

Title of thesis

**Expert System with an Embedded Imaging Module for
Diagnosing Lung Diseases**

I, KAVITHA SHAGA DEVAN

hereby allow my thesis to be placed at the Information Resource Center (IRC) of Universiti Teknologi PETRONAS (UTP) with the following conditions:

1. The thesis becomes the property of UTP.
2. The IRC of UTP may make copies of the thesis for academic purposes only.
3. This thesis is classified as

Confidential

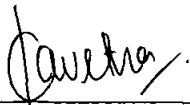
Non-confidential

If this thesis is confidential, please state the reason:

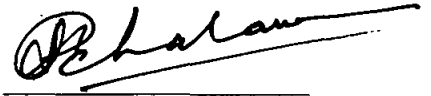
The contents of the thesis will remain confidential for _____ years.

Remarks on disclosure:

Endorsed by



Signature of Author



Signature of Supervisor

Permanent : 18, Lorong Cherry 1,
address Taman Cherry, 30100 Ipoh,
Perak, Malaysia.

Name of Supervisor
Prof. Dr. P.A Venkatachalam

Date : 16/1/06

Date : 16/1/06

Prof. Dr. P.A. Venkatachalam
Professor
Electrical & Electronic Engineering
Academic Block No. 22
Universiti Teknologi PETRONAS
Bandar Seri Iskandar
31750 Tronoh, Perak Darul Ridzuan, MALAYSIA.

UNIVERSITI TEKNOLOGI PETRONAS

Approval by Supervisor

The undersigned certify that they have read, and recommend to The Postgraduate Studies

Programme for acceptance, a thesis entitled

Expert System with an Embedded Imaging Module

for Diagnosing Lung Diseases

submitted by


Kavitha Shaga Devan

for the fulfilment of the requirements for the degree of

Masters of Science in Electrical and Electronics Engineering

16/1/06

Date

Signature : 
Main Supervisor : Prof. Dr. P.A Venkatachalam
Date : 16/1/06
Co-Supervisor 1 : A/P Ir. Dr. Ahmad Fadzil Mohd. Hani
Co-Supervisor 2 : _____

Prof. Dr. P.A. Venkatachalam
Professor
Electrical & Electronic Engineering
Academic Block No. 22
Universiti Teknologi PETRONAS
Tandjar Seberang
3150

11/1/06

UNIVERSITI TEKNOLOGI PETRONAS
Expert System with an Embedded Imaging Module
for Diagnosing Lung Diseases

By
Kavitha Shaga Devan

A THESIS
SUBMITTED TO THE POSTGRADUATE STUDIES PROGRAMME
AS A REQUIREMENT FOR THE
DEGREE OF MASTERS OF SCIENCE IN ELECTRICAL AND ELECTRONICS
ENGINEERING

Electrical and Electronics Engineering

BANDAR SERI ISKANDAR,
PERAK

JANUARY, 2006

DECLARATION

I hereby declare that the thesis is based on my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degree at UTP or other institutions.

Signature : 

Name : Kavitha Shaga Devan

Date : 16/1/06

ACKNOWLEDGEMENT

First and foremost, I would like to thank my supervisor, Prof. Dr. P.A Venkatachalam for developing my knowledge in the field of image processing and Artificial Intelligence. He has been very supportive and has shown much perseverance in an attempt to help me in my research. His guidance and support has given me much confidence to accomplish my research successfully.

I would like to express my gratitude to my co-supervisor, A/P Dr. Ir. Ahmad Fadzil Mohd. Hani who is also the Director of Postgraduate Studies Programme for his kind assistance and guidance throughout my research. Thank you to A/P Dr. Mohd Noh Karsiti, Head of Electrical & Electronics Engineering Programme for his help in providing the facilities and equipments needed for the successful completion of my research. My gratitude to him and A/P Dr. Varun Jeoti for providing me with valuable suggestions and recommendations during the course of my research. I wish to thank the staff of Postgraduate Studies Programme and Research Enterprise Office for assisting and helping me in time of need and support.

I would like to thank the Dato' Dr. Selliah, Head of Diagnostic Imaging of Ipoh General Hospital for his invaluable contribution to this research. He has given lots of assistance and advice on my research despite his hectic schedule. He has kindly provided me with the medical images used in this work. Gratitude is also expressed to the staff of the Ipoh General Hospital for the help they provided me. I would also like to thank the library staff of Perak Medical College for allowing me to borrow the medical books needed for my research.

I would like to take this opportunity to thank my parents and sisters for giving me all the support I needed to complete my research. No amount of gratitude would suffice for what they have done for me. I want like thank a few of my special friends who have helped me so much in my life. They have been a constant source of support for me.

ABSTRACT

Lung diseases are one of the major causes of suffering and death in the world. Improved survival rate could be obtained if the diseases can be detected at its early stage. Specialist doctors with the expertise and experience to interpret medical images and diagnose complex lung diseases are scarce. In this work, a rule-based expert system with an embedded imaging module is developed to assist the general physicians in hospitals and clinics to diagnose lung diseases whenever the services of specialist doctors are not available. The rule-based expert system contains a large knowledge base of data from various categories such as patient's personal and medical history, clinical symptoms, clinical test results and radiological information. An imaging module is integrated into the expert system for the enhancement of chest X-Ray images. The goal of this module is to enhance the chest X-Ray images so that it can provide details similar to more expensive methods such as MRI and CT scan. A new algorithm which is a modified morphological grayscale top hat transform is introduced to increase the visibility of lung nodules in chest X-Rays. Fuzzy inference technique is used to predict the probability of malignancy of the nodules. The output generated by the expert system was compared with the diagnosis made by the specialist doctors. The system is able to produce results which are similar to the diagnosis made by the doctors and is acceptable by clinical standards.

ABSTRAK

Penyakit paru-paru adalah sejenis penyakit yang telah menyebabkan banyak keperitan dan kematian di dunia. Kadar sembuh yang lebih tinggi boleh diperolehi jika penyakit ini dapat dikesan pada peringkat awalnya. Doktor pakar dengan pengalaman dan kepakaran untuk menganalisa imej-imej perubatan dan mendiagnosi penyakit paru-paru adalah amat kurang. Dalam pengkajian ini, satu sistem pakar berintegrasikan modul pemprosesan imej telah dibina untuk membantu doktor-doktor di hospital dan klinik untuk mendiagnos penyakit-penyakit paru-paru pada bila-bila masa perkhidmatan doktor pakar tidak dapat diperolehi. Sistem mahir ini mengandungi data-data perubatan daripada empat kategori iaitu maklumat peribadi dan perubatan pesakit, gejala-gejala penyakit, keputusan-keputusan ujian perubatan dan maklumat radiologi. Satu module pemprosesan imej telah diintegrasikan dalam sistem pakar untuk peningkatan imej X-Ray dada. Matlamat modul ini adalah untuk meningkatkan ciri-ciri imej X-Ray supaya ia boleh memberi maklumat yang sama seperti teknik-teknik imej perubatan yang lebih mahal seperti MRI dan CT Scan. Satu algoritma baru yang merupakan pengubahsuaian 'morphological top hat transform' telah dibina untuk meningkatkan pengesanan bintil-bintil paru-paru di imej X-Ray dada. Teknik 'Fuzzy Inference' telah digunakan untuk menganggar 'Malignancy Factor' bintil-bintil paru-paru. Keputusan yang diberi oleh sistem pakar telah dibandingkan dengan diagnosis doktor pakar. Sistem ini mampu menghasilkan keputusan yang hampir sama dengan diagnosis doktor pakar.

TABLE OF CONTENT

STATUS OF THESIS.....	i
APPROVAL PAGE.....	ii
TITLE PAGE.....	iii
DECLARATION.....	iv
ACKNOWLEDGEMENT.....	v
ABSTRACT.....	vi
ABSTRAK.....	vii
TABLE OF CONTENT.....	viii
LIST OF TABLES.....	xiii
LIST OF FIGURES.....	xii
ABBREVIATIONS.....	xvi
CHAPTER 1: INTRODUCTION.....	1
1.1 General Background.....	1
1.2 Objectives.....	3
1.3 Scope of Work.....	5
1.3 Organization of Thesis.....	6
CHAPTER 2: LITERATURE REVIEW.....	7
2.1 Lung Diseases.....	7
2.1.1 Risk Factors Causing Lung Diseases.....	10
2.1.2 Symptoms of Lung Diseases.....	11
2.1.3 Diagnosis of Lung Diseases.....	12
2.1.3.1 Chest X-Ray.....	16
2.1.3.2 Chest X-Ray Abnormalities.....	18
2.1.4 Lung Cancer.....	19
2.1.4.1 Lung Cancer Indicators: Lung Nodules.....	20
2.1.4.2 Diagnosis of Lung Cancer: Diagnostic Imaging.....	21

2.2	Computer Aided Image Analysis for Chest Radiography	22
2.2.1	General Image Processing Functions	22
2.2.1.1	Enhancement Filters	22
2.2.1.2	Edge Detection Filters	24
2.2.1.3	Geometrical Operations	24
2.2.1.4	Contrast Enhancement	25
2.2.1.5	Arithmetic Operations	27
2.2.1.6	Segmentation	27
2.2.1.7	Feature Extraction	29
2.2.1.8	Mathematical Morphology	30
2.2.2	Previous Work on Detection and Diagnosis of Lung Nodules.....	31
CHAPTER 3: EXPERT SYSTEMS		34
3.1	General Concept of Expert Systems.....	34
3.2	Features of Expert Systems	35
3.3	Structure of Expert Systems	37
3.4	Inference Mechanism	40
3.4.1	Fuzzy Inference Mechanism.....	41
3.5	Advantages of Expert Systems.....	42
3.6	Medical Expert Systems	43
3.6.1	Expert System for Lung Diseases.....	46
CHAPTER 4: DEVELOPMENT OF CHEST X-RAY IMAGE PROCESSING		
	TOOL.....	49
4.1	Digital Enhancement and Analysis of Chest X-Ray	49
4.2	Image Processing Tasks.....	52
4.2.1	Low-level Processing	53
4.2.1.1	Image Acquisition.....	53
4.2.1.2	Preprocessing.....	53
4.2.2	Intermediate-level Processing.....	55
4.2.2.1	Segmentation	55

4.2.2.2 Feature Extraction.....	56
4.2.3 High-level Processing.....	57
4.2.3.1 Classification	57

CHAPTER 5: DEVELOPMENT OF LUNG NODULES DETECTION AND

DIAGNOSIS TOOL	59
5.1 Overview of Lung Cancer	59
5.2 Methodology.....	60
5.2.1 Lung Nodule Detection	61
5.2.1.1 Morphological Transform.....	61
5.2.1.2 Region Growing Technique.....	65
5.2.2 Lung Nodule Diagnosis	65
5.2.2.1 Shape Feature Extraction.....	67
5.2.2.1.1 Compactness.....	67
5.2.2.1.2 Dispersion.....	67
5.2.2.1.3 Solidity.....	68
5.2.2.1.4 Moments	68
5.2.2.1.5 Eccentricity.....	71
5.2.2.2 Fuzzy Inference System.....	72
5.2.2.2.1 Fuzzification of Inputs.....	73
5.2.2.2.2 Application of Fuzzy Operators	79
5.2.2.2.3 Implication.....	79
5.2.2.2.4 Aggregation of Outputs	80
5.2.2.2.5 Defuzzification	81

CHAPTER 6: DEVELOPMENT OF AN EXPERT SYSTEM WITH AN

EMBEDDED IMAGING MODULE	83
6.1 Development of Expert Systems	83
6.1.1 Knowledge Acquisition	84
6.1.2 Knowledge Representation.....	85
6.1.2.1 Rules.....	85

6.1.2.2	Certainty Factor	87
6.1.3	Inference	89
6.1.4	Development of Knowledge Base	90
6.1.4.1	Patient's Personal and Medical Background	90
6.1.4.2	Clinical Symptoms	91
6.1.4.3	Clinical Test Results	93
6.1.4.4	Radiological Findings	94
6.1.4.5	List of Diseases	95
6.1.4.6	Decision Table	95
6.2	Expert System with an Embedded Imaging Module	97
6.2.1	Imaging Module Functions	98
6.2.2	System Operation	99
6.2.3	System Interface	101
CHAPTER 7: RESULTS AND DISCUSSION		106
7.1	Diagnoses of Benign Diseases by LUNEX	107
7.1.1	Patient A: Case Study	107
7.1.2	Patient B: Case Study	109
7.1.3	Patient C: Case Study	111
7.1.4	Case Studies of Patients D, E, F and G	112
7.2	Diagnoses of Malignant Diseases by LUNEX	115
7.2.1	Patient H: Case Study	116
7.2.2	Patient I: Case Study	117
7.2.3	Case Studies of Patients J, K and L	118
CHAPTER 8: CONCLUSION AND RECOMMENDATIONS		123
8.1	Conclusion	123
8.2	Recommendations	125
8.2.1	Recommendations for LUNEX	125
8.2.2	General Recommendations	126

REFERENCES	128
PUBLICATIONS.....	140
APPENDICES.....	141
Appendix A.....	141
Appendix B.....	166

LIST OF TABLES

Table 3.1:	Comparison of the expert system features with human experts and conventional programs	37
Table 5.1:	Shape factors.....	75
Table 5.2:	FIS rule base	79
Table 6.1:	Confidence factor categories	88
Table 6.2:	Confidence factor categories for clinical test results.....	88
Table 6.3:	Risk factors for lung diseases	91
Table 6.4:	List of clinical symptoms	92
Table 6.5:	List of clinical tests results	93
Table 6.6:	Chest X-Ray abnormalities.....	94
Table 6.7:	List of lung diseases	96

LIST OF FIGURES

Figure 2.1:	Chest CT Scan	14
Figure 2.2:	Chest MRI.....	14
Figure 2.3:	PET image	15
Figure 2.4:	Perfusion and ventilation scan.....	16
Figure 2.5:	Normal chest X-Ray.	17
Figure 2.6:	Lung diffuse abnormality	18
Figure 2.7:	Subtle diffuse abnormality.....	18
Figure 2.8:	Linear scarring.....	19
Figure 2.9:	Infiltrate	19
Figure 2.10:	Calcified lymph node.....	19
Figure 2.11:	Calcification.....	19
Figure 2.12:	Cyst with a cavity	19
Figure 2.13:	Abnormal hilum.....	19
Figure 3.1:	Structure of an expert system	38
Figure 4.1:	Image processing task.....	52

Figure 5.1:	Lung nodules detection and diagnosis process flow-chart	60
Figure 5.2:	Sample benign nodules	66
Figure 5.3:	Sample malignant nodules	66
Figure 5.4:	Fuzzy inference system	72
Figure 5.5:	Compactness membership function	76
Figure 5.6:	Dispersion membership function	76
Figure 5.7:	Eccentricity membership function	76
Figure 5.8:	Solidity membership function	77
Figure 5.9:	Moment 1 membership function	77
Figure 5.10:	Moment 2 membership function	77
Figure 5.11:	Moment 3 membership function	88
Figure 5.12:	Probability of malignancy membership function	78
Figure 5.13:	Aggregation for the set of crisp inputs	80
Figure 6.1:	Prototyping model	84
Figure 6.2:	Example of LUNEX rule	86
Figure 6.3:	Sample rule on patient's personal and medical history	91
Figure 6.4:	Sample rule on clinical symptoms	92
Figure 6.5:	Sample rule on clinical test results	93
Figure 6.6:	Sample rule on radiological findings	95
Figure 6.7:	Structure of the expert system with an embedded imaging module	97
Figure 6.8:	Functions in CIPT	98
Figure 6.9:	Process flow-chart of LUNEX	99
Figure 6.10:	LUNEX title screen	102
Figure 6.11:	Medical questions category	102
Figure 6.12:	Question screen on patient's history	103
Figure 6.13:	Question screen on clinical symptoms	103
Figure 6.14:	Question screen on clinical tests	103
Figure 6.15:	Question screen on radiology	104
Figure 6.16:	CIPT user interface	104
Figure 6.17:	Expert system with CIPT interface	105

Figure 7.1(a): Question screen.....	108
Figure 7.1(b): Original X-Ray of A.....	108
Figure 7.1(c): Processed image.....	109
Figure 7.1(d): Diagnosis by LUNEX.....	109
Figure 7.2(a): Question screen.....	109
Figure 7.2(b): Original X-Ray of B.....	109
Figure 7.2(c): Processed image.....	110
Figure 7.2(d): Diagnosis by LUNEX.....	110
Figure 7.3(a): Question screen.....	111
Figure 7.3(b): Original X-Ray of C.....	111
Figure 7.3(c): Processed image.....	111
Figure 7.3(d): Diagnosis by LUNEX.....	111
Figure 7.4(a): Original X-Ray of D.....	112
Figure 7.4(b): Processed image.....	112
Figure 7.4(c): Diagnosis by LUNEX.....	112
Figure 7.5(a): Original X-Ray of E.....	113
Figure 7.5(b): Processed image.....	113
Figure 7.5(c): Diagnosis by LUNEX.....	113
Figure 7.6(a): Original X-Ray of F.....	114
Figure 7.6(b): Processed image.....	114
Figure 7.6(c): Diagnosis by LUNEX.....	114
Figure 7.7(a): Original X-Ray of G.....	114
Figure 7.7(b): Processed image.....	114
Figure 7.7(c): Diagnosis by LUNEX.....	114
Figure 7.8(a): Original X-Ray of H.....	116
Figure 7.8(b): MGTH processed image.....	116
Figure 7.8(c): Region grown image.....	116
Figure 7.8(d): Malignancy factor display.....	116
Figure 7.8(e): Diagnosis by LUNEX.....	116
Figure 7.9(a): Original X-Ray of I.....	117
Figure 7.9(b): MGTH processed image.....	117

Figure 7.9(c): Region grown image.....	117
Figure 7.9(d): Malignancy factor display	117
Figure 7.9(e): Diagnosis by LUNEX.....	117
Figure 7.10(a):Original X-Ray of J	118
Figure 7.10(b): MGTH processed image.....	118
Figure 7.10(c):Region grown image.....	118
Figure 7.10(d): Malignancy factor display	119
Figure 7.10(e):Diagnosis by LUNEX.....	119
Figure 7.11(a):Original X-Ray of K	119
Figure 7.11(b): MGTH processed image.....	119
Figure 7.11(c):Region grown image.....	119
Figure 7.11(d): Malignancy factor display	120
Figure 7.11(e):Diagnosis by LUNEX.....	120
Figure 7.12(a):Original X-Ray of L	121
Figure 7.12(b): MGTH processed image.....	121
Figure 7.12(c):Region grown image.....	121
Figure 7.12(d): Malignancy factor display	121
Figure 7.12(e):Diagnosis by LUNEX.....	121

ABBREVIATIONS

ARDS	Adult Respiratory Distress Syndrome
ABG	Arterial Blood Gas
AI	Artificial Intelligence
ANN	Artificial Neural Networks
BTH	Black Top Hat
CAD	Computer-Aided Diagnosis
CBC	Complete Blood Count
CIPT	Chest X-Ray Image Processing Tool
CLAHE	Contrast Limited Adaptive Histogram Equalization
CO ₂	Carbon dioxide

CT	Computed Tomography
FDA	Food and Drug Administration
FEV1	Forced Expiratory Volume
FIS	Fuzzy Inference System
FVC	Forced Vital Capacity
HRCT	High Resolution CT
MGTH	Modified Grayscale Top Hat
MRI	Magnetic Resonance Imaging
O ₂	Oxygen
PA	Post Anterior
PET	Positron Emission Tomography
PFT	Pulmonary Function Testing
PPI	Pixels Per Inch
LUNEX	Expert System with an Embedded Imaging Module for Lung Diseases
ROI	Region of Interest
RV	Residual Volume
TLC	Total Lung Capacity
WTH	White Top Hat

CHAPTER 1

INTRODUCTION

1.1 General Background

One of the major diseases that people are suffering in both Malaysia and globally is lung diseases. Lung diseases are responsible for a considerable burden of suffering and death in all age groups. They currently rank among the 10 leading causes of death worldwide. In the year 2002, lung diseases caused more than 11 million deaths worldwide. This translates as 183 deaths per 100 000 population and accounts for 20% of all deaths. It is not only a fatal condition but also chronic, making it a constant struggle to stay alive. Projections for 2010 predict an increase in mortality rate. Lung conditions account up to one third of the demand for healthcare (World Health Organization, 2005 & Loddenkemper et al., 2003). The incidence of lung diseases is steadily increasing every year. This rise is caused by an increase in a number of risk factors such as cigarette smoking habits, pollution, heavy industrialization and the deterioration of socioeconomic conditions in certain countries.

Early detection could greatly improve the survival rate and generally raise the health quality of the people. Common detection methods for lung diseases include the detailed analysis of risk factors and clinical symptoms, physical and clinical lab tests as well as diagnostic imaging procedures. X-Rays, Computed Tomography (CT) scans and Magnetic Resonance Imaging (MRI) are some of the imaging procedures used for diagnosis. These imaging methods, with the exception of X-Ray, are very expensive and require the use of specialized equipments and personnel. This makes routine mass screening for lung diseases rather costly and impractical.

On the contrary, X-Ray is a very commonly used technique. It is non-invasive, cost-effective and widely available. However, diagnosing lung diseases using chest X-Rays is

a very challenging task for the radiologists as many lung abnormalities are not visible on the x-rays, especially at the initial stage of the disease (Come et al., 1999).

Specialist doctors and radiologists with the experience and expertise to diagnose complex lung problems and interpret medical images are scarce, especially in small hospitals and rural clinics. In Malaysia, the ratio of doctors to the general population is 1: 1377 (Ministry of Health Malaysia, 2004). Most often, only general physicians are available in most of the small medical centers. Difficulties are faced by the middle-class and rural population in obtaining early and effective treatment for lung related diseases due to the inadequate availability of specialized medical consultants and the ever increasing cost of specialist treatment.

Computer-Aided Diagnosis (CAD) systems have a very huge potential in the medical field. Humans and computers have complementary strengths that, when combined, have the potential to surpass the abilities of either alone. Humans can reason inductively, recognize patterns, apply multiple strategies to solve a problem, and adapt to unexpected events. Computers can store large quantities of information, recall data accurately, perform complex calculations, and execute repetitive actions reliably without tiring.

The utilization of Artificial Intelligence (AI) in the form of expert system and computer aided image analysis in the field of medicine is quite prevalent. Through the utilization of specific rules and algorithms derived from experts in the medical field, an expert system can simulate the thought process of medical specialists to make diagnoses with higher degree of consistencies. Computer-aided image analysis programs are able to aid doctors in enhancing and analyzing medical images for a detailed and comprehensive interpretation. The use of image processing techniques on X-Rays may render it to be more useful for early detection of diseases.

In this work, a rule-based expert system with an embedded imaging module is developed to diagnose lung diseases. The system uses rule based reasoning and fuzzy inferencing

for the diagnoses. A new algorithm, a modified version of morphological top hat transform is introduced for the enhancement of chest X-Ray images to increase the visibility of lung abnormalities such as lung nodules. This system has the potential to help general physicians to diagnose lung diseases in situations where specialist doctors and advanced facilities are not available. The system might be able to aid inexperienced doctors in producing consistent diagnoses. It also serves as a good training tool for the general physicians in the interpretation of chest X-Ray images. The system has been tested with about 60 test cases and the result obtained coincides with the diagnoses made by the specialist doctors.

1.2 Objectives

The objectives of this research work are outlined below.

- **To develop an expert system using rule-base approach and fuzzy inference to diagnose complex lung disease.**
- **To develop an imaging module for the enhancement and analysis of chest X-Ray images.**
- **To integrate the imaging module into the expert system.**
- **To test the diagnosis performance of the expert system in comparison to the diagnosis made by the human expert.**

The diagnosis of lung diseases requires the detailed analysis of a large amount of medical information. General physicians may not possess the expertise and experience to accurately diagnose a vast number of complex lung diseases, but he or she may have the basic knowledge to interpret the data. The expert system contains a large knowledge base of data from various categories such as patients' personal and medical history, clinical

symptoms, clinical tests and radiological information. These data are used by the expert system to arrive at accurate diagnosis. This helps the general physician to diagnose diseases without the need to consult specialist doctors. This will help in saving time, cost and travel expenses for the patients as well as make diagnostic expertise more widely available in the clinical community.

In this work, we have developed an imaging module called Chest X-Ray Image Processing Tool (CIPT) embedded as a component of the expert system. The goal of the CIPT is to enhance the chest X-Ray images to increase the visibility of abnormalities. CT scan and MRI are the preferred imaging procedures for diagnostic evaluations. However, these methods are very expensive and require the use of specialized equipments and personnel. X-Ray is a very common and low-cost procedure. CIPT contains many image processing functions that will enable the general physicians to detect and analyze abnormalities in the chest X-Ray. Thus, the doctors do not have to resort to more expensive methods such as CT scan and MRI.

The integration of the imaging module into the expert system is a very important feature. This feature enables the physicians to process the X-Ray images and answer the expert system questions at the same time. In this way, the user could provide more accurate answers to the expert system in questions relating to radiology as they would have much clearer images for their analysis. Thus, the user will just have to use single software to diagnose the diseases and enhance the X-Ray images. This promotes user-friendliness, and saves cost.

The diagnostic performance of the expert system will be assessed by comparing the diagnosis made by the system with the diagnosis of the specialist doctors. The system should be able to yield diagnosis which is similar to the diagnosis of the specialist doctors.

1.3 Scope of Work

This research work is focused on developing an expert system with an embedded imaging module for diagnosing lung diseases. This system is called LUNEX which means 'Lung Expert'. This expert system uses a rule-based approach and is built using an expert system shell called EXSYS CORVID.

The knowledge base contains about 500 rules pertaining to lung diseases. These rules are classified into three major categories which are patients' personal and medical history, clinical symptoms, clinical tests results and radiological information. The system is designed to diagnose 60 types of lung diseases. The system uses the information provided by the user and the rules in the knowledge base to perform its inferencing. The output will be a list of diseases with the corresponding probability of occurrences. The system has an explanation facility to provide reasoning to the user on how and why the system has reached a particular conclusion.

LUNEX diagnoses a large number of lung diseases in general but lung cancer in particular. Cancer of the lung is the most common killer among all malignancies (Parkin et al., 2005). The important application of LUNEX is to assist in the early detection and diagnosis of lung cancer even by a general physician in a rural clinic.

This imaging module contains image processing functions for enhancement, restoration, segmentation and analysis of the chest X-Ray images to help the general physicians to enhance the original X-Ray images in order to detect any abnormalities which are initially not clearly visible to the doctor's eye. The wide variety of functions provided enable the doctors to select suitable functions for the output that they require.

As the system focuses on lung cancer, a computer-aided lung nodules detection and diagnosis tool is developed and incorporated into CIPT. A new algorithm for the detection of lung nodules in chest X-Ray images is introduced. This algorithm uses

morphological transforms to detect lung nodules. Fuzzy inferencing technique is used to predict the probability of malignancy of lung nodules.

1.4 Organization of Thesis

Chapter 2 focuses on the various types of lung diseases and the methods used to diagnose them. Special focus is given on the characteristics and diagnosis of lung cancer. The formulation of general image processing methods is presented. A review of techniques used in the detection and diagnosis of lung diseases is given.

Chapter 3 gives an introductory theory on expert systems. The structure and components of the expert system are described in detail. The characteristics and functions of the expert system are also explained and the merits of expert system are also covered. A brief study on popular medical expert systems is also included.

Chapter 4 explains the development of CIPT. The various image processing tasks incorporated into CIPT is described.

Chapter 5 describes the formulation of the lung nodule detection and diagnosis tool. Detailed methodology and the associated functions are discussed.

The software development process of the imaging module embedded expert system is explained in chapter 6. The operation of the system is also included in this chapter.

Chapter 7 focuses on results and discussion. In this chapter, the results obtained from LUNEX for a number of test cases are discussed. The comments of specialists' doctors on the results of LUNEX are also included.

Chapter 8 concludes the research work. Recommendations for further improvements are also given.

CHAPTER 2

LITERATURE REVIEW

Introduction

Literature review is fundamental part of any research. Literature review enables a researcher to have a deeper understanding of the problem domain of the research. In this chapter, a brief overview on the various types of lung diseases and their diagnostic procedures are presented. Lung cancer is probably the most fatal of all the lung diseases. Thus, detailed explanation on lung cancer as well as its current methods of diagnosis is included. This is followed by some literature on the various diagnostic imaging procedures with special focus on chest X-Ray. The formulation of general image processing methods is described. The various types of techniques used in the detection and diagnosis of lung nodules are also presented.

2.1 Lung Diseases

During a normal day, we breathe nearly 25,000 times, and inhale large amounts of air. The air that we inhale constitute mostly of oxygen and nitrogen. The oxygen in the air travels from the lungs through the bloodstream to the cells in all parts of the body. The cells use the oxygen as fuel and produce carbon dioxide as a waste gas. The waste gas is carried by the bloodstream back to the lungs to be eliminated or exhaled. The lungs accomplish this vital process which is called gas exchange by using an automatic and quickly adjusting control system.

In addition to gas exchange, the lungs and the other parts of the respiratory system perform important functions such as regulating the air intake to suit the body's humidity and temperature and protecting the body from harmful substances by initiating coughing, sneezing, filtering or swallowing actions. The respiratory system defends the lungs with

some built in protective mechanisms such as cilia which are microscopic hairs along the air passages, phlegm (mucus or sputum) which collects dirt and germs inhaled into the lungs and moves them out to be coughed up or swallowed and macrophages which are scavenger cells in the lungs that literally eat up dirt and germs invading the lungs (Murray et al., 2000).

Lung diseases occur when there are damaging conditions in the lungs or when the respiratory system is unable to perform its normal functions optimally. Lung diseases can be classified into many categories depending on how they affect the lung and the respiratory system on the whole (Farzan et al., 1997).

Infectious Diseases of the Lung: These diseases are caused by bacteria, mycoplasma, fungi and virus infections that affect the lungs. In most instances, the infective agents enter the respiratory tract through the airways during breathing. Occasionally, infection occurs through the bloodstream. Examples of this disease are Acute Bronchitis, Pneumonia, Lung Abscess and Tuberculosis.

