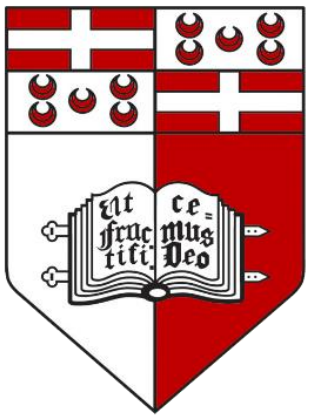
# Image Processing Techniques for Lung Lesion Detection in Chest CT Scans

Daniel Bonanno

Primary Supervisor: Prof. Ing. Carl J. Debono

Co-Supervisors: Dr. Paul Bezzina, Dr. Francis Zarb




**Department of Communications and Computer Engineering**
**Faculty of ICT**

**University of Malta**

**May 2016**

*Submitted in partial fulfilment of the requirements for the degree of B.Sc. (Hons.) in Computer Engineering*

# Copyright Notice

# Plagiarism Declaration Form

Plagiarism is defined as “the unacknowledged use, as one's own, of work of another person, whether or not such work has been published, and as may be further elaborated in Faculty or University guidelines" (University Assessment Regulations, 2009, Regulation 39 (b)(i), University of Malta).

I, the undersigned, declare that the Final Year Project report submitted is my work, except where acknowledged and referenced.

I understand that the penalties for committing a breach of the regulations include loss of marks; cancellation of examination results; enforced suspension of studies; or expulsion from the degree programme.

Image Processing Techniques for Lung Lesion Detection in Chest CT Scans
Title of Work Submitted

ICT3903
Course Code

Daniel Bonanno

26/07/2016
Date

# Abstract

Imaging is an essential component in modern medicine, especially in areas that deal with nonvisible internal organs, such as the lung. Lung lesion detection is an important step in the diagnosis of several diseases, including lung cancer, which is considered to be one of the deadliest cancers.

With that in mind, Computer Aided Detection (CAD) tools are becoming more popular in medical diagnosis. Such tools aim to support practitioners in their jobs by reducing human error as much as possible. Image processing is the backbone of CAD systems. It is used by such systems to enhance the quality of the image, segment the region of interest, highlight features within the image, and support anomaly detection techniques. Computed Tomography (CT) is an X-Ray based medical imaging modality which generates hundreds of images per patient. Therefore, developing CAD systems for this particular type of imaging modality is crucial.

This study applies image processing techniques on CT Scans of the thorax, designs and implements an automatic CAD tool aimed at detecting lesions in lungs. This is done by pre-processing the CT images to remove noise through the use of a Bilateral Filter and enhance the contrast by using Gamma Correction and Contrast Limited Adaptive Histogram Equalisation. In both these pre-processing steps, the parameters required for the techniques implemented, were verified by an expert radiologist to ensure that no anatomical details are removed by these techniques. The lungs are segmented by means of a marker-based watershed algorithm. Detection is performed by applying a priori knowledge about lesion intensities and shapes. Lesion characteristics are used by the CAD system developed to accurately separate lesions from vessels. This helps in reducing false positives, that is, the detection of a lesion in the absence of one.

The CAD system designed was tested on 25 cases, achieving a sensitivity of 76.3%, an average false positive rate of 7.88 lesions per scan and an average false negative rate of 0.56 lesions per scan. This is done in an average time of 74.3ms per slice.

# Acknowledgements

I would like to express my gratitude to my supervisors: Prof. Ing. Carl J. Debono, Dr. Paul Bezzina and Dr. Francis Zarb for their continuous help with this project. I could not have wished for better guidance.

A big thank you goes to my parents Mark and Rose and my siblings: Marie Claire, Francesca, Matthew and Luke. They were always there to provide a helping hand in my time of need.

Lastly, I would like to thank my Aunt Carmen for all her support and encouragement, not just during this project, but throughout all my studies.

# Contents

**Copyright Notice** ..... **i**

**Plagiarism Declaration Form** ..... **ii**

**Abstract** ..... **iii**

**Acknowledgements** ..... **iv**

**List of Figures** ..... **vii**

**List of Tables** ..... **x**

**List of Abbreviations** ..... **xi**

**Chapter 1 : Introduction** ..... **1**

1.1 What is Computer Aided Detection? ..... 1

1.2 What is Computed Tomography? ..... 2

1.3 The Inner Workings of a CT System ..... 2

1.3.1 X-Rays and Absorption ..... 2

1.3.2 Obtaining an Image ..... 3

1.4 Focus of this Study ..... 6

1.5 Structure of this Report ..... 7

**Chapter 2 : Background and Literature Review** ..... **8**

2.1 Noise Removal ..... 8

2.2 Contrast Enhancement ..... 12

2.3 Segmentation ..... 15

2.4 Detection ..... 21

**Chapter 3 : Implementation** ..... **26**

3.1 General Overview of the Implemented CAD System ..... 26

3.2 Reading the DICOM Files ..... 27

3.3 Noise Removal ..... 29

3.3.1 Bilateral Filtering ..... 29

3.3.2 Wiener Filtering ..... 29

3.3.3 Anisotropic Diffusion ..... 31

3.3.4 Choosing the Noise Reduction Method and its Parameters ..... 32

3.4 Contrast Enhancement ..... 33

3.5 Segmentation ..... 34

3.5.1 Removing the Background ..... 35

3.5.2 Extracting the Lungs ..... 38

3.6 Detection ..... 43

3.6.1 Obtaining an Intensity Mask ..... 43
3.6.2 False Positive Reduction ..... 44
3.7 Output ..... 51
**Chapter 4 : Testing and Evaluation ..... 52**
4.1 Results ..... 52
4.2 Comparison with other CAD Systems ..... 55
4.3 Observations ..... 55
4.3.1 Detection of Smaller Lesions ..... 55
4.3.2 Irregularly Shaped Lesions ..... 58
4.3.3 Pleural Lesions ..... 58
4.3.4 Vessels Detected as False Positives ..... 60
**Chapter 5 : Conclusion and Future Work ..... 62**
5.1: Conclusion ..... 62
5.2 Future Work ..... 63
5.2.1 Blob Enhancement Filter ..... 63
5.2.2 Segmentation and Mapping of the Respiratory Tree ..... 63
5.2.3 Limiting the Number of Planes along the *Z* Axis ..... 63
5.2.4 Consideration of Other Features ..... 64
5.2.5 Output ..... 64
5.2.6 Parallelisation Techniques ..... 64
**References ..... 65**
**Appendix ..... 71**
A – A Whole CT Scan of 99 Images ..... 71
B – Lesion having an Irregular Shape, pointed to by Arrows ..... 88
C – Documents Pertaining to the Approval of use of CT Scans and Medical Reports from a General Hospital in Malta ..... 90
Letter of Permission Sent ..... 90
Permission from the Hospital's Authorities ..... 91
Data Protection Declaration ..... 92
Permission by the University of Malta's Research Ethics Committee ..... 93
D – Contents of the Attached CD ..... 94

# List of Figures

**Figure 1.1:** *Different views of the same object. [5]* ..... 4

**Figure 1.2:** *Reconstruction using (a) simple backprojection and (b) filtered backprojection [5].* ..... 5

**Figure 1.3:** *(a) Narrow beam and detector pair [5] and (b) Fan beam and detector pair, coloured blue [13]* ..... 5

**Figure 1.4:** *An example of a lesion, indicated by the arrow* ..... 6

**Figure 2.1:** *Resulting image after the reconstruction filter, with (b) being sharper but noisier than (a). [18]* ..... 8

**Figure 2.2:** *A visual example of the marker based watershed algorithm [36]* ..... 18

**Figure 2.3:** *A visual representation of the graph (a) and graph cut (b), where node S represents the foreground and node T the background [40]* ..... 20

**Figure 3.1:** *General Flow Diagram of the CAD System* ..... 26

**Figure 3.2:** *A typical DICOM image before (a) and after (b) windowing* ..... 28

**Figure 3.3:** *Resulting images after filtering; (a) original image, (b) applying a Bilateral filter having a neighbourhood radius of 5, colour standard deviation of 40 and co-ordinate space standard deviation of 70, (c) using a Weiner filter having constant (inverse PSNR) of 0.018, and (d) utilising Anisotropic Diffusion with lambda set to 0.25, k_ns and k_we set to 45, and 10 iterations.* ..... 30

**Figure 3.4:** *The frequency response of the image in (a) Figure 3.3(a) and (b) Figure 3.3(c)* ..... 31

**Figure 3.5:** *The image of Figure 3.3(b) after contrast enhancement* ..... 34

**Figure 3.6:** *Background removal showing (a) the contrast enhanced image, (b) the image after morphological opening using a disc radius of 10, (c) outside background removal and (d) final result.* ..... 36

**Figure 3.7:** *Lung image (a) after removal of the background around the chest cavity, and (b) after the second part of the background removal* ..... 37

**Figure 3.8:** *The resulting image after dilation (a) and the binary image after thresholding (b).* ..... 38

**Figure 3.9:** *Mask containing lesion picked up as background (a) and the corrected binary image (b).* ..... 39

**Figure 3.10:** *Image after dilation and erosion (a) and the thresholded binary mask including pixels that may be foreground (b).* ..... 40
**Figure 3.11:** *The final segmentation process; (a) the final markers used by the watershed algorithm, (b) the mask obtained from the watershed algorithm, (c) smoothed mask following morphological closing, (d) the binary mask, and (e) the extracted lungs.* ..... 41
**Figure 3.12:** *Further segmentation scenarios: (a) Two correctly segmented, connected lungs; (b) background region between the lungs incorrectly included in the output; (c) white region wrongly included by the hole filling algorithm.* 42
**Figure 3.13:** *The segmented image (a) and the intensity mask (b).* ..... 44
**Figure 3.14:** *The transformation performed on the axes.* ..... 45
**Figure 3.15:** *A thresholded shape index mask. The arrow points towards a lesion. Note how a lot of false positives still exist and the lesion region is disjoint, making further false positive reduction difficult* ..... 45
**Figure 3.16:** *Demonstration on how the radius of the lesion varies with slice number.* ..... 47
**Figure 3.17:** *The result of applying false positive reduction along the xy-plane (b) of the intensity mask (a).* ..... 48
**Figure 3.18:** *An example of a vessel, denoted by an arrow, that has an elliptical shape but is not removed by false positive reduction in the xy plane, since the vessel is parallel to the y axis.* ..... 48
**Figure 3.19:** *The final lesion mask obtained by reducing false positives along the z axis.* ..... 50
**Figure 3.20:** *Sample output for the image in Figure 3.13(a) with the detected lesion marked* ..... 51
**Figure 4.1:** *Typical Output of the CAD system.* ..... 54
**Figure 4.2:** *Misclassification of a lesion as a False Negative.* ..... 57
**Figure 4.3:** *An example of an irregularly shaped lesion, with an enlarged bounding box due to the spur. This spur will cause the lesion to be considered a false positive.* ..... 58
**Figure 4.4:** *An example of a pleural lesion (a), (c) and the segmentation result (b), (d).* ..... 59
**Figure 4.5:** *False Positive Detection of Ellipsoidal Vessels; (a) and (b) show the vessel marked as a lesion, while (c) confirms that it is a vessel.* ..... 60

**Figure 4.6:** *False Positive Detection of a vessel which shows minimum movement between slices; (a), (b) and (c) show the regions which were erroneously detected while in (d) it is no longer detected as a lesion.*..... 61

# List of Tables

**Table 3.1:** *The parameter values considered for Bilateral Filtering* ..... 32
**Table 3.2:** *The parameter values considered for Wiener Filtering* ..... 32
**Table 3.3:** *The parameter values considered for Anisotropic Diffusion* ..... 32
**Table 3.4:** *The parameter values considered for Contrast Enhancement* ..... 33
**Table 4.1:** *Confusion Matrix showing the Results* ..... 52
**Table 4.2:** *Comparison between this work and other CAD systems in literature* ..... 55

# List of Abbreviations

| | |
|---|---|
| CT | Computed Tomography |
| AHE | Adaptive Histogram Equalisation |
| BMHE | Bin Modified Histogram Equalisation |
| CAD | Computer Aided Detection |
| CLAHE | Contrast Limited Adaptive Histogram Equalisation |
| HE | Histogram Equalisation |
| MBPHE | Mean Brightness Preserving Histogram Equalisation |
| MRI | Magnetic Resonance Imaging |
| MSE | Mean Square Error |
| NGCCLAHE | Normalised Gamma-Corrected Contrast Limited Adaptive Histogram Equalisation |
| POSHE | Partially Overlapped Sub-Block Histogram Equalisation |
| PSNR | Peak Signal-to-Noise Ratio |
| SNR | Signal-to-Noise Ratio |
| WHE | Weighted Histogram Equalisation |

# Chapter 1 : Introduction

Lung lesion detection is a crucial part in the diagnosis of a wide range of health problems such as infections, tuberculosis or even cancer. Lung cancer, which was one of the five most common cancers in men and women in 2012 and the most deadly cancer that year [1], starts off as a lesion before evolving into a tumour. With that being said, the earlier it is detected, the lower the mortality rate. Thus, it is clear that lung lesion detection is of high importance and it is natural that advancements in technology be applied to an area of such relevance.

Image processing techniques have evolved in a direction that allows high resolution processing in near real-time, with Medical Diagnosis being one of the applications that can benefit from such tools. These advancements generally result in the development of new Computer Aided Diagnosis (CAD) systems or the improvement of existing ones to help medical practitioners, in this case radiologists, in their day to day work. Such systems aim to reduce any human error present, which can be caused by various reasons, including radiologist fatigue.

In this project, image processing techniques are applied to Computed Tomography (CT) scans of the thorax, thereby developing a CAD System, to aid in identifying lesions in lungs.

## 1.1 What is Computer Aided Detection?

Computer Aided Detection (CAD) involves tools developed for medical applications with the aim to support physicians in the process of identifying specific pathologies in an area of interest. Computer systems are common in digital data acquisition processes, such as Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) scans [2]. CAD is the next step, in which the acquired data is analysed and brings any suspicious features to the attention of the physicians, who use the computer output as a "second opinion". Using this type of CAD systems, rather than fully automated CAD systems where the physician is non-existent, the performance does not have to be better than or comparable to that of physicians, but rather complementary to aid in the diagnosis [3]. Doi also notes that CAD systems are already routinely in use in some areas of medicine, such as in breast cancer detection on mammograms [3].

A test was carried out by Roberts [4], where a CAD system was used as a second reader. The results of the test show that the CAD system was able to detect an additional amount of lesions that were not observed by the physician, meaning that the inclusion of CAD systems in diagnosis is able to increase detection rates. Roberts also points out that no CAD system is 100% sensitive and therefore, no CAD system is fully autonomous. False positives, that is, the identification of a lesion when it is not the case, are also expected from every CAD system. Such a limitation may also increase the overall reading time. Depending on the way a CAD system is used, Roberts reports that a CAD system can either improve detection or decrease interpretation time but not both, while also stating that these results require verification. A medical imaging modality that can make use of such CAD systems is Computed Tomography.

## 1.2 What is Computed Tomography?

CT is a non-invasive technique that makes use of ionising radiation in the form of X-Rays to obtain cross-sectional images of the patient's organs. CT is a response to the limitations of Classical Tomography. The most significant limitation is the fact that 3D volumes are mapped onto a 2D image, resulting in structures in front of and behind the plane of interest being superimposed on the image, although blurred [5, 6]. CT solves this issue by producing cross-sectional images which do not include superimposed structures. The images generated by CT scanners can be fed to a CAD system to be analysed for any harmful pathologies. Appendix A shows the typical output slices of a CT scanner.

## 1.3 The Inner Workings of a CT System

### 1.3.1 X-Rays and Absorption

The intensity of an electromagnetic wave can be defined as "*The rate of flow of electromagnetic radiation energy through unit area perpendicular to the direction of motion of waves*" [7]. The intensity can therefore be considered as the electromagnetic power per unit area.

X-Rays enter tissue with an initial intensity $I_0$ and emerge from the tissue with a lower intensity $I$. These two quantities are related by the following equation, known as Beer's Law:

$$I = I_0 e^{-\mu x} \tag{1.1}$$

where $x$ is the distance through which the X-Rays propagate and $\mu$ is a proportionality factor called the linear absorption co-efficient [8, 9].

$\mu$ is proportional to the density ($\rho$) of the material through which the X-Rays are moving and is also dependent on the X-Ray wavelength [8, 10]. Thus, each material has a different value for $\mu$ for a particular value of wavelength. The denser the material, the larger the value of $\mu$ and, as a result, the lower the intensity of the emerging X-Rays. This property is called the radiodensity or the radiopacity of a material. For example, bones are much denser than air and therefore have a much larger radiopacity. This results in bones absorbing much more X-Rays. Therefore different tissue has different shades on the image: bones appear white whereas air appears black [11].

### 1.3.2 Obtaining an Image

The problem of superimposition in classical tomography is solved by using computerised techniques to reconstruct an image by taking multiple viewpoints of the same image. Multiple projections would give more information about the object. In radiology, these multiple projections obtained at different angles are called views. Given that X-Rays penetrate the human body, multiple views would not only provide information about the surface but also about the contents. Therefore, such views define the absorption profile from a given angle.

A single point in the resulting view would be the sum of the absorption values of all the tissue through which the x-ray passed, called a raysum [12]. Filler [6] uses an example of shining rays over a perspex cylinder with a dense dowel at its centre. The receiver will experience high intensity when the rays pass over the perspex, whereas the intensity at the receiver will be reduced once the beam is over the dowel. The cylinder is then rotated and the process is repeated, therefore obtaining multiple views. Figure 1.1, taken from [5], illustrates how each view measures different parts of the same object, thereby giving different information.

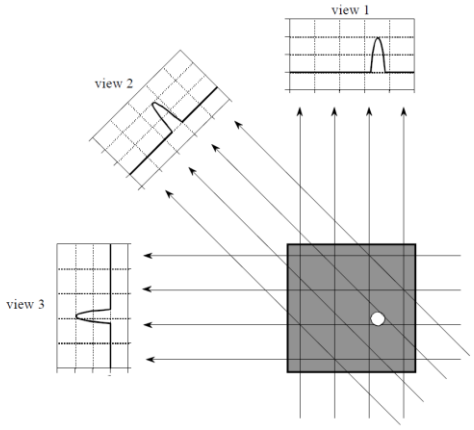
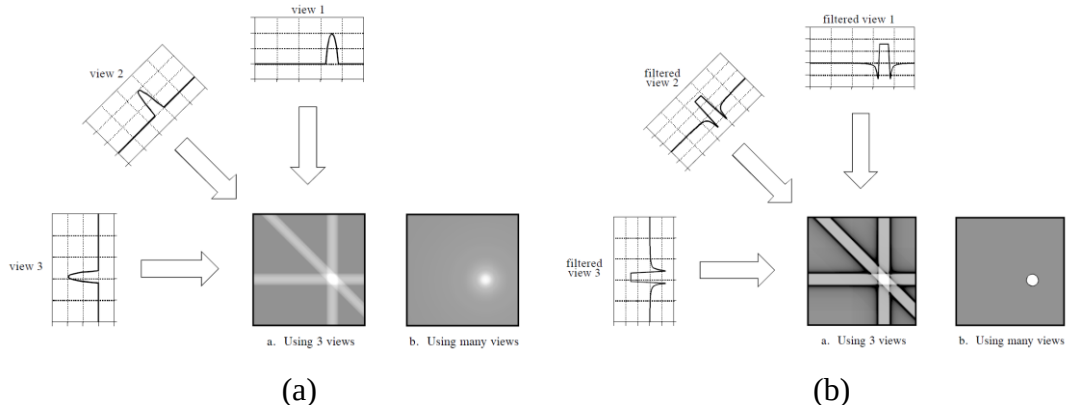



**Figure 1.1:** *Different views of the same object. [5]*

According to [5], reconstruction of the original object can be done with four main CT reconstruction algorithms. The first method uses computers to solve a considerable amount of simultaneous equations to obtain the output image, whilst the second method described, makes use of iterative techniques. Both are quite inefficient and slow.

The last two techniques are based on mathematical solutions and digital signal processing. A method called simple backprojection can be used, where each view obtained is smeared back through the image plane in the same direction which created it. The final result is the sum of all the backprojections. However, the resulting image is blurry, as can be observed in Figure 1.2(a). Filtered backprojection [5, 12] is an improvement over simple backprojection, where this blurring effect is corrected by filtering each view before it is backprojected, as can be seen in Figure 1.2(b). This technique is capable of creating an image identical to the actual object if an infinite amount of views are taken [5]. Note also the changes the filtering invokes on the views: the pulse is flattened, such that a uniform signal is observed inside, and negative spikes are introduced. Such negative spikes counteract the blurring.

This new method of obtaining CT information is based on two main principles: the Radon transform and the Fourier Slice Theorem. Radon transform mathematics is the technique behind the multiple-view approach to obtain a better image, whereas the Fourier Slice Theorem is what inspired Filtered Back Projection to reduce the blur [5, 12].

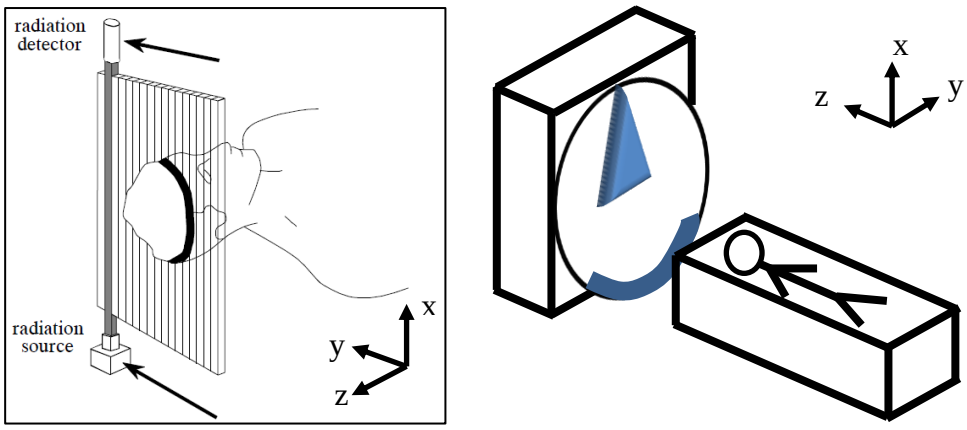



(a) (b)

**Figure 1.2:** *Reconstruction using (a) simple backprojection and (b) filtered backprojection [5].*

Whereas the first prototypes made use of a pencil beam of X-Rays and a single detector, newer generations replaced this narrow beam and single detector pair with a fan beam configuration and an arc array of detectors. This also helped to reduce the time needed to scan an object, since the translation of the beam in the *y* direction to obtain a single viewpoint was eliminated altogether: this new configuration provided a wide enough beam and detector pair to map one view instantaneously. These different configurations can be observed in Figures 1.3(a) and 1.3(b) below.

These improvements resulted in the helical scanners which are used today, where the patient is placed on a motorized bed inside a circular opening of a structure called the gantry. The gantry, containing the X-Ray source and detectors, would keep rotating whilst the patient is moving in the *z* direction. The X-Rays therefore, form a spiral around the patient, which is where the scanners get their name.




(a) (b)

**Figure 1.3:** *(a) Narrow beam and detector pair [5] and (b) Fan beam and detector pair, coloured blue [13]*

## 1.4 Focus of this Study

The area of focus of this study is the development of a CAD system that aids in the detection of lesions in thoracic CT scans. The thorax, also called the chest, is *"...the upper part of the trunk or cage of bone and cartilage containing the principal organs of respiration* [lungs] *and circulation* [heart and blood vessels]*..."* [14].

The same source [14] defines a lesion as *"A wound, injury, or pathological change in body tissue."*, an example of which is shown in Figure 1.4. In CT scans, lesions generally present themselves as spots [15]. A lesion's main characteristics, when imaged using CT, are its intensity and its shape. Lesions tend to have high intensities, that is, the areas that correspond to a lesion generally have whiter intensities than normal lung tissue, which appears darker. Having said that, other vessels present in the lung also have high intensities and therefore, the shape of the volume under observation is also of importance. Lesions tend to have spherical shapes, that is, they are enclosed volumes. This is in contrast to vessels, which are continuous along a large number of slices. Therefore, lesions tend to appear and disappear from CT slices abruptly. Thus, in order to actually know if the region corresponds to a lesion or not, the same region in previous and following slices must be observed. With reference to Appendix A, the same lesion in Figure 1.4 can be observed with respect to the CT slices before and after it. This lesion is indicated by a red arrow in Appendix A.

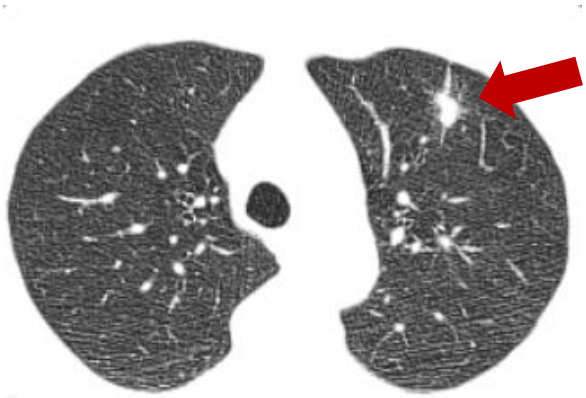


**Figure 1.4:** *An example of a lesion, indicated by the arrow*

As already discussed, early detection of lesions in the chest is of utmost importance. Thus, the development of a CAD system, that takes as input CT scan images and produces as output an indication of where these lesions may be present, is beneficial to patients and radiologists. Therefore, the main objectives of this study are to:

- Develop a CAD system which makes use of image processing techniques to detect lesions in thoracic CT Scans.
- Analyse the output of the developed system and evaluate its outcome.
- Propose any future work that might improve the CAD system developed.

## 1.5 Structure of this Report

In this report, Chapter 2 gives a brief background about image processing techniques that are of importance to this project. It also analyses literary works pertaining to the area. Chapter 3 discusses the specific implementation of the CAD system developed in this study. In Chapter 4, the results obtained are analysed and the CAD system is evaluated against other CAD systems proposed in literature. Finally, Chapter 5 contains a summary of the study and concluding remarks. It also gives light to any future work that might improve the results of the CAD system.

# Chapter 2 : Background and Literature Review

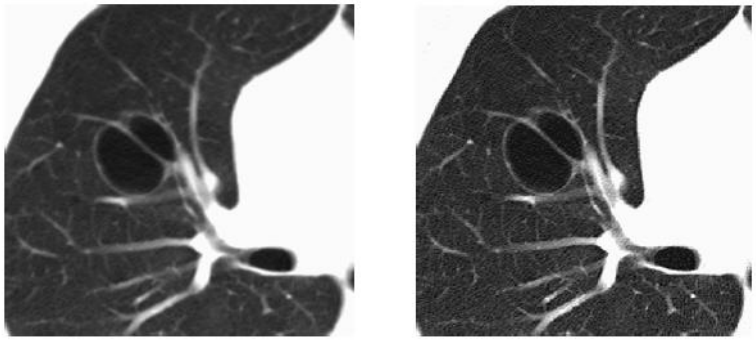
This chapter gives a brief background about the techniques involved in CAD system development. Such techniques include noise removal, contrast enhancement, segmentation, and finally the detection of lesions. Literature which compares and analyses the different techniques is also discussed in light of the problem to be tackled.

## 2.1 Noise Removal

In every measurement in the real world, random noise is a degrading factor which is fundamentally present [16]. Images obtained by CT scans are therefore, like all other signals, prone to noise. Thus, filtering out noise is one of the first pre-processing steps taken when the CT images are fed to the CAD system.

A small contributor to the image noise found in CT scans is electronic noise. The random motion of electrons is what causes noise from electronics, called thermal noise [5, 17]. Having said that, it is quite irrelevant when compared to other sources of noise in CT.

Some noise is also present based on the reconstruction parameters chosen. As discussed in previous sections, the backprojections used to reconstruct the image must first be filtered. The sharper the filter, also called the kernel number, the sharper the resulting image will be, but the more noise it will contain [17, 18]. This can be clearly seen in Figure 2.1 where (a) is not as sharp as (b) but it is less noisy.



(a) (b)

**Figure 2.1:** *Resulting image after the reconstruction filter, with (b) being sharper but noisier than (a). [18]*

The largest noise in CT scan images is related to the radiation dose used in the CT scan procedure. In general, this form of noise provides the worst signal-to-noise ratio (SNR) and is therefore called the quantum sink [5]. It can be shown that the noise level is inversely proportional to the root of the number of photons used in the scanning process [17]. The larger the dose, the more photons available and therefore, the lower

the noise. However, it is important to keep in mind that radiation can be harmful to humans.

One of the contributing factors to the dose level is the tube voltage, measured in kV [17, 18]. The tube voltage defines the beam energy. The larger the tube voltage, the lower the noise in the image. Another parameter which affects the amount of noise is the mAs [17, 18]. The mAs is the current-time product. Therefore, noise is not only dependent on the tube current, but also on the time the tube takes to perform one rotation. The higher the mAs, the larger the radiation dose and the less noise present in the CT images. The dose is also dependent on the pitch [17]. The pitch is defined as the ratio between patient table feed per gantry rotation. A low pitch will result in a lower effective radiation dose into the patient. This would therefore generate more noise.

Slice width also plays a role in noise generation [17, 18]. The wider the reconstructed slice, the more photons would be available for image reconstruction and therefore, the less noise in the image. However, a wide slice would result in a lower resolution in the *z* axis, since fewer slices would be available in the reconstruction process.

As Toth [17] notes, the kV, mAs and pitch are highly dependent on the anatomical region being scanned and the clinical indications for the scan.

A low signal-to-noise ratio can make low contrast objects indistinguishable [17]. Therefore, filtering noise out is important to aid in further processing in the CAD system. Various techniques exist to filter noise. However, they are all based on smoothing or blurring principle. This is equivalent to a low pass filter. Noise presents itself as sharp transitions, resulting in high frequency content. Unfortunately, edges in an image are also considered as high frequencies and therefore, the negative effect of filtering is the blurring of edges [12, 16].

One type of filter that aids in removing noise is a **mean** filter. A mean filter is based on the principle that noise is random and therefore, the probability that erroneous pixels are clustered together is low. Thus, a mean filter works on an $m \times n$ neighbourhood, centred at a pixel $P$ with coordinates $(x, y)$. Different mean filters can be applied, such as an **arithmetic mean filter** or a **geometric mean filter** [12]. Another variation of the mean filter is an **α-trimmed mean filter**, where the average is calculated without taking into consideration the α largest and smallest values. This should be more effective against shot noise [19].

In the above mentioned filters, all the pixels in the neighbourhood are given equal weighting. However, it is likely that pixels closer to *P* would be more correlated and therefore they would provide much more valuable information than pixels farther away. Thus, the idea of a **Gaussian Filter** exists, where each pixel in the neighbourhood is given a weighting: the closer the pixel is to *P*, the higher its weight and its influence on the new value of *P* [16]. This distribution of weights resembles a Gaussian distribution, which is where the filter gets its name. Vijaya *et al* [20] compare both the α-trimmed mean filter and the Gaussian Filter with other filters discussed below. The study concluded that the other filters provided better results.

Another family of filters that can help reduce noise are **median filters**. Like mean filters, median filters work on the neighbourhood surrounding a pixel. Median filters replace the pixel *P* by the median pixel intensity in the neighbourhood. The idea is that values which are affected by noise, such as impulse (shot) noise, are eliminated since they would be at the extremes. It can also be observed that median filters blur edges less, which is a useful property [16, 19]. However, Szeliski notes that such a filter might not be very effective to remove Gaussian noise [19]. To get around this, he proposes the use of a **weighted median filter**, where each pixel intensity is used *n* times when calculating the mean, where *n* depends on the spatial distance between pixel *P* and the pixel in the neighbourhood [19]. **Adaptive median filtering** is an improvement over median filtering, where the filter changes according to the local image properties. A 5×5 adaptive median filter was used by Ashwin *et al* [21] in their CAD implementation. Adaptive median filtering was also compared to other filters by Vijaya *et al* [20]. However filtering using this technique did not achieve the best results.

