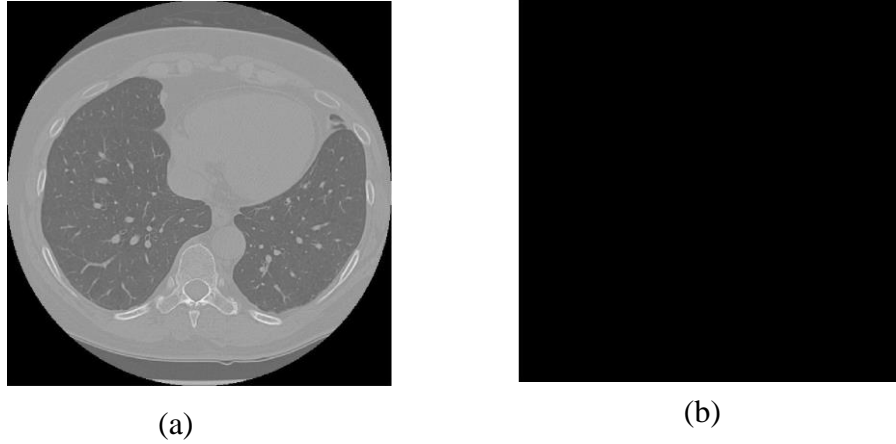


Figure 6-16 (a) Original image b) Final segmentation by the proposed algorithm

Figure 6-17 shows the 2D output of the patch comparison performed between Nodule-3 and (Atlas-5 nodule 7) shown in Table 6-2. The segmented image turned out to be blank, which proves that there is no lung nodule.

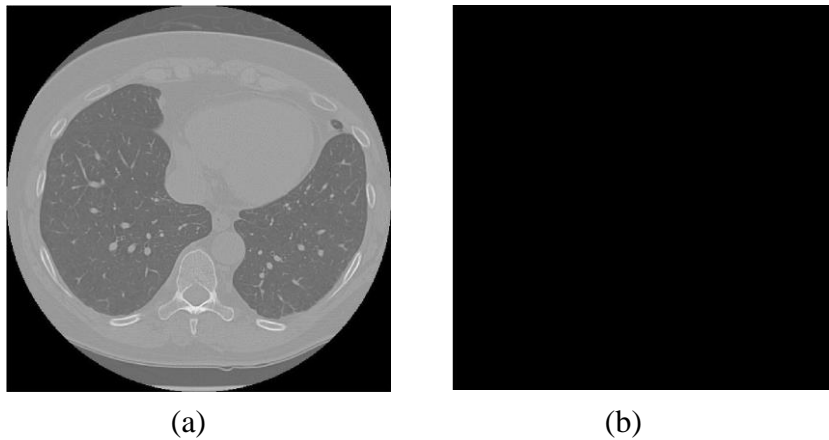


Figure 6-17 (a) Original image b) Final segmentation by the proposed algorithm

These experiments prove that the algorithm successfully and efficiently locates a suspected cancerous nodule where in fact it is present and does not produce false positives.

CHAPTER 7

CONCLUSIONS AND FUTURE WORK

7.1 Comparisons of the 2D and the 3D method:

In this dissertation, two different algorithms were developed to detect and segment lung nodules. In real patients, a lung nodule is a 3D element. Thus, in principle, only in a 3D implementation can all the different features be made available for detection and segmentation purposes. In the 2D image, a partial view and some of the features can be extracted for detection purpose. In this dissertation, at first a 2D method is implemented using those 2D features for computationally efficiency and the its development provided a prototype for a 3D lung nodule detection algorithm.

The 3D method is developed after successfully implementing the 2D method. Both the 2D and 3D algorithm have three main steps and they used size and shape features of the nodule candidates. Both methods use patch comparison for nodule segmentation once the atlases are selected. The 2D algorithm broadly consists of atlas election, image segmentation, and the post-processing step. The 3D algorithm on the other hand, has nodule candidate detection, atlas selection, and image segmentation step. Even though both algorithms have two common steps (atlas selection and image segmentation using patch comparison) the 3D algorithm, which was developed later is implemented differently to improve the efficiency. In the next paragraphs, a more detailed comparison of the 2D and 3D algorithms is presented.

Two preprocessing steps are common for both algorithms, they are: thresholding and morphological operation. In the 2D algorithm, the obvious threshold is selected as the midpoint

between the high intensity and the low intensity of a 2D image. In the 3D algorithm, iterative thresholding method is developed, which is clearly an improvement. In both 2D and 3D algorithms, thresholding methods are chosen based on the assumption that the image contains two classes of pixels and the intensity difference between them is significant, which is true for lung images. For the 2D image, the number of pixels is small compared to the 3D image. Thus, iterative method proved unnecessary and therefore saved computational time. In both algorithms, similarly shaped structuring element is applied for morphological erosion. For the 2D algorithm, a circular structuring element is used and a sphere-shaped structuring element is used for the 3D algorithm.

In the 3D algorithm, contour correction method is implemented for finding the ROI for possible nodule candidates. In the 2D algorithm, no ROI selection process is designed. The reason for this in the 2D algorithm, all the connected components are going through the possible nodule candidacy test and for 2D images the process is not computationally heavy.

In the 2D algorithm, a size and shape based feature vector is created for the atlas selection process, which shows the highest and lowest of different features of nodule candidates in a 2D image. The same method couldn't be applied for the 3D image, because in a 3D image, a) the number of possible nodule candidates is high, b) many of them have similar features and c) one feature vector cannot encompass all the different variations of nodule candidates. Thus, a nodule based *atlas selection process* is developed for the 3D algorithm.

In the 3D method, *nodule based patch comparison* is developed for image segmentation. In the 2D algorithm, the entire 2D image is used for patch comparison in the patch-segmentation process. For the 3D algorithm, only a cube around each nodule is selected for patch

comparison. It reduces the computational burden without losing accuracy and reduces the chance of false positives.

In the 3D method, no post-processing step is developed. The reason for this is the LoG blob detection is computationally expensive. It is also not necessary because the position of the nodule candidate is detected in the atlas selection step and segmentation is completed in image segmentation step, which can efficiently find the approximate size and shape of nodule candidate.

In the 3D algorithm, contour correction process is implemented to find the likely ROI. ROI selection is necessary for 3D process to remove the organs such as the heart where there is no lung nodule. Through the contour correction process, attached nodules in the wall are included in the ROI, which improves the output of the 3D algorithm.

7.2 Contribution:

In this dissertation, a novel approach is taken to detect and segment lung nodules. This work is the first attempt to use a patch based, multi-atlas method to this end. Two different algorithms are developed for 2D implementation and another for 3D implementation respectively. In the 2D implementation, a new size and shape-based feature vector is developed and, in 3D implementation, a nodule-based atlas search is developed and a new parameter *nodule similarity index* is defined. In the 3D implementation, a novel weight calculation is introduced, where both the size and similarity feature of an atlas nodule candidate is utilized. Both algorithms detected lung nodules very efficiently and false positive rate is found to be very small.