Obstructive Airways Diseases: A group of lung diseases characterized by the chronic obstruction to airflow within the lungs. This causes a decrease in the exhaled air flow and is normally caused by a narrowing or blockage of the airways. This type of disease is by far the most common chronic pulmonary disorder. Some examples of this disease are Asthma, Chronic Bronchitis, Cystic Fibrosis, Emphysema and Bronchiectasis.

Restrictive Lung Diseases: Diseases caused by decrease in the total volume of air that the lungs are able to hold. This type of disease is often characterized by abnormalities in the gas exchange mechanism in the respiratory system. Diseases of this type are Idiopathic Pulmonary Fibrosis, Sarcoidosis and Hypersensitivity Pneumonitis.

Environmental and Inhalation Lung Diseases: These diseases are pathologic conditions of the respiratory tract that result directly from the inhalation of various gaseous or

particulate matters in the air. Most often, these group of diseases is occupational related and is called as occupational lung diseases. Examples of this type of disease include Silicosis, Pneumoconiosis and Asbestosis.

Pulmonary Aspiration and Atelectasis: Aspiration is caused by the inhalation of secretions into the airways after the vocal chords. Atelectasis is the partial or complete collapse of lung tissue that has previously been expanded. Some examples of this disease are Aspiration Pneumonitis and Pulmonary Atelectasis.

Disorders of the Pulmonary Circulation: The pulmonary circulation includes the right heart chamber, a system of pulmonary arteries and arterioles, capillary bed and a system of venules and veins. Disorders of the pulmonary circulation are the result of various abnormalities in the lung, its blood vessels or the heart. The disorders include Pulmonary Edema, Pulmonary Embolism, Pulmonary Hypertension and Chronic Cor Pulmonale.

Diseases of the Pleura and the Thoracic Wall: Pleura and the thoracic wall are part of the respiratory system. Due to the close anatomic and physiologic relationship between the lung, pleura and thoracic wall, the pathologic condition of one often affects the other. Significant diseases of the pleura impair the function of the underlying lung. Some examples of this disease are Pleural Effusion, Pleural Emphysema and Pneumothorax.

Disorders of the Respiratory Control: Numerous muscles under various circumstances contribute to the mechanical respiratory function. They include the diaphragm, muscles of the rib cage, abdominals, scalenes and accessory muscles. This type of disease is caused by disorders in the neuromuscular and central nervous system that affect the normal respiration process. Diseases that can be categorized into this group include Muscular Dystrophies, Inflammatory Myopathies and Myasthenia Gravis.

Respiratory Failure: This disease is characterized by the impairment of the respiratory function. Respiratory failure occurs when the respiratory system fails to maintain its

principle function, that is, the adequate oxygenation of arterial blood and/or proper elimination of carbon dioxide (CO₂) with normal activities. This disease can be acute, chronic or both. Asthma, Pneumonia, and Adult Respiratory Distress Syndrome (ARDS) can lead to respiratory failure.

Neoplastic Diseases of the Lung: Neoplastic disease of the lung refers to malignant neoplasms that start in the lung tissues. The most common malignant neoplasm in the lung originates from the bronchial mucosa and is called bronchogenic carcinoma. Malignant neoplasms are damaging cell conditions that are very harmful. The most common neoplastic lung disease is Lung Cancer.

2.1.1 Risk Factors Causing Lung Diseases

Some people might have a higher chance of contracting lung diseases. This is due to genetics, lifestyle as well as personal health condition. Some of the major risk factors associated with lung diseases are described below.

Smoking: One of the major causes of lung diseases is smoking. Cigarette smoking accounts for 82% of obstructive lung disease deaths. Being around second-hand smoke also increases the chances of contracting lung disease. Smoking includes not only just cigarettes, but also tobacco, cigars and pipes.

Air Pollution: Particles and gases in the air can be a source of lung irritation. People who are constantly inhaling highly polluted air are at a high risk of lung diseases.

Occupational Hazards: Substances constantly inhaled at work can cause lung problems. Workers who are exposed to industrial fumes and dusts such as coal, silica, asbestos, raw cotton as well as metal fumes or chemical vapors are at a high risk of developing lung diseases.

Virus, Fungus and Bacterium: These organisms when inhaled into the lungs can cause colds, influenza, pneumonia, and other respiratory infections. Early vaccination can prevent the attack of many infectious lung diseases.

Genetics: Certain types of lung diseases are prevalent in people who have a history of lung disease in their family. The risk is increased if the person who has the disease is a first degree relative.

2.1.2 Symptoms of Lung Diseases

The common and important symptoms of lung diseases are listed below.

Chronic Cough: Any cough that last more than a month is chronic. This is an important early symptom indicating something is wrong with the breathing system.

Dyspnea: It is breathing problems such as chest-tightness, shortness of breath that continues after a brief rest following normal exercise, or comes after little or no exertion or breathing difficulties.

Chronic Phlegm Production: Phlegm, or sputum, is produced by the lungs as a defense response to infection or irritants. Phlegm or mucus production that last more than a month could indicate a lung problem.

Wheezing: Noisy breathing or wheezing is a sign that something unusual is blocking the airways of the lungs or making the airways too narrow.

Coughing Blood (Hemoptysis): The cough may consist of pure blood or it may be mixed with the sputum, causing blood-tinged sputum.

Frequent Chest Infections: Chest colds that occur more than twice a year could indicate a weak respiratory system that can cause many lung disorders.

2.1.3 Diagnosis of Lung Diseases

Proper care of patients with lung diseases necessitates identifying their specific problems and diagnosing the disorder. This can be achieved only after adequate information is obtained from various sources and by various means. Taking a patient's history, doing a physical and clinical examination, obtaining radiographic studies, assessing clinical symptoms and performing diagnostic procedures are methods used to gain this information (Ward et al., 2002 & Meholic et al., 1996).

Patients' personal and medical history: This should include both the patient's and his or her family's history. Essential information include the state of the patient's health prior to the present illness; previous diseases, surgeries, and injuries; occupational, environmental and travel history; allergies; health of family; smoking and other habits; and intake of medications. History taking is quite important especially occupational and environmental history which is crucial in evaluating patient with respiratory disorder.

Clinical Symptoms: Clinical symptoms are the patient's complaints on the clinical problems they are having as well as the associated symptoms. Detailed information about the symptoms which includes the time and date of the onset of the symptom, their severity and duration, circumstances leading to them and the factors alleviating them are analyzed.

Clinical Tests: This includes both the physical examination as well as the laboratory tests. A physical examination can yield valuable information about a person's general as well as respiratory health. Some of the physical characteristics noted are development and nutritional state, color, complexion, posture gait, chest movement, chest shape and etc. Other types of physical examinations include palpation and auscultation. Medical

laboratory tests are conducted in order to obtain an accurate functional assessment of the patient's situation. The common laboratory tests for lung diseases include Arterial Blood Gas (ABG), Complete Blood Count (CBC), Pulmonary Function Testing (PFT), Serum and Sputum test.

Radiographic Examination: Radiographic examination plays a very important role in diagnosing lung diseases and studying patients with respiratory disorders. A number of lung conditions would either be undiagnosed or misdiagnosed without radiographic study. In patients with respiratory diseases, radiographic examination enables better identification of abnormalities, correct assessment of its extent and its precise localization. It also helps in following a patient's progress and response to therapeutic measures. There are a few types of diagnostic imaging methods used for lung diseases.

X-Ray: The chest X-Ray is the most common diagnostic imaging method. The chest X-Ray procedure often involves a view from the back to the front of the chest as well as a view from the side. Like any X-Ray procedure, chest X-Rays expose the patient briefly to a minimum amount of radiation. The chest is very suitable for X-Ray because of differences in the density of its various structures. There is sufficient contrast between most of the structures to allow for their delineation on the X-Ray film. The characteristics of any abnormalities observed on a chest X-Ray film is described in terms of its location, density, size, configuration and its effect on other adjacent structures. A detailed description about chest X-Ray is given in section 2.1.3.1.

CT scan: It is an imaging technique in which radiologist and computer technologies are combined to enable the reconstruction of pictures showing transverse cross section of the body as thin slices. The concept of CT is based on contrasting various structures in the slice that have different specific gravities. This contrast may be enhanced by the intravenous injection of radiopaque material, which increases the densities of tissues proportional to their vascularity. Chest and Thoracic CT scans are used for the diagnoses of lung diseases. Further refinement of CT technology in the form of High Resolution CT

(HRCT) has improved the capabilities of CT imaging. Figure 2.1 shows chest CT scan image.

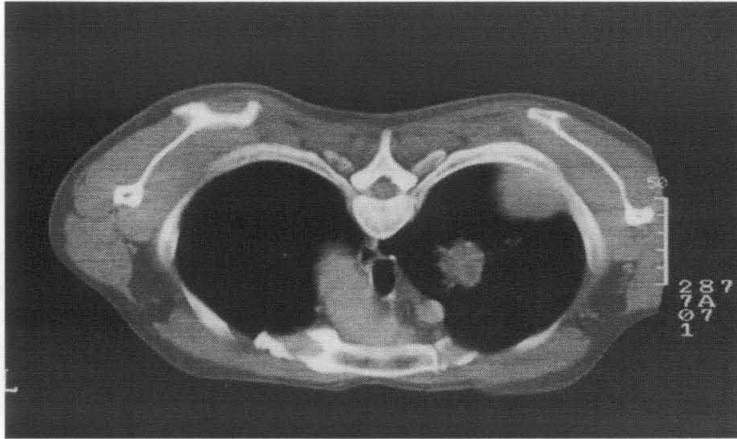


Figure 2.1: Chest CT Scan. (Obtained from Diagnostic Imaging Department, General Hospital of Ipoh, 2005)

MRI: The MRI technology is based on the characteristics of nuclei of atoms of certain elements in body tissues that can be magnetized when placed in a strong magnetic field. These magnetized nuclei are then energized by a radiofrequency pulse. The stored radio signals emitted by the protons is used by highly specialized equipment to make sectional images of the body. Contrast between the images of different types of tissues made by MRI is the result of variation in their composition and concentration of protons. Figure 2.2 shows a chest MRI image.

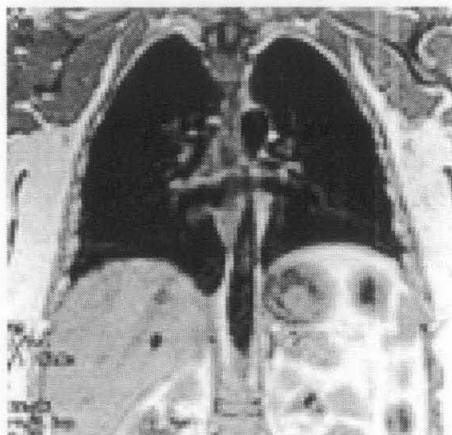


Figure 2.2: Chest MRI. (Taken from Hornack, 1996)

Positron emission tomography (PET): This is a specialized imaging technique that uses short-lived radioactive substances to produce three-dimensional colored images of those substances functioning within the body. PET scans measure metabolic activity and functioning of tissue. PET scans are often used to reveal metastatic lung diseases. PET scans can determine whether a tumor tissue is actively growing and can aid in determining the type of cells within a particular tumor. Figure 2.3 shows a PET image.

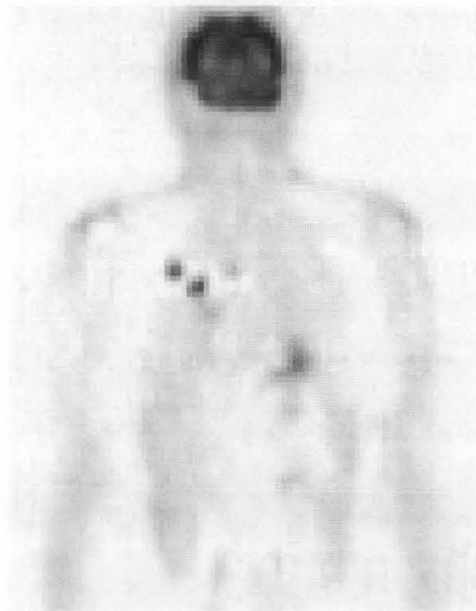


Figure 2.3: PET image. (Obtained from “PET”, 2005)

Radionuclide Lung Scanning: This type of imaging is normally used for the diagnosis of thromboembolic disorders. Perfusion lung scanning examines the distribution of blood flow to different lung regions. A ventilation lung scan is used in determining the distribution of ventilation in various lung regions. The scanning is usually done by intravenous injection of a substance tagged with radioactive material. Figure 2.4 shows a ventilation and perfusion lung scan.

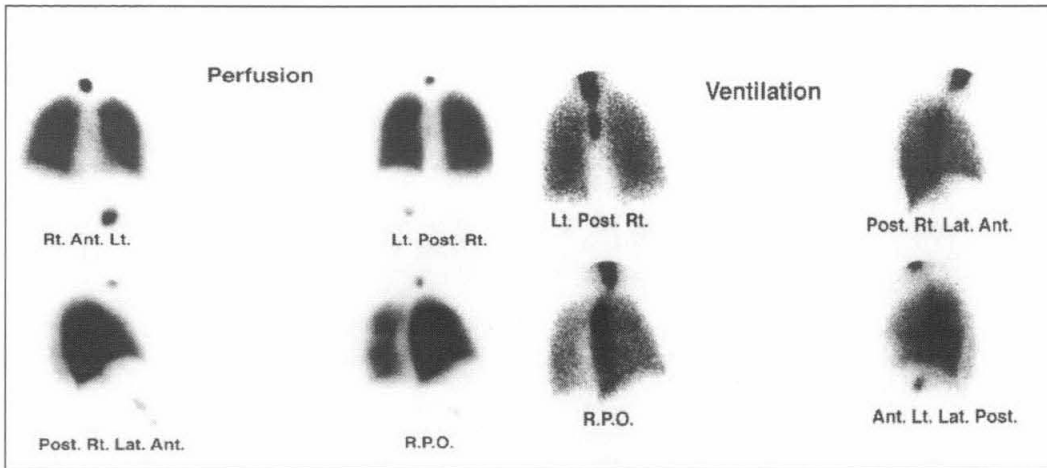


Figure 2.4: Perfusion and ventilation scan. (Obtained from “Lung”, 2005)

2.1.3.1 Chest X-Ray

The chest X-Ray is the most commonly performed radiographic examination. Approximately 45% of all radiographic exams performed in hospitals are chest X-Rays. A chest X-Ray is usually done for the evaluation of lungs, heart and chest wall. Routine chest radiography consists of post anterior (PA) and lateral views (Gurney, 2002).

Radiographic images depend on the fact that X-Rays are absorbed to a variable extent as they pass through the body. The visibility of normal structures and of diseases depends on this differential absorption. With conventional radiography there are four basic densities which are gas, fat, soft tissues and calcified structures. X-Rays that pass through air are least absorbed and therefore cause the most blackening of the radiograph, whereas calcium absorbs the most and so bones and other calcified structures appear virtually white. The soft tissues, with the exception of fat, e.g. the solid viscera, muscle, blood, fluids, bowel wall, etc., all have similar absorptive capacity and appear the same shade of gray on the conventional radiograph. Fat absorbs slightly fewer X-Rays and therefore appears a little blacker than the other soft tissues (Farzan et al., 1997).

The chest X-Ray shown in Figure 2.5 is a normal chest X-Ray taken in the post anterior view which means that the radiation passes through the patient from back to front. The patient always faces the observer, thus, the left side of the image shows the right lung. The lungs are radiolucent regions of air; therefore they show up black in the image. Within the lung fields, only bony structures and blood vessels are visible. The anatomy of the chest X-Ray is marked with numbers in Figure 2.5. The anatomies are:

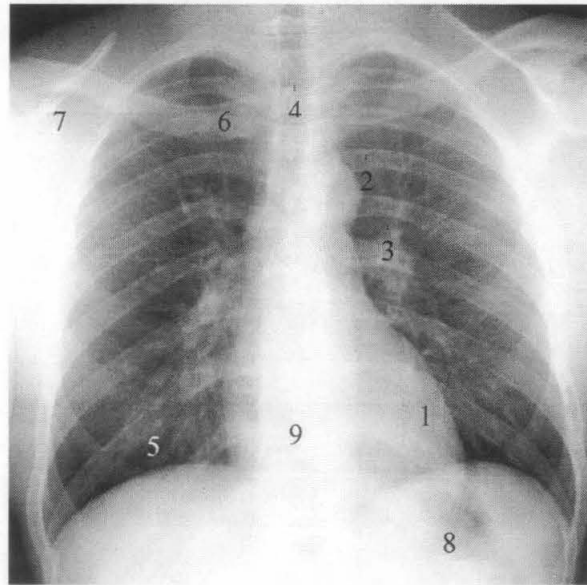


Figure 2.5: Normal chest X-Ray. (Obtained from Ginneken, 2001)

- | | |
|----------------|-------------------|
| 1. Heart | 6. Clavicle |
| 2. Aortic arch | 7. Shoulder blade |
| 3. Hilum | 8. Stomach gasses |
| 4. Trachea | 9. Spine |
| 5. Diaphragm | |

Unfortunately, although chest X-Ray is the most commonly performed examination, it often is inadequately performed. In a US Food and Drug Administration (FDA) department survey, over 40% of the chest X-Rays were judged inadequate by a panel of radiation physicists and physicians. Poor quality films markedly reduce the chance of finding abnormalities (Gurney, 2002).

Chest X-Rays are the most difficult of all radiographic examinations to perform because the chest contains tissues of such different consistency. The lungs are composed nearly entirely of air next to thick soft tissue and bone. Producing an image which adequately provides clear definition of all structures in the chest requires meticulous technique and attention to detail. The machine (film processor) used to develop the film must be working properly. This is rarely achieved. In the same FDA study, over 50% of the film processors surveyed were judged inadequate. Thus, this goes to prove that distinguishing normal lung structures from abnormal signs is a very challenging task to the doctors.

2.1.3.2 Chest X-Ray Abnormalities

There are various types of abnormalities that can be observed in a chest X-Ray film. Various types of abnormalities which can be seen in a chest X-Ray are shown in figures 2.6 to 2.13 (Ginneken, 2001). Figure 2.6 shows obvious diffuse abnormalities in the lower right lung. There are more subtle diffuse abnormalities in Figure 2.7. Figure 2.8 shows linear scarring, where there are abnormal lines that run from top-left to lower right. The denser area in the center of Figure 2.9 is an obvious infiltrate. Figure 2.10 shows a round dense object in the middle which is a calcified lymph node. There is a small calcification located behind a rib in Figure 2.11. Figure 2.12 shows an image with a cyst with a cavity. In Figure 2.13, the hilum of the left lung is abnormal where the lymph nodes are swollen.

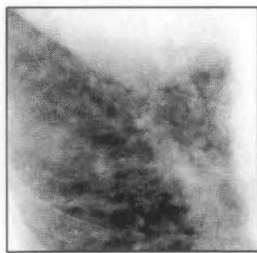


Figure 2.6: Lung diffuse abnormality

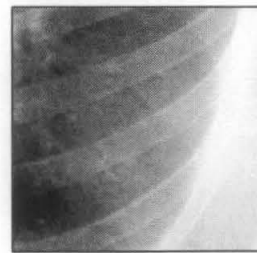


Figure 2.7: Subtle diffuse abnormality



Figure 2.8: Linear scarring



Figure 2.9: Infiltrate

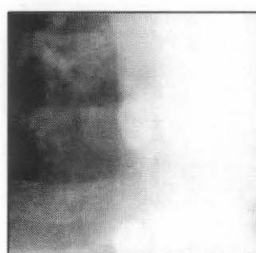


Figure 2.10: Calcified lymph node



Figure 2.11: Calcification



Figure 2.12: Cyst with a cavity

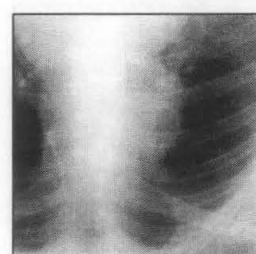


Figure 2.13: Abnormal hilum

2.1.4 Lung Cancer

Lung cancer is the most common killer among all malignancies with 1.18 million deaths per year (Parkin et al., 2005). In Malaysia, lung cancer comprises of 13.8% of all male cancer and 3.8% of all female cancer (Ministry of Health Malaysia, 2003).

Lung cancer develops when cells inside the lungs multiply at an uncontrollable rate. These abnormal tissue masses are called tumors. Some tumors grow and enlarge only at

the site where they began and these are referred to as benign tumors. Other tumors not only enlarge locally but also have the potential to invade and destroy the normal tissue around them and to spread to distant parts of the body. This kind of tumors are designated as malignant or cancerous. Benign tumors may grow causing discomfort and bleeding, but they do not spread to other parts of the body and generally are not life-threatening. Malignant cells detach themselves from the original tumor, are carried to other parts of the body through the blood or lymphatic vessels and establish themselves in the new site as an independent (secondary) cancer. A tumor that has spread in this manner is said to have metastasized and the secondary tumors are called as metastases. Lung cancer can spread to any organ in the body (Scott et al., 2001).

Lung cancer survival rate remain dismally low at 14% at 5 years. It has been reported that the 5-year survival rate for all lung cancer stages combined is 14%. This rate increases to 49% if the tumor is localized and decreases to 2% if it has metastasized. However, early detection, when the primary tumor is less than 3 cm in diameter, may lead to 5-year survival rates of 70-80% (Daffner et al., 1999). Therefore, prompt diagnosis and management of early lung cancer manifesting as pulmonary nodules may be the only chance for cure. As the global problem of lung cancer is a growing one, this research work gives much importance to the diagnosis of lung cancer.

2.1.4.1 Lung Cancer Indicators: Lung Nodules

Lung Nodules are one of the major indicators of lung cancer. The diagnosis of lung cancer is performed similar to the diagnosis of other lung diseases which was described in section 2.1.3. However, in the case of lung cancer, an accurate diagnosis for a majority of patients is normally achieved through chest radiography (Muellor et al., 2001)..

The most important manifestation of lung cancer in chest radiography is in the form of a lung nodule. A lung nodule is defined as a single or multiple lung lesions that is less than

3 cm in diameter and surrounded by normal lung tissue. Approximately 40 to 50% of these nodules are malignant.

The most important question to be considered in the diagnosis of lung nodules is whether the nodules are benign or malignant. Chest radiographs can provide information regarding the size, shape, cavitation, growth rate, and calcification pattern of the nodule. All of these radiological features can help determine whether the nodule is benign or malignant.

Specific radiological characteristics may help establish the diagnosis with reasonable certainty. The size and shape of the nodule plays a very crucial role in the differential diagnosis. A lesion more than 3 cm in diameter is very likely to be cancerous. The shape of a nodule is generally very helpful in determining its malignancy. A very irregular edge may indicate a malignant condition. A well-defined, smooth, non-lobulated edge may indicate a benign lesion or metastasis.

2.1.4.2 Diagnosis of Lung Cancer: Diagnostic Imaging Methods

The most common diagnostic imaging method used for lung cancer detection and diagnosis are chest X-Ray, CT scan and MRI. The chest X-Ray is the most common routine diagnostic step taken when any symptoms of lung cancer are present. The chest X-Ray procedure often involves a view from the back to the front of the chest as well as a view from the side. Like any X-Ray procedure, chest X-Rays expose the patient briefly to a minimum amount of radiation. The chest X-Ray procedure is inexpensive and widely available.

Generally, a lung nodule must reach 1 cm in diameter before it can be identified on a chest X-Ray by naked eyes. Thus, chest X-Ray is not an ideal method for diagnosing lung cancer at its early and curable stage. However, there is a huge potential of detecting and diagnosis lung nodules on chest X-Rays by using image processing techniques on

digitized X-Ray images (Sutton, 1997). This is a major part of this research and is explained in detail in chapter 5.

CT scans may be performed on the chest to examine for both primary and metastatic nodule. A CT scan of the chest is suggested when the X-Ray images do not yield sufficient information. One advantage of CT scan is that it is more sensitive than standard chest X-Rays in the detection of lung nodules.

Magnetic resonance imaging (MRI) detailed and can detect tiny changes of structures within the tumors. People with heart pacemakers, metal implants, artificial heart valves, and other surgically implanted structures cannot be scanned with an MRI because of the risk that the magnet may move the metal parts of these structures.

2.2 Computer Aided Image Analysis for Chest Radiography

2.2.1 General Image Processing Functions

2.2.1.1 Enhancement Filters

Median filter is used for noise reduction and image smoothing. The filter replaces the intensity of the center pixel in an $N \times N$ neighborhood with the median value of the neighboring intensities. This filter is effective in removing isolated pixel noise while retaining the sharp edges of an image (Umbough, 1998).

A Gaussian filter mask has the form of a bell-shaped curve. The Gaussian filter produces, for each pixel in the image, a weighted average such that the central pixels contribute more significantly to the result than pixels at the mask edges (Aguado, 2002). For a filter mask of size $k \times k$, Gaussian filter coefficients are obtained in the form of

$$h_G(i, j) = \exp\left[\frac{-1}{2(d/\sigma)^2} \right] \quad (4.1)$$

where $d = \sqrt{i^2 + j^2}$, $-(k-1)/2 \leq i, j \leq (k-1)/2$ and σ is the standard deviation.

Weiner filter is normally used for noise suppression in noisy images. Wiener filter, also called a minimum mean-square estimator, attempts to model the error in the restored image through the use of statistical methods. After the error is modeled, the average error is mathematically minimized (Gonzalez et al., 2002). Wiener filter is modeled by

$$\hat{F}(u, v) = \left[\frac{1}{H(u, v) \left(|H(u, v)|^2 + S_n(u, v) / S_f(u, v) \right)} \right] G(u, v) \quad (4.2)$$

where (u, v) are spatial frequencies, $H(u, v)$ is the degradation function, $S_n(u, v)$ is the power spectrum of the noise and $S_f(u, v)$ is the power spectrum of the original image.

Unsharp masking filter technique entails low pass filtering, or unsharpening, of an input image and subtraction of the filtered image from the original (Seul et al., 2000). If $f(x, y)$ represents the original image and $f_{LP}(x, y)$ the low pass filtered image, unsharp masking yields an output image $g(x, y)$, as follows:

$$g(x, y) = af(x, y) - bf_{LP}(x, y) \quad (4.3)$$

where a and b are positive constants with $(a, b) \geq 1$

2.2.1.2 Edge Detection Filters

Edges characterize object boundaries and are useful for segmentation purposes, registration and object recognition purposes. Edge points are pixel locations of abrupt gray-level changes (Seul et al., 2000).

The Laplacian edge detector yields the rate of change of intensity variations. Laplacian edge detector for an image function $f(x,y)$ is characterized by the equation 4.4.

$$\Delta f(x, y) \equiv \nabla[\nabla f(x, y)] = \partial_{xx}^2 f(x, y) + \partial_{yy}^2 f(x, y) \quad (4.4)$$

The sign of the Laplacian determines the location of a given pixel relative to an edge. The sign of laplacian is positive for pixels located on the darker side of a nearby edge, while it is negative for pixels located on the brighter side (Jain, 1989).

2.2.1.3 Geometrical Operations

Geometric transforms include those processes that change dimensional properties of images. These include resizing, cropping and rotating. Image rotation function enables the input images to be rotated either 90 degrees to the left, 90 degrees to the right or 180 degrees. The function also allows the user to rotate at a user specified angle and direction, whether clockwise or anti clockwise. This will help the user to analyze the x-ray from different angles. The transformation $x' \equiv F_x(x, y)$ and $y' \equiv F_y(x, y)$ can be specified in the form of an orthogonal transformation as below.

$$F(x, y) = \begin{bmatrix} F_x \\ F_y \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \quad (4.5)$$

2.2.1.4 Contrast Enhancement

Contrast stretching is an image enhancement technique that attempts to improve the contrast in an image by ‘stretching’ the range of intensity values it contains to span the full range of pixel values that the image type concerned allows (Seul et al., 2000). The formula concerned with contrast stretching is shown by equation 4.6

$$b_{out} = \begin{cases} 0 & , b_{in} \leq b_{low} \\ \frac{b_{in} - b_{low}}{b_{hi} - b_{low}} & , b_{low} \leq b_{in} \leq b_{hi} \\ 255 & , b_{in} \geq b_{hi} \end{cases} \quad (4.6)$$

where b_{low} is the lower intensity limit, b_{hi} is the upper intensity limit, b_{in} is the input intensity of a pixel and b_{out} is the output intensity of a pixel.

Histogram equalization technique is to obtain a uniform histogram for the output image (Jain, 1989). An image pixel value $u \geq 0$ is considered to be a random variable with a continuous probability density function $p_u(a)$ and cumulative probability density distribution $F_u(a) \triangleq P[u \leq a]$. Then the random variable

$$v \triangleq F_u(u) \triangleq \int_0^u p_u(a) da \quad (4.7)$$

will be uniformly distributed over (0, 1). To implement this transformation on digital images, suppose that the input u has L gray levels $x_i, i=0,1,\dots,L-1$ with probabilities $p_u(x_i)$. These probabilities can be determined from the histogram of the image that gives $h(x_i)$, the number of pixels with gray level value x_i .

$$p_u(x_i) = \frac{h(x_i)}{\sum_{i=0}^{L-1} h(x_i)}, \quad i = 0, 1, \dots, L-1 \quad (4.8)$$

The output \dot{v} , also assumed to have L levels, is given as follows

$$v = \underline{\underline{\Delta}} \sum_{x_i=0}^u p_u(x_i)$$

$$\dot{v} = \underline{\underline{\Delta}} \text{Int} \left[\frac{(v - v_{\min})}{1 - v_{\min}} (L - 1) + 0.5 \right] \quad (4.9)$$

where v_{\min} is the smallest positive value of v . Now \dot{v} will be uniformly distributed only approximately because v is not a uniformly distributed variable.