Similar to Gaussian Filtering, **Bilateral Filtering** aims to denoise an image by taking a weighted average of neighbouring pixels. However, unlike Gaussian filters, bilateral filters do not only take into consideration the spatial distance of pixels, but also give a weight to the variation of intensity of pixels. This, therefore, aims to reject pixels whose intensity differs too much from the pixel *P* and results in better preservation of edges. In order to weigh the spatial distance and the intensity variations of pixels separately, two different functions can be used [19]. It should be noted that, since Bilateral filtering is a non-linear technique, it is not separable, making it slower when compared to linear filtering. Having said that, [19] cites several accelerated versions, at the expense of using more memory. Another disadvantage of the Bilateral filter, is the fact that certain noise,

such as shot noise, might not be removed, since the noisy pixel's intensity would differ too much from the intensity of the neighbouring pixels. It can be noted that in Vijaya *et al*'s works [20], Bilateral filtering produced the best results. However, when compared to other techniques, such as anisotropic diffusion and a sparse representation based method [22], this was not the case.

**Wiener Filtering** is another technique that can be incorporated to reduce image noise and blur. It works by trying to minimise the mean square error (MSE) of the reconstructed image and the original image. This is done by taking into account information about the degradation function which caused the blur, such as camera motion, and the noise that degraded the image. Wiener filtering, therefore, takes a statistical approach to remove noise. The resultant equation is the following:

$$\hat{F}(u,v) = \frac{H^*(u,v)}{|H(u,v)|^2 + \left[\frac{S_\eta(u,v)}{S_f(u,v)}\right]} G(u,v) \qquad (2.1)$$

where:

- $\hat{F}(u,v)$ is the image in the Fourier Domain,
- $H(u,v)$ is the Fourier Transform of the degradation, with $H^*(u,v)$ being its complex conjugate,
- $S_\eta(u,v)$ is the noise spectral density
- $S_f(u,v)$ is the signal spectral density
- $G(u,v)$ is the corrupted image's Fourier Transform

A full derivation of (2.1) can be found in [12] and [19]. In this study, it can be assumed that no degradation, other than noise, is present such that $H(u,v) = H^*(u,v) = 1$. From this equation, it can be noted that for high frequencies, that is, where noise is usually present, the attenuation is large since $S_f(u,v) \ll S_\eta(u,v)$ [19]. The opposite is true for low frequencies. One of the main disadvantages of Wiener filtering is the fact that the mean square error is used, which weighs all errors equally [16]. The Wiener filter is used in several CAD implementations in literature, including [23] and [24].

**Anisotropic diffusion** takes yet another approach to reduce noise in an image. Although blurring an image is effective in reducing noise, it is clear that edges are of high importance and should not be blurred. Anisotropic diffusion introduces a function to

regulate the amount of blurring that should happen: if no edges are present, the function should have a high result, such that blurring occurs. If edges are present, the output of this function should be a low value, such that blurring is limited [22]. The gradient of an image is quite a good detector to indicate if a pixel is part of an edge or not. Thus, a function dependent on the gradient of an image is suitable to control the smoothing intensity. Perona and Malik give two definitions for this function which can be found in [25].

In [22], the authors compare a sparse representation based method, bilateral filtering and anisotropic diffusion filtering. They note that the sparse representation based method provides the best peak signal-to-noise ratio (PSNR), at the expense of more loss of detail and increased computation time. This is not the case with anisotropic diffusion filtering, which appears to provide a good balance between time, noise removal and detail preservation.

## 2.2 Contrast Enhancement

Another processing step that is performed in a CAD system is contrast enhancement. Contrast refers to the difference between intensity values, and is defined as a ratio of the average intensity of a foreground object and the background [16]. The larger the contrast, the more pronounced the difference between the object and the background.

An image's histogram shows the number of pixels present in the image having a particular intensity level. Given that contrast is based on the intensities in an image, a histogram is useful when enhancing contrast. An image with low contrast will have its intensities centred in a small band in the histogram, that is, most pixel values will have similar intensities. An image with well-balanced contrast has intensities spanning over a large range of the possible intensities, therefore, making details more visible. If the contrast is high, the histogram is said to be bi-modal, since two defined peaks will be shown. The extreme case of high contrast is a binary (black and white) image, where the intensities can only be one of two values [26]. In cases were the contrast is high, details might be lost. This is because objects having slightly different pixel intensities might have their values saturated to one of the peaks.

To differentiate between structures, it is important that CT scan images have a good contrast level. Image contrast is so important in CT, that patients may be administered a contrast enhancing dye before being scanned, even though such a dye can, on rare

occasions, have negative side effects [27]. Obtaining a good contrast level in CT images will therefore aid in the detectability of any harmful pathologies. By increasing the contrast, light areas will become lighter and dark areas will become darker.

Intuitively, enhancing the contrast of an image requires stretching the histogram. **Contrast stretching (or Normalisation)** refers to a process that increases the dynamic range of intensities of an image, such that they span the full possible range [12]. By applying contrast stretching, the histogram of intensities of the original image, which may only span over a small range of intensities, is mapped over the whole range. This can be achieved by multiplication. It is important that, when defining the multiplication factor, it is ensured that saturation does not occur. Saturation takes place when, due to the multiplication, the new intensity is outside of the range of intensities that can be represented, and is therefore clipped to the minimum or maximum value. Saturation could result in a loss of contrast in some parts of an image. Contrast stretching was used in [23] to improve the CT image. In general, contrast stretching is performed using the following equation:

$$Output\ Image = \big(Image - Min(Image)\big) \times \frac{New\ Max - New\ Min}{Old\ Max - Old\ Min} + New\ Min \qquad (2.2)$$

where *New Min* and *New Max* determine the new histogram range. Note that, due to noise, it is possible to take values for *Old Min* and *Old Max* which are the $5^{th}$ and $95^{th}$ percentiles respectively [28].

A similar technique is **Histogram Equalisation** (HE), which aims to not only stretch the image over the whole intensity range, but also tries to distribute the intensities such that the histogram is uniform [12, 16]. This aids to avoid obtaining a bi-modal histogram, which would result in a very high contrast image. It should be noted that Gonzales *et al* [12] acknowledge the fact that a uniform distribution might not be the best distribution for every type of image. Therefore, the concept of **histogram matching (or histogram specification)** is introduced where a non-uniform histogram shape may be defined. Note that, however, Gonzales *et al* also emphasize that this method is, for the most part, based on trial and error [12]. It can also be noted that HE, although simple, is not able to adapt to the local features of an image, since it makes use of a global histogram. This may therefore, result in contrast losses and loss of detail [12, 21, 29, 30]. Kong *et al* [30] note that another disadvantage is the fact that HE can have a bias when enhancing the

contrast, such that more common intensities are enhanced whereas other intensities may experience loss of contrast. This technique can also cause a saturation effect, which may lead to loss of information, whilst also enhancing any noise present [30].

In order to obtain better results, **Adaptive Histogram Equalisation** (AHE) operates on a local histogram, allowing it to cater for different local content, such as small details [30]. The local histogram of each pixel is calculated from a neighbourhood surrounding this pixel, a sub-block of the image, with the pixel located at the centre of the sub-block [12, 21, 29, 30]. Although AHE is better than HE, it can enhance noise [30, 31] and is computationally expensive, with the cost increasing as the image size increases. This is because, for each pixel, a new sub-block and histogram must be defined and computed [29]. Therefore, to try and improve on AHE, various improvements of it exist, including **Contrast Limited Adaptive Histogram Equalisation** (CLAHE) to reduce the noise enhancement factor [21]. **Partially Overlapped Sub-Block Histogram Equalisation Method** (POSHE) aims to reduce the computational complexity of CLAHE, but still achieve similar results [29]. These versions of AHE and others can be found in [30].

Kong *et al* [30] also mention two other groups of contrast enhancement techniques which are improvements over HE. **Mean Brightness Preserving Histogram Equalisation (MBPHE)** techniques are used to avoid unnatural enhancement of the input image, thereby maintaining the artistic value of the image. However, given that this study works with scientific images rather than artistic images, these techniques will not be considered. The second group of techniques is **Bin Modified Histogram Equalisation (BMHE)** techniques, which modify the histogram before it is expanded by increasing or decreasing the counts per bin according to some thresholds. These techniques can be considered to be dataset dependent which is not desirable, since this study aims to develop an automated system which works on all the slices in a CT scan. For these reasons, this group of techniques will also not be considered.

A **Power Law Transformation** (more generally called **Gamma Correction**) is a non-linear technique which changes image values by using a power law transformation [12]. The power law transformation can be written as follows:

$$Image_{Corrected} = c \times Image^{\gamma} \tag{2.3}$$

where *c* is a constant which controls the brightness and can result in loss of detail [12, 32]. Therefore it can be removed, resulting in:

$$Image_{Corrected} = Image^{\gamma} \tag{2.4}$$

With the above definition in mind, for a $\gamma < 1$, any normalised image (that is, having image intensities in the range [0, 1]) would lighten up, whereas for a $\gamma > 1$, the image would darken. A $\gamma$ equal to 1 results in no change in the image [12]. In reality, Gamma Correction was developed to correct for differences between voltage-intensity relationships between different displays [12, 16]. However Gonzales *et al* [12] also note its use in contrast enhancement. In fact [31] made use of such a technique to enhance medical imagery by automatically calculating the value of gamma from each image's cumulative histogram. Also, Al-Ameen *et al* developed a **Normalised Gamma-Corrected Contrast Limited Adaptive Histogram Equalisation (NGCCLAHE)** method [32], which makes use of both Gamma Correction and CLAHE to enhance CT images, with the technique developed being compared against several contrast enhancing techniques. They report that it outperformed the other techniques.

Pizer *et al* [33] tested **Weighted Histogram Equalisation** (WHE), which is another form of AHE, on CT images and noted that it produced relatively the same results as AHE but took more time. They also suggest that CLAHE is used on medical imagery, compared to WHE and another AHE adaptation: **Interpolated Adaptive Histogram Equalisation.**

## 2.3 Segmentation

Segmentation divides an image into different regions having distinct characteristics. One of its main aims is identifying a region of interest in the image. In the case of this study, the region of interest, which is the pulmonary structure, is extracted from the CT images after they are pre-processed. Gonzales *et al* [12] note that segmentation is a challenging task, with its accuracy playing an important role in the success or failure of an algorithm. Segmentation will not only allow the detection phase to focus only on the region of interest, thereby allowing it to obtain better results, but will also have the added benefit of reducing the detection phase's computation time, since it will not need to perform an exhaustive search over the whole image.

**Thresholding** is one of the most simple and computationally inexpensive segmentation techniques [12, 16]. It splits an image into regions based on a comparison between pixel intensities and some threshold value. Selecting a correct threshold value is crucial in this method of segmentation. Having said that, it is quite unusual for a single threshold value to be used for the whole image. When a single threshold is used for the whole image, the process is called **global thresholding**. However, the more common approach is to make use of a variable thresholding value. This is called **regional thresholding**, where the threshold value is dependent on the local image characteristics [12, 16]. To separate an image into more than two regions, more than one threshold value needs to be selected.

Since thresholding directly involves the pixel intensities, an image's histogram is particularly useful. From the histogram, some important properties can be observed, such as the peaks and their separation, which may aid in the threshold value selection [12]. However, selecting threshold values is generally not trivial. Otsu's method [12] is a threshold selection method based on the idea that a good threshold value will create regions that are distinct, such that the variance between regions is large. It is further described in [12].

Silva *et al* [34] propose three methods to segment lungs from CT images efficiently. The first method adapts a B-form smoothing spline to the intensity histogram. A spline is a piecewise polynomial interpolation representation of a curve. A threshold value is then selected, based on the fact that lung tissue corresponds to approximately 25% of the intensity range. With this threshold, a binary mask is obtained which is improved by morphological operators. More information regarding morphological operators can be found in [16]. Silva *et al* note that this method does not separate the left and right lungs if the boundary between them is small, and it is sensitive to structures other than the lungs, such as vascular structures and the trachea. They further improved upon this first method by applying the Radon Transform in both the horizontal and the vertical directions of the binary image to obtain two mass centres corresponding to each lung. They then apply the morphological fill operator, using these centres as seed points. If the two resulting regions connect, they are separated. Silva *et al* note that this method always separates the lungs, even if they are not actually separated. However, compared to the previously discussed method, it is less sensitive to other structures.

**Region Based Segmentation** takes a different approach to obtain regions in an image, aiming to find regions based on some homogeneity criteria [12, 16]. Homogeneity criteria can be based on various properties such as intensities, textures, or a combination of properties.

**Region Growing** is a region-based segmentation technique that starts with a set of initial seed points. Regions are formed from these points by adding more pixels to them based on some homogeneity criteria. These criteria also act as a stopping rule, such that the region stops growing. Gonzales *et al* [12] go on to describe a basic region growing algorithm. Two other region based segmentation techniques are also described in [12] and [16]: **region splitting** and **region merging**. Before describing these methods, it is important to understand the following region properties. Let $H$ represent the homogeneity criterion. For a region $R$:

$$H(R) = True \tag{2.5}$$

Also for two regions $R_1$ and $R_2$:

$$H(R_1 \cup R_2) = False \tag{2.6}$$

where $\cup$ is the union operator. The second property means that a region is as big as it can be, that is, adjacent regions cannot be merged.

The region merging method starts by considering each pixel as a set. Regions are merged together as long as (2.5) is satisfied and (2.6) is not. This procedure goes on until both (2.5) and (2.6) are satisfied.

On the other hand, region splitting works in an opposite way: the whole region is considered as a single region, for which (2.5) will surely not be satisfied. Thus, the region is split. This procedure is repeated until both (2.5) and (2.6) are satisfied.

Both [12] and [16] define algorithms where merging and splitting are used simultaneously.

Memon *et al*'s algorithm for segmentation [35] starts by assuming a reasonable threshold value: the mean grey value of the pixels. This threshold is improved upon iteratively until it converges, after which it is used to separate lung tissue from lung cavities. The authors note that this process must be repeated for all the images in a CT

scan. Although this is time consuming, Memon *et al* believe that this yields optimal results. Once thresholding is complete, a seed pixel is chosen from the area surrounding the lung structure. This corresponds to a pixel having a dark value (generally 0). From this seed pixel, region growing is performed to select the border surrounding the lungs. These are then assigned a white value. White pixels already present in the image are tagged before region growing, so as not to be confused with background pixels which were just set to white.

**Watershed** segmentation is based on the concept of viewing an image as a 3D topographical map. Apart from 2D spatial co-ordinates, the intensities of the pixels represent a 3$^{rd}$ dimension. With this 3D concept in mind, the notions of catchment basins and watershed lines are introduced. Catchment basins hold within them pixels that obey the same homogeneity rule and therefore conform to a region. Watershed lines, on the other hand, are used to separate regions from one another [12, 16].

In order to obtain watershed lines, both [12] and [16] introduce a concept called "flooding", which works as follows. Assume that water is pumped though the minima (catchment basins) of the image. When the water from separate catchment basins is about to merge, a dam is created. The dams represent the watershed lines.

Direct use of watershed algorithms will not always render a usable result, since it generally leads to over segmentation: many unusable regions will be present. To counteract this problem, markers are used to control segmentation. Internal markers are used to define catchment basins and external markers define watershed lines. Thus, fewer regions will be present since not all the minima will contribute to a region. This can be seen in Figure 2.2, where defining two markers has led to the creation of two regions rather than three. More details can be found in [12].

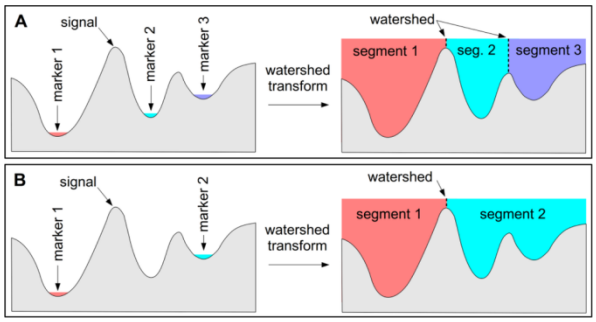



**Figure 2.2:** *A visual example of the marker based watershed algorithm [36]*

A marker-based watershed algorithm is used in [37]. Alirezaie *et al* note that by this method, the problem of finding an optimal threshold is eliminated. A marker-based watershed transform is run on the gradient image of the CT image. The gradient image was obtained by using a Sobel operator [16] in both the vertical and horizontal directions. Markers are used to avoid over segmentation. A rolling ball filter is used to smooth and fill any cavities in the mask.

**Active Contours**, also called snakes, can be used in image segmentation. A snake may be defined as an energy-minimising spline [16]. Since snakes aim to minimise energy, image properties represent minima in the energy function, such that the deformable snake takes the required shape based on image information. It can be noted that the energy function is also associated to features of the curve. Further details as to how energy functions are modelled can be found in [16], which gives an insight of what an energy function is composed of and how to model terms so as to obtain the desired behaviour from the contour.

Zhou *et al* [38] note that different CT images might have different contrast. Thus, setting a fixed threshold is not robust enough to perform adequate segmentation, whereas obtaining an optimal threshold iteratively is time consuming. Zhou *et al,* therefore, propose to find the optimum threshold by histogram stretching and Otsu's algorithm [12]. Once thresholding is performed, the region with the largest contour is retained, removing the background. Within this region, objects with a large enough contour, that is, larger than the trachea, and having values lower than the previous threshold found, are considered as texture regions and therefore, smoothed. An active contour which is texture aware is employed to refine the segmentation.

Silva *et al* [34] model the energy function for the active contour using forces, including the elastic force of the contour and the image intensity, gradient and expansion forces. They also note that the forces must be calculated for each point on the contour. The algorithm is iterative, in the sense that it must be repeated several times for the contour to move towards the lung shape. Once a specific criterion is met, the iterations stop. Silva *et al* note that the contours are initialised as circular contours, with the seed points found by the Radon Transform. They also state that, the number of points on the contour is related to the computational load. With that in mind, they made their algorithm adaptive, such that the contour is resampled as it expands. It can be noted, however, that

the algorithm did not always segment the lungs correctly, and is more time consuming than the thresholding methods proposed by the same authors in [34].

In **Graph Cut Segmentation**, each pixel is treated as a node in a graph, with two extra terminal nodes added: one for the foreground and one for the background. Every pixel node is connected to its neighbouring pixels, the foreground and the background node by a weighted edge [16, 39]. A visual representation is shown in Figure 2.3(a). A cut is obtained by severing edges such that no path exists between the background and foreground nodes. This means a cut will result in two separate graphs, one with the root node being the foreground and the other with the root node being the background. With the creation of the two graphs, the foreground and background are separated, such that, every pixel that remains attached to the background node after the cut, is considered as a background pixel. Any pixel connected to the foreground node is a foreground pixel. A visualisation of a cut can be seen in Figure 2.3(b).

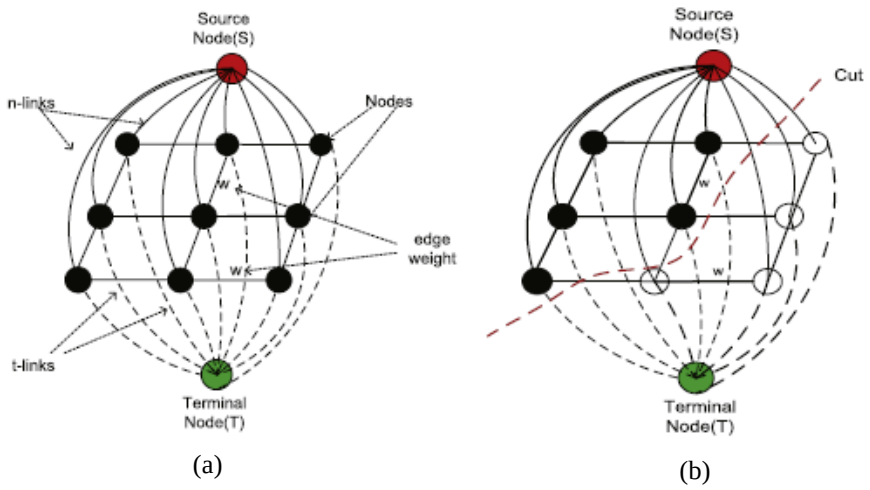



**Figure 2.3:** *A visual representation of the graph (a) and graph cut (b), where node S represents the foreground and node T the background [40].*

Several cuts exists, thus the minimum cut is desirable. The minimum cut has the lowest costs over all the other possible cuts. The total cost of a cut is equal to the sum of all the weights of the severed edges. Thus, the weights assigned to each edge are of key importance. Weights between pixels can be assigned based on local image information. For example, if adjacent pixels have similar colours, then the probability that both belong to the same region is high. Thus, the edge between them should have a large weight, so that it is unfavourable to cut. The weights between a pixel node and the background and foreground nodes are generally calculated based on some probability

that a pixel is in the background or the foreground. To obtain such probabilities, seed points representing the foreground and background are generally required [16].

To solve the minimum cut problem, also called by its dual, maximum flow, several algorithms exist, as described by [16] and [39]. Sonka *et al* [16] also note that the choice of algorithm will, undoubtedly, affect the segmentation outcome.

Dai *et al* [40] use a graph cut approach to obtain the lung structure from CT images. They note that, although thresholding methods are fast, their dependency on the threshold value is an obstacle, whilst region growing is dependent on the growing criterion, which is too sensitive to parameters and is time consuming. Thus, they first smooth the CT image with Gaussian smoothing and obtain a probability model from the seed points. The probability model is obtained by a Gaussian Mixture Model, with an Estimation Maximisation Algorithm being used to learn the model parameters. Segmentation is then performed using a minimum cut algorithm. Dai *et al* note that, compared with region growing, this graph cut method does not over segment to include other regions, such as the bronchus. When compared with thresholding, their method performed full segmentation whereas thresholding did not.

A similar graph cut approach is used by [41]. However, the weight of an edge is calculated based on the grey value of the pixel, using a Linear Combination of Gaussians model, whose parameters are determined by Estimation Maximisation. Spatial interaction is also considered, modelled using a Markov Gibbs Random Field. The minimum cut algorithm is then run on an energy function formulated using both these models. Also, the authors in [41] make use of a 3D graph rather than just a 2D one. When compared with two iterative segmentation techniques, including an iterative threshold technique, Ali *et al* report that their algorithm outperforms the other methods.

## 2.4 Detection

Once all the previous steps have been completed, the images can enter the detection phase. The previous steps were performed to prepare the images for this phase, such that the best detection rate possible will be achieved. The main focus of a CAD system is, of course, the detection phase. Having said that, detecting lesions in a CT scan is not a simple task, since lesions and vessels may appear similar in CT scan images. Thus, 3D information is generally required to differentiate between them.

Antonelli *et al* [42] note that lesions have intensities similar to vessels and therefore, lesion detection based only on intensity is not enough. To cater for this, they propose a method that detects spherical regions in 3D within a given grey level range. To start with, they denote the function to represent the convolution of a voxel's intensity with a 3D Gaussian filter. To calculate the shape index, partial derivatives are calculated with the convolution property. Using the first and second fundamental forms of differential geometry, the curvature is found. With these values, the shape index is calculated. To account for any derivatives that may result in an undefined shape index, they adjusted their formulas by axis rotation. Antonelli *et al* note that shape indexes close to 1 indicate spherical shapes, whereas a shape index closer to 0.75 would indicate a cylindrical structure which would resemble a vessel. With the intensity and shape index known for each voxel, a region growing algorithm is applied, using both the intensity and shape index as growing criteria. As a seed point, a voxel which has both the intensity and shape index within the specified range is chosen. Antonelli *et al* note that this algorithm gave a small number of false positives, which resulted due to s-shaped blood vessels that were not distinguishable from lesions attached to the pleura.

After pre-processing, Namin *et al* [43] obtain the lung structure and enhance the image by Gaussian smoothing. They use Gaussian kernels with different sizes to cater for a range of nodule sizes. After the convolution with the different kernels, the maximum intensity value obtained is kept. According to Namin *et al*, this emphasises spherical nodules and blurs cylindrical vessels. Once this is done, the shape index is calculated, similar to [42], with values around 1 representing spherical shapes and values around 0.75 representing cylindrical shapes. A threshold of 0.8 for the shape index, determined experimentally, is used to produce 3D connected voxels. Voxel-sets containing a minimum of seven voxels are kept. Once again, the value of seven was found experimentally. Unlike [42], Namin *et al* note that not all voxels pertaining to a lesion candidate will be present in the above selection, since not all voxels will meet the sphericity criterion. With this in mind, they calculate the mean intensity value for a voxel-set. Neighbour voxels having intensity in a range centred around this mean value are added to the voxel set. The range is also determined experimentally, with Namin *et al* stating that the best range is ±0.05 times the mean value. The inclusion of neighbouring voxels to the voxel-set is iterated until no more voxels are added. Once this procedure is finished, the whole candidate object is contained within a voxel-set.

From the voxel-set, several features are extracted and normalised in the range [-1, 1]. These are then fed to a fuzzy k-nearest neighbours classifier, which classifies the candidates as malignant nodules, semi-malignant nodules, benign nodules and non-nodules using a Euclidean distance measure. A classical k-nearest neighbours classifier [44] chooses *k* neighbours to the input object. These neighbours have maximal similarity between their features and the input object's features, as defined by some distance measure. The input object is classified as belonging to the class that the majority of these *k* neighbours belong to. In fuzzy k-nearest neighbours [44], the membership of the input object is tested for every class which has at least one of the *k* neighbours belonging to it. Thus, classification is no longer based only on a majority vote system, but is also weighted by the similarity distance between the input and the neighbours as well. Namin *et al* note that the optimal number of neighbours to be considered were seven.

In [45], Suiyuan *et al* note that, after obtaining the lung structure, the histogram of intensities is characterised by a sharp peak, with only the last part of the tail containing intensities that may lead to lesion seed points. They choose a threshold value to obtain seed points for a region growing algorithm. The threshold is chosen by applying a method inspired by the decay of radioactive materials. Suiyuan *et al* note that the parameters used in the 'half-life' method are flexible, since the seed points will be fed to the region growing algorithm. The algorithm's region growing criterion is based on the pixel intensity and is used to connect voxels in 3D. The next step is to filter out candidates which are not lesions by observing shape information. To do this, they use image moments to obtain the centroid of the candidate. Once this is done, the distance from the centroid to each point on the contour is calculated. The standard deviation to the mean is used to distinguish between lesions and vessels. They also define an equation such that, for a lesion, the result is higher than that for other tissue. When these values are plotted, the highest value is chosen, since this indicates the lesion. Suiyuan *et al* note that this method yields good results for solitary nodules, with false positives being present because long vessels are disconnected by the region growing process.

After obtaining the lung structure, Tong *et al* [46] start detecting lesions by applying a shape filter. They note that the shape filter is dependent on the values of the Hessian matrix. They also note that this is applied for various scales, since lesion sizes may be different. Once the filter is applied, identification of true positives and false positives is obtained by considering features such as diameter, maximum and minimum intensity

values, sphericity and more. In order to classify the candidate lesions, rule based classification takes place: for each feature, a threshold range is obtained based on training data and diagnostic rules. If the feature value of a candidate object is within this threshold, then the output is true. Tong *et al* note that rule based classification is more efficient than neural networks and k-nearest neighbours classification.

A similar rule-based approach is used by Sharma *et al* [24]. The candidate lesions are segmented from the lung structure by means of Sobel [16] edge detection. Once candidate lesions are obtained, features such as shape, size and area are obtained from candidates. Rules are formed for each feature to filter out false positives from candidate lesions. For example, if the area is too small, the segmented object might have been the result of noise. Thus, it is eliminated from the candidate objects. Similar rules for each feature are formed to obtain lesions.

In [23], Banupriya *et al* also take a shape-based approach to find possible candidate lesions. The circular structure of lesions is found by using a circular Hough Transform. However, the technique is not well described and a method as to how false positives were reduced is not presented. Given that vessels might also have circular structures in a CT scan slice, the amount of false positives generated is expected to be high. Also, further literature to extract lesions from lung CT scans using a circular Hough Transform could not be found.

Support vector machines [47] can be used to classify data. The main aim is to use training data to train the classifier to a function which would separate the given data. Such a function, however, should generalise well enough such that it will also result in the target output when unseen data is used. Further reading on support vector machines can be found in [47].

Choi *et al* [48] make use of a support vector machine in their approach to detect lesions. The CT scan images are split into 3D blocks. Five different sizes were applied, to cater for varying lesion sizes. By applying the block sizes in decreasing order, the analysis is performed from rough to fine in a hierarchical manner. By placing a threshold on the entropy information of each block, blocks are filtered out such that low entropy blocks are eliminated. Choi *et al* note that, if blocks are filtered out at a large size, they do not need to be re-analysed at a smaller level. This should therefore aid to reduce the computational complexity. To select the threshold, they observe the histogram of regions

that are marked as lesions by radiologists. Once the blocks are selected, they are enhanced by a 3D filter. An iterative thresholding method is utilised to obtain candidate lesions. Whilst obtaining candidate lesions, the block location is also updated, such that the candidate object is at the centre. From the candidate objects, different features are extracted, including geometric features. A support vector machine is trained by selecting an equal amount of lesion and non-lesion feature vectors from the candidate lesions in the training data. These vectors, along with their target outputs are fed to the classifier, which obtains the optimal hyperplane that separates the lesions from non-lesions based on the feature vectors provided. The trained classifier is then used to distinguish between lesions and non-lesions from the candidate objects which were not used for the training process.

An artificial neural network [49] consists of several simple processing units called neurons, each with their own inputs and outputs. They are connected together via weighted connections. The weight determines the effect one neuron has on another. Three types of neural units exist: input units which receive input from outside the network, output units which deliver data out of the network, and hidden units whose input and output remain in the network. Neural units may also have a bias input. Every neuron has a propagation rule, which evaluates a single input from all the inputs. This singular input is then fed to an activation function, which determines the output. A neural network requires some learning function, in order to be trained to produce the desired output when given an input. The neural network can be laid out in one of two ways: in Feed-Forward networks, the input is only fed in the forwards direction towards the output, that is, no feedback is present. Conversely, in Recurrent networks, feedback exists.

Ashwin *et al* [21] implemented a neural network to detect lesions. The inputs to the neural network are the area and grey level intensities extracted after pre-processing the CT images, to reduce noise and enhance contrast, and segment candidate lesions by thresholding. A two layer, feed-forward network with sixty hidden neurons is used. The neural network makes use of a sigmoid function as the activation function. The iterative learning mechanism used is the Broyden, Fletcher, Goldfarb and Shanno algorithm, which is a quasi-Newton method. Backpropagation was used to update the weights. More information about this technique can be found in [21], whilst further literature on neural networks and techniques used in neural networks can be found in [49].

# Chapter 3 : Implementation

This chapter details the different phases of the CAD system implemented. It explains how the DICOM files are read and the pre-processing techniques used to reduce noise and enhance the contrast. The segmentation algorithm applied to extract the lung regions, and the detection methods used to obtain the lung lesion locations, are explained. It also describes the output of the system.

## 3.1 General Overview of the Implemented CAD System

It is clear that, as outlined in the previous chapter, the main steps of any CAD system involve reading in the images, pre-processing them, segmenting the lung structures from the images, detecting lesions and finally outputting the results. This results in the flow diagram shown below in Figure 3.1. The CAD system was implemented in C++, with the use of additional libraries. The implementation of each of these steps and the libraries used will be discussed in this chapter.

