COMPUTERIZED DETECTION OF LUNG NODULES WITH AN ENHANCED FALSE POSITIVE REDUCTION SCHEME

by

Negar Memarian

B.Sc., University of Tehran, 2004

A thesis

presented to Ryerson University

in partial fulfillment of the

requirements for the degree of

Master of Applied Science

in the program of

Electrical and Computer Engineering

Toronto, Ontario, Canada, 2006

UMI Number: EC53524

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy
submitted. Broken or indistinct print, colored or poor quality illustrations and
photographs, print bleed-through, substandard margins, and improper
alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if unauthorized
copyright material had to be removed, a note will indicate the deletion.

UMI®

I hereby declare that I am the sole author of this thesis.

I authorize Ryerson University to lend this thesis to other institutions or individuals for the purpose of scholarly research.

_____
Signature

I further authorize Ryerson University to lend this thesis by photocopying or other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

_____
Signature

Ryerson University requires the signatures of all persons using or photocopying this thesis. Please sign below, and give address and date.

COMPUTERIZED DETECTION OF LUNG NODULES WITH AN ENHANCED FALSE POSITIVE REDUCTION SCHEME

Master of Applied Science

2006

Negar Memarian

Electrical and Computer Engineering

Ryerson University

Abstract

This thesis is composed based on the original investigations of author in the field of computerized lung nodule detection in computed tomography (CT) images. The methodologies discussed in this thesis include two main topics of region of interest detection and enhanced false positive (FP) reduction. The system, which is developed to be a supplementary diagnostic tool for radiologists, first spots all the regions suspected to be nodules in lung. Then it pins down the candidates with highest possibility of being nodule through a series of rule based filtering stages. Finally an enhanced false positive reduction system, which is in fact designed as a hybrid scheme based on learning algorithms, reduces the false positive detections further. The overall system performs with 72% sensitivity and 2.42 FP/slice, which competes with state-of-the-art methods. The system was tested on a database consisting of 24 pediatric clinical subjects with 1190 images and 154 metastatic nodules.

Acknowledgements

Dedication

To my beloved mother for all her sacrifices and backup to help me achieve my lifetime dream, which is nothing but to live and learn similar to her and my father.

Table of Contents

List of Tables

List of Figures

1 Introduction

Advancements in technology have had great impact on medical sciences. Nowadays, from examination and diagnosis steps to treatment and post-treatment stages, engineering and applied sciences are exploited to reach faster and more accurate results. The developing field of biomedical engineering is a broad area of knowledge providing the link between medicine and engineering. According to the Whitaker Foundation, which is a premier supporter of biomedical research, biomedical engineering is a discipline that advances knowledge in engineering, biology and medicine, and improves human health through cross-disciplinary activities that integrate the engineering sciences with the biomedical sciences and clinical practice [1].

One important area of biomedical engineering is the field of medical imaging. Medical imaging can be categorized into two main divisions of medical image acquisition and medical image processing. The former includes the apparatus and devices (hardware) of acquiring static or dynamic medical images, while the latter deals with algorithms and systems (software) that process, analyze and interpret imaging data. Imaging devices work with a variety of modalities, such as computed tomography (CT), radiography, magnetic resonance imaging (MRI), ultra sound, etc.

1.1 Computed Tomography

Computed Tomography (CT) is defined as a non-invasive procedure that takes cross-sectional images of the brain or other internal organs to detect any abnormalities that may not show up on an ordinary x-ray. CT simply generates images, which in fact shows sequential layers of a specific organ. When this series of images is viewed together, it represents a three dimensional view of the examined organ. While most of the other methods of imaging have the disadvantage of superimposed structures due to the projection of the three dimensional human body onto a two dimensional image, the CT images do not suffer from this problem [2].

Conventional CT involved an x-ray source that rotated around the patient, who was enclosed in a doughnut-shaped gantry. The housing was designed so the x-ray source rotated over the patient to obtain a single transverse image and then unwound to prepare for another rotation and another scan. The patient was required to suspend respiration for each scan. Then, patient and operator had to wait for realignment of the table before the next scan could begin. In recent years, with more powerful computers and higher energy x-ray tubes, a process known as helical CT has developed, which consists of continuous activation of the x-ray source and continuous movement of the tabletop through the gantry, resulting in volumetric acquisition. In other words, whereas conventional CT required a stop-start maneuver to acquire a single slice, helical CT uses similar, but much faster and nonstop, technology to acquire multiple transverse slices and volumetric results [3]. With emergence of new multi slice CT scanners, acquisition of thin slice scans in a very short period of time has become possible. In the last decade, spatial resolution achievable with state-of-the-art CT scanners has increased more than 10 fold [4].

The intensity of pixels in a CT image is expressed in Hounsfield Units (HU), which is a quantitative scale for describing radiodensity. According to this scale the radiodensity of distilled water at standard temperature and pressure (STP) is defined as zero HU. The radiodensity of air at STP is defined as -1000 HU.

1.2 CT and Lung Diseases

Computed tomography has proved to be a useful tool for diagnosis and surveillance of pulmonary abnormalities, specially lung cancer. Imaging protocols on single detector scanners typically generate about 40 images in a thoracic CT exam, while multi-slice protocols may generate 300-600 high resolution axial images [5]. New scanners, which now have a spatial resolution of less than a millimeter, routinely provide detailed images of early-stage lung cancer [4]. Figure 1.1(a) shows the chest radiograph of a patient, while figure 1.1(b) presents a visualization of the stack of CT images for a patient. As it is apparent in these images, in the case of radiograph, some details are hidden or difficult to observe due to the nature of radiography which superimposes the whole 3D structure of organ onto a 2D image. Therefore, distinguishing different structures (e.g., lung parenchyma, bones, nodules,

vessels, etc.) from a radiograph is not as accurate as when multiple cross sectional images are acquired from the lungs as in CT. Figure 1.1(c) shows the image of one section (one CT slice) of the thorax. When viewed and compared together the series of CT images can provide useful information about the location, shape and connectivity of various structures. This fact in addition to the following points can be considered as main reasons for the significant effect of CT technology on diagnosis and surveillance of lung diseases, particularly lung nodules:

1) Increase of the spatial resolution achievable with state-of-the-art CT scanners

2) The advent of multi-detector row scanners combined with gantry rotation times of less than 500 milliseconds per rotation [2].

3) Growing awareness of lung cancer screening by using a low dose helical CT protocol [6].
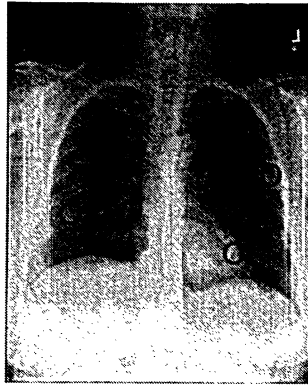
(a)
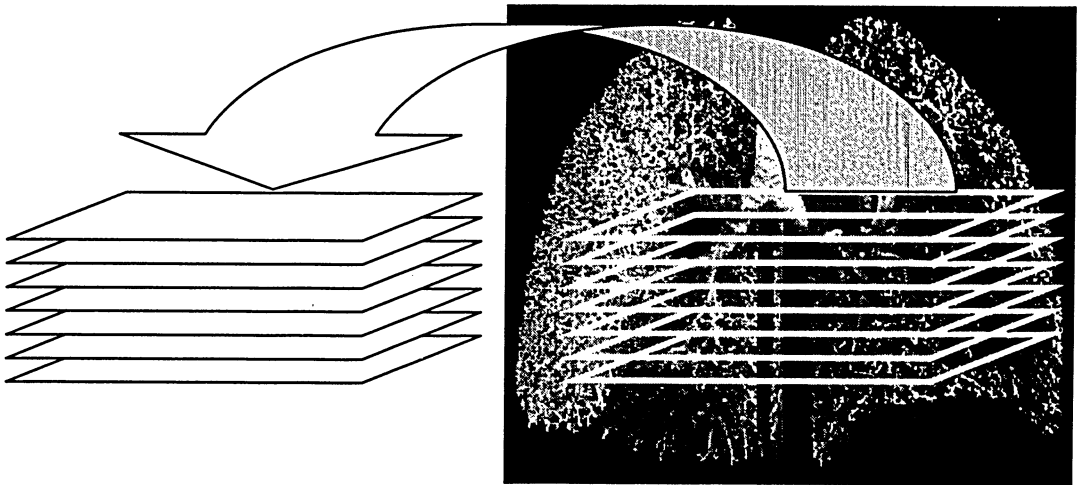
(b)
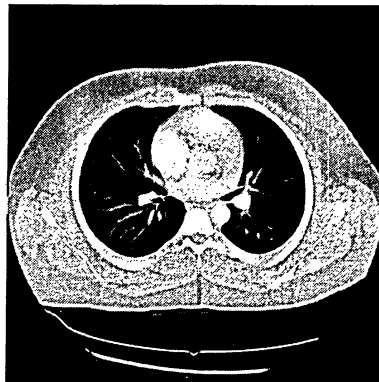
(c)

Figure 1.1: Chest radiograph and CT
(a) A sample chest radiograph. The lung region containing air is darker and traces of ribs (light curves) can be seen in that region. Some lung nodules have been marked on this radiograph with black circles. (b) Three dimensional visualization of lung and the way stack of CT slices are generated by capturing images from different depths of lung volume. (c) A sample CT image. The air-filled lungs appear in dark (about -1000 HU), while body tissue and vessels appear in lighter intensities (-500 HU and above).

Since the following sections of this thesis deal with identification and classification of lung nodules, it is important to define the term lung nodule before moving on.

1.3 Lung Nodule

It is not easy to define a single definition that includes deterministic values for size, shape and location of nodules. A couple of more prominent definitions explain nodule as:

1) Round opacity, at least moderately well margined and not greater than 3 cm in maximum diameter [7].

2) Small, approximately spherical, circumscribed focus of the abnormal tissue. [7]

While the first definition emphasizes more on the size and shape of a nodule, the second definition reflects the three-dimensional nature of the physical lesion diagnosed as nodule on the CT scans. Although these statements can be considered good descriptions for the term nodule, but the natural intricacies of biologic systems make it difficult sometimes to judge the existence of a nodule based on such definitions. It is therefore more comprehensive to apply the term nodule to a spectrum of abnormalities, which is itself subset of a broader spectrum of abnormalities, termed *focal abnormalities* [2]. As it can be seen in figure 1.2, focal abnormalities include a wider range from scars to non-specific opacities. Nodules lie somewhere in between this range.

Nodules may be solid or sub-solid (part solid and nonsolid nodules). It has been discovered by recent research that subtle nodules must be more clinically investigated, since there seems to be a higher malignancy risk associated with sub-solid nodules [8].

While different studies choose different criteria for nodule inclusion in their research, the Lung Image Database Consortium (LIDC) has decided to include only nodules (both calcified and non-calcified) with diameter not exceeding 30 mm, in their database [2]. The minimum diameter for included nodules has been set at 3 mm [7]. Nodules with less than 5 mm diameter seem to be less clinically significant. When suspicious to cancer though, nodules even smaller than 3 mm diameter would be identified and sent for further examination.

Some examples of lung nodules can be seen in figure 1.3.
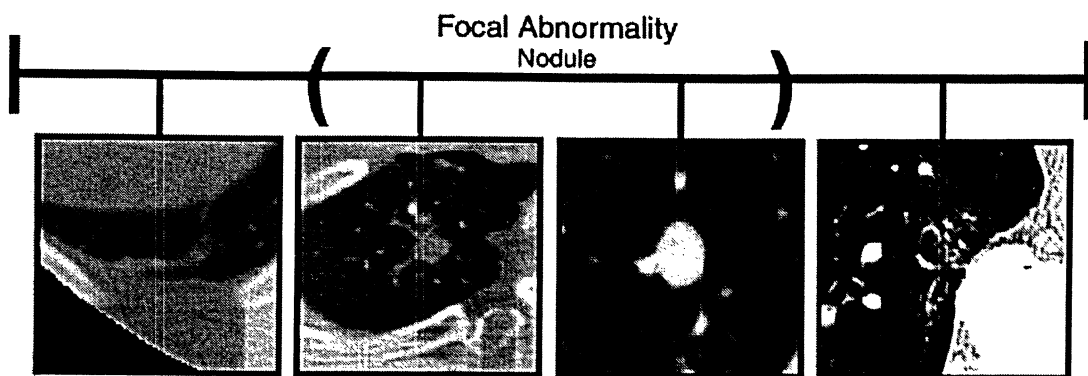
Figure 1.2: Spectrum of term nodule
The term *nodule* refers to a spectrum of abnormalities that is itself part of a larger spectrum of abnormalities termed *focal abnormalities*. This scar (far left) is a focal abnormality that would not be considered a nodule, as is the nonspecific opacity (far right), which is likely due to previous infection. The spiculated lesion (left-center) and the compact lesion (right-center) are two representative examples of lesions within the nodule spectrum. All images were obtained from transverse CT examinations [2].
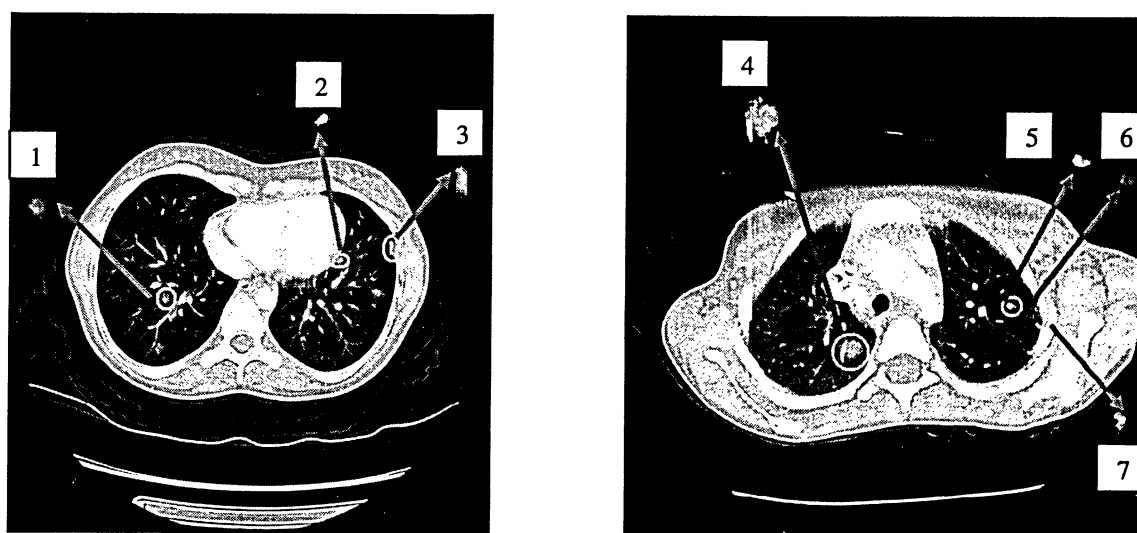
Figure 1.3: Examples of a variety of nodules
Object 1 is a round distinguishable nodule. Objects 2, 3 and 7 are examples of nodules attached to lung wall (juxtapleural nodule). Object 4 is example of a large interior nodule and object 5 is a small but not quite circular nodule. Object 6 is a pale nodule hardly recognizable from background.

1.4   Computer Aided Diagnosis Systems (CAD)

CT, specially multi-slice exam protocols, provide the radiologists with large volume of data, which is both time-consuming and laborious when manually assessed. This calls for

emergence of Computer Aided Diagnostic (CAD) systems. CAD has been defined as a diagnosis made by a radiologist with the benefit of information generated by computerized image analysis [9]. Computerized detection of lung nodules is now an important area of research, with the ultimate goal of enhancing the diagnosis made by the radiologist. Development and enhancement of lung nodule CAD systems have been among the popular areas of medical imaging research in recent years.

1.5 Contributions

The major contributions of the author in this thesis are as follows:

- A multi-stage and target-specific region of interest (ROI) detection system for identification of lung nodules in CT images is developed. The system recognizes and deals with airway tube cross sections, elongated in-plane vessels and soft tissue, and cross sections of in-depth vessels specifically and independently.

- Performance of the proposed ROI detection system is fast and with minor computational complexity. The system is time efficient and the result data it generates is easy to store and restore. These are significant advantages for a CAD system, which is meant to be used in clinical environments.

- New features are designed and extracted in the feature extraction stage.

- A novel false positive reduction scheme, which classifies pulmonary data, based on a combination of learning algorithms, is developed. The hybrid structure and adaptive nature of this scheme make it an absolutely new method for FP reduction.

- FP reduction is achieved more effectively because various features (multidimensional data) are converged into a single classification decision, in a way that misclassification errors are minimized.

1.6 Outline

This thesis is composed based on the original investigations of the author on two main topics of (a) detection of regions of interest (ROI) and (b) false positive reduction. There have been two major initiatives for the undertaken research:

- Developing a computerized lung nodule detection system
- Developing an effective false positive reduction scheme to enhance the performance of any rule based lung nodule CAD system.

After thoroughly reviewing the literature on the concept of lung nodule false positive reduction in chapter 2, chapter 3 explains a system designed by the author for detection of lung nodules (ROI detection). A novel learning FP reduction scheme which has also been proposed and implemented by the author will be discussed in chapter 4. A detailed presentation and discussion of results will be made in chapter 5. Finally the conclusions and future works will be presented in chapter 6.

2 Literature Review

2.1 Main Challenges in CAD

The ultimate goal of computerized lung nodule detection systems is enhancing the diagnosis made by radiologist. The main challenges in developing a useful CAD system are achieving:

1. High true positive (TP) detection rate
2. Low false positive (FP) detection rate

The first point refers to the fact that for a CAD system to be considered beneficial to radiologist, its overall detection results must include most of (ideally all) those regions that are truly nodules. If the CAD is incapable of highlighting true nodules for the radiologist (low sensitivity performance), then there would be little motive for its application.

The second point has also been a major challenge for researchers and developers of lung nodule CAD systems. Existence of large number of false positives in results would slow down the diagnosis process for the radiologist. The ROI (regions of interest) detection stage in CAD systems normally marks up a substantial number of candidates as potential nodules. However, not all of those candidates are true nodules. They can be traces of in-plane or in-depth vessels, bones, heart, diaphragm, or other lung abnormalities except lung nodules.

2.2 Overview

Identification of potential nodule candidates is a primary step in any lung nodule CAD system, but reducing the number of false positive objects and pinpointing the most suspicious regions on CT is perhaps the more critical part of work. Therefore the main emphasis of this literature review chapter will be on the broad concept of FP reduction. While ROI detection methods discussed in the literature will be briefly introduced in each section, more detail will be provided about the false positive reduction procedures investigated by different researchers.

2.3 False Positive Reduction Methods in the Literature

FP reduction is often applied as a post processing step to the results produced by a computerized ROI detection system. The FP reduction methods published in the literature can be categorized into four main groups. In the following sections of this chapter, examples of research from each category will be discussed.

2.3.1 Rule Based

In this method, false positive reduction is accomplished through setting certain rules (conditions) and testing each candidate in the ROI for those conditions. The rules are often set for one or more features of the candidates, such as geometrical and intensity related features.

An early attempt for FP reduction through rule based feature analysis was carried out by Kanazawa *et al.* [10]. After segmentation, fuzzy clustering and surface curvature analysis (for removing those artifacts that are near the circumference of the lung fields), total of 8 features were extracted for ROI candidates. The features were area, thickness, circularity, gray-level, variance of gray-level, localization, variance of gradient, and distance from the lung wall. Then several diagnostic rules were set that put conditions (thresholds) on some of the extracted features. The candidates were eliminated if they satisfied those conditions. The main parameters checked by the rules were ROI thickness, circularity and existence of bone in the adjacent slice images in the location of the current ROI. Total of 230 true nodules were detected by physicians in the dataset of [10] and their suggested algorithm resulted in sensitivity of 90% with 8.6 FP/case. The sensitivity performance of CAD was compared with diagnosis sensitivity of three physicians. When only the results in which the three physicians were in agreement were regarded as correct, the sensitivity of CAD system was above that of the expert physician. A major drawback for system of [10] however is that it only targeted nodules larger than a 4 mm diameter circle.

Zhao *et al.* also attempted rule-based FP reduction in their research work for automatic detection of small lung nodules [11]. In their work, three ratios were calculated for each ROI candidate. These ratios consisted of shape and size parameters associated with each

candidate. The defined rules ensure that the remaining objects have relatively compact shapes in both three dimensions and two dimensions and fall into the size range of nodule. There were a total of 266 simulated small nodules added onto eight chest CT scans, each scan having 60 to 80 slices. The final detection sensitivity of system discussed in [11] is 84.2% and the average number of false positives per case was reduced to 5 (ranged 1 to 9). The algorithm, without any modification, was directly applied to a clinical subject with four actual small lung nodules. Three out of the four nodules were successfully detected by the algorithm (75% sensitivity). There were nine false positive results, and one nodule was not detected [11]. However this algorithm performed poorly when it was tested on other datasets of clinical subjects [12].

In [13] three dimensional shape and connectivity analysis was accomplished to reduce FP lung nodule detections at low-dose CT. In order to classify vessels (major source of FPs) and nodules, the shape and gray value of each nodule candidate was considered. 3D CT data was reconstructed from each CT slice handling only the vessel related data. Nodule candidates were extracted by 3D region growing, and each was labeled using the connected component labeling technique. For each candidate, volume, elongation factor, and compactness were calculated. These parameters characterized the 3D shape of a nodule candidate. Based on the knowledge that nodules are sphere-like and highly compact, while blood vessels are tube-like and much less compact, the classification was automatically performed and regions labeled as nodules were highlighted. However the authors of [13] have not explained about how they set the rules (thresholds) on the shape, size and compactness of objects. This method also had some other pitfalls including its poor sensitivity performance when relatively thick section CT data was used. Thus, only nodules greater than 5 mm in diameter were included in performance analysis. Even with such nodule size range, the sensitivity of system was only 65% with 8 ± 5.2 false-positive CAD results per CT study.

In a recent research on 5 mm CT slices [14], FPs were reduced by using unique features such as vessel and lung wall connectivity, a modified bounding box and 3D compaction to compensate for partial volume artifacts due to thick CT slice. The dataset used in that research consisted of 239 true nodules, with 41 nodules smaller than a 4 mm diameter circle. Minimum size of nodule which was detected by system of [14] was 2.1 mm in

diameter. This is counted as a plus for that system, as there are few CAD systems in the literature that are capable of detecting very small nodules, i.e., nodule diameter less than 3 mm. The algorithm resulted in 80% sensitivity and 3 FP/slice. Further false positive reduction was suggested to be the first action to be taken for improvement of that system.

Rule based techniques are sometimes carried out together with another method to reach more effective FP reduction results. A popular paradigm found in the literature is the combination of rule based and linear discriminant analysis [15-18]. Those examples will be discussed in the following section.

2.3.2 Linear Discriminant Analysis

Linear discriminant analysis (LDA) is a supervised learning method which predicts in which group an individual belongs (classification) based on a pattern learnt from labeled training data. The goal of this mathematical tool is to develop optimal classification i.e., minimizing the total probability of misclassification and minimizing the costs of misclassification. A number of examples of applying LDA for lung nodule FP reduction can be found in the literature [15-18], [20].