An additional contribution of this dissertation is that these algorithmic steps of the proposed methods can be separately applied. For example, for the 2D algorithm, the first step is the atlas selection step, where a small set of atlases is selected. Thus, if a radiologist wants to only reduce her workload, but rely on her own expertise to actually detect the nodule, she can choose a relatively small set of atlases, which are similar to the current patient as determined by the automated algorithm, but then manually examine only those atlases to make an informed decision.


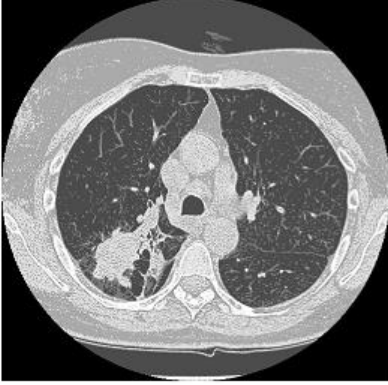
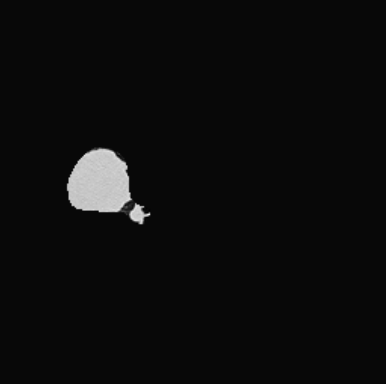
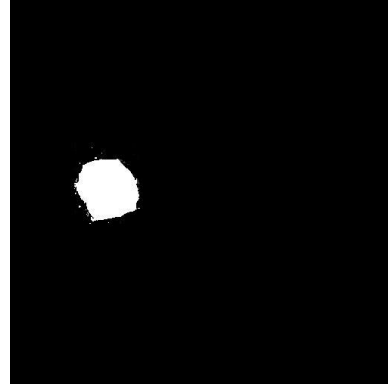
7.3 Future work:

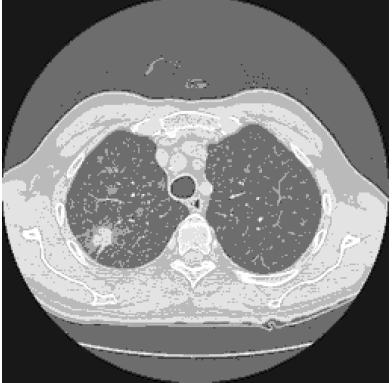

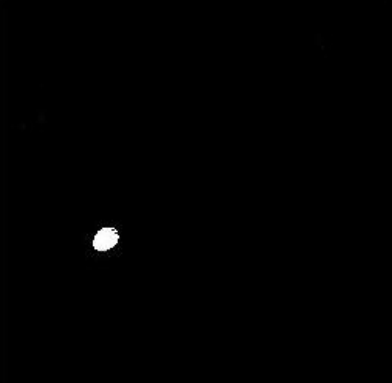
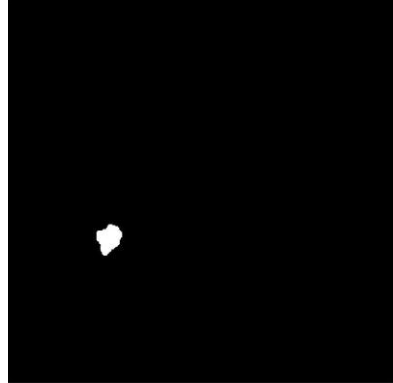

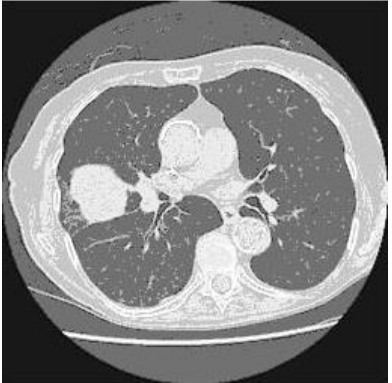
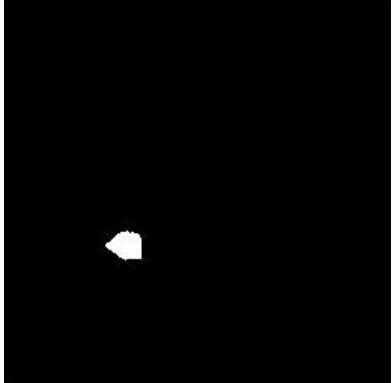
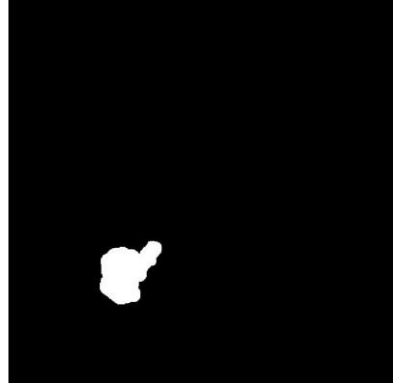
For future work the following attempts can be made:

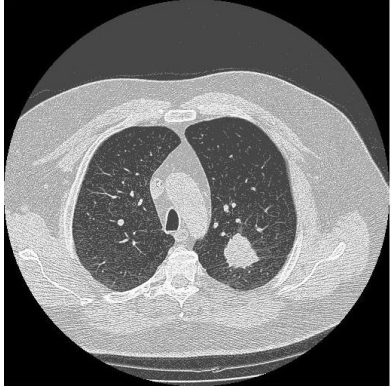

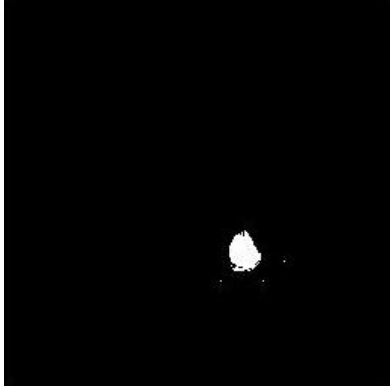
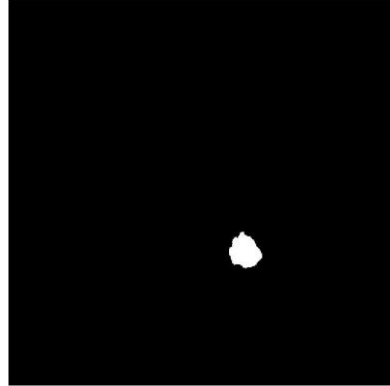