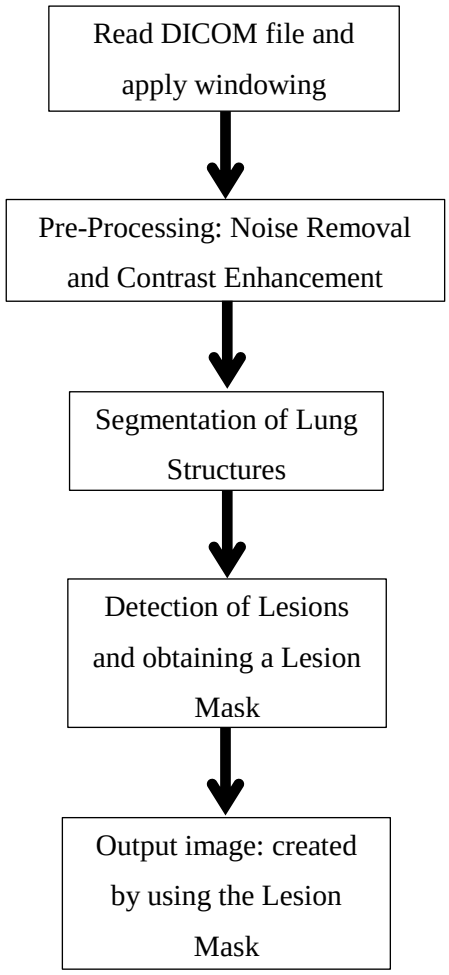


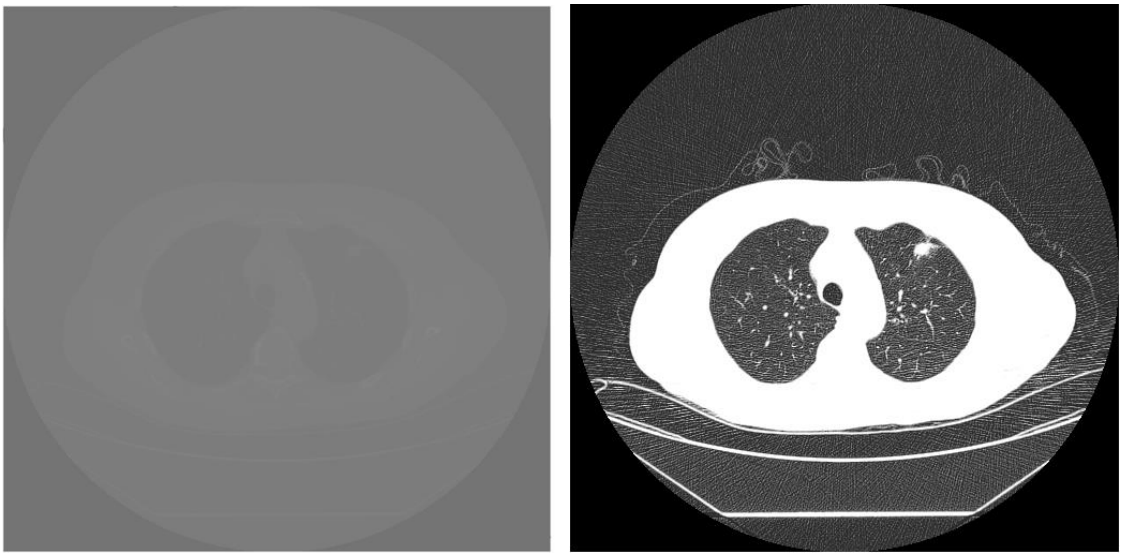

**Figure 3.1:** *General Flow Diagram of the CAD System*

## 3.2 Reading the DICOM Files

The DICOM standard [50] is the RAW file format used in medical imagery. Other than the actual image data, it contains other information such as patient details and scanning parameters. In order to read the image from DICOM files, the DCMTK library [51] is used. The DCMTK library is a popular and free DICOM toolkit which provides an API to read and manipulate DICOM files.

DCMTK also allows the user to apply a windowing function to the DICOM image. The windowing function, specified in the DICOM standard [50], is used by radiologists to ensure that only volumes of interest to them are observed in the image when viewed on a DICOM reader. This is possible since different tissue has different absorption coefficients, but the same tissue has a consistent absorption profile across all CT scan slices for a single individual as well as different individuals. The function saturates pixel intensities, both towards the upper and the lower limits, as well as applies a transformation to the other intensities in the range. It can therefore be thought of as a normalising function, which allows the same structures to be shown with the same intensities across different CT scan images obtained by different CT scanners. The windowing function is made up of two parameters: the window centre and the window width.

Figures 3.2(a) and 3.2(b) show a typical image before and after a windowing function, having a window width of 1000 and a window centre of -700, was applied. In the windowed image, only the lung structures are shown, with the bones removed to improve readability. Applying the same windowing function to any CT image will give the same result.



(a) (b)

**Figure 3.2:** *A typical DICOM image before (a) and after (b) windowing*

Although both the window centre and window width parameters can be found in the DICOM file, it was noted that not all scans have the same windowing parameters specified. However, the volumes not required in this study can be removed by applying a windowing function having a window width of 1000 and a window centre of -700, as seen above. These values were obtained from a DICOM file which had the window width and window centre tags already set to show only the lungs. Since all DICOM images will have the same window applied, thus implying that they will also have the same intensities for the same lung tissue, any values obtained experimentally can be applied to all the DICOM images which are fed to the CAD system.

After applying the windowing function, the image is read from the DICOM file and is converted to an OpenCV Mat [52] object, which is the standard container for OpenCV images. OpenCV version 3.0 is the image processing library used in this study. It can be noted that the image read is a single channel, unsigned, 8-bit grayscale image.

In order to place the images in the correct order, the instance number tag is also read from the DICOM file, which notes the sequence in which the images must be placed. The slice thickness and pixel spacing, that is the resolution of each slice along the *z* axis in mm and the resolution of each pixel along the *x* and *y* axes in mm, were also read. These parameters were used later on in the implementation, such that uniformity between different scans could be achieved.

## 3.3 Noise Removal

In the studies where denoising techniques are compared, comparative measures were required. Generally, values for PSNR and Mean Square Error (MSE) are used to compare between techniques. However, given that a 100% noiseless image cannot be obtained, such measures might not give a true picture.

With that being said, the idea for this project was to implement some of the noise removal techniques mentioned, to try and find a clear indicator as to which performs best. The subjective opinion of an expert in the medical field was used to determine which filtering technique should be applied. The techniques which were implemented are the Bilateral Filter, the Wiener Filter and Anisotropic Diffusion. Mean and Median filters, along with their variations such as the Gaussian Filter and Adaptive Median Filter, were not considered, since they might not give results as good as those of the techniques chosen. This is because they only take into consideration the distance between pixels in the neighbourhood when assigning weights, if any weights are assigned.

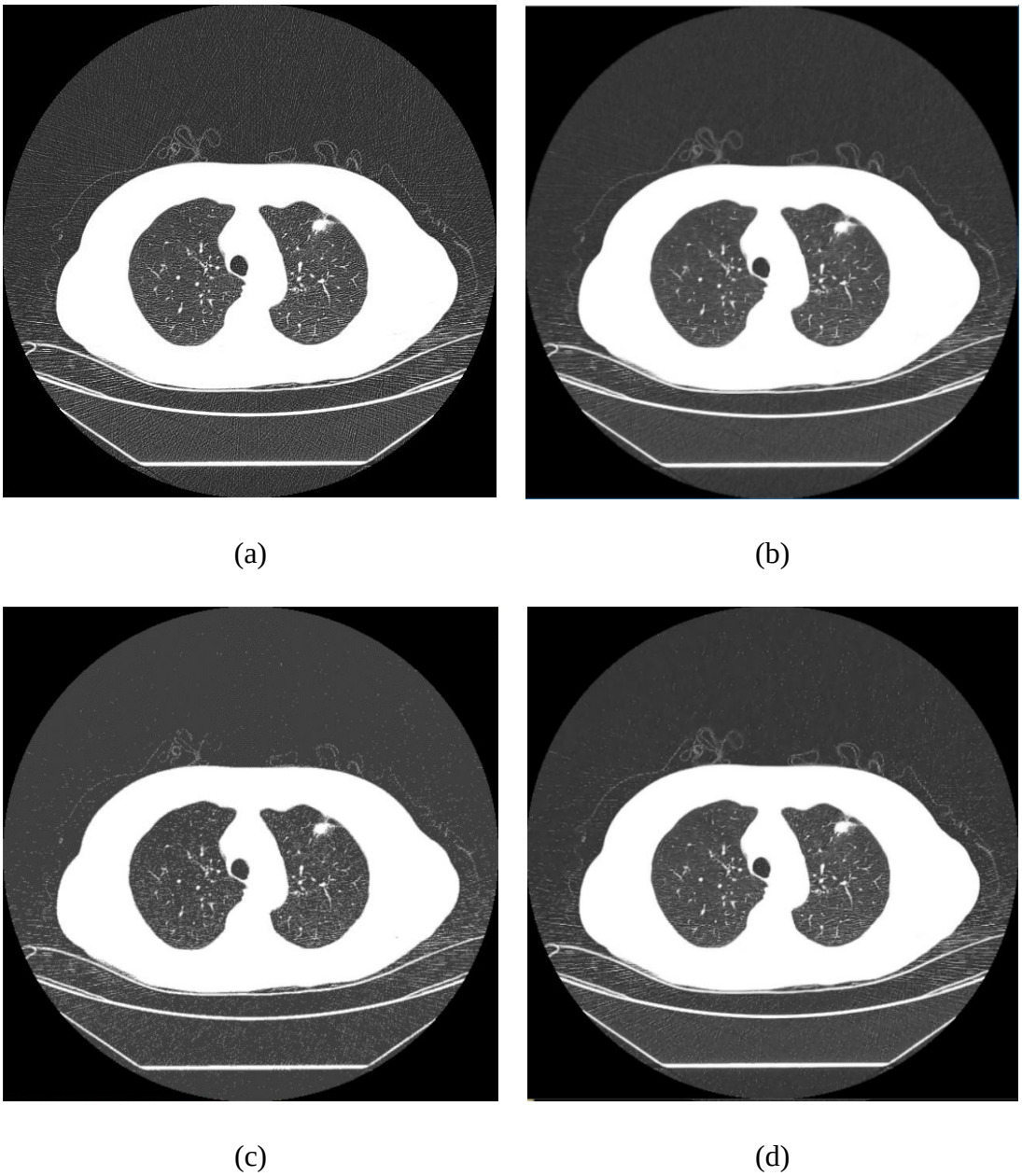
### 3.3.1 Bilateral Filtering

OpenCV provides its own implementation of bilateral filtering. Other than the input image, the other parameters required are the pixel neighbourhood diameter and the standard deviation values for the colour and co-ordinate space. The standard deviations are used to determine the Gaussian functions which will describe the weights for the pixels in the neighbourhoods. Figure 3.3(b) shows an image after it was filtered using a bilateral filter.

### 3.3.2 Wiener Filtering

To apply Wiener filtering, the input image must first be converted into the frequency domain. To do this, the Discrete Fourier Transform function provided by OpenCV is used to obtain a complex output. After modifying equation (2.1) such that the degradation function is set to 1, it can be noted that the result is the multiplication of the image's frequency domain by some constant which is proportional to the inverse of the signal-to-noise ratio of the image. Having said that, since a noiseless image is not available, this constant is determined experimentally. It can be noted that this value must be less than 1 in order to reduce noise, since the signal-to-noise ratio must be large if the image can be recovered. Once the multiplication is done, the Inverse Discrete Fourier

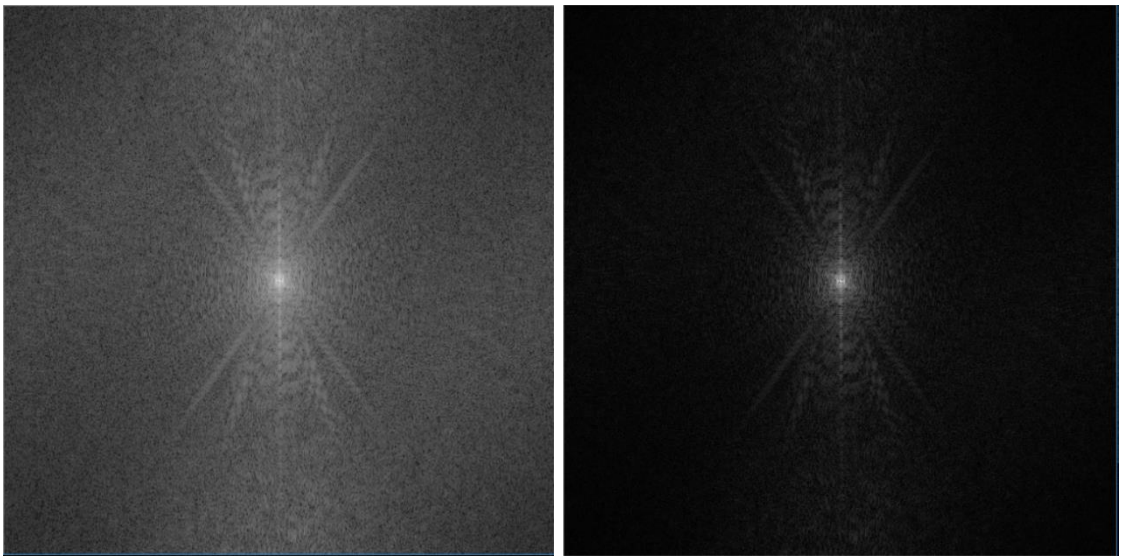
Transform function implemented in OpenCV is used to obtain a normalised real output, since an image is not made up of complex numbers. Figure 3.3(c) shows an image after it was filtered with a Wiener Filter. Figures 3.4(a) and 3.4(b) show the frequency domain of the images in Figures 3.3(a) and 3.3(c) respectively. Note how the high frequency content has been heavily attenuated due to the fact that the values obtained for high frequencies are very small and only 8 bits are used to represent the data, meaning they will be quantised to a very small value.



(a) (b)

(c) (d)

**Figure 3.3:** *Resulting images after filtering; (a) original image, (b) applying a Bilateral filter having a neighbourhood radius of 5, colour standard deviation of 40 and co-ordinate space standard deviation of 70, (c) using a Weiner filter having constant (inverse PSNR) of 0.018, and (d) utilising Anisotropic Diffusion with lambda set to 0.25, k_ns and k_we set to 45, and 10 iterations.*



(a) (b)

**Figure 3.4:** *The frequency response of the image in (a) Figure 3.3(a) and (b) Figure 3.3(c)*

### 3.3.3 Anisotropic Diffusion

The final technique to be implemented is Anisotropic Diffusion. To perform Anisotropic Diffusion, 3×3 edge detecting kernels for the North, South, East and West directions need to be specified. These were created as specified in equation (8) in [25]. The kernels created are similar to the Laplacian edge detection function along each direction. For reference, the kernel for the North direction is as follows: $N = \begin{pmatrix} 0 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

Anisotropic Diffusion runs for a number of iterations, and updates the output image after each iteration, based on some conductance function, as shown in equation (7) in [25]. The conductance function determines how much blurring should occur in the image. The conductance function used in this study is the one found in equation (11) in [53], which is a modified version of one of the conductance functions provided by Perona and Malik in [25]. Tsiotsios *et al* [53] indicate that this function performs better than the other functions which were tested. In this function, the parameter *K* is the gradient threshold parameter, which defines the threshold which separates edges from noise. As in [53], two gradient threshold parameters are used: one for edges in the North or South direction and another for edges in the West or East direction. Another parameter present in the equation which updates the output after each iteration is λ. λ determines how much each iteration affects the result and must be some value between 0 and 0.25. The number of

iterations required, the values for lambda and the two gradient threshold parameters were determined experimentally. Figure 3.3(d) shows the image shown in Figure 3.3(a) after being filtered using Anisotropic Diffusion.

### 3.3.4 Choosing the Noise Reduction Method and its Parameters

Choosing between the different methods implemented and finding the optimal parameters is not an easy task. Considering the fact that noise must be removed but anatomical structures must be maintained, the help of an expert radiologist was employed. Tables 3.1, 3.2 and 3.3 show the range of parameter values used for every technique. All the possible combinations of these values were tested. The filters were employed on different images from different scans, to obtain the best result possible. The Wiener filter parameter was only varied up to 0.06 since any values after this were found to have no change on the image.

After all the images were observed, the filtering technique of choice is Bilateral Filtering with a neighbourhood of 5 pixels, a colour standard deviation of 40 and a space standard deviation of 70.

**Table 3.1:** *The parameter values considered for Bilateral Filtering*

| Bilateral Filter | | |
|---|---|---|
| **Parameter** | **Range** | **Step** |
| Neighbourhood | 5 - 15 | 2 |
| Space Variance | 20 - 100 | 10 |
| Colour Variance | 20 - 100 | 10 |

**Table 3.2:** *The parameter values considered for Wiener Filtering*

| Wiener Filter | | |
|---|---|---|
| **Parameter** | **Range** | **Step** |
| Multiplication Factor | 0.01 - 0.06 | 0.002 |

**Table 3.3:** *The parameter values considered for Anisotropic Diffusion*

| Anisotropic Diffusion | | |
|---|---|---|
| **Parameter** | **Range** | **Step** |
| λ | 0.15 - 0.25 | 0.05 |
| Iterations | 10 - 100 | 10 |
| K_ns | 35 - 65 | 10 |
| K_we | 35 - 65 | 10 |

## 3.4 Contrast Enhancement

NGCCLAHE [32] adopts the advantages presented by both Gamma Correction and CLAHE, and was specifically tested on CT imagery. Unlike NGCCLAHE, when WHE was tested on CT images [33] it was noted that it produced relatively the same results as AHE but took more time. Thus, it would be wise to consider a variation of AHE that has already been tested on CT imagery and produced notable results. Pizer *et al* [33] also suggest the use of CLAHE on medical imagery. Although CLAHE might be slower than some other implementations of AHE, such as POSHE, it should not produce any artefacts which may hinder further processing. Therefore, the method used in [32] seemed ideal for this study.

The first step in this contrast enhancement technique is Gamma Correction. In order to do this, a particular value of gamma needs to be provided. Since the lungs are shown as grey on a white background, it was decided that they should be darkened. Also, any potential lesions would be white as well, so it made sense to darken the lung tissue. Thus, it was decided that gamma values in the range 1.5 to 4 were to be tested. To perform Gamma Correction, a lookup table was created such that the power operation would only need to be applied 256 times (from 0 to 255) rather than calculated for every pixel. This should aid to reduce the speed needed to perform Gamma Correction. In order to apply the new values to all the pixels in the image, OpenCV's LUT function was used. Once the image has been gamma corrected, OpenCV's CLAHE class was used to perform contrast correction. The required parameters for CLAHE are the clip limit, which aids to prevent enhancement of noise, and the neighbourhood size. The range of values tested for these parameters can be seen in Table 3.4.

**Table 3.4:** *The parameter values considered for Contrast Enhancement*

| Contrast Enhancement | | |
|---|---|---|
| **Parameter** | **Range** | **Step** |
| γ | 1.25 – 4 | 0.25 |
| Clip Limit | 0.5 – 2.5 | 0.5 |
| Neighbourhood | 4 – 10 | 2 |

The chosen parameters were the following: γ of 2, clip limit of 2, neighbourhood of 6. These were also verified by an expert radiologist, to ensure that anatomical details are not lost. Figure 3.5 shows the image of Figure 3.3(b) with contrast enhancement applied.

**Figure 3.5:** *The image of Figure 3.3(b) after contrast enhancement*

## 3.5 Segmentation

Once the image has been pre-processed, it is forwarded to the segmentation phase, such that the lungs can be extracted. It is evident from literature that threshold segmentation techniques are too dependent on the threshold value [35, 38]. Such a value can take a long time to be computed and it must be computed for every CT image in a particular scan [37, 40]. Similarly, Active Contours are dependent on the stopping criteria. It might be difficult to define general stopping criteria that work for all the images in a patient's scan and for all the scans belonging to different patients. As noted in [34], if the stopping criteria are not well defined, poor segmentation results will be achieved. Active Contours can also be quite slow due to their iterative nature. By comparison to the other algorithms mentioned, watershed based segmentation appears to be simple, since it does not require explicit specification of functions to define relationships between pixels, as is the case with Region Merging or Splitting and Graph Cuts. Thus, due to its simple implementation, watershed segmentation was used to extract the lung structures from the CT images. Markers to avoid over segmentation which might result from this technique were also created.

### 3.5.1 Removing the Background

The first step in obtaining the lung regions was to remove the background. To achieve this, the image was blurred, such that no anatomical details could be observed. This was achieved by morphological opening, using a disk with a radius of 10 pixels as a morphological structure. This resulted in the image shown in Figure 3.6(b). This was done so that any residual noise, which was not removed in the Noise Removal phase, would not be considered in the following few steps. Also, this gave the main regions of the image a more uniform intensity, which is helpful when removing the background.

Once that was done, the area around the chest cavity was removed. It can be noted that the background is dark compared to the chest cavity, which is white. Thus, to remove this area, the following technique similar to region growing was developed. A column wise search was started, from the top left corner of the image. If a pixel had less than *counter_threshold* pixels in *search_area* pixels below it, having intensity higher than *intensity_threshold*, then that pixel was changed to white (intensity value 255). It was found experimentally that *counter_threshold* should be set to 3 pixels, *search_area* should be set to 5 pixels and *intensity_threshold* should be 200. Therefore, if a pixel had less than 3 out of 5 pixels below it with an intensity less than 200, it was set to white. Once a pixel having 3 or more pixels out of the 5 below it with an intensity larger than 200 was found, the procedure stops along that column and moves to the adjacent column. Therefore, upon detection of a white region, which is the chest cavity, by a majority vote, the deletion of the background stops. The process was repeated starting from the bottom left corner as well. This resulted in the image shown in Figure 3.6(c). It was decided that the search should be done column wise since the largest contouring of the image is along the width. Thus, performing a column wise search gave better results.

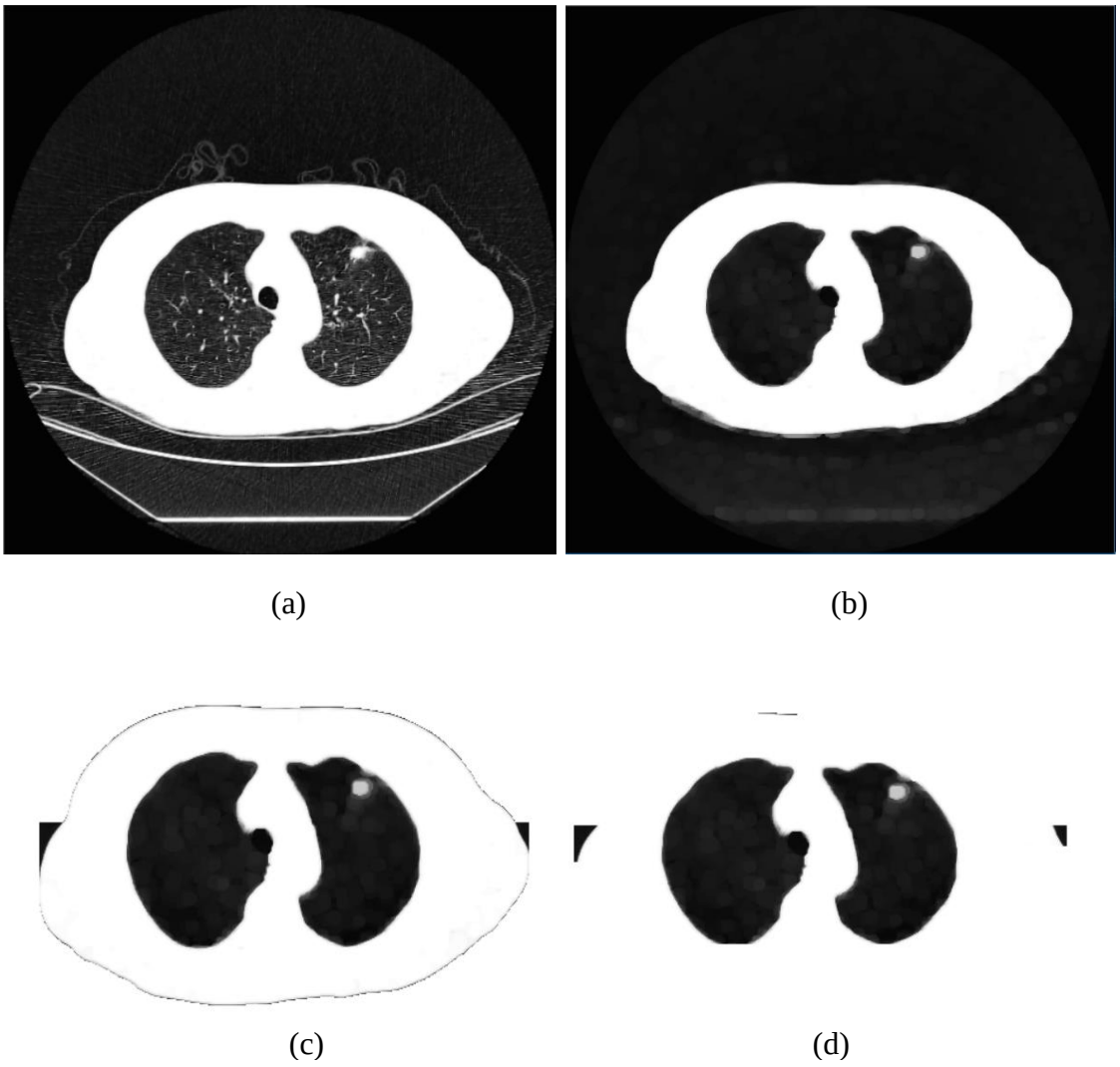
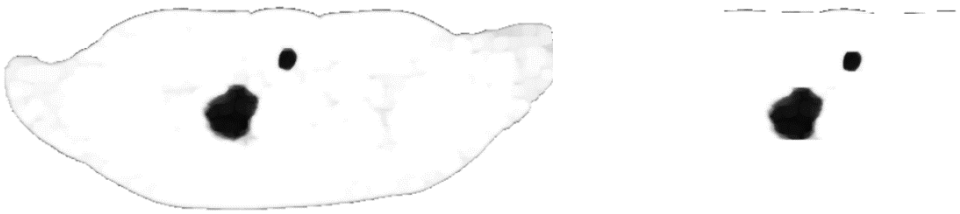



**Figure 3.6:** *Background removal showing (a) the contrast enhanced image, (b) the image after morphological opening using a disc radius of 10, (c) outside background removal and (d) final result.*

The image now contained the lungs, which have a dark intensity, on top of a mostly white background. It was also observed that there was a thin black line around the chest cavity and, in some images, some residual noise was left inside the chest cavity, as can be seen in Figure 3.7(a). Thus, to obtain a better image, a procedure similar to the one above was utilised to remove the faint line and the residual noise as much as possible. This time, however, the search was performed row wise, since the largest contour is presented by the lungs, that is, along the length of the image. The *counter_threshold* was set to 3 pixels, the *search_area* to 50 pixels and *intensity_threshold* to 70. These values were obtained experimentally. This resulted in the image in Figure 3.7(b). A low intensity threshold was used since the logic is reversed: pixels less than the *intensity_threshold* increment the counter. Since the lungs are dark, a low threshold is required. The low counter and large search area ensure that the lungs are not modified. Figure 3.6(d) shows the same image in Figure 3.6(c) after it undergoes the same procedure.



(a) (b)

**Figure 3.7:** *Lung image (a) after removal of the background around the chest cavity, and (b) after the second part of the background removal*

### 3.5.2 Extracting the Lungs

With most of the background clearly defined as white, the markers for the watershed algorithm were needed. In the marker mask, three values represent three different regions. Pixels which are sure foreground pixels are set to white, pixels which are sure background pixels are set to black, whilst pixels which are undetermined are set to grey.

#### *3.5.2.1 Obtaining the Foreground Mask*

To obtain the foreground mask, morphological dilation was performed on the output of the previous steps with a disc of radius 25 pixels. Dilation achieved three purposes:

1. It removed any parts of the remaining faint line along the chest cavity.
2. It reduced the area occupied by the lungs in the image, since it shrinks dark regions.
3. In some images, it helped to clearly separate the trachea from the lungs or removed it altogether. The trachea appears as a circular structure outside of the lungs.

The resulting image can be observed in Figure 3.8(a). This image is then inverted and a threshold of 0 is applied, such that any value greater than 0 is transformed to white. This resulted in a binary image with the white parts representing the sure foreground pixels, as observed in Figure 3.8(b). The threshold value of 0 is utilised since any values larger than 0 must be foreground pixels, due to the previous dilation.

(a) (b)

**Figure 3.8:** *The resulting image after dilation (a) and the binary image after thresholding (b).*

In some images, where the lung lesion is quite large and well defined, it was observed that the lesion was not included in this mask, resulting in a hole in the image. This was due to the fact that the blurring achieved was not enough to obscure it. Thus, in order to ensure that such parts of the lung were included, OpenCV's findContours function was used on this binary image to obtain a list of the external points which make up the white region. Using OpenCV's drawContours function, the area enclosed by these points was set to white. Thus, any holes in the mask were filled. An example of this process can be seen in Figures 3.9(a) and (b).

(a) (b)

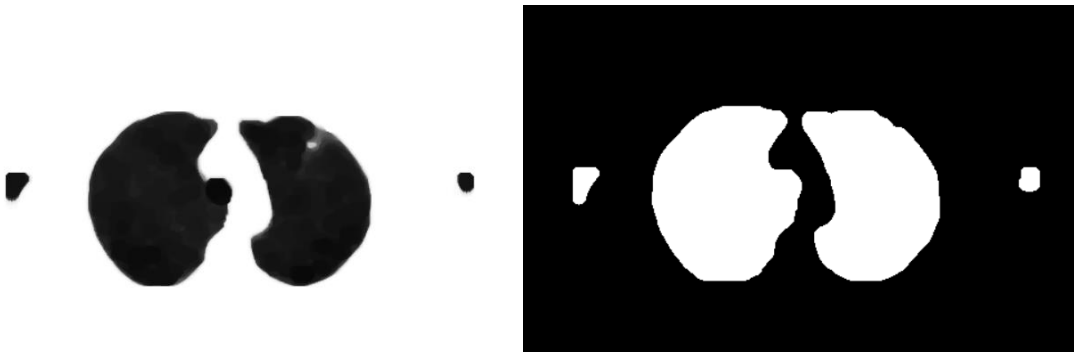
**Figure 3.9:** *Mask containing lesion picked up as background (a) and the corrected binary image (b).*

#### *3.5.2.2 Obtaining the Uncertain Pixels Mask*

The next step was to get a mask of pixels that might be part of the foreground. To do this, a similar procedure to obtaining the foreground mask was utilised. At first, the same dilation took place, but in this case, a smaller 7 pixel radius disk was utilised. This ensured that less of the lung region was removed, compared to the previous procedure, whilst still removing the black line along the chest cavity. Once this border was removed, erosion with a disk having a radius of 12 pixels was performed. This resulted in a larger lung region than in the original image, as seen in Figure 3.10(a), since erosion enlarges dark regions. Thus, the region contains the lungs, the lung edges, as well as part of the background.

Finally, the image required is a binary mask, where white represents potential foreground pixels and black represents sure background pixels. To achieve this, the same procedure as with the foreground mask was utilised, that is, the image in Figure 3.10(a) was inverted and a threshold was applied to obtain a binary image. The threshold was

once again set to 0. Filling, using the same procedure as before, was also used. The mask obtained can be seen in Figure 3.10(b).



(a) (b)

**Figure 3.10:** *Image after dilation and erosion (a) and the thresholded binary mask including pixels that may be foreground (b).*

#### *3.5.2.3 Obtaining the Final Mask*

The two masks described above were combined to produce a single mask which was fed to the watershed algorithm. The single mask was obtained by averaging out the two masks: any pixels which were white in both masks must remain as such, since they represent sure foreground pixels. Any pixels which are black in both masks must be set to grey (128), since they are sure background pixels. On the other hand, pixels which were white in one mask but black in the other are unsure pixels and must therefore be set to black (0). An example of this single mask with all the foreground markers, background markers and unsure markers can be seen in Figure 3.11(a).