The group of researches in University of Chicago has been among the very active groups in the field of computerized detection of pulmonary abnormalities. They used different types of linear discriminant analysis for the purpose of FP reduction in a series of their studies [15-17]. In [15], after ROI detection through segmentation, cumulative gray level analysis, histogram analysis and multi level thresholding (MLT), substantial FP reduction was achieved by using linear discriminant analysis. Total of nine 2D and 3D features were computed for each nodule candidate. That included six geometric features (volume, sphericity, radius of the equivalent sphere, maximum compactness, maximum circularity, and maximum eccentricity), and three gray level features (mean gray level within the structure, standard deviation for the gray level, and the gray level threshold at which the volume of the structure first decreased below the upper volume bound). The values of these features for all nodule candidates were analyzed by the LDA classifier. A leave-one-out scheme was chosen, in which the LDA classifier was trained by all except one nodule candidate and the omitted candidate was subsequently used to test the trained classifier. The

process of training and testing the classifier was independently repeated until all nodule candidates had been used as the left out candidate for testing. The database in [15] comprised a total of 17 patients (493 sections) with 187 pulmonary nodules. The effective diameter of the nodules ranged from 3.1 to 27.8 mm. The system resulted in an overall sensitivity of 70% with an average of 3 FP findings per section. Although the FP figure is not very eye-catching, it corresponded to an 89% reduction in the number of FP findings after the application of LDA.

In a continuation of their research, the same group utilized LDA in conjunction with rule based FP reduction [16]. The huge pool of ROI candidates produced by the MLT stage was filtered by testing for three rule based conditions: eccentricity, pixel value standard deviation and cross-sectional area. The intension of this step was to eliminate "obvious" false positive detections. After the rule based scheme was applied, the remaining candidates were input to an automated LDA classifier. The feature vector consisted of the same nine features used in [15] but a leave-one-out-by-case analysis was performed this time, in which the nodule candidates from 42 cases were used to train the classifier, which was then applied to the nodule candidates from the remaining case (patient). This process was repeated until scans of all 43 patients in the database had been used as the left out case. The 171 true nodules in the database ranged between 3 to 25 mm in size. The automated method yielded an overall nodule detection sensitivity of 70% with an average of 1.5 FP/section. The study also presented statistics about the dependence of nodule detection sensitivity on nodule size at a fixed FP rate of 1.5/section. When only those nodules in the database with effective diameter of 5 mm or greater were considered, 80% of the nodules were accurately detected by the automated method. Effective diameter was defined as the diameter of a circle with area equal to the maximum area of a nodule in any section in which it appears.

In a lung cancer screening program [17], Armato *et al.* once again applied a combination of rule based and discriminant analysis methods for FP reduction, however this time jackknife model LDA was used. In the jackknife approach, features extracted from half of the database were used for training the linear discriminant function and the other half of database was used for testing. The nodule candidates that emerged from the LDA with values below a specified LDA output value were considered detected nodules. The overall performance of the method was obtained from the average performance of ten random

partitions of the nodule candidate set into training and testing subsets. An overall detection sensitivity of 70% (330 of 470) was attained with an average of 1.6 FP/section. The effective diameter of true nodules ranged between 3 to 25 mm and nodules with effective diameter between 5 to 7 mm were most frequent. In that study, effective diameter of each nodule was computed as the mean bounding region dimension (one-half of the sum of the short and long axis lengths). When a nodule was present in more than one section, the bounding region with the greatest area was used for the effective diameter calculation. This study discussed the relative performance of the method in detection of lung nodules in various malignancy status, size, subtlety, and radiographic opacity categories. The overall performance of the automated lung nodule detection method decreased for nodules of increased subtlety and for nodules with an increased non-solid component.

Gurcan *et al.* [18] also applied linear discriminant analysis for FP reduction both with and without a rule based FP reduction step. They used a stepwise [19] method to extract 14 features, including six 3D and nine 2D features. The 3D features were volume, surface area, average gray value, standard deviation, skewness and kurtosis of the gray value histogram. In addition, area, perimeter, circularity, compactness, major and minor axes and their ratio, and eccentricity of a fitted ellipse in each cross section of the 3D object were calculated. The maximums of these values were used as the object's 2D feature values. In the first stage of FP reduction a rule based analysis was performed that categorized the candidates into nodules and non-nodules. The rules were regarding the eccentricity (2D), area (2D), dimension of bounding box (3D), maximum circularity (3D) and relation of object location and object size (3D). After rule based classification, LDA classifier was used to further reduce the number of FP objects. The most discriminating features were determined to be the volume, surface area, average and standard deviation of gray values and the maximum values of the cross sectional area, perimeter, circularity, major and minor axes, and eccentricity. The LDA was designed using a leave-one-case-out training and test resampling scheme. Outcome of applying only the LDA classification on a dataset of 34 patients with 63 true lung nodules was 84% sensitivity with 5.48 FP objects per slice. When the rule based classification was also used, the FP rate decreased to 1.74 objects per slice at the same sensitivity. This revealed the effectiveness of the rule based stage in the FP reduction scheme. However it should be noted that the rules were designed based on knowledge gained from radiologists such as the

anatomical characteristics and morphological features of lung nodules (size, location and shape). The size of nodules in the dataset of [18] ranged between 2 to 25 mm with the average size of 8.89 mm (diameter).

In another work [20], after ROI detection by means of applying the quantized convergence index (QCI) filter on CT image pixels, Fisher line discriminant analysis was used to categorize the candidates into groups of true positives and false positives. Five features were involved which consisted of effective diameter, elongatedness, goodness-of-fit with an ellipse, target contrast-to-noise ratio, and perimetrical gradient convergence. The Fisher line discriminant analysis of the five feature quantities generated the final binary decision of the scheme, i.e., either 0 ("no nodule") or 1 ("nodule"). The coefficients for the discriminant function were obtained using the datasets used in a study by Matsumoto *et al.* (personal communications, June 2003), which contained 50 nodules as defined by the consensus of three radiologists. The threshold value for the discriminant function was set such that the scheme had a nodule detection sensitivity of 80% for the above dataset. The suggested CAD system identified 81 loci of potential nodules consisting of 42 "true nodules" and 39 "false nodules" in five clinical CT datasets (four cases) ranging from 2.5 mm to 12 mm in size. The 81 potential nodules were then evaluated in 2 dimensions by four radiologists. Mixture distribution analysis of the results of the four radiologists demonstrated a relative proportion agreement of 0.84. The kappa statistic was used to compare the agreement of the computational scheme with the results of the four radiologists. Kappa statistic is an index which compares the agreement against that which might be expected by chance. Kappa can be thought of as the chance-corrected proportional agreement, and possible values range from +1 (perfect agreement) via 0 (no agreement above that expected by chance) to -1 (complete disagreement).The kappa value of 0.65 calculated in [10] was shown to be significantly different from chance.

In addition to the purpose of FP reduction, discriminant analysis has been employed with the goal of classifying benign and malignant nodules. For instance the work by Shah *et al.* [21], investigates the utility of a computer-aided diagnosis in the task of differentiating malignant nodules from benign nodules based on single thin-section computed tomography image data. For each ROI contour/outline of the solid portion of the nodule, image analysis techniques were used in order to extract mathematical descriptors of attenuation, size, shape,

and texture. This was followed by feature selection by a stepwise model selection search [22]. The selected features were then input into four different classifiers. These classifiers included linear discriminant analysis, quadratic discriminant, logistic regression, and rpart, which closely followed the CART algorithm [23]. Receiver operator characteristic (ROC) analysis revealed that the best performing classifier when tested using a leave-one-out method was the linear discriminant analysis classifier with an A_z (area under ROC curve) of 0.92. There were a total of 33 benign and 48 malignant true nodules in the dataset. Most of these nodules had diameter bigger than 12.5 mm and there was only one nodule with diameter less than 7.5 mm in the dataset. This is a drawback of this system, as it does not include any small nodules.

2.3.3 Artificial Neural Networks

An Artificial Neural Network (ANN) is an information processing paradigm that is inspired by the way biological nervous systems, such as the brain, process information. The key element of this paradigm is the novel structure of the information processing system. It is composed of a large number of highly interconnected processing elements (neurons) working in unison to solve specific problems [24]. The ability of neural networks to accumulate knowledge about objects and processes using learning algorithms makes their application in pattern recognition very promising and attractive [25]. Various kinds of ANNs have been developed for a variety of applications. Reviewing the literature reveals that a few groups challenged the problem of lung nodule FP reduction through application of artificial neural networks, e.g., [26-29].

An early attempt for lung nodule detection and FP reduction through the use of ANN was made in [26]. Two level artificial neural networks were utilized for computer-aided diagnosis of lung nodules in chest radiographs. First ANN level was for detecting suspicious regions within the lung (ROI detection). Second level of ANN was intended for FP reduction. After ROI detection, each suspected nodule area was transferred to a curvature peak space and then introduced to the second ANN for classification. The ANN composed of multilayer perceptron structure with hidden layers. The system was tested by a series of simulated and real nodules and provided results of 89%-96% sensitivity and 5-7 FPs/image, depending on

the size of the nodules. Size of 50 real true nodules in the database ranged between 8 mm and 20 mm, for which the CAD provided 94% sensitivity and 5 FP/image. No result was reported for small real nodules.

A neural fuzzy model was designed in [27] to extract suitable diagnosis rules and classify true nodules from false candidates in CT images. Three features (average brightness, size of area and circularity) were input to a multilayered neural network system. The model suggested in [27] was organized layer by layer including input layer, rule inference layer, and defuzzification layer. The learning mechanism and high computational power of neural networks were brought into fuzzy inference system. The back-propagation algorithm was used as a learning tool in that fuzzy inference system. That model can revise the membership function of each feature and the corresponding weights of defuzzification by itself. The system was applied to 29 clinical cases containing 393 physician confirmed nodules and resulted in 89.3% sensitivity and 0.3 FP/image. Although the result numbers seem promising, there are two downsides associated with that system. The threshold used for differentiating between nodules and non-nodules (false positives) in the defuzzification layer is a fixed number. There is no guarantee that this number will perform as efficiently for other test datasets as it did for the original dataset used to achieve the experimental results. The other disadvantage is that the size of nodules that the CAD system can detect is between 1 cm and 5 cm in diameter within lung field. This conveys the fact that the suggested performance of CAD system corresponds to only when big nodules are present in the dataset.

In [28], Suzuki et al. used massive training artificial neural network (MTANN) for reduction of false positives in computerized detection of lung nodules in low-dose CT images. The MTANN consisted of a modified multilayer ANN and is trained by a large number of subregions extracted from input images together with a teacher image containing the distribution for the "likelihood of being nodule". In that method, the original image including a nodule or non-nodule was divided pixel by pixel into a large number of overlapping subregions. All pixel values in each of the subregions were entered as input to the MTANN, whereas a pixel value of each single pixel from the teacher image was used as the teacher value. In order to eliminate various types of non-nodules, the capability of each MTANN was extended and a multiple MTANN was developed. Each MTANN acted as an expert for distinction between nodules and a specific type of non-nodule. The output of each

17

MTANN was integrated by the logical AND operation. The discussed structure was applied both with and without the presence of a series of rule based and linear discriminant classifiers in the FP reduction system. The results are summarized in Table 2.1. It can be seen that in terms of FP/section, best result was achieved when all three stages (rule based, LDA and MTANN) were employed. Yet the FP rate was not similarly low when only the neural network based component (MTANN) was used. Nodule size in [28] ranged from 4 mm to 27 mm. Independent nodules were used for training an testing of the LDA stage and the mean diameter of the 50 nodules used for training was 12.7 ± 6.1 mm while the mean diameter of the 71 nodules in the test set was 13.5 ± 4.7 mm.

Stages of FP reduction	Sensitivity	FP/slice
Rule based + LDA + MTANN	80.3%	0.18
Rule based + LDA	81.7%	0.98
MTANN	80%	1.85

Table 2.1: Sensitivity and FP/slice results acquired in [28]

Recently authors of [28] applied a similar method to chest radiographs rather than low-dose CT [29]. They challenged the issue that chest radiograph is a projection image whose characteristics differ from those of an axial image such as that acquired by CT (e.g., a nodule can overlap with ribs in a chest radiograph, whereas a nodule and ribs are separated in a CT image). A couple of preprocessing steps such as background trend correction and contrast normalization were added to the system discussed in [28]. They finally obtained an overall sensitivity of 81.3% and 1.4 false positives per image. The dataset contained 91 solitary pulmonary nodules and the size of nodules ranged from 8.9 mm to 29.1 mm.

2.3.4 Other Mathematical/Statistical Methods

In addition to the main categories of FP reduction techniques mentioned in earlier sections, examples of other mathematical/statistical methods can be found here and there in the literature. The following will point out a few of those methods.

A computer-based algorithm was suggested in [30] that combined traditional image-processing techniques with rigorous pattern-recognition and receiver-operator characteristic curve analyses. The goal was to select an appropriate, but adjustable, sensitivity/FP operating point. For false positive reduction, the classifier module used a multi-feature quadratic classifier based on Eigen value and gray-level analyses. The following features were used:

- The ratio of minimum and maximum Eigen values of the covariance matrix of the pixels making up each candidate nodule
- The maximum Eigen value of the covariance matrix
- The average gray value of the pixels in the candidate nodule

The Eigen value features were used to distinguish (and eliminate) from true nodules any long, thin structures more indicative of bronchial false positives. The average gray level feature was used to remove false positives that were either brighter or darker than typical nodules. The entire set of candidate nodules emerging from the detection module was divided into true and false classes. Each class was modeled as a multivariate Gaussian distribution. Then, using the probability density function of those multivariate Gaussian distributions, the log likelihood ratio (LLR) for every detection was found. A threshold was then applied to the LLR values to determine which nodules should be rejected as false positives, creating a quadratic decision boundary in the multi-feature space. An optimal value for LLR could then be selected, using receiver-operator characteristic (ROC) analysis. A total of 12 three dimensional image sets of lung CT images were examined in this study. An experienced chest radiologist identified 145 nodules by examining all slices of all images. The nodule sizes expressed as the radii of equivalent spheres ranged between 4.43 mm and 18.02 mm. At a sensitivity of about 90% a corresponding FP rate of 2 false positives per CT slice was obtained. The authors of [30] claimed that the quadratic Bayesian classifier employed in their technique could more accurately determine which features are important. They also believed that due to the fact that their technique used far fewer features than other algorithms, it has a greater likelihood of being generalizable to a larger dataset [30].

In a work by Dehmeshki *et al.* [31], after morphology based techniques for ROI detection, 12 features are extracted and selected. Based on those features, the dataset is divided into two groups of nodules and non-nodules. Then for purpose of accurate

classification (FP reduction) a sampling mechanism was used to overcome the imbalanced data problem by reducing the non-nodule samples without compromising the classification ability. After that, a support vector machine (SVM) was trained for nodule classification. In the sampling stage, k-means clustering was used to group all non-nodule samples into a number of clusters. Then a Gaussian mixture model was used to improve the cluster distribution for a more accurate clustering performance. The support vector machine classifier constructed a hyperplane as the decision surface in such a way that the margin between two classes was maximized. The dataset used in that research contained 85 solid nodules with maximum intensity higher than -650 HU in 47 CT scans, detected by at least one radiologist. No information about the size of nodules was expressed. According to [31], no true nodule was missed in the midst of data classification, however since 4 large true nodules were missed by the ROI detection step, the average sensitivity of overall system was about 95% with 39.83 FP/scan or 0.27 FP/slice. The article asserted that the SVM has superiority over most other classifiers including neural networks because while NN only find a local solution to the problem of maximizing the margin between two classes, SVM finds a unique solution.

Finally in [32], logistic regression model and linear combination of Gaussians (LCG) with positive and negative components were used to eliminate false positive nodules. After detecting potential lung nodules using deformable 2D and 3D templates (a combination of cross-correlation template matching and genetic optimization algorithm), the post-classification of the candidate nodules was performed with the following three textural and geometric features of each detected nodule: (a) radial non-uniformity of its borders, (b) mean gray level over the 3D or 2D nodular template; and (c) the 10%-tile gray level for the marginal gray level distribution over the 3D or 2D nodular template. To distinguish between the false positive nodules (FPNs) and true positive nodules (TPNs), a supervised Bayesian classifier was used on a training set of false and true nodules. The three features (a)-(c) were used to classify the FPNs in the lungs while only the last two features were applied to the lung wall nodules. The density estimation required in the Bayes classifier was performed, for each feature, using logistic regression model and linear combination of Gaussians (LCG) with positive and negative components. The technique was applied to a dataset of 200 subjects, which contained a total of 30 true nodules and yielded an overall correct detection

rate of 82.3% with 9.2% FP rate. A major step that can improve the system of [32] is selection of features that distinguish between small lung nodules and normal objects.

2.4 Summary

In the previous sections of this chapter, the schemes in the scientific literature for computer-aided reduction of lung nodule false positive detections were categorized into four major categories: (a) rule based, (b) linear discriminant analysis, (a) artificial neural networks, (d) other mathematical/statistical methods. Examples of each category were also discussed.

There are several issues that make the comparison of different FP reduction methods difficult. Among the primary reasons is that the number of patients used in each study is different with others. Also the scans are acquired by various CT scanners, thus having different thickness and resolution. The number of true nodules is another factor of discrepancy among the studies. Level of experience and number of radiologists who provided the truth sets for each dataset is of course variable. A dataset and performance summary of the methods discussed in previous sections of this chapter is shown in Table 2.2. As it is evident in the table, nodule size range in various studies is dissimilar. All this has obstructed the formation of a standard reference, based on which, performance of different lung nodule CAD systems can be compared.

No.	Research	Year	Nodule size range	Sensitivity	FP
1	Kanazawa *et al.* [10]	1998	> 4mm	90%	8.6 FP/case
2	Zhao *et al.* [11]	2003	2mm-7mm	75%	9/case
3	Goo *et al.* [13]	2003	> 5mm	65%	(8 ± 5.2)/case
4	Dajnowiec *et al.* [14]	2005	> 2.1 mm	80%	3/slice
5	Armato *et al.* [15]	1999	3.1mm-27.8mm	70%	3/slice
6	Armato *et al.* [16]	2001	3mm-25mm	70%	1.5/slice
7	Armato *et al.* [17]	2005	3mm-25mm	70%	1.6/slice
8	Gurcan *et al.* [18]	2002	2mm-25mm	84%	1.74/slice
9	Kung *et al.* [20]	2004	2.5mm-12mm	80%	9.75/case
10	Penedo *et al.* [26]	1998	8mm-20mm	94%	5/slice
11	Lin *et al.* [27]	2002	1cm-5cm	89.3%	0.3/slice
12	Suzuki *et al.* [28]	2003	4mm-27mm	80%	1.85/slice
13	Suzuki *et al.* [29]	2005	8.9mm-29.1mm	81.3%	1.4/slice
14	Sivaramakrishna *et al.* [30]	2002	4.43mm-18.02mm	90%	2/slice
15	Dehmeshki *et al.* [31]	2004	<30mm	95%	0.27/slice
16	Farag *et al.* [32]	2004	>12mm	82.3%	0.92/slice

Table 2.2: Dataset and performance summary of several lung nodule CAD systems in the literature

3 Detection of Regions of Interest

3.1 Overview

The main focus of this chapter is to discuss a computerized method for spotting suspicious regions on CT images of lung. Suspicious regions or regions of interest (ROI), as they are usually referred to in the literature, are those locations within the lung volume that can be signs of lung cancer or some kind of pulmonary disease. In the case of current research lung nodules constitute the ROI. Of course, not all candidates marked up with the computerized detection system are true nodules and not all true nodules are necessarily malignant.

As for every comprehensive ROI detection approach, the current method also involves the three main stages of:

- Extraction of lungs
- Detection of potential nodule candidates within the lungs
- Removing false positive detections

Lungs are extracted using a 3D segmentation algorithm which includes optimal thresholding and region growing. That method is explained in detail in [12]. The present thesis focuses on the second and third steps of lung nodule detection process: (a) Identification of potential nodule candidates within the lungs, and (b) false positive reduction.

3.2 Identification of Potential Nodule Candidates within the Lungs

Figure 3.1 depicts the main stages of the proposed nodule detection algorithm. First, through the use of contrast enhancement and histogram analysis, the system extracts some parts of the lungs, which fall in nodule range in terms of intensity. Then by applying a series of shape, size and intensity analysis techniques, the objects with highest chance of being nodule are sorted out of the initial candidates.

```
┌─────────────────────────────┐
│    Contrast Enhancement     │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│     Histogram Analysis      │
└─────────────────────────────┘
              │
              ▼
┌───────────────────────────────────┐
│ Image Thresholding and Translation│
└───────────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│   Morphological Operations  │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│    Shape and Size Analysis  │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│      Intensity Analysis     │
└─────────────────────────────┘
```
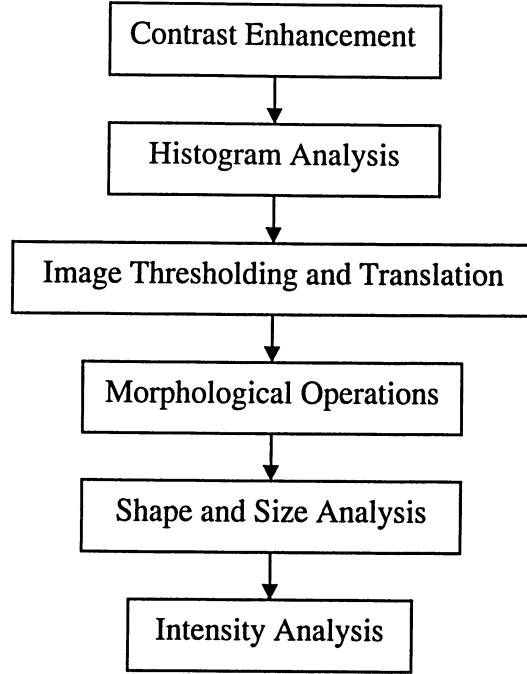
Figure 3.1: Block diagram of the proposed lung nodule detection system

3.2.1   Contrast Enhancement

CT scans are acquired with different devices and doses. Hence, CT images do not look visually similar in terms of contrast. A CT image with poor contrast may lead into loss of useful information. As the whole idea of the nodule detection process is to differentiate between nodule and non-nodule objects, it is therefore important to emphasize those regions of lung that may carry nodule information. A primary step to have those regions stand out is to enhance the contrast of CT image.

In order to enhance the contrast quality, every 2D CT image goes through a histogram equalization process. This process enhances the contrast of image by transforming the values in the intensity image so that the histogram of the output image approximately matches a specified histogram. The specified histogram is calculated as the flat histogram H:

$$H = ONES \times prod \ (size \ (A))/n; \tag{3.1}$$

24

where A is the 2D array of digital image to be contrast enhanced. n is the number of discrete gray levels in the transformed (contrast enhanced) image. The default value used for n is 64. ONES is a vector of all 1s with dimension $1 \times n$. The operator *prod* (...) multiplies the terms inside the brackets by each other. In this case the multiplier and multiplicand are the dimensions of the original image A. The histogram equalization algorithm chooses a grayscale transformation T to minimize $|c_1(T(k)) - c_0(k)|$. c_0 is the cumulative histogram of A; c_1 is the cumulative sum of H for all intensities k. This minimization is subject to the constraints that T must be monotonic and $c_1(T(a))$ cannot overshoot $c_0(a)$ by more than half the distance between the histogram counts at a. The following transformation is used to map the gray levels to their new values [33].