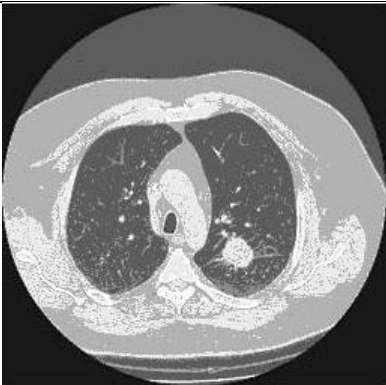
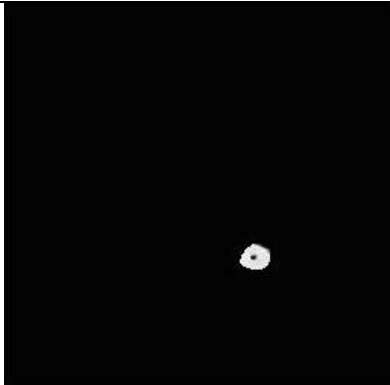
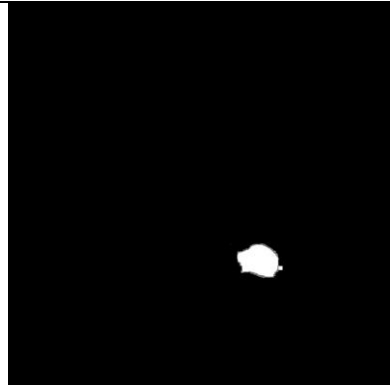
1. In these two algorithms, unsupervised learning (*K*-NN method) method is used to reduce the computation load. In future, supervised learning methods can be used to see if the efficiency can be increased.
2. In both algorithms, only size and shape based features of the lung nodules are used. In future, texture features can also be added.
3. Currently, the algorithm is in MATLAB code. The code can be optimized to run significantly faster to satisfy commercial demands.

APPENDIX A

TEST IMAGES AND OUTPUTS FOR THE FIVE CASES WHERE WE KNOW NODULES EXIST

PATIENT/TARGET IMAGE	SELECTED ATLAS	CORRECT FINAL OUTPUT	MANUAL SEGMENTATION
			

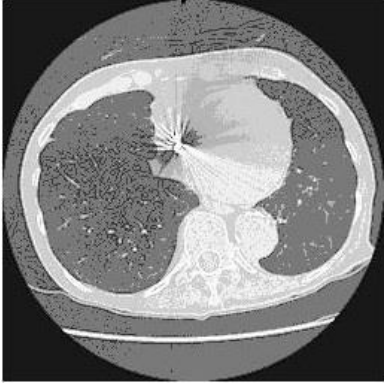
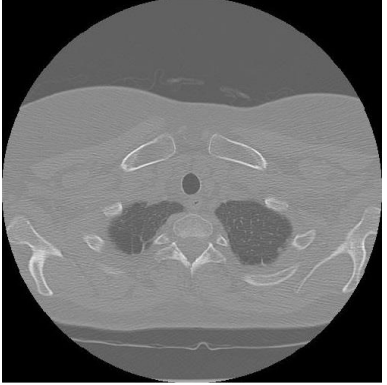
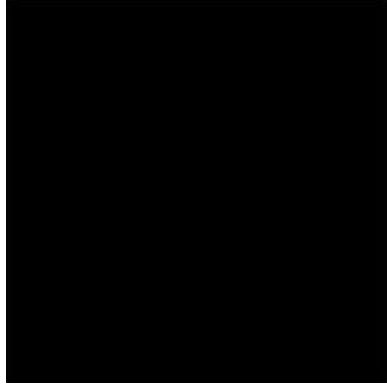
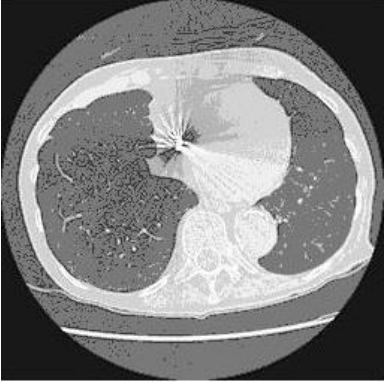
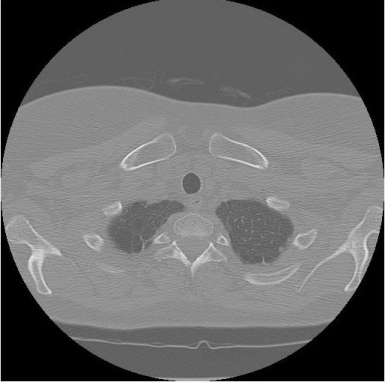
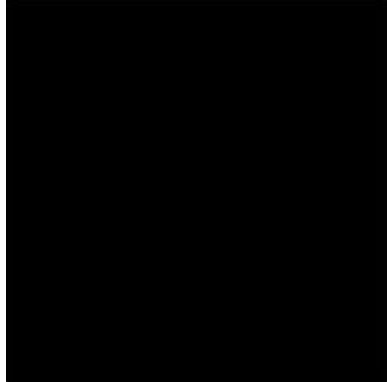
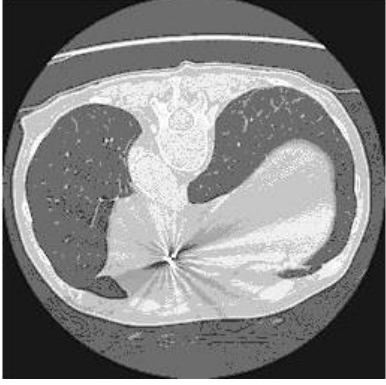
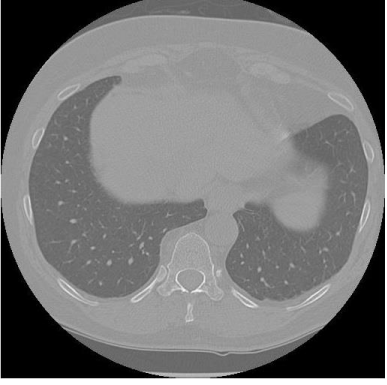
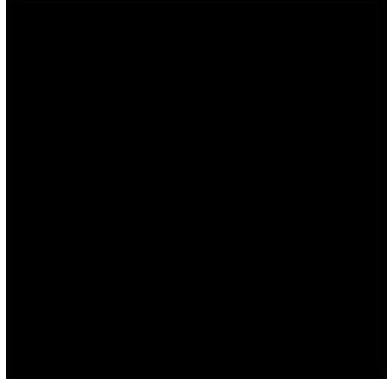
PATIENT/TARGET IMAGE	SELECTED ATLAS	CORRECT FINAL OUTPUT	MANUAL SEGMENTATION
			
			