This mask was then used in the watershed algorithm which was performed utilising OpenCV's watershed function. The watershed was run on a copy of the pre-processed image acquired from the previous phase in the CAD system. The final mask obtained by the watershed algorithm can be seen in Figure 3.11(b). This mask was then smoothed using morphological closing with a disc of radius 25 pixels, resulting in the image in Figure 3.11(c). After this morphological operation, a threshold of 200 was applied to obtain a binary image, since the watershed algorithm sets the background to grey. This final mask defines the lung structure that is to be segmented. An example of the mask

can be seen in Figure 3.11(d) whist Figure 3.11(e) shows the mask applied to the pre-processed image.

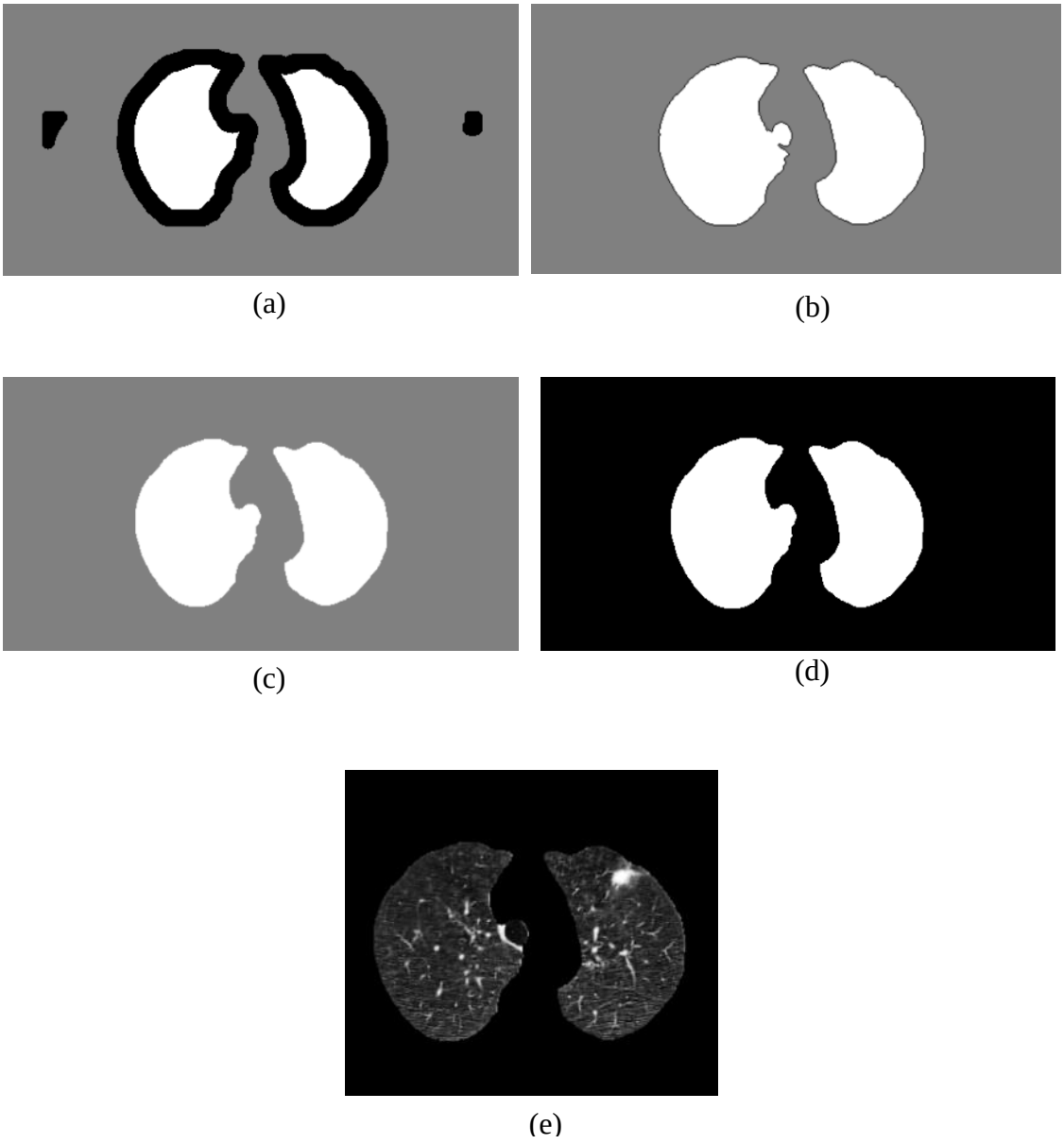
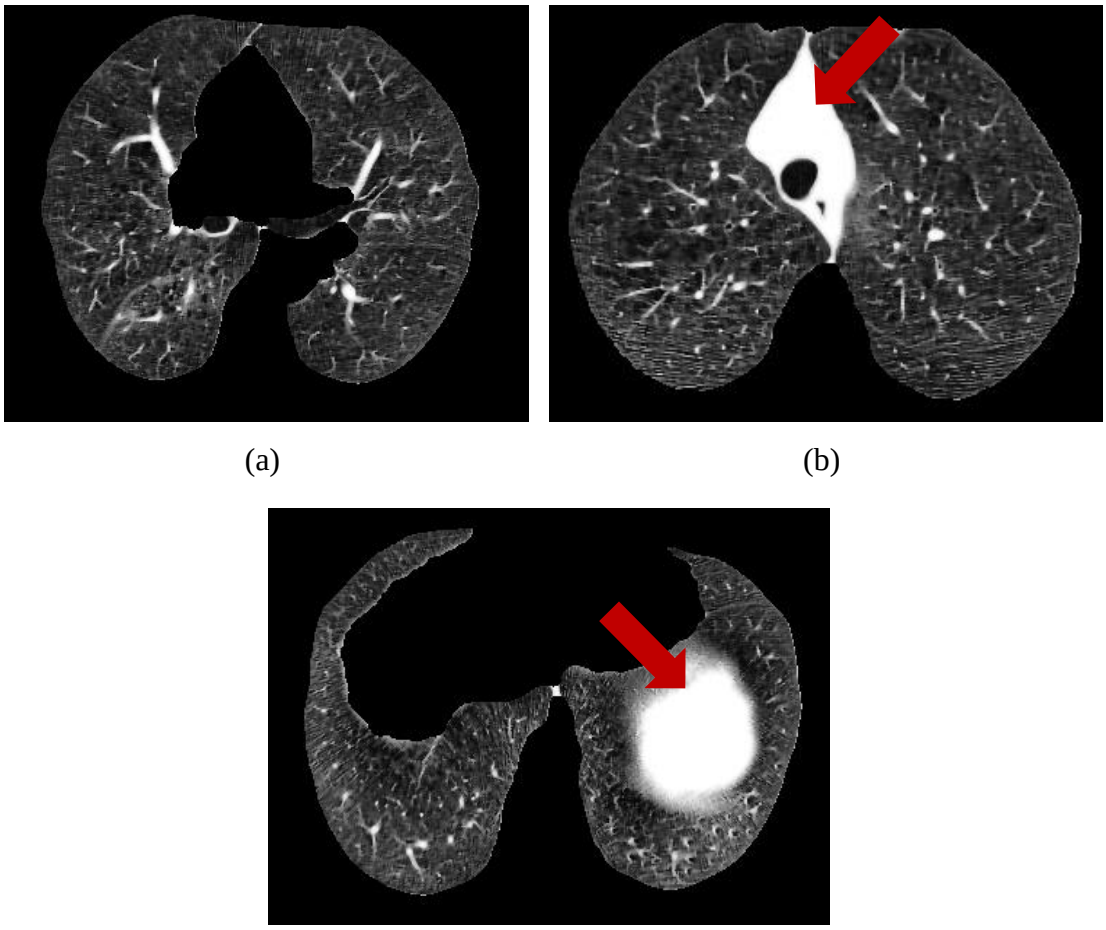



**Figure 3.11:** *The final segmentation process; (a) the final markers used by the watershed algorithm, (b) the mask obtained from the watershed algorithm, (c) smoothed mask following morphological closing, (d) the binary mask, and (e) the extracted lungs.*

Note that the heavy dilation performed in the beginning, when obtaining the foreground markers, ensured that the two lung structures were not connected together when creating the foreground markers mask, even if they are connected together in the CT scan. Any connections between the lungs are discovered by the watershed algorithm, rather than enforced in this mask. An example of a pair of correctly segmented, touching lungs can

be observed in Figure 3.12(a). This is done on purpose, since, given that a hole filling technique is used when obtaining the markers, the internal region between both lungs would have been included in the sure foreground pixel mask, had the two lungs been connected. This would have resulted in an erroneous segmentation output, similar to that in Figure 3.12(b), where the extra background region is pointed to by an arrow.

Having said that, for images where the lung is moving out of the field of view, the hole filling technique included large areas of the background, in some cases. An example of such an error is indicated by an arrow in Figure 3.12(c). These regions, which are generally very large, were dealt with in the detection phase.



(a) (b)

(c)

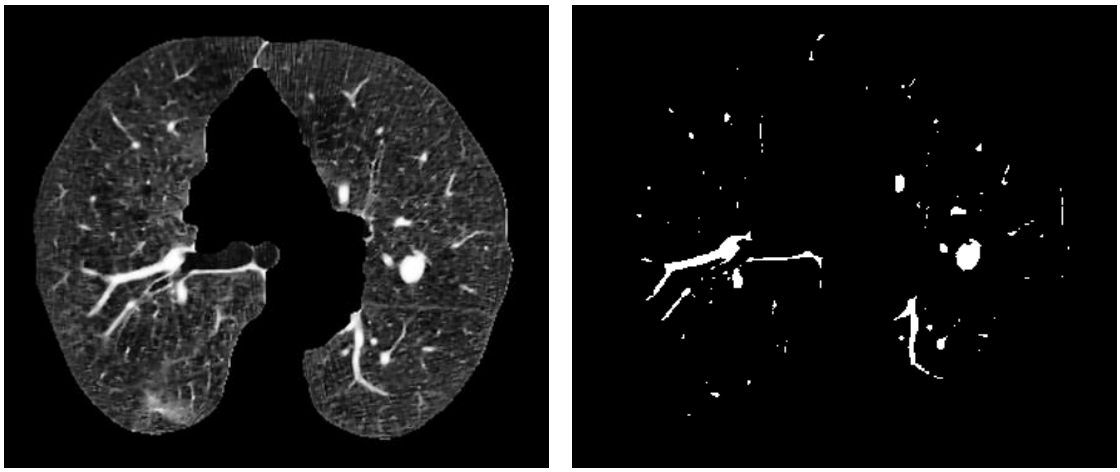
**Figure 3.12:** *Further segmentation scenarios: (a) Two correctly segmented, connected lungs; (b) background region between the lungs incorrectly included in the output; (c) white region wrongly included by the hole filling algorithm.*

## 3.6 Detection

After obtaining the lung masks, detection must take place to find any lesions. The methods employed by Choi *et al* [48] and Ashwin *et al* [21] require training to learn parameters. However, unlike in the fuzzy k-nearest neighbours method used by Namin *et al* [43], these algorithms can be over trained, such that they over fit their parameters to the training data. Hence, these methods were not used. A fuzzy k-nearest neighbour classifier is used by Namin *et al* [43]. However, their aim was not only to detect lesions, but also to classify their malignancy. This is not the aim of this study and therefore, the only features that were considered were the intensity and shape, which should be sufficient in detecting lesions. This also means that the fuzzy k-nearest neighbours classifier was not needed.

### 3.6.1 Obtaining an Intensity Mask

One of the main features of lesions is the fact that they have a high intensity value when compared to other lung tissue. Therefore, an intensity mask is created to indicate any potential lesion candidate locations. To obtain this intensity mask, an intensity threshold was applied to the segmented lung regions. The threshold was chosen experimentally and was found to be 185. The value was chosen with the thought that the edges of a lesion might not be very well defined, especially after the initial noise removal. Therefore, edges might have a lower grey level than the centre of the lesion. Given that all the lung slices had the same window applied, this threshold is valid for all lung slices. Also, given the previous contrast enhancement performed, each potential lesion was distinct from the surrounding tissue. An example of a typical intensity mask can be seen in Figure 3.13(b), with Figure 3.13(a) depicting the segmented image from which the intensity mask was derived.



(a) (b)

**Figure 3.13:** *The segmented image (a) and the intensity mask (b).*

### 3.6.2 False Positive Reduction

Having said that, as observed in Figure 3.13(b), the mask is full of false positives, resulting from imperfectly segmented lungs and lung vessels. With that in mind, another criterion needed to be added to reduce the false positives.

#### *3.6.2.1 Shape Index*

Initially, a shape index was calculated for each pixel in the intensity mask. The shape index was calculated by using a method similar to the works of Antonelli *et al* [42] and Namin *et al* [43], that is, by obtaining the 2nd order derivatives of each image along the *x*, *y* and *z* directions. The *x* and *y* derivatives were obtained by using the inbuilt Scharr function in OpenCV which, according to the OpenCV documentation [54], provides better results than the OpenCV Sobel function for a 3×3 kernel. A 3×3 kernel was used since local image information should define the shape index. The smaller the kernel, the more importance is given to local information. Thus, using a larger kernel reduced performance. OpenCV does not natively handle filtering along the *z* direction. Therefore, to work around this problem, a function was developed to switch the *x* and *z* axes, as shown in Figure 3.14. Edge detection along the *z* axis was then performed as though it were the *x* axis using OpenCV's Scharr function. Once this was done, the axes were reverted back.


**Figure 3.14:** *The transformation performed on the axes.*

A Hessian matrix was then created using the 2nd order derivatives as follows: $H = \begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix}$, where $I_{ab}$ is the 2nd derivative of the image $I$ first along the $a$ direction and then along the $b$ direction. Its eigenvalues were used as parameters in the equation for the shape index:

$$Shape\ Index = 0.5 - \frac{1}{\pi}\tan^{-1}\frac{k_1(p) + k_2(p)}{k_1(p) - k_2(p)} \tag{3.1}$$

where $k_1(p)$ and $k_2(p)$ represent the principle curvatures and are equal to the largest and smallest eigenvalues of the Hessian matrix [42, 43]. The eigenvalues were obtained by using the Eigen library [55]. The resultant mask was then thresholded with a threshold of 0.8 to obtain the lesions. The threshold value was obtained from the works of Antonelli *et al* [42] and Namin *et al* [43]. An example of a thresholded shape index mask is shown in Figure 3.15. This is the shape index mask obtained from the image in Figure 3.13(a).

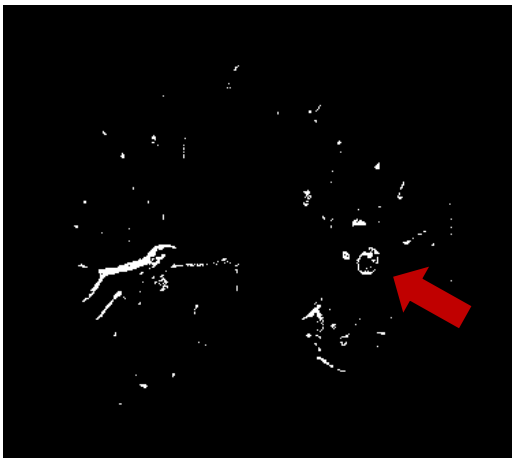


**Figure 3.15:** *A thresholded shape index mask. The arrow points towards a lesion. Note how a lot of false positives still exist and the lesion region is disjoint, making further false positive reduction difficult*

However, upon thresholding the resultant mask, it was noted that too many false positives were still present and that the points are somewhat spurious, that is, the regions pertaining to a lesion are not always very well connected. In an attempt to mitigate this, the threshold was varied but similar results still emerged. Trying to reconnect regions back together using morphology resulted in false positives also being connected together, which meant reducing them further based on the size of a region was not possible. For these reasons, the shape index parameter was not used.

#### *3.6.2.2 False Positive Reduction along the XY Plane*

In order to reduce the number of potential lesion candidates, the physical size of the potential lesions was considered. Lesions larger than 5mm are considered in this study, since research shows that lesions less than that size are likely to be benign [56]. In fact, as per the recommendations of the Fleischner Society, an international multidisciplinary medical society for thoracic radiology, lesions having sizes between 4-6mm require a follow up in 12 months in low risk patients and a follow up in 6-12 months in high risk patients. This is compared to the 3 month follow up required for lesions larger than 8mm [56]. With this 5mm minimum diameter in mind, it is known that, for a candidate to be an actual lesion, it must have a diameter equal to, or larger than, a fixed number of pixels along all axes.

Since different scans have different resolutions along both the *x* and *y* directions, the resolutions for each axis were read when reading the DICOM file. The number of pixels that must be in the length and width of each region along the *x* and *y* directions is given by:

$$Number\ of\ Pixels = \frac{5\text{mm}}{Resolution_{Axis}} \tag{3.2}$$

By iterating through each region in an image and taking the minimum and maximum *x* and *y* co-ordinate values, the length and width of a bounding box around each region can be determined. Using the values obtained by equation (3.2) as the minimum width and length, small regions are removed.

However, a two pixel delta is allowed along the length and width of a region. This is because of two main reasons. Firstly, although the threshold value chosen for the intensity mask should take into consideration the grey level of the edges, some pixels

pertaining to edges, might still not have been included in the intensity mask. Furthermore, if a lesion's diameter is 5mm, it does not mean that in all the slices in which the lesion is present it will appear with a 5mm diameter. This is because, since the lesion is spherical, slices which may contain only the topmost or bottommost parts of the lesion will have a region with a smaller diameter. Figure 3.16 demonstrates this fact.

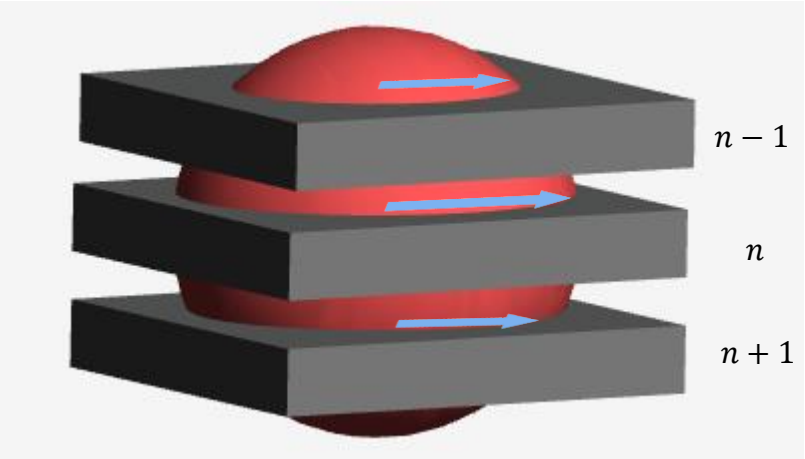



**Figure 3.16:** *Demonstration on how the radius of the lesion varies with slice number.*

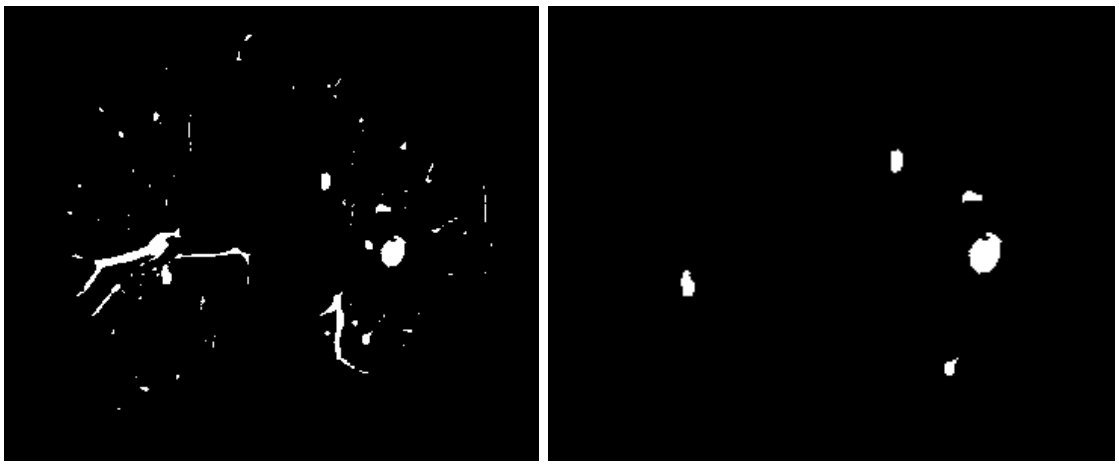
However, it was noted that vessels such as arteries, might form elliptical regions, which might still have a large enough width and height by this method, and thus, they might not be removed by this threshold. To remove these false positives, the area of the region inside the bounding box created must be checked.

Consider a candidate region in the binary image having a bounding box of width *w* and height *h* pixels. The area of the box is given by multiplying both these parameters. Consider now a circle having a radius *r*. Its area is given by $\pi r^2$. If a perfect box is to be drawn around the circle, it will result in a square, having both *w* and *h* equal to 2*r*, with an overall area of $4r^2$. The ratio of the area of the circle and its bounding box, is therefore given by:

$$Ratio = \frac{\pi r^2}{4r^2} = \frac{\pi}{4} \approx 0.785 \tag{3.3}$$

Thus, this value can be used to threshold the ratio of the area occupied by the regions to the area of their bounding box in order to remove most non-spherical regions. In reality,

the threshold used is 0.7, to allow for some error, since lesions might not present themselves as perfect circles. This results in the mask shown in Figure 3.17(b). The removal of such elliptical regions is highly dependent on the angle which they form with the *x* and *y* axes: if they are horizontal or vertical, this technique will not work. An example of such a case in shown in Figure 3.18. It is also dependent on the thickness of the vessel: the thicker it is, the more area it will cover.



(a) (b)

**Figure 3.17:** *The result of applying false positive reduction along the xy-plane (b) of the intensity mask (a).*

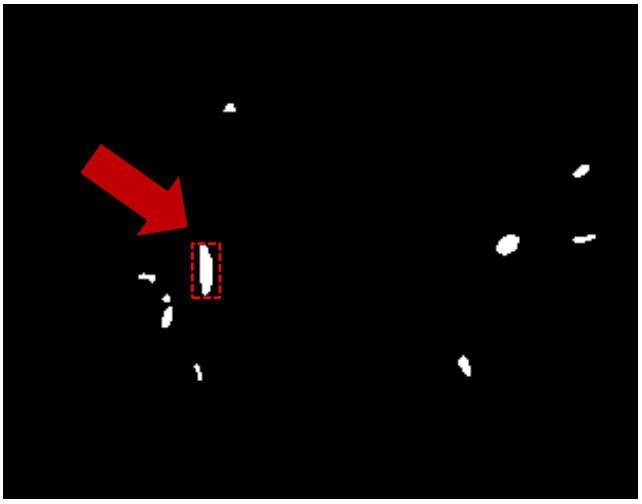


**Figure 3.18:** *An example of a vessel, denoted by an arrow, that has an elliptical shape but is not removed by false positive reduction in the xy plane, since the vessel is parallel to the y axis.*

Recall that in the previous section, when segmentation was performed, the extracted regions might end up a large white part of the background in them, as per Figure 3.12(c). This will also be considered here as a candidate lesion. In order to remove such regions, a limit is created on the number of pixels that can constitute to a lesion's region, the limit being 2700 pixels. This value is realistically large enough to cater for lesions of different sizes which may be detected by a CAD system.

#### *3.6.2.3 False Positive Reduction along the Z Axis*

For the *z* axis, equation (3.2) can be used to obtain the minimum height, that is, the minimum number of slices that the region must be present in, to be considered as a lesion. The resolution used is equal to the slice thickness. However, equation (3.2) must be modified such that for different regions of different lengths and widths, a different amount of slices are considered. Assume a scan is performed with a slice thickness of 2.5mm: if a lesion has a diameter of 5mm, then it must be present in at least two slices, whereas if a lesion has a diameter of 10mm, it must be present in at least four. The minimum between the length and width of the region is used in the numerator of the equation, such that the minimal number of slices is checked. Thus, equation (3.2) is modified as follows:

$$Slices = \frac{\text{Min}(Length_{Region}, Width_{Region})}{Slice\ Thickness} \tag{3.4}$$

From then, a similar procedure as with the *x* and *y* false positive reduction is used: assume that the minimum number of slices evaluates to *s*. Half of this amount, that is, $\frac{s}{2}$ slices above the region are checked and a binary AND is performed between these slices, since the region must be present in <u>all</u> slices. Similarly, a binary AND is performed on $\frac{s}{2}$ slices below the region. $\frac{s}{2}$ is used since the region under consideration might be in the middle of a lesion. Assume the result of the $\frac{s}{2}$ top slices is stored in $R_{top}$ and the result of the $\frac{s}{2}$ bottom slices is stored in $R_{bot}$. A binary OR is performed between $R_{top}$ and $R_{bot}$, since the region under consideration might not have $\frac{s}{2}$ slices in which the lesion is present on top of it, but it might have them below it. This would be the case if the region under consideration would be at the topmost or bottommost slice of a lesion.

By the above procedure, connectivity between regions along the *z* direction was ensured. However, it might be the case that two false positives might overlap by just a few pixels, which would be enough to ensure connectivity. Having said that, it can be safely assumed that regions forming a lesion must have high connectivity between them, the ideal case being when they lie exactly on top of each other. The region is only kept if 85% of the pixels between slices overlap. Otherwise, it is most probably the result of overlapping false positives. Assume that the result of the previous binary OR between $R_{top}$ and $R_{bot}$ is stored in $R_{total}$. Note that the condition is that either the region in the current slice has 85% of its pixels overlapping with $R_{total}$ or $R_{total}$ has 85% of its pixels overlapping with the region in the current slice. Thus, it is not necessary that both have 85% overlap. This was done so that the condition takes into consideration the fact that one of the regions might have a much lower count than the other region. For example, if the region in $R_{total}$ has 500 pixels, whereas the current region only has 300 pixels, yet they overlap fully, it would not make sense to remove the regions because the threshold was not met for both regions. Given that for one region it would be 60% connectivity (<85%), but for the other it is 100% connectivity (>85%), the regions will be kept. Figure 3.19 shows the final lesion mask obtained after applying this false positive reduction on the image in Figure 3.17(b).

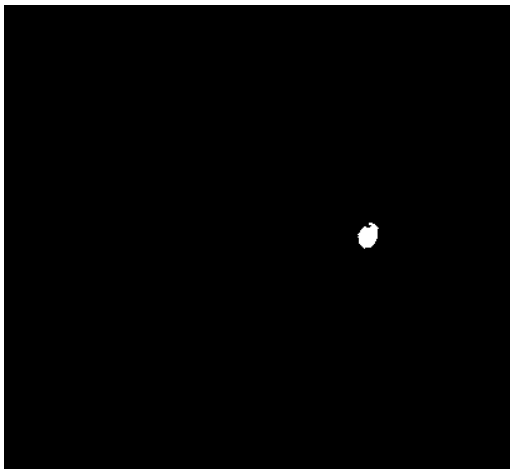


**Figure 3.19:** *The final lesion mask obtained by reducing false positives along the z axis.*

Note also that the order of operations here is important. First, reduction along the *xy* plane was performed, before moving to reduction along the *z* axis. This order must be maintained since, whereas *xy* reduction depends only on local image properties, reduction along the *z* axis is dependent on the information present in more than one

image. Had the order been reversed, it might be the case that there is an overlap along the *z* axis with a region that would be removed later on in the *xy* reduction, resulting in a larger number of false positives. This situation was avoided by reducing false positives in this order.

## 3.7 Output

Once the lesion mask is obtained, the output must be set so that a marker is added to the place where lesions have been detected. To do so, the original image was first converted to a 3 channel RGB image. The lesion mask was then iterated such that, every pixel that was set to white in the mask, and was therefore considered to be a lesion by the CAD system, had its corresponding pixel in the original image changed, to denote a marker. The modification increases the intensity of the red channel, so that a marker appeared over the area where the lesion was detected. An example of the output can be seen in Figure 3.20. This simple indicator should be of use to radiologists when trying to identify lesions, since it clearly distinguishes them from other tissue. The images were then saved in .PNG format.

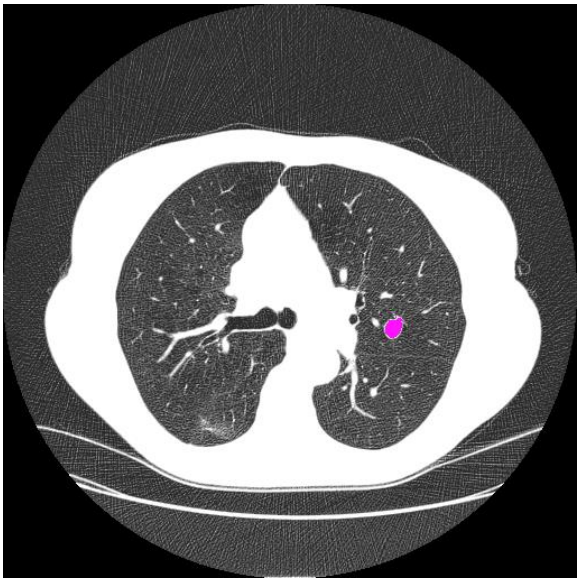


**Figure 3.20:** *Sample output for the image in Figure 3.13(a) with the detected lesion marked*

# Chapter 4 : Testing and Evaluation

Testing was performed on 25 Chest CT Scans. Permission was obtained from a general hospital in Malta to obtain a set of anonymised thoracic CT Scans and their accompanying medical report. The medical report was used as a baseline for the output of the CAD system developed. Out of the CT Scans provided, five were eligible for this study. To increase the dataset, additional CT scan data was obtained from the Lung Image Database Consortium Database [57], which is a public online database. Each scan downloaded from the database has an XML file attached which contains details about the scan. This XML file was used to evaluate the output of the CAD system.

## 4.1 Results

The 25 scans under test contained a total of 59 lesions of different shapes and sizes and in different locations in the lung. Out of these 59 lesions, 45 were successfully detected. Using the equation:

$$Sensitivity = \frac{True\ Positives}{True\ Positives + False\ Negatives} \qquad (4.1)$$

the sensitivity of the system (or the true positive detection rate) works out to be 76.27%, with an average of 0.56 false negatives per scan. A true positive is a lesion which was detected as such by the CAD system, whereas a false negative is a lesion which was not successfully detected. A false positive refers to any anatomical structure which is not a lesion, but was mistakenly detected as a lesion by the CAD system. The average amount of false positives is 7.88 lesions per scan. These results can be observed in the confusion matrix below.

**Table 4.1:** *Confusion Matrix showing the Results*

| | | **Actual** | |
|---|---|---|---|
| | | **Lesion** | **Not a Lesion** |
| **Predicted** | **Lesion** | 45 | 197 |
| | **Not a Lesion** | 14 | / |

Figure 4.1 shows a typical output from the CAD system. The reality is that a total of three lesions are present in this test case. The CAD system marks three regions as being lesions. By analysing the output, it can be observed that two of the marked regions are true positives whereas the other is a false positive.

Of the two true positives detected, Figures 4.1(b) and 4.1(c) show the smaller detected lesion, whilst Figures 4.1(d), 4.1(e), 4.1(f) and 4.1(g) show the larger detected lesion. In the same scan, a 3mm lesion is also observed, indicated by red arrows in Figures 4.1(b), 4.1(c) and 4.1(d). Given that this size is below the threshold, it was not detected by the CAD system. The vessel wrongly detected by the system as a lesion, that is, the false positive, is denoted by a blue arrow in Figures 4.1(d) and 4.1(e). In fact, its continuation can be observed in the other figures, meaning that it is not a lesion.