$b = T(a)$, such that:

$$b = \left[\sum_{j=0}^{a} N_j \right] \times \text{(Maximum intensity level)} / \text{(Number of pixels)} \tag{3.2}$$

The summation inside square brackets indicates the number of pixels having the intensity below the intensity level a or equal to it. The numerator of the second term of multiplication is the maximum intensity level that a pixel can get.

Figure 3.2(a) shows two instances, where the true nodules are indistinguishable due to low contrast of the CT image. In such cases, it is difficult both for naked eye and for CAD system to pull out the nodule candidate from the surrounding lung parenchyma. In figure 3.2(b), the same two images as 3.2(a) are demonstrated after contrast enhancement. The nodules appear remarkably brighter in this version and a perceptible difference between the intensity of nodule and its surroundings can be observed.

3.2.2 Histogram Analysis

A labeled example of lung CT image is shown in figure 3.3. From the ROI detection standpoint, we are concerned about the inner lung and lung wall regions. The gray level intensity of these regions can be categorized into two main divisions of (a) lung parenchyma, and (b) nodules, blood vessels and bronchial wall [11]. Based on this *a priori* knowledge, histogram of the contrast enhanced lung image is automatically divided into two sections.

Through this step, the elements of image are binned into two equally spaced containers and the number of elements in each container along with the location of each bin center are outputted for the image.
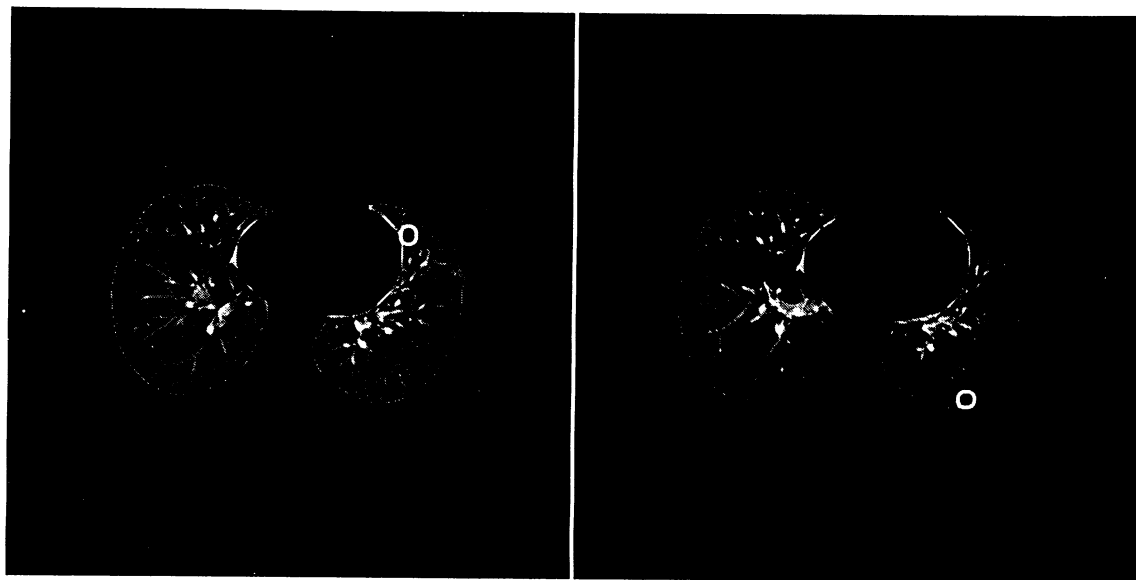
(a)
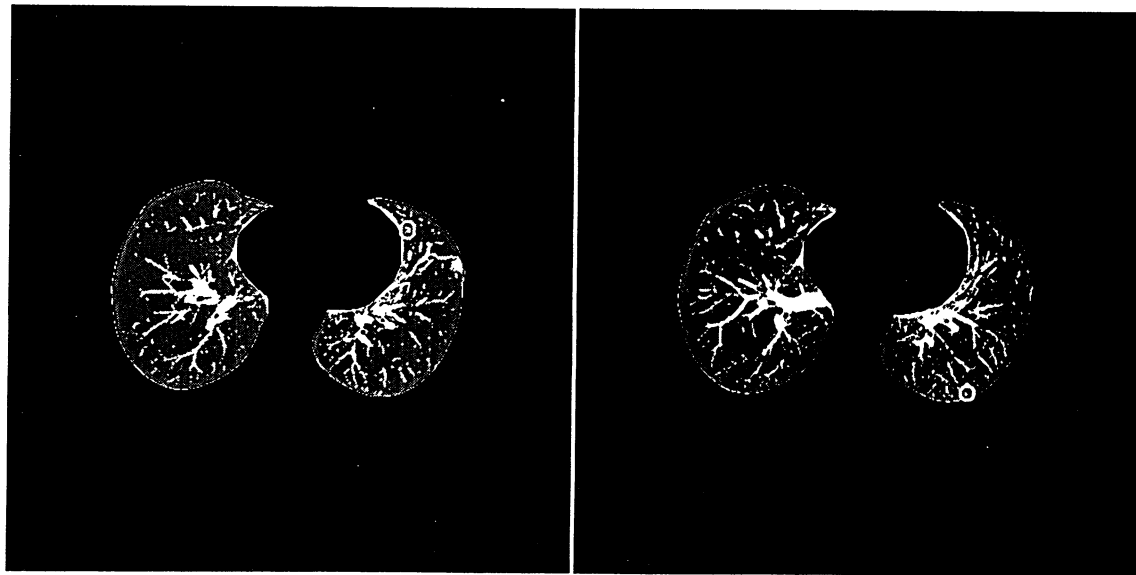
(b)

Figure 3.2: Effect of contrast enhancement on CT images
(a) Instances of lung nodules with very low contrast. (b) The same images after contrast enhancement. The nodules are much more distinguishable in (b).

Figure 3.4 illustrates a lung image and its corresponding two bin histogram. It can be seen that most of the pixels fall in the first bin, admitting what can also be observed by eye; i.e., darker pixels (corresponding to lung parenchyma) are the majority and brighter pixels (corresponding to vessels and nodules) are the minority. Location of each bin center is in fact the mid intensity value for that bin. Center of the second bin is chosen as the threshold value for each image.
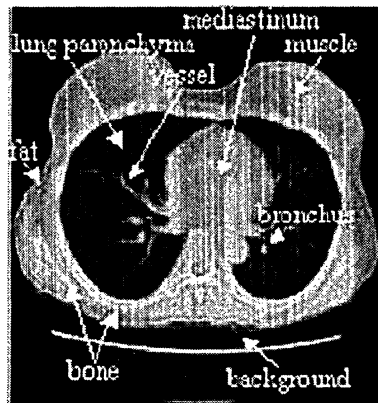
Figure 3.3: A labeled CT scan of lung
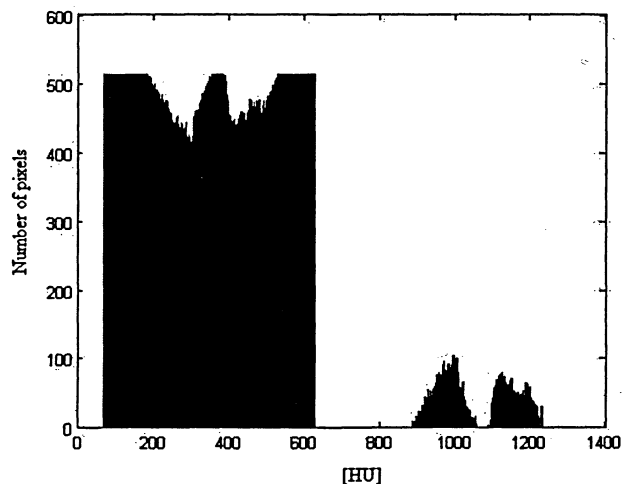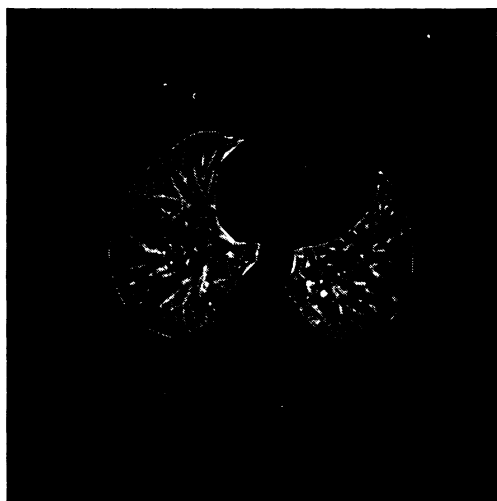
Figure 3.4: Segmented lungs from a sample CT slice and gray level histogram of lungs after the contrast enhancement step

3.2.3 Image Thresholding and Translation

Once the threshold value is determined, the slice is thresholded with that value. The result is a binary (black and white) image, where pixels with intensity greater than the threshold appear in white and the rest appear in black. Thus, in addition to nodules which may be present in the slice, all other objects with intensities in the same range as nodules also show up white in the thresholded image. Those objects may include branching in-plane vessels, cross section of vertical vessels and airways, mediastinum and other non-nodule elements. An example can be seen in figure 3.5.
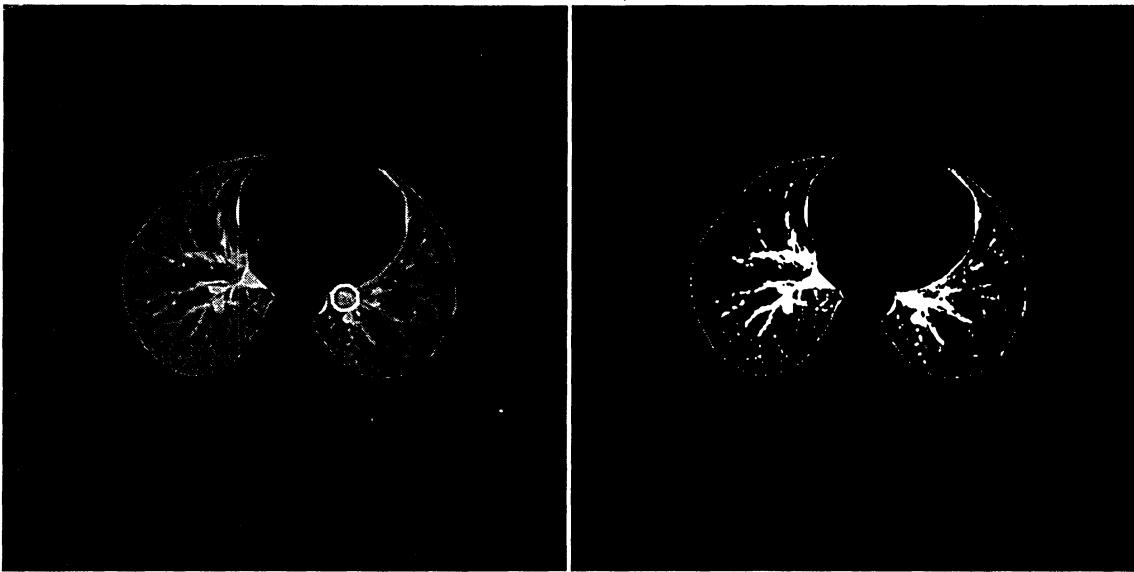
Figure 3.5: Lungs after thresholding
The region enclosed inside the circle in the left image is a nodule. In the binary image of right, the nodule and other objects with intensity in the same range as the nodule appear in white.

As an initial step to eliminate some of the unwanted objects, the thresholded image is shifted one pixel to left and one pixel to right and then the resulting translated images are ANDed together. The premise behind this task is to remove any object which is less than two pixels thick and more importantly eliminate the lung boundary. The overall shift is chosen two pixels because in this study we are only looking for nodules greater than 3 mm in diameter. The pixel spacing in digital CT images is usually equal or less than one, i.e., distance between two adjacent pixels is either 1 mm or a fraction of 1 mm. This is true for both x and y axes, as the CT images are isotropic. Hence two pixels would be equal to or less

than 2 mm, and therefore a nodule with 3 mm diameter will not be completely removed due to the translation. Some investigators maintain that nodules smaller than 5 mm may be of limited clinical importance, while others contend that since the benefits of detecting small lung cancers are presently unknown, to exclude from consideration nodules in the 1-2 mm size range may limit the relevance of the database [2]. The 3 mm inclusion criterion in current research is selected based on the Lung Image Database Consortium (LIDC) standard, which strikes a compromise between the mentioned views and also takes into consideration the practical issue that all lesions identified as nodules in the database will require effort to define spatial location and extent and to follow through subsequent examinations [2].

The simple translation step downsizes the number of objects in the binary image. Figure 3.6 depicts an example. 3.6(a) is the binary image after thresholding and 3.6(b) is the result of logical AND of left-translated and right-translated versions of that image. Figures 3.6(c) and 3.6(d) magnify a small region of slice that is removed through the translation and AND process.
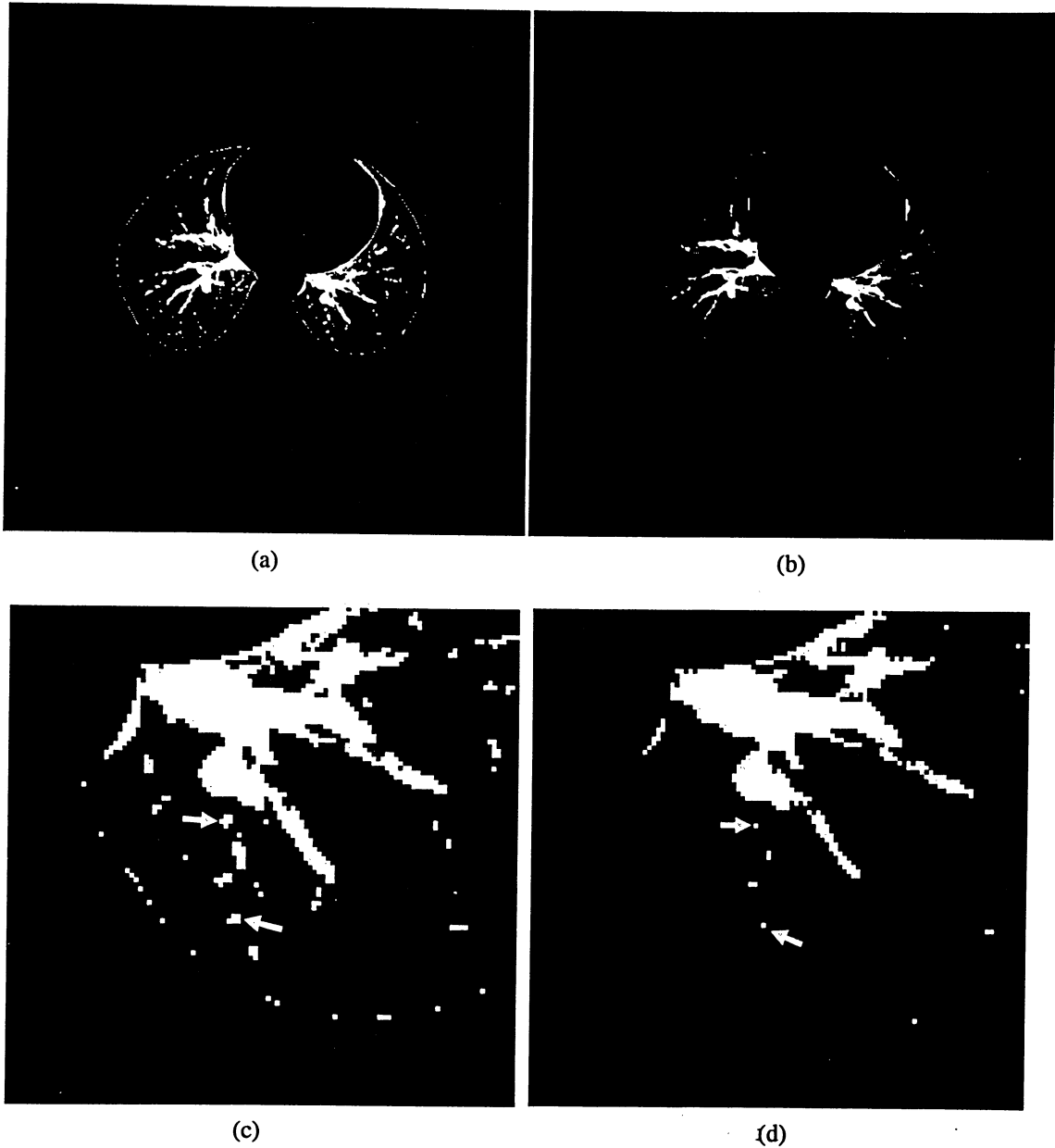
Figure 3.6: Effect of image translation

(a) Binary image of lungs after thresholding. (b) Image in (a) after translation and AND process. (c) and (d) A portion of (a) and (b) magnified respectively. It is evident that as a result of the translation stage, many pixels on lung border in (c) do not exist in (d). Also small objects are shrunk to isolated pixels (arrows).

3.2.4 Morphological Operations

A series of morphological operations are utilized for further separating the connected objects and removing objects with low chance of being nodule. However before starting the

morphological techniques, objects with size less than 6 mm diameter circle are sought and copied in another 2D array. This is done to prevent losing small objects in the midst of morphological operations which will follow. The subsequent sections will elucidate this issue.

First the boundary of each object is subtracted from it, leaving the interior pixels of objects on. Object boundary is found by setting a pixel to 0 if all its 4-connected neighbors are 1, thus leaving only the boundary pixels on. This boundary removal step is functional for connectivity purging, where two or more objects are tangentially connected to each other. This is crucial for detaching nodule candidates that are attached to another object, mainly the lung wall or the bronchial tree. Figure 3.7 presents several examples of juxtapleural nodules (nodules connected to lung wall) and nodules attached to a lung interior object. The effect of boundary removal step is also shown for each case. Candidate size will of course be reduced after the boundary elimination step, thus bringing up the chance of small entities getting eliminated. However this is compensated by inserting the small objects retained earlier.

As a result of non-ideal lung segmentation or over-cautious segmentation, sometimes soft tissue containing the cross section of airways (trachea) or parts of the bronchial tube also appear within the lung region. Those cases are automatically diagnosed by testing for the Euler number. The Euler number is a scalar whose value is the total number of objects in the image minus the total number of holes in those objects and is computed by considering patterns of convexity and concavity in local 2-by-2 neighborhoods [34]. In the present research, each object that has Euler number less than 1 and greater than -4 is removed (figure 3.8). The -4 limit is set to prevent elimination of cases where lung parenchyma appears as a single object with many holes at places of abrupt intensity difference.

Another group of lung nodules that are difficult to handle by CAD –and are therefore often missed- are those nodules that attach to vessels or elongated traces of soft tissue. To better perceive such nodules, lengthy objects are broken into pieces. Maximum end to end distance of object along each dimension (x axis and y axis) is calculated. This is equivalent to finding the dimensions of the smallest rectangle that contains the region occupied by the
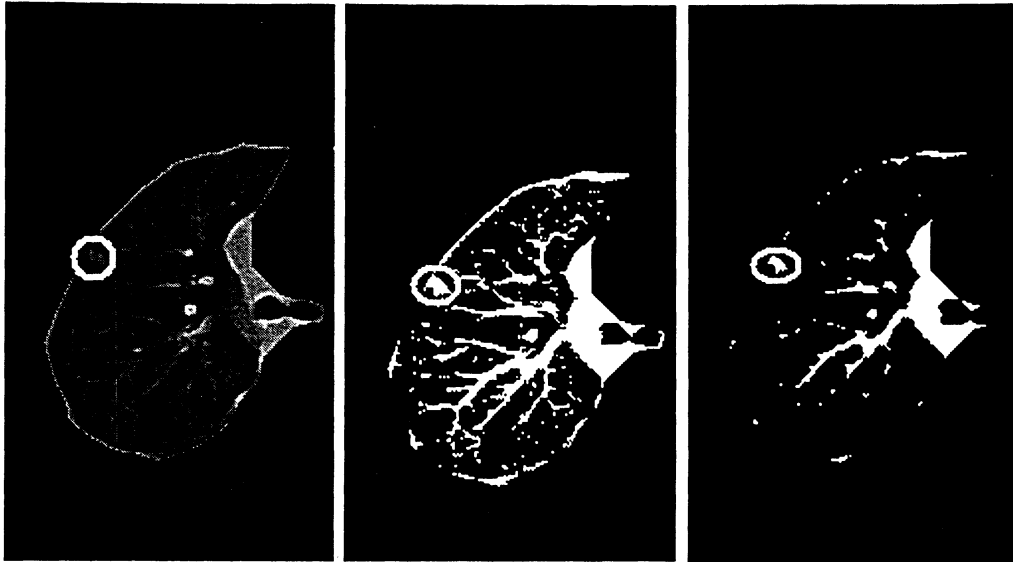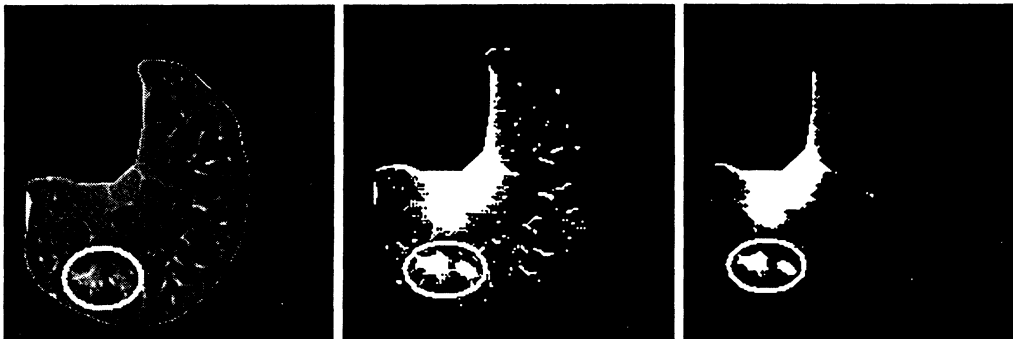
Figure 3.7: Examples showing effectiveness of the object boundary removal step
Each section from left to right shows original CT, after thresholding and after boundary elimination
respectively; (a) Wall nodule. (b) Nodule trapped between wall and thick vessels. (c) Two interior nodules
attached to each other.