Contrast Limited Adaptive Histogram Equalization (CLAHE) algorithm partitions the images into contextual regions and applies the histogram equalization to each one. This evens out the distribution of used grey values and thus makes hidden features of the image more visible. The full grey spectrum is used to express the image. Sharp field edges can be maintained by selective enhancement within the field boundaries. Selective enhancement is accomplished by first detecting the field edge in an input image and then only processing those regions of the image that lie inside the field edge (Pratt, 1991). The formula for CLAHE implementation is shown in equation 4.10.

$$P_{out} = [P_{\max} - P_{\min}]P(f) + P_{\min} \quad (4.10)$$

where P_{out} is output pixel value, P_{\min} is minimum pixel value, P_{\max} is maximum pixel value and $P(f)$ is cumulative probability distribution.

2.2.1.5 Arithmetic Operations

Arithmetic operations are used extensively in image processing. Addition takes as input two identically sized images and produces as output a third image of the same size as the first two, in which each pixel value is the sum of the values of the corresponding pixel from each of the two input images. This function is denoted by the equation 4.11

$$P_{out}(i, j) = P_1(i, j) + P_2(i, j) \quad (4.12)$$

where P_{out} is the output image and P_1 and P_2 are the input images.

Subtraction takes two images as input and produces as output a third image whose pixel values are simply those of the first image minus the corresponding pixel values from the second image. Image subtraction is normally used to subtract background variations in illumination from a scene so that the foreground objects in it may be more easily analyzed (Baxes, 1994). This function is denoted by the equation 4.12

$$P_{out}(i, j) = P_1(i, j) - P_2(i, j) \quad (4.13)$$

where P_{out} is the output image and P_1 and P_2 are the input images.

2.2.1.6 Segmentation

In the analysis of the objects in images it is essential to distinguish between the objects of interest and the background. The segmentation techniques are used to find the objects of interest.

Seed-based region growing is a procedure that groups pixels or sub regions into larger regions based on some pre-defined criteria. The basic approach is to start with a set of seed pixels and from these grow regions by adding to each seed those neighboring pixels that have properties similar to the seed. These seed pixels are chosen by user's opinion on what should be the regions to extract on the image.

A seed pixel's gray level is compared with the statistics (e.g. mean, variance etc.) of its neighborhood pixels. If some growing conditions are fulfilled from the comparison, the seed pixel will grow towards its neighbors. The process continues recursively until the region cannot grow anymore or all the pixels have been considered. When the growth of one region stops, another seed pixel which does not yet belong to any region is chosen and the step is repeated. This whole process is continued until all pixels belong to some region are considered.

The process is first started with a number of seeds which have been grouped into n sets: A_1, A_2, \dots, A_n . Sometimes individual sets could consist of single point. At each step of the algorithm one pixel is added to some of the sets $A_i, i=1, \dots, n$. The state of sets A_i , after m steps is considered. Let T be the set of unallocated pixels which border at least one of the regions, which means that T is the set of all pixels which are on the borders of the formed regions.

$$T = \left\{ x \notin \bigcup_{i=1}^n A_i \mid N(x) \cap \bigcup_{i=1}^n A_i \neq \emptyset \right\} \quad (4.14)$$

$N(x)$ is the set of immediate neighbors of the pixel x . In this case, the 8-connectivity approach is used. Each step in the algorithm takes one pixel from the T set and adds it to one of the regions with which neighbors $N(x)$ of the pixel intersect and labels it with the label of that region. Once the growing process is terminated two regions are obtained. One is the seed-grown region and the other is the surrounding region.

The region growing method can be applied to images with variations in their grayscale range. However, suitable preprocessing might be necessary before the application of region growing in order to obtain a correct region. The positions of the initial seeds are very crucial. It is possible that user makes mistake selecting seed pixels. A proper homogeneity criteria ensure that correct results is obtained. Incorrect selection of seed points and homogeneity criteria can lead to less or more regions on the images. In some cases, post-processing can be done to correct result of the segmentation (Nghah et al., 2000 & Ilic et al., 2000).

2.2.1.7 Feature Extraction

Feature extraction is used to obtain specific features of objects in an image. These characteristics are used to describe the object.

The area of an object in an image, is defined by

$$A(S) = \int_x \int_y I(x, y) dy dx \quad (4.16)$$

where $I(x, y)$ is 1 if the pixel is within a shape, $(x, y) \in S$, and 0 otherwise.

The perimeter of an object, P is the sum of all the edge pixels of an object (Aguado, 2002). If $x(t)$ and $y(t)$ denote the parametric co-ordinates of a curve enclosing a region S , then the perimeter of the region is defined as below.

$$P(S) = \int_t \sqrt{x^2(t) + y^2(t)} dt \quad (4.17)$$

2.2.1.8 Mathematical Morphology

Morphology relates to the structure or form of objects. Dilation allows object to expand, thus potentially filling in small holes and connecting disjoint objects. Erosion shrinks objects by eroding their boundaries (Gonzalez et al., 2002).

If A is the original image and B is the structuring element as sets in Z^2 where Z is the elemental array, then the dilation of A by B , denoted by $A \oplus B$ is defined in the equation 4.17.

$$A \oplus B = \{z | [(B)_z \cap A] \subseteq A\} \quad (4.18)$$

The erosion of A by B , denoted by $A \otimes B$ is defined as

$$A \otimes B = \{z | (B)_z \subseteq A\} \quad (4.19)$$

Opening operation is the application of erosion followed by dilation. Closing operation is the application of dilation followed by erosion (Gonzalez et al., 2002).

The opening of A by B , denoted by $A \circ B$ is defined in the equation 4.19.

$$A \circ B = (A \otimes B) \oplus B \quad (4.20)$$

The closing of A by B , denoted by $A \bullet B$, is defined in the equation 4.20.

$$A \bullet B = (A \oplus B) \otimes B \quad (4.21)$$

2.2.2 Previous Work on Detection and Diagnosis of Lung Nodules

The automatic detection of lung tumors such as masses and nodules is one of the most studied problems in computer analysis of chest radiographs. Early detection of lung abnormalities in chest X-Rays is a rather challenging task. Many methods have been attempted to overcome this problem.

Subtraction techniques are used to remove normal anatomical structures in chest radiographs, so that the abnormalities are more apparent for both computer detection and radiologists' observation. Kano et al. (1994) used a temporal subtraction method. A current image is registered with a previous radiograph of the same patient. The displacement of small regions of interest is computed based on cross correlation and a smooth deformation field using elastic matching. Both the images are subtracted from one another and the lung nodules stand out clearly.

Zhao et al. (2001) also used a temporal subtraction technique that performs elastic registration on rib border segments and reported that the contrast around lung nodules substantially increased in the subtracted images. Temporal subtraction techniques have the potential to be used on a routine basis since periodical is one of the main functions of chest radiographs. In situations where a previous radiograph is not available, subtraction can be achieved by made by mirroring the left/right lung field, performing elastic registration on the right/left lung field and subtracting. This technique is called *contralateral subtraction*.

Giger et al. (1988) introduced a method where a subtraction image is produced by filtering with the image with a spherical kernel and subtracting it from a median filtered image. The nodules are obtained through thresholding. Sklansky et al. (1973) used a generalized Hough transform method. This method utilizes edge detection. The centers of circles of different radii which represent the nodules are detected using gradient magnitude and non-maxima suppression.

Yoshida et al. (1995) used a wavelet based method for nodule detection. A nodule enhanced image is obtained by computing the least asymmetric Daubechies wavelet transform and amplifying responses from intermediate levels before back-transformation. A subtraction image is produced and the nodules enhanced image is combined with the subtracted image.

The above described systems use mainly subtraction, filtering and edge detection methods for nodule detection. In the methods utilizing subtraction, the quality and accuracy is affected by variations in patient positioning and inaccurate image registration methods.

Classification of a nodule plays a major role in diagnosing lung diseases. The nodule can be classified as either benign or malignant. Benign nodules are generally harmless while malignant nodules are serious. Thus, correct classification is very crucial in the diagnostic process. Different types of method have been used in the classification process.

Nakamura et al. (2000) used feed-forward neural network to classify nodules as benign or malignant. The neural network performs the classification on the basis of subjective or objective features. The performance of this method was compared with that of the radiologists by means of receiver operating characteristic (ROC) analysis and was found to yield better results.

Gurney et al. (1995) used a Bayesian classifier and a feed-forward neural network to discriminate benign and malignant nodules based on a set of specific nodule characteristics. They reported good results and the Bayesian classifier out-performed the neural network.

Xu et al. (1997) used feature extraction to extract specific shape values of the nodules such as the diameter, circularity and irregularity by means of region growing. The slopes of these features, profile measures, size of regions as a function of thresholds, contrast

and gradient measures were also extracted. Artificial neural network (ANN) was used as the preferred classifier.

Li et al. (2001) used a template matching algorithm with nodule and non-nodule candidates. Meanwhile, Floyd et al. (1996) and Vititoe et al. (1997) used fractal dimensions estimated from power spectrum. A threshold was set based on these values for classifying the nodules.

Summary

Image processing techniques can be used to significantly enhance and improve the chest X-Ray images. The detection of lung nodules mainly utilizes subtraction, filtering and edge detection methods. In subtraction, the quality and accuracy of the results are often affected by variations in patient positioning and inaccurate image registration methods. Thus, a new method of detection that uses a modified grayscale top hat transform has been introduced in this work. This method might have the capability to give results equivalent or even better than the above described method. Various types of method have been used in the classification of lung nodules. Most of the work described uses Artificial Neural Network for the classification based on a set of nodule characteristics. Apart from that, Bayesian classifier, template matching and fractal dimensions have also been tried. In this work, a new method of classification using Fuzzy Inference System has been proposed. This method is a new approach that is able to perform the classification in an accurate and consistent manner.

CHAPTER 3

EXPERT SYSTEMS

Introduction

Expert systems are computer programs that are able to think and reason like a human in order to make decisions. This research work is about the development of an expert system to diagnose lung diseases. The mechanism and functions of the expert system are introduced. A brief overview on the concept of a rule-based expert system is presented. This is followed by the objectives as well as the features of the system. The structure of the system as well as its various components is described in detail. The merits of using expert systems are also indicated. Currently, there are many types of expert systems which are being used for various applications. In accordance to this, some attention is given to medical expert systems and their use in clinical environment.

3.1 General Concept of Expert Systems

An expert system is an intelligent computer program that uses knowledge and inference procedures to solve problems that are complex enough to require significant human expertise for their solution. It simulates the judgment and behavior of a human or an organization that has expert knowledge and experience in a particular field.

Expert systems solve problems using specialized knowledge at the level of a human expert. Knowledge is the theoretical or practical understanding of a subject or domain. A person is considered as a domain expert if he or she has a deep knowledge and strong practical experience in a particular domain. Experts are those who will be able to solve complex problems which could not be easily solved by most of the people.

The knowledge of an expert system is obtained through various sources and coded in a form that is suitable for the system for its inference process. The knowledge is mostly obtained from human experts of a particular field or other reliable sources such as books, journals, publications, internet websites and databases.

An expert's knowledge is normally focused on one particular area. Thus, an expert system is also designed to be an expert in a specific area or solve a specific problem. The expert system makes inferences in the same way a human expert would infer the solution to a problem. The system takes the knowledge programmed into it as well as the information provided to it by the user and makes its inferences. The system applies heuristics to guide the reasoning process (Giarratano et al., 1998).

3.2 Features of Expert Systems

Expert system has special features that make them the perfect solution for certain applications (Ignizio, 1991 & Negnevitsky, 2005). The features are described below.

Natural knowledge representation

An expert usually explains the problem-solving procedure with such expressions as this: "In such-and-such situation, I do so-and-so". These expressions can be *If-Then* structure. Each rule is an independent piece of knowledge. The syntax of the rules enables them to be easily understood without much documentation.

Separation of knowledge from its processing

The structure of a rule-based expert system provides an effective separation of the knowledge base from the inference engine. This makes it possible to develop different applications using the same expert system shell. The knowledge can be continuously added, deleted or modified without affecting the main program. Thus, improving the expert system would be a rather hassle free task.

Explanation capability

The system should be able to provide reasons of how it came up with a particular solution. Thus, the system justifies its conclusion the same way as a human expert can. This is very crucial as most often, the user might want to know why the system made a specific decision. This occurs when the user is not satisfied with the system's conclusion.

Specific problem domain

The expertise of the system is focused on a narrow area. This is because human experts are also most often knowledgeable in a specific field. The knowledge in the system is specific to a particular problem.

Dealing with incomplete and uncertain knowledge

Most rule-based expert systems have an additional functionality where it tries to infer and reason with incomplete and uncertain knowledge. It provides its conclusions with a level of certainty. This is important when dealing with fuzzy rules as certain applications especially those in the medical field that deal with inaccurate data.

Problem solving quality

The system has a mechanism to add, change or delete knowledge in the knowledge base. This is an important feature of rule-based system. Rules can be added or modified to improve the performance of the system.

A comparison of the features of expert system with conventional systems and human experts is presented in Table 3.1.

Table 3.1: Comparison of the expert system features with human experts and conventional programs. (Adapted from Negnevitsky, 2005)

Human Experts	Expert Systems	Conventional Programs
Knowledge exists in a compiled form.	Knowledge separate from its processing.	No separation of knowledge from the control structure for processing.
Knowledge in the form of heuristics or rules of thumb is used to solve problems in a narrow domain.	Knowledge is expressed in the form of rules and symbolic reasoning is used to solve problems in a narrow domain.	Data processing is done by the use of algorithms.
Capable of providing explanation for reaching a particular conclusion.	Explanation facility enables the system to explain how a particular conclusion was reached and why specific data was needed.	Do not explain the results or why specific inputs were needed.
Capable of making mistakes when the information is incomplete or uncertain.	Might have the capability to infer from inexact or incomplete information to deliver a solution.	Provide no solution or the wrong results when the information is incomplete or inexact.
Quality of problem solving is enhanced through years of learning and experience. The process is slow, inefficient and expensive.	Quality of problem solving is enhanced by adding new rules or modifying existing ones in the knowledge base.	Quality of problem solving is enhanced through changing the program code, which affects the knowledge and its processing, making changes tedious.

3.3 Structure of Expert Systems

In the early seventies, Newell and Simon (1976) from Carnegie-Mellon University proposed a production system model, which laid the foundation of the modern rule-based expert systems. The model is based on the idea that humans solve problems by applying their knowledge (expressed as production rules) to a given problem represented by problem-specific information. Rule-based systems consist of a set of *If-Then* statements. The knowledge of an expert is encoded into a set of rules. Rule-based systems are feasible for problems for which the knowledge in the problem area can be represented in the form of *If-Then* rules. Figure 3.1 shows the structure of a rule-based expert system.

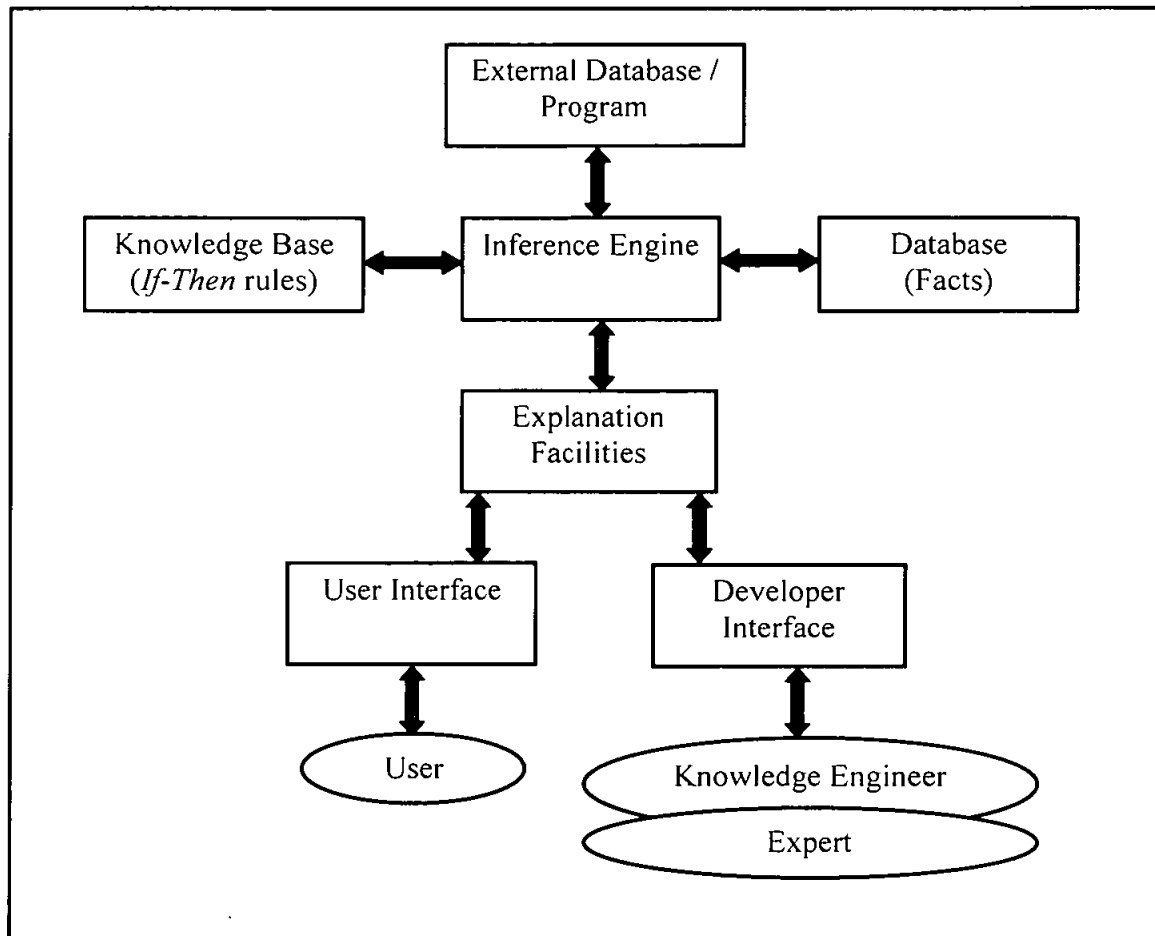


Figure 3.1: Structure of an expert system. (Adapted and modified from Negnevitsky, 2005)

The components of an expert system may vary from one system to another. The components described below are commonly found in most of the expert systems (Negnevitsky, 2005 & Jackson, 1999).

User

The user is the person who uses the system. In this work, the user may be a general physician.

User Interface

The user interface is the means of communication between a user seeking a solution to the problem and an expert system. The system asks for information through questions displayed on the screen and the user answers by typing or clicking on the provided options.

Knowledge Base

Knowledge base contains the knowledge acquired from the experts and reliable sources such as books, journals, websites, etc. The knowledge might be of object descriptions, problem-solving behaviors, experience, constraints, judgment, heuristics, facts and uncertainties. In a rule-based expert system, the knowledge is represented as a set of *If-Then* rules. The success of an expert system relies on the completeness and accuracy of its knowledge base.

Inference Engine

The inference engine uses the knowledge stored in the form of rules to produce solutions. The engine examines the rules and combines them with new facts in the knowledge base to induce inferences. The rules are often heuristics and the facts might be only true to a certain degree, thus the decisions carry some level of probability. Most often, the system may not yield a single conclusion, but a number of possibilities with different degrees of probabilities. The system may use either backward chaining or forward chaining for its inference. This is dependent on its specific domain.

Explanation Facility

The explanation facilities enable the user to ask the expert system how a particular conclusion is reached and why a specific fact is needed. Such explanation facilities provide the user with a means of understanding the system behavior. This is similar to a consultation with a human expert who will often provide some explanation for his or her conclusions. Many people would not always accept the answers of an expert without

some form of justification. An expert system must be able to explain its reasoning and justify its advice, analysis or conclusion.

Knowledge Engineer

Knowledge engineer is the person who is capable of designing, building and testing an expert system. The knowledge engineer is also responsible for testing, revising and integrating the expert system into the workplace.

Expert

The expert is the expert sources of information relevant to the expert system. The experts can be human experts in specific fields, books, journals, websites, etc.

External Database / Program

The external database is a database system from which the expert system retrieves facts and information externally. External programs are any programs which are linked to the expert system.

3.4 Inference Mechanism

Inference is the process of drawing a conclusion by means of a set of rules, for a specific set of facts about a given situation. Thus inference is the knowledge processing element of an expert system. The inference engine contains rules and facts that pertain to the general control and the search strategy employed by the system in the development of a solution. The inference engine compares each rule stored in the knowledge base with facts contained in its database. When the *If* (condition) part of the rule matches a fact, the rule is fired and its *Then* (action) part is executed. The matching of the *If* parts to the facts produces inference chains. The inference chain indicates how an expert system applies the rules to reach a conclusion. There are two popular reasoning processes used in the inferencing processes which are forward chaining and backward chaining (Ignizio, 1991 & Awad, 1996).

Forward Chaining

Forward chaining is a data-driven reasoning process. The reasoning starts from the known data and proceeds forward with that data. Every time, only the topmost rule is executed. When fired, the rule adds a new fact in the database. Any rule can be executed only once. This cycle stops when no further rules can be fired. Forward chaining is a technique for gathering information and then inferring from it whatever can be inferred. However, in forward chaining, many rules that have nothing to do with the established goal may be executed. Forward chaining should be the preferred technique if the system needs to gather some information and then infers from it. This method is suitable to infer many facts.

Backward Chaining

Backward chaining is a goal-driven reasoning process. In backward chaining, an expert system starts with a goal (a hypothetical solution) and the inference engine attempts to find relevant supporting information. First, the knowledge base is searched to find rules that might have the desired solution. Such rules must have the matching goal in their *Then* (action) parts. If such a rule is found and its *If* (condition) part matches the data in the database, then the rule is fired and the goal is proven. However, this is rare case. Most often, the inference engine puts aside the rule it is working with (the rule is said to stack) and sets up a new goal, a sub-goal, to prove the *If* part of this rule. Then the knowledge base is searched again for rules that can prove the sub-goal. The inference engine repeats the process of stacking the rules until no rules are found in the knowledge base to prove the current sub-goal. Backward chaining is preferred when the system begins with a hypothetical solution and then attempts to find facts to prove it.

3.4.1 Fuzzy Inference Mechanism

Fuzzy inference is the process of formulating the mapping from a given input to an output using fuzzy logic. The mapping then provides a basis from which decisions can be made, or patterns discerned (*Matlab Fuzzy Logic Toolbox Documentation*, 2001). The

main concept of fuzzy sets is that many problems in the real world are imprecise rather than exact. It is believed that the effectiveness of the human brain is not only from precise cognition, but also from fuzzy concept, fuzzy judgment, and fuzzy reasoning. An advantage of fuzzy classification techniques lies in the fact that they provide a *soft decision*, a value that describes the degree to which a pattern fits with a class, rather than only a *hard decision*, a pattern matches a class or not.

Fuzzy inference systems have been successfully applied in fields such as automatic control, data classification, decision analysis, expert systems, and computer vision. Because of its multidisciplinary nature, fuzzy inference systems are associated with a number of names, such as fuzzy-rule-based systems, fuzzy expert systems, fuzzy modeling, fuzzy associative memory, fuzzy logic controllers, and simply fuzzy systems.

The most popular models of fuzzy systems are the Mamdani model and the Takagi-Sugeno-Kang (TSK) model. Both of these models use fuzzy rules in the form of *If-Then*. The *if* part of the rule is called *antecedent* and the *then* part is called *consequence*. The main difference between them is the consequence part of fuzzy rules. The Mamdani model describes the consequent part using the linguistic variables, while the Takagi-Sugeno-Kang model uses the linear combination of the input variables. Both models use linguistic variables to describe the antecedent part of fuzzy rules (Meesad, 2004).

3.5 Advantages of Expert Systems

Some of the advantages of the expert systems are listed below (Nghah, 1996 & Awad, 2003).

- Enable expert services to be readily available when needed
- Retain scarce knowledge
- Substitute for an unavailable human expert
- Increase timeliness in decision making
- Serve as a training tool for inexperienced people

- Reduce the high cost associated with expert services
- Improved consistency in decisions
- Formalization of knowledge
- Easy knowledge transfer
- Utilize incomplete and uncertain information
- Improve work performance and productivity

3.6 Medical Expert Systems

Known as a branch of artificial intelligence, expert systems offer a way to capture and encode knowledge from experts. These computer programs emulate the way people think. Born in the labs of Stanford University, United States of America over forty years ago to help diagnose diseases, expert systems have since gained huge popularity in the medical field. In this section, some expert systems developed to aid in the diagnoses of diseases are highlighted.

The DENDRAL Project (Buchanan et al., 1993) was one of the earliest expert systems. It was started at Stanford University in 1965. DENDRAL began as an effort to explore the mechanization of scientific reasoning and the formalization of scientific knowledge by working within a specific domain of science. The program's task was chemical analysis. Starting from spectrographic data obtained from a substance, DENDRAL would hypothesize the substance's molecular structure. DENDRAL's performance rivaled that of human chemist expert at this task, and the program was used in the industry and in universities.

MYCIN project (Buchanan et al., 1984) was started in 1972 as collaboration between the medical and AI communities at Stanford University. It is probably the most widely known of all expert systems developed so far. MYCIN is an interactive program that diagnoses certain infectious diseases, prescribes antimicrobial therapy, and can explain its reasoning in detail. In a controlled test, its performance equaled that of specialists.

Although MYCIN was never used routinely by physicians, it has substantially influenced other AI research.

The medical expert system PNEUMON-IA (Verdaguer et al., 1992) was developed to provide expert advice to physicians about the likelihood of presence of certain diseases. This system uses results of blood tests to detect the presence of diseases.

Iliad (Lincoln et al., 1991) is a computerized, expert system for internal medical diagnosis. This expert system provides expert diagnostic consultations in the field of internal medicine for over 200 diseases. The system is designed to teach diagnostic skills by means of simulated patient case presentations. It also serves as a medical reference for users.

GermAlert (Kahn et al., 1993) was developed by the Washington University School of Medicine Department. This system assists the Infection Control Departments of Barnes and Jewish Hospitals (teaching hospitals affiliated with the university) with their infection control activities. GermAlert applies local hospital culture-based criteria for detecting “significant” infections, which require immediate treatment. Microbiology culture data from the hospital’s laboratory system are monitored by GermAlert. Using a rule-base consisting of criteria developed by local infectious diseases experts, GermAlert scans the culture data and generates an “alert” to the infection control staff when a culture representing a “significant” infection is detected.

DermaDex (“Medical Expert System”, n.d.) is an expert system designed to assist in the diagnosis of skin disorders. This system was developed for the use of primary care physician. A step-by-step diagnostic tool allows a physician to select from a variety of conditions. Results are rank ordered by relative priority. This system has demonstrated significant improvements in decreasing referrals while improving therapy provided at the level of primary care. Dermadex provides rapid access to a comprehensive dermatology

database of over 220 skin conditions and more than 725 prototypic illustrations. This is commercial software and is available from the Australian Medical Information Services.

GIDEON: Global Infectious Disease and Epidemiology Network (Felitti, 2005) is an expert system for diagnosis and reference in the field tropical and infectious diseases, epidemiology, microbiology and antimicrobial chemotherapy. It was designed to diagnose all the worlds' infectious diseases, based on symptoms, signs, laboratory testing and dermatological profile. Gideon's infectious disease network pays special attention to the country of origin. The database incorporates 327 diseases, 205 countries, 806 bacterial taxa and 185 antibacterial agents. GIDEON has been developed in the past 10 years by specialists in infectious diseases; biostatistics and computer sciences at University based Medical schools in the United States and Israel. Gideon is commercially available software distributed by C.Y. Informatics LTD.

Dxplain (Barnett et al., 1987) is a diagnostic decision support system in general medicine which was first developed in 1985. It is used to assist in the process of diagnosis, taking a set of clinical findings including signs, symptoms, and laboratory data and then produces a ranked list of diagnoses. It provides justification for each of differential diagnosis, and suggests further investigations. The system contains a data base of crude probabilities for over 4,500 clinical manifestations that are associated with over 2,000 different diseases. DXplain is in routine use at a number of hospitals and medical schools mostly for clinical education and also for clinical consultation.

Quick Medical Reference (QMR) (Sumner, 1993) is an in-depth information resource that helps physicians to diagnose adult diseases. It provides electronic access to more than 750 diseases representing the vast majority of the disorders seen by medical interns in daily practice as well a compendium of less common diseases. The system uses more than 5,000 clinical findings to describe the features of diseases in the QMR knowledge base. Findings include medical history, symptoms, physical signs, and laboratory test results. Every disease profile included in the QMR knowledge base is the result of an

extensive review of the primary medical literature. Consultation with experts is used to resolve any inconsistencies or deficiencies found in published reports. QMR is used in in-hospital and office practice.

MAMMEX (Ngah et al., 1996) is an image integrated expert system for the diagnosis of breast diseases. This system contains integrated image processing functions for the enhancement and analysis of mammograms. It is a rule-based system that uses patient's history, physical findings and mammographic signs for the diagnosis of breast diseases. The system covers seven level of benign/malignancy and 33 diseases.

Alhady et al. (2000) developed a noiseless ECG monitoring system with integrated expert system. This expert system is developed to detect certain types of heart problems. The user will have to interpret the ECG signals and answer the expert system questions. The integrated ECG monitoring system filters out interference to the signals in order to obtain a correct interpretation.

XBONE (Hatzilygeroudis et al., 1997) is a hybrid expert system supporting diagnosis of bone diseases. The diagnosis of this expert system is based on various patient data and is performed in two stages. In the early stage, diagnosis is based on demographic and clinical data of the patient, whereas in the later stage it is mainly based on nuclear medicine image data. Knowledge is represented via an integrated formalism that combines production rules and the Adaline artificial neural unit.

3.6.1 Expert Systems for Lung Diseases

There has been substantial work done in the area of expert system development for lung disease diagnoses.

PUFF (Aikins et al., 1983) is a medical expert system designed to determine the presence and severity of lung disease in a patient by interpreting measurements from pulmonary

function tests. The knowledge of the system is represented in the form of rules. The system consists of a rule interpreter, explanation facility and knowledge acquisition facility. PUFF had 400 rules and 75 clinical parameters. The system utilizes the backward chaining method of inferencing.