The time taken for the CAD system to complete each test, including file input and output, was also recorded. It was observed that all the scans under test had images which were 512 pixels wide and 512 pixels long. CT scans with images having different resolutions were not available. The number of slices for each scan differed. As expected, the time taken for the algorithm to run is dependent on the number of slices that the CT scan contained.

The smallest number of slices of all the CT scans under test was 99 slices, whilst the largest was 266 slices. The CAD system took approximately 6 seconds to run on a CT scan having 99 slices, whereas it takes approximately 21 seconds to run on a CT scan having 266 slices. The average of all 25 tests having a different amount of slices is 74.3 milliseconds per slice. All tests were performed using an Intel Core i7-6700HQ CPU, clocked at 2.60GHz.

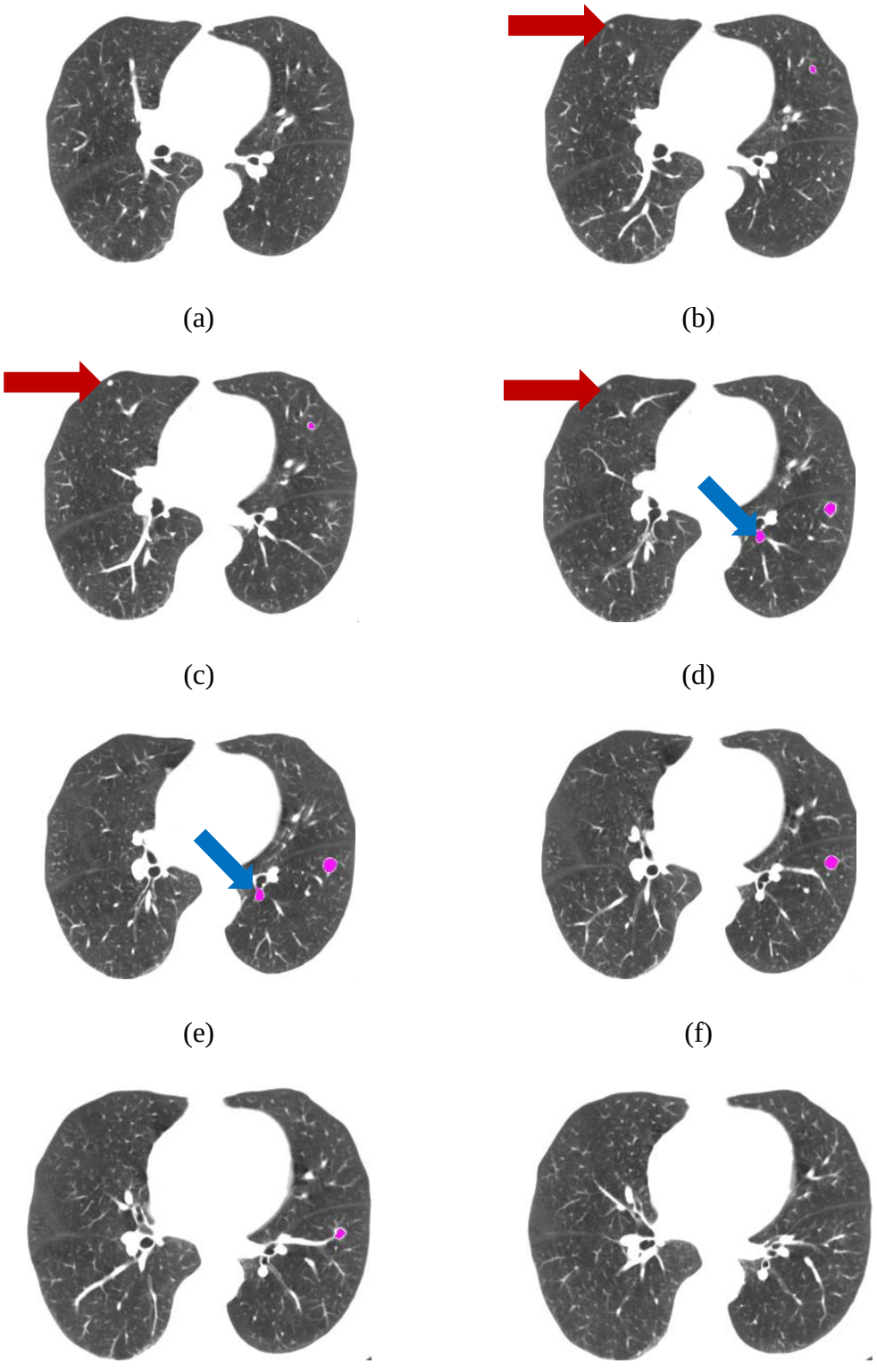



**Figure 4.1:** *Typical Output of the CAD system.*

## 4.2 Comparison with other CAD Systems

The table below shows the results of some other CAD systems found in literature. Comparing such systems can be quite difficult, especially since the dataset considered is generally different.

It can be noted that Namin *et al* [43] managed to achieve a higher sensitivity rate, at the expense of increasing the number of false positives per scan compared to the work in this project. The increase in the sensitivity rate may be due to the larger amount of features considered.

However, Tong *et al* [46], achieved higher sensitivity and a low false positive rate compared to the CAD system implemented in this study. This might be due to the fact that vessels were segmented and removed before detection.

**Table 4.2:** *Comparison between this work and other CAD systems in literature*

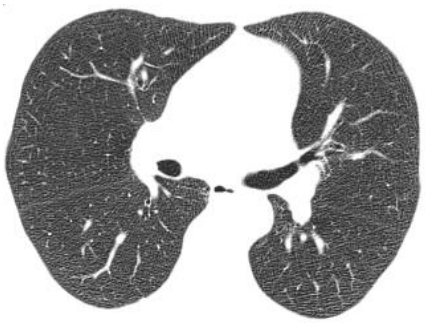
| | Sensitivity | False Positives |
|---|---|---|
| **Namin *et al* [43]** | 88% | 10.3 per scan |
| **Tong *et al* [46]** | 85% | 2 per scan |
| **This work** | 76.30% | 7.88 per scan |

## 4.3 Observations

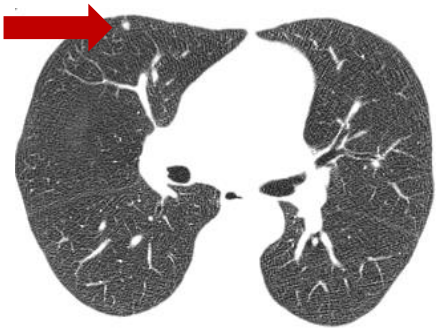
### 4.3.1 Detection of Smaller Lesions

It was noted that small size lesions (<= 6mm) had a low detection rate. By taking note of the output in all the different stages of the whole system, it was observed that, although the lesions were included in the intensity mask, they were considered to be false positives by the system. This misclassification was mainly due to the connectivity checks along the *z* axis. Since smaller lesions have less volume, this translates to fewer pixels in the *xy* plane and fewer slices along the *z* axis, especially if thick slices are used to scan the patient. Thus, a shift by a single pixel in the *xy* plane between slices, possibly due to poor edge definition, causes the overlap ratio to be heavily reduced, resulting in the region being removed by the system. Although the parameters used to distinguish between true and false positives were changed to be more lenient, it was noted that the false positive rate per scan would increase drastically to include such lesions.

Figures 4.2 (a) to (d) show part of a scan containing a lesion, indicated by an arrow in 4.2(b) and 4.2(c). Figure 4.2(e) and Figure 4.2(g) show the magnified intensity masks after false positive reduction along the *xy* plane for the images in Figure 4.2(b) and Figure 4.2(c) respectively. Figures 4.2(f) and Figure 4.2(h) show the intensity values and the co-ordinates of Figure 4.2(e) and Figure 4.2(g) respectively. The area in Figure 4.2(e) has 25 pixels whereas that in Figure 4.2(g) consists of 30 pixels. This means that, in order for the regions to be kept by the false positive reduction along the *z* plane, either 22 pixels (85% of 25) from the first region or 26 pixels (85% of 30) from the second region must overlap. However, it can be noted that there are only 18 overlapping pixels, meaning that the region is not considered to be a lesion.



(a)



(b)

(c)

(d)

(e)

| | 171 | 172 | 173 | 174 | 175 | 176 | 177 |
|---|---|---|---|---|---|---|---|
| 140 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 141 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 142 | 0 | 0 | 0 | 255 | 255 | 0 | 0 |
| 143 | 0 | 255 | 255 | 255 | 255 | 255 | 0 |
| 144 | 0 | 255 | 255 | 255 | 255 | 255 | 0 |
| 145 | 0 | 255 | 255 | 255 | 255 | 255 | 0 |
| 146 | 0 | 255 | 255 | 255 | 255 | 255 | 0 |
| 147 | 0 | 0 | 0 | 255 | 255 | 255 | 0 |
| 148 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

(f)

(g)

| | 171 | 172 | 173 | 174 | 175 | 176 | 177 | 178 |
|---|---|---|---|---|---|---|---|---|
| 140 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 141 | 0 | 0 | 255 | 255 | 0 | 0 | 0 | 0 |
| 142 | 0 | 255 | 255 | 255 | 255 | 255 | 0 | 0 |
| 143 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 0 |
| 144 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 0 |
| 145 | 0 | 255 | 255 | 255 | 255 | 255 | 255 | 0 |
| 146 | 0 | 0 | 0 | 255 | 255 | 255 | 0 | 0 |
| 147 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

(h)

**Figure 4.2:** *Misclassification of a lesion as a False Negative.*

### 4.3.2 Irregularly Shaped Lesions

Some lesions have spurious points which result in a large bounding box around their region in the *xy* plane due to the spike-like structure. The longer the spur, the larger the bounding box. An example of such a lesion can be observed in Figure 4.3. Thus, when it comes to considering the region area to bounding box area ratio, the region will occupy a small area of the box, since in reality, the box was elongated due to a thin pixel line. Therefore, when reducing false positives along the *xy* plane, such regions would be removed, resulting in a lower overall detection rate. Had the edge of the lesion been better defined, the outcome might have been different, since either the lesion would be detected as a circular region or more edge pixels would be present in the intensity mask, increasing the area occupied by the region in the bounding box. Appendix B shows 10 slices of the scan which contains the lesion in Figure 4.3, with the lesion marked with arrows. Appendix B, image (c), shows the exact same slice shown in Figure 4.3, with the lesion marked with a blue arrow. The CAD system does not detect this lesion, meaning it is classified as a false negative.

**Figure 4.3:** *An example of an irregularly shaped lesion, with an enlarged bounding box due to the spur. This spur will cause the lesion to be considered a false positive.*

### 4.3.3 Pleural Lesions

The segmentation phase is a critical phase in the CAD system, since it defines what areas should be processed in the lesion detection phase. Pleural lesions are lesions that are connected to the lung wall and therefore, they might be totally or partially mistaken as part of the background when segmenting. An example of a pleural lesion can be seen in Figure 4.4(a) and Figure 4.4(c). Note that this is the same lesion in consecutive slices. Figure 4.4(b) and Figure 4.4(d) show the segmentation output. The same lesion can be observed in Appendix A, where it is denoted by a blue arrow. It was observed that

pleural lesions are not always correctly included in the segmentation mask. As a result, volumes containing pleural lesions are not always passed on to the detection phase or, due to erroneous segmentation, the region passed on to the detection phase would have an irregular shape. Also, if a pleural lesion is included in one slice and not included in the next due to segmentation, connectivity along the *z* axis would be disrupted. This resulted in pleural lesions only being partially detected or not being detected at all.

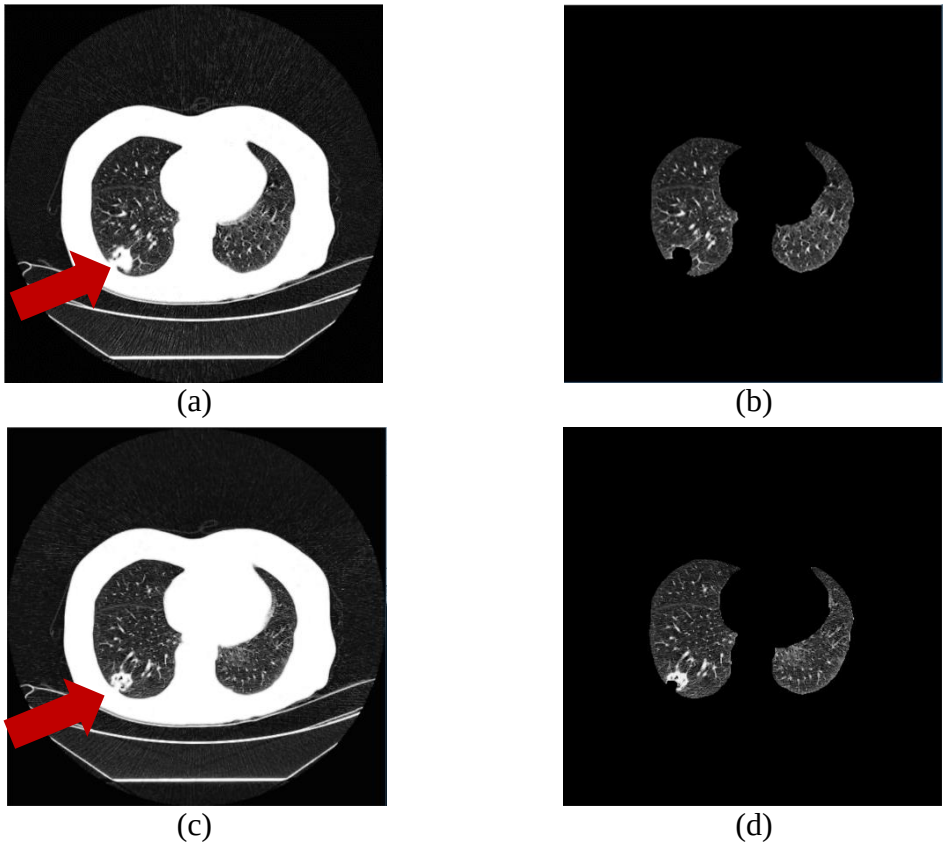
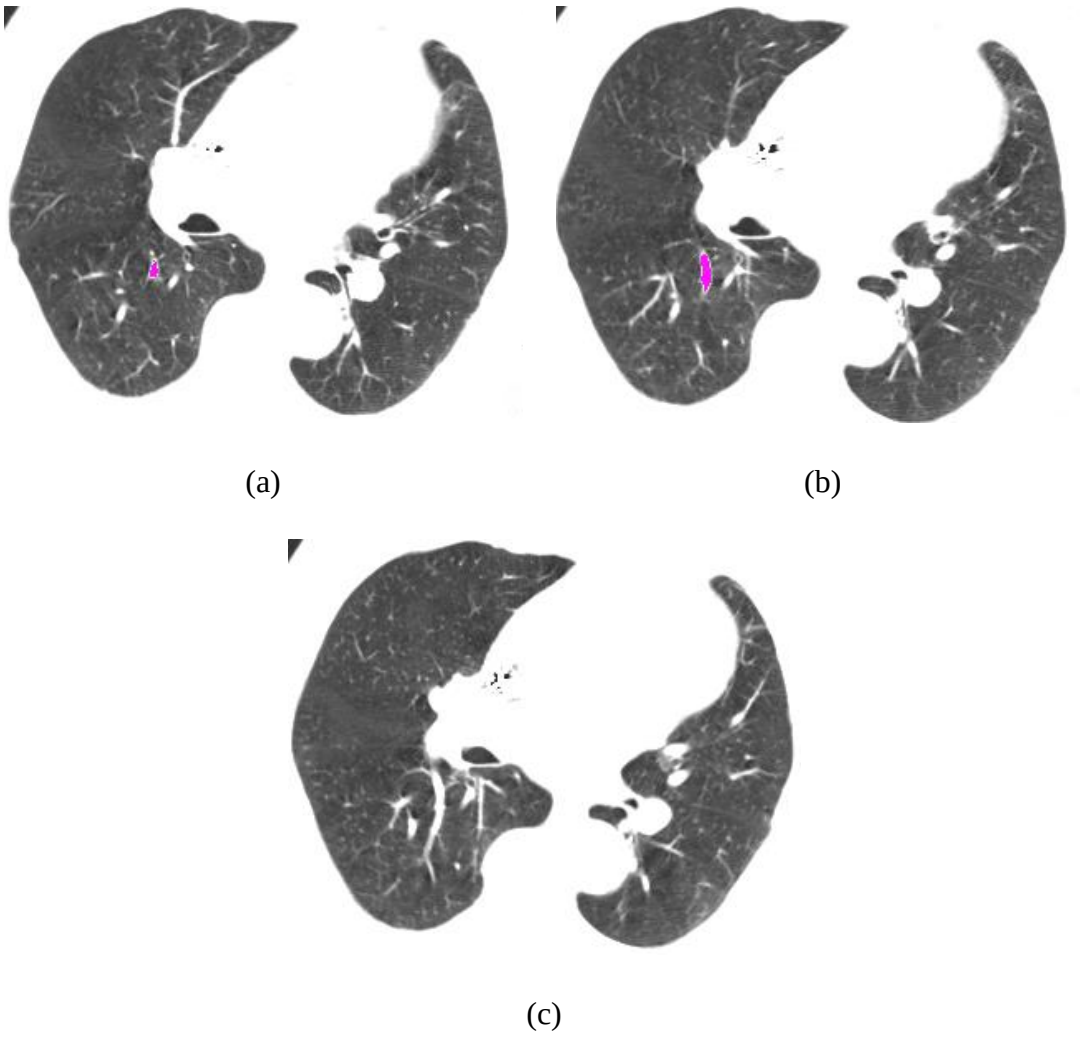



**Figure 4.4:** *An example of a pleural lesion (a), (c) and the segmentation result (b), (d).*

### 4.3.4 Vessels Detected as False Positives

As previously discussed, certain vessels will not be eliminated by the false positive reduction along the *xy* plane. This can be observed in Figure 4.5 below. From Figure 4.5(c), it is clear that the structure identified as a lesion is a vessel. However, since the regions formed in Figures 4.5(a) and 4.5(b), although elliptical, are parallel to the *y* axis, they are not removed.



(a) (b)

(c)

**Figure 4.5:** *False Positive Detection of Ellipsoidal Vessels; (a) and (b) show the vessel marked as a lesion, while (c) confirms that it is a vessel.*

Having said that, it was observed that most of the false positives being detected were vessels which show minimal movement between slices. This may be the result of the scan slices being very thin or because of the vessel being almost parallel to the *z* axis, that is, it deviates minimally between slices. Such vessels would satisfy the overlapping threshold set when reducing false positives along the *z* direction, adding to the number of false positives detected. If the overlapping threshold was increased, the detection of small lesions would have suffered, as discussed in a previous section.

Figure 4.6 below shows an example of a vessel that only just moves between slices. This vessel is erroneously detected as a lesion, since the movement is small enough for the overlapping threshold to be exceeded. In fact, the movement between Figure 4.6(c) and 4.6(d) appears to be enough for the system to no longer classify it as a lesion in Figure 4.6(d).

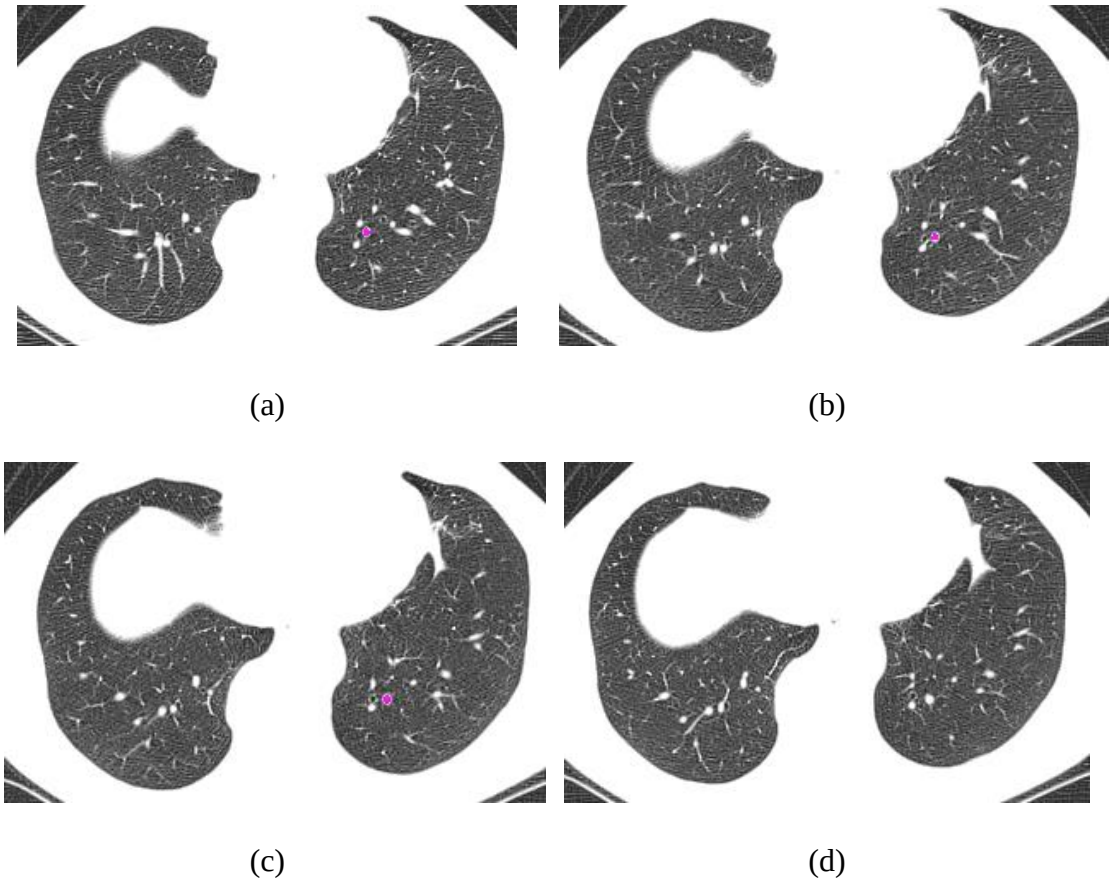



**Figure 4.6:** *False Positive Detection of a vessel which shows minimum movement between slices; (a), (b) and (c) show the regions which were erroneously detected while in (d) it is no longer detected as a lesion.*

# Chapter 5 : Conclusion and Future Work

This chapter gives a brief summary of the CAD system developed, whist also highlighting the results achieved. Any methods that might improve the CAD system's usability and overall results are also discussed.

## 5.1: Conclusion

With lung lesion detection playing a key role in the diagnosis of various health problems, it is imperative that improvements in technology be applied to an area of such relevance. The focus of the work described in this report has been the development of a CAD System which makes use of image processing techniques to detect lung lesions in thoracic CT Scans. The CAD System's steps can be summarised as follows:

- The DICOM data is read into the system, and a windowing function is applied to the images;
- Noise is removed through the use of a Bilateral Filter;
- The contrast of each image is enhanced by first applying Gamma Correction, followed by Contrast Limited Adaptive Histogram Equalisation;
- The lung structure is segmented from the images by means of marker based watershed segmentation;
- Detection of lesions is performed by taking into consideration the intensity and the shape of lesions;
- The system outputs a set of images with markers to indicate where it has detected lesions.

The system was tested on 25 thoracic CT Scans. A sensitivity of 76.27% was achieved, with an average false negative rate of 0.56 lesions per scan and an average false positive rate of 7.88 lesions per scan. This was done in an average time of 74.3 milliseconds per slice. It was noted that improvements can be made to the system to better handle small lesions (<= 6mm), irregularly shaped lesions and pleural lesions. To this aim, potential ideas that might improve the system are also presented, including the use of enhancement filters and mapping of the respiratory tree.

## 5.2 Future Work

### 5.2.1 Blob Enhancement Filter

To detect smaller lesions and to obtain a more circular shape out of spurious lesions, a blob enhancement filter can be applied in the pre-processing stage. Such a filter would help to pinpoint lesions better, since their edges would be much better defined. Of course, various sizes of such a filter would need to be applied to cater for lesions of different sizes. A blob enhancement filter might also aid to filter out elliptical regions which would not be eliminated by the technique employed in the *xy* plane due to their angle.

### 5.2.2 Segmentation and Mapping of the Respiratory Tree

Better segmentation would define pleural lesions better, which would in turn increase their detection rate. To this aim, it might be beneficial to map the lower respiratory tree, that is, the part of the respiratory system that supplies air to the lungs. It was observed that the lung contour in these regions, where the respiratory tree enters the lungs, could be better defined. Thus, removing the respiratory tree from these regions could aid in the segmentation process. Parts of the respiratory tract also show up as white structures in CT scans and therefore, false positives may be reduced if these areas are not considered in the detection phase. Mapping of the vessels was used by Tong *et al* [46], whose system has a low false positive rate.

### 5.2.3 Limiting the Number of Planes along the *Z* Axis

As observed in the results, vessels which are quasi-parallel to the *z* axis are one of the causes of the high false positive rate. To reduce such a rate, it might be beneficial to limit the number of planes that can be included in the connectivity analyses along the *z* axis, based on the lesion candidate width and height in the *xy* plane. For example, if a lesion is 5mm wide, and the slice thickness is 2.5mm, then it should not be connected in more than 3 planes. Any region that is 5mm wide, but has connectivity in more than 3 planes, is most likely a vessel.

The current implementation only specifies the minimum number of planes that must be connected. Defining the maximum number of planes that can be connected for a particular lesion is not a trivial task. This is because consideration of which part of the

lesion is in the *xy* plane needs to be determined. For example, if it is the topmost part of the lesion, planes above it should not be considered, and the amount of planes to be considered below it should be higher than if it were a central part.

### 5.2.4 Consideration of Other Features

It would be interesting to try and extract more features of the detected lesions such as the size and volume occupied by the region. Such features might not only be useful to radiologists in their diagnosis, but may also be helpful when removing false positives from the output. This would require the use of some classifier, such as the fuzzy k-nearest neighbours classifier used by Namin *et al* [43].

### 5.2.5 Output

The current system produces a series of 2D images, similar to the output of a CT scanner. It might be more beneficial for radiologists to observe the output in a 3D manner, such as a point cloud. The same segmentation masks developed can be used to extract the lungs from the images and obtain the 3D model.

### 5.2.6 Parallelisation Techniques

The current implementation makes use of serial techniques to detect lesions. Having said that, there are a lot of steps in the system that can be implemented in parallel, due to the large amount of data independency present in the same CT scan. This would result in a speedier version of the CAD system, possibly achieving results in near real-time. This would in turn result in faster reading times by radiologists.

# References


[1] "Cancer", World Health Organization, 2015. [Online]. Available: http://www.who.int/mediacentre/factsheets/fs297/en/. [Accessed: 19- May- 2016].

[2] R. Castellino, “Computer aided detection (CAD): an overview”, Cancer Imaging, vol. 5, no. 1, pp. 17-19, 2005.

[3] K. Doi, “Computer-aided diagnosis in medical imaging: Historical review, current status and future potential”, Computerized Medical Imaging and Graphics, vol. 31, no. 4-5, pp. 198-211, 2007.

[4] H. Roberts, “Use of computer-aided detection in thoracic CT”, Applied Radiology, vol. 36, no. 12, 2007.

[5] S. Smith, “Special Imaging Techniques” in *The scientist and engineer's guide to digital signal processing*, 2nd ed. San Diego, Calif.: California Technical Pub., 1999, ch. 25, pp. 423-449

[6] A.G. Filler, “The History, Development and Impact of Computed Imaging in Neurological Diagnosis and Neurosurgery: CT, MRI, and DTI”, The Internet Journal of Neurosurgery, vol. 7, no. 1, 2010.

[7] Lecture Notes, Stanford University. [Online]. Available: http://web.stanford.edu/group/glam/xlab/MatSci162_172/LectureNotes/01_Properties%20&%20Safety.pdf

[8] Y. Waseda, E. Matsubara and K. Shinoda, “Fundamental Properties of X-rays” in *X-Ray diffraction crystallography*, 1st ed. Berlin: Springer, 2011, ch. 1, pp. 1-5

[9] M. Newville, “Fundamentals of XAFS”, University of Chicago, Chicago, Ver. 1.7, 2004

[10] "X-ray absorption, A", Pd.chem.ucl.ac.uk, 2006. [Online]. Available: http://pd.chem.ucl.ac.uk/pdnn/diff2/abs.htm. [Accessed: 19- May- 2016].

[11] K. Patil and M. Patil, “Principle of Radiographic Interpretation,” in Textbook of Oral Radiology, New Delhi, India: Elsevier, 2008, ch. 21, sec. 4, p. 214

[12] R. Gonzalez and R. Woods, *Digital image processing*, 3rd ed. New Jersey: Prentice-Hall, Inc., 2006.

[13] T. Allmendinger, "X-ray CT scan and dual source CT system" DE Patent App. DE201,310,203,541, September 4, 2014

[14] D. Anderson, *Mosby's medical, nursing, & allied health dictionary*, 6th ed. St. Louis: Mosby, 2002.

[15] My.clevelandclinic.org, "Pulmonary Nodules (Lung Nodules) | Cleveland Clinic", 2015. [Online]. Available: https://my.clevelandclinic.org/health/diseases_conditions/hic_Pulmonary_Nodules. [Accessed: 19-May-2016].

[16] M. Sonka, V. Hlavac and R. Boyle, *Image processing, analysis, and machine vision*, 3rd ed. Toronto: Thompson Learning, 2008.

[17] T. L. Toth, "Image quality in CT: challenges and perspectives" in *Radiation dose from multidetector CT*, edited by D. Tack, M. Kalra and P. Gevenois, 2nd ed. Berlin: Springer, 2012, ch. 2, sec. 3, pp. 81-99.

[18] Siemens Medical Solutions, "Basics of CT," [Online]. Available: http://www.schreisupport.com/Schreisupport/GENERAL%20INFORMATION/CT/Basics%20of%20CT.pdf

[19] R. Szeliski, "Image Processing" in *Computer Vision Algorithms and Applications*, 1st ed. London: Springer, 2011, ch. 3, pp. 87-180.

[20] G. Vijaya and A. Suhasini, "An Adaptive Preprocessing of Lung CT Images with Various Filters for Better Enhancement", Academic Journal of Cancer Research, vol. 7, no. 3, 2014.

[21] S. Ashwin, J. Ramesh, S. Kumar and K. Gunavathi, "Efficient and reliable lung nodule detection using a neural network based computer aided diagnosis system", in Proceedings of the International Conference on Emerging Trends in Electrical Engineering and Energy Management (ICETEEEM), 2012

[22] L. Zhang, J. Chen, Y. Zhu and J. Luo, "Comparisons of Several New De-Noising Methods for Medical Images", in Proceedings of the 3rd International Conference on Bioinformatics and Biomedical Engineering, 2009.

[23] V. Banupriya and J. Bagyamani, “Lung Cancer Detection Using Circular Hough Transform Technique”, International Journal of Advance Research in Science and Engineering, vol. 4, no. 03, 2015.

[24] D. Sharma and G. Jindal, "Identifying Lung Cancer Using Image Processing Techniques", in Proceedings of the International Conference on Computational Techniques and Artificial Intelligence (ICCTAI)

[25] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629-639, Jul. 1990.

[26] E. Berry, “Image processing Basics” in *A practical approach to medical image processing*, 1st ed. London: Taylor & Francis, 2008, ch. 1, pp. 1-35.

[27] "Information About Intravenous and Oral Contrast Used in CT | CT Scan | Imaginis - The Women's Health & Wellness Resource Network", Imaginis.com, 2010. [Online]. Available: http://www.imaginis.com/ct-scan/information-about-intravenous-and-oral-contrast-used-in-ct-1. [Accessed: 19- May- 2016].

[28] "Contrast Stretching", Academic.mu.edu, 2001. [Online]. Available: http://academic.mu.edu/phys/matthysd/web226/Lab01.htm. [Accessed: 19- May- 2016].