object. For each dimension, if the width of the objects is greater than 3 cm, i.e., the maximum size criterion for lung nodule set by the LIDC [2], the object is split every 2 cm along that dimension. Figure 3.9 demonstrates this process. Applying the latter step has two advantages. The central advantage, as mentioned earlier, is that it detaches the nodules from long non-nodule objects. Secondly, it breaks up large non-nodule components that have round opaque shape similar to nodule morphology. Parts of the mediastinum or bronchus that are included in the lung region can often look like this. Such objects remain unaffected through the prior filtering steps.
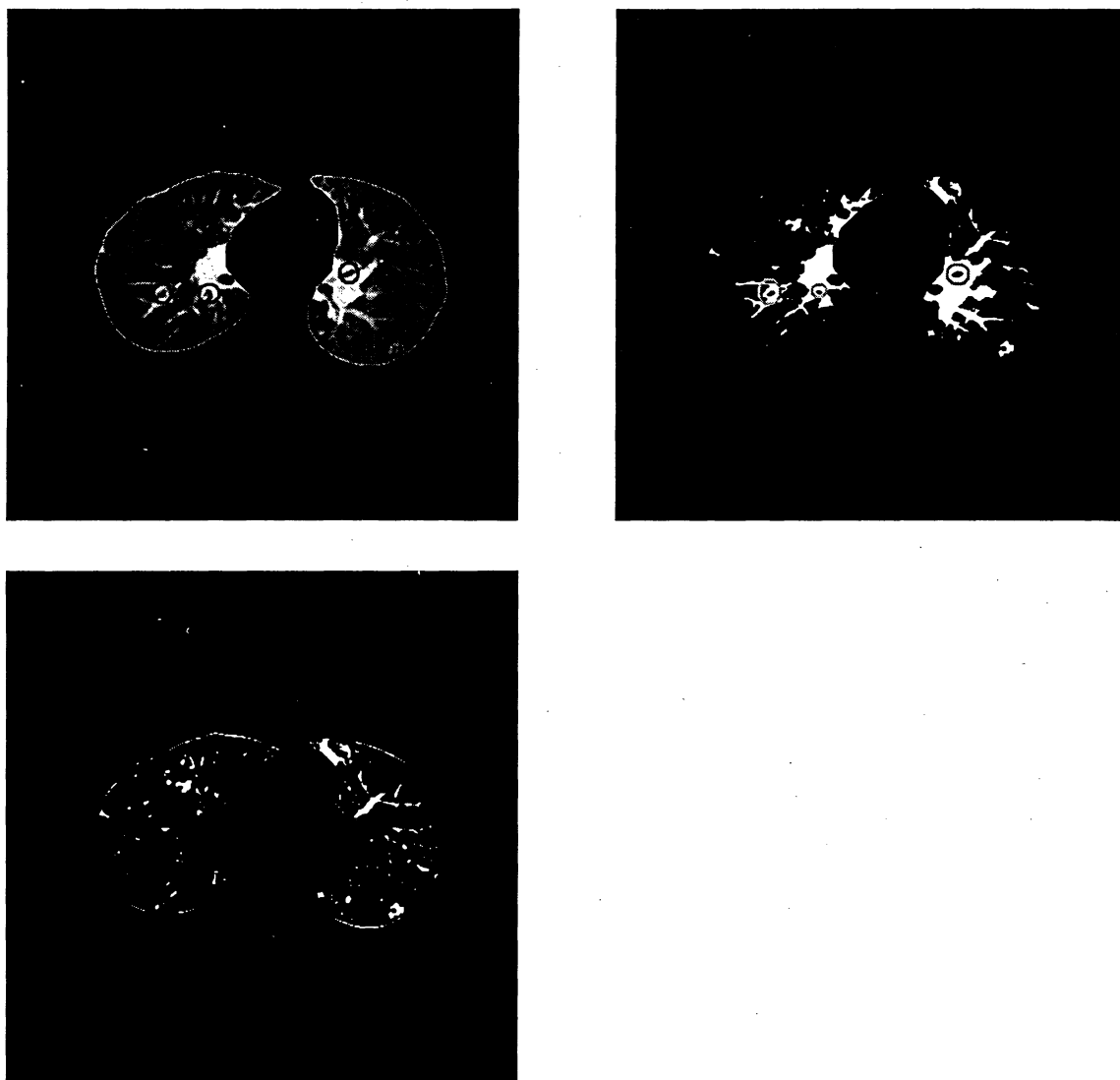
Figure 3.8: Using Euler number for filtering nodules from non-nodule objects
Left to right, top to bottom: intensity change within the bronchial vessels or airway cross-sections appear as holes in the binary image before testing for Euler number. Objects with such interior holes are removed after the Euler number test.

Figure 3.9: Effect of breaking long objects into pieces
(a) Original CT slice. (b) Segmented lungs; a nodule is located inside trace of soft tissue in left lung (arrow). (c) Lungs before the splitting step. Nodule is trapped in the arc shape tissue. If not broken into pieces, this large irregular shape will be eliminated in subsequent steps and consequently the nodule will be missed. (d) Lungs after the splitting step.

Figure continued on next page
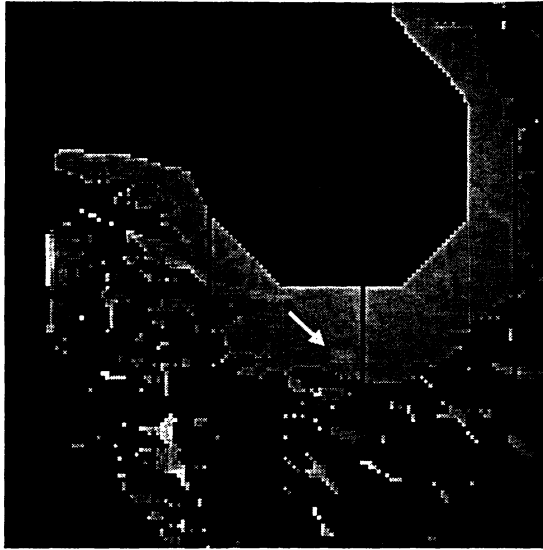
(e)

Figure 3.9 (Continued from last page): (e) Image in (d) magnified about where the nodule is located. The piece of tissue containing the nodule is now separated from the rest of the irregular shape.

3.2.5 Shape and Size Analyses

At this stage, the objects are filtered based on their area and axis ratio. Objects with area smaller than a 3 mm diameter circle (the minimum effective diameter for nodule inclusion set by LIDC [2]) are confiscated. The axis ratio is defined as candidate's major axis length to its minor axis length within its tight bounding box. The idea comes from [14], in which tight fitted bounding box is obtained by rotating the object based on its orientation using bi-cubic interpolation. With this modified bounding box, an object could be at any orientation and the box dimensions and compactness would be the same. The tight bounding box contains a new rotated version of object in a way that the angle between x axis and the major axis of the ellipse with same second moments as the object is zero. In the current work, objects with axis ratio greater than 2 are set to zero. The shape and size analyses stage trims down the entities which are too small or too long in order to be considered nodule.

3.2.6 Intensity Analysis

Last stages of the ROI extraction process are two intensity-based filtering steps. The first step is intended to remove cross section of in-depth vessels. These objects appear in

round well-formed shapes, with very close size and morphology to that of small nodules. Their main distinction from nodules is that they are very bright points with sharp intensity disparity with their surroundings. Figure 3.10 shows several examples. When nodule detection is performed manually, radiologists ignore this type of false positives by recognizing their very bright and sharp pointed appearance.
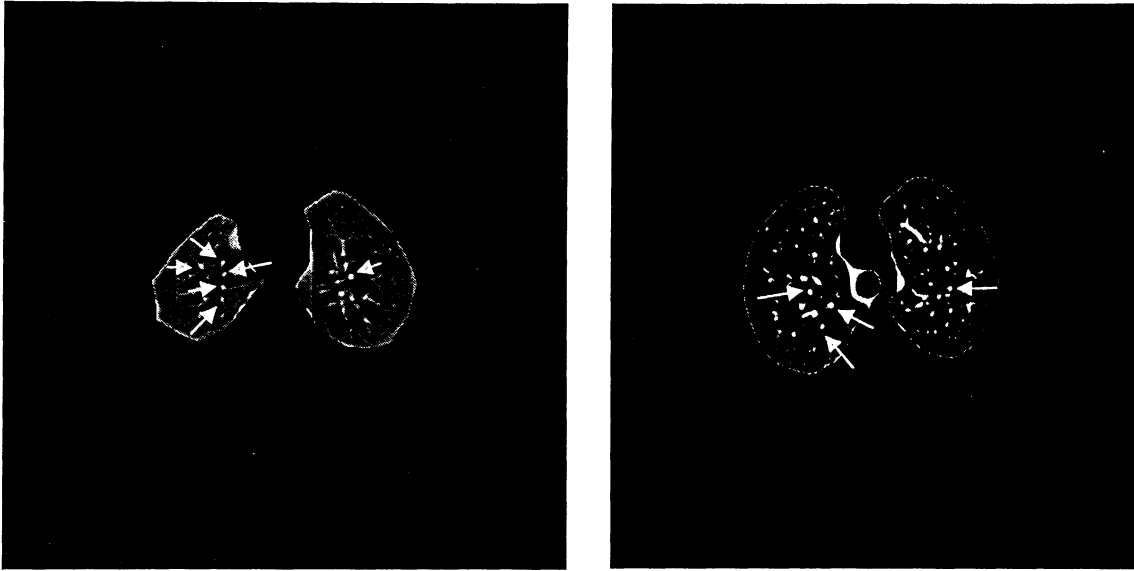
Figure 3.10: Density profile analysis for removing cross section of vertical vessels
Arrows point to cross section of in-depth vessels. They appear as bright round objects and have distinct intensity difference with their surrounding lung parenchyma.

As no more shape or size filtering will be performed, the image is dilated with a disk structuring element of radius 1 to attain information about gray level intensity of neighboring pixels of each object. Then the intensity of points on the border of the dilated image at $0°$, $90°$, $180°$ and $270°$ are compared with the intensity of object's centroid. For the object to remain un-removed, all of the next four inequalities must be satisfied:

$$I_c - I_{0°} < 220 \text{ HU} \tag{3.3}$$

$$I_c - I_{90°} < 220 \text{ HU} \tag{3.4}$$

$$I_c - I_{180°} < 220 \text{ HU} \tag{3.5}$$

$$I_c - I_{270°} < 220 \text{ HU} \tag{3.6}$$

Parameter I represents the intensity of pixel at the location indicated in the subscript. Subscript c corresponds to centroid. The value 220 HU is selected empirically based on the

minimum observable intensity change in vessel cross-section objects. Figure 3.11(a)-(d) magnifies a sample vessel cross-section and its density profile along its vertical and horizontal axes. The density profiles are sharp and the intensity peak is concentrated at the centroid. This is against the more uniform intensity distribution of a true nodule with similar shape and size, as can be witnessed in 3.11(e)-(h).

(a)   (b)

(c)   (d)
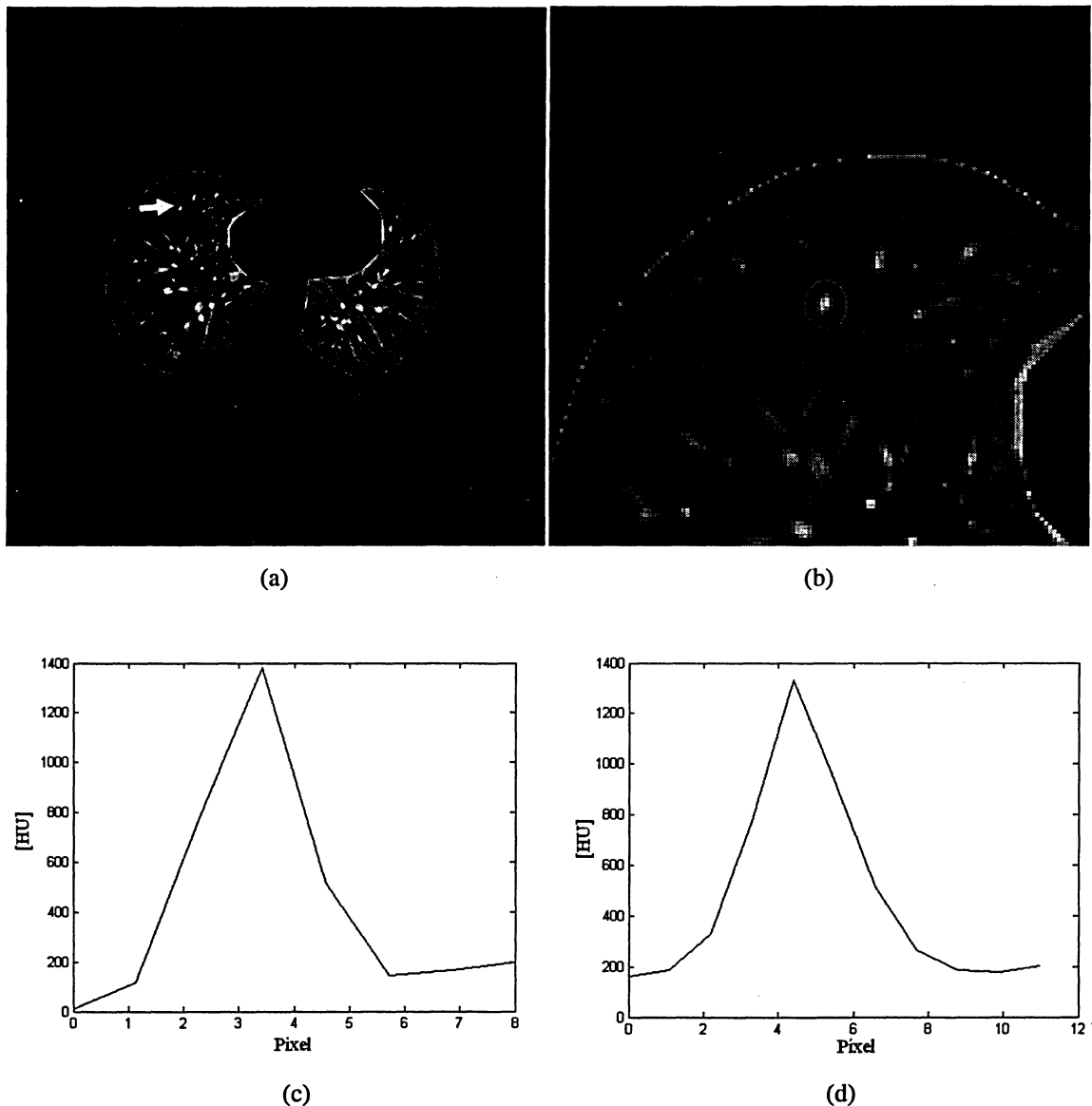
Figure 3.11: Density profile of nodule and non-nodule objects along their horizontal and vertical axes
(a) Cross-section of a vertical in-depth vessel. (b) Object in (a) zoomed in. (c) Density profile of vessel cross section along its horizontal axis. (d) Density profile of vessel cross section along its vertical axis.

37
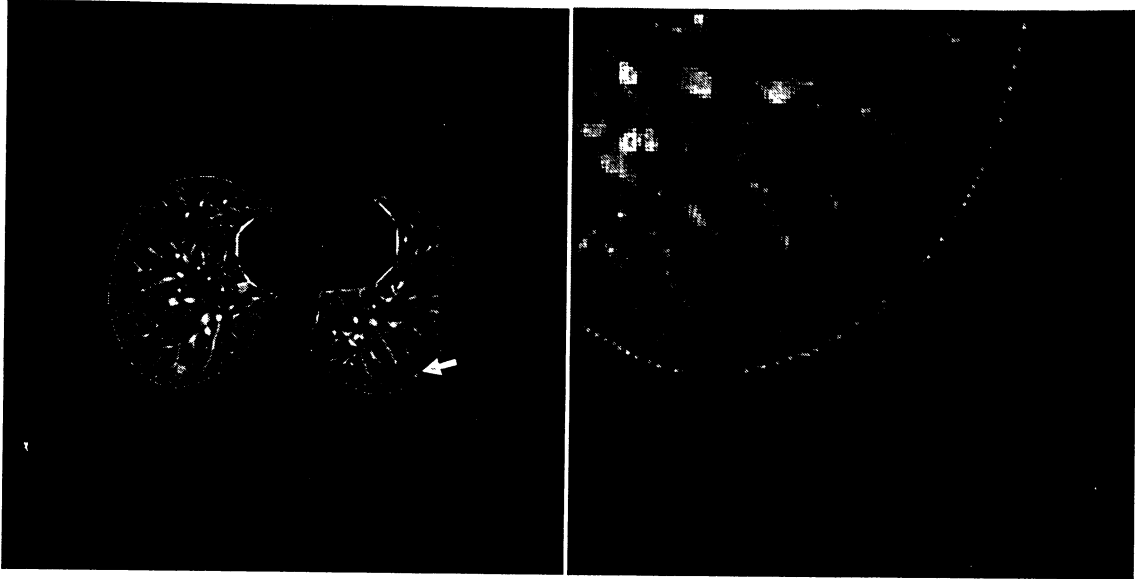
(e)   (f)
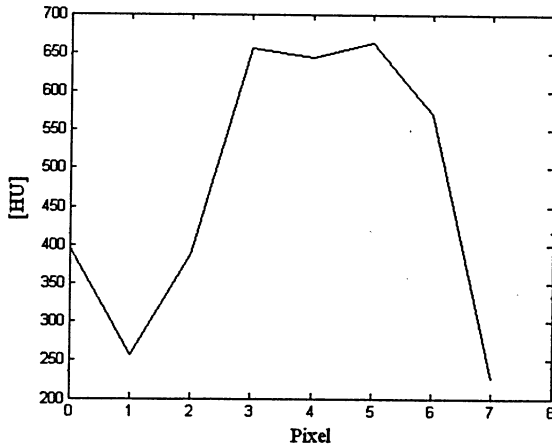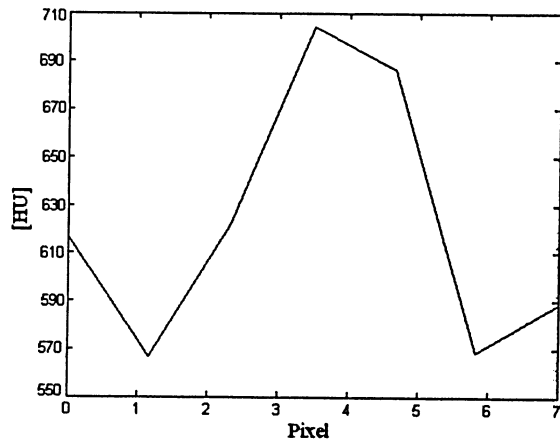
(g)   ·   (h)

Figure 3.11 (continued from last page): (e) A true nodule. (f) Nodule in (e) zoomed in. (g) Density profile of nodule along its horizontal axis. (h) Density profile of nodule along its vertical axis.

The last step is removing objects with gray level intensity too low to be considered potential nodules. All objects with mean intensity lower than value T expressed in (3.7) are removed.

$$T = min (450\ HU, t) \tag{3.7}$$

If the intensity histogram of original CT image is divided into four equal sections, t is the upper boundary of the first quarter. In other words t is equivalent to 25% of the

38

maximum intensity range of CT image. The other component inside brackets of equality (3.7), 450 HU, originates from the fact that in the CT image data, air will appear with a mean intensity of approximately -1000 Hounsfield units (HU), most lung tissue will be in the range of -910 HU to -500 HU, while the chest wall, blood, and bone will be much more dense (well above -500 HU) [35]. With this in mind and with the assumption that in the current research all intensities have been shifted by +1024 HU for rescaling purpose, most lung tissue will be in the approximate range of 0 HU to +500 HU. Hence there is minor chance that candidates with mean intensity less than 500 HU are of nodule nature. This cutoff intensity is lowered by another 50 HU to be even more cautious about possible nodules. This results in the value 450 HU. Figure 3.12 illustrates three low intensity false positive objects that continued to remain after all the preceding filtering stages. These objects are caught by the very last step discussed and consequently removed.
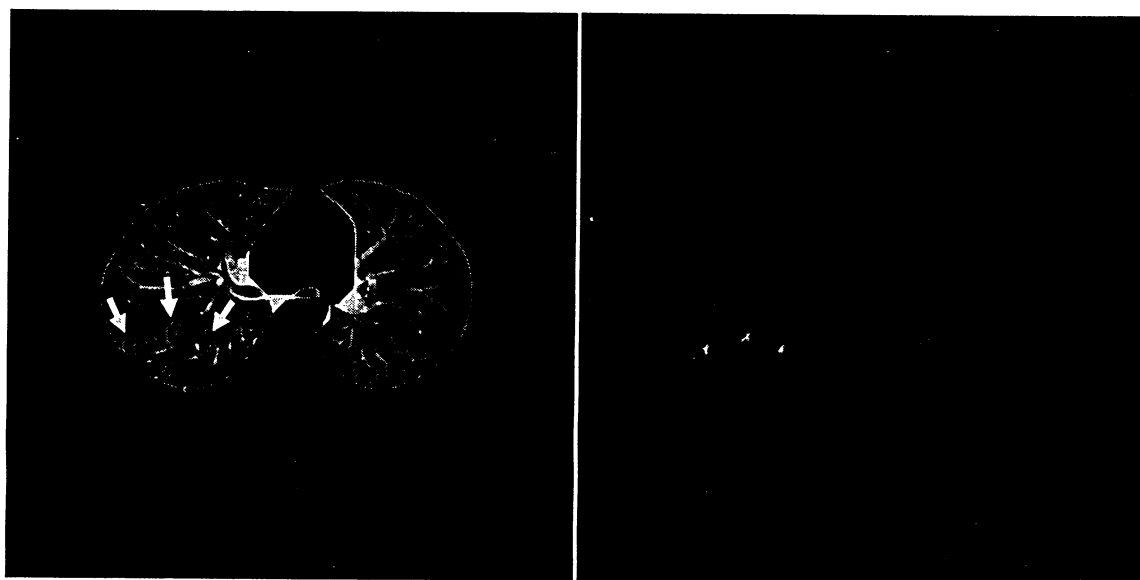
Figure 3.12: Effect of filtering based on object's mean intensity
(Left) Circles enclose low intensity non-nodule objects. (Right) These objects continue to remain until the very last filtering step, through which they are removed.

4 Enhanced False Positive Reduction Scheme

4.1 System Overview

The contribution presented in this chapter is focused on deign and development of a novel hybrid learning method for reducing the number of false positives in results acquired by an automatic lung nodule detection system in CT images.

In spite of the filtering steps performed in the ROI detection stage, still many false positives exist in the final ROIs. It is thus imperative to design and employ a more enhanced means of FP reduction. In fact, effective FP reduction is an universal requirement for all CAD systems. This reason motivated the author to develop the following hybrid learning FP reduction scheme.

The hybrid scheme was originally developed in order to trim down the number of candidates picked out as potential nodules by an existing rule-based CAD system [12]. However use of the proposed FP reduction system is not limited to the specific system of [12]. As will be discussed in the results chapter, the enhanced learning FP reduction scheme was also applied to results acquired by the ROI detection system introduced in chapter 3 of this thesis and proved to be rewarding.

4.2 Hybrid Learning FP Reduction Scheme

Block diagram of the hybrid learning FP reduction algorithm is depicted in figure 4.1. The algorithm consists of three main modules: features extraction, fuzzy c-means clustering and iterative linear discriminant analysis. The proposed scheme is hybrid in the sense that combines the discrimination power of fuzzy c-means clustering with classification ability of a modified version of linear discriminant analysis to achieve false positive reduction.