PXDES (Yue et al., 1986) is a Pneumoconiosis X-Ray diagnosis expert system. Pneumoconiosis is an occupational lung disease. This system uses the analysis of chest X-Rays for diagnosis. PXDES examines the shadows present on chest X-Rays and determines the type and degree of pneumoconiosis. This system contains three major components which are the knowledge base, the explanation interface, and the knowledge acquisition mode. The knowledge base contains the radiological information of the various stages of the disease. The inferencing is performed using approximate reasoning. The knowledge acquisition mode allows medical experts to add or change information in the system.

Balko et al. (1994) developed an expert system for teaching cytopathologic diagnosis of Lung Cancer. The system uses cytologic criteria as feature parameters for comparison against a set of known carcinomas. The system's diagnosis corresponds to the degree of agreement between the cytologic criteria and the carcinomas. The system allows the user to compare his or her description with that of the expert for any of the tumor entities. The knowledge base is represented by a numerical matrix within the system. This matrix is used for parameter comparison and score interpretation.

Pereira et al. (2002) developed an expert system to diagnose Pneumonia. The system uses patient's chest X-Ray information and clinical symptoms to perform its diagnosis. This expert system utilizes fuzzy logic for its inferencing. The diagnosis is performed using defuzzification technique. Pereira et. al studied to different method of defuzzification which are the *max-min relation* and the *Godel's implication* to determine the method that yield's the best result.

TUBERDIAG (Phuong et al., 1998) is an expert system for the diagnosis of Tuberculosis. It is a rule-based system where each rule is assigned a weight of 0 or 1. This system combines both positive and negative knowledge in its knowledge base. The final diagnosis is a total degree of confirmation and exclusion of tuberculosis.

Economou et al. (1994) developed a medical expert system for the diagnoses of pulmonary diseases. It covers the full spectrum of pulmonary diseases. The system uses Artificial Neural Networks (ANNs) method for its inferencing procedure. The ANNs were trained using real patients' clinical data.

Summary

Expert system is an intelligent computer system that simulates the judgment and behavior of a human expert in order to solve complex problems. Natural knowledge representation as well as high quality problem solving capability have made expert system to be widely used a variety of applications. The system has many features that make it suitable medical diagnosis applications. The use of expert systems is quite prevalent in the medical field. Currently, many expert systems are being used in hospitals and clinical care facilities for a variety of purposes such as diagnosis, treatment, monitoring and consultation. Most of the systems are focused on diagnosing a particular disease. Few expert systems for diagnosing lung diseases have been developed. These systems mostly diagnose a particular lung disease only. Apart from that, the expert systems use only one type of medical information such as blood test, PFT data, radiological information or symptoms for its diagnosis. However, the expert system proposed in this work is able to diagnose 60 types of lung diseases using four types of medical information. The integration of the imaging module into the expert system is another new feature that is not available in most of the expert systems.

CHAPTER 4

DEVELOPMENT OF CHEST X-RAY IMAGE PROCESSING TOOL

Introduction

This chapter presents the image processing tasks for chest X-Ray image enhancement and analysis. This imaging module has been named as Chest X-Ray Image Processing Tool (CIPT). Image enhancement is useful for feature extraction, image analysis and improvement of visual clarity. It can emphasize the characteristics of certain features in the raw image. This module is embedded into the expert system in order to extract certain features clearly. The imaging module is divided into two main parts. The first part deals with the general enhancement and analysis of chest X-Ray images using image processing functions which are discussed in this chapter. The second part is a semi-automated computer aided lung nodule detection and diagnosis tool for lung cancer which is discussed in detail in chapter 5. The main functions of the imaging module are to assist in detecting lung diseases in general and lung cancer in particular.

4.1 Digital Enhancement and Analysis of Chest X-Ray

Chest X-Ray is one of the most commonly performed medical diagnostic examination. It is used for the evaluation of a wide range of clinical problems. Chest X-Ray is widely preferred because it is inexpensive, commonly available and non-invasive. Apart from that, the x-rays which are a type of electromagnetic radiation, are invisible and create no sensation when they pass through the body. Thus, there is no discomfort caused by the procedure. The chest X-Ray is one of the lowest radiation exposure medical examinations performed today. Modern, state-of-the-art X-Ray systems have tightly controlled X-Ray beams with significant X-Ray beam filtration and dose control methods. Thus, stray or

scatter radiation is minimized, and those parts of a patient's body not being imaged receive minimal exposure.

The chest X-Ray is very useful for examination, but it has its limitations. Some types of abnormalities are not visible in chest X-Rays. Technical manipulation of the X-Ray equipment may increase the likelihood one type of abnormality being detected but most often can decrease the likelihood of the detection of another abnormality. For example, if the X-Ray equipment is adjusted so that it can improve the visibility of un-obscured lung, at the same time, it may tend to diminish the visibility of lung projecting behind the heart or mediastinum. Features of clinical importance in chest radiography that often escape detection include small, faint, opacities resulting from an early lung cancer or faint linear opacities caused by early interstitial disease. Therefore, a normal chest X-Ray does not necessarily rule out all problems in the chest. Most often, a chest CT is performed to further clarify a finding seen on the chest X-Ray or to look for an abnormality in order to diagnose the clinical problem.

The quality of the X-Ray image is also another main concern. Image contrast, sharpness and noise are the mainstays in the quantification of image quality for chest radiography. Most often, the X-Ray film is corrupted by noise. The major type of noises corrupting X-Ray is quantum noise and structured noise.

Quantum noise is caused by the X-Ray equipment itself. It is most prevalent in old X-Ray systems. The noise in the image can be reduced with increasingly efficient recording systems such as those in new digital detectors. However, most small hospitals and clinics do not have high-end X-Ray equipments.

Normal anatomic structures are the chief source of structured noise. The noise may camouflage subtle abnormalities or distract the observer's attention. Structures noise increases when the image sharpness is increased in the digitizing equipment. Thus, there is always a trade-off between image quality and noise corruption.

Low contrast in the X-Ray film is another major challenge in detecting abnormalities. Due to the nature of the X-Ray, most often there might not be sufficient contrast between the different anatomical structures in the chest region, making detection a tedious task. The contrast for regions of interest can be improved by x-ray scatter rejection technology. However, this is again rather costly (“Image quality in chest radiography”, 2003).

Digital enhancement of chest X-Rays allows a more thorough and accurate interpretation of complex cases without resorting to more expensive procedure like CT scan, follow-up patient examinations and other secondary procedures. This also allows for quicker diagnoses to be made for routine cases. This is most useful for general physicians who do not possess in-depth knowledge and experience in medical image interpretation.

Through the application of image processing techniques, chest X-Rays images which are considered unsatisfactory can be rendered useful. Many limitations of chest X-Ray films can be overcome by applying imaging techniques. X-Ray imaging is commonly done using film. Digitizing the X-Ray images by means of medical scanner or digital camera allows the images to be manipulated by imaging software in a variety of ways.

When an image is transformed to its digital equivalent, it is represented by a two-dimensional numerical matrix, each corresponding to pixels. These pixels have information regarding the light intensity (brightness) and their location (x, y coordinates). The intensity value of each pixel would be proportionate to the X-Ray absorbed by the structure in the anatomical region that intersects that pixel.

Digital imaging systems allow image acquisition, display, storage, enhancement and analysis. Various types of image processing functions could be used to enhance and analyze the images as needed by the doctors. Most often, the imaging functions are used for improving the perceptive quality of the image, extraction of objects of interest in the

image by an automated or semi-automated approach and performing feature measurements such as area, perimeter and shape factors (Nghah, 1996).

Thus, computer aided image enhancement of chest X-Rays have the potential to improve the detection rate of lung diseases as well as improving the diagnostic performance of doctors.

4.2 Image Processing Tasks

Image processing functions have been developed for the enhancement of chest X-Ray images. The abnormalities associated with lung diseases form a very large number. Thus, in this module, the imaging functions suitable for the enhancement and analysis of general lung abnormalities have been included. The users can use any combination imaging functions as suited for their individual requirements. It is assumed that the system would be used by general physicians and radiologists who do not have in-depth knowledge in the utilization of image processing functions. Thus, a simple and user-friendly interface is adopted. The image processing task involved in this work is shown in Figure 4.1.

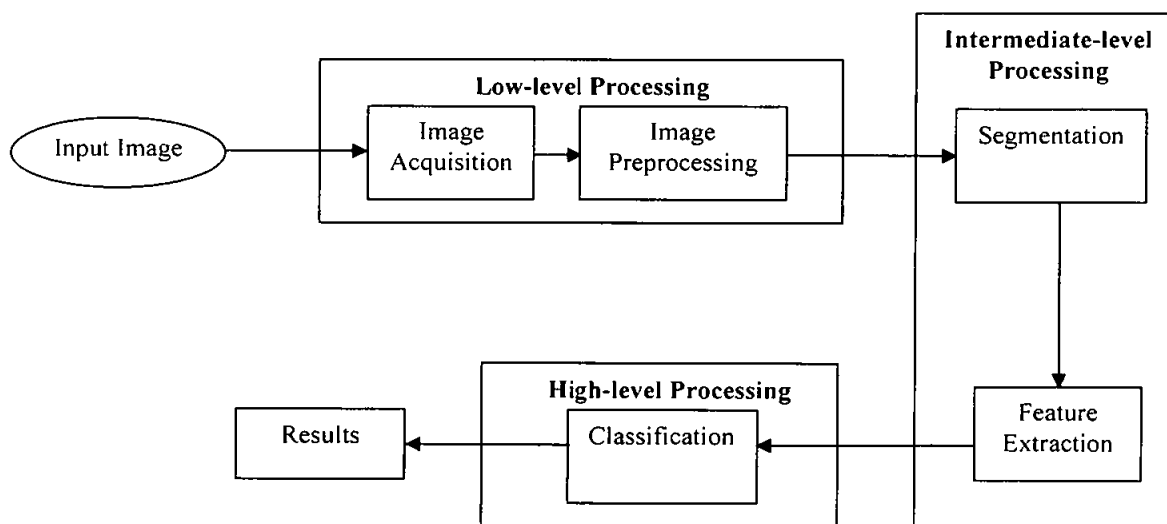


Figure 4.1: Image Processing Task.

There are many functions associated with each of the image processing task. The functions used in each of the task are described in the section below. The theoretical formulation of the image processing functions is available in Chapter 2.

4.2.1 Low-level Processing

4.2.1.1 Image Acquisition

Image acquisition involves the capturing of the image and its subsequent digitization. The images used in the imaging module are initially in the form of chest X-Ray films. Digital images can be acquired by a variety of methods such as frame grabber, digital camera and scanner.

In this work, the images were acquired by using a high resolution medical image scanner. The films were scanned with a resolution of 2000 ppi. The images are of 8 bit grayscale. The size of the image may vary depending on the computational power of the computer, computational time and image aspect ratio.

4.2.1.2 Image Preprocessing

Preprocessing is a very crucial aspect in image processing. Quite often, the image obtained might not be satisfactory. This is due to many factors such as noise, improper contrast and bad lighting. Through preprocessing, unwanted variations and noise can be reduced and the desired features can be enhanced. There are many image processing functions that can be applied for preprocessing.

A low pass filter can be effectively used to suppress the high frequency noise. The high pass filter contributes to the reduction in overall contrast and average intensity and apparent sharpening of edges and sharp details (Russ, 1998). There are many types of low-pass and high-pass filter masks available. This imaging tool allows the user to

specify own filter mask and thus customize the filtering process suitable for the user's need. This would be very useful in situations where the user need to apply a specific customized mask.

Median filter is a very popular filter used for noise reduction and image smoothing. The median filter can be effectively used to remove 'salt & pepper' noises in the chest X-Ray images. Median filtering also tends to smooth the X-Ray image by reducing the intensity of abnormalities such as masses and nodules (Chiou et al., 1994).

The Gaussian filter is used to enhance high intensity grayscale structures such as the irregularly shaped masses found in chest X-Rays (Munday et al., 2002). Weiner filter is normally used for noise suppression in noisy images. Wiener filtering is quite useful for signal dependent images. As noise in low-dose X-Ray images is inherent signal-dependent, a significant improvement in noise reduction can be obtained through wiener filtering (Hensel et al., 2005). Unsharp masking technique is very useful in enhancing the details in chest X-Ray images. The sharpening of the X-Ray image will make the edges of the anatomical structures in the image clearer for visual inspection.

Image contrast enhancement is important in medical applications. This is due to the fact that visual examination of medical images is essential in the diagnosis of many diseases. In applications such as chest radiography, the image contrast is inherently low due to the small differences in the X-Ray attenuation coefficients. The problem is further complicated if an image consists of several regions with different X-Ray attenuation characteristics. For example, in chest X-Rays, the mediastinum and the lung field have different exposures. It is usually desirable to enhance the details in both regions simultaneously. Contrast stretching and histogram equalization are two widely utilized methods for global image enhancement.

CLAHE is a new method that is able to effectively enhance an image by using both contrast stretching and histogram equalization approach. CLAHE is able to produce

X-Ray images in which the noise content is not excessively enhanced, but sufficient contrast is provided for the visualization of structures within the image. X-Ray images processed with CLAHE have a more natural appearance and facilitate the comparison of different regions within the image. Negation of digital images is normally useful in applications where a positive film is used to obtain a negative film as an output.

Geometric transforms include those processes that change dimensional properties of images. These include resizing, cropping and rotating. Rotation enables an image to be resized according to a user specified size. The resizing function is performed using three different interpolation methods which are Bicubic Interpolation, Bilinear Interpolation and Nearest Neighbor method. The user can select whichever interpolation method that gives the best visual results for the chest X-Ray (Burdick, 1997). Cropping is the process of selecting a small portion of an image and cutting it away from the rest of the image. This function crops an image to a specified rectangle. The function displays the input image and waits for the user specify the region of interest (ROI) with the mouse (Castleman, 1995).

4.2.2 Intermediate-Level Processing

4.2.2.1 Segmentation

In the analysis of the objects in images it is essential to distinguish between the objects of interest and the background. The segmentation techniques are used to find the objects of interest. For segmentation, we use thresholding, seed-based region growing and edge detection techniques.

In thresholding, a threshold value is selected based on the image histogram or some other user specific calculation and applied to an image. The resultant image is in binary form. The seed-based region growing method can be applied to images with variations in their grayscale range. However, suitable preprocessing might be necessary before the

application of region growing in order to obtain a correct region. The positions of the initial seeds are very crucial. It is possible that user makes mistake selecting seed pixels. A proper homogeneity criteria ensure that correct results is obtained. Incorrect selection of seed points and homogeneity criteria can lead to less or more regions on the images. In some cases, post-processing can be done to correct result of the segmentation (Nghah et al., 2000 & Ilic et al., 2000).

Edge detection filters segment edges in an image. Laplacian filter could be used to detect the abnormalities such as cracks in the ribs, the shape of abnormal masses and nodules and etc. Prewitts and Roberts are directional edge detectors. These filters are used in highlighting and analyzing specific areas in the chest X-Ray such as a particular rib.

4.2.2.2 Feature Extraction

Feature extraction seeks to identify inherent characteristics or features of objects within an image. These characteristics are used to describe the attributes of the object. These features greatly assist in object description.

Statistical features such as area and perimeter of an object in an image is an important indicator of object size. Moments provides a method of describing the properties of an object in terms of its area, position, orientation and other precisely defined parameters. Detailed discussion on moments is given in Chapter 5.

Morphology relates to the structure or form of objects. Morphological filtering simplifies a segmented image to facilitate the search for objects of interest. This is done by smoothing out object outlines, filling small holes, eliminating small projections, and using other similar techniques (Bovik, 2000).

Dilation allows object to expand, thus potentially filling in small holes and connecting disjoint objects. Erosion shrinks objects by eroding their boundaries. Erosion is

commonly used to remove unwanted structures from the chest X-Ray images. One of the uses of erosion is for eliminating irrelevant details in a binary or grayscale image. Both of these operations can be customized for specific operations by the use of suitable structuring element (Gonzalez et al., 2002).

Opening generally smoothes the contours of an object, breaks narrow isthmuses and eliminates thin protrusions. Closing tends to smooth contours, but it generally fuses narrow breaks and long thin gulfs, eliminates small holes, and fills gaps in the contour (Gonzalez et al., 2002). Boundary extraction is applied on binary images for the extraction of an object's boundary.

Arithmetic operations are used extensively in image processing. These functions serve a variety of purposes in image analysis. Addition takes as input two identically sized images and produces as output a third image of the same size as the first two. This function is very useful in adding two images, one a normal image and the other an abnormal image, so that the abnormality such as masses, nodules or shadows is superimposed on one another.

Subtraction takes two images as input and produces as output a third image whose pixel values are simply those of the first image minus the corresponding pixel values from the second image. Image subtraction is popularly used for the detection of tumors and nodules in the chest X-Ray. A current image is subtracted with a previously taken image and the resultant image would be an image with a visible tumor or nodule.

4.2.3 High-Level Processing

4.2.3.1 Classification

Classification is the process of classifying an object in an image. It involves object recognition and pattern matching. In this imaging module, Fuzzy Inference System has

been introduced as a classifier for the diagnosis of the lung nodules. Detailed information about Fuzzy Inference System development is described in Chapter 5.

Summary

Image processing techniques have the ability to improve the quality of chest X-Ray films tremendously. Images which were deemed useless can be rendered useful through the application of imaging functions. The imaging task involved in the enhancement and analysis of an image include low-level processing such as image acquisition and preprocessing followed by intermediate-level processing such as segmentation and feature extraction and finally high-level processing such as classification. The imaging functions which incorporate each of the tasks are available in CIPT. The processed images would be able to show details and abnormalities that were originally not visible in the initial X-Ray. The processed images help the general physicians to interpret the chest X-Ray more easily and enable them to answer the questions relating to radiological findings in the expert system in a more accurate manner.

CHAPTER 5

DEVELOPMENT OF LUNG NODULES DETECTION AND DIAGNOSIS TOOL

Introduction

This chapter describes the development of the computer-aided lung nodule detection and diagnosis tool. This tool is part of the CIPT imaging module. Lung cancer is a fatal condition with a rapidly increasing incidence rate. Thus, the diagnosis of lung cancer is given importance in this research work. This tool has been developed to aid in the detection and diagnosis of lung nodules in chest X-Rays. Lung nodules are the major indicators of lung cancer. This semi-automated tool enables the general physicians to detect nodules in chest X-Rays and diagnose them. The detection of lung nodules is achieved by using morphological transform and seed-based region growing method. The diagnosis uses feature extraction and fuzzy inference techniques. Diagnosis is done by determining a malignancy factor which indicates the probability of malignancy caused by the nodules. This tool is specifically for the diagnosis of lung cancer and thus can be used either separately or with the expert system.

5.1 Overview of Lung Cancer

Chest radiography such as X-Ray and CT scan are crucial sources in diagnosing lung cancer. In both chest X-Rays and chest CT scans, lung cancer can be detected by the presence of lung nodules. Both solitary and multiple lung nodules can be an indication of lung cancer. Lung nodules are round or oval, sharply defined lung lesions. These nodules can either be benign (non-cancerous) or malignant (cancerous). Nodules show up as relatively low-contrast white circular objects within lung fields.

The challenge for this proposed system is to increase the visibility of these nodules and distinguish the true nodules from overlapping structures such as vessels and ribs. This diagnosis system proposed in this work will use normal chest X-Rays to detect the nodules and predict the severity of the cancer.

5.2 Methodology

The computer aided detection and diagnosis tool has two major parts which are the:

- Lung Nodule Detection
- Lung Nodule Diagnosis

The flowchart of the tool is shown in Figure 5.1.

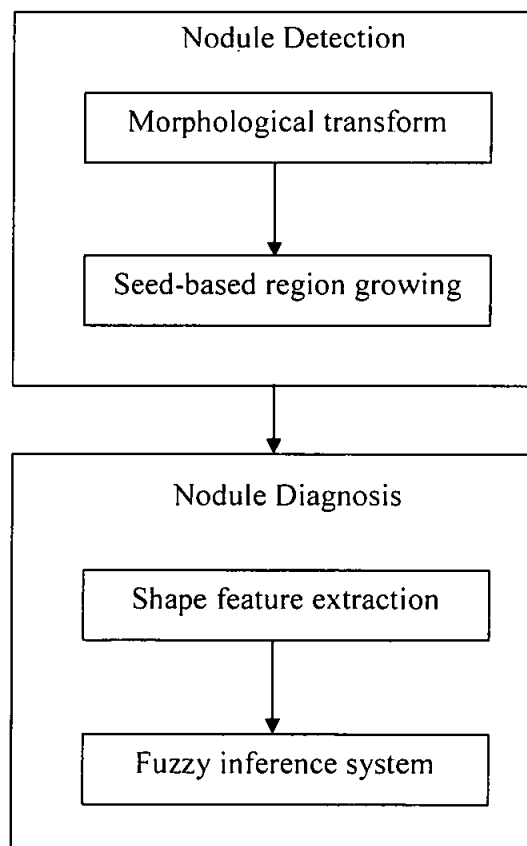


Figure 5.1: Lung nodules detection and diagnosis process flowchart.

5.2.1 Lung Nodule Detection

Most often lung nodules are not visible in the 2-dimensional chest X-Rays. This is because the nodules are often hidden between structures such as the heart, ribs, sternum and etc. Enhancement is performed in order to remove the unnecessary structures and large layers of connected tissues so that the underlying nodules are visible to the eye of the doctor. This is done by a morphological transform. Then the nodules are segmented from the other structures in the image for further analysis. This is achieved by a seed-based region growing technique.

5.2.1.1 Morphological Transform

The first stage in the detection of the lung nodules is the enhancement of the nodules in order to increase their visibility in chest X-Rays. For this purpose, a grayscale morphological transform is found to be well suited (Wróblewska et al., 2003, Fahmy et al., 2004 & Lee et al., 2004).

In order to carry out the enhancement process, a new algorithm called Modified Grayscale Top Hat (MGTH) transform has been developed. This algorithm uses a combination of Dilation, Erosion, Opening and Closing operations to define a modified Top Hat Transform (Gonzalez et al., 2002).

Grayscale dilation of the original image f by structuring element b is denoted by

$$(f \oplus b)(s, t) = \max \{f(s - x, t - y) + b(x, y) \mid (s - x), (t - y) \in D_f; (x, y) \in D_b\} \quad (5.1)$$

where x and y are the coordinate of the structuring element, s, t are the coordinates of the input image, and D_f and D_b are the domains of f and b respectively (Gonzalez et al., 2002).

Grayscale erosion of f by b is denoted by

$$(f \otimes b)(s, t) = \min \{f(s + x, t + y) - b(x, y) \mid (s + x), (t + y) \in D_f; (x, y) \in D_b\} \quad (5.2)$$

Grayscale Opening is denoted by $f \circ b$ and is given below.

$$f \circ b = (f \otimes b) \oplus b \quad (5.3)$$

Similarly, grayscale closing is denoted by $f \bullet b$ and is given below.

$$f \bullet b = (f \oplus b) \otimes b \quad (5.4)$$

A White Top Hat (WTH) transform is defined as the subtraction of an opened image from the original image and is given below.

$$WTH = f - (f \circ b) \quad (5.5)$$

A Black Top Hat (BTH) transform is defined as the subtraction of the original image from a closed image and is given below.

$$BTH = (f \bullet b) - f \quad (5.6)$$

Using a combination of the processes listed above, the MGTH transform is developed.

For an image, f , the MGTH is defined as shown in equation 5.7.

$$\begin{aligned} \text{MGTH} &= WTH - BTH \\ &= (f - \text{Opening}) - (\text{Closing} - f) \\ &= (f - (f \circ b)) - ((f \bullet b) - f) \\ &= 2(f) - (f \circ b) - (f \bullet b) \end{aligned} \quad (5.7)$$

Equation 5.7 can be further expanded as below.

$$\text{MGTH} = 2(f) - ((f \otimes b) \oplus b) - ((f \oplus b) \otimes b) \quad (5.8)$$

The intensity of the pixels in the image is directly proportional to the amount of X-Ray absorbed by the anatomical structures. The X-Ray absorption is proportional to tissue density of the structures. The denser the tissue, the whiter it appears on an X-Ray. The less dense the tissue, the blacker it appears.

The chest X-Ray contains a few major regions such as the rib bones, spine, heart, lungs and abnormalities such as the nodules. Structures such as the rib bones, spine and heart are high in density and thus absorb high amount of X-Ray. This causes the structures to have high pixel intensity regions in the digitized X-Ray image. The lung region is of low density and thus shows up in the X-Ray image as a low intensity region or as a back region. Lung nodules have high tissue density and thus constitute a high intensity region.

The rib bones, spine and heart appear connected in the X-Ray image and thus have a large size while the nodules are mostly isolated and are of small size. The region of interest in the X-Ray is the lung nodules. The other structures such as the rib bones, spine and heart are not needed. The objective of the MGTH algorithm is to highlight the nodules in the X-Ray while removing the unwanted structures.

The original chest X-Ray is first subtracted with its opened image. This is the WTH transform operation. This operation will enhance the bright features that are smaller than the size of the structuring element in an image (Mukhopadhyay et al., 2002 & Barata et al., 2002). In this operation, the original chest X-Ray image will first be eroded. Any structure smaller than the size of the structuring element, such as the nodules would be removed while leaving the larger bright regions undisturbed. The subsequent dilation would increase the overall intensity of the image without reintroducing the structures removed by the erosion process. This is the opened image. Subtracting this image from

the original image will highlight the bright nodule areas while suppressing the large bone and heart area.

The structuring element selected would be disc shaped with a size slightly larger than the largest expected nodule size (Goatman, 1997). A disc shaped structuring element is used as its shape is similar to the shape of the nodules. This ensures that there would not be any distortion to the nodule shape. The largest expected nodule size in chest X-Rays are about 3 cm in diameter (Leef et al., 2002). Thus, the structuring element used would be larger this size. The exact size of the structuring element will be dependent on the size of the original digital image and the calibration factor used in the digitization. This will vary within chest X-Ray films of different sources. The WTH operation will remove all the connected structures as they are significantly larger than the nodule size.

BTH transform will highlight all the dark structures in the X-Ray image (Mukhopadhyay et al., 2002 & Barata et al., 2002). First, the chest X-Ray would be dilated. This will remove all the dark details in the image and brighten the image. The subsequent erosion darkens the image without reintroducing details removed by dilation. This will remove all the dark details in the image such as the background lung field while leaving the bright regions such as the bone structures intact. This is the closed image. Subtracting the original image from the closed image will remove all the bright bone structures. Thus, in the BTH operation, all the bright regions are suppressed while the dark regions are highlighted.

When the BTH image is subtracted from WTH image, only the bright nodules and few isolated bright regions are highlighted and quite visible to the eye while all the large bone structures are mostly removed. Thus, the MGTTH processed image will only constitute of the region of interest which is the nodules and some isolated bright regions.

5.2.1.2 Region Growing Technique

In the MGTH processed image, in addition to the nodules, there might be some unwanted small bright structures still present. Apart from that, not all the nodules visible in the processed image are lung nodules. There may be nodule-like structures such as infiltrates and small consolidation present in the X-Ray. Thus, the physician or radiologist has to select the regions which they suspect of being lung nodules.

This is achieved through the use of seed-based region growing method. The physician will mark the suspicious regions in the X-Ray image and a seed pixel is chosen within the suspected region and then the seed-based region growing algorithm is applied. The resultant grown regions will be the suspicious nodules. The seed-based region growing technique used was described in detail section 4.2.6.1 of Chapter 4.

The resultant image is a binary image with the suspected nodules as white objects and the rest as the black background. The homogeneity criteria used for growing the nodule region maybe any intensity tolerance value τ . The value of τ may differ based on the texture and the grayscale distribution of the image. For the set of test cases used in this work, the tolerance value of 10 was chosen through a trial and error method to include as much as the suspicious region without expanding or growing other parts.

5.2.2 Lung Nodule Diagnosis

Lung nodules are the prime indicators of lung cancer in chest X-Ray. Lung nodules can either be benign or malignant. Benign nodules are not dangerous while malignant nodules pose a very serious health threat. Benign nodules are small (less than 20mm in diameter), have smooth, well defined margins, and tend to be symmetrical and spherical. Malignant nodules tend to have irregular, lobulated or shaggy borders (Farzan et al., 1997). Figure 5.2 shows some of the shapes of benign nodules and Figure 5.3 shows some sample

shapes of malignant nodules. Commonly the diagnosis of lung nodules involves the determination of whether the nodule is malignant or benign.



Figure 5.2: Sample benign nodule shapes.



Figure 5.3: Sample malignant nodule shapes.

In this work, the diagnosis of the lung nodules is the determination of the level of malignancy. This translates into how dangerous the nodules are in term of percentage. Thus, the diagnosis is made in terms of a malignancy factor. The malignancy factor is in percentage form. Low percentage indicates low probability of malignancy and vice versa.

The shape of the nodule strongly correlates with the level of malignancy of the lung nodules (Farzan et al., 1997). Shape information of the nodules is used to obtain the level of malignancy of the nodules. Thus, feature extraction is performed to extract shape information of the lung nodules. The features extracted from the region grown image should be ideally translation, rotation and scale invariant and provide a defined difference between regular and irregular shapes. The shape features used are compactness (circularity), dispersion, eccentricity, solidity and moments. These shape factors is quite suitable to be used for the analysis of the shape of the lung nodules (Angelina et al., 2003, Ongun et al., 2001 and Gonzalez et al., 2002). These features values obtained are used as input to a fuzzy inference system to obtain the malignancy factor.

5.2.2.1 Shape Feature Extraction

5.2.2.1.1 Compactness

Compactness provides a measure of how the shape of an object approaches to a circle. A circle has a value of 4π and this value will be different for different shapes.

It is defined by the formula

$$C = P^2 / A \quad (5.9)$$

where P is the perimeter of the region and A is the area of the region.