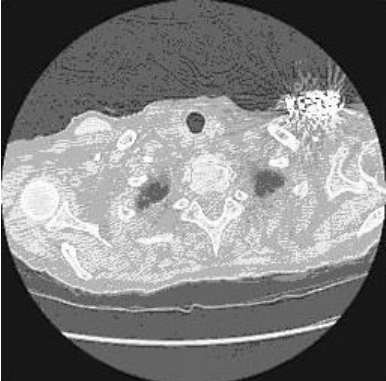
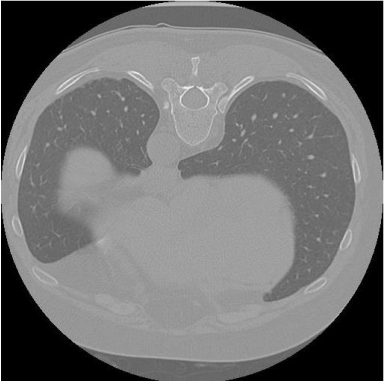
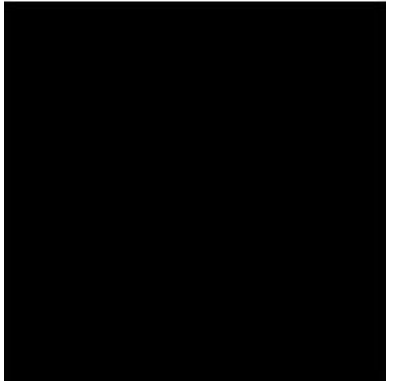
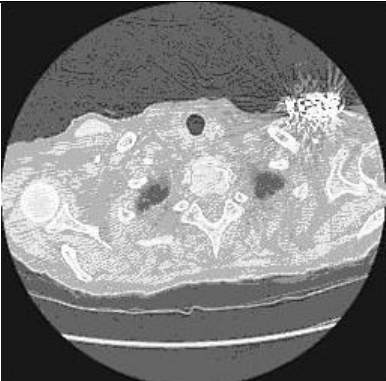

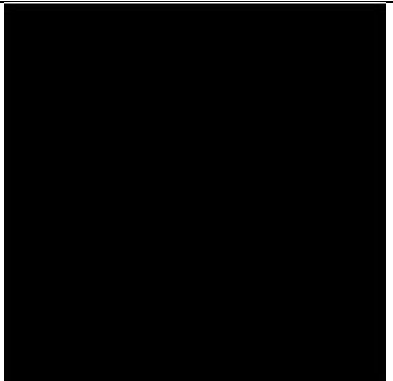

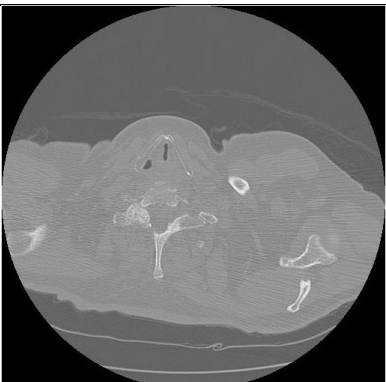
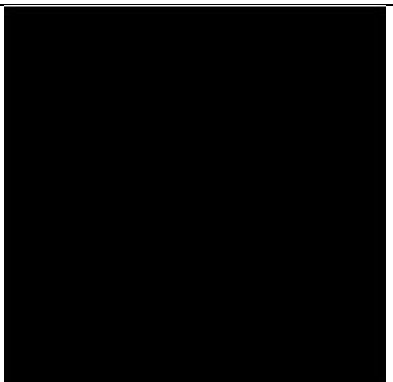

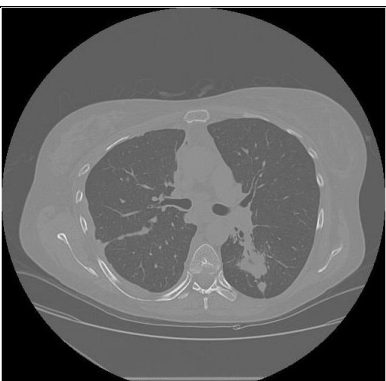
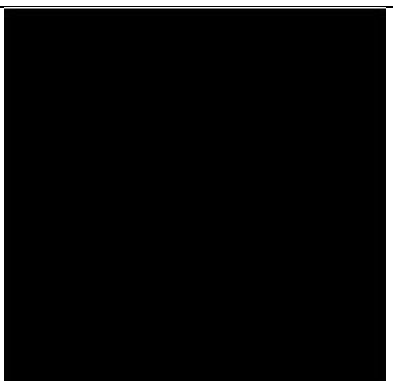
PATIENT/TARGET IMAGE	SELECTED ATLAS	CORRECT FINAL OUTPUT	MANUAL SEGMENTATION
			
			

APPENDIX B

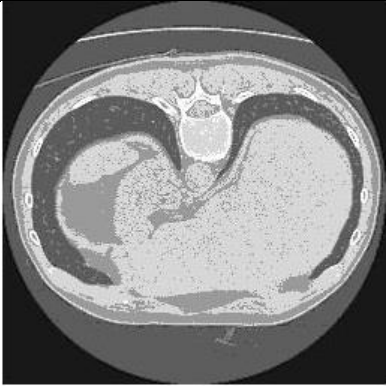
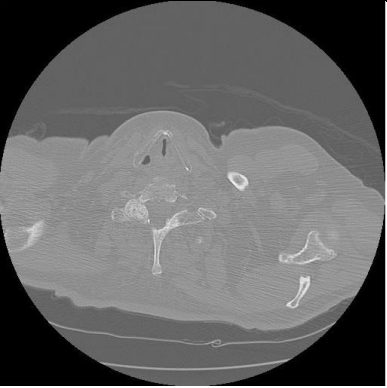
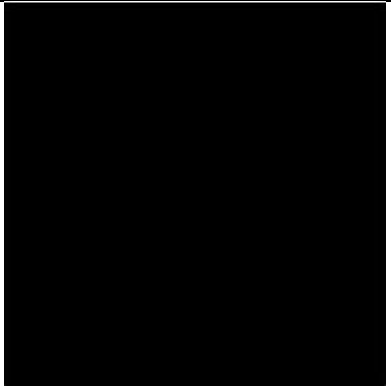
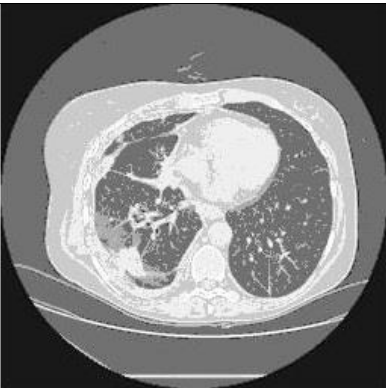
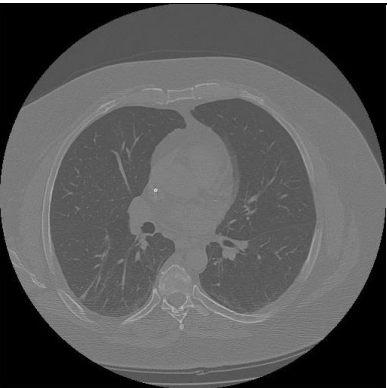
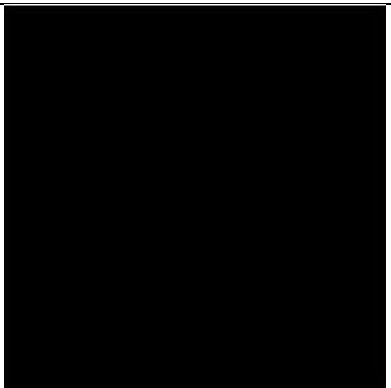