The learning methods classify sample data by extracting and integrating information from samples. That information is called features and is acquired in the first module of system. The second module, fuzzy c-means clustering (unsupervised learning), makes an approximate decision about the case being nodule free or containing multiple nodules. The third unit is what has been exclusively designed for this research and is named adaptive

iterative LDA (supervised learning). This step is used for further reduction of FPs when the fuzzy c-means clustering unit has determined the existence of multiple nodules in a given case.
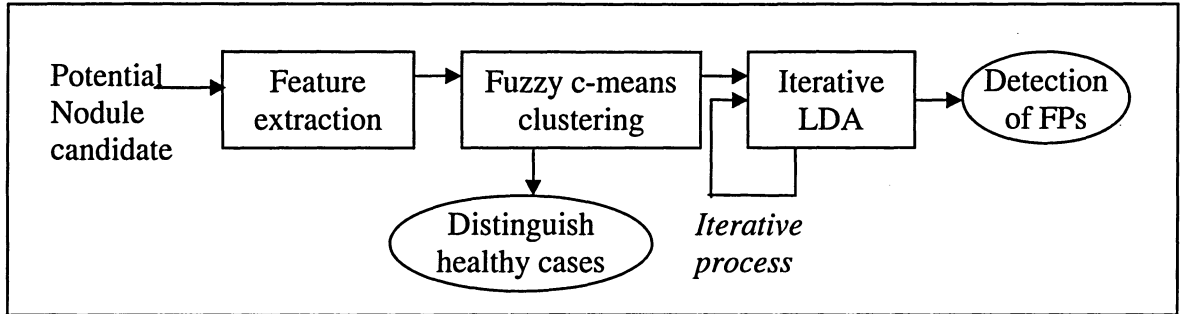
Figure 4.1: Block diagram of the hybrid FP reduction scheme

One might wonder why two learning algorithms are used -instead of one- in the design of the proposed enhanced FP reduction scheme. Or, with the presence of a supervised learning module such as the iterative LDA, benefit of incorporating an unsupervised learning method like fuzzy c-means clustering, which attempts to learn something of value from unlabeled samples may seem of little practicality. The answer is that in early stages of an investigation it may be valuable to perform exploratory data analysis and thereby gain some insight into the nature or structure of the data. The discovery of distinct subclasses –clusters or groups of patterns whose members are most similar to each other than they are to other patterns- or of major departures from expected characteristics may suggest significant alterations in the approach for designing the classifier [36]. Realization of this fact in current research is that the nodule-free or almost nodule-free cases are recognized in the unsupervised learning module. The supervised learning module, which is more data and time consuming is applied only when there is need for more cautious classification. This will be further clarified when the three modules of the enhanced FP reduction scheme will be explained in the following sections.

4.2.1 Feature Extraction

The primary stage of work is selection and extraction of appropriate features for further classification. For each potential candidate, total of nine features are extracted. Two

gray-level related, four morphological and three contextual features are calculated for each candidate. Table 4.1 lists the features and the category they belong to. As it is evident from the table, the gray level rated features contain information about the intensity value of pixels within and around the nodule candidate. The morphological features include information about the shape and size of candidate and contextual features provide knowledge about the relative location of candidate within the lung context. The nine features are extracted and used in the subsequent classification modules of the hybrid approach. Four of these features (mean intensity of candidate interior, mean intensity of candidate surrounding, volume and candidate's position within lung context) have been designed and utilized by the author specifically for the current research, while the rest of features were inspired by previous works in this field. Each feature will be further explained in the following. The features are 3D except mentioned otherwise. By 3D it is meant that 3D connected components with 26 point connectivity are counted as one object. In other words for a given pixel, any other pixel contained within a 3×3 cube centered at the given pixel is connected.

Gray level related	Morphological	Contextual
Mean intensity of candidate interior	Axis Ratio	Perimeter
Mean intensity of candidate surrounding	Compactness	Vertical Ratio
-	Area	Position of nodule in lung context
-	Volume	-

Table 4.1: Categorization of nine features extracted for each potential nodule candidate

Mean intensity of candidate interior: This feature relates to the average intensity of nodule candidate's interior pixels. To obtain this feature, the mean gray level value of pixels that fall within the object's tight bounding box is calculated. The reader is reminded that the tight bounding box is obtained by rotating the object based on its orientation using bi-cubic interpolation [14].

Mean intensity of candidate surrounding: This feature relates to the average intensity of nodule candidate's surrounding pixels. It is calculated through applying an exclusive OR

(XOR) operation on the tight bounding box surrounding the candidate and its dilated version. The result of XOR is only non-zero where the tight bounding box and its dilated version do not overlap, i.e., along the boundary of candidate. Figure 4.2 presents a visual example.

Figure 4.3 shows several nodule samples and their corresponding mean interior and mean surrounding intensities extracted by the system.

| (a) | (b) | (c) |
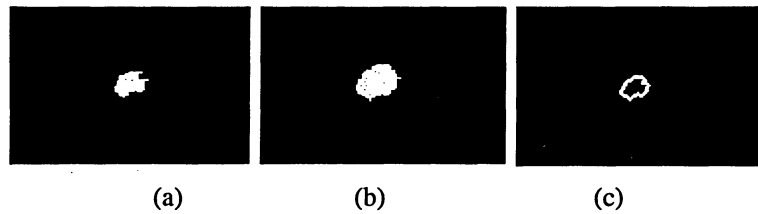
Figure 4.2: Extracting the surrounding pixels of a nodule
(a) A sample nodule, which appears only in one slice. (b) Nodule in (a) dilated with a disk element of radius 2.
(c) Result of applying XOR operation on (a) and (b).
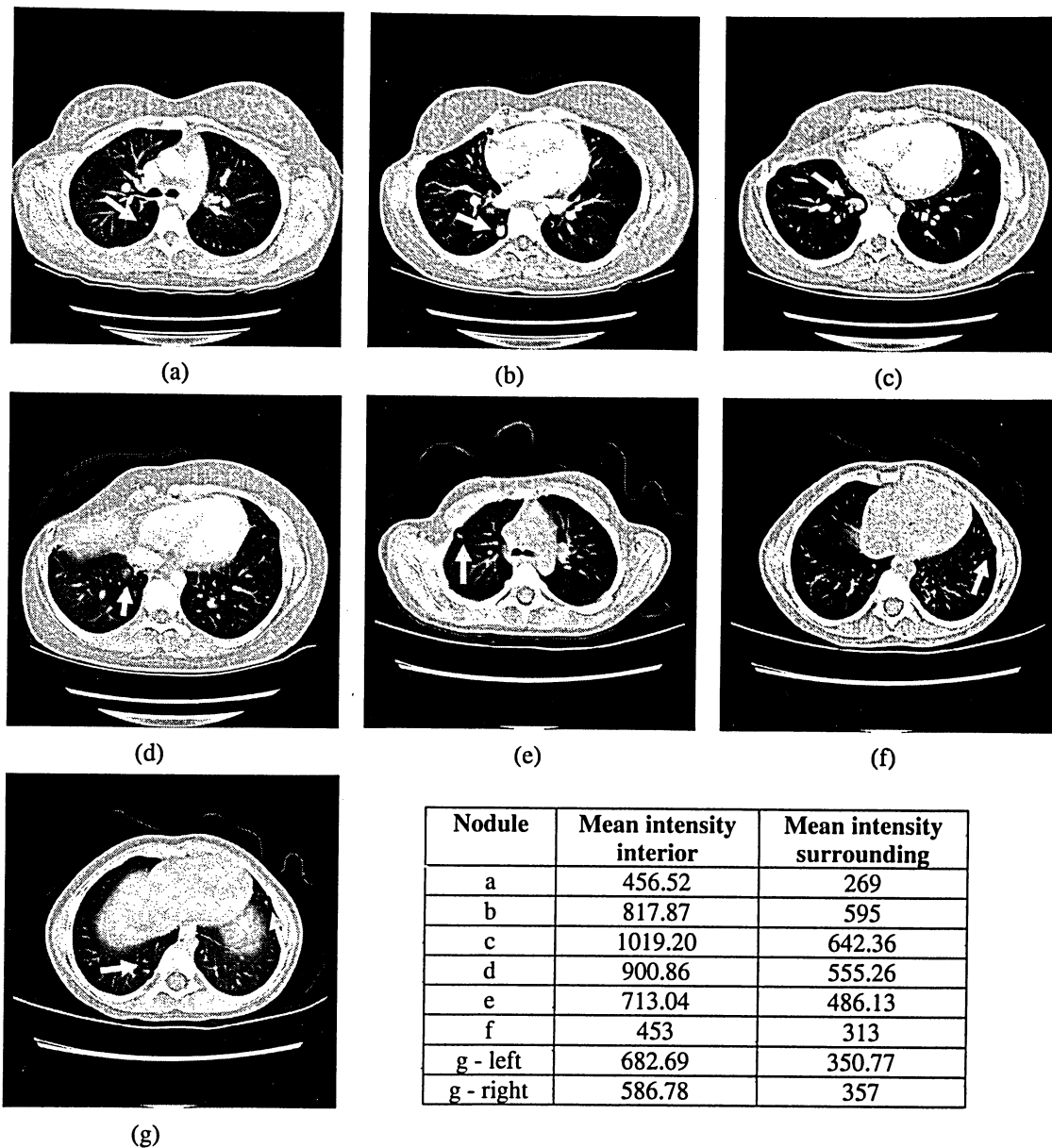
Figure 4.3: Examples of a variety of lung nodules and their corresponding interior and surrounding mean intensities

Nodule	Mean intensity interior	Mean intensity surrounding
a	456.52	269
b	817.87	595
c	1019.20	642.36
d	900.86	555.26
e	713.04	486.13
f	453	313
g - left	682.69	350.77
g - right	586.78	357

Axis Ratio (2D): This feature is calculated as the ratio of major axis length to minor axis length of 3D candidate's biggest cross section, within its tight bounding box. Figure 4.4(a) shows an example of a true nodule, while figure 4.4(b) shows a vessel. The dimensions of nodule in part (a) are almost similar along both x and y axis, resulting in an axis ratio of close to 1. On the other hand in part (b), the vessel's major axis dimension

(along vertical axis) is quite longer than its minor axis (along horizontal axis). Hence the axis ratio for this case is large.

(a)   (b)
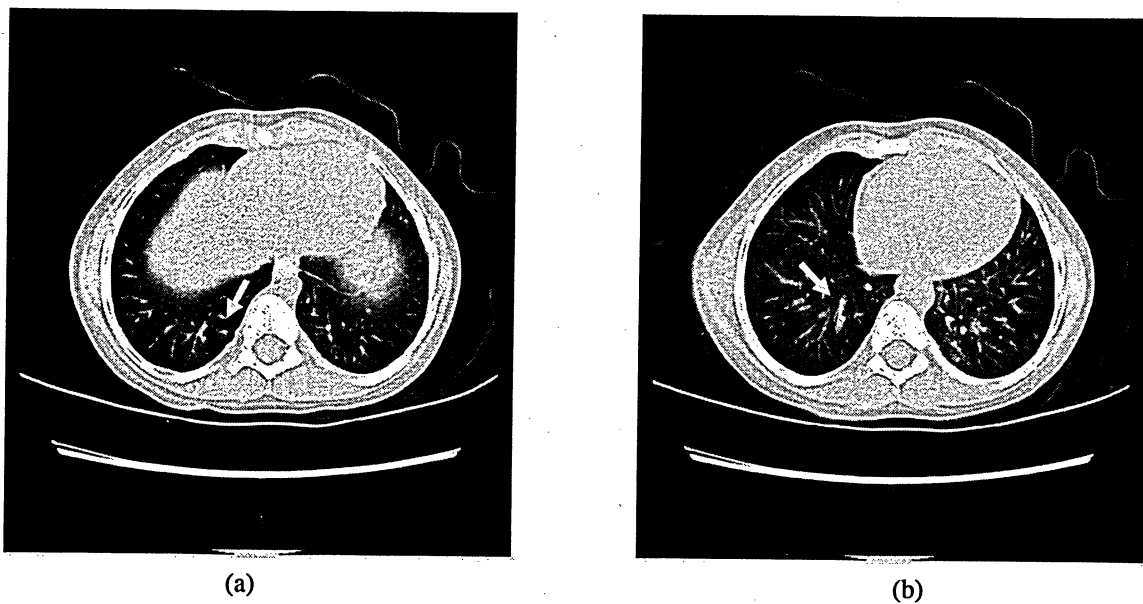
Figure 4.4: Axis ratio of a nodule and a non-nodule object
(a) Example of a well-formed round nodule with axis ratio close to 1. (b) Example of long vessel with axis ratio greater than 1.

Area (2D): This feature relates to the area of nodule candidate's biggest 2D cross section. It is determined as the total number of pixels belonging to that region. This feature does not provide any information regarding the shape of candidate. For example object in figure 4.5(a) has the same area as the object in figure 4.5(b), but shape-wise they are clearly dissimilar.

(a)   (b)
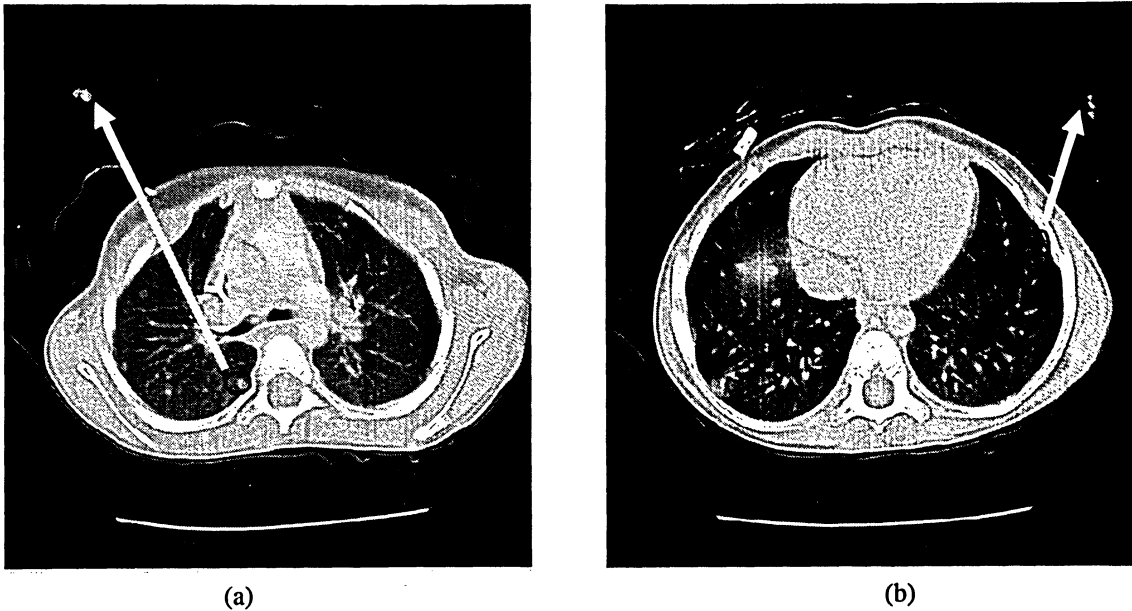
Figure 4.5: Nodule candidates and the area feature
Objects indicated in (a) and (b) have identical area, although their shapes are different.

Compactness (2D): Feature compactness is calculated as the area of candidate's biggest cross section divided by the area of its tight bounding box. Figure 4.6 presents some nodule candidates and the compactness values associated with them. It can be seen that the largest compactness value relates to part (a), which illustrates the most opaque object of the three.
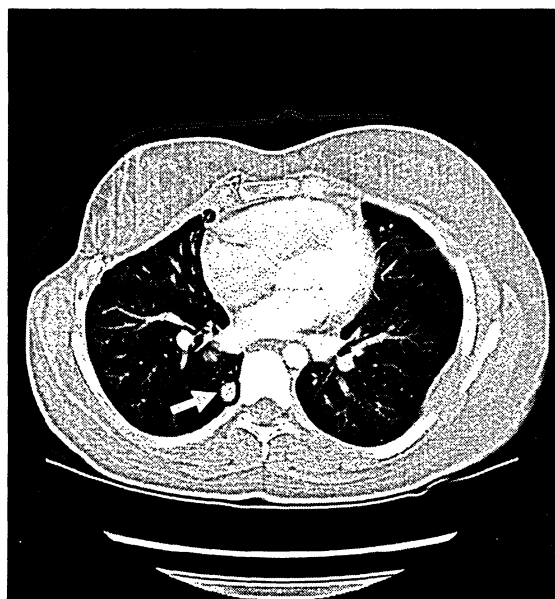
(a)
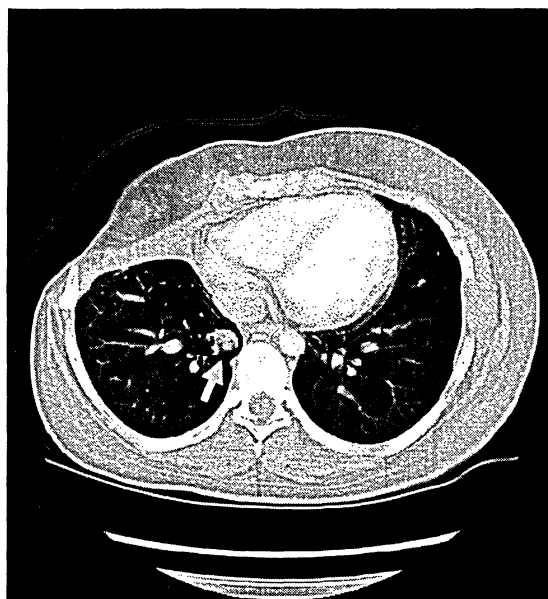
(b)

(c)

Nodule	Compactness
a	0.61
b	0.03
c	0.02

Figure 4.6: Some instances of nodule candidates and their compactness value

Volume: Volume is a 3D feature and it is calculated by summing up the number of pixels present in all cross sections of candidate and then multiplying this total by the unit volume of a voxel. A voxel (VOlume piXEL) represents a quantity of 3D data just as a pixel represents a point or cluster of points in 2D data [37]. Thus, the thickness of CT slices has a direct impact on the amount of this feature. Figure 4.7 shows the histogram of candidate volume –acquired by the current feature extraction system- for all those true nodules that were detected by the rule based CAD system of [12].
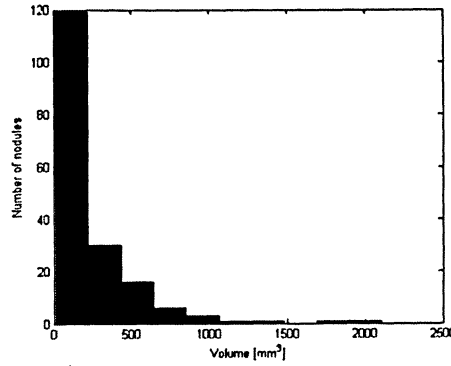
Figure 4.7: Histogram of nodule volume for all those true nodules that were detected by the rule based CAD system of [12]

Perimeter: For each candidate, the perimeter pixels are defined as those pixels that are new to the candidate after dilating it using a disc element of radius 2. The perimeter feature is then described as the ratio of perimeter pixels that are outside of the lung volume to those perimeter pixels which are internal to the lungs [12]. This feature quantifies the strength of connection between each candidate and lung wall. Figure 4.8(a) shows interior lung nodules with no attachment to lung wall, thus having a perimeter value of zero. Example of a nodule strongly connected to lung wall can be observed in Figure 4.8(b). The perimeter feature for the latter case is greater than zero.
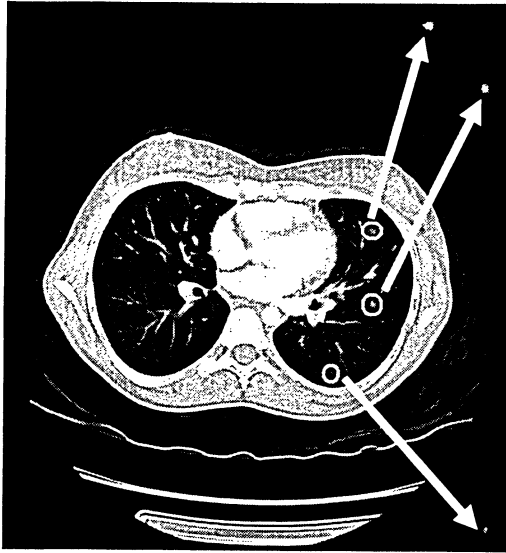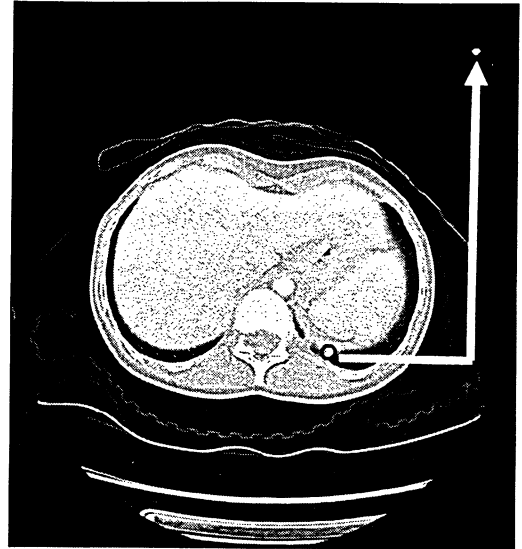
<div align="center">(a)</div> <div align="right">(b)</div>

Figure 4.8: Nodule candidates and the perimeter feature
(a) Examples of interior nodules with perimeter feature equal to zero. (b) Example of wall nodule with perimeter value greater than zero.

Vertical Ratio: This feature, which can be visualized as the 3D equivalent of axis ratio, is the depth of nodule candidate less one slice divided by the longest length of all the tight bounding boxes for each 2D cross section [12]. Of course this feature is nonzero for those nodule candidates that are more than one slice thick. This feature can be helpful in discriminating between long vertical vessels and the more spherical nodules.

Position of candidate in lung context (2D): If centroid of the 3D candidate is located on slice number n, the shortest distance between that centroid and centroid of lung cross section, which also appears in slice n and encompasses the candidate is the position feature. This feature indicates how far from the center of lung the potential nodule candidate is located and is of importance for two main reasons: a) Based on experimental observations, the majority of true nodules detected by the automated nodule detection system suggested in previous chapter are located close to lung boundaries and far from the centroid of lung. b) According to the expert radiologist, who assessed the scans for nodule verification, the nodules which are situated close to the center of lungs (around the boundary of upper and lower lobes) are more likely to be missed. The position feature is therefore an interesting feature for both the purpose of classification and also for further nodule CAD research. Figure 4.9 shows the histogram of candidate position feature for all those true nodules that

were detected by the rule based CAD system of [12]. It can be seen that in agreement with the point mentioned in reason (a) above, the greater population of true nodules is accumulated around the higher distance values.
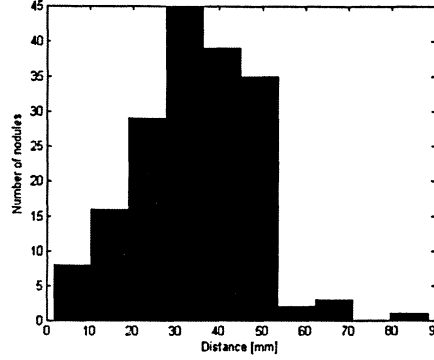
Figure 4.9: Histogram of nodule position feature
Histogram of distance between centroid of nodule and centre of the lung cross section, which surrounds that nodule for all those true nodules that were detected by the rule based CAD system of [12].