[29] J. Kim, L. Kim and S. Hwang, "An advanced contrast enhancement using partially overlapped sub-block histogram equalization", *IEEE Transactions Circuits Systems for Video Technology*, vol. 11, no. 4, pp. 475-484, Apr. 2001.

[30] N. Kong, H. Ibrahim and S. Chun Hoo, "A Literature Review on Histogram Equalization and Its Variations for Digital Image Enhancement", International Journal of Innovation, Management and Technology, 2013.

[31] K. Somasundaram and P. Kalavathi, "Medical Image Contrast Enhancement Based On Gamma Correction", International Journal of Knowledge Management & e-Learning, vol. 3, no. 1, pp. 15-18, 2011.

[32] Z. Al-Ameen, G. Sulong, A. Rehman, A. Al-Dhelaan, T. Saba and M. Al-Rodhaan, "An innovative technique for contrast enhancement of computed tomography images using normalized gamma-corrected contrast-limited adaptive histogram equalization", EURASIP Journal on Advances in Signal Processing, vol. 2015, no. 1, 2015.

[33] S. Pizer, E. Amburn, J. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. ter Haar Romeny, J. Zimmerman and K. Zuiderveld, "Adaptive histogram equalization and its variations", Computer Vision, Graphics, and Image Processing, vol. 39, no. 3, pp. 355-368, 1987.

[34] J. Silva, A. Silva and B. Souza, "Lung Segmentation Methods in X-ray CT Images", in Proceedings of the SIARP2000 5th Iberoamerican Symposium on Pattern Recognition, pp. 583-598, 2000.

[35] N. Memon, A. Mirza and S. Gilani, "Segmentation of Lungs from CT Scan Images for Early Diagnosis of Lung Cancer", World Academy of Science, Engineering and Technology 20, 2006.

[36] A. Fisher, "Cloud and Cloud-Shadow Detection in SPOT5 HRG Imagery with Automated Morphological Feature Extraction", Remote Sensing, vol. 6, no. 1, pp. 776-800, 2014.

[37] R. Shojaii, J. Alirezaie and P. Babyn, "Automatic lung segmentation in CT images using watershed transform", in Proceedings of the IEEE International Conference on Image Processing, 2005.

[38] H. Zhou, D. Goldgof, S. Hawkins, L. Wei, Y. Liu, D. Creighton, R. Gillies, L. Hall and S. Nahavandi, "A Robust Approach for Automated Lung Segmentation in Thoracic CT", in Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, 2015.

[39] Y. Boykov and O. Veksler, "Graph Cuts in Vision and Graphics: Theories and Applications", in *Handbook of Mathematical Models in Computer Vision*, 1st ed. New York: Springer, 2006, ch.5, pp. 79-96.

[40] S. Dai, K. Lu, J. Dong, Y. Zhang and Y. Chen, "A novel approach of lung segmentation on chest CT images using graph cuts", Neurocomputing, vol. 168, pp. 799-807, 2015.

[41] A. Ali and A. Farag, "Automatic Lung Segmentation of Volumetric Low-Dose CT Scans Using Graph Cuts", Advances in Visual Computing, pp. 258-267, 2008.

[42] M. Antonelli, G. Frosini, B. Lazzerini and F. Marcelloni, “Lung Nodule Detection in CT Scans”, International Journal of Medical, Health, Biomedical, Bioengineering and Pharmaceutical Engineering, vol. 1, no. 1, 2007.

[43] S. Namin, H. Abrishami Moghaddam, R. Jafari, M. Esmaeil-Zadeh and M. Gity, "Automated detection and classification of pulmonary nodules in 3D thoracic CT images", in Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, 2010.

[44] W. Shang, H. Huang, H. Zhu, Y. Lin, Z. Wang and Y. Qu, "An Improved kNN Algorithm – Fuzzy kNN", Computational Intelligence and Security, pp. 741-746, 2005.

[45] W. Suiyuan and W. Junfeng, "Pulmonary Nodules 3D Detection on Serial CT Scans", 2012 Third Global Congress on Intelligent Systems, 2012.

[46] J. Tong, W. Ying and W. Dong, "A lung cancer lesions detection scheme based on CT image", in Proceedings of the 2nd International Conference on Signal Processing Systems, 2010.

[47] S. Gunn, *Support vector machines for classification and regression*, Image Speech and Intelligent Systems Group, Department of Electronics and Computer Science, University Southampton, 1998

[48] W. Choi and T. Choi, "Automated Pulmonary Nodule Detection System in Computed Tomography Images: A Hierarchical Block Classification Approach", *Entropy*, vol. 15, no. 2, pp. 507-523, 2013.

[49] B. Krose and P. van der Smagt, An Introduction to Neural Networks, 8th ed. 2016.

[50] “DICOM Homepage”, *Dicom.nema.org*, 2016. [Online]. Available: http://dicom.nema.org/standard.html. [Accessed: 19- May- 2016].

[51] Offis Team, "dicom.offis.de - DICOM Software made by OFFIS - DCMTK - DICOM Toolkit", Dicom.offis.de, 2016. [Online]. Available: http://dicom.offis.de/dcmtk.php.en. [Accessed: 19- May- 2016].

[52] "OpenCV 3.0 | OpenCV", Opencv.org, 2016. [Online]. Available: http://opencv.org/opencv-3-0.html. [Accessed: 19- May- 2016].

[53] C. Tsiotsios and M. Petrou, "On the choice of the parameters for anisotropic diffusion in image processing", Pattern Recognition, vol. 46, no. 5, pp. 1369-1381, 2013.

[54] "Image Filtering — OpenCV 3.0.0-dev documentation", Docs.opencv.org, 2016. [Online]. Available: http://docs.opencv.org/3.0-beta/modules/imgproc/doc/filtering.html#sobel. [Accessed: 19- May- 2016].

[55] "Eigen", Eigen.tuxfamily.org, 2016. [Online]. Available: http://eigen.tuxfamily.org/index.php?title=Main_Page. [Accessed: 19- May- 2016].

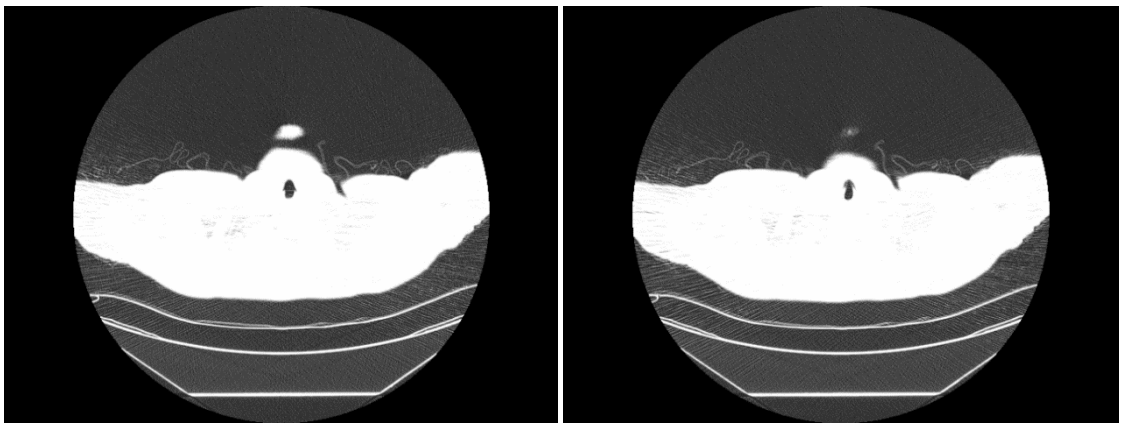
[56] H. MacMahon, J. Austin, G. Gamsu, C. J. Herold, J. R. Jett, D. P. Naidich, E. F. Patz, Jr, S. J. Swensen, "Guidelines for Management of Small Pulmonary Nodules Detected on CT Scans: A Statement from the Fleischner Society", Radiology, vol. 237, no. 2, pp. 395-400, 2005.

[57] "Lung Image Database Consortium (LIDC)", *Cancer Imaging Program - National Cancer Institute*, 2016. [Online]. Available: http://imaging.cancer.gov/programsandresources/informationsystems/lidc. [Accessed: 19- May- 2016].

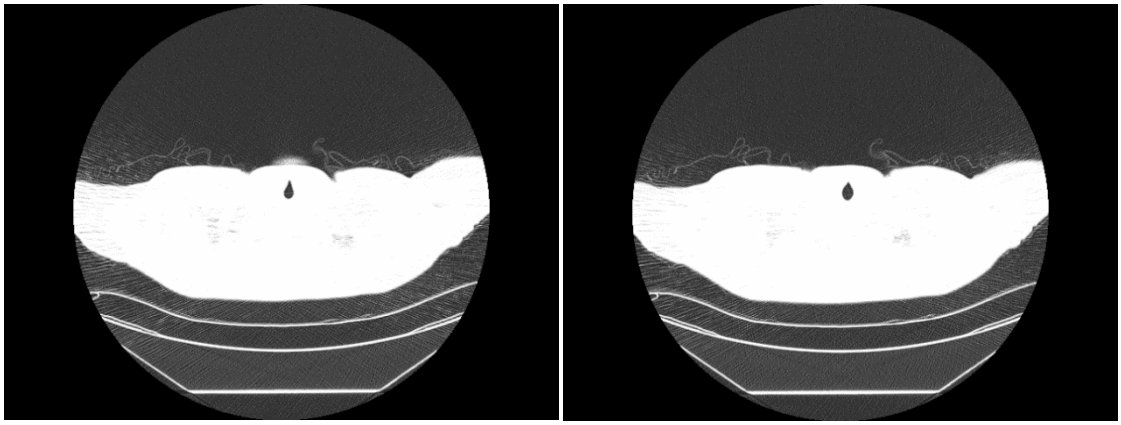
# Appendix

## A – A Whole CT Scan of 99 Images

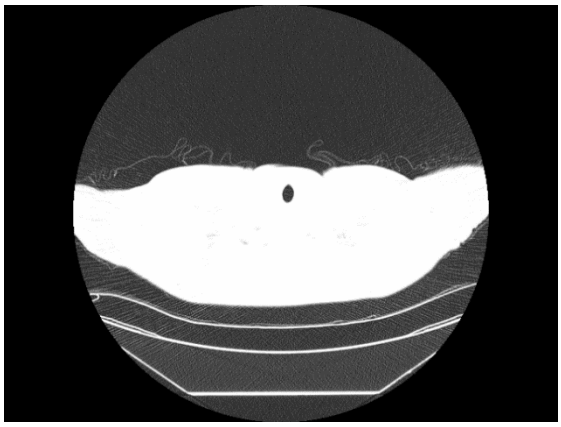
The following is a typical CT scan consisting of 99 slices. Each slice is 2.5mm thick. Each image is 512 × 512 pixels, with the resolution along both the *x* and *y* axes equal to 0.816mm. Two lesions are marked with arrows, with the blue arrow showing a pleural lesion.



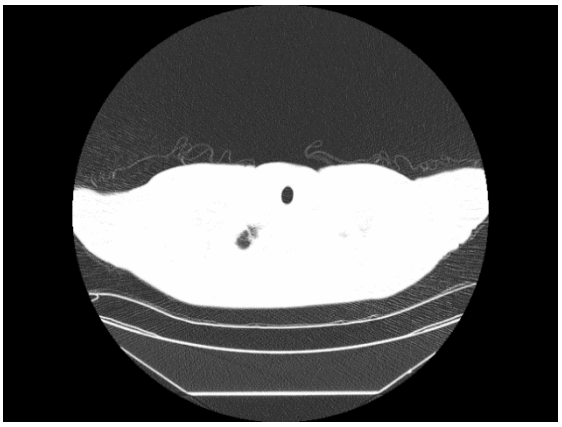
(1) (2)



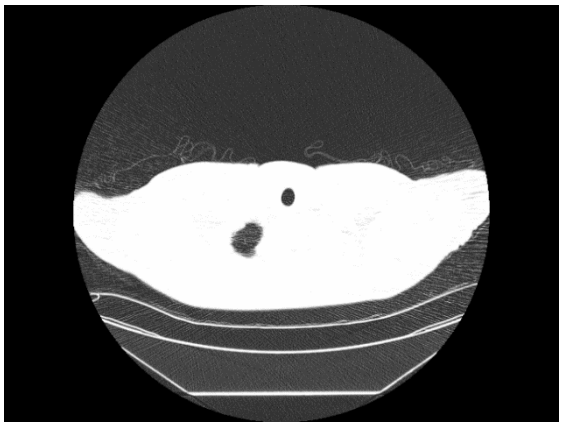
(3) (4)



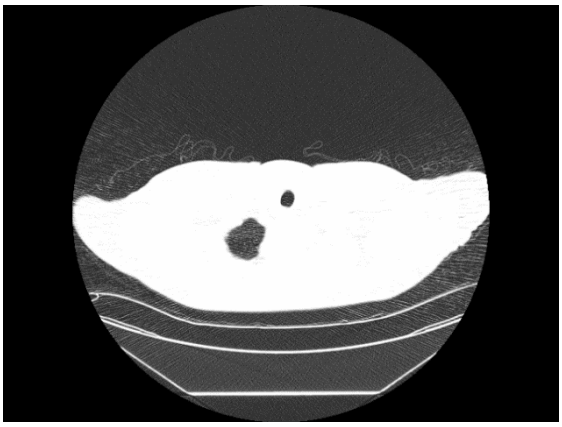
(5)



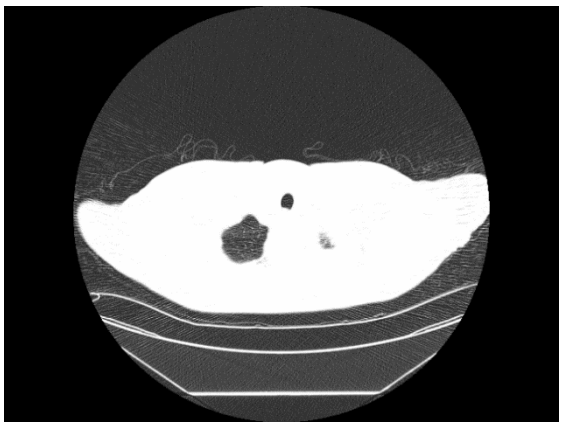
(6)



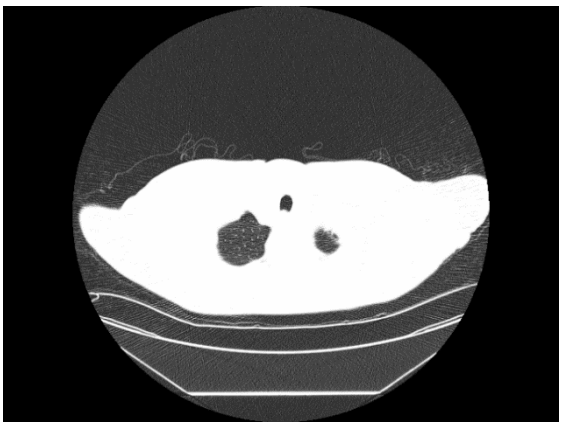
(7)



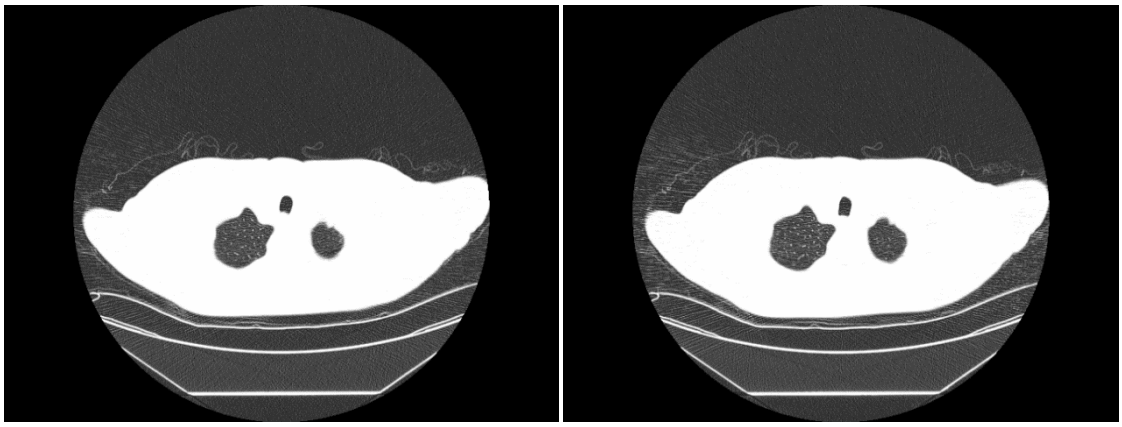
(8)



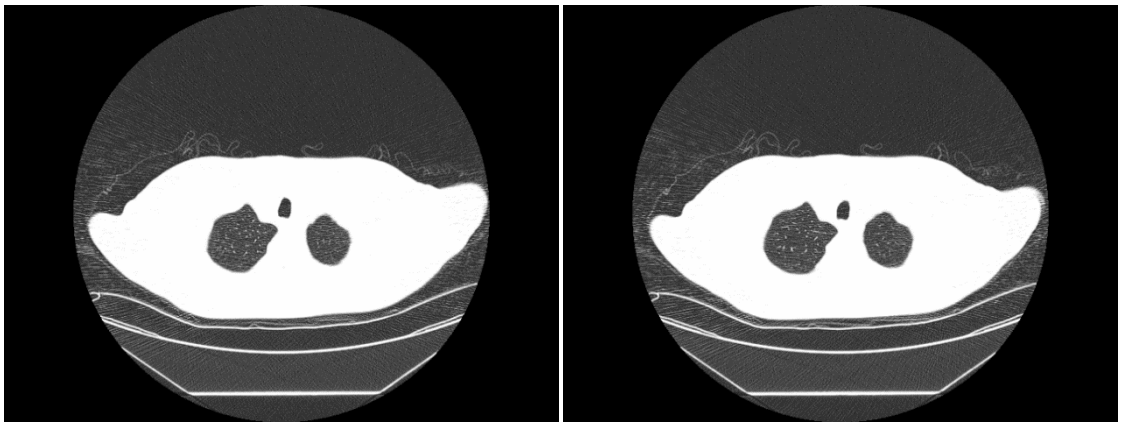
(9)



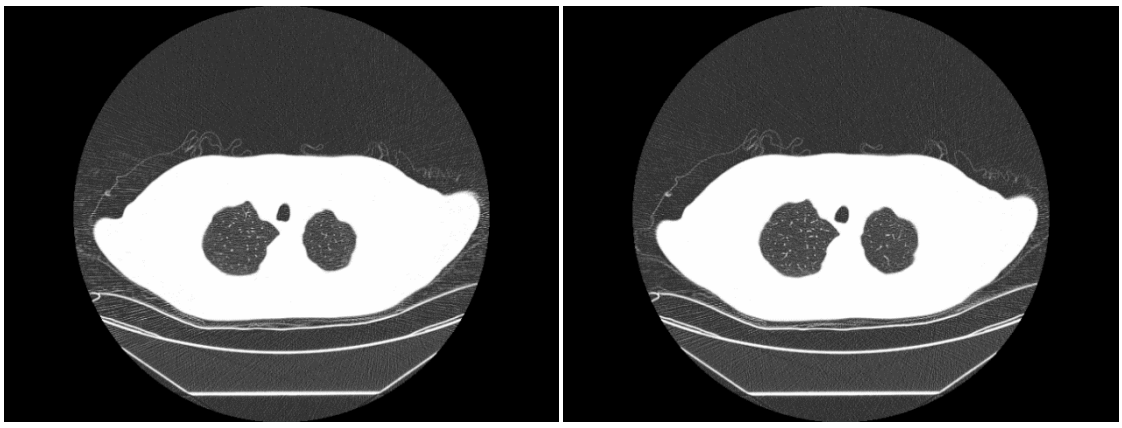
(10)



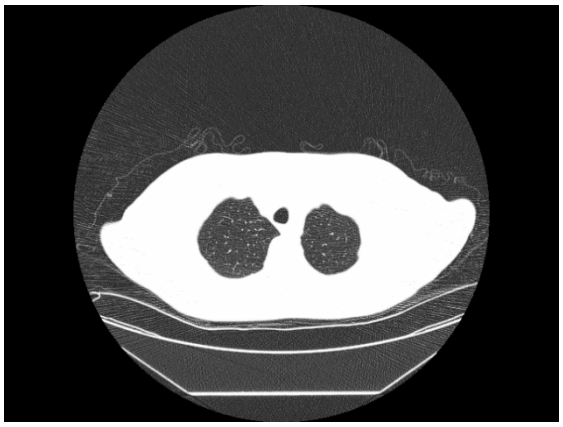
(11) (12)



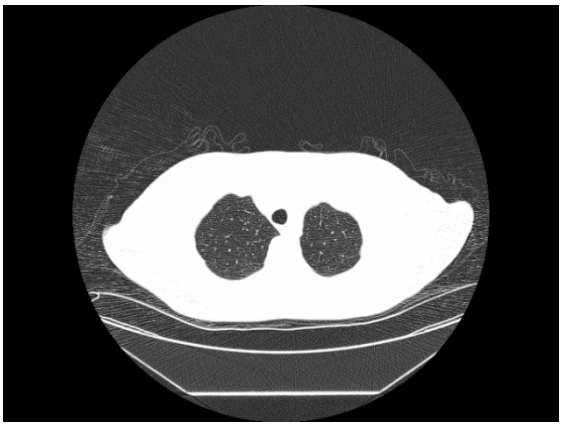
(13) (14)



(15) (16)



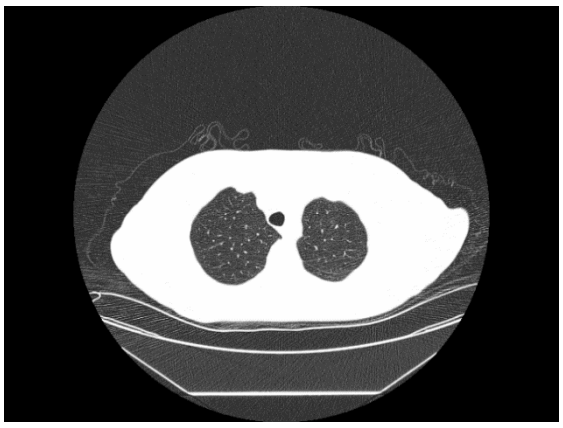
(17)



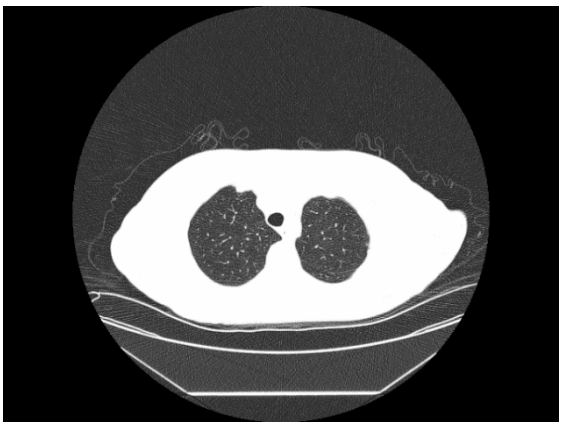
(18)

(19)

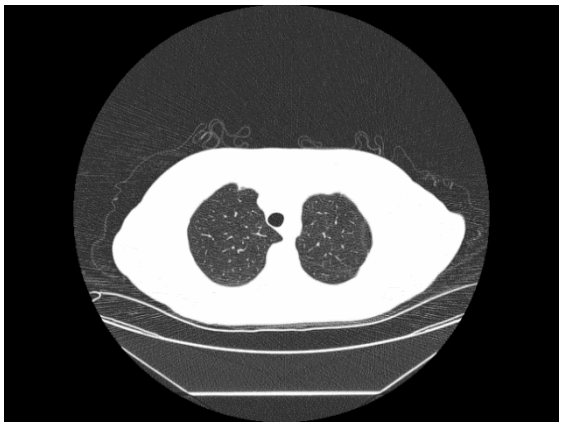
(20)



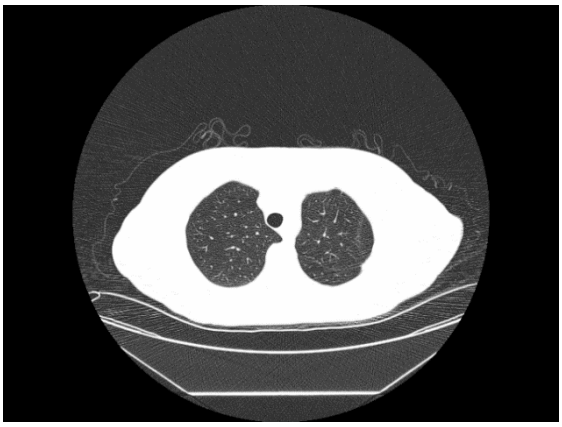
(21)



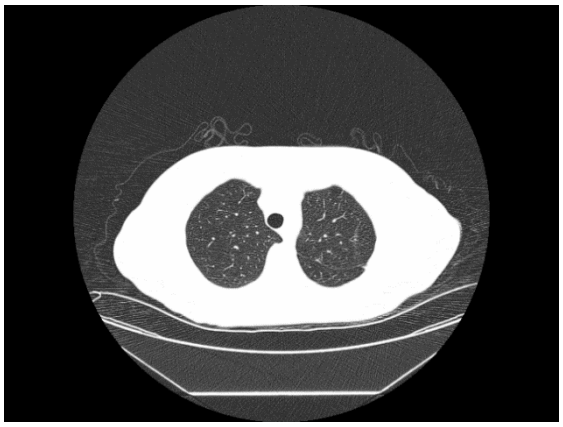
(22)



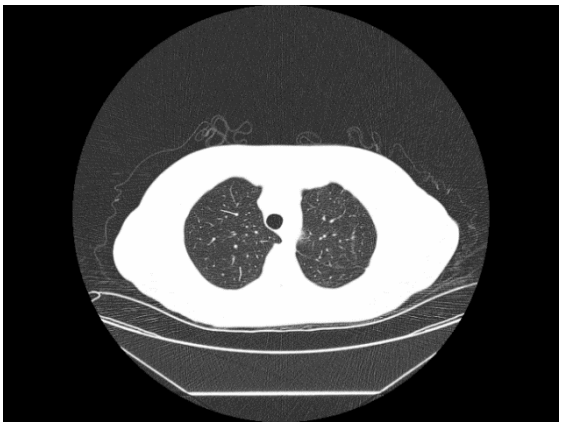
(23)



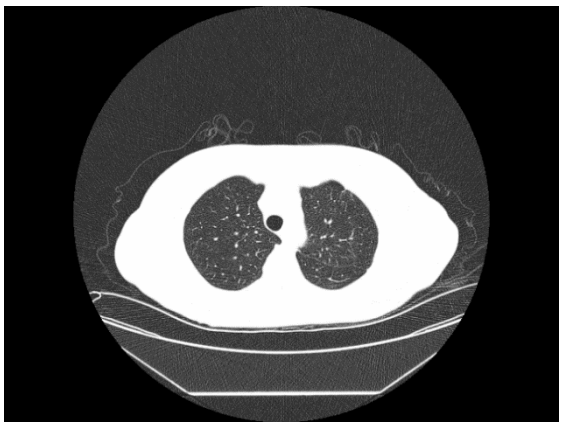
(24)



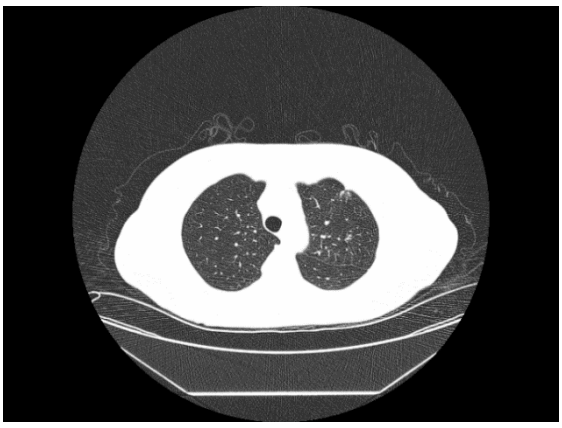
(25)



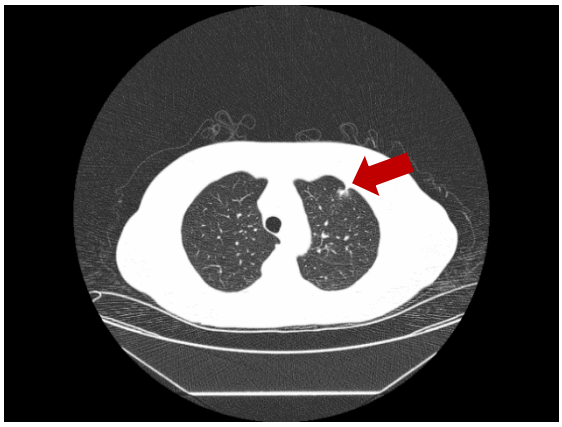
(26)



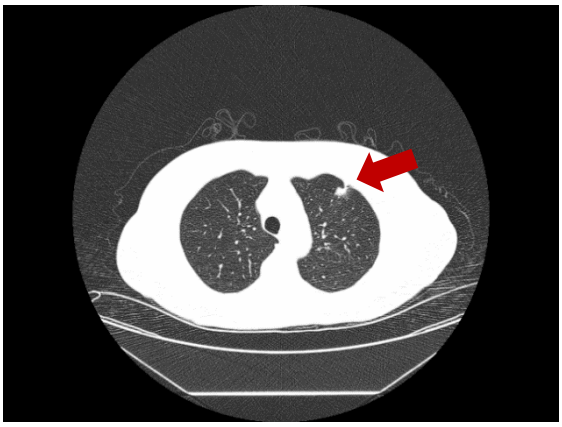
(27)



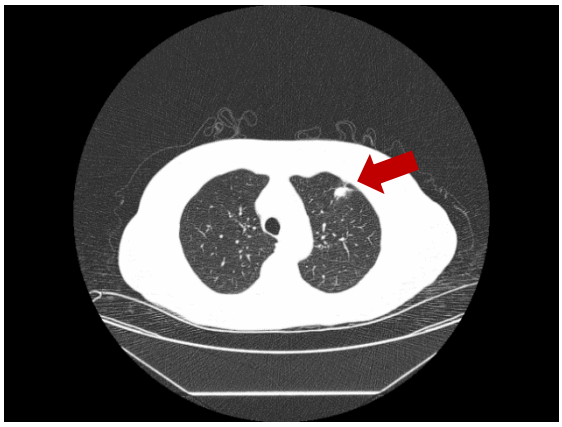
(28)



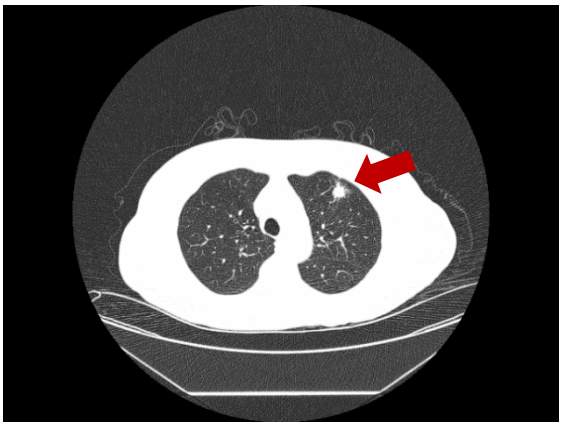
(29)