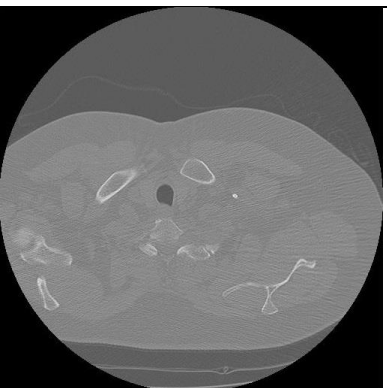
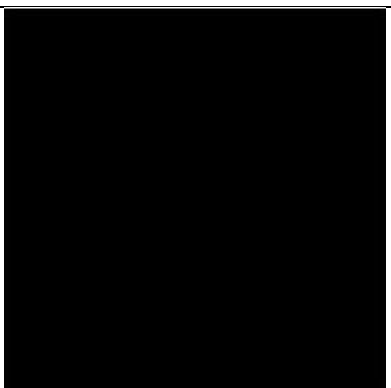
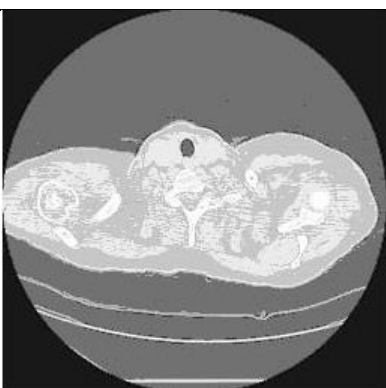
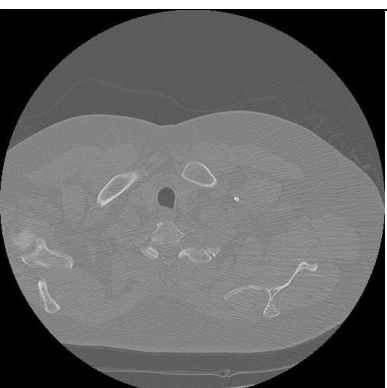
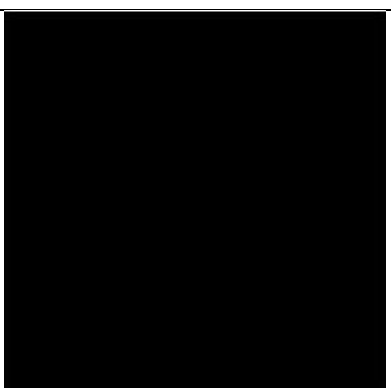
2D FALSE POSITIVE TEST

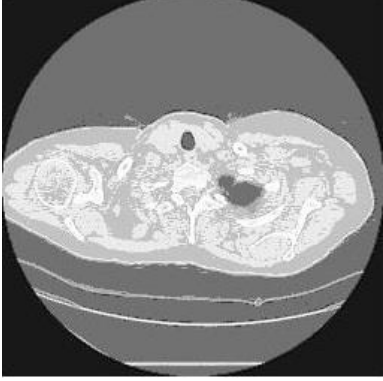
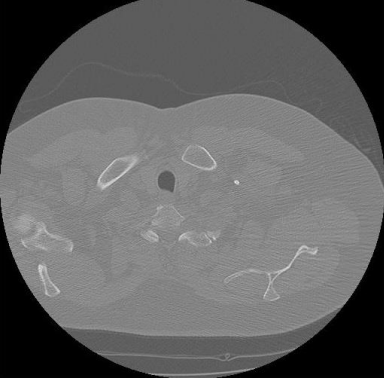
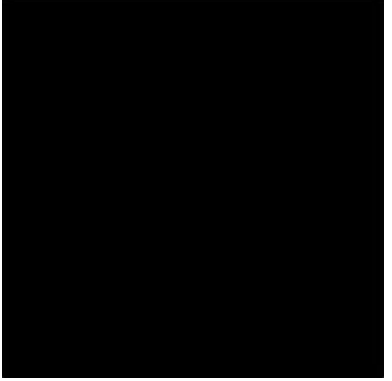
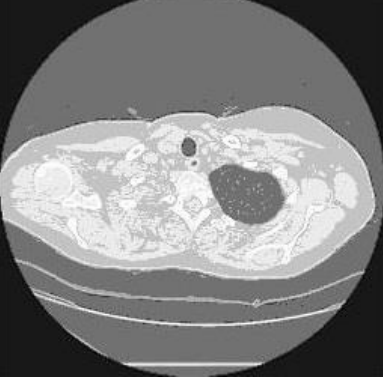
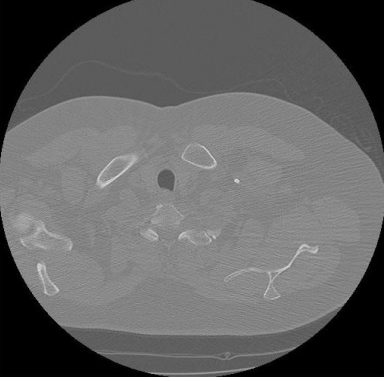
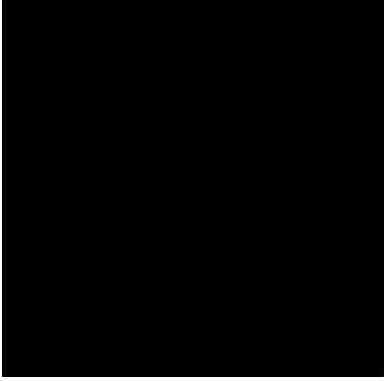
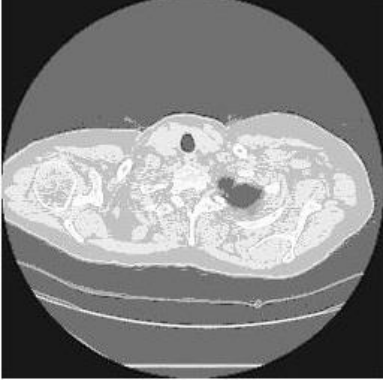

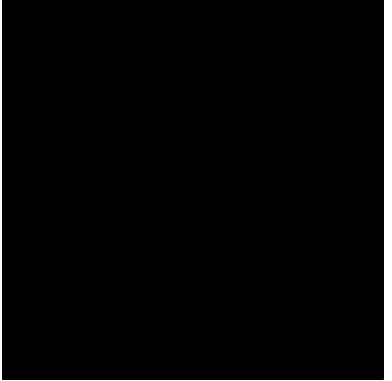


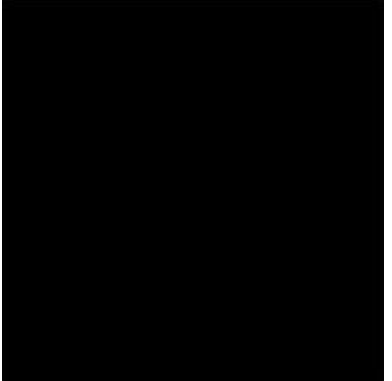
These 20 images are tested for false positives. Notice that no false positives were detected in the final outputs.