5.2.2.1.2 Dispersion

Dispersion measures how unevenly an object's mass is distributed around the center of the mass of the object region. The measure of dispersion (irregularity) can be expressed as the ratio of the maximum radius to the minimum radius. That is,

$$D = \max(\sqrt{(x_i - \bar{x})^2 + (y_i - \bar{y})^2}) / \min(\sqrt{(x_i - \bar{x})^2 + (y_i - \bar{y})^2}) \quad (5.10)$$

where x_i and y_i are pixel coordinates and \bar{x} and \bar{y} are coordinates of the centroid.

This measure defines the ratio between the radius of the maximum circle enclosing the region and the minimum circle that can be contained in the region. The measure will increase as the region spreads. Irregular shapes have more spread than regular shapes (Aguado, 2002). Thus, irregular shapes will have higher values for D than regular shapes.

5.2.2.1.3 Solidity

Solidity describes the extent to which the shape is convex or concave. Solidity deals with the proportion of pixels in the convex hull, which was also in the object. Convex hull is the smallest convex polygon that contains the object (Angelina et al., 2003). Regular shapes are more solid and will have high values while irregular shapes are less solid and thus have comparatively low values. Solidity is given by

$$S = A / A_c \quad (5.11)$$

where A is the area of the object and A_c is the area of the convex hull polygon

5.2.2.1.4 Moments

Moments are global descriptors of shape. It describes a shape's layout with invariance properties (Aguado, 2002). Moments have the capability to provide a distinct difference between regular and irregular values (Shen et al., 1994).

The central moment is defined as (Gonzalez et al., 2002)

$$\mu = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y) \quad (5.12)$$

where $f(x, y)$ is the image, p and q are non-negative numbers with the moment of order is defined as $(p+q)$ for $p, q = 0, 1, 2, \dots$ and \bar{x} and \bar{y} are the coordinates of the centroid.

The central moments of order up to three are defined in the equations below.

$$\begin{aligned}
 \mu_{00} &= \sum_x \sum_y (x - \bar{x})^0 (y - \bar{y})^0 f(x, y) \\
 &= \sum_x \sum_y f(x, y) \\
 &= m_{00}
 \end{aligned} \tag{5.13}$$

$$\begin{aligned}
 \mu_{10} &= \sum_x \sum_y (x - \bar{x})^1 (y - \bar{y})^0 f(x, y) \\
 &= m_{10} - \frac{m_{10}}{m_{00}} (m_{00}) \\
 &= 0
 \end{aligned} \tag{5.14}$$

$$\begin{aligned}
 \mu_{01} &= \sum_x \sum_y (x - \bar{x})^0 (y - \bar{y})^1 f(x, y) \\
 &= m_{01} - \frac{m_{01}}{m_{00}} (m_{00}) \\
 &= 0
 \end{aligned} \tag{5.15}$$

$$\begin{aligned}
 \mu_{11} &= \sum_x \sum_y (x - \bar{x})^1 (y - \bar{y})^1 f(x, y) \\
 &= m_{11} - \frac{m_{10} m_{01}}{m_{00}} \\
 &= m_{11} - \bar{x} m_{01} \\
 &= m_{11} - \bar{y} m_{10}
 \end{aligned} \tag{5.16}$$

$$\begin{aligned}
 \mu_{20} &= \sum_x \sum_y (x - \bar{x})^2 (y - \bar{y})^0 f(x, y) \\
 &= m_{20} - \frac{2m_{10}^2}{m_{00}} + \frac{m_{10}^2}{m_{00}} \\
 &= m_{20} - \frac{m_{10}^2}{m_{00}} \\
 &= m_{20} - \bar{x} m_{10}
 \end{aligned} \tag{5.17}$$

$$\begin{aligned}
\mu_{02} &= \sum_x \sum_y (x - \bar{x})^0 (y - \bar{y})^2 f(x, y) \\
&= m_{02} - \frac{m_{01}^2}{m_{00}} \\
&= m_{02} - \bar{y}m_{01}
\end{aligned} \tag{5.18}$$

$$\begin{aligned}
\mu_{21} &= \sum_x \sum_y (x - \bar{x})^2 (y - \bar{y})^1 f(x, y) \\
&= m_{21} - 2\bar{x}m_{11} - \bar{y}m_{20} + 2\bar{x}\bar{y}m_{01}
\end{aligned} \tag{5.19}$$

$$\begin{aligned}
\mu_{12} &= \sum_x \sum_y (x - \bar{x})^1 (y - \bar{y})^2 f(x, y) \\
&= m_{12} - 2\bar{y}m_{11} - \bar{x}m_{02} + 2\bar{x}\bar{y}m_{10}
\end{aligned} \tag{5.20}$$

$$\begin{aligned}
\mu_{30} &= \sum_x \sum_y (x - \bar{x})^3 (y - \bar{y})^0 f(x, y) \\
&= m_{30} - 3\bar{x}m_{20} + 2\bar{x}^2m_{10}
\end{aligned} \tag{5.21}$$

$$\begin{aligned}
\mu_{03} &= \sum_x \sum_y (x - \bar{x})^0 (y - \bar{y})^3 f(x, y) \\
&= m_{03} - 3\bar{y}m_{02} + 2\bar{y}^2m_{01}
\end{aligned} \tag{5.22}$$

The normalized central moments, denoted by η_{pq} , are defined as

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^\gamma}$$

where $\gamma = \frac{p+q}{2} + 1$, for $p+q = 2, 3, \dots$

A set of seven invariant moments can be derived from the second and third moments.

$$\begin{aligned}
\phi_1 &= \eta_{20} + \eta_{02} \\
\phi_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\
\phi_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\
\phi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\
\phi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\
\phi_6 &= (\eta_{20} - 3\eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{21})(\eta_{21} + \eta_{03}) \\
\phi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]
\end{aligned}
\tag{5.23}$$

This set of moments is invariant to translation, rotation and scale changes. For this work, the first three moments are used for describing the lung nodule's shape (Forero et al., 2004). Thus, ϕ_1 , ϕ_2 and ϕ_3 are subsequently called as Moment 1, Moment 2 and Moment 3.

5.2.2.1.5 Eccentricity

Eccentricity measures the degree to which mass is concentrated along a particular axis. The eccentricity of an object is the ratio of the distance between the foci length and the long axis length of its best fitting ellipse (Liu et al., 2000). An ellipse whose eccentricity is 0 is actually a circle, while an ellipse whose eccentricity is 1 is a line segment. Eccentricity, E can be expressed in terms of second-order central moments as below.

$$E = \frac{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2}{\mu_{00}}
\tag{5.24}$$

The eccentricity measure gives distinctly different values for regular and irregular objects (Forero et al., 2004).

The goal here is to differentiate between regular and irregular shapes. The shape factors described above are good discriminators for differentiating between the shapes (Angelina et al, 2003, Liu et al., 2000 & Forero et al., 2004). These factors are invariant to rotation, translation and scaling making it suitable for this application.

5.2.2.2 Fuzzy Inference System

The data (shape factors) extracted from the image are used as the inputs to the Fuzzy Inference System (FIS) to predict the percentage of malignancy. The percentage of malignancy is called as Malignancy Factor. For solving any problem, a person applies fuzzy reasoning process sub-consciously. A doctor’s thought process in determining a diagnosis for a disease also uses fuzzy reasoning. This can be modeled efficiently by the FIS (Serpen et al., 2000).

The process of fuzzy inference involves membership functions, fuzzy logic operators, and if-then rules. The components of a fuzzy inference system can be depicted as shown in Figure 5.4.

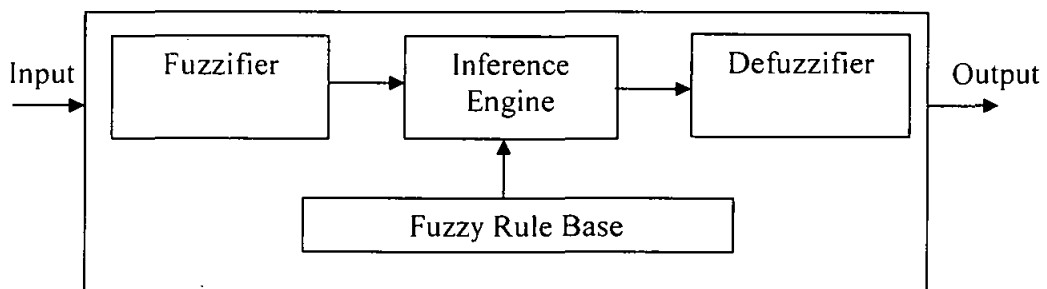


Figure 5.4: Fuzzy Inference System.

The FIS has four major components which are as listed below.

- Fuzzifier
- Fuzzy Rule Base
- Inference Engine
- Defuzzifier

The function of the fuzzifier is to convert the crisp numerical inputs into the corresponding appropriate fuzzy inputs. The fuzzy rule base contains a set of fuzzy rules. The role of the inference engine is to allow the application of the rules in the fuzzy rule base to the fuzzy input parameters whereby producing the output. The defuzzifier converts the output produced by the inference engine into user understandable terms.

Mamdani systems are widely accepted for capturing expert knowledge. It allows the expertise to be described in a more intuitive and human-like manner. As this system has to emulate the reasoning of a doctor, this model is used for the development of the FIS (Popoola, 2004). The FIS has five major steps. These steps are described in the following sections.

5.2.2.2.1 Fuzzification of Inputs

The inputs to the FIS are Compactness, Dispersion, Eccentricity, Solidity and Moments (Moment 1, Moment 2 and Moment 3). The inputs are crisp numerical values limited to the universe of discourse (input data range) of the input variables. The output of the fuzzification process is a fuzzy degree of membership in the qualifying linguistic set which is between the intervals of 0 to 1 (*Matlab Fuzzy Logic Toolbox Documentation*, 2001). The membership functions are set in order to determine to what extent an input value correlates to the malignancy of the nodules.











A membership function is a curve that defines how each point in the input space is mapped to a membership value (or degree of membership) between 0 to 1. The Gaussian membership function is used. Gaussian function provides a smooth mapping which is suitable for fuzzy inference. Apart from that, the data distributions tend to be normal, which can be approximated well by the Gaussian function. The curve creates a smooth transition and shows the degree of membership for the various inputs and output.

The data input range for each of the shape variables was determined by taking a set of images ranging from regular (benign) to irregular (malignant) and calculating their shape factors. The minimum and maximum values obtained for a set of test images was used as the limits for the membership functions (Serpen et al., 2000). Table 5.1 shows a subset of test images and their corresponding shape factors.

Each input has three fuzzy sets which are linguistically labeled as *Low*, *Medium* and *High*. The *Low* fuzzy set means a low probability of malignancy; *Medium* corresponds to a medium probability of malignancy and *High* to a high probability of malignancy. The membership functions of the inputs are shown in Figure 5.5 to 5.11.

The output of this FIS system is called Malignancy Factor. This system will give in terms of percentage the probability of malignancy factor of the lung nodules. The fuzzy sets for the output membership function are also divided into three which are *Low*, *Medium* and *High*. This is shown in Figure 5.12. The y-axis shows the fuzzy degree of membership which is set between 0 to 1. The x-axis shows the input data range.

Table 5.1: Shape factors.

	Image	Compactness	Dispersion	Eccentricity	Solidity	M1	M2	M3
shape 1		13.4868	1.0008	0.1994	0.9443	-3.8932	-2.9227	0.5858
shape 2		13.9563	1.0159	0.2102	0.9393	-3.4798	-2.0543	0.7499
shape 3		14.5628	1.2028	0.3125	0.9285	-2.8416	-1.0245	0.6905
shape 4		15.0586	1.5336	0.4185	0.8225	-2.4477	-1.0284	1.1024
shape 5		17.4562	1.7584	0.5213	0.7582	-1.8976	0.4186	1.5215
shape 6		26.8759	2.693	0.8409	0.4516	1.5283	3.1847	3.0546
shape 7		27.5628	2.3177	0.7789	0.4126	1.2784	2.4729	4.9867
shape 8		31.1256	2.9158	0.9371	0.4218	3.4519	4.6467	5.9341
shape 9		37.5691	2.9452	0.9215	0.3124	3.7195	5.0079	6.9514
shape 10		38.4168	2.9785	0.9412	0.2416	3.8316	5.6497	7.3511

M1: Moment 1, M2: Moment 2 and M3: Moment 3

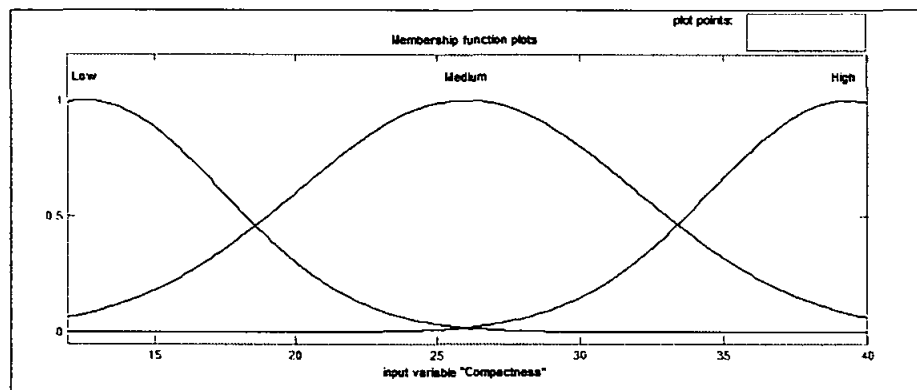


Figure 5.5: Compactness membership function.

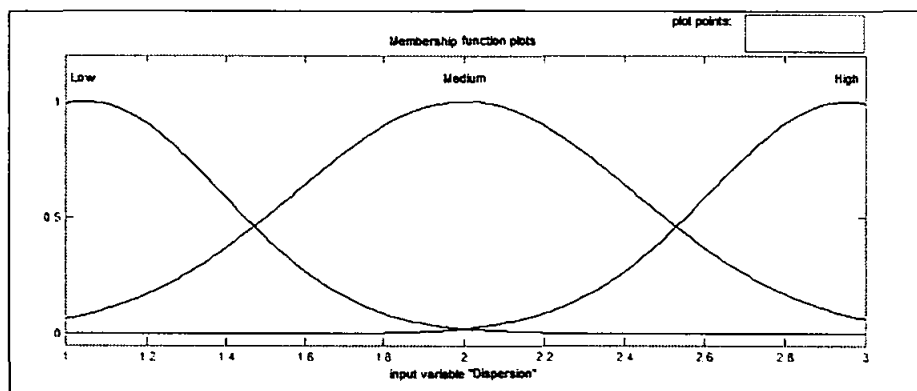


Figure 5.6: Dispersion membership function.

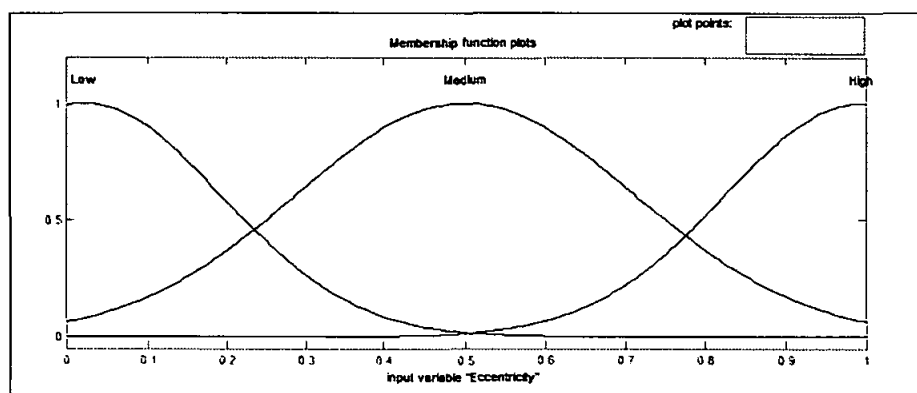


Figure 5.7: Eccentricity membership function.

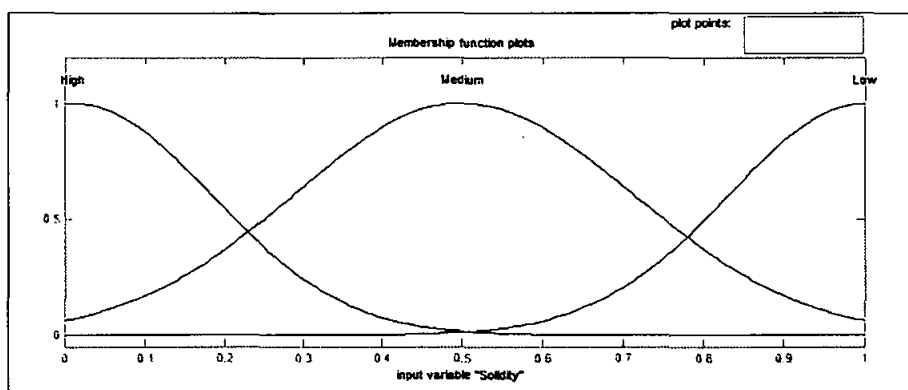


Figure 5.8: Solidity membership function.

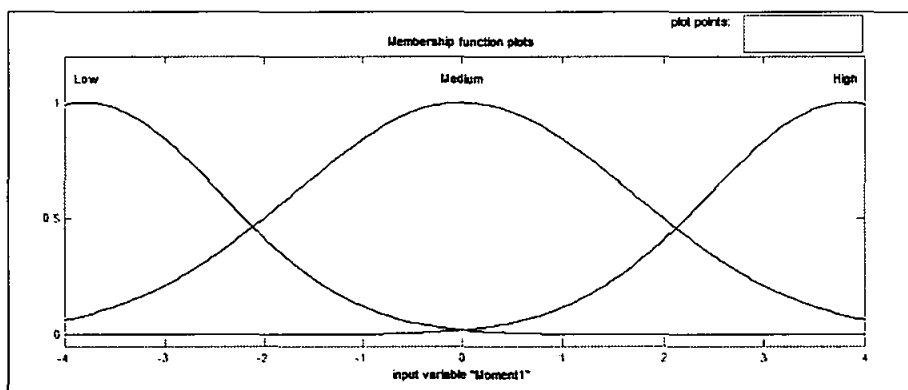


Figure 5.9: Moment 1 membership function.

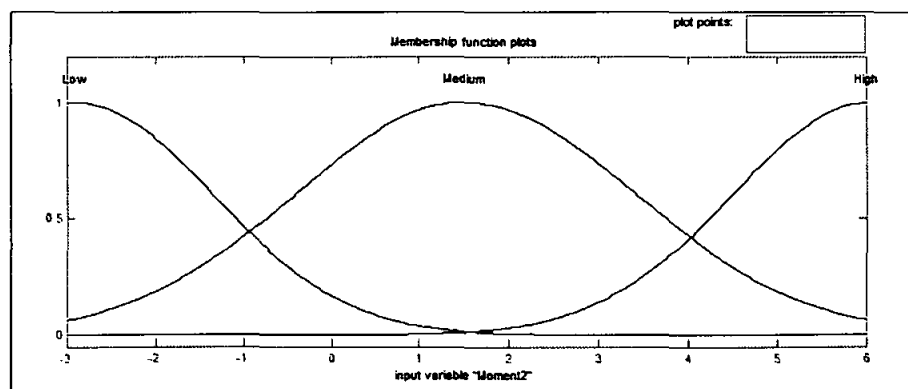


Figure 5.10: Moment 2 membership function.

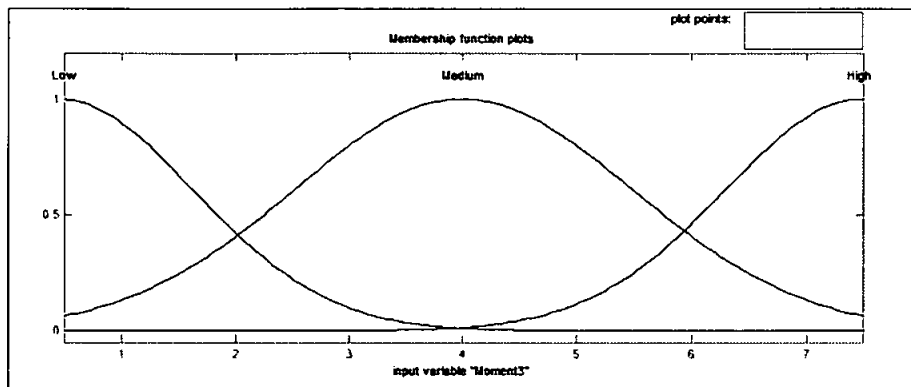


Figure 5.11: Moment 3 membership function.

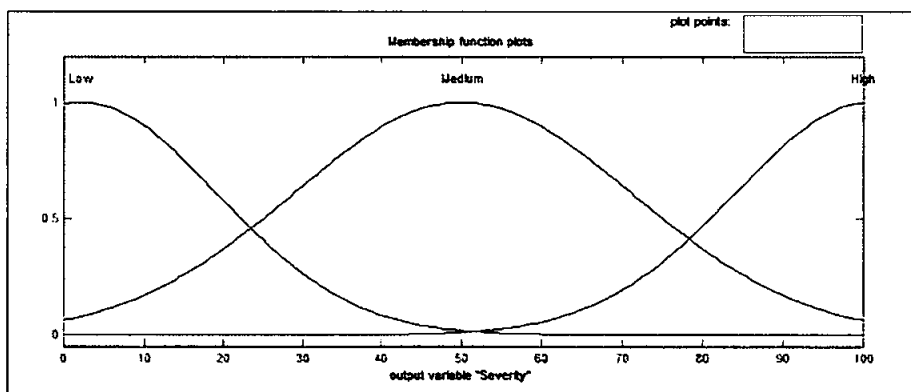


Figure 5.12: Probability of Malignancy membership function.

A fuzzy rule base is established for the diagnosis of the lung nodules based on the shape factors. The fuzzy rule base contains a set of fuzzy rules. The statements in the antecedent of the fuzzy rules are mapped to a degree of membership (*Matlab Fuzzy Logic Toolbox Documentation*, 2001). The rule base pertaining to the membership functions shown in Figure 5.5 to 5.12 is shown in Table 5.2.

The inputs are fuzzified. In the fuzzification process, the degree of membership of the input variables is computed. This is done by finding the intersection of the crisp input value with the input membership function.

Table 5.2: FIS rule base.

If Compactness is <i>Low</i> then Probability of Malignancy is <i>Low</i>
If Compactness is <i>Medium</i> then Probability of Malignancy is <i>Medium</i>
If Compactness is <i>High</i> then Probability of Malignancy is <i>High</i>
If Dispersion is <i>Low</i> then Probability of Malignancy is <i>Low</i>
If Dispersion is <i>Medium</i> then Probability of Malignancy is <i>Medium</i>
If Dispersion is <i>High</i> then Probability of Malignancy is <i>High</i>
If Eccentricity is <i>Low</i> then Probability of Malignancy is <i>Low</i>
If Eccentricity is <i>Medium</i> then Probability of Malignancy is <i>Medium</i>
If Eccentricity is <i>High</i> then Probability of Malignancy is <i>High</i>
If Solidity is <i>Low</i> then Probability of Malignancy is <i>Low</i>
If Solidity is <i>Medium</i> then Probability of Malignancy is <i>Medium</i>
If Solidity is <i>High</i> then Probability of Malignancy is <i>High</i>
If Moment 1 is <i>Low</i> then Probability of Malignancy is <i>Low</i>
If Moment 1 is <i>Medium</i> then Probability of Malignancy is <i>Medium</i>
If Moment 1 is <i>High</i> then Probability of Malignancy is <i>High</i>
If Moment 2 is <i>Low</i> then Probability of Malignancy is <i>Low</i>
If Moment 2 is <i>Medium</i> then Probability of Malignancy is <i>Medium</i>
If Moment 2 is <i>High</i> then Probability of Malignancy is <i>High</i>
If Moment 3 is <i>Low</i> then Probability of Malignancy is <i>Low</i>
If Moment 3 is <i>Medium</i> then Probability of Malignancy is <i>Medium</i>
If Moment 3 is <i>High</i> then Probability of Malignancy is <i>High</i>

5.2.2.2.2 Application of Fuzzy Operators

Once the inputs are fuzzified, the degree of support for each rule is computed by applying fuzzy operators to combine all parts of its antecedent and generate a single crisp value. This step is performed when there is more than one antecedent in the rule. However, the fuzzy rules defined for this system has only one antecedent. Thus, fuzzy operators were not applied in this situation.

5.2.2.2.3 Implication

The values obtained from the fuzzification indicate the firing strength of the rule. The implication method is defined as the shaping of the output membership function on the basis of the firing strength of the rule. The input for the implication process is a single

number given by the antecedent and the output is a fuzzy set. Implication is implemented for each rule. For this FIS, the minimum function is used as the implication method. It involves cutting the consequent membership function at the level of the antecedent truth. This method is preferred because it involves less computational complexity, and generates an aggregated output surface that is easier to defuzzify.

5.2.2.2.4 Aggregation of Outputs

In this stage, all implicated fuzzy sets that represent the outputs of each rule are combined to generate a single fuzzy set (*Matlab Fuzzy Logic Toolbox Documentation, 2001*). The aggregation method is commutative, thus, the order in which the rules are executed is unimportant. The aggregation is performed by combining the sum of each rule’s output set. Figure 5.13 shows the aggregation for the set of crisp input values.

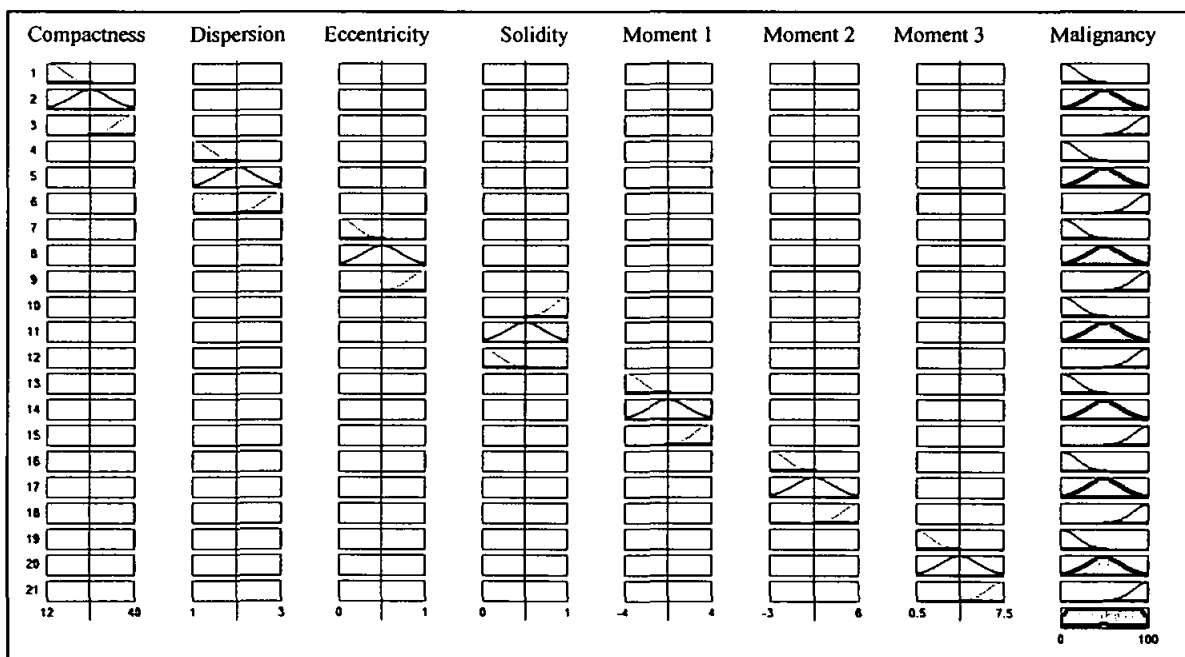


Figure 5.13: Aggregation for the set of crisp inputs.

5.2.2.2.5 Defuzzification

The output needed from the FIS is a single crisp value. The input for the defuzzification process is the fuzzy set and the output would be a single number. The aggregate of a fuzzy set encompasses a range of output values, and so it must be defuzzified in order to resolve a single output value from the set.

The centroid defuzzification method is employed for this system. This is the most commonly used method in FIS. In this method, the defuzzifier determines the center of gravity (centroid) and uses that value as the output of the FIS. This is computed as given by the equation below

$$y' = \frac{\sum_{i=1}^n y_i \mu_B(y_i)}{\sum_{i=1}^n \mu_B(y_i)} \quad (5.25)$$

where y' is the centroid, μ_B is the membership in a fuzzy set B at the value y_i .

This method finds the balance point of the solution fuzzy region by calculating the weighted mean of the output fuzzy region. When this method is used, the defuzzified values tend to move smoothly around the output fuzzy region (Kulkarni, 2001). The defuzzified value would be the malignancy factor.

Summary

The lung nodule detection and diagnosis tool is used to detect and determine the severity of lung cancer. Lung nodules are the major indicators of lung cancer in chest radiology. Through the use of MGTH transform, the visibility of the lung nodules in the chest x-ray could be greatly increased. This is followed by the application of seed-based region growing in order to obtain the shape of the nodules clearly. The diagnosis of the nodules involves the determination of the severity of the lung cancer. FIS is used to determine the malignancy based on shape characteristics of the nodules. The result is in the form of a Malignancy Factor which indicates the level of malignancy of the nodules in the form of percentage. This is a semi-automated tool where the physicians and radiologists have some control on the output of the system. The use of FIS technique in the development of this tool increases the accuracy and reliability of the system on the whole.

CHAPTER 6

DEVELOPMENT OF AN EXPERT SYSTEM WITH AN EMBEDDED IMAGING MODULE

Introduction

This chapter details the development of the expert system with an embedded imaging tool. The expert system is developed using EXSYS CORVID. It is a windows-based knowledge automation expert system development tool. The expert system developed in this work is named as LUNEX. The main performance criterion for this system is that its diagnoses should be consistent and similar with those of an expert doctor. The imaging module, CIPT is embedded into the expert system.