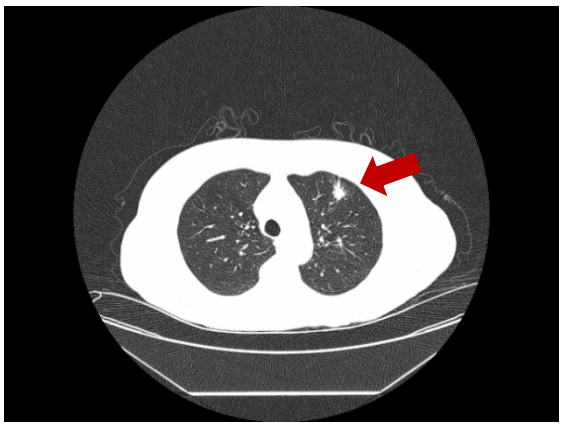
(30)



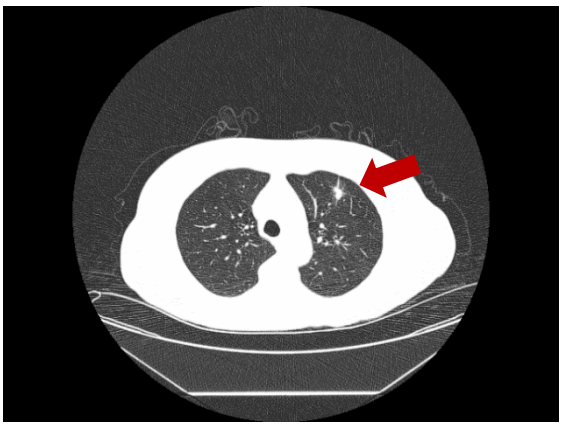
(31)



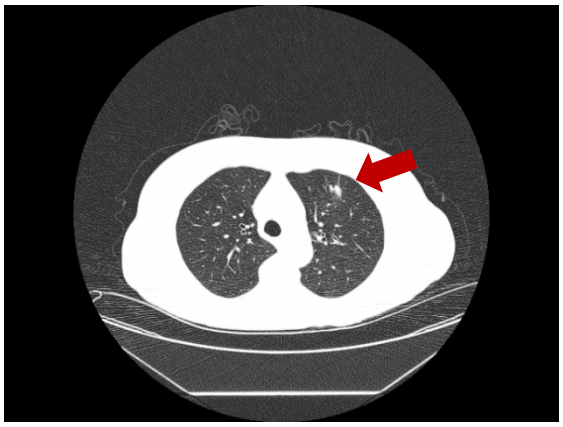
(32)



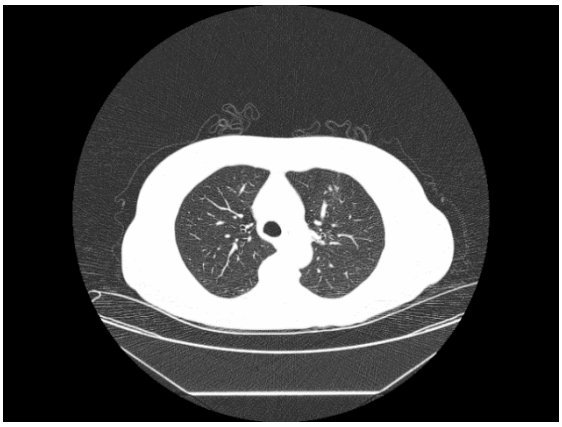
(33)



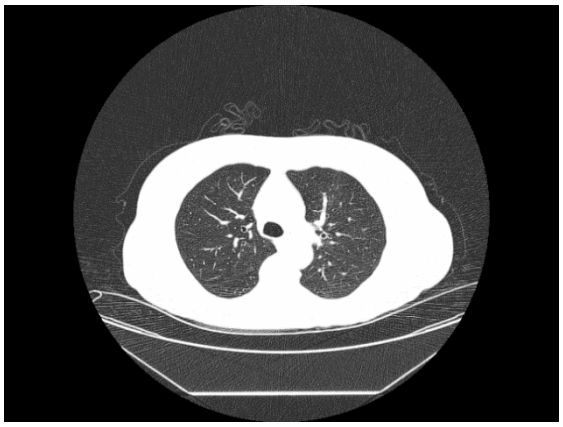
(34)



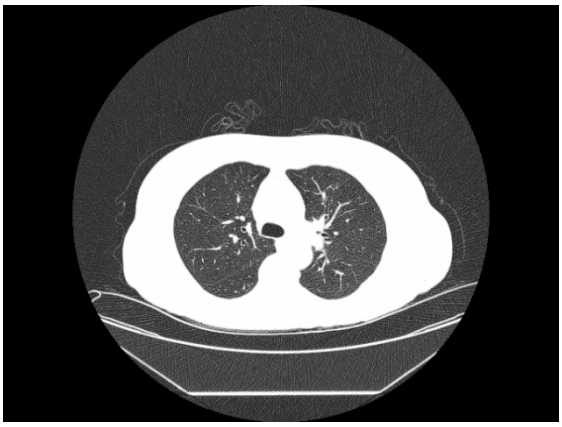
(35)



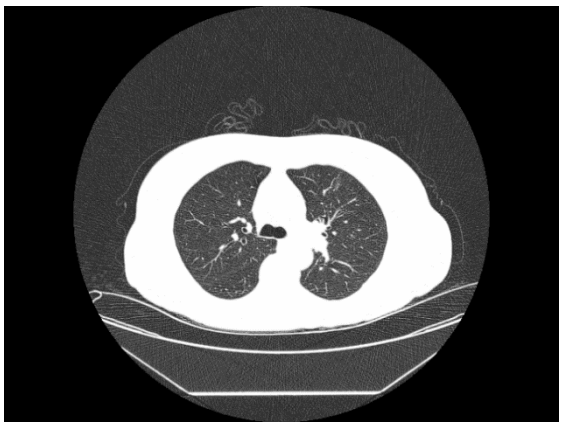
(36)



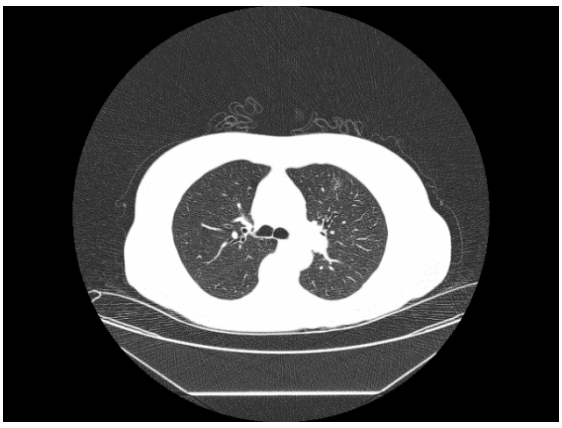
(37)



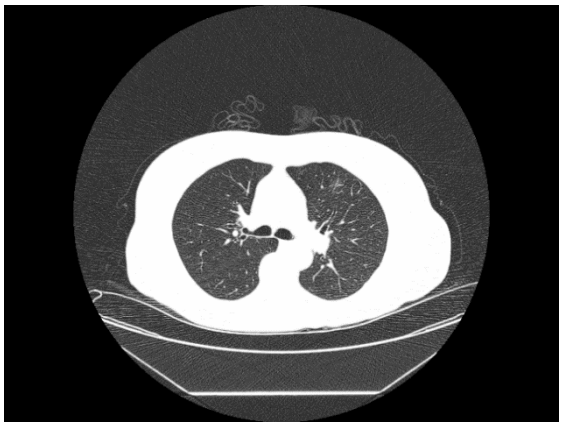
(38)



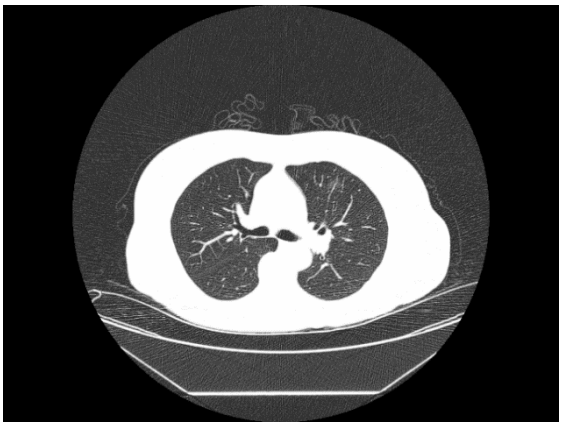
(39)



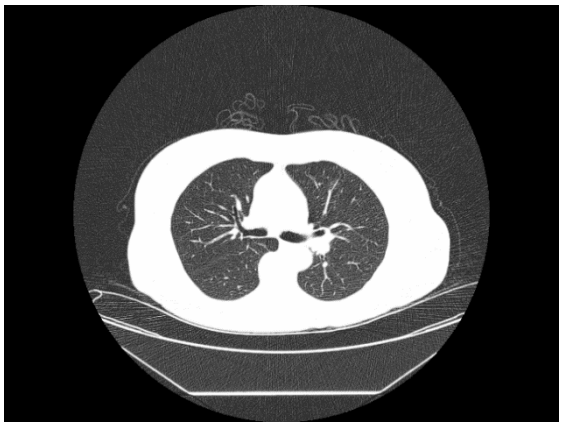
(40)



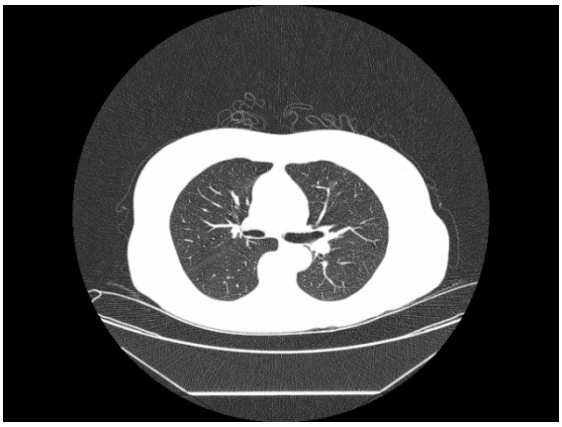
(41)



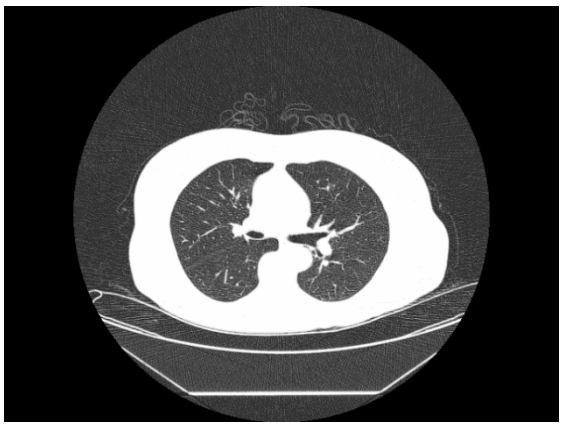
(42)



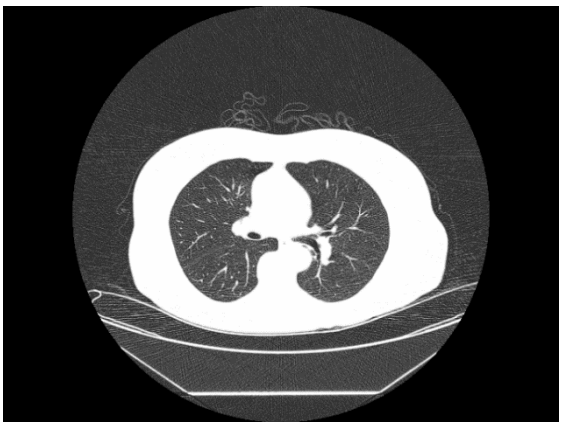
(43)



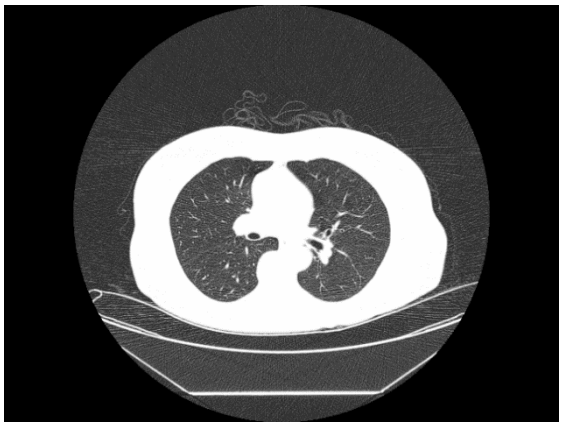
(44)



(45)

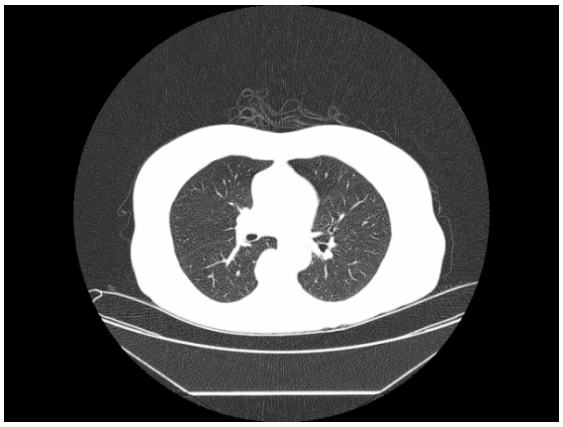


(46)

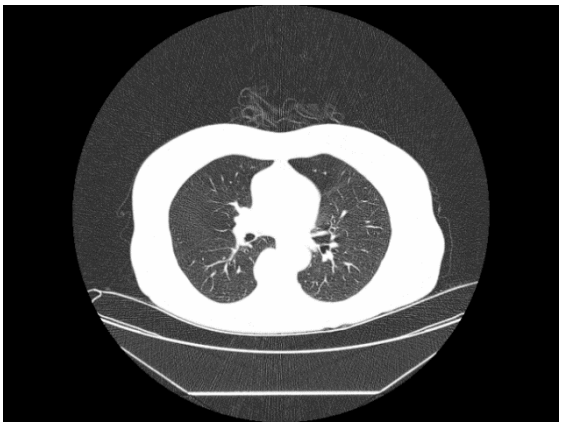


(47)

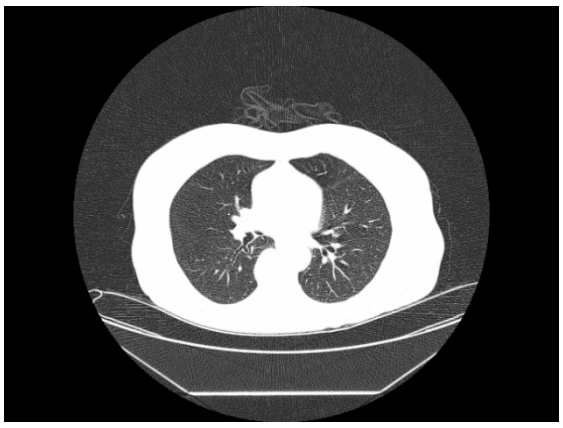
(48)



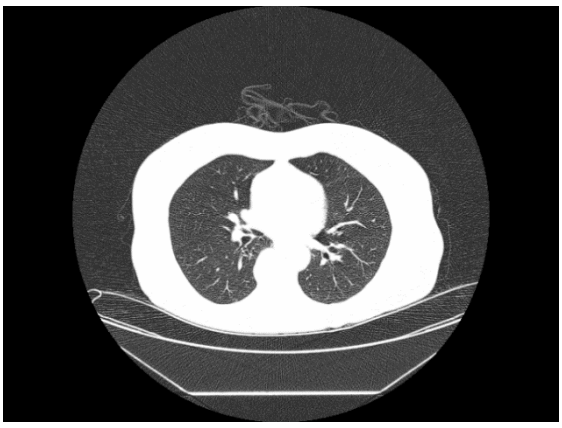
(49)



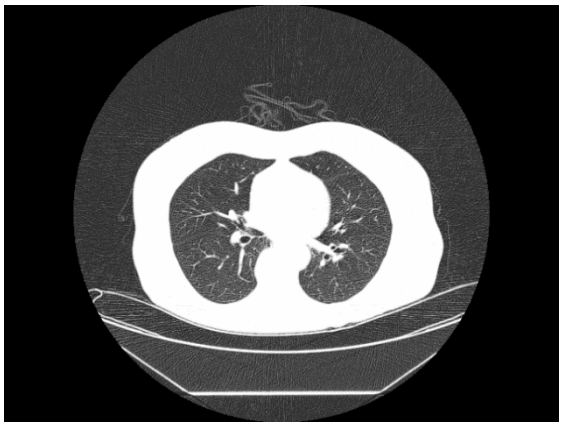
(50)



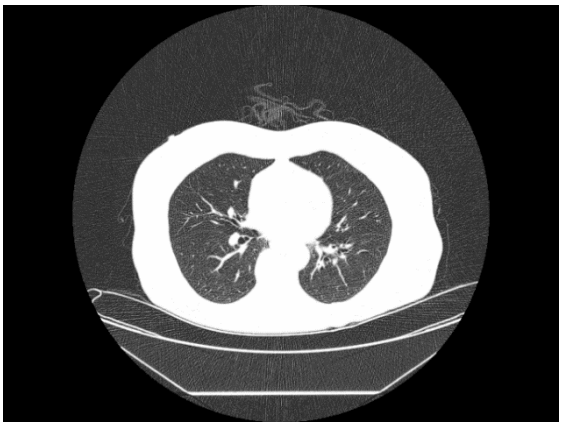
(51)



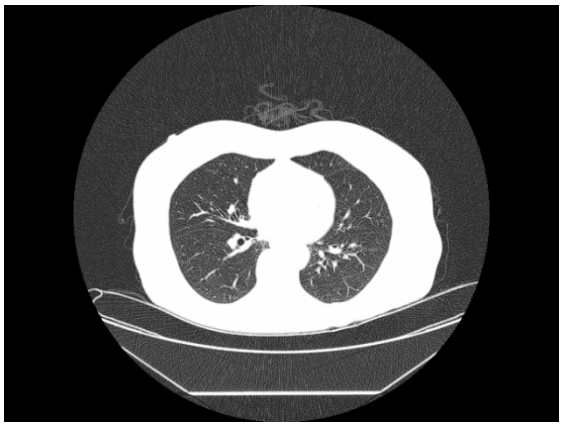
(52)



(53)

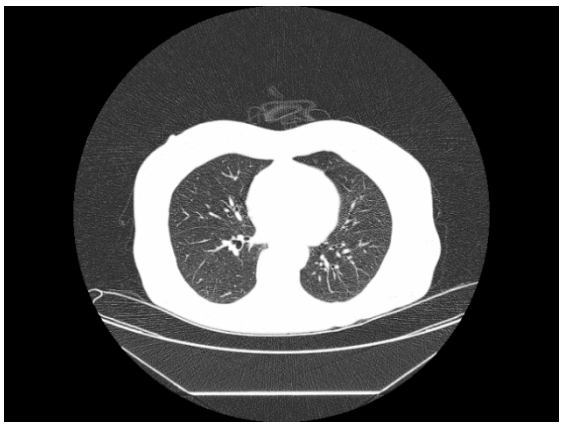


(54)

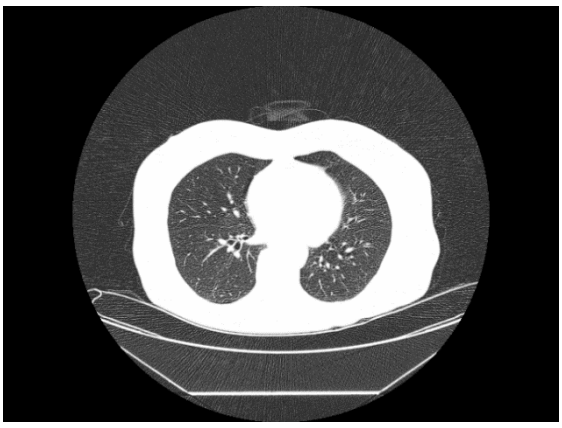


(55)

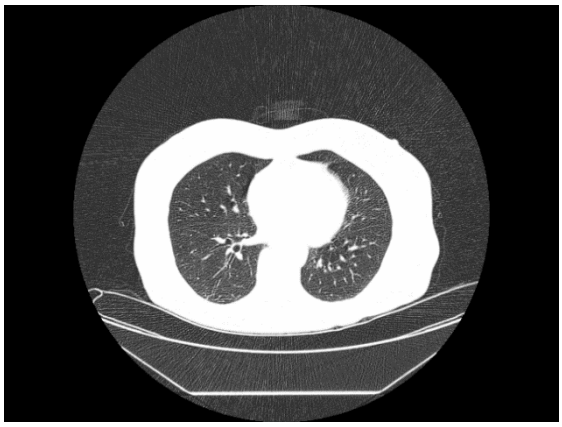
(56)



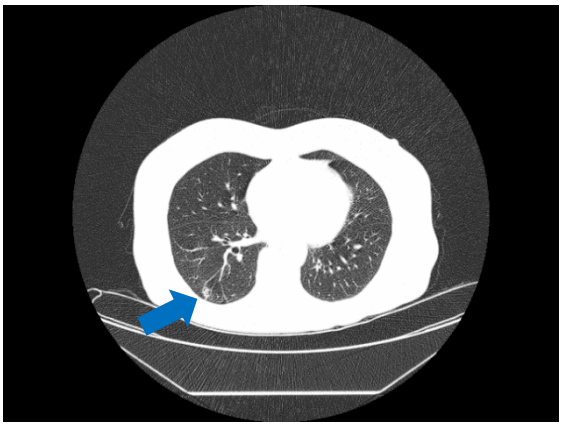
(57)



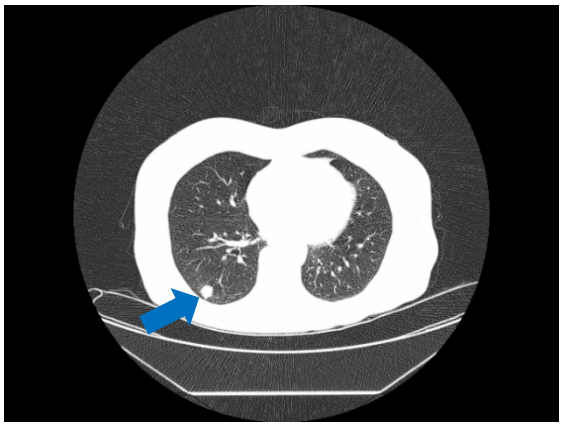
(58)



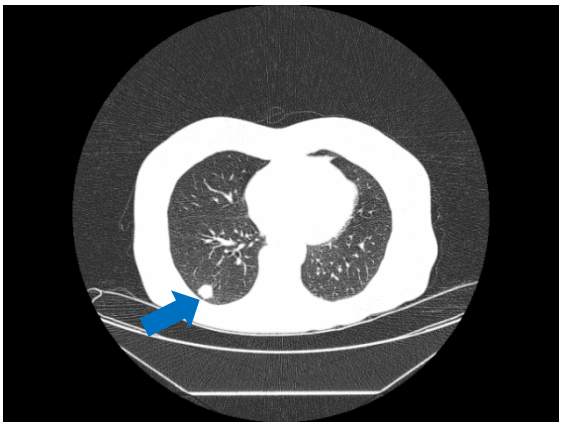
(59)



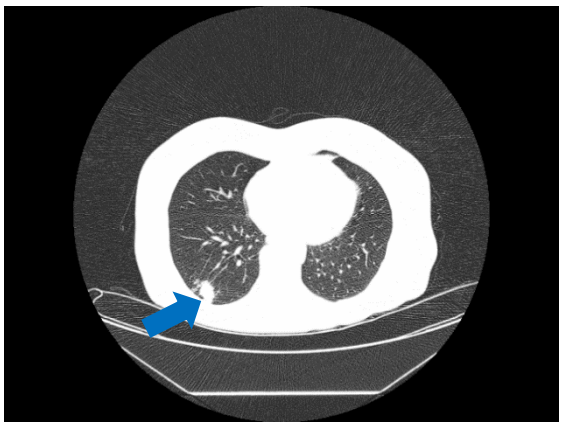
(60)



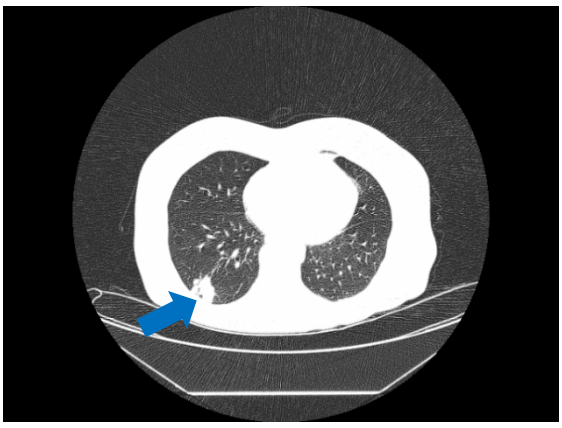
(61)



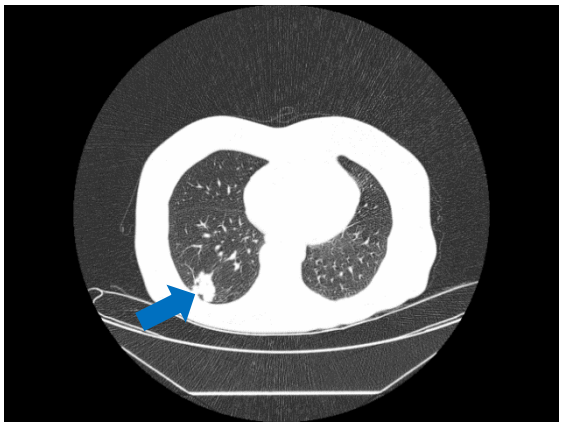
(62)



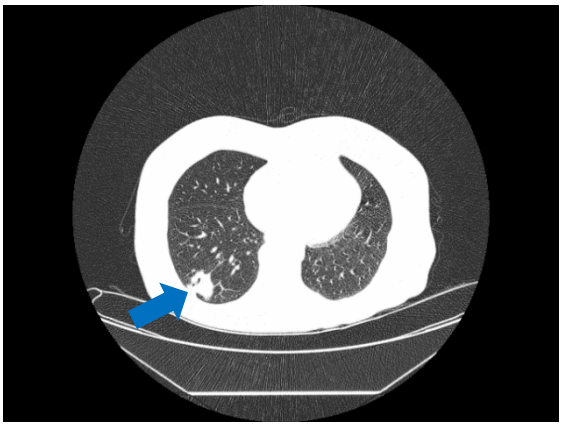
(63)



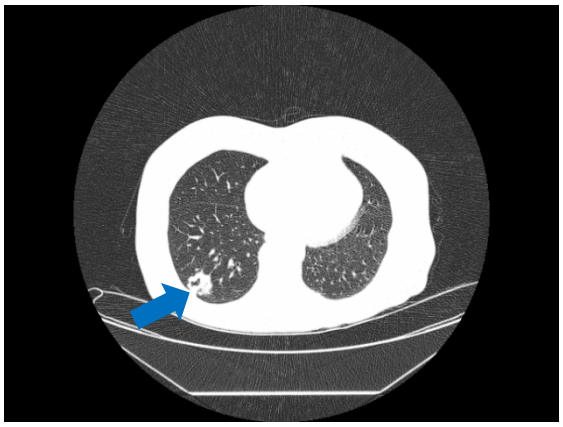
(64)



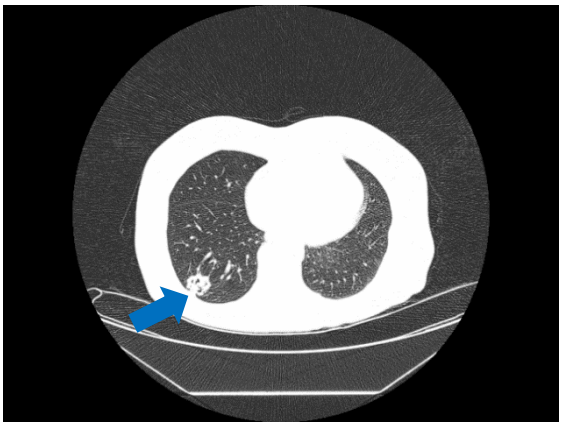
(65)



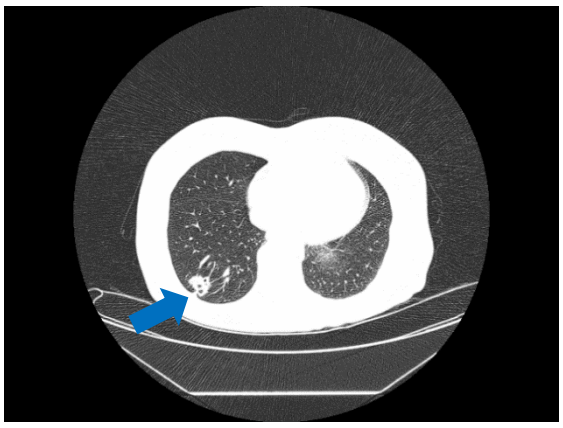
(66)



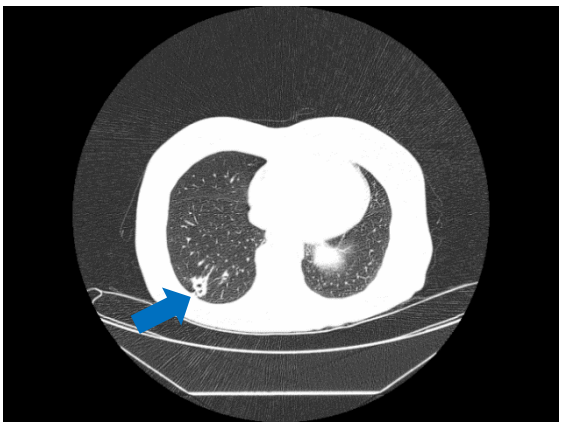
(67)



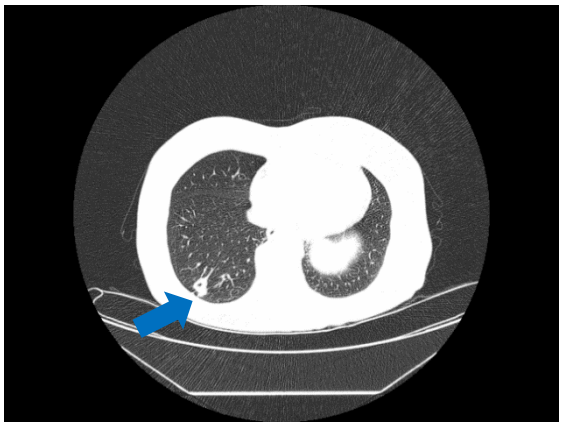
(68)



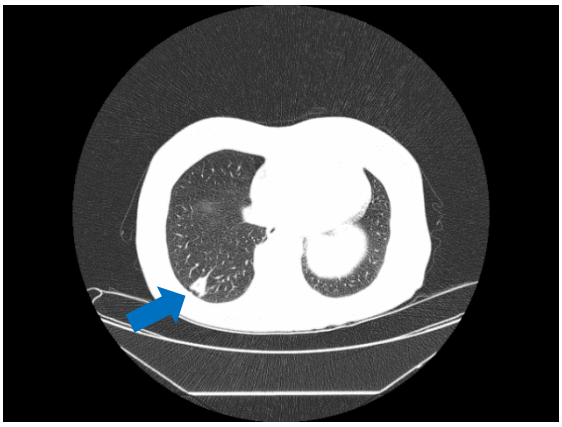
(69)



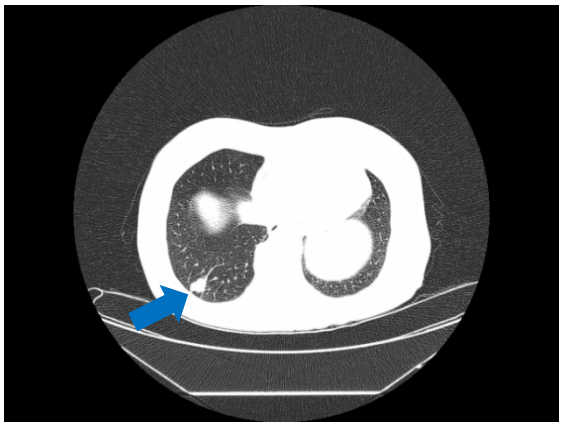
(70)