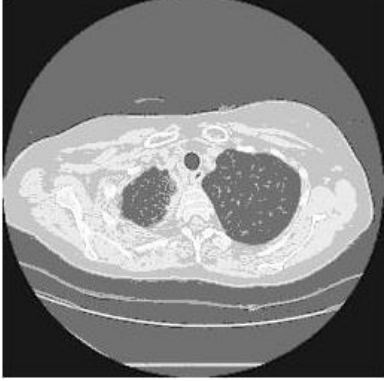

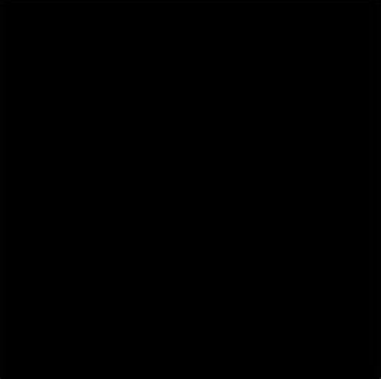


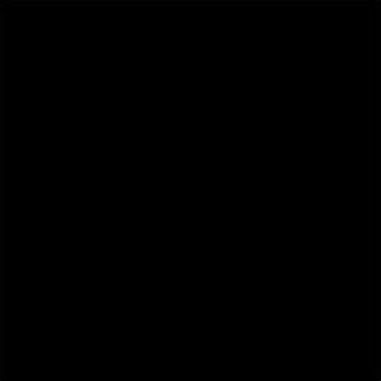


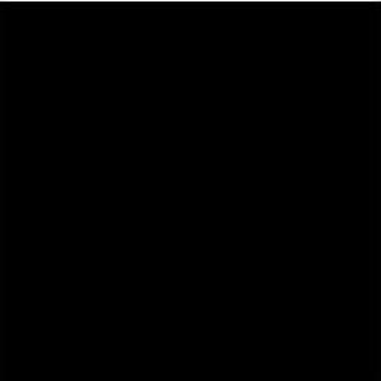
PATIENT/TARGET IMAGE	SELECTED ATLAS	FINAL OUTPUT
		
		
		

PATIENT/TARGET IMAGE	SELECTED ATLAS	FINAL OUTPUT
		
		
		
		

PATIENT/TARGET IMAGE	SELECTED ATLAS	FINAL OUTPUT

PATIENT/TARGET IMAGE	SELECTED ATLAS	FINAL OUTPUT
		
		
		
		

PATIENT/TARGET IMAGE	SELECTED ATLAS	FINAL OUTPUT
		
		
		
		

PATIENT/TARGET IMAGE	SELECTED ATLAS	FINAL OUTPUT
		
		
		

REFERENCES

- [1] A. C. Society, *Cancer facts and figures 2013*. American Cancer Society Atlanta, 2013.
- [2] “Cancer Facts & Figures 2017.” [Online]. Available: <https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2017.html>. [Accessed: 09-Apr-2017].
- [3] D. M. Hansell, A. A. Bankier, H. MacMahon, T. C. McLoud, N. L. Muller, and J. Remy, “Fleischner society: glossary of terms for thoracic imaging 1,” *Radiology*, vol. 246, no. 3, pp. 697–722, 2008.
- [4] I. R. S. Valente, P. C. Cortez, E. C. Neto, J. M. Soares, V. H. C. de Albuquerque, and J. M. R. Tavares, “Automatic 3D pulmonary nodule detection in CT images: A survey,” *Computer methods and programs in biomedicine*, 2015.
- [5] S. L. A. Lee, A. Z. Kouzani, and E. J. Hu, “Automated detection of lung nodules in computed tomography images: a review,” *Machine vision and applications*, vol. 23, no. 1, pp. 151–163, 2012.
- [6] S. G. Armato III *et al.*, “The Reference Image Database to Evaluate Response to Therapy in Lung Cancer (RIDER) Project: A Resource for the Development of Change Analysis Software,” *Clinical pharmacology and therapeutics*, vol. 84, no. 4, p. 448, 2008.
- [7] M. F. McNitt-Gray *et al.*, “The Lung Image Database Consortium (LIDC) data collection process for nodule detection and annotation,” *Academic radiology*, vol. 14, no. 12, pp. 1464–1474, 2007.
- [8] “Public Lung Image Database.” [Online]. Available: <http://www.via.cornell.edu/lungdb.html>. [Accessed: 21-Mar-2017].
- [9] D. Cascio, R. Magro, F. Fauci, M. Iacomì, and G. Raso, “Automatic detection of lung nodules in CT datasets based on stable 3D mass–spring models,” *Computers in Biology and Medicine*, vol. 42, no. 11, pp. 1098–1109, 2012.
- [10] S. Soltaninejad, M. Keshani, and F. Tajeripour, “Lung nodule detection by KNN classifier and active contour modelling and 3D visualization,” in *Artificial Intelligence and Signal Processing (AISP), 2012 16th CSI International Symposium on*, 2012, pp. 440–445.
- [11] H. Kim, T. Nakashima, Y. Itai, S. Maeda, J. kooi Tan, and S. Ishikawa, “Automatic detection of ground glass opacity from the thoracic MDCT images by using density features,” in *Control, Automation and Systems, 2007. ICCAS'07. International Conference on*, 2007, pp. 1274–1277.
- [12] J. Pu, J. Roos, A. Y. Chin, S. Napel, G. D. Rubin, and D. S. Paik, “Adaptive border marching algorithm: automatic lung segmentation on chest CT images,” *Computerized Medical Imaging and Graphics*, vol. 32, no. 6, pp. 452–462, 2008.
- [13] I. Gori *et al.*, “Lung nodule detection in screening computed tomography,” in *Nuclear Science Symposium Conference Record, 2006. IEEE*, 2006, vol. 6, pp. 3489–3491.
- [14] G.-Q. Wei, L. Fan, and J. Qian, “Automatic detection of nodules attached to vessels in lung CT by volume projection analysis,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2002, pp. 746–752.
- [15] A. Retico, P. Delogu, M. E. Fantacci, I. Gori, and A. P. Martinez, “Lung nodule detection in low-dose and thin-slice computed tomography,” *Computers in biology and medicine*, vol. 38, no. 4, pp. 525–534, 2008.