4.2.2   Fuzzy C-Means Clustering

Among the dataset being analyzed by the CAD system, there are cases that contain no nodules. These cases are often the cause for large overall false positive per slice rate. Fuzzy c-means clustering is exploited in the system with the aim of identifying such cases.

Fuzzy c-means (FCM) is a method of clustering which allows one piece of data to belong to two or more clusters. This method (developed by Dunn in 1973 [38] and improved by Bezdek in 1981 [39]) is frequently used in pattern recognition. Fuzzy c-means clustering is an unsupervised leaning method and clusters sample data based on their given features by assigning some graded or "fuzzy" cluster membership to each sample. These memberships are equivalent to probability $\hat{P}(\omega_i | x_j, \hat{\theta})$ where x_j is the jth of d-dimensional measured data, ω_i represents cluster i and $\hat{\theta}$ is the parameter vector for the membership function. The fuzzy c-means clustering algorithm seeks a minimum of a heuristic global cost function J_{fuc} as in equation (4.1).

50

$$J_{fuz} = \sum_{i=1}^{c} \sum_{j=1}^{n} [\hat{P}(\omega_i | x_j, \hat{\theta})]^b \|x_j - \mu_i\|^2 \tag{4.1}$$

where $\hat{P}(\omega_i | x_j, \hat{\theta})$ is the degree of membership of x_j in the cluster ω_i, b is a parameter chosen to adjust the blending of different clusters, μ_i is the cluster center for cluster i and $\|*\|$ is any norm expressing the similarity between any measured data and the center. It can be shown that at the solution (i.e., the minimum of J_{fuz}) the cluster center to which the sample belongs and the probability of membership can be calculated as shown in (4.2) and (4.3).

$$\mu_i = \frac{\sum_{j=1}^{n} [\hat{P}(\omega_i | x_j)]^b x_j}{\sum_{j=1}^{n} [\hat{P}(\omega_i | x_j)]^b} \tag{4.2}$$

$$\hat{P}(\omega_i | x_j) = \frac{(1/d_{ij})^{1/(b-1)}}{\sum_{r=1}^{c} (1/d_{rj})^{1/(b-1)}} \quad \text{and} \quad d_{ij} = \|x_j - \mu_i\|^2 \tag{4.3}$$

In general, the J_{fuz} criterion is minimized when the cluster centers μ_i are near those points that have high estimated probability of being in cluster j [36]. The fuzzy c-means algorithm can be summarized in the following steps:

1. *Initialize P=[p_{ij}] matrix, $P^{(0)}$*

2. *At k-step: calculate the centers vectors $\mu^{(k)}=[\mu_i]$ with $P^{(k)}$*

$$\mu_i = \frac{\sum_{j=1}^{n} p_{ij}^b \cdot x_j}{\sum_{j=1}^{n} p_{ij}^b}$$

3. *Update $P^{(k)}$, $P^{(k+1)}$*

$$p_{ji} = \frac{1}{\sum_{r=1}^{c} \left(\frac{\|x_j - \mu_i\|}{\|x_j - \mu_r\|} \right)^{\frac{2}{b-1}}}$$

4. *If (|| $P^{(k+1)}$ - $P^{(k)}$ ||< ε) OR (the maximum number of iterations (K_{max}) is reached) then STOP; otherwise return to step 2.*

Table 4.2 displays the selected parameter values while applying the fuzzy c-means clustering in the current lung nodule false positive reduction research. These parameters were introduced earlier in equations (4.1)-(4.3) and the algorithm summary above.

Symbol	Description	Value
x_j	Sample j	2 features for each sample candidate: (1) compactness and (2) area
c	Number of clusters (i.e., nodule and non-nodule)	2
n	Number of sample data	Variable for each case
$P^{(0)}$	Initial fuzzy partition matrix	A randomly generated cxn matrix.
b	Exponent for the matrix P	2
K_{max}	Maximum number of iterations	100
ε	Minimum amount of improvement	1e-5

Table 4.2: Fuzzy c-means clustering parameters for the specific application of lung nodule FP reduction in current research

For each sample candidate, features area and compactness are chosen to be used in the fuzzy c-means clustering step. These features proved to be effective in terms of clustering the sample candidates into two different groups of nodules and non-nodules. Based on these features, each sample is given a membership grade. The values 0 and 1 indicate no membership and full membership respectively. Grades between 0 and 1 indicate that the data point has partial membership in a cluster. In order to find the best location for the clusters, the algorithm updates itself by minimizing the cost function J_{fuz} (4.1) at each iteration. The fuzzy c-means clustering algorithm continues the clustering procedure until the results converge, that is when the improvement of cost function between two consecutive iterations is less than a specified minimum value or when the maximum number of iterations is reached.

For each patient, if the number of objects clustered as nodule after convergence is smaller than a threshold, that case is considered to be healthy or almost healthy. Hence, except those few candidates clustered as nodules all the other objects, i.e., objects in the non-nodule cluster, are considered false positives and discarded. The threshold is equal to 5% of the total number of candidates in the case with least number of detected objects. This means

that in worst case, which would be the case of patient with minimum number of detected candidates, for the case to be counted nodule free, at least 95% of the samples should be clustered as non-nodule after fuzzy c-means clustering. Figure 4.10(a) depicts the result of fuzzy c-means clustering on the two dimensional feature space of candidate area and compactness, for a patient with no true nodules. Figure 4.10(b) on the other hand shows an example of clustering result for a case with multiple nodules. As it is apparent from these figures, the fuzzy c-means clustering does a good job of separating suspicious (potential true nodule) samples from healthy ones. Those cases that do not satisfy the mentioned threshold constraint are assumed to contain multiple nodules and are therefore forwarded to the next FP reduction module for adaptive false positive analysis and amputation.

(a)   (b)
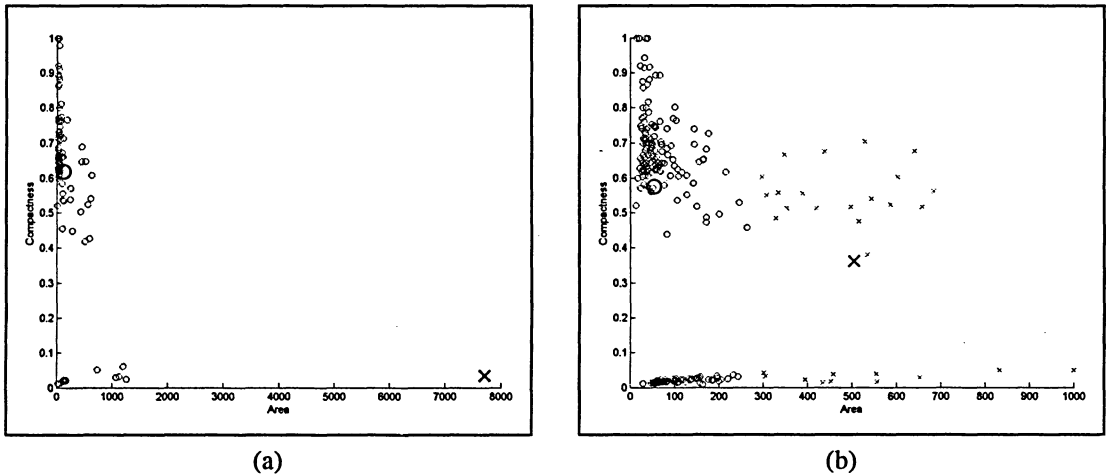
Figure 4.10: Result of fuzzy c-means clustering based on candidate's area and compactness
Circles represent members of non-nodule cluster and crosses represent members of nodule cluster. Bold markings point out center of each cluster type. (a) Patient with no nodules. (b) Patient with multiple nodules.

4.2.3   Iterative Linear Discriminant Analysis

After the fuzzy clustering step, those cases that are designated to contain multiple nodules are processed by a novel iterative linear discriminant approach.

In general, discriminant analysis (DA) is a supervised learning method that predicts in which group an individual belongs based on a pattern learnt from labeled training data and prior probabilities for the groups. The DA combines (i.e., weighs) the features so that a single

new composite variable, the discriminant function, is produced. Discriminant functions are the uncorrelated linear combinations of the original features:

$$Z = W_1 X_1 + W_2 X_2 + \ldots + W_i X_i \qquad ; X_i = \text{Vector containing feature } i \qquad (4.4)$$
$$; W_i = \textit{Weights}$$

Linear discriminant analysis (LDA) seeks a linear combination that has a maximal ratio of the separation of the class means to the within-class variance. This can be expressed in mathematical notations as follows.

For a labeled set of training vectors:

$$\tau_{XY} = \{(x_1, y_1), \ldots, (x_l, y_l)\}, \qquad x_i \in \Re^n, y \in Y = \{1, 2, \ldots, C\} \qquad (4.5)$$

The within-class S_W and between-class S_B scatter matrices are described as

$$S_W = \sum_{y \in Y} S_y, \qquad (4.6)$$

$$S_B = \sum_{y \in Y} |I_y| (\mu_y - \mu)(\mu_y - \mu)^T, \qquad (4.7)$$

where $S_y = \sum_{i \in I_y} (x_i - \mu_i)(x_i - \mu_i)^T$, and the total mean vector μ and the class mean vectors μ_y are defined as:

$$\mu = \frac{1}{l} \sum_{i=1}^{l} x_i, \qquad \mu_y = \frac{1}{|I_y|} \sum_{i \in I_y} x_i, y \in Y \qquad (4.8)$$

$|I_y|$ is the number of samples labeled y.

The goal of LDA is to train the linear data projection

$$z = W^T x, \qquad (4.9)$$

such that the class separability criterion

$$F(W) = \frac{\det(S_B)}{\det(S_W)} \qquad (4.10)$$

is maximized [40].

In the current study, the proposed iterative LDA algorithm is adaptive in the sense that it seeks the appropriate level of discrimination for each patient. Similar to the work done by Armato *et al.* [16], the nature of our LDA is a leave-one-out-by-case scheme, in which nodule candidates of all except one of the cases are used to train the classifier, which is then utilized to classify the nodule candidates from the remaining case. The linear classifier is fed by the 9 features discussed in section 4.2.1. The novelty of our approach however comes from the fact that unlike previous studies which use fixed classification thresholds for discriminant analysis of all candidates, the present approach updates its discrimination level based on the information attained from previous iterations of LDA. Figure 4.11 depicts a more detailed view of the iterative module. For all cases, the discriminant analysis begins with assigning a prior probability of 95% for membership in nodule class and 5% for membership in non-nodule class. This is a very cautious allocation, as we desire to retain as many true positive detections as possible. At each iteration, the linear classifier classifies the test candidates into nodule and non-nodule classes. If the total number of candidates in nodule class at one iteration (e.g., iteration i in figure 4.11) is different from that of the iteration before it (e.g., iteration i-1 in figure 4.11), and that difference is less than a cutoff threshold (T_2 in figure 4.11), it is implied that the LDA with current membership probabilities has not yet been able to detect many false positives. Therefore the membership probabilities are altered accordingly, i.e., the prior probability for nodule class membership is decreased by 5% while the prior probability for non-nodule class membership is increased by 5%. The LDA is applied on the candidates with these new values once again. In other words the LDA starts affecting the test candidates with very low aggressiveness and it becomes more and more aggressive as the number of iterations increases. The more aggressive LDA acts, the greater number of candidates it classifies as false positives. This iterative process is terminated when a big difference (cutoff threshold) is detected between size of nodule class in two consecutive runs of LDA. At that point the system considers the groupings arrived at iteration one before last to be the final classification result of the iterative LDA module. The cutoff threshold is chosen to be 20% of the total number of candidates for the case with least number of detected objects. This means that in the worst case, which would be the case with least number of candidates, if more than one fifth or 20% of the total number of candidates is

removed through an iteration of the iterative classification process, the process is halted at that stage. Experiments showed that this is a reasonable threshold as it is not too small to immediately stop the iterative process (under estimation of FPs), while it is not also too large to allow for over detection of false positives.
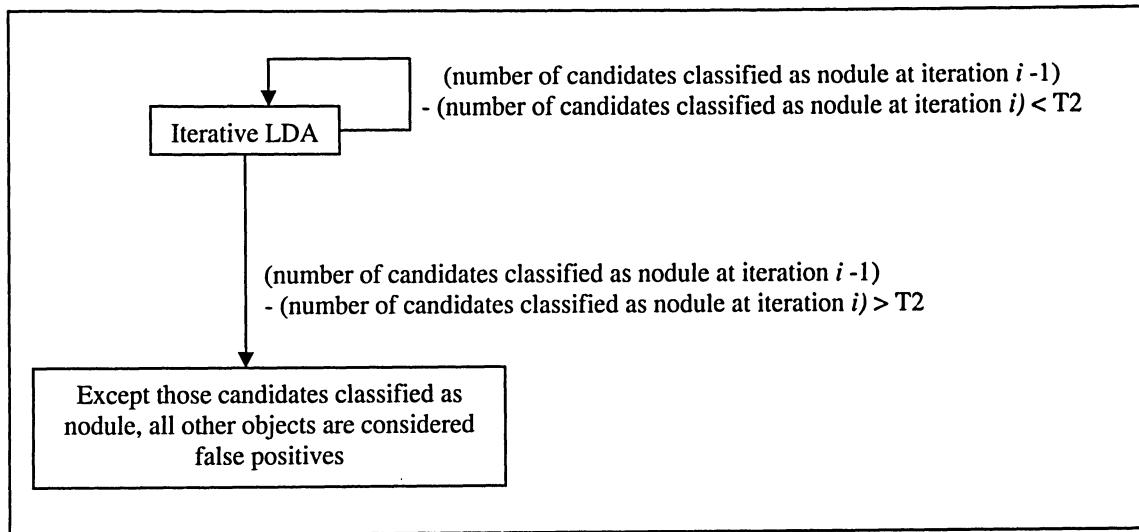
Figure 4.11: A closer look at the iterative linear discriminant analysis module

Iterative LDA

(number of candidates classified as nodule at iteration i -1) - (number of candidates classified as nodule at iteration i) < T2

(number of candidates classified as nodule at iteration i -1) - (number of candidates classified as nodule at iteration i) > T2

Except those candidates classified as nodule, all other objects are considered false positives

5 Results

This chapter provides experimental results of the methodologies discussed in previous two chapters. The dataset used for experiments is introduced in section 5.1. Results of the discussed ROI detection system is presented in section 5.2. The results acquired by using the hybrid learning FP reduction scheme is explained in section 5.3. Finally in section 5.4, performance results are compared with some of the previous research found in the literature.

5.1 Dataset

Dataset composition may have a significant impact on the functionality of system. Lack of an universal dataset to use as gold standard for testing various thoracic CAD methods is still a significant pitfall that researchers have to face. Some researchers have tested their detection approaches on simulated (synthetic) nodules. While this can be considered as a good initial step for getting insight about system performance, it can never replace the requisite for system performance assessment on real human nodules. The biological intricacies and variations associated with clinical subjects are significant challenges that need to be considered, resolved and refined by a CAD system.

Database of this research consists of 24 thoracic helical CT scans acquired from 24 different child patients in The Hospital for Sick Children, Toronto, Canada. 12 patients were female and 12 were male and the mean age of patients at time of scan was 12 years. The CT examinations were performed on a GE Medical Systems CT scanner and were acquired with a protocol of 120mA to 160mA (mean 138mA), 120 kVp. Each 5 mm thick reconstructed CT slice has an image matrix size of 512 x 512 pixels[1]. The database consists of total of 1190 slices, and the number of slices per scan ranges from 29 to 68 (average 50 slices per scan).

It has been tried to include a broad range of nodule size and distribution in the dataset of this pediatric study. The dataset contains 154 metastatic true nodules, with 20 nodules being smaller than a 4 mm diameter circle. Size of 10 nodules fall below the area of 3 mm circle, which is assumed as the minimum nodule size criterion [2]. This criterion is smaller

---

[1] Images have been resized in the figures of this thesis for space limitation.

57

than what has been used in some of the other studies [10], [27], [28], [30] and [32]. Nodule population per scan ranges from 0 to 58 for different patients in the database (average 6.42 nodules per patient).

An experienced thoracic radiologist indicated the location of each nodule using a computer interface. The radiologist marked up the location, which he believed to be a nodule by a mouse click on each CT slice in which the nodule appeared.

5.2 ROI Detection Results

The 24 clinical subjects of this database were in fact three groups of patients, with CT scans performed at various dates (Table 5.1, column 1). Group A and B with total of 18 patients were used for developing the ROI detection system. The system was later tested with cases of group C. As will be presented soon in this chapter, the sensitivity and false positive rate attained by utilizing the ROI detection system discussed in chapter 3 was compared against the performance of a recently developed rule based CAD system [12] for each group of patients.

Tables 5.1 to 5.3 demonstrate the results attained in different phases of the ROI detection system discussed in chapter 3. The first three columns of each table show tag name for each case, number of nodules per case, and total number of CT slices per case respectively. The last four rows of every table, express the overall value of each performance parameter for patients in group A, B, C and all three groups combined.

Columns four to seven of Table 5.1 relate to system performance at the end of the morphological operations stage (refer to block diagram figure 3.1) and before any of the rule based FP reduction phases are performed. It can be observed from Table 5.1 that while the system achieves high sensitivity at this phase of work, it also detects a very large pool of false positive candidates. This is quite expected, as the combination of histogram analysis, thresholding and morphological operations specify many entities as potential nodule candidates. This calls for further refinement of the designated candidates, which is accomplished through the application of candidate size, shape and intensity analysis phases as were described in sections 3.2.5 and 3.2.6. Numerical results of those phases can be viewed in Tables 5.2 and 5.3.

Without breaking big objects into pieces, 3 nodules would have been missed and also many independent objects would have erroneously become connected, due to existence of long objects, which linked the pieces three dimensionally.

Case Tag	# of nodules	# of slices	# true detected	Sensitivity	# objects detected	FPs/slice
A1	1	40	0	0	860	21.50
A2	0	67	0	N/A	1953	29.15
A3	34	57	27	79.41	1986	34.37
A4	1	50	1	100	1381	27.60
A5	1	61	0	0	1785	29.26
A6	13	40	10	76.92	1335	33.13
A7	0	37	0	N/A	1722	46.54
A8	1	53	1	100	2046	38.58
A9	8	59	6	75	1930	32.61
A10	1	35	1	100	1056	30.14
B1	58	65	47	81.03	1910	28.66
B2	0	32	0	N/A	988	30.88
B3	13	61	10	76.92	2396	39.11
B4	1	46	1	100	2558	55.59
B5	0	51	0	N/A	917	17.98
B6	0	56	0	N/A	2026	36.18
B7	3	55	3	100	1177	21.35
B8	1	44	0	0	1574	35.77
C1	9	68	8	88.89	1377	20.13
C2	2	53	1	50	1800	33.94
C3	1	37	1	100	1453	39.24
C4	3	35	1	33.33	895	25.54
C5	0	59	0	N/A	1920	32.54
C6	3	29	3	100	1592	54.79
Set A	60	499	46	76.67	16054	32.08
Set B	76	410	61	80.26	13546	32.89
Set C	18	281	14	77.78	9037	32.11
A & B & C	**154**	**1190**	**121**	**78.57**	**38637**	**32.37**

Table 5.1: System performance at the end of the morphological operations stage
Data in this table is collected before any of the rule based filtering phases are performed. N/A stands for not applicable.

Sensitivity and false positive results of the system after the size and shape filtering phase are presented in Table 5.2. The total number of detected objects for each case and consequently the overall number of detected objects drops significantly (87% reduction in overall FP/slice rate compared to Table 5.1). Three Nodules are missed in the midst of this phase and hence the sensitivity performance is decreased by approximately 2% (76.62% versus 78.57%). The missed nodules did not satisfy the conditions set for area and axis ratio explained in 3.2.5.

Case Tag	# of nodules	# of slices	# true detected	Sensitivity	# objects detected	FPs/slice
A1	1	40	0	0	127	3.18
A2	0	67	0	N/A	209	3.12
A3	34	57	30	88.24	259	4.02
A4	1	50	0	0	119	2.38
A5	1	61	0	0	195	3.20
A6	13	40	9	69.23	178	4.23
A7	0	37	0	N/A	230	6.22
A8	1	53	1	100	212	3.98
A9	8	59	6	75	256	4.24
A10	1	35	0	0	150	4.29
B1	58	65	48	82.76	487	6.75
B2	0	32	0	N/A	141	4.41
B3	13	61	9	69.23	337	5.38
B4	1	46	1	100	299	6.48
B5	0	51	0	N/A	105	2.06
B6	0	56	0	N/A	191	3.41
B7	3	55	2	66.67	170	3.05
B8	1	44	0	0	205	4.66
C1	9	68	7	77.78	191	2.71
C2	2	53	0	0	201	3.79
C3	1	37	1	100	242	6.51
C4	3	35	1	33.33	138	3.91
C5	0	59	0	N/A	234	3.97
C6	3	29	3	100	368	12.59
Set A	60	499	46	76.67	1935	3.79
Set B	76	410	60	78.95	1935	4.57
Set C	18	281	12	66.67	1374	4.85
A & B & C	154	1190	118	76.62	5244	4.31

Table 5.2: System performance after size and shape filtering
FP/slice rate drops significantly while sensitivity is decreased less than 2%.

Table 5.3 illustrates the figures for final performance of ROI detection system. This includes the final intensity based filtering phase. The final sensitivity is 73.38% and FP rate is 3.39 FP/slice. This stage proved to be particularly effective in refining cases with higher FP rates. For instance consider case A3, B3 and B4. By comparing the numbers in Table 5.2 and 5.3 it can be observed that the FP/slice rate for these subjects reduced from 4.02, 5.38 and 6.48 to 2.47, 3.34 and 4.13 FPs/slice respectively. This translates to the fact that these cases contained many non-nodule objects which had comparable shape and size to that of nodules, because they did not get filtered in the shape and size analysis stage. Filtering based on local gray level density profile and mean pixel intensity helped to distinguish those objects from true nodules. Figure 5.1 plots the number of false positive objects per case after each filtering step. Between the second step (filtering based on local density profile) and the

third step (filtering based on object mean intensity), the latter almost always has a greater share in downsizing the number of non-nodule objects. According to figure 5.1, in cases A1, C4 and C6, the number of false positive objects detected by system after the third phase of FP reduction is larger than the number of false positives after the second phase. This is because objects are counted three dimensionally. i.e., 3D connected components are counted as one object, and hence some objects appear in more than one slice. What happened in cases A1, C4 and C6 is that through the third filtering step, a part of such thick object is removed from one of the 2D slices in which it appears. Consequently the object is broken into more than one object in the 3D space.