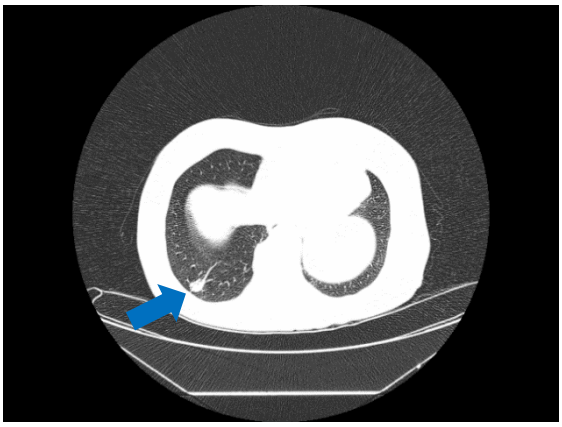
(71)



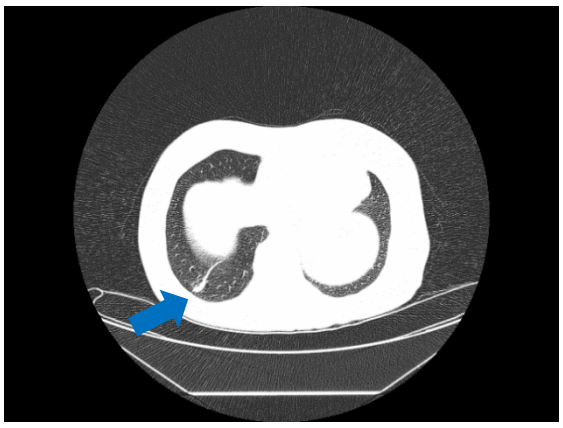
(72)



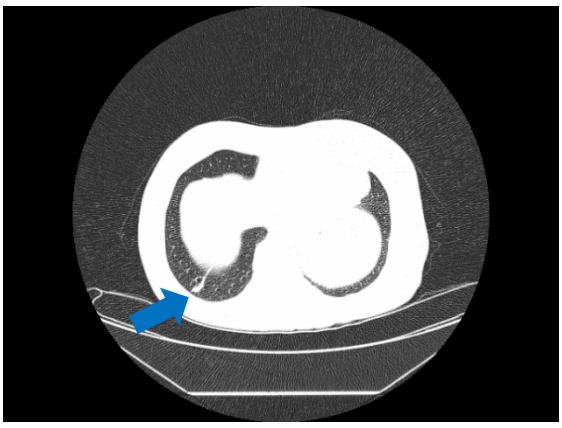
(73)



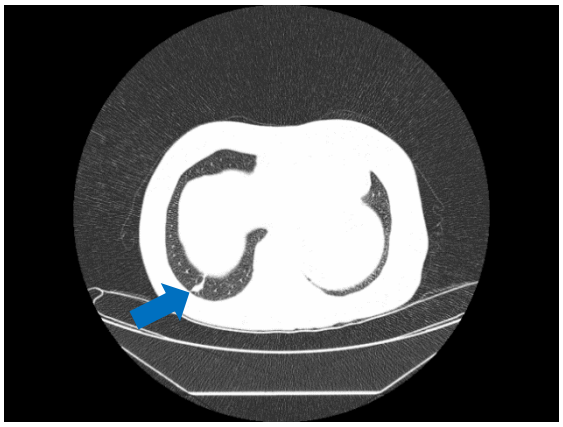
(74)



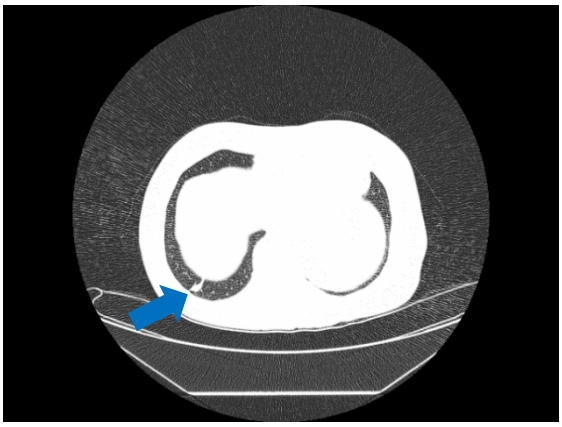
(75)



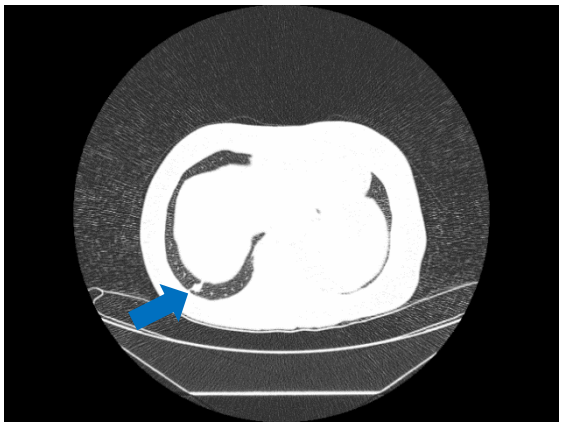
(76)



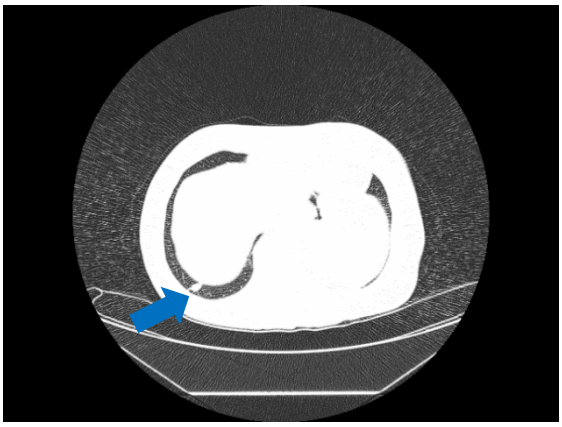
(77)



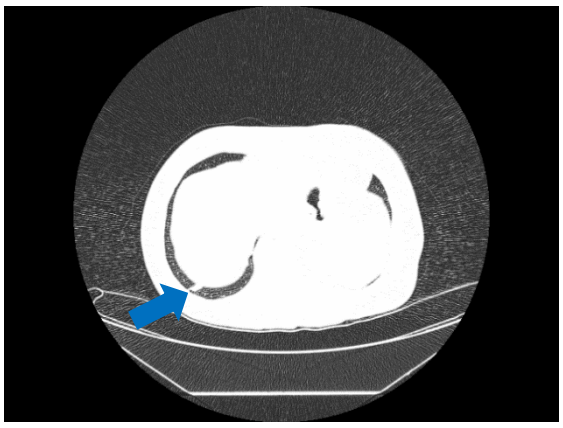
(78)



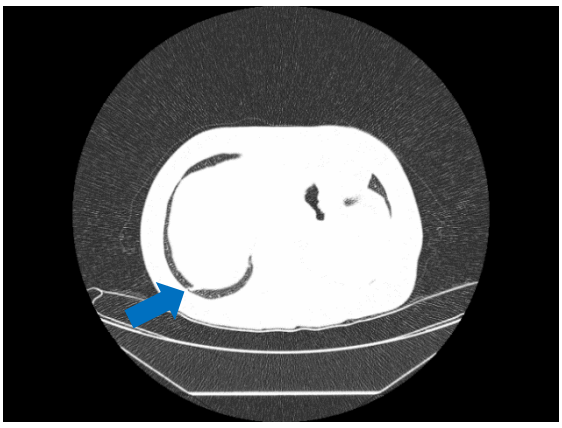
(79)



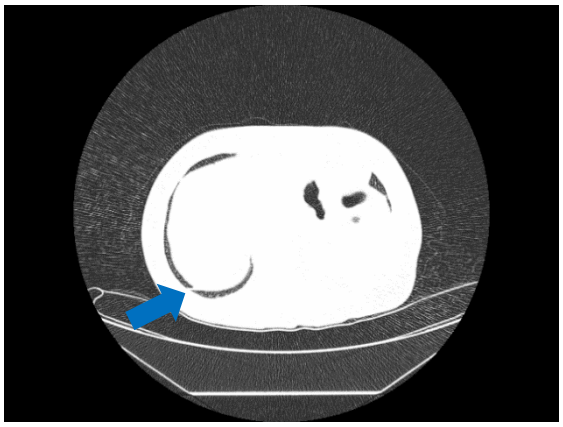
(80)



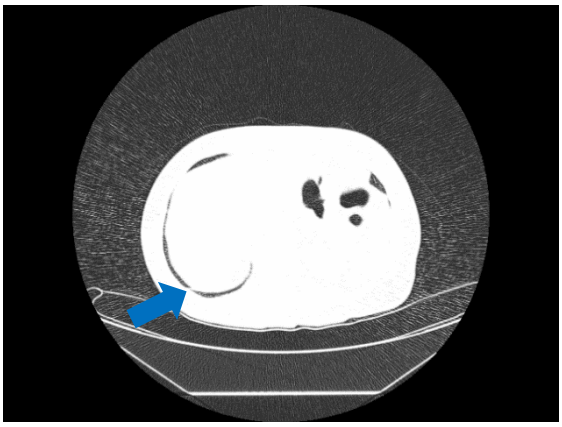
(81)



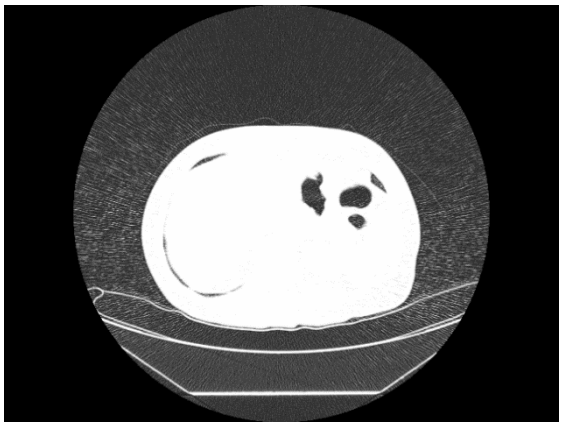
(82)



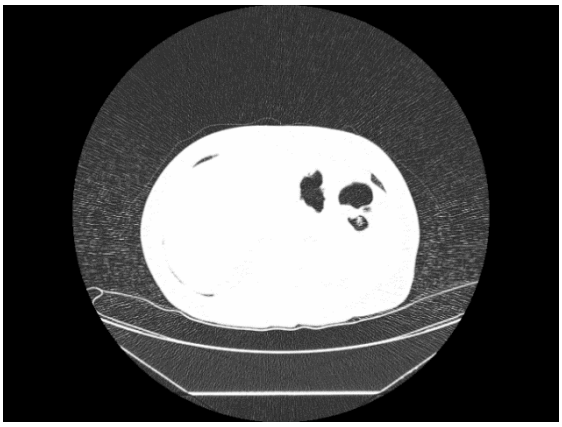
(83)



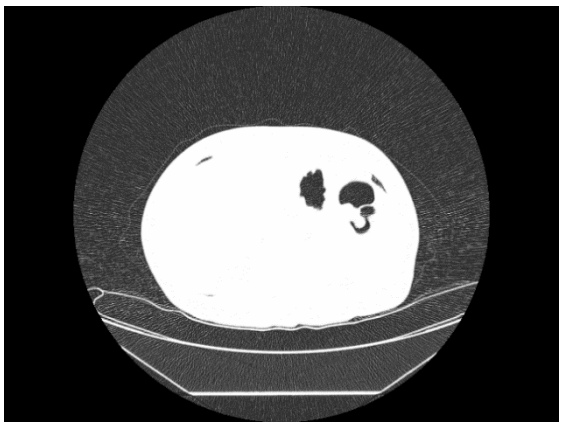
(84)



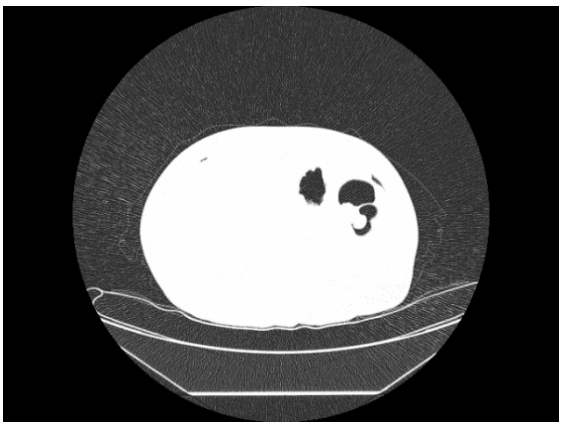
(85)



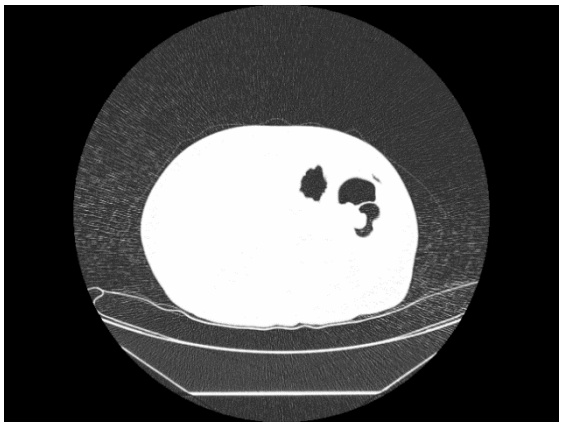
(86)



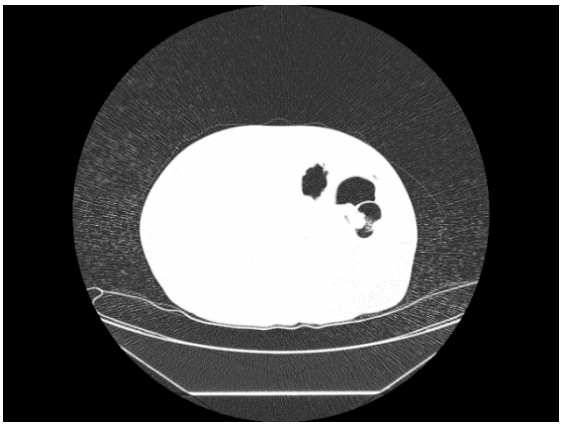
(87)



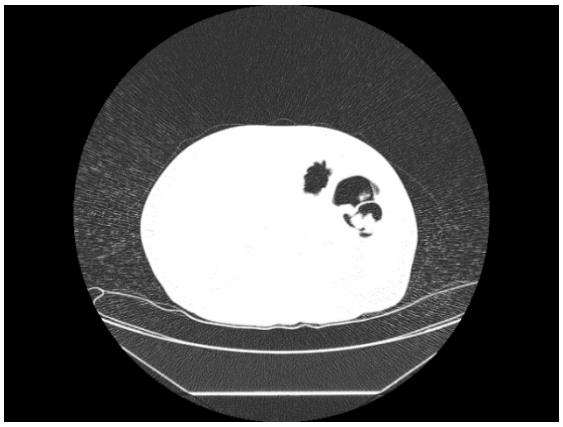
(88)



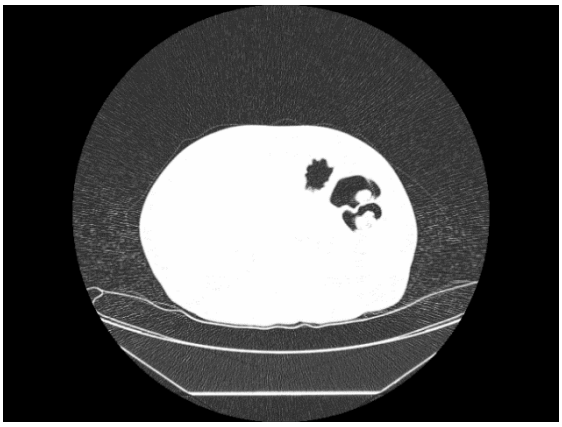
(89)



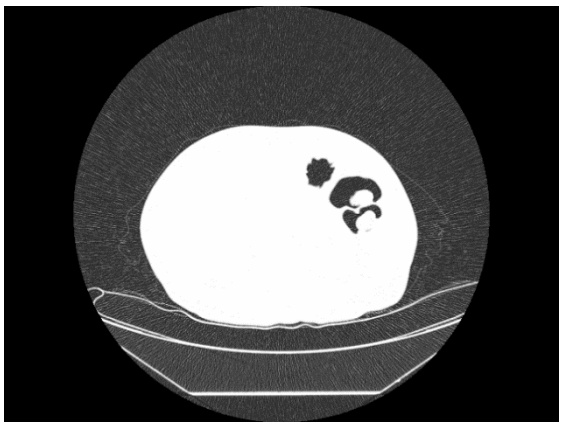
(90)



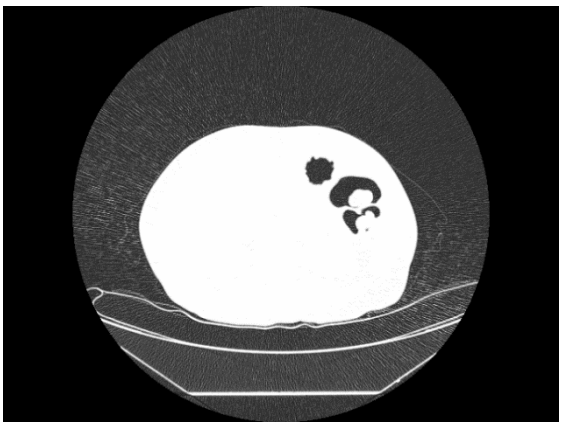
(91)



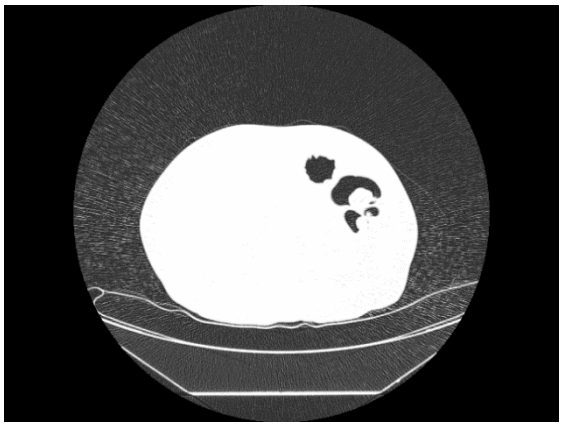
(92)



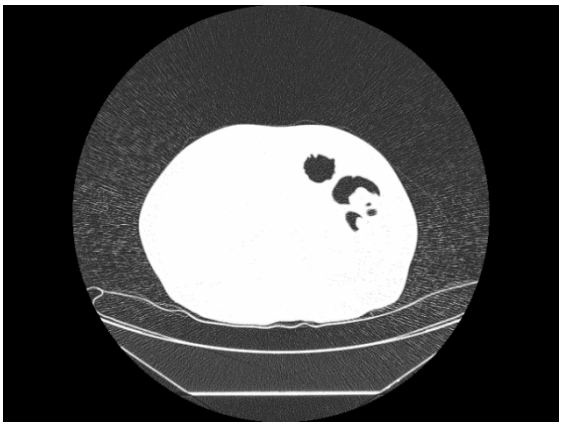
(93)



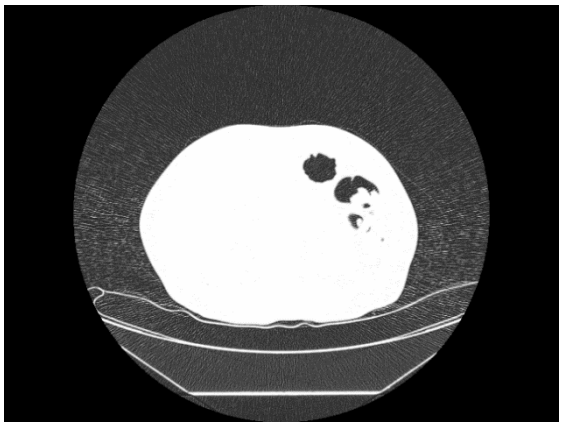
(94)



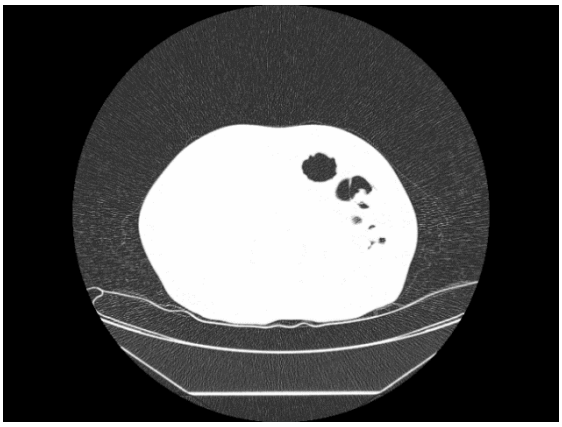
(95)



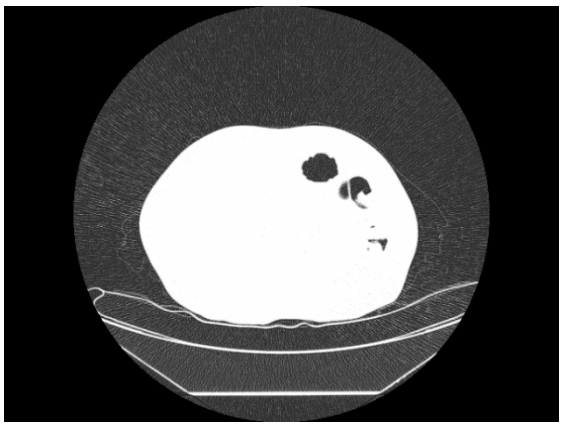
(96)



(97)



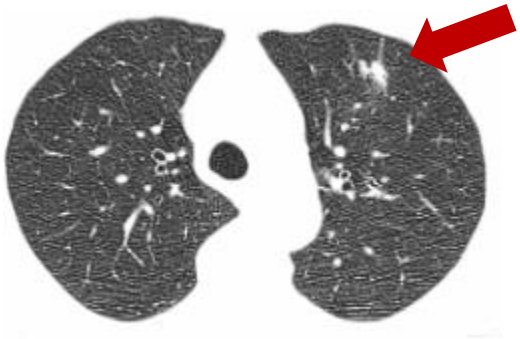
(98)



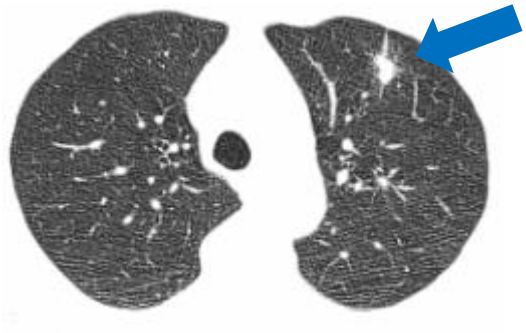
(99)

**B – Lesion having an Irregular Shape, pointed to by Arrows**

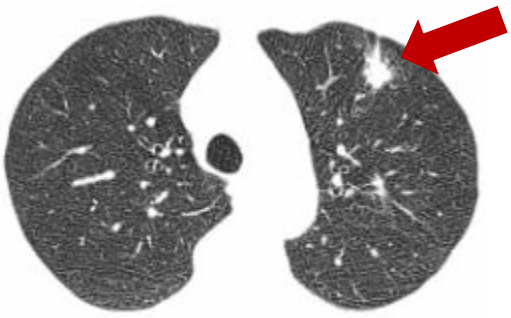
(a)



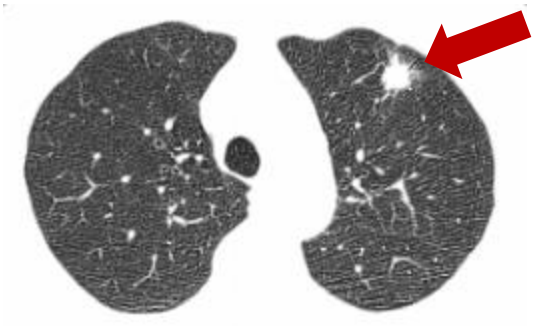
(b)



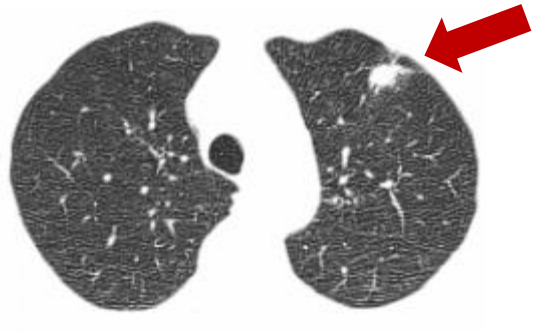
(c)



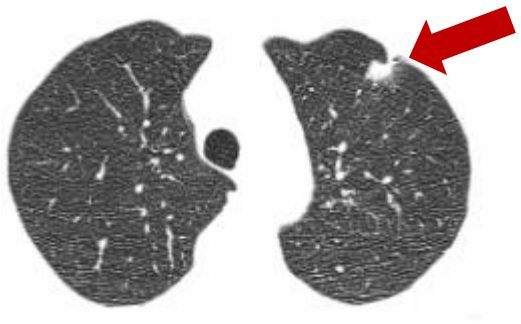
(d)



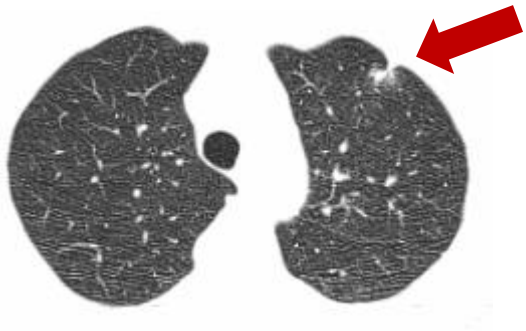
(e)



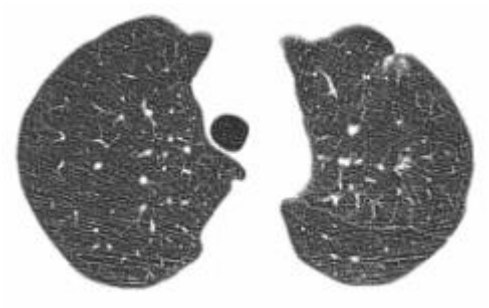
(f)



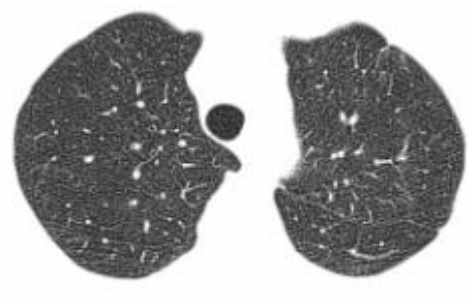
(g)



(h)



(i)



(j)

# C – Documents Pertaining to the Approval of use of CT Scans and Medical Reports from a General Hospital in Malta

## Letter of Permission Sent

**Undergraduate Final Year Project 2015/16 – Permission Letter for Use of Anonymised Data**

My name is Daniel Bonanno, and I am currently reading for a Bachelor's Degree in Computer Engineering at the University of Malta. I am presently preparing for my final year special project, entitled, *"Image Processing Techniques for Lung Lesion Detection in Chest CT Scans"*, under the supervision of Prof. Ing. Carl James Debono and co-supervised by Dr. Paul Bezzina and Dr. Francis Zarb. Attached please find the proposal for the project.

As part of the study I require anonymised CT images of the thorax. Your permission and approval is being sought to obtain anonymised Chest CT Scan images (in DICOM format) of patients and the corresponding anonymised examination reports. The images will be used in the study to develop an algorithm which identifies any lesions in the lungs whereas the examination reports will allow me to verify whether the algorithm developed works as intended.

As part of the study I will observe the highest possible ethical standards and strict confidentiality. Permission is also being sought from the Ethics committee of the University of Malta to undertake this study.

Should you require further information or clarification please contact me or any member of the supervisory team using the following details:

| | | |
|---|---|---|
| Prof. Ing. Carl James Debono | carl.debono@um.edu.mt | Tel No: 23402076 |
| Dr Paul Bezzina | paul.bezzina@um.edu.mt | Tel No: 2340 1824 |
| Dr Francis Zarb | francis.zarb@um.edu.mt | Tel No: 2340 1833 |
| Daniel Bonanno | daniel.bonanno.13@um.edu.mt | Tel No: 21657501 / 79657501 |

I thank you for your assistance in this matter.

Sincerely,

Daniel Bonanno

## Permission from the Hospital's Authorities

**Data Protection at**
to Aquilina, Buhagiar, me

27 Oct

Dear Mr Bonanno

Good Morning
On the basis of the documentation you submitted, from the data protection point of view you have been cleared to proceed with your study provided that you obtain approval from CEO and the University Ethics Committee.
Please contact Ms. Nadine Buhagiar on 2545 5334 or Ms. Graziella Aquilina on 2545 5346 to present a copy of your approvals and fill in the appropriate Data Protection Form.
Remember that in no way should you retain any personal details you obtain from your research and this should be destroyed at the end of your study.
All medical records are to be viewed at the Medical Records Department
You are requested to submit a copy of your findings to this office at the end of your study.
Regards

Sharon Young

Data Protection Officer

Daniel Bonanno
97, The Burrows
Arznell Street
Marsaxlokk
MXK 1405
ID No: 0141295(M)

16th October, 2015

Ms. Sharon Young
Data Protection Officer,

Mr. Ivan Falzon
CEO,

10/11/2015

AM 33 on behalf of Dr S. Zrinzo

Dr. Salvina Zrinzo
Chairperson, Medical Imaging Department,

Dr. Adrian Mizzi
Consultant Radiologist
Medical Imaging Department

Mr Joseph Castillo
Professional Lead, Medical Imaging Department,

Mr. Victor Micallef
Team Leader, CT Unit, Medical Imaging Department,

**Undergraduate Final Year Project 2015/16 – Permission Letter for Use of Anonymised Data**

## Data Protection Declaration

Human Resources & Administration Directorate

Telephone: 2545 4137

DECLARATION BY VISITING STUDENTS

I hereby declare that I will respect the confidentiality and privacy of any personal data or information that I might come across during my attachment at and will in no circumstance disclose any such information. I also confirm that I am aware of the provisions of the Data Protection Act and that I will abide by all Government and hospital regulations related to data.

I am also aware that I will be assigned a mentor for the duration of my attachment and will abide by the directions given by the same mentor.

Details of student

Signature:

Full name: Daniel Bonanno

ID number: 0141295 (M)

Educational Institution: University of Malta

Date: 28th October 2015

Endorsement by Parent/Guardian (where applicable)

Signature:

Full name:

ID number:

Date:

Disclaimer: will not be, under any circumstance, liable for consequential or incidental damages that might arise out of or related to the behaviour of the said student vis-à-vis the above instructions.

Data Protection Statement: The personal data supplied on this form will be processed in accordance with the Data Protection Act, XXXVI of 2001.

HRA/1043/01.0

INFORMATION MANAGEMENT & TECHNOLOGY DIRECTORATE

28/10/15

## Permission by the University of Malta's Research Ethics Committee

**Helen Grech** <helen.grech@um.edu.mt> 28/12/2015

to me, UNIVERSITY

Dear Mr Bonanno,
Following your submission of the requested documentation, I am pleased to inform you that your proposed study has been approved by UREC.
Good luck

## D – Contents of the Attached CD

The CD attached with this report contains the following:

- A soft copy of this report in PDF format
- The source code for the CAD system implemented and any additional libraries required.
- Two executables for the CAD system: a Debug and Release version
- A sample CT scan from the LIDC database
- A user manual (README file)