6.1 Development of Expert System

The development of a computer system can be viewed as a life cycle that begins with a problem and ends in a solution. In developing the expert system, the knowledge developer works with experts to identify the problem domain, capture the expertise that provide the solution and plans the development strategy and process from beginning to end. The first phase is capturing knowledge from an expert, someone with proven expertise in the particular field. The knowledge is then represented in the form of rules which forms the system's knowledge base. The final step would be the testing and implementation of the system. Several approaches are available for the development of expert systems. Some approaches are based on conventional system development life cycle while some are unique to software engineering. In this research work, the Prototyping Model (Pressman, 2001) is utilized for the development of this system. This model is commonly used in the development of conventional computer systems. The model is shown in Figure 6.1.

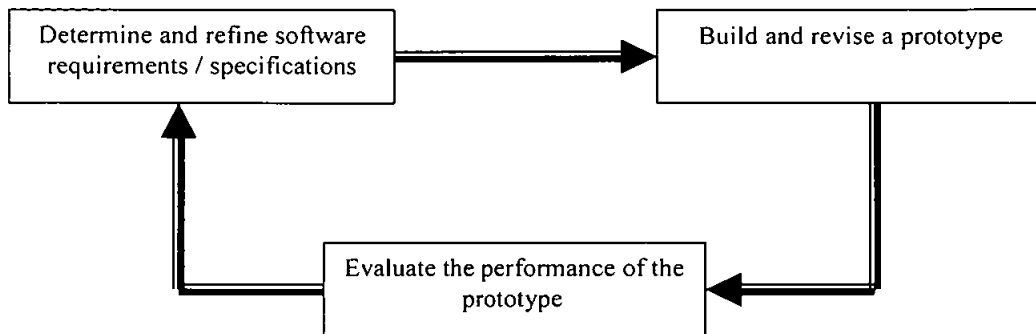


Figure 6.1: Prototyping Model.

6.1.1 Knowledge Acquisition

Expert systems derive their power from the knowledge they use. Knowledge acquisition involves eliciting, analyzing and interpreting the knowledge that a human expert uses to solve a particular problem. In expert system development, the developer captures heuristic knowledge from the experts in order to build the knowledge base that emulates the experts' thought process. Knowledge acquisition deals with choosing the expert, tapping the expert's knowledge, re-tapping the knowledge for verification and validation of the rules and updating and modifying the knowledge base after the system is operational. Thus, it is a continuously evolving step. Much of the experts' knowledge tends to be qualitative rather than quantitative.

The important steps in the knowledge acquisition process are listed below (Awad, 2003).

- Using an appropriate tool to elicit information from the experts.
- Interpreting the information and inferring the expert's underlying knowledge and reasoning process.
- Using the interpretation to build rules that represent the expert's thought processes or solutions.

The knowledge is acquired through interviews, brainstorming sessions, observation and from publications such as books, journals and websites. For the LUNEX system, the medical knowledge was acquired through 3 main sources as listed below.

- Interviews with chest specialists, radiologists and general physicians
- Medical books and journals
- Reputable medical websites

6.1.2 Knowledge Representation

Knowledge representation is the single most critical phase in building an expert system. The main goal in knowledge representation is the accuracy of the information represented. The problem that might arise in achieving accuracy is the lack of alignment between a computer's operation and the way the human mind thinks. Human judgment is mainly based on perception, cognition, memory, recollections and heuristics. Thus, factors affecting expert judgment are influenced by beliefs and assumptions which make it difficult to ensure accuracy in knowledge representation in the same way of an arithmetic or mathematical calculation. There are many forms of knowledge representation schemes for expert systems such as rules, semantic networks, frames, formal logic, decision trees and etc. In this work, the rule method has been chosen as the preferred representation scheme. Rules are the most popular representation of knowledge in expert systems. A rule-based representation is selected because of the widespread availability of rule-based expert systems shells, flexibility in modifying and updating the system and rule-based shells cost less than other shells which utilize alternative modes of representation (Ignizio, 1991).

6.1.2.1 Rules

Rules are referred to as *If-Then* or production rules. In some instances, rules are extended to include *If-Then-Else*. However, in this expert system, only rules in the form of *If-Then*

were used. The designation of *If-Then* rules is that of condition-action or premise-conclusion statement. There is a distinct difference between a rule premise and the conclusion. In the rule premise, the value in the statement is compared with any value provided. However, in the conclusion, a value is assigned to an attribute. There may be several premises and conclusions within a single rule. Each of these are termed clauses (i.e., premise clauses and conclusion clauses). The premises might be connected by AND as well as OR operators, while the conclusion clauses can only be connected by AND statements. This denotes that all conclusions in a rule must be true. Example of a rule used in the LUNEX system is shown in Figure 6.2.

```
If fever present is high
AND chest pain is severe
AND on physical examination – blood pressure is high
AND on clinical ABG test – oxygen level is low
AND on radiological information – heart silhouette is enlarged
Then it is Pulmonary Hypertension
```

Figure 6.2: Example of LUNEX rule.

The attribute values may be in the form of symbolic or numeric. The rules used in this system are both represented both symbolically and numerically depending on the nature of the attributes. For example age is represented numerically while chest pain is represented symbolically by *slight, moderate* or *severe*.

Knowledge is a continuously evolving entity. Knowledge addition and modification is a very important attribute in an expert system. LUNEX has the capability to support the addition of new rules as well as the modification of existing rules. Redundant or

inaccurate rules can also be removed. This can be performed without affecting the processing capability and performance of the system.

6.1.2.2 Certainty Factor

In knowledge representation, there might always be uncertainty in the conclusions (i.e., the conclusion attribute values assigned). There are two sources of uncertainty in an expert system which is uncertainty with regard to the validity of the rules and uncertainty with regard to the validity of a user response. In the LUNEX system, the main source of uncertainty is due to the response of the user of the system, specifically the answers provided in response to the question posed by the system.

In medicine, most medical concepts are imprecise. Rules in medicine include words like 'severe pain' or 'moderate pain' that are difficult to formalize and measure. For example, if the expert system question is:

Is the person having chest pain?

A *yes* or *no* would not be sufficient. In this case, giving a number of possible choices of pain level such as *minimal*, *moderate* or *severe*, where the user needs to select the most appropriate response, is a much better approach. Thus, in the expert system, a number of choices are provided to the user. Each of the choices is associated with a confidence factor. A confidence factor (C.F) is a numerical probability value assigned to a choice and indicates the subjective estimates of the relative level of confidence of the selected choice.

LUNEX system has a range of confidence factors which are assigned to the choices in each rule in the knowledge base. The confidence factors are based on five categories indicating varying degrees of successes that could be expected when consultation takes place (Ngah et al., 1996). The system utilizes 0-10 confidence factor system which is

permitted by the EXSYS CORVID shell. The final conclusion would be in the form of a disease or a list of diseases with their corresponding certainty factors. The confidence factors for each true rule is summed together to form the final confidence factor. The diseases with the highest confidence factor indicate the highest possibility of occurring. The confidence factor categories are given in Table 6.1.

Table 6.1: Confidence factor categories.

Definitely Yes	–	10
Most Likely	–	7
Likely	–	5
Not Likely	–	3
Definitely No	–	0

The choices for the questions posed by the expert system may vary depending on the nature of the attributes addressed by the question. However, the confidence factor categories listed above is maintained for most of the questions with the exception of few. The exemptions in the case of clinical test results as this questions require a specific numerical answer without any level of uncertainty.

For example, in the case of blood test results, the question might be in the form

What is the red blood cell count?

The user has to enter the numerical blood count. In this case, if the cell count is within a specific range, it is positive for a particular disease. In such situations, the confidence factor assigned would be as shown in Table 6.2.

Table 6.2: Confidence factor categories for clinical test results.

Definitely Yes	–	10
Definitely No	–	0

Where, 'definitely yes' indicates that the data is within range and is positive for a disease while 'definitely no' indicates the data does not coincide with the range associated with the disease.

6.1.3 Inference

During system operation, the status of each rule is expected to change. Keeping track of such changes is an essential part of the inference process.

- A rule is true whenever the premise is satisfied.
- A rule is false whenever the premise is not satisfied.
- If the rule is true, then the rule is said to be triggered.
- If the rule is false, the rule remains inactive.

The information received is compared with the premise part of the rule to determine the truth or falsity of the premise. If the system determines that the premise is true, then the rule executes or fires.

The EXSYS CORVID expert system shell allows two main inference strategies which are the forward chaining and backward chaining. In this work, the forward chaining method is employed. This method is used because it is quite suitable for diagnostic applications where there is a set of data to begin with.

Forward chaining is data-driven where it begins with a set of known data to come up with a conclusion. This method can also provide explanation for any conclusions in terms of the rules that was used to deduce it. In forward chaining the main focus is on the premise. The mechanism chains forward from the premise of one rule, looking for the premise of another rule that matches the conclusion part of the first rule. In the next step, the process proceeds to the premise of the third rule, which matches the conclusion part of the second rule, and so on until a conclusion is reached. The rules are tested in the order they occur.

If a variable is needed in order to determine if a rule is true or false and the value of the variable is not known, the variable is asked for the user. If a rule is found to be false, it is discarded. The system actually executes the conclusion statements for each rule whose premise is true (Awad, 2003).

6.1.4 Development of Knowledge Base

The diagnosis of lung diseases requires a thorough analysis of a large amount of medical information. The information encoded in the knowledge base of LUNEX system is classified into four categories as listed below.

- Patient's personal and medical background
- Clinical symptoms
- Clinical test results
- Radiological findings

6.1.4.1 Patient's Personal and Medical Background

A patient's personal and medical background is very important in assessing the risk factors involved for the patient in contracting a particular disease. The personal and medical background involves medical history of family, patient's medical history, habits and hobbies and common activities. Table 6.3 lists the risks factors common for lung diseases.

Table 6.3: Risks factors.

Age	Infections
Gender	Choking/Inhaling foreign objects
History of lung diseases	Obesity
History of other diseases	Exposure to nature
Family history of disease	Activities
Prolonged exposure to smoke/fumes	Drug consumption
Smoking	Malnutrition
Exposure to chemicals	Exposure to people with diseases
Trauma/Injury	Diseases triggered by some elements
Organ Failure	Allergies

A sample rule of patient’s personal and medical background is given in Figure 6.3.

If bacterial or viral infection *Most Likely*: C.F = 7
 And prolonged exposure to smoke is *Likely*: CF = 5
 And smoking for a long period of time *Definitely Yes*: C.F = 10
 And patient has respiratory infections *Most Likely*: C.F = 7
Then Chronic Bronchitis: C.F = 29

Figure 6.3: Sample rule on patient’s personal and medical history.

The final conclusion’s confidence factor is the summation of all the confidence factors of the rule premises which are true.

6.1.4.2 Clinical Symptoms

Clinical symptoms are symptoms that appear at the onset of a disease. The symptoms are experienced by a patient. Clinical symptoms are extremely crucial in diagnosing diseases. The symptoms are either told by the patient or discovered by the doctor during the medical examination. Table 6.4 below lists the common symptoms for lung diseases.

Table 6.4: List of Clinical Symptoms.

Coughing	Sore throat
Coughing blood	Chest retraction
Dyspnea	Pain/aches
Wheezing	Stuffy nose
Weight Loss	Sneezing
Phlegm expectoration	Nasal discharge
Fatigue	Dizziness
Cyanosis	Sweating
Clubbing	Abdominal problems
Bad breath	Anxiety
Chest pain	Chills
Headaches	Itchiness
Rapid heart beat	Apetite loss
Leg and ankle swelling	Sleeping problems
Runny nose	Eye problems
Fainting	Skin problems
Chest tightness	Itchiness
Facial swelling	Appetite loss
Nose bleeds	Sleeping problems
Easily bruise	Eye problems
Speaking difficulties	Skin problems
Poor coordination	Enlarged veins
Disease trigger effect	Enlarged lymph nodes

A sample rule of clinical symptoms is given in Figure 6.4.

If Dyspnea is *Definitely Yes*: C.F = 10
 And wheezing is *Most Likely*: CF = 7
 And Cyanosis is *Not Likely*: C.F = 3
 And weight loss is *Most Likely*: C.F = 7
 And Anxiety is *Likely*: C.F = 5
Then Emphysema: C.F = 30

Figure 6.4: Sample rule on clinical symptoms.

6.1.4.3 Clinical Test Results

Clinical tests are performed in the hospital or clinic laboratory and the results are used for diagnosis purposes. Some of the tests are specific for lung diseases while some are routine medical examinations. Table 6.5 below shows the lists of medical tests commonly performed when diagnosing lung problems. A sample rule of clinical test results is given in Figure 6.5.

If Fever is $\geq 38.30C$: C.F = 10
 And ABG oxygen level < 88 mm Hg: CF = 10
 And ABG carbon dioxide level < 34 mm Hg: C.F = 10
Then Pneumonia: C.F = 30

Figure 6.5: Sample rule on clinical test results (Farzan, et al. 1997).

Table 6.5: List of Clinical Tests.

Tests	Sub-tests	Tests	Sub-tests
Auscultation	Bronchial sound	Serum	CPK
	Rales		Calcium
	Ronchi		Amylase
	Friction rub		LDH
	Egophony sound	Sputum	Color
	Breath sound		Culture
	Vocal fremitus		PFT
Blood Pressure	-	RV	
Percussion	-	FEV1	
ABG	O ₂	FVC	
	CO ₂	FEV1/FVC	
CBC	White blood cell count	Other tests	Sweat test
	Red blood cell count		Skin test
	ESR		Blood culture
	Eosinophilic		

6.1.4.4 Radiological Findings

Radiological findings deal with the chest X-Ray analysis. Chest X-Ray is a major indicator for many lung diseases. Normally, abnormalities in the chest X-Ray may point to a particular disease. Table 6.6 below lists the radiological abnormalities that are normally found in people suffering from lung diseases.

Table 6.6 : Chest X-Ray abnormalities.

Hyper-radiolucency	Consolidation
Parenchymal infiltration	Cavitation
Homogenous density	Lobe/segment retraction
Patches of density	Patchy parenchymal density
Diffuse reticular or nodular density	Miliary pattern
Radiolucent area	Fungus ball
Parenchymal lesion	Abscess formation
Cardiac silhouette	Air build up
Hilum enlargement	Fluid build up
Hilum displacement	Atelectasis
Flattening of diaphragm	Volume loss
Elevation of diaphragm	AP chest diameter increase
Mediastinal lymph node	Hyperinflation
Mediastinal shift	Tramline shadows
Bronchovascular marking	Transient shadows
Lung markings	Sternum space
Honeycombing	Mucoid impaction
Hazy lung	Migratory infiltrates
Calcifications	Bilateral diffuse mottling
Lesion/mass	Lung fissures
Nodular shadows	Prominent Artery
Bullous lesions	Lung detachment
Pleural effusion	Heart deviation
Pleural thickening	Malignancy factor
Fibrosis	

A sample rule for radiological findings is given in Figure 6.6.

<p><i>If</i> parenchymal lesion in lung periphery is <i>Most Likely</i>: C.F = 7 And hilar enlargement is <i>Most Likely</i>: CF = 7 And patchy airspace densities is <i>Not Likely</i>: C.F = 3 And cavitation inside parenchymal density is <i>Definitely Yes</i>: C.F = 10 And nodular densities is <i>Likely</i>: C.F = 5 And military pattern is <i>Definitely No</i>: C.F = 0 <i>Then</i> Pneumonia: C.F = 32</p>

Figure 6.6: Sample rule for radiological findings.

6.1.4.5 List of Diseases

The LUNEX system is designed to diagnose 60 types of lung diseases. These include both diseases of the lung and diseases affecting the respiratory systems on the whole. Table 6.7 lists all the diseases that can be diagnosed by this system.

6.1.4.6 Decision Table

A decision table is a technique used to represent data in a tabular form, prior to actually creating a rule base. The tabular representation of the decision table is characterized by the separation between premises and conclusions. The table column indicates the premises and the row indicates the conclusions (Preece, 1993). For this work, four decision tables were created which are patient's personal and medical history, clinical symptoms, clinical test results and radiological findings. The column of the decision table indicates the medical conditions and the rows indicate the list of diseases. The condition that coincides with a disease is marked. Appendix A contains the decision tables for the four categories mentioned above.

Table 6.7: List of Lung Diseases.

Acute Bronchitis	Influenza
Adult Respiratory Distress Syndrome (ARDS)	Kaposi's Sarcoma
Allergic Bronchopulmonary Aspergillosis	Khyphoscoliosis
Alpha-1 related Emphysema	Laryngitis
Asbestosis	Lung Absess
Aspergilloma	Lung Cancer
Aspiration Pneumonitis	Lymphangioliomyomatosis
Asthma	Lymphatic Interstitial Pneumonitis
Atelectasis	Mesothelioma
Blastomycosis	Mountain Sickness
Bronchiectasis	Pertussis
Bronchiolitis	Pleural Effusion
Bronchitis Obliterans	Pleurisy
Chronic Bronchitis	Pneumoconiosis
Chronic Cor Pulmonale	Pneumonia
Chronic Necrotizing Pulmonary Aspergillosis	Pneumonitis
Croup	Pneumothorax
Coccidioidomycosis	Primary Cilia Dyskinesia
Common Cold	Progressive Systemic Sclerosis
Cystic Fibrosis	Pulmonary Edema
Diphtheria	Pulmonary Embolism
Emphysema	Pulmonary Fat Embolism
Eosinophilic Pneumonias	Pulmonary Hypertension
Gaucher Disease	Pulmonary Parenchymal Injury
Goodpasture Syndrome	Respiratory Diphtheria
Hantavirus Pulmonary Syndrome	Sarcoidosis
Hayfever	Silicosis
Hemothorax	Systemic Lupus Erythematosis
Histoplasmosis	Tuberculosis
Idiopathic Pulmonary Fibrosis	Wegener's Granulomatosis

6.2 Expert System with an Embedded Imaging Module

The imaging module embedded expert system has two major components as listed below.

- Expert system
- Imaging Module

The structure of the system is shown in Figure 6.7.

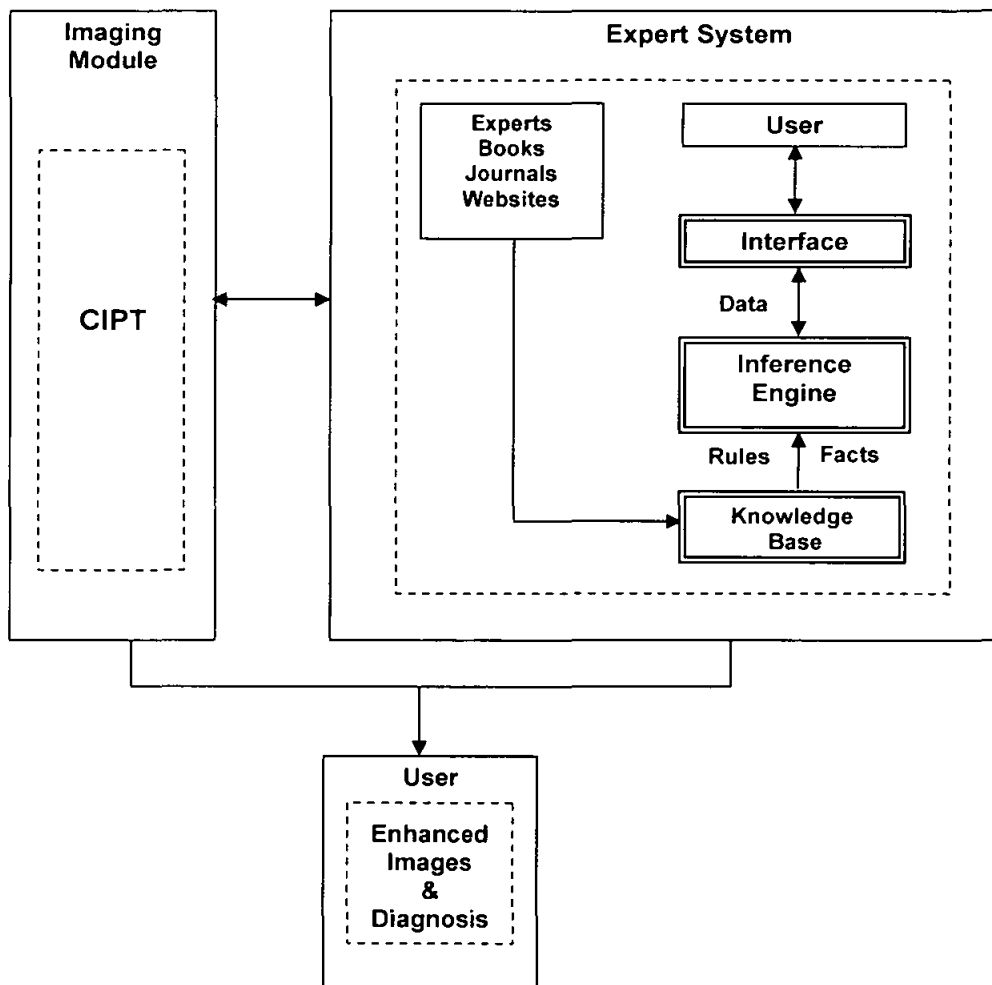


Figure 6.7: Structure of the expert system with an embedded imaging module.

6.2.1 Imaging Module Functions

The imaging module, CIPT contains two main parts which are as below.

- Imaging Functions
- Lung Nodule Detection and Diagnosis Tool

The formulation of the imaging functions was described in chapter 4. The development of the lung nodule detection and diagnosis tool was described in chapter 5. The imaging module is standalone windows-based software. The imaging functions are general function that is used to enhance and analyze the chest X-Ray images. The lung nodule detection and diagnosis tool is specifically for the diagnosis of Lung Cancer. Figure 6.8 shows the functions available in CIPT.

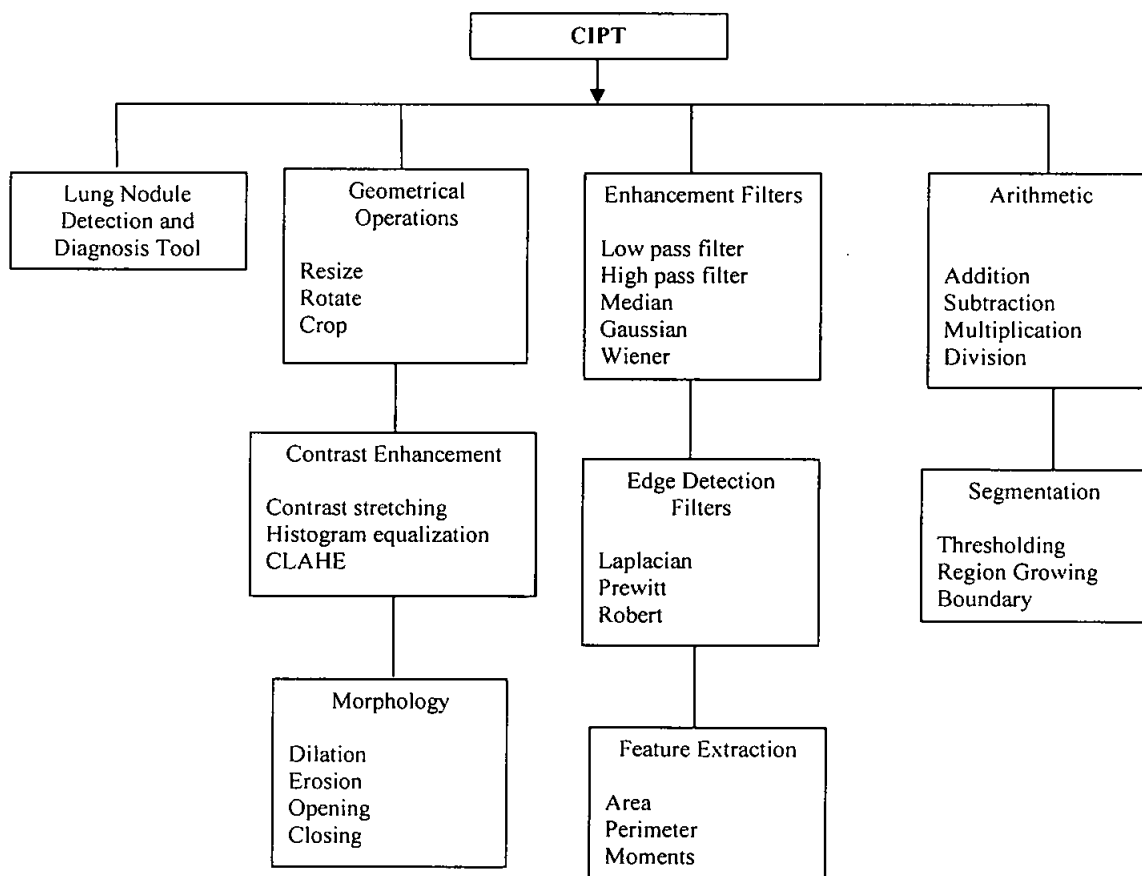


Figure 6.8: Functions in CIPT.

6.2.2 System Operation

LUNEX is a windows-based software, thereby, it its user friendly and can run in any window-based computer. The system is standalone software and it does not require any expert system shell to run. The process chart of the system is given in Figure 6.9.

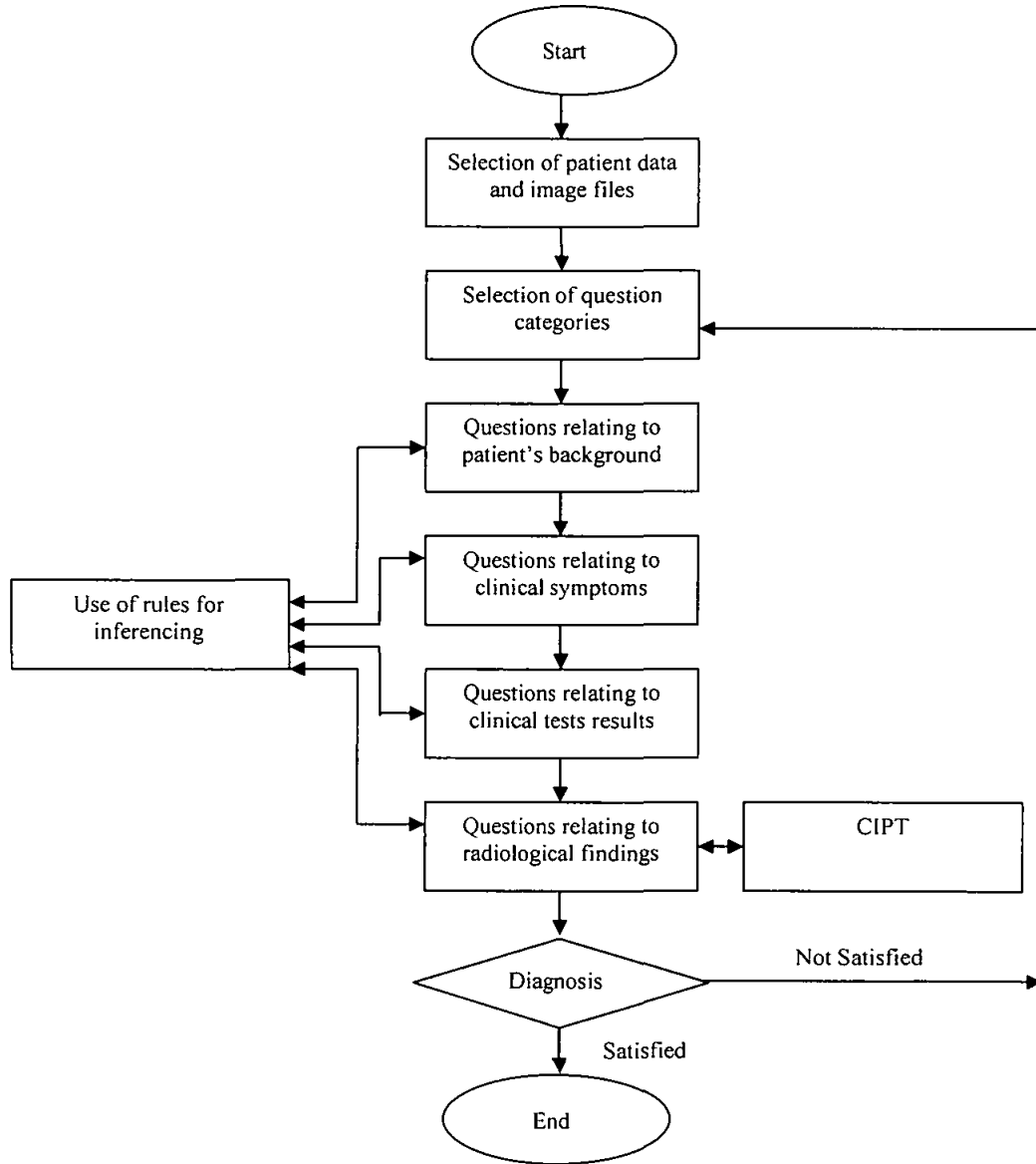


Figure 6.9: Process Flow-Chart of LUNEX.

The system will first prompt the user to select the category of question the user wants to answer. There are four categories that the user can choose from which are as below.

- Patient's personal and previous medical history
- Clinical symptoms
- Clinical test results
- Radiological findings

The system is divided into categories as most often, not all categories of medical data might be available to the doctor, but still, a diagnosis is needed to be made. The system enables the user to perform an initial diagnosis with a small amount of data or a thorough diagnosis with large amount of data. The user can select one or more categories of question. All four categories can also be selected. If more than one category is chosen, the system will ask question relating to one category and then will proceed to the next category.

The system will ask the user a series of questions relating to the category the user selected and these questions would be displayed on screen. The types of questions asked are of multiple choices. A question will display a statement with a list of choices. The user has to click on the radio button next to the preferred choice and click the OK button. For some question, the user has to enter a numerical input. This is done by keyboard input. If the user does not want to answer a question, he or she can click on the Not Applicable radio button. The user can go back and change the answer to any question by clicking on the BACK button. The RESTART button enables the user to restart the system again.