Case Tag	# of nodules	# of slices	# true detected	Sensitivity	# objects detected	FPs/slice
A1	1	40	0	0	127	3.18
A2	0	67	0	N/A	147	2.19
A3	34	57	28	82.35	169	2.47
A4	1	50	0	0	71	1.42
A5	1	61	0	0	127	2.08
A6	13	40	9	69.23	170	4.03
A7	0	37	0	N/A	180	4.86
A8	1	53	0	0	151	2.85
A9	8	59	6	75	192	3.15
A10	1	35	0	0	134	3.83
B1	58	65	48	82.76	440	6.03
B2	0	32	0	N/A	139	4.34
B3	13	61	9	69.23	213	3.34
B4	1	46	0	0	190	4.13
B5	0	51	0	N/A	90	1.76
B6	0	56	0	N/A	137	2.45
B7	3	55	2	66.67	111	1.98
B8	1	44	0	0	185	4.20
C1	9	68	6	66.67	143	2.01
C2	2	53	0	0	141	2.66
C3	1	37	1	100	217	5.84
C4	3	35	1	33.33	137	3.89
C5	0	59	0	N/A	181	3.07
C6	3	29	3	100	359	12.28
Set A	60	499	43	71.67	1468	2.86
Set B	76	410	59	77.63	1505	3.53
Set C	18	281	11	61.11	1178	4.15
A & B & C	154	1190	113	73.38	4151	3.39

Table 5.3: Final performance of ROI detection system, after the intensity based filtering phase
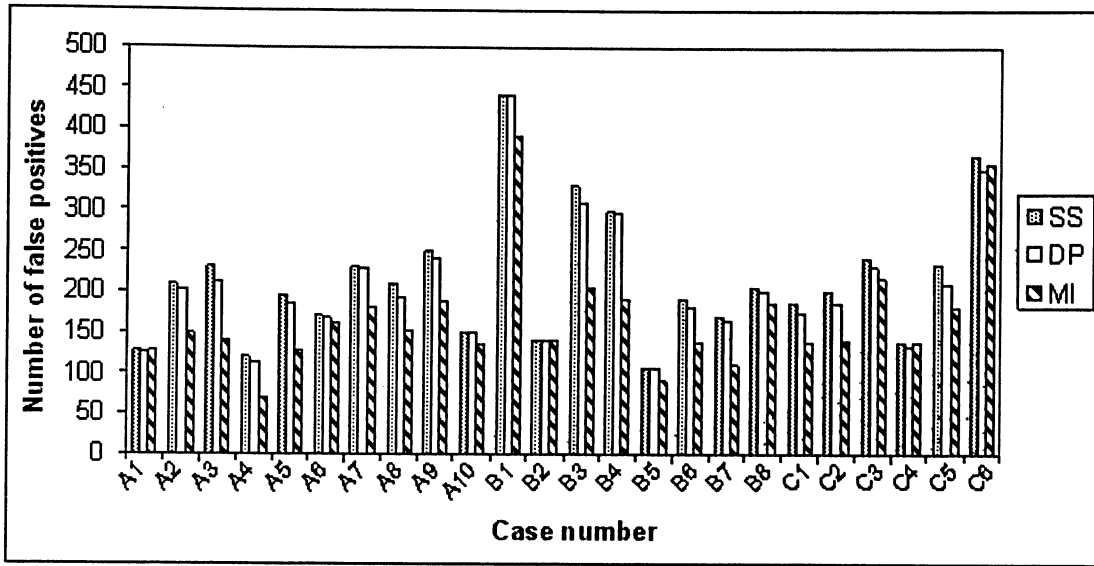
Figure 5.1: Number of detected objects after each rule based filtering step
SS stands for shape and size analysis, DP stands for density profile analysis, and MI stands for mean intensity analysis.

The final results of the ROI detection system proposed in this thesis, is comparable to some of the well-known lung nodule CAD studies in the literature [15-17]. The system performance is compared to that of a recently in-house developed rule based CAD system [12] in a case by case basis. Table 5.4 shows that comparison for the three groups of patients. The system discussed in this thesis outperforms the system of [12] in overall sensitivity. The FPs/slice rate in the current system is 0.24 (7%) higher than that of [12]. This is minor and acceptable considering the better sensitivity current system offers.

In terms of sensitivity, system in [12] beats the current system only for patients of group B. This is mainly because of the higher detection sensitivity of [12] in case B1, which is the case with greatest number of nodules. System of [12] detected 52 of the 58 true nodules existing in case B1, while the current system detected 48. Performance of the current system however is better than [12] both for patients of group A and C. The superiority of current system is clearly evident for patients of group C, for which the overall sensitivity of current system is 38.89% higher than that of [12] (61.111% versus 22.222%). Inspection showed that system of [12] was unsuccessful in detecting nodules which were either very small or had inconspicuous intensity difference with their surrounding lung context. The current ROI

detection system recognized such cases because of its consideration of local intensity variations. Several examples are shown in figure 5.2.

Case Tag	Sensitivity Current System	FPs/slice Current System	Sensitivity System of [12]	FPs/slice System of [12]
A1	0	3.16	0	2.23
A2	N/A	2.19	N/A	1.31
A3	82.35	2.47	73.53	2
A4	0	1.42	100	1.82
A5	0	2.08	0	2.07
A6	69.23	4.03	69.23	3.48
A7	N/A	4.86	N/A	4.14
A8	0	2.85	100	3.23
A9	75	3.153	75	2.34
A10	0	3.83	0	3.60
B1	82.76	6.03	89.66	5.51
B2	N/A	4.34	N/A	4.19
B3	69.23	3.34	76.92	3.18
B4	0	4.13	0	3.59
B5	N/A	1.76	N/A	1.96
B6	N/A	2.45	N/A	2.04
B7	66.67	1.98	66.67	2.18
B8	0	4.20	0	5.14
C1	66.67	2.01	11.11	1.19
C2	0	2.66	0	2.23
C3	100	5.84	0	6.86
C4	33.33	3.89	33.33	4.97
C5	N/A	3.07	N/A	2.58
C6	100	12.28	66.67	11.31
Set A	71.67	2.86	70	2.47
Set B	77.63	3.53	84.21	3.44
Set C	61.11	4.15	22.22	3.94
A & B & C	**73.38**	**3.39**	**71.43**	**3.15**

Table 5.4: Performance comparison for ROI detection system suggested in current thesis and system discussed in [12]

Another key advantage of the current ROI detection system over system of [12] and perhaps some other lung nodule CAD systems, is its fast performance or time efficiency. Figure 5.3 depicts the time elapsed for the ROI detection process in [12] and in the current system, on a case by case basis. The average elapsed time for processing one case (patient) by [12] is 1080 seconds (18 minutes). This is compared to 287 seconds (4.8 minutes) of average ROI detection time by the present system[2]. Therefore the current system showed 3.75 fold increase in performance time, while it produced similar –if not better- results.

---

[2] Both CAD systems were run on a computer with 1.6 GHz Pentium M processor and 1 GB of RAM.
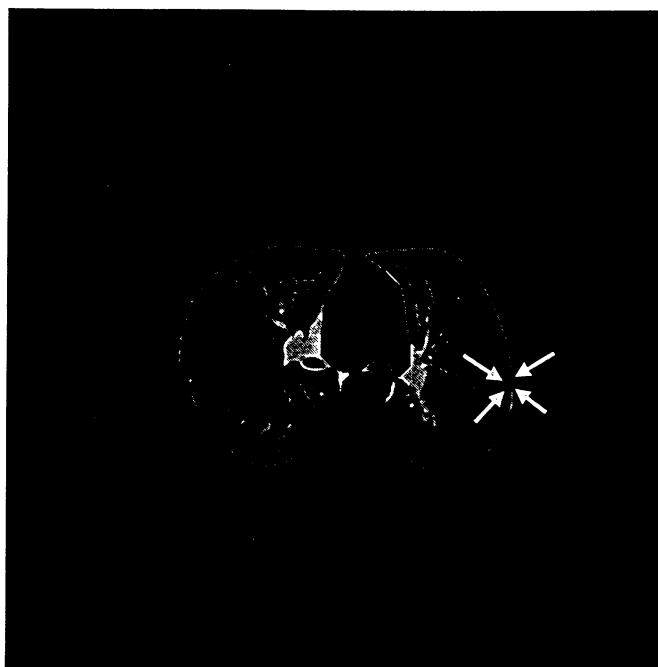
(a)
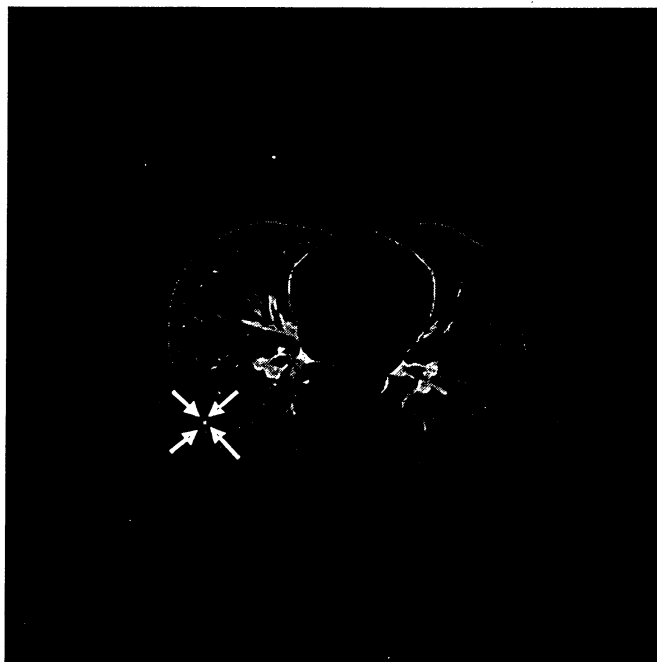
(b)

Figure 5.2: Examples of nodules missed by system of [12] but detected by the current ROI detection system

(a) Small nodule with inconspicuous intensity difference with its surrounding lung context. (b) Small well-formed nodule.

(c)
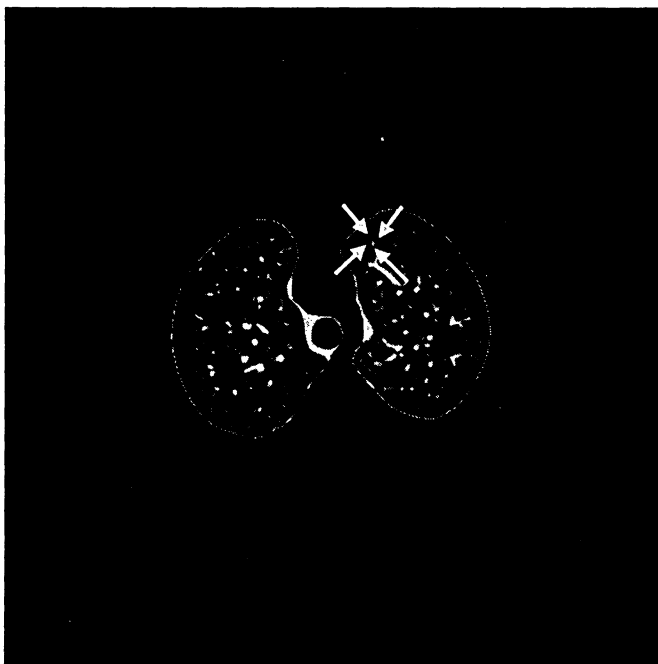
(d)

Figure 5.2 (continued): (c) Small pale nodule hardly recognizable from the background and nearby thin vessels. (d) Small irregular shaped nodule.
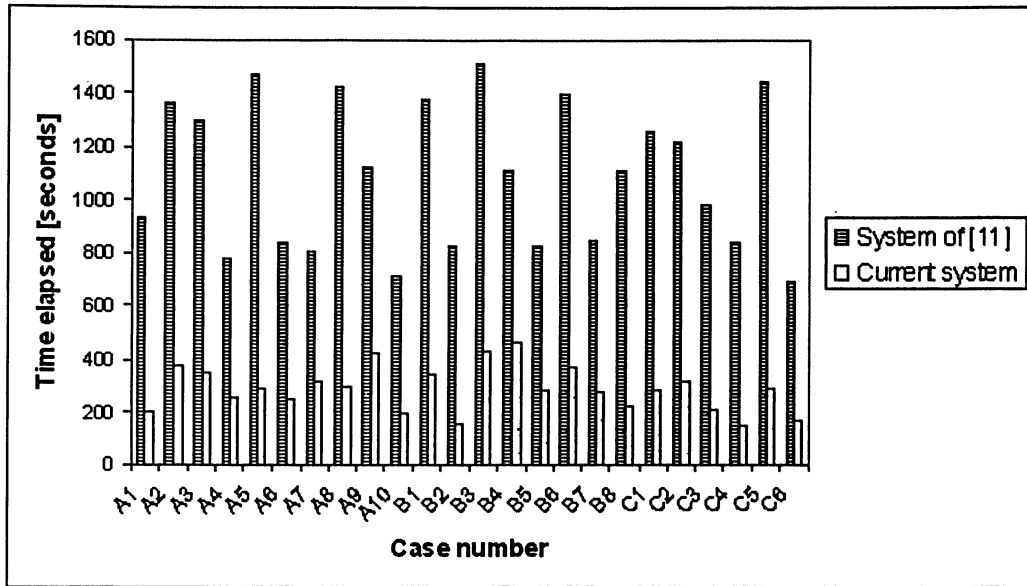
Figure 5.3: Time elapsed in seconds by the system of [12] and the current system for the ROI detection process, on a case by case basis

5.3 Enhanced FP Reduction Scheme Results

The enhanced FP reduction system, discussed in chapter 4, was applied to potential nodule candidates detected by two different ROI detection methods. The following sections explain each test.

5.3.1 Results of Applying the Enhanced FP Reduction Scheme on the ROI Detection System Discussed in Chapter 3

As explained in chapter 4, the enhanced FP reduction scheme is a hybrid approach based on learning algorithms. In order to test the proposed scheme, it was applied as an enhanced FP reduction stage to final candidates designated by the ROI detection system of chapter 3. The effect on each case together with the overall effect is summarized in Table 5.5. The overall FP/slice rate downgrades from 3.39 to 2.42, a 29% improvement in terms of false positive reduction. The overall sensitivity is decreased 1.3% (73.38% to 72.08%), due to missing two nodules in the course of the enhanced FP reduction stage.

66

Case Tag	Sensitivity - BEFORE enhanced FP reduction scheme	FPs/slice - BEFORE enhanced FP reduction scheme	Sensitivity - AFTER enhanced FP reduction scheme	FPs/slice - AFTER enhanced FP reduction scheme
A1	0	3.18	0	0.05
A2	N/A	2.19	N/A	0.01
A3	82.35	2.47	79.41	2.07
A4	0	1.42	0	0.04
A5	0	2.08	0	1.82
A6	69.23	4.03	69.23	3.78
A7	N/A	4.86	N/A	0.08
A8	0	2.85	0	2.75
A9	75	3.15	75	2.59
A10	0	3.83	0	3.83
B1	82.76	6.03	82.76	6.03
B2	N/A	4.34	N/A	0.09
B3	69.23	3.34	69.23	2.92
B4	0	4.13	0	3.48
B5	N/A	1.76	N/A	0.04
B6	N/A	2.45	N/A	1.96
B7	66.67	1.98	66.67	1.95
B8	0	4.20	0	3.57
C1	66.67	2.01	55.56	1.62
C2	0	2.66	0	2.38
C3	100	5.84	100	5.30
C4	33.33	3.89	33.33	0
C5	N/A	3.07	N/A	2.86
C6	100	12.28	100	11.93
Set A	71.67	2.86	70	1.65
Set B	77.63	3.53	77.63	2.70
Set C	61.11	4.15	55.56	3.37
A & B	75	3.16	74.26	2.15
A & B & C	73.38	3.39	72.08	2.42

Table 5.5: Impact of the enhanced FP reduction scheme (hybrid learning scheme) on the performance of the lung nodule detection system of chapter 3

The results are even superior, when only sets A and B are looked at (row one before last, Table 5.5). With the price of less than 1% decrement in sensitivity (74.26% versus 75%), the FP/slice rate is lowered 32%. Also according to the table, between the three sets, set A benefits the greatest FP reduction (2.86 FP/slice down to 1.65 FP/slice or 42% reduction).

Experiments show that without the fuzzy c-means clustering module, the overall FP/slice rate of Table 5.5 will jump up by about 8% while one more true nodule will be missed. This highlights another advantage of the hybrid structure of the proposed scheme.

Most of the features used to train the classifier in the enhanced FP reduction system are similar to features used earlier in the ROI detection system for purpose of filtering nodule objects from non-nodules ones. This might raise the argue that how can the enhanced FP reduction stage reduce the number of false positives even further? The answer refers to the ability of learning algorithms to converge multidimensional data into a single classification decision. While the components of the proposed enhanced FP reduction scheme, i.e., fuzzy c-means clustering and linear discriminant analysis are different in the sense that the former is an unsupervised learning method and the latter is a supervised one, the notion of both is to group data points that populate some multidimensional space into a specific number of different clusters. In the case of current research, although the candidates are filtered in the ROI detection phase based on the criteria set on shape, size or intensity features, the filtering steps are independent from each other. In the enhanced FP reduction scheme however, the learning algorithms fuse the various features in order to minimize the group membership error and maximize the differences between groups. The groups are nodules and non-nodules of course.

At the fixed false positive rate of 2.4 FP/slice, overall sensitivity of the computer-aided lung nodule detection system discussed in this thesis when nodules smaller than 3 mm in diameter are excluded is 77.08% and when nodules smaller than 4 mm in diameter are excluded, it is 82.84%.

5.3.2 Results of Testing the Enhanced FP Reduction Scheme on a Previously Developed ROI Detection System

As a second test, the enhanced FP reduction system was applied to nodule candidates detected by the rule based lung nodule CAD system developed by [12]. Table 5.6 shows the performance numbers of that system before and after applying the enhanced FP reduction scheme. The overall FP/slice rate is enhanced by almost 38% (3.15 FP/slice declined to 1.96 FP/slice), while the overall sensitivity dropped 2.6%.

For patients of group B, no nodule was missed through the enhanced FP reduction stage, thus the sensitivity for that group remained at 84.2%. The other two groups of patients (A and C) lost two nodules each in the midst of the enhanced FP reduction stage. Without

group C, which suffered from poor sensitivity even before the enhanced FP reduction stage, the final sensitivity and FP/slice figures would be 76.47% and 1.67 respectively. Comparing these numbers to analogous quantities (i.e., sensitivity and FP/slice for groups A and B only) before the enhanced FP reduction stage, shows a 43% drop in FP rate and only 1.47% cutback in sensitivity. This comparison can be viewed in row one before last in Table 5.6. These results are quite promising when judged against results reported by other popular lung nodule CAD systems [11], [17], and [18].

Case Tag	Sensitivity - BEFORE enhanced FP reduction scheme	FPs/slice - BEFORE enhanced FP reduction scheme	Sensitivity - AFTER enhanced FP reduction scheme	FPs/slice - AFTER enhanced FP reduction scheme
A1	0	2.23	0	0.10
A2	N/A	1.31	N/A	0.13
A3	73.53	2	70.59	1.47
A4	100	1.82	0	0.04
A5	0	2.07	0	0.07
A6	69.23	3.48	69.23	2.95
A7	N/A	4.14	N/A	0.08
A8	100	3.23	100	2.51
A9	75	2.34	75	1.86
A10	0	3.60	0	0.06
B1	89.66	5.51	91.38	5.23
B2	N/A	4.188	N/A	0.03
B3	76.92	3.18	76.92	2.97
B4	0	3.59	0	3.39
B5	N/A	1.96	N/A	0.10
B6	N/A	2.04	N/A	1.18
B7	66.67	2.18	33.33	1.60
B8	0	5.14	0	4.82
C1	11.11	1.19	0	0.04
C2	0	2.23	0	0
C3	0	6.86	0	6.57
C4	33.33	4.97	33.33	4.57
C5	N/A	2.58	N/A	2.25
C6	66.67	11.31	33.33	9.31
Set A	70	2.47	66.67	0.94
Set B	84.21	3.44	84.21	2.56
Set C	22.22	3.94	11.11	2.88
A & B	77.94	2.91	76.47	1.67
A & B & C	71.43	3.15	68.83	1.96

Table 5.6: Impact of the enhanced FP reduction scheme (hybrid learning scheme) on the performance of the lung nodule detection system of [12]