- [16] S. T. Namin, H. A. Moghaddam, R. Jafari, M. Esmaeil-Zadeh, and M. Gity, "Automated detection and classification of pulmonary nodules in 3D thoracic CT images," in *Systems Man and Cybernetics (SMC), 2010 IEEE International Conference on*, 2010, pp. 3774–3779.
- [17] Y. Liu, J. Yang, D. Zhao, and J. Liu, "A method of pulmonary nodule detection utilizing multiple support vector machines," in *2010 International Conference on Computer Application and System Modeling (ICCA SM 2010)*, 2010, vol. 10, pp. V10–118.
- [18] R. Garnavi, A. Baraani-Dastjerdi, H. A. Moghaddam, M. Giti, and A. A. Rad, "A new segmentation method for lung HRCT images," in *Digital Image Computing: Techniques and Applications, 2005. DICTA'05. Proceedings 2005*, 2005, pp. 52–52.
- [19] S. Diciotti, G. Picozzi, M. Falchini, M. Mascacchi, N. Villari, and G. Valli, "3-D segmentation algorithm of small lung nodules in spiral CT images," *IEEE transactions on Information Technology in Biomedicine*, vol. 12, no. 1, pp. 7–19, 2008.
- [20] I. C. Sluimer, P. F. van Waes, M. A. Viergever, and B. van Ginneken, "Computer-aided diagnosis in high resolution CT of the lungs," *Medical physics*, vol. 30, no. 12, pp. 3081–3090, 2003.
- [21] K. T. Bae, J.-S. Kim, Y.-H. Na, K. G. Kim, and J.-H. Kim, "Pulmonary Nodules: Automated Detection on CT Images with Morphologic Matching Algorithm—Preliminary Results 1," *Radiology*, vol. 236, no. 1, pp. 286–293, 2005.
- [22] R. A. Ochs *et al.*, "Automated classification of lung bronchovascular anatomy in CT using AdaBoost," *Medical image analysis*, vol. 11, no. 3, pp. 315–324, 2007.
- [23] W.-J. Choi and T.-S. Choi, "Automated pulmonary nodule detection based on three-dimensional shape-based feature descriptor," *Computer methods and programs in biomedicine*, vol. 113, no. 1, pp. 37–54, 2014.
- [24] M. Keshani, Z. Azimifar, F. Tajeripour, and R. Boostani, "Lung nodule segmentation and recognition using SVM classifier and active contour modeling: A complete intelligent system," *Computers in biology and medicine*, vol. 43, no. 4, pp. 287–300, 2013.
- [25] D.-Y. Kim, J.-H. Kim, S.-M. Noh, and J.-W. Park, "Pulmonary nodule detection using chest CT images," *Acta Radiologica*, vol. 44, no. 3, pp. 252–257, 2003.
- [26] R. Bellotti *et al.*, "A CAD system for nodule detection in low-dose lung CTs based on region growing and a new active contour model," *Medical Physics*, vol. 34, no. 12, pp. 4901–4910, 2007.
- [27] A. El-Baz, A. Elnakib, M. Abou El-Ghar, G. Gimel'farb, R. Falk, and A. Farag, "Automatic detection of 2D and 3D lung nodules in chest spiral CT scans," *International journal of biomedical imaging*, vol. 2013, 2013.
- [28] W. Suiyuan and W. Junfeng, "Pulmonary nodules 3D detection on serial CT scans," in *Intelligent Systems (GCIS), 2012 Third Global Congress on*, 2012, pp. 257–260.
- [29] N. Xu, N. Ahuja, and R. Bansal, "Automated lung nodule segmentation using dynamic programming and EM-based classification," in *Medical Imaging 2002*, 2002, pp. 666–676.
- [30] M. Aoyama, Q. Li, S. Katsuragawa, F. Li, S. Sone, and others, "Computerized scheme for determination of the likelihood measure of malignancy for pulmonary nodules on low-dose CT images," *Medical Physics*, vol. 30, no. 3, pp. 387–394, 2003.
- [31] J. Wang, R. Engelmann, and Q. Li, "Segmentation of pulmonary nodules in three-dimensional CT images by use of a spiral-scanning technique," *Medical Physics*, vol. 34, no. 12, pp. 4678–4689, 2007.

- [32] L. Fan *et al.*, “Automatic segmentation of pulmonary nodules by using dynamic 3D cross-correlation for interactive CAD systems,” in *Medical Imaging 2002*, 2002, pp. 1362–1369.
- [33] W. J. Kostis, A. P. Reeves, D. F. Yankelevitz, and C. I. Henschke, “Three-dimensional segmentation and growth-rate estimation of small pulmonary nodules in helical CT images,” *IEEE transactions on medical imaging*, vol. 22, no. 10, pp. 1259–1274, 2003.
- [34] A. A. Enquobahrie, A. P. Reeves, D. F. Yankelevitz, and C. I. Henschke, “Automated detection of pulmonary nodules from whole lung helical CT scans: performance comparison for isolated and attached nodules,” in *Medical Imaging 2004*, 2004, pp. 791–800.
- [35] Q. Li, F. Li, and K. Doi, “Computerized detection of lung nodules in thin-section CT images by use of selective enhancement filters and an automated rule-based classifier,” *Academic radiology*, vol. 15, no. 2, pp. 165–175, 2008.
- [36] Y. Kawata *et al.*, “Hybrid classification approach of malignant and benign pulmonary nodules based on topological and histogram features,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2000, pp. 297–306.
- [37] S. Matsumoto, H. L. Kundel, J. C. Gee, W. B. Geftter, and H. Hatabu, “Pulmonary nodule detection in CT images with quantized convergence index filter,” *Medical Image Analysis*, vol. 10, no. 3, pp. 343–352, 2006.
- [38] T. Jia, D.-Z. Zhao, J.-Z. Yang, and X. Wang, “Automated detection of pulmonary nodules in HRCT images,” in *Bioinformatics and Biomedical Engineering, 2007. ICBBE 2007. The 1st International Conference on*, 2007, pp. 833–836.
- [39] G. Fukano *et al.*, “Recognition method of lung nodules using blood vessel extraction techniques and 3D object models,” in *Medical Imaging 2003*, 2003, pp. 190–198.
- [40] L. Zhao, L. Boroczky, and K. Lee, “False positive reduction for lung nodule CAD using support vector machines and genetic algorithms,” in *International Congress Series*, 2005, vol. 1281, pp. 1109–1114.
- [41] J. Dehmeshki, J. Chen, M. V. Casique, and M. Karakoy, “Classification of lung data by sampling and support vector machine,” in *Engineering in Medicine and Biology Society, 2004. IEMBS’04. 26th Annual International Conference of the IEEE*, 2004, vol. 2, pp. 3194–3197.
- [42] A. M. Santos, A. O. de Carvalho Filho, A. C. Silva, A. C. de Paiva, R. A. Nunes, and M. Gattass, “Automatic detection of small lung nodules in 3D CT data using Gaussian mixture models, Tsallis entropy and SVM,” *Engineering Applications of Artificial Intelligence*, vol. 36, pp. 27–39, 2014.
- [43] S. Matsumoto, Y. Ohno, H. Yamagata, D. Takenaka, and K. Sugimura, “Computer-aided detection of lung nodules on multidetector row computed tomography using three-dimensional analysis of nodule candidates and their surroundings,” *Radiation medicine*, vol. 26, no. 9, pp. 562–569, 2008.
- [44] K. Clark *et al.*, “The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository,” *Journal of digital imaging*, vol. 26, no. 6, pp. 1045–1057, 2013.
- [45] T. Rohlfing, R. Brandt, R. Menzel, and C. R. Maurer, “Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains,” *NeuroImage*, vol. 21, no. 4, pp. 1428–1442, 2004.
- [46] A. Klein, B. Mensh, S. Ghosh, J. Tourville, and J. Hirsch, “Mindboggle: automated brain labeling with multiple atlases,” *BMC medical imaging*, vol. 5, no. 1, p. 1, 2005.