When the user is answering radiological questions, he or she can click on a X-RAY icon on the window of the expert system screen. Clicking this icon will open the CIPT in a child window. This CIPT is integrated into the expert system and can be accessed through a simple mouse click. The user can use this tool to enhance the X-Ray image, in

order to answer the questions more accurately. Thus, the user does not have to open two separate softwares.

A user is supposed to answer all the questions from the category that he or she has selected. The system will continue asking the questions until it has reached a conclusion. When the user has answered the last question, the system will process all the answers given by the user and arrive at a conclusion.

The conclusion may be the selection of a single disease or a list of diseases according to their confidence factors. The disease with the highest confidence factor indicates the most probable disease; disease with the second highest confidence factor indicated the second most probable disease and so on. In this way, the doctor can just narrow down his diagnosis to a few diseases and make an accurate decision.

The explanation facility can provide reason of how and why it has reached a specific diagnosis. The explanation facility can be enabled or disabled according to the preference of the user.

6.2.3 System Interface

The LUNEX system has a windows-based interface. When the user opens the system, a title screen would be displayed. Figure 6.10 shows the title screen of the system. Clicking the OK button will take the user to the next screen. The user will be then asked to select the category of medical questions he or she would like to answer. This is shown in Figure 6.11.

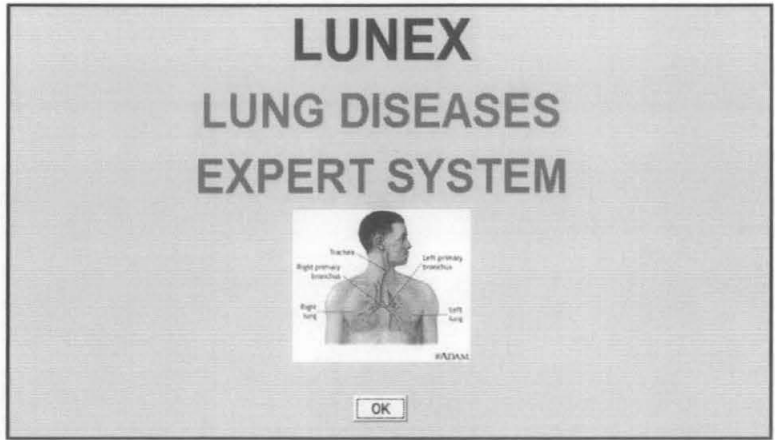


Figure 6.10: LUNEX title screen.

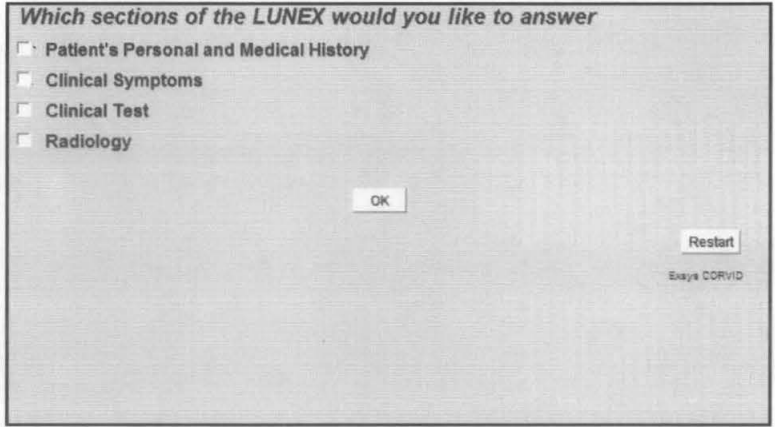


Figure 6.11: Medical questions categories.

Depending on the category selected, the system will start posing the questions. If more than one category was selected, questions from one category will be asked before displaying questions of the next category. Figure 6.12 shows a sample screen for patient's personal and medical history. Figure 6.13 shows a sample screen for clinical symptoms. Figure 6.14 shows a sample screen for clinical tests. Figure 6.15 shows a sample screen for radiology.

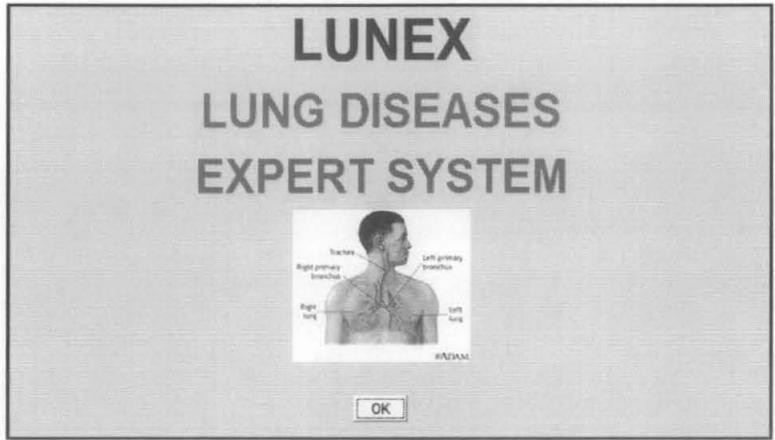


Figure 6.10: LUNEX title screen.

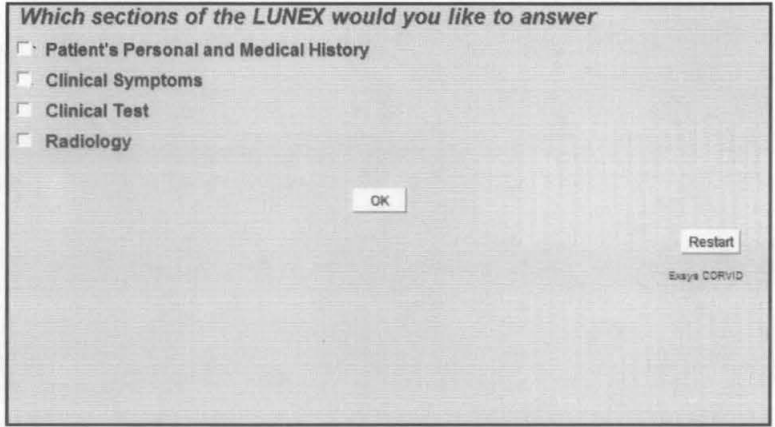


Figure 6.11: Medical questions categories.

Depending on the category selected, the system will start posing the questions. If more than one category was selected, questions from one category will be asked before displaying questions of the next category. Figure 6.12 shows a sample screen for patient's personal and medical history. Figure 6.13 shows a sample screen for clinical symptoms. Figure 6.14 shows a sample screen for clinical tests. Figure 6.15 shows a sample screen for radiology.

How long has the patient been smoking

- Less than a year
- 1 to 5 years
- 5 to 10 years
- 10 years or more
- Not applicable

Exsys CORVID

PERSONAL & MEDICAL HISTORY




Figure 6.12: Question screen on patient's history.

Does the patient have a stuffy or blocked nose?

- Definitely Yes
- Most Likely
- Likely
- Not Likely
- Definitely No
- Not Applicable

Exsys CORVID

CLINICAL SYMPTOMS




Figure 6.13: Question screen on clinical symptoms.

Was an Arterial Blood Gas (ABG) test done?

- Yes
- No

Please enter the oxygen level of the patient in units of mm Hg

Please enter the Carbon Dioxide level of the patient in units of mm Hg

Exsys CORVID

CLINICAL TEST




Figure 6.14: Question screen on clinical tests.

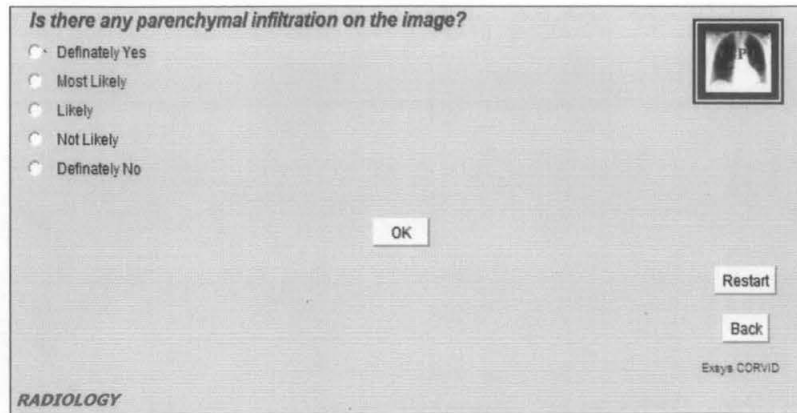


Figure 6.15: Question screen on radiology.

Figure 6.16 shows a sample of user interface of CIPT imaging module. This module can be opened by clicking on the X-Ray icon on the expert system question screens. Figure 6.17 shows the expert system with the CIPT interface.

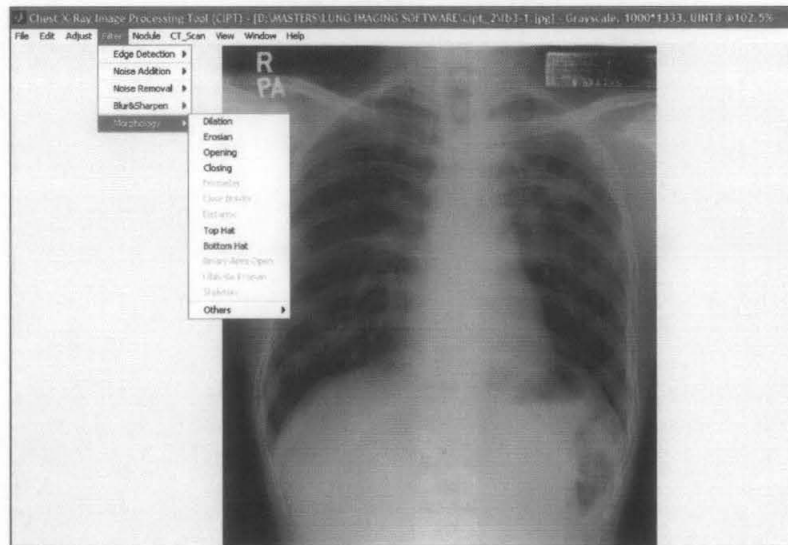


Figure 6.16: CIPT user interface.

CHAPTER 7

RESULTS AND DISCUSSION

Introduction

The expert system with an embedded imaging module developed in this work was applied on many test cases using medical data and X-Ray images obtained from hospitals. The various image processing functions developed under the CIPT were applied on X-Ray images. The sample X-Rays collected from hospitals spans a variety of lung diseases. Lung nodule detection and diagnosis part of the CIPT is applied on the suspected regions of the digitized X-Ray images. The results are then fed into the expert system for deciding the probability of occurrence of the lung diseases. The LUNEX system uses four categories of data namely, i) patients' personal and medical information, ii) clinical symptoms, iii) clinical test results and iv) radiological findings using varieties of chest X-Rays for the diagnoses.

The chest X-Ray images used in this work were obtained from hospitals in Ipoh. The set of test images used were 8 bit grayscale images. The spatial resolution used is 1000 x 1333. This size is selected in accordance to the computational power of the computer, computational time and image aspect ratio. The actual names of patients and medical reports are not included due to confidentiality. All the results presented in this chapter have been verified by specialist doctors.

One very important factor to be taken into consideration is that, the diagnoses made by the system are very much dependent on the information provided by the user. Two different physicians might provide two quite different set of responses, even when dealing with the same patient and medical data set. Thus, there may be always a small amount of uncertainty in the final results. Apart from that, a general physician may miss some of the details or abnormalities in interpreting the digitized X-Ray images.

7.1 Diagnoses of Benign Diseases by LUNEX

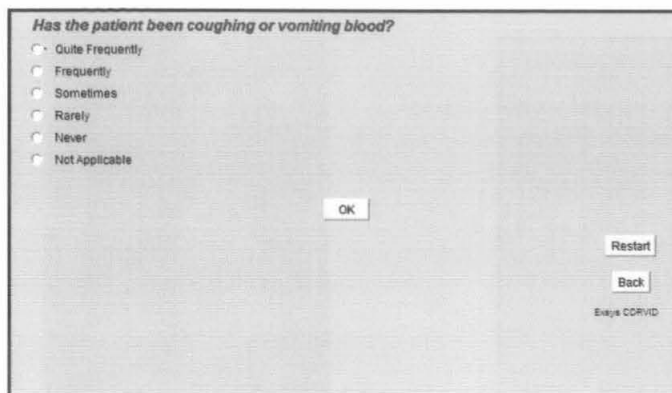
LUNEX can diagnose both benign and malignant diseases. This system was tested with 60 patient test cases obtained mostly from the hospitals and few from the internet. Since, the expert system can diagnose with very little information, it is not necessary to have all four categories of data to perform the diagnosis. The system can perform the diagnosis with just information from one category. However, more accurate diagnosis can be obtained by providing as much information as possible to LUNEX.

The user can select one or more categories of medical data and then answer the questions posed by LUNEX. When all the questions have been answered, the system will come to a conclusion which will be in the form of probabilities of occurrences of diseases and a list of such diseases with their corresponding confidence factors will be displayed. The disease with the highest confidence factor is the most probable disease and this will be followed by the second most probable and so on.

The results of the system are very much dependent on the answer provided by the doctor. Thus, two doctors working on the data of the same patient might obtain different results. The system has been tested by general physicians and the results have been verified by specialist doctors. About 88% of the results from LUNEX match with the diagnoses given by the specialist doctors. A few case studies are presented and discussed. The first three case studies are presented in detail while the next few cases are given in a brief manner. The rest of the results are given in a table form in Appendix B.

7.1.1 Patient A: Case Study

The LUNEX was tested using the medical data relating to patient A. The X-Ray image was not very clear for a general physician's diagnosis. Figure 7.1(a) shows a sample screen of a question asked. Figure 7.1(b) shows the original X-Ray image.



Has the patient been coughing or vomiting blood?

- Quite Frequently
- Frequently
- Sometimes
- Rarely
- Never
- Not Applicable

OK

Restart

Back

Exiya COVID

Figure 7.1(a):
Question screen.



Figure 7.1(b):
Original X-Ray of A.

Image processing techniques was applied to the original chest X-Ray in order to enhance the image and to detect any abnormalities. In the Figure 7.1(b), MGTH transform was used to improve the image. The output image is shown in Figure 7.1(c). Figure 7.1(d) shows the display of results on the screen by LUNEX.

LUNEX findings:

Here, the system concludes that Tuberculosis is the major disease with a confidence factor of 81 and in addition, it displays a list of other probable diseases. The probable diseases are listed according to descending confidence factors. The disease with the highest confidence factor is the most likely disease to occur. In the Figure 7.1(d), it is found that Tuberculosis is the possible disease. Acute Bronchitis is the next possible disease as it has the second highest confidence factor. Thus, the doctor will just have to focus on the first few diseases only. Diseases with very small confidence factors have a very minimal chance of occurring.

Image processing techniques was applied to the original chest X-Ray in order to enhance the image and to detect any abnormalities. A combination of imaging functions was used to enhance the image. Figure 7.2(c) shows the resultant image on applying contrast stretching and brightness control on the original image. Figure 7.2(d) shows the results as displayed on the screen by LUNEX.



Figure 7.2(c):
Processed image.

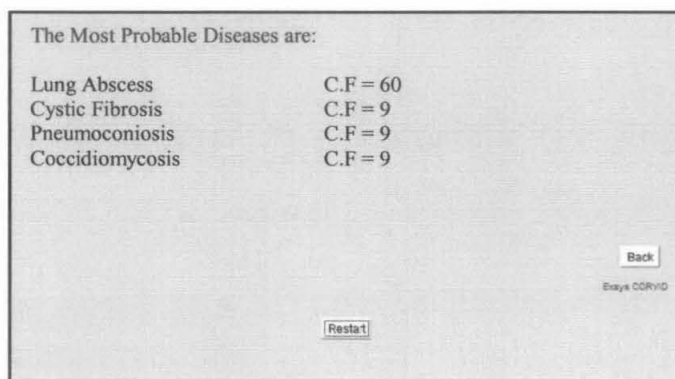


Figure 7.2(d):
Diagnosis by LUNEX.

LUNEX findings:

From the displayed results, it can be seen that Lung Abscess has the highest confidence factor and thus, Lung Abscess is the most likely disease. Cystic Fibrosis is the next possible disease.

Doctor's comments:

In the processed image, a large area of radiolucency containing fluid was observed very clearly. There seems to be an abscess formation in this area. An area of consolidation can also be seen. These findings are in accordance with the symptoms for Lung Abscess.

7.1.3 Patient C: Case Study

The LUNEX system was tested on patient C. Figure 7.3(a) shows an example screen of a question asked. Figure 7.3(b) shows the original X-Ray image.

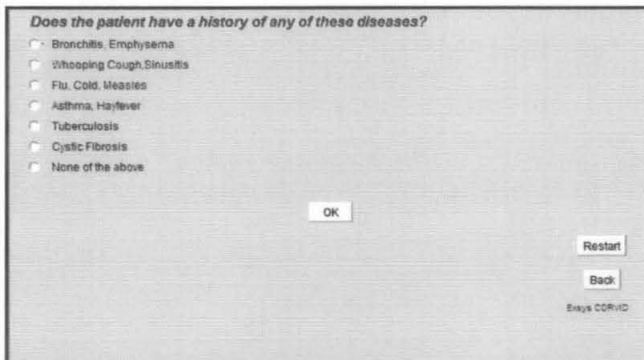


Figure 7.3(a):
Question screen.



Figure 7.3(b):
Original X-Ray of C.

This image is enhanced using a combination of contrast enhancement and editing tools. The processed image is shown in Figure 7.3(c). The results of the diagnosis are displayed in Figure 7.3(d).



Figure 7.3(c):
Processed image.

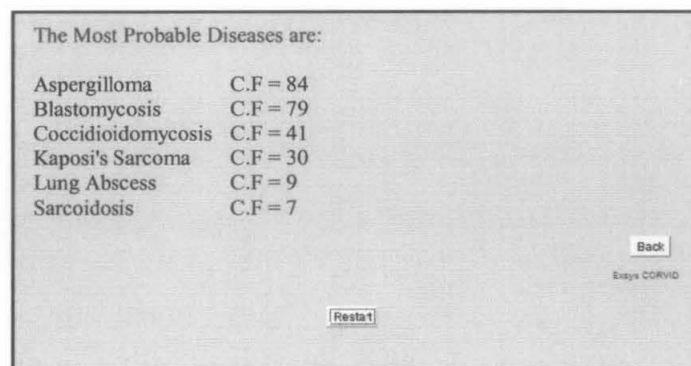


Figure 7.3(d):
Diagnosis by LUNEX.

LUNEX findings:

From the results it is found that Aspergilloma is the most probable disease.

Doctor's comments:

The processed image shows an area of localized consolidation and a hollow region of fungal ball. These abnormalities are often caused by fungal infectious diseases such as Aspergilloma and Blastomycosis.

7.1.4 Case Studies of Patients D, E, F and G

In the same way, many case studies were analyzed using LUNEX. Among them, we display the findings of LUNEX for four more case studies in a brief manner.

The system was tested using the medical data of patient D. The original X-Ray image of the patient, the processed image and the display of diagnosis by LUNEX are shown in Figures 7.4(a), 7.4(b) and 7.4(c) respectively.



Figure 7.4(a):
Original X-Ray of D.



Figure 7.4(b):
Processed image.

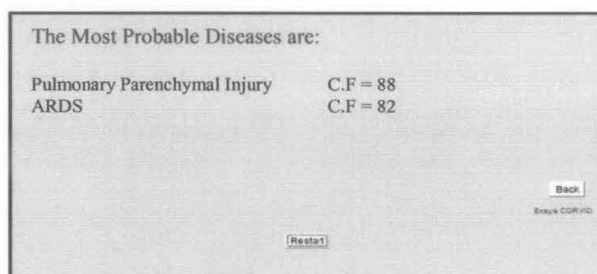


Figure 7.4(c):
Diagnosis by LUNEX.

LUNEX findings:

Most probable diseases are Pulmonary Parenchymal Injury and ARDS.

Doctor's comments:

There is an abrupt discontinuity in the smooth outline of the rib which indicates a broken rib. This indicates a rib injury or trauma. This abnormality could cause Parenchymal Injury or ARDS.

The system was tested using the medical data of patient E. The original X-Ray image of the patient, the processed image and the display of diagnosis by LUNEX are shown in Figures 7.5(a), 7.5(b) and 7.5(c) respectively.



Figure 7.5(a):
Original X-Ray of E.

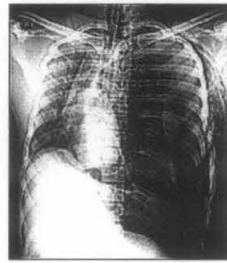


Figure 7.5(b):
Processed image.

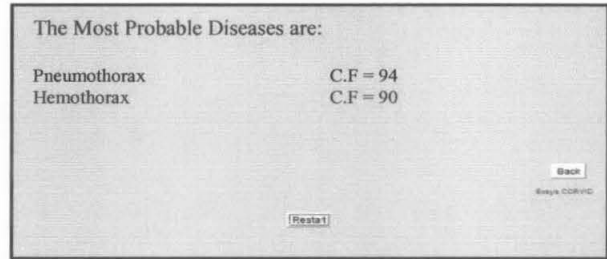


Figure 7.5(c):
Diagnosis by LUNEX.

LUNEX findings:

The most probable diseases are Pneumothorax and Hemothorax.

Doctor's comments:

The abnormalities such as the deviation of the trachea away from the side of the Pneumothorax, shift of the mediastinum and depression of the diaphragm are clearly seen in the processed image. This patient is having Pneumothorax and Hemothorax is another probable conclusion.

Pneumothorax is the progressive build-up of air within the pleural space, usually due to a lung laceration which allows air to escape into the pleural space but not to return. The Hemothorax is the build up of fluid in the lungs.

For patient F, the original X-Ray image of the patient, the processed image and the display of diagnosis by LUNEX are shown in Figures 7.6(a), 7.6(b) and 7.6(c) respectively.



Figure 7.6(a):
Original X-Ray of F.



Figure 7.6(b):
Processed image.

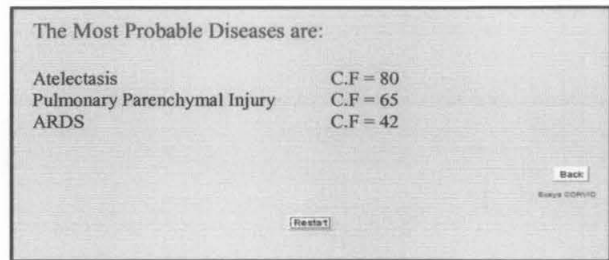


Figure 7.6(c):
Diagnosis by LUNEX.

LUNEX findings:

The most probable diseases are Atelectasis followed by Parenchymal Injury and ARDS.

Doctor's comments:

On comparison, in the processed image, fracture and callous formation on the whole rib can be observed clearly. These formations are due to rib fractures. There is retraction of lung lobe. These signs indicate Atelectasis. Rib fractures may cause ARDS and Pulmonary Parenchymal Injury.

The diagnosis for patient G is given below. The original X-Ray image of the patient, the processed image and the display of diagnosis by LUNEX are shown in Figures 7.7(a), 7.7(b) and 7.7(c) respectively.



Figure 7.7(a):
Original X-Ray of G.



Figure 7.7(b):
Processed image.

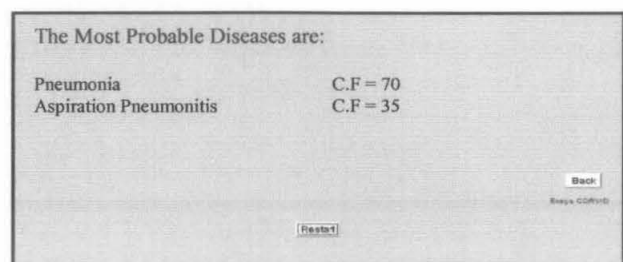


Figure 7.7(c):
Diagnosis by LUNEX.



Figure 7.6(a):
Original X-Ray of F.

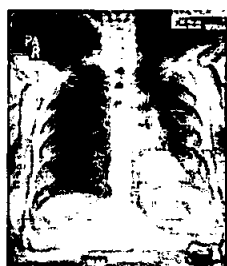


Figure 7.6(b):
Processed image.

The Most Probable Diseases are:	
Atelectasis	C.F = 80
Pulmonary Parenchymal Injury	C.F = 65
ARDS	C.F = 42

Back

[Home]

Easy CDROM

Figure 7.6(c):
Diagnosis by LUNEX.

LUNEX findings:

The most probable diseases are Atelectasis followed by Parenchymal Injury and ARDS.

Doctor's comments:

On comparison, in the processed image, fracture and callous formation on the whole rib can be observed clearly. These formations are due to rib fractures. There is retraction of lung lobe. These signs indicate Atelectasis. Rib fractures may cause ARDS and Pulmonary Parenchymal Injury.

The diagnosis for patient G is given below. The original X-Ray image of the patient, the processed image and the display of diagnosis by LUNEX are shown in Figures 7.7(a), 7.7(b) and 7.7(c) respectively.



Figure 7.7(a):
Original X-Ray of G.

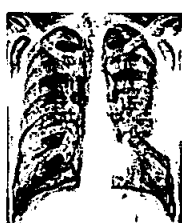


Figure 7.7(b):
Processed image.

The Most Probable Diseases are:	
Pneumonia	C.F = 70
Aspiration Pneumonitis	C.F = 35

Back

[Home]

Easy CDROM

Figure 7.7(c):
Diagnosis by LUNEX.

LUNEX findings:

Pneumonia is the most probable disease followed by Aspiration Pneumonitis.

Doctor's comments:

After processing, it can be seen clearly that there is a poorly defined abnormal opacity in the upper left lobe. Increased bronchovascular markings can be seen clearly in the processed image. These symptoms are synonymous with airspace diseases such as Pneumonia.

7.2 Diagnoses of Malignant Diseases by LUNEX

The seven cases studied above are relating to selected diseases of benign nature. In this section, the LUNEX is tested with test cases which are suspected of having diseases of malignant nature such as lung cancer. For malignant diseases, we have to apply high level image processing and analyzing techniques such as region growing and feature extraction. Since lung nodules are major indicators of malignancies, a powerful lung nodule detection and diagnosis tool has been developed and applied on the chest X-Ray images. This tool has the capability of determining the malignancy condition of lung diseases. The malignancy factor obtained from the tool is then fed into the expert system along with medical data from the four categories mentioned earlier in order to obtain the final diagnosis.

The malignancy factor introduced here is solely dependent on the shape features of the lung nodules obtained from processed images. The final diagnosis of malignant conditions, however, is determined using both the medical information obtained under the four categories considered above and shape features provided by the lung nodule detection and diagnosis tool. The first two case studies are presented in detail and the subsequent case studies are discussed briefly.

7.2.1 Patient H: Case Study

LUNEX was tested on the medical information and chest X-Ray of patient H. The original X-Ray image is shown in Figure 7.8(a). Initially, the nodules are not clear for diagnosis. The image is then processed with the MGTH function. The output is shown in Figure 7.8(b). The region grown image is shown in Figure 7.8(c).



Figure 7.8(a):
Original X-Ray of H.



Figure 7.8(b):
MGTH processed image.

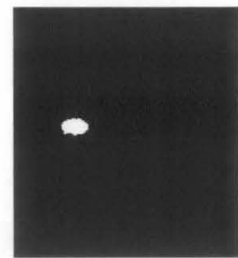


Figure 7.8(c):
Region grown image.

The malignancy factor displayed by the system is shown in Figure 7.8(d) and the value is 22.78 %. This malignancy factor is then fed into the expert system along with other available medical data. The system then arrives at a diagnosis which is displayed in Figure 7.8(e).

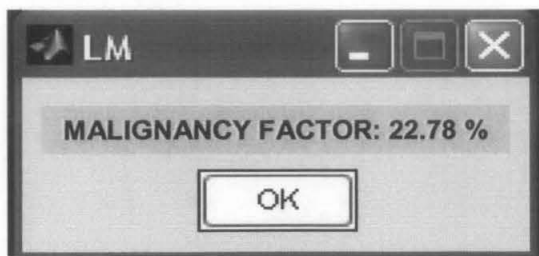


Figure 7.8(d):
Malignancy factor display.

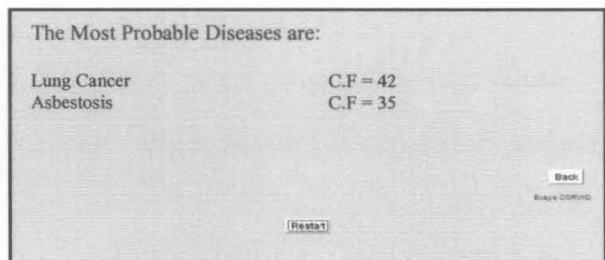


Figure 7.8(e):
Diagnosis by LUNEX.

LUNEX findings:

The system has diagnosed Lung Cancer as the most probable disease followed by Asbestosis.

Doctor's comments:

The MGTH processed image clearly shows the lung nodule. The region grown image has shown the shape of the nodule clearly. The shape of the nodule is smooth and circular indicating a benign condition. This justifies the low malignancy value. But however, the clinical data are also to be taken into account for overall malignancy.

7.2.2 Patient I: Case Study

The original image shown for the patient I in Figure 7.9(a) is processed with the MGTH function. The output of MGTH is shown in Figure 7.9(b). The region grown image is shown in Figure 7.9(c).

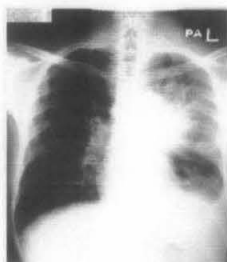


Figure 7.9(a):
Original X-Ray of I.



Figure 7.9(b):
MGTH processed image.

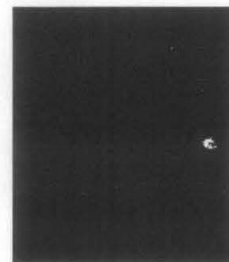


Figure 7.9(c):
Region grown image.