The enhanced FP reduction system showed very good performance in retaining those nodules that are hardly distinguishable from non-nodule components. Several instances can be seen in figures 5.4-5.6. In all of these figures the candidates enclosed in circle present true nodules, that were correctly detected by the initial ROI detection system and were also retained after the enhanced FP reduction stage. The objects appearing inside rectangles relate to false positives erroneously detected by the initial ROI detection system, which were successfully recognized as false positives (i.e., were classified as belonging to the group of non-nodules) after the enhanced FP reduction stage.
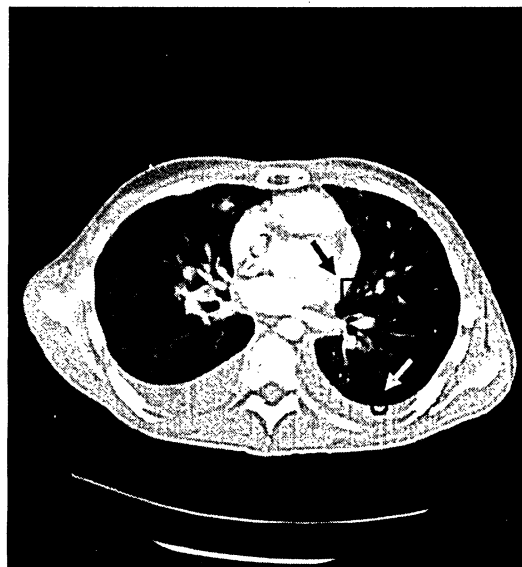
Figure 5.4: An example of successful nodule classification
The circle encloses a true wall nodule. The rectangle surrounds a very similar wall object that is in fact a false positive erroneously detected by the original CAD system. The current hybrid FP reduction approach recognizes and removes this.
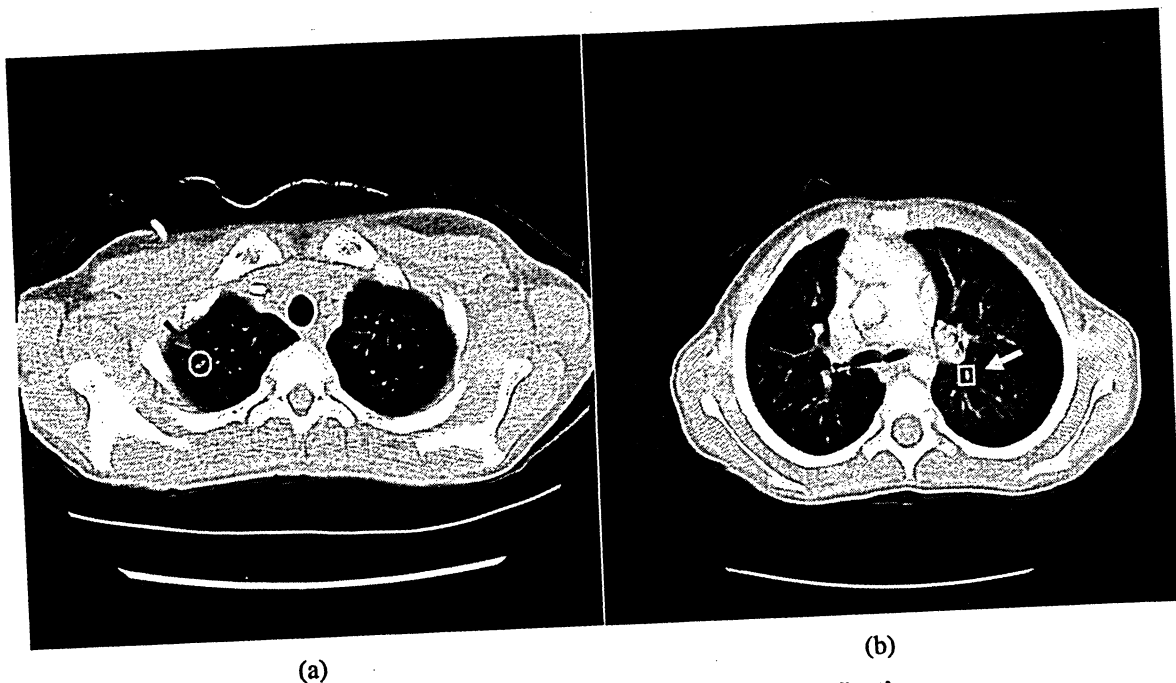
Figure 5.5: An example of successful nodule classification
(a) Circle surrounding a true interior nodule appearing with low gray level on dark background. (b) Rectangle surrounding a non-nodule candidate with very similar appearance as (a). The non-nodule object is correctly classified as a false positive after the hybrid FP reduction stage
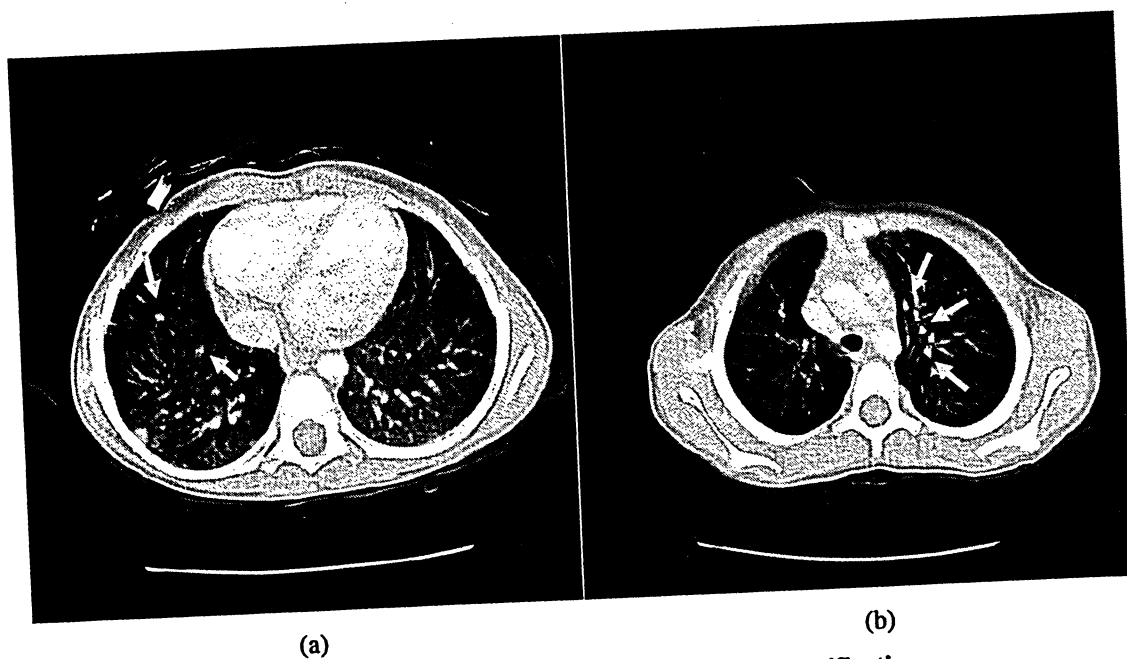
Figure 5.6: An example of successful nodule classification
(a) Bright well-formed true nodules marked inside circles. (b) Very similar non-nodule objects marked inside rectangles. The hybrid FP reduction stage categorizes these objects as false positives accurately.

5.4 Comparison of Results with Previous Works

A summarized comparison of performance results is demonstrated in Table 5.7. All the studies in this table followed a rule based or rule based combined with linear discriminant analysis method for computer aided lung nodule detection.

| Research	Year	Sensitivity	FP/slice
Armato *et al.* [15]	1999	70%	3.0
Gurcan *et al.* [18]	2003	84%	1.7
Armato *et al.* [17]	2005	70%	1.6
Dajnowiec *et al.*	2005	71%	3.2
Current Research	2006	72%	2.4

Table 5.7: Comparison of performance results with previous works

FP performance of current research is better than two of the previous studies (Armato 1999 and Dajnowiec 2005). The sensitivity of present research is better than all other works in the table except the work of Gurcan *et al*. The main reason for the superiority of results of that work is because the rules were designed based on *a priori* knowledge gained from radiologists such as the anatomical characteristics and morphological features of lung nodules (size, location and shape). Also total of 15 2D and 3D features were extracted and used in that study, which is greater than the number of features (9 features) used in the current thesis. The reader is advised that along with the possible benefit of result improvement that may be achieved by adding more features to the classification system also comes more time consumption and computation complexity.

Armato *et al.*, 2005 used jackknife LDA, in which candidates from half of the database, randomly chosen, were used for training and the other half for testing. Final performance was assessed based on the average result of multiple such jackknife runs. Training and testing of the automated nodule detection method were thus performed separately on subsets of nodule candidates from independent patients. Their database contained 393 patients, total of 11029 slice images and 470 lung nodules. Each case demonstrated at least one lung nodule [17]. A better result is expected from that study, as

their database is much broader than what was used for the current thesis. For a supervised learning method, the greater number of training samples, the more cases the classifier will be exposed to or the more it will learn. The great number of patients in the work of Armato *et al.*, 2005 also allowed them to train and test the classifier independently, which is definitely a fairer method compared to the leave-one-out-by-case (Round Robin) scheme used in the current thesis. The reader is reminded that in the leave-one-out-by-case scheme, LDA classifier is trained by all except one nodule candidate and the omitted candidate is subsequently used to test the trained classifier. Hence there is a big difference between when the left out test case has many nodules and when it has no or little number of nodules. When a case with many nodules is out for testing, the classifier has less to learn from compared to when a case with no/little number of nodules is out as the test case. However, exploiting the leave-one-out-by-case method was inevitable in the case of this thesis, due to dataset's limited number of patients.

It should also be noted that except current research and work of Dajnowiec 2005, which worked on a pediatric database, all other studies in Table 5.7 worked with adult patients.

6 Conclusions

A computerized system for detection of pulmonary nodules in computed tomography images is developed. The initiative for implementing such a system was to facilitate the diagnosis process for thoracic radiologists. Toward this objective, a collaborative research between the Medical Imaging Department of The Toronto Hospital for Sick Children and Department of Electrical Engineering of Ryerson University has been undertaken. The system developed for this thesis is the second practicable outcome of this collaboration after the rule based system developed by [12].

This research first started as a continuation to the system of [12]. As part of the future works suggested in [12], false positive reduction was mentioned as the main area that needed to be improved. The idea of design and development of an enhanced false positive reduction scheme that can be used as an improving step after any lung nodule CAD system was inspired from that need. Later a new ROI detection system was developed by the author. The ROI detection system together with the enhanced false positive reduction system makes a comprehensive lung nodule detection tool.

The ROI detection system developed for this thesis designates regions of interest or possible nodule candidates through gray level intensity and morphology analysis. By incorporating techniques such as contrast enhancement, histogram analysis, and thresholding, the system spots all the regions suspected to be nodules in lung. Then it pins down the candidates with highest possibility of being lung nodule through a series of rule based filtering stages. Those stages include applying morphological operations on the initial candidates and filtering the objects by analyzing their shape, size, local density profile and mean intensity. Results show the effectiveness of the ROI detection system in pulling out nodules from the lung context. After running the system on 24 pediatric cases, an overall sensitivity of 73.38% with 3.39 FP/slice was achieved. The cases belonged to three different groups of patients. ROI detection results for group A and B in particular (72% with 2.86 FP/slice and 78% with 3.53 FP/slice respectively) are quite encouraging. The database used in this research contains a broad variety of nodules both in terms of effective size and intensity and also in terms of distribution of nodules among subjects. Also, nodules appear in

74

various shapes and locations in different cases of the database. They include both interior and juxtapleural nodules, which range from round well-formed objects to irregular lesions. This is in fact considered a plus for the developed system, as it is not biased to a specific type of nodule. The multistage target specific structure of the discussed ROI detection system is unique. This is the only rule based system that targets each and every source of false positive findings specifically. That includes recognition of airway tube cross sections, elongated in-plane vessels and soft tissue, and cross sections of in-depth vessels. Another advantage of this ROI detection system is its processing time efficiency, which makes it suitable for use in clinical environments.

The enhanced false positive reduction scheme is a hybrid approach for reduction of false positive detections produced by any rule based lung nodule detection system. The scheme consists of two main modules, the fuzzy c-means clustering and the iterative linear discriminant analysis. Fuzzy clustering shows effectiveness in recognizing those cases which have no true nodules. The iterative LDA step does a very good job in adapting the level of discrimination with the aggressiveness of the classifier. A combination of morphological, gray level related, and contextual features constitute the features for training and testing of the classifier. This hybrid approach results in efficient drop of FP/slice rate, while maintaining the level of sensitivity. Testing this hybrid scheme on two different ROI detection systems (ROI detection system developed for this thesis and ROI detection system of [12]) resulted in 30% and 38% more FP drop respectively.

Lack of a universal dataset to use as gold standard for testing and comparing various thoracic CAD methods is a major issue. Many different lung nodule detection methods have been investigated in the literature. While some methods have reported extremely encouraging results [27], [31, 32], it is difficult to generalize the effectiveness of those methods. From studying the evolution of lung nodule computer aided detection and diagnosis, it appears that rule based approaches complemented with discriminant analysis are the most robust of all [15-18]. The results achieved by the computerized lung nodule detection and enhanced FP reduction system discussed in this thesis are comparable with research carried out by the well-known Kurt Rossmann Laboratories for Radiologic Image Research in University of Chicago [15-17]. In their recent publication, Armato *et al.* have claimed a sensitivity of 70% and 1.6 FP/slice [17]. System of current thesis is stronger in terms of sensitivity (72%) with

slightly higher FP rate (2.42/slice). The capability of the enhanced FP reduction scheme to converge multidimensional data into a single classification decision makes it a constructive procedure for recognizing false positive candidates in findings of rule based lung nodule CAD systems. The hybrid structure and adaptive nature of the enhanced FP reduction system are novel contributions exclusive to the present research.

All in all, the system developed in this Master's thesis for computerized detection of lung nodules with an enhanced false positive reduction scheme is definitely one step forward for applying concepts such as image processing, computer vision and machine learning in order to facilitate the process of conventional thoracic radiology.

6.1 Future Work

It has been tried to implement a practical lung nodule CAD system in the condensed time frame of a Master's research. However the current system can be further developed both in ROI detection and false positive reduction aspects. The primary venue of development is expanding the database. Growth of nodule database both in number and variety certainly adds to the universality and robustness of the system.

One useful step is perhaps to survey the CT scans in detail and seek image acquisition requirements that can upgrade the final image quality of scans. This can be achieved by imposing requirements on the scanner hardware or on the dose, setup or procedures through which the scans are acquired.

Investigations show that 7 nodules were missed in the lung extraction stage. As mentioned previously in chapter 3, lung extraction was performed by a 3D segmentation algorithm suggested in [12]. Thus another venue for development would be improving the performance of that algorithm or creating new segmentation algorithms in order to lose less – ideally no- true nodules in the midst of the lung segmentation process.

In the ROI detection part of work, the shape and intensity analyses can be extended to three dimensions. Comparison of the local density profile of similar locations in adjacent slices can provide insight about the nature of objects. For example, if density profiles of a single object in adjacent slices contain a sharp pulse at a certain coordinate, it can be concluded that the object is most likely a vertical vessel running through more than one slice.

In the enhanced false positive reduction section, more features can be added to the system for training and testing the classifier. Threshold, at which the fuzzy c-means clustering module or the iterative LDA module terminates, is another area to explore. Generally speaking, after discriminant analysis, each class will ideally have a normal distribution of discriminant scores; the degree of overlap between the discriminant score distributions can then be used as a measure of the success of the technique. It has been suggested [41] that when small samples are used and the normality assumption is violated, linear classification functions may be more reliable than quadratic functions [42]. In spite of all this, other types of discriminant analysis such as the quadratic form can be tested in the enhanced FP reduction scheme. Exploring the degree of effectiveness of other classification methods such as support vector machine (SVM) or relevance vector machine (RVM) may also be worthwhile.

Bibliography

[1] Definition of Biomedical Engineering, [online document] 2006 [cited 2006 March 15] Available at: http://www.whitaker.org/glance/definition.html

[2] S. G. Armato *et al.*, *"Lung image database consortium: developing a resource for the medical imaging research community"*, Radiology, Vol. 232, No. 3, pp 739-748, 2004.

[3] D. C. Rappaport, *"Helical CT applications in the thorax and abdomen"*, Postgraduate Medicine, Vol. 104, No. 5, p 105, November 1998.

[4] C. C. McCulloch *et al.*, *"Model-based detection of lung nodules in computed tomography exams"*, Academic Radiology, Vol. 11, No. 3, pp 258-266, March 2004.

[5] M. S. Brown *et al.*, *"Patient-specific models for lung nodule detection and surveillance in CT images"*, IEEE Trans. Medical Imaging, Vol. 20, No. 12, pp 1242-1250, December 2001.

[6] S. Sone *et al.*, *"Mass screening for lung cancer with mobile spiral computed tomography scanner"*, Lancet, Vol. 351, pp 1242-1245, 1998.

[7] J. H. Austin *et al.*, *"Glossary of terms for CT of the lungs: recommendations of the Nomenclature Committee of the Fleischner Society"*, Radiology, Vol. 200, pp. 327-331, 1996.

[8] C. Henschke *et al.*, *"CT screening for lung cancer: frequency and significance of part-solid and nonsolid nodules"*, AJR Am J Roentgenol, Vol. 178, pp 1053-1057, 2002.

[9] ML. Giger *et al.*, *"Current issues in CAD for mammography"*, Proceedings of the 3rd International Workshop on Digital Mammography, Elsevier Publishing, pp. 53-59, 1996.

[10] K. Kanazawa, Y. Kawata, N. Niki, H. Satoh, H. Ohmatsu, R. Kakinuma, M. Kaneko, N. Moriyama and K. Eguchi. *"Computer-aided diagnosis for pulmonary nodules based on helical CT images"*, Computerized Medical Imaging and Graphics, 22:157-167, 1998.

[11] Binsheng Zhao *et al.*, *"Automatic detection of small lung nodules on CT utilizing a local density maximum algorithm"*, Journal of Applied Clinical Medical Physics, Vol. 4, Number 3, pp. 248-259, 2003.

[12] Maciej Dajnowiec, *"A context based automated system for lung nodule detection in CT images"*, MASc. Thesis, Department of Electrical Engineering, Ryerson University, 2005.

[13] J. M. Goo *et al.*, *"Automated lung nodule detection at low-dose CT: preliminary experience"*, Korean Journal of Radiology, Vol. 4, No. 4, pp. 211-216, 2003.

[14] M. Dajnowiec, J. Alirezaie and P. Babyn, *"An adaptive rule based automatic lung nodule detection system"*, Pattern Recognition and Image Analysis, LNCS 3678:773-782, 2005.

[15] S. G. Armato III *et al.*, *"Computerized detection of pulmonary nodules on CT scans"*, Radiographics, Vol. 19, No. 5, pp. 1303-1311, 1999.

[16] S. G. Armato III, M. L. Giger and H. MacMahon. *"Automated detection of lung nodules in CT scans: preliminary results"*, Medical Physics, Vol. 28, No. 8, pp.1552-1561, 2001.

[17] S. G. Armato III, A. S. Roy, H. MacMahon, F. Li, K. Doi, S. Sone and M. B. Altman. *"Evaluation of automated lung nodule detection on low-dose computed tomography scans from a lung cancer screening program"*, Academic Radiology, Vol. 12, No. 3, pp. 337-346, 2005.

[18] M. N. Gurcan, B. Sahiner, N. Petrick, H. Chan, E. A. Kazerooni, P. N. Cascade and L. Hadjiiski, *"Lung nodule detection on thoracic tomography images: preliminary evaluation of a computer-aided diagnosis system"*, Medical Physics, Vol. 29, No. 11, pp. 2552-2558, 2002.

[19] H. P. Chan *et al.*, *"Computer-aided classification of mammographic masses and normal tissue: Linear discriminant analysis in texture feature space"*, Phys. Med. Biol. 40, pp. 857-876, 1995.

[20] J. W. Kung, S. Matsumoto, I. Hasegawa, B. Nguyen, L. C. Toto, H. Kundel and H. Hatabu, *"Mixture distribution analysis of a computer assisted diagnostic method for the evaluation of pulmonary nodules on computed tomography scan"*, Academic Radiology, Vol. 11, No. 3, pp. 281-285, 2004.

[21] S. K. Shah *et al.*, *"Computer-aided diagnosis of the solitary pulmonary nodule"*, Acad. Radiol., Vol. 12, No. 5, pp. 570-575, May 2005.

[22] H. Akaike, *"A new look at the statistical model identification"*, IEEE Trans Automatic Control, pp. 716-723, 1974.

[23] L. Bieman *et al.*, Classification and Regression Trees, Belmont, CA: Wadsworth International Group, 1984.

[24] I. Aleksander and H. Marton, An Introduction to Neural Computing, Chapman & Hall. 1990.

[25] Haykin S., Neural Networks. A Comprehensive Foundation, 2nd ed., New York: Macmillan College Publishing Company, 1994.

[26] M. G. Penedo *et al.*, *"Computer-aided diagnosis: a neural-network-based approach to lung nodule detection"*, IEEE Transactions on Medical Imaging, Vol. 17, No. 6, 1998.

[27] D. Lin and C. Yan, *"Lung nodules identification rules extraction with neural fuzzy network"*, Proceedings of the 9th International Conference on Neural Information Processing, Vol. 4, pp. 2049-2053, 2002.

[28] K. Suzuki *et al.*, *"Massive training artificial neural network (MTANN) for reduction of false positives in computerized detection of lung nodules in low-dose computed tomography"*, Medical Physics, Vol. 30, No. 7, pp. 1602-1617, 2003.

[29] K. Suzuki *et al.*, *"False-positive reduction in computer-aided diagnostic scheme for detection of nodules in chest radiographs by means of massive training artificial neural network"*, Academic Radiology, Vol. 12, No. 2, pp. 191-201, 2005.

[30] R. Sivaramakrishna *et al.*, *"Computer-aided detection of lung cancer from computed tomography images"*, Technology Review Journal, pp. 79-92, Fall/Winter 2002.

[31] J. Dehmeshki, J. Chen, M. V. Casique and M. Karakoy, *"Classification of lung data by sampling and support vector machine"*, Proceedings of the 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS), Vol. 2, pp. 3194-3197, 2004.

[32] A. Farag, A. El-Baz, G. Gimel'farb and R. Falk, *"Detection and recognition of lung nodules in spiral CT images using deformable templates and Bayesian post-classification"*, Proceedings of IEEE International Conference on Image Processing (ICIP), Vol. 5, pp. 2921-2924, 2004.

[33] The MathWorks R2006a Documentation [online document] [cited 2006 May], Available at: http://www.mathworks.com/access/helpdesk/help/toolbox/images/histeq.html

[34] Pratt, William K., Digital Image Processing, New York: John Wiley & Sons, Inc., 1991, p. 633.

[35] M. S. Brown *et al.*, *"Method for segmenting chest CT image data using an anatomical model: preliminary results"*, IEEE Trans. Med. Imag., Vol. 16, pp. 828–839, Dec. 1997.

[36] Duda, R.O., P.E. Hart and D.G. Stork, Pattern Classification, 2nd ed., John Wiley & Sons, 2001.

[37] Answers.com [online document] [cited February 2006], Available at: http://www.answers.com/voxel

[38] J. C. Dunn, *"A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters"*, Journal of Cybernetics, Vol. 3, pp. 32-57, 1973.

[39] J. C. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms, New York: Plenum Press, 1981.

[40] V. Franc, V. Hlavac, *"Statistical pattern recognition toolbox for Matlab, user's guide"*, Research Reports of CMP, Czech Technical University in Prague, No. 8, 2004.

[41] Stevens, J. P. Applied Multivariate Statistics for the Social Sciences (4th ed.), Mahwah, NJ: Lawrence Erlbaum Associates, 2002.

[42] Huberty, C. J. Applied Discriminant Analysis, Athens, Georgia: John Wiley & Sons, Inc., 1994.

Publications

Memarian, N., Alirezaie, J., and Babyn, P., "Computerized Detection of Lung Nodules with an Enhanced False Positive Reduction Scheme", Accepted to appear in Proceedings of IEEE International Conference on Image Processing (ICIP), Atlanta, USA, 8-11 October, 2006.

Memarian, N., Alirezaie, J. and Babyn, P., "A Novel Hybrid Learning Scheme for Computerized False Positive Reduction of Lung Nodule Candidates in CT images", Accepted to appear in Computer Assisted Radiology and Surgery, June-July 2006, Japan.

Memarian, N., Alirezaie, J., and Golshani, A., "Automated System for Image Analysis of Yeast Colonies: A Novel Application in Functional Genomics", Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 1120-1123, Toulouse, France, 14-19 May, 2006.

Memarian, N., Alirezaie, J., and Golshani, A., "Application of Computerized Image Processing in Functional Genomics: Preliminary Results", Proceedings of the IASTED International Conference on Biomedical Engineering (Biomed 2006), pp. 7-12, Innsbruck, Austria, 15-17 February, 2006.

Glossary

ANN Artificial Neural Network

CAD Computer Aided Diagnosis

CT Computed Tomography

DA Discriminant Analysis

FCM Fuzzy C-Means

FP False Positive

HU Hounsfield Unit

LCG Linear Combination of Gaussians

LDA Linear Discriminant Analysis

LLR Log Likelihood Ratio

MLT Multi Level Thresholding

MRI Magnetic Resonance Imaging

MTANN Massive Training Artificial Neural Network

NN Neural Network

QCI Quantized Convergence Index

ROC Receiver Operator Characteristic

ROI Region of Interest

RVM Relevance Vector Machine

STP Standard Temperature and Pressure

SVM Support Vector Machine

TP True Positive