- [47] R. A. Heckemann, J. V. Hajnal, P. Aljabar, D. Rueckert, and A. Hammers, “Automatic anatomical brain MRI segmentation combining label propagation and decision fusion,” *NeuroImage*, vol. 33, no. 1, pp. 115–126, 2006.
- [48] J. E. Iglesias and M. R. Sabuncu, “Multi-atlas segmentation of biomedical images: a survey,” *Medical image analysis*, vol. 24, no. 1, pp. 205–219, 2015.
- [49] E. Heiberg, J. Sjögren, M. Ugander, M. Carlsson, H. Engblom, and H. akan Arheden, “Design and validation of Segment-freely available software for cardiovascular image analysis,” *BMC medical imaging*, vol. 10, no. 1, p. 1, 2010.
- [50] L. Ramus and G. Malandain, “Assessing selection methods in the context of multi-atlas based segmentation,” in *Biomedical Imaging: From Nano to Macro, 2010 IEEE International Symposium on*, 2010, pp. 1321–1324.
- [51] S. Candemir *et al.*, “Lung segmentation in chest radiographs using anatomical atlases with nonrigid registration,” *Medical Imaging, IEEE Transactions on*, vol. 33, no. 2, pp. 577–590, 2014.
- [52] J. Lee, I. Lyu, and M. Styner, “Multi-atlas segmentation with particle-based group-wise image registration,” in *SPIE Medical Imaging*, 2014, pp. 903447–903447.
- [53] P. Coupé, J. V. Manjón, V. Fonov, J. Pruessner, M. Robles, and D. L. Collins, “Patch-based segmentation using expert priors: Application to hippocampus and ventricle segmentation,” *NeuroImage*, vol. 54, no. 2, pp. 940–954, 2011.
- [54] Z. Wang, C. Donoghue, and D. Rueckert, “Patch-Based Segmentation without Registration: Application to Knee MRI,” in *Proceedings of the 4th International Workshop on Machine Learning in Medical Imaging-Volume 8184*, 2013, pp. 98–105.
- [55] Z. Wang, R. Wolz, T. Tong, and D. Rueckert, “Spatially aware patch-based segmentation (saps): an alternative patch-based segmentation framework,” in *Medical Computer Vision. Recognition Techniques and Applications in Medical Imaging*, Springer, 2012, pp. 93–103.
- [56] V. Fonov, P. Coupé, S. Eskildsen, J. Manjon, and L. Collins, “Multi-atlas labeling with population-specific template and non-local patch-based label fusion,” in *MICCAI 2012 Workshop on multi-atlas labeling*, 2012, pp. 63–66.
- [57] P. Aljabar, R. Heckemann, A. Hammers, J. V. Hajnal, and D. Rueckert, “Classifier selection strategies for label fusion using large atlas databases,” in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2007*, Springer, 2007, pp. 523–531.
- [58] E. M. van Rikxoort *et al.*, “Adaptive local multi-atlas segmentation: Application to the heart and the caudate nucleus,” *Medical image analysis*, vol. 14, no. 1, pp. 39–49, 2010.
- [59] X. Artaechevarria, A. Muñoz-Barrutia, and C. Ortiz-de-Solorzano, “Efficient classifier generation and weighted voting for atlas-based segmentation: Two small steps faster and closer to the combination oracle,” in *Medical Imaging*, 2008, p. 69141W–69141W.
- [60] K. D. Fritscher, M. Peroni, P. Zaffino, M. F. Spadea, R. Schubert, and G. Sharp, “Automatic segmentation of head and neck CT images for radiotherapy treatment planning using multiple atlases, statistical appearance models, and geodesic active contours,” *Medical physics*, vol. 41, no. 5, p. 051910, 2014.
- [61] V. Kulasingam and E. P. Diamandis, “Strategies for discovering novel cancer biomarkers through utilization of emerging technologies,” *Nature clinical practice Oncology*, vol. 5, no. 10, pp. 588–599, 2008.

- [62] S. Saien, A. H. Pilevar, and H. A. Moghaddam, “Refinement of lung nodule candidates based on local geometric shape analysis and Laplacian of Gaussian kernels,” *Computers in biology and medicine*, vol. 54, pp. 188–198, 2014.
- [63] M. S. Brown *et al.*, “Method for segmenting chest CT image data using an anatomical model: preliminary results,” *IEEE transactions on medical imaging*, vol. 16, no. 6, pp. 828–839, 1997.
- [64] D. T. Larose, *Discovering knowledge in data: an introduction to data mining*. John Wiley & Sons, 2014.
- [65] X. Ye, X. Lin, J. Dehmeshki, G. Slabaugh, and G. Beddoe, “Shape-based computer-aided detection of lung nodules in thoracic CT images,” *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 7, pp. 1810–1820, 2009.
- [66] A. K. Jain, *Fundamentals of Digital Image Processing*, 1 edition. Englewood Cliffs, NJ: Pearson, 1988.
- [67] V. Alexandrov, “Using critical points in contours for segmentation of touching characters,” in *presented at the Int. Conf. Comput. Syst. Technol.*, 2004.
- [68] “Pixel Connectivity :: Morphological Operations (Image Processing Toolbox).” [Online]. Available: <https://www.mathworks.com/help/releases/R13sp2/toolbox/images/morph12.html>. [Accessed: 29-Mar-2017].
- [69] Y. Lee, T. Hara, H. Fujita, S. Itoh, and T. Ishigaki, “Automated detection of pulmonary nodules in helical CT images based on an improved template-matching technique,” *IEEE Transactions on medical imaging*, vol. 20, no. 7, pp. 595–604, 2001.
- [70] M. Alam, G. Sankaranarayanan, and V. Devarajan, “Lung Nodule Detection and Segmentation Using a Patch-Based Multi-Atlas Method,” in *Computational Science and Computational Intelligence (CSCI), 2016 International Conference on*, 2016, pp. 23–28.
- [71] “3Dim Laboratory s.r.o.” [Online]. Available: <http://www.3dim-laboratory.cz/en/software/3dimviewer>. [Accessed: 13-Apr-2017].
- [72] J. He and C. Window, “MATLAB Quick Reference.”