The malignancy factor is shown in Figure 7.9(d). The system's diagnoses are shown in Figure 7.9(e).

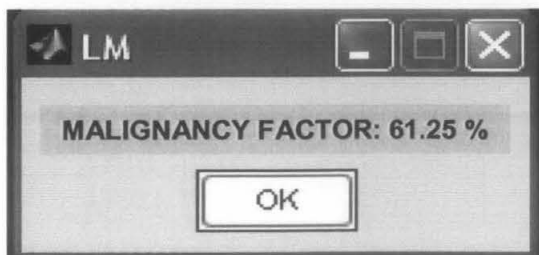


Figure 7.9(d):
Malignancy factor display.

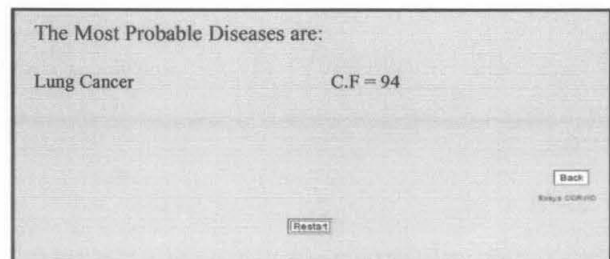


Figure 7.9(e):
Diagnosis by LUNEX.

LUNEX findings:

The malignancy factor is shown as 61.25% in Figure 7.9(d). The system diagnoses Lung Cancer as the most probable disease with a high confidence factor as in Figure 7.9(e) after taking into account patient's personal and medical background and clinical data.

Doctor's comments:

In the original image, a large infiltrative mass can be observed. This is a major indicator of lung cancer. The MGTH processed image shows an irregular lung nodule. The malignancy factor further adds to confirm our observation. Thus, lung cancer is the most probable disease.

7.2.3 Case studies of Patient J, K and L

In the same way, a few more case studies were analyzed using LUNEX. The findings of LUNEX for the selected three more case studies are discussed briefly.

The original X-Ray image of patient J, the processed image by MGTH and the region grown image are shown Figures 7.10(a), 7.10(b) and 7.4(c) respectively.



Figure 7.10(a):
Original X-Ray of J.



Figure 7.10(b):
MGTH processed image.

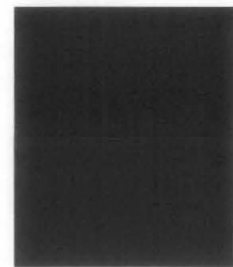


Figure 7.10(c):
Region grown image.

The malignancy factor and the diagnosis by LUNEX are displayed in Figures 7.10(d) and 7.10(e) respectively.

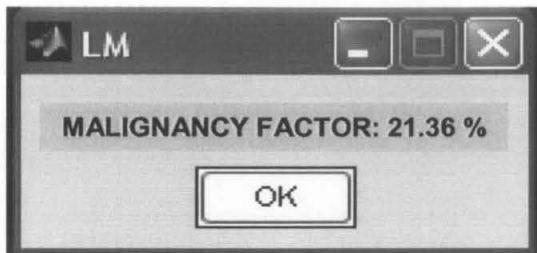


Figure 7.10(d):
Malignancy factor display.

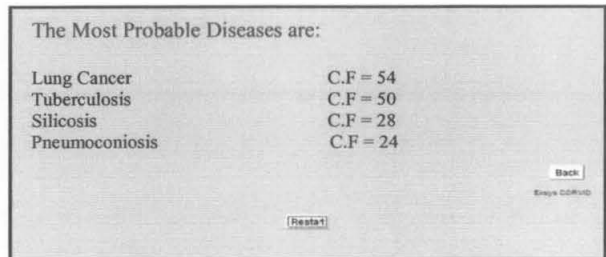


Figure 7.10(e):
Diagnosis by LUNEX.

LUNEX findings:

The malignancy factor is 21.36 %. The system diagnoses Lung Cancer as the most probable disease followed closely by Tuberculosis. Silicosis and Pneumoconiosis are also listed as probable diseases with lower confidence factors.

Doctor's comments:

In the processed image, there are two visible nodules which are almost circular and smooth in shape. The shape of the modules indicates a benign condition which is in line with the malignancy factor. Lung cancer is a likely disease to occur. From clinical data, it is suspected that there may be a possibility of Tuberculosis. Tuberculosis also may cause nodule like structures and hence the diagnosis of the system seems to be acceptable.

The original X-Ray image of the patient K, the processed image by MGTH and the region grown image are shown in Figures 7.11(a), 7.11(b) and 7.11(c) respectively.

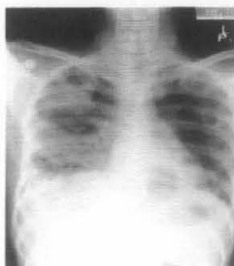


Figure 7.11(a):
Original X-Ray of K.

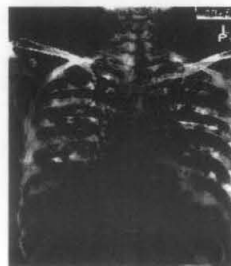


Figure 7.11(b):
MGTH processed image.



Figure 7.11(c):
Region grown image.

The malignancy factor is shown in Figure 7.11(d) and the LUNEX diagnosis is displayed in Figure 7.11(e).

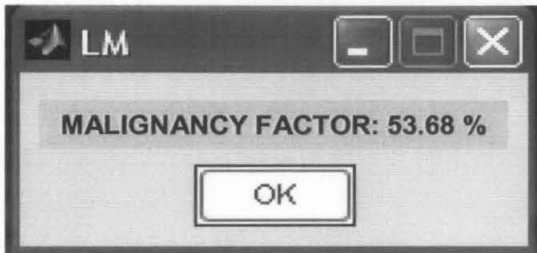


Figure 7.11(d):
Malignancy factor display.

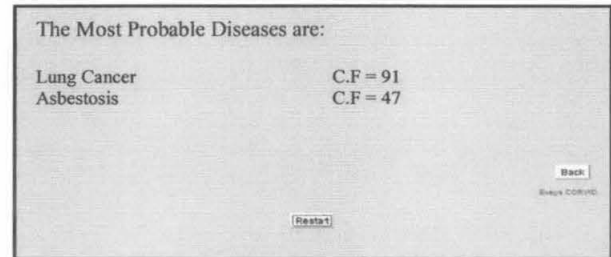


Figure 7.11(e):
Diagnosis by LUNEX.

LUNEX findings:

The malignancy factor is 53.68 %. The system has diagnosed Lung Cancer as the most probable disease. This is followed by Asbestosis.

Doctor's comments:

Several nodule-like structures are observed in the MGTH image. But it is found that four predominating nodules are only displayed. Since there is more than one nodule, it can be concluded that the nodules are metastasis (secondary cancer). The irregular nodules indicate a probable malignant condition. The malignancy factor displayed by the system seems to be reasonable. The diagnosis indicates that lung cancer may be the most probable disease.

The diagnosis for patient L is given below. The original X-Ray image of the patient, the MGTH processed image and the region grown image are shown in Figures 7.12(a), 7.12(b) and 7.12(c) respectively.



Figure 7.12(a):
Original X-Ray of L.



Figure 7.12(b):
MGTH processed image.

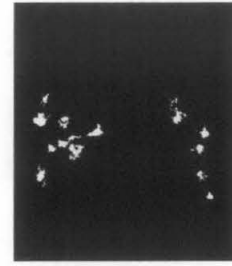


Figure 7.12(c):
Region grown image.

The malignancy factor is shown in Figure 7.11(d) and the LUNEX diagnosis is displayed in Figure 7.11(e).

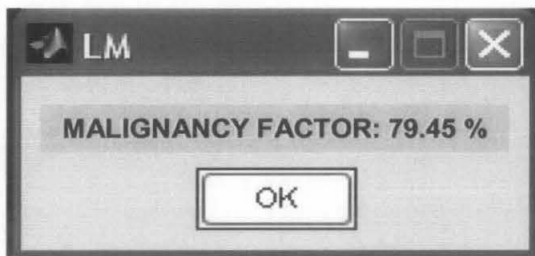


Figure 7.12(d):
Malignancy factor display.

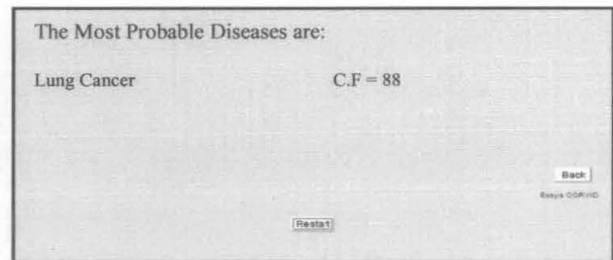


Figure 7.12(e):
Diagnosis by LUNEX.

LUNEX findings:

The malignancy factor is 79.45%. The probability of malignancy is quite high. The system has diagnosed Lung Cancer as the most probable disease.

Doctor's comments:

After processing, many lung nodules can be clearly observed. The nodules seem to have irregular shapes. That is why the malignancy factor seems to be on the higher side. Since, more than one nodule is present; this is also a case of metastasis. Diagnosis of the system as lung cancer is acceptable. Even though the system has not indicated, it is felt that the patient is having the symptoms of Lymphoma.

Summary

The images from a chest X-Ray are invaluable in detecting problems with the lungs. However, due to the limitations of the X-Ray, details might be missed. The use of imaging functions has greatly enhanced the X-Ray images to increase the visibility of subtle details. LUNEX was tested by general physicians and the results were verified by specialist doctors. The system's diagnosis was compared with the diagnosis of specialist doctors. About 88% of the diagnosis made by LUNEX matches the diagnosis of the specialist doctors. The system is able to give good results which are satisfactory to the specialist doctors and is acceptable by clinical standards. LUNEX can be effectively used by general physicians in rural places where specialists are not available to quickly diagnose the lung diseases and their conditions of severity. Based on severity, the physician can make decisions whether the patients can be treated locally or to be sent to specialist centers. This will help to raise the healthcare quality in the primary level itself.

CHAPTER 8

CONCLUSION AND RECOMMENDATIONS

8.1 Conclusion

A rule-based expert system with an embedded imaging module, named LUNEX for diagnosing lung diseases has been developed. The expert system contains a large knowledge base which incorporates many categories of patient information such as the patients' personal and medical background, clinical symptoms, clinical test results and digital radiological information. A large knowledge base would enable the system to generate more accurate results as the diagnosis of lung diseases often require a comprehensive analysis of various types of medical data.

An imaging module containing image processing functions is integrated into the expert system. This is a new feature which is not available in many disease diagnosing systems. The diagnosis of lung diseases often require the analysis of medical images such as X-Rays which is a common diagnostic imaging procedure. The various integrated image processing functions of the module will enable the doctors to enhance and analyze the X-Ray images in order to come up with an accurate interpretation.

The imaging tool called CIPT also consists of a lung nodule detection and diagnosis function which is specific for lung nodules in chest X-Rays. Lung nodules are the primary indicators of lung cancer which is a malignant disease. Thus, a special focus is given to the diagnosis of lung cancer using high level image processing and FIS technique.

A lung nodule detection and diagnosis tool is achieved using a new morphological method which is named as Modified Grayscale Top Hat (MGTH) transform. This tool

incorporating a fuzzy inference system with shape feature extraction capability that can diagnose and determine the probability of malignancy. A new measurement technique called malignancy factor which indicates the probability of the severity of the lung nodules has been introduced.

LUNEX is developed to provide an interactive and real-time disease diagnosis approach. LUNEX is designed to diagnose about 60 types of lung diseases. The diseases covered range from commonly occurring diseases to rare lung conditions.

The system has been tested with more than 60 patient test cases and X-Ray images obtained from hospitals. The test set contains a variety of patient information that covers many diseases. The information can be fed into the expert system by general physicians in rural and other clinics where specialist doctors are not available. In most cases, LUNEX is able to make diagnoses which are similar to those of the specialist doctors. The confidence factor provided by LUNEX for various test cases is acceptable for the specialist doctors.

The integrated imaging module in the system is very useful for the general physicians for online enhancement of the relevant segments of the images. However, some basic image processing knowledge is essential to the user of the LUNEX in order to understand the various functions available in the imaging module. This can be remedied by providing a user manual with complete details and simple training on the various imaging functions available.

As the system gives a probability of malignancy, the radiologists may be able to pre-screen the patients' X-Ray images in order to determine who needs immediate medical treatment.

The diagnosis is heavily subjected to the users' interpretation of a medical condition. The choices given to the users are in various levels of uncertainties. Thus, different users

might give different answers for the questions posed by the system for the same test case. Thus, two users might obtain slightly different diagnosis on the same patient. This is a normal occurrence in the medical field as the diagnosis is very much dependent on the doctor's experience, interpretation and heuristics. The medical data are also often shrouded with impreciseness and inaccuracies. However, the results by LUNEX have been tested and are found to be within the clinically acceptable standards.

This expert system was designed to help general physicians especially those based in rural areas to perform complex diagnoses without the need for the services of specialists' doctors and advanced facilities. This would be very helpful in saving cost and time for both the doctors and patients and generally raise the standard of health services in rural areas.

LUNEX will not only aid the doctors in making pre-diagnosis of the severity of the diseases but also serve as a training tool for graduating doctors. This system even though can make diagnoses, it is necessary to confirm the results of the system by real specialist doctors depending upon the severity of the diseases.

8.2 Recommendations

8.2.1 Recommendations for LUNEX

The capability of expert systems and computer aided image analysis tools can be further enhanced by incorporating some new features to overcome certain drawbacks. The analyses of X-Ray images need high resolution processing and display systems. However, most of the normal computers available do not have high resolution processors and display units. Thus, very subtle abnormalities and details may not be very visible. Higher resolution monitors are necessary especially for displaying medical images in order to maximize the capability of the system.

It is suggested the system is expanded to include the analysis of other medical images and tests such as CT scan, MRI and cytology (the study of the cells of a specimen taken from a biopsy). The image processing functions could be applied to enhance these images in order to obtain more accurate diagnosis.

A medical text archive can be included into LUNEX. This archive will contain detailed information on all the diseases covered by the system. The user will just have to click on a disease name to obtain information regarding a particular disease. The archive can also contain a dictionary of medical terms that can be used anytime during the consultation. The system should be linked to the internet so that the user can obtain more details about the various medical conditions that he or she encounters.

8.2.2 General Recommendations

Expert systems can emulate the thought process of medical specialists in making complex decisions through the use of AI techniques. Thus, medical expert systems can help the doctors in making timely and accurate decision at the primary care level itself and generally raise the health quality of the people.

There are many different types of clinical tasks to which expert systems can be applied. Some of them are highlighted below.

Diagnostic assistance. In the medical field, complex medical decisions are needed to be made on a day to day basis. Most often the patient's case might be very complex or the person making the decision is inexperienced. In this situation, an expert system can help the doctor to come up with a quick and an acceptable diagnosis.

Therapy critiquing and planning. Expert systems can also incorporate the therapeutic recommendations according to the diagnosis made by them. Systems can either look for inconsistencies, errors and omissions in an existing treatment plan, or can be used to formulate a treatment based upon a patient's specific condition and accepted treatment guidelines.

Generating alerts and reminders. In so-called real-time situations, an expert system can warn of changes in a patient's condition. In less acute circumstances, it might scan laboratory test results or drug orders and send reminders or warnings through an e-mail system.

REFERENCES

- Aikins, J. S., Kunz, J.C., Shortliffe, E.H., & Fallat, R.J. (1983). PUFF: An expert system for interpretation of pulmonary function data. *Journal of Computers and Biomedical Research, Vol. 16*, pp199-208.
- Alhady, S.S.N., Venkatachalam, P.A., & Sulaiman, M. (2000). *Noiseless ECG monitoring system with integrated expert system: Proceedings of First International Conference on Advances in Medical Signal and Information Processing, 2000*. IEE Conf. Publ. No. 476, pp. 79 – 87.
- Angelina, A., Tzacheva, B.S., Najarian, K., & Brockway, J.P. (2003). Breast cancer detection in gadolinium enhanced MR images by static region descriptors and neural networks. *Journal of Magnetic Resonance Imaging, Vol. 17(3)*, pp. 337-342.
- Awad, E. (1996). *Building Expert Systems: Principles, Procedures and Applications*. Minneapolis/St.Paul: West Publishing Company.
- Awad, E. (2003). *Building Knowledge Automation Expert Systems: With Exsys Corvid, (2nd ed.)*. USA: EXSYS Inc.
- Balko, J.A., & Tao, L.C. (1994). An expert system for teaching cytopathologic diagnosis of lung cancers. *Anal Quant Cytol Histol., Vol. 16(5)*, pp. 321-331.
- Barata, T., & Pina, P. (2002). Morphological segmentation of remotely sensed forest covers in high spatial resolution images. In H. Talbot & R. Beare (Eds.), *Mathematical Morphology* (pp. 147-156). Sydney: CSIRO Publishing.

- Barnett, G.O., Cimino, J.J., Hupp, J.A., & Hoffer, E.P (1987). DXplain: An evolving diagnostic decision support system. *Journal of the American Medical Association*. Vol. 258(1), pp. 67-74.
- Baxes, G.A. (1994). *Digital Image Processing: Principles and Applications (1st ed.)*. New York: John Wiley.
- Bovik, A. (2000). *Handbook of Image and Video Processing (1st ed.)*. USA: Academic Press.
- Buchanan, B.G., & Shortliffe, E.H. (1984). *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*. Reading: Addison-Wesley.
- Burdick, H.E. (1997). *Digital Imaging: Theory and Applications*. USA: McGraw-Hill.
- Castleman, K.R. (1995). *Digital Image Processing (1st ed.)*. USA: Prentice Hall.
- Chiou, Y.S.P., Lure, Y.M.F., & Ligomenides, P.A. (1994). Neural network image analysis and classification in hybrid lung nodule detection (hlnd) system. *Cancer Letters*, Vol. 77(2-3), pp. 119-26.
- Corne, J., Carroll, M., & Delany, D. (1999). *Chest X-ray Made Easy (2nd ed.)*. Edinburgh: Churchill-Livingstone.
- Daffner, R. (1999). *Clinical Radiology: The Essentials (2nd ed.)*. Baltimore: Williams & Wilkins.

- Economou, G.P.K., Spiropoulos, C., Economopoulos, N.M., Charokopos, N., Lymberopoulos, D., Spiliopoulou, M., Haralambopulu, E., & Goutis, C.E. (1994). *Medical diagnosis and artificial neural networks: A medical expert system applied to pulmonary diseases : Proceedings of the IEEE Workshop of Neural Networks for Signal Processing , 1994*, pp. 482-489.
- Fahmy, G., Nassar, D., Haj-Said, E., Chen, H., Nomir, O., Zhou, J., Howell, R., Ammar, H.H., Abdel-Mottaleb, M., & Jain, A.K. (2004). *Towards an automated dental identification system (ADIS): Proceedings of International Conference on Biometric Authentication (ICBA), 2004*, pp. 789-796.
- Farzan, S., & Farzan, D. (1997). *A Concise Handbook of Respiratory Diseases (4th ed.)*. USA: Apple & Lange.
- Feigenbaum, E.A., & Buchanan, B.G. (1993). DENDRAL and Meta-DENDRAL: Roots of knowledge systems and expert system applications. *Artificial Intelligence, Vol. 59*, pp. 233-240.
- Felitti, V.J. (2005). GIDEON: The global infectious disease and epidemiology network. *Journal of American Medical Association, Vol. 293*, pp. 1674-1675.
- Floyd, C., Patz, E., Lo, J., Vittitoe, N., & Stambaugh, L. (1996). Diffuse nodular lung disease on chest radiographs: A pilot study of characterization by fractal dimension. *American Journal of Roentgenology, Vol. 167*, pp. 1185-1187.
- Forero, M., Sroubek, F., & Cristobal, G. (2004). Identification of tuberculosis bacteria based on shape and color. *Real-Time Imaging, Vol. 10(4)*, pp. 251-262.
- Giarratano, J., & Riley, G. (1998). *Expert Systems: Principles and Programming (3rd ed.)*. Boston: PWS Publishing.

- Giger M., Doi, K., & MacMahon, H. (1988). Image feature analysis and computer-aided diagnosis in digital radiography: Automated detection of nodules in peripheral lung fields. *Medical Physics*, Vol. 15(2), pp. 158–166.
- Ginneken, B.V. (2001). *Computer-Aided Diagnosis in Chest Radiography* (Doctoral dissertation, Image Sciences Institute, University Medical Center Utrecht, 2001).
- Goatman, K. (1997). Automated detection of microaneurysms. Retrieved May 14, 2005 from <http://www.biomed.abdn.ac.uk/Abstracts/A07890/>.
- Gonzalez, R., & Woods R. (2002). *Digital Image Processing (3rd ed.)*. New Jersey: Prentice Hall.
- Gurney, J.W., & Swensen, S. (1995). Solitary pulmonary nodules: Determining the likelihood of malignancy with neural network analysis. *Radiology*, Vol. 196, pp. 823–829.
- Gurney, J.W. (2002). *Chest X-ray: Your thoracic imaging resource*. Retrieved July 12, 2005 from <http://www.chestx-ray.com/GenPublic/GenPubl.html>.
- Hatzilygeroudis, P. J., Vassilakos, A., & Tsakalidis (1997). *XBONE: A hybrid expert system supporting diagnosis of bone diseases: Proceedings of the Medical Informatics Europe, 1997*, Greece, pp. 259-299.
- Hensel, M., Wiesner, G., Kuhrmann, B., Pralow, T., & Grigat, R. (2005). *Motion and noise detection for adaptive spatio-temporal filtering of medical x-ray image sequences: Proceedings of 9th Annual Conference on Medical Image Understanding and Analysis (MIUA), 2005*, pp. 219-222.

- Hornack, J.P. (1996). *Basics of MRI*, Retrieved May 18, 2005, from <http://www.cis.rit.edu/htbooks/mri/chap-14/chap-14.html>.
- Ignizio, J. (1991). *Introduction to Expert Systems: The Development and Implementation of Rule-Based Expert Systems (1st ed.)*. USA: McGraw-Hill.
- Ilic, S., Ulicny, B. (2000). *Seeded region growing method for image segmentation*. Retrieved May 23, 2005 from <http://ligwww.epfl.ch/~silic/predocschool/ComputerVision/cvision.html>.
- Image quality in chest radiography. (2003). *The Journal of the International Committee of Radiographic Units and Measurements, ICRU Report No. 70, Vol. 3(2)*, pp. 1-176.
- Jackson, P. (1999). *Introduction to Expert Systems (3rd ed.)*. USA: Addison-Wesley.
- Jain, A.K. (1989). *Fundamentals of Digital Image Processing*. New Jersey: Prentice Hall.
- Jardins, T.D, & Burton, G.G. (2001). *Clinical manifestation and Assessment of Respiratory Disease (4th ed.)*. USA: C.V Mosby.
- Judith, F. (n.d.). *Medical expert systems*. Retrieved June 15, 2005, from http://www.computer.privateweb.at/judith/special_field3.html.
- Kahn, M.G., Steib, S.A., Fraser, V.J., & Dunagan, W.C. (1993). *An expert system for culture-based infection control surveillance: Proceedings of Symposium on Computer Applications in Medical Care, 1993*, New York, pp. 171-175.

- Kano, K.D., MacMahon, H., Hassell, D., & Giger, M. (1994). Digital image subtraction of temporally sequential chest images for detection of interval change, *Medical Physics*, Vol. 21(3), pp. 453–461.
- Leef, J., & Klein, J. (2002). The solitary pulmonary nodule. *Journal of Clinical Radiology of North America*, Vol. 40, pp. 123-143.
- Li, Q., Katsuragawa, S., Engelmann, R., Armoto, S., MacMahon, H., & Doi, K. (2001). *Development of a multiple-templates matching technique for removal of false positives in computer-aided diagnostic scheme : Proceedings of SPIE, 2001, Vol. 4322*, pp. 1763–1770.
- Lincoln, M.J., Turner, C.W., Haug, P.J., Warner, H.R., Williamson, J.W., Bouhaddou, O., Jessen, S.G., Sorenson, D., Cundick, R.C., & Grant, M. (1991). Iliad training enhances medical students' diagnostic skills. *Journal of Medical Systems*, Vol. 15(1), pp. 93-110.
- Liu, W.Y., Wang, T., & Zhang, H.J. (2000). *A hierarchical characterization scheme for image retrieval: Proceedings of the IEEE International Conference on Image Processing (ICIP), 2000, Vol. 3*, pp. 42-45.
- Loddenkemper, R., Gibson, G.J., & Sibille, Y. (2003). The burden of lung disease in Europe: Why a European white book on lung disease?. *European Respiratory Journal*, Vol. 22(6), pp. 869.
- Lung. (2005). Library of Radiopharmaceutic Images, University of Kentucky. Retrieved August 12, 2005 from <http://www.uky.edu/Pharmacy/research/Lung.html>.
- Matlab Fuzzy Logic Toolbox Documentation (5th ed.)*. (2001). USA: MathWorks Inc.

- Meesad, P. (2004). *Fuzzy systems*. Retrieved March 19, 2005 from <http://kmitnb05.kmitnb.ac.th/~pym/fuzzy.html>.
- Meholic, A., Ketai, L., & Lofgren, R. (1996). *Fundamentals of Chest Radiology (1st ed.)*, USA: W.B Saunders Publishing.
- Ministry of Health, Malaysia. (2003). *Second report of the national cancer registry: Cancer incidence in Malaysia 2003*.
- Ministry of Health, Malaysia. (2004). *Health facts 2003*.
- Muellor, N.L., Paret, P., Fraser, R., & Colman N. (2001). *Radiologic Diagnosis of Diseases of the Chest*. New York: Elsevier-Health Sciences.
- Muhm, J., Miller, W., Fontana, R., Sanderson, D., & Uhlenhopp, M. (1983). Lung cancer detected during a screening program using four-month chest radiographs. *Radiology*, Vol. 148, pp. 609–615.
- Mukhopadhyay, S., & Chanda, B. (2002). *Hue preserving color image enhancement using multi-scale morphology: Proceedings of Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP), 2002, India*, pp.179-184.
- Munday, E., & Taylor, C. (2002). Pre-processor location of possible masses in mammograms. *British Machine Vision Association*, pp. 145-148.
- Murray, J.F., Nadel, J.A., Mason, R.J., & Broaddus, C. (2000). *Textbook of Respiratory Medicine (3rd ed.)*. USA: W.B Saunders.

- Nakamura, K., Yoshida, H., Engelmann, R., MacMahon, H., Katsuragawa, S., Ishida, T., Ahizawa, K., & Doi, K. (2000). Computerized analysis of the likelihood of malignancy in solitary pulmonary nodules with use of artificial neural networks. *Radiology, Vol. 214*, pp. 823–830.
- National Cancer Registry. (2003). *The first report of the national cancer registry cancer incidence in Malaysia 2002*. Ministry of Health, Malaysia.
- Negnevitsky, M. (2005). *Artificial Intelligence: A Guide to Intelligent Systems (2nd ed)*. USA: Addison Wesley.
- Newell, A., & Simon, H. A. (1976). Computer science as empirical enquiry. *Communications of the Association for Computing Machinery, Vol. 19(3)*, pp. 113-126.
- Ngah, U.K. (1996). An expert system with integrated image processing features for the diagnoses of breast diseases (Thesis, Universiti Sains Malaysia, 1996).
- Ngah, U.K., Venkatachalam, P.A., & Hani, A.F.M. (1996). Mammographic image integrated expert system. *International Journal of Computers and Their Applications, Vol. 3(3)*, pp. 122-129.
- Ngah, U.K., Hai, O.T., Khalid, N.E.A., & Venkatachalam, P.A. (2000). *Mammographic calcification clusters using region growing techniques : Proceedings of New Millenium International Conference on Pattern Recognition and Robot Vision (TATI), 2000*, Malaysia.
- Nixon, M., & Aguado, A. (2002). *Feature Extraction and Image Processing (1st ed.)*. Oxford: Newnes Press.

- Ongun, G., Halici, U., Leblebicioglu, K., Atalay, V., Beksac, M., & Beksac, S. (2001). Feature extraction and classification of blood cells for an automated differential blood count system. *IEEE Publication*, pp. 2461-2466.
- Parkin, M., Bray, F., Ferlay, J., & Pisani, P. (2005). Global cancer statistics 2002. *CA: A Cancer Journal for Clinicians*, Vol. 55, pp. 74-108.
- Pereira, J.C.R., Tonelli, P.A., Barros, L.C., & Ortega, N.R.S. (2002). Defuzzification in medical diagnosis. *Advances in Logic, Artificial Intelligence and Robotics*, pp. 202-207.
- PET: Position Emission Tomography, Power of molecular imaging*. (2005). Jennifer Jones Simon Foundation, UCLA School of Medicine, Institute for Clinical PET & U.S. Department of Energy. Retrieved May 12, 2005 from <http://www.nuc.ucla.edu/pet/pdf/petbrochure.pdf>.
- Phuong, N.H, Hung, D.H., Tuan, D.T., Co, N.V., Duong, B.D., Thinh, P.T., Hoa, N.P., & Linh, N.N. (1998). *TUBERDIAG: An expert System for pulmonary tuberculosis diagnosis : Proceedings of IEEE International Conference on Systems, Man, and Cybernetics, 1998, Vol. 2*, pp. 1587-1590.
- Popoola, D. (2004). *Fuzzy expert systems*. Retrieved July 14, 2005 from www.computing.surrey.ac.uk/courses/cs364/fuzzy_examples_Negnevitski.ppt.
- Pratt, W.K. (1991). *Digital Image Processing (2nd ed.)*. New York: John Wiley.

- Preece, A.D. (1993). Specifications and tools for building reliable expert systems. *International Journal of Software Engineering and Knowledge Engineering*, Vol. 3(1), pp. 17-52.
- Pressman, R.S. (2001). *Software Engineering : A Practitioner's Approach (5th ed.)*. New York: McGraw-Hill.
- Russ, J. (1998). *The Image Processing Handbook (3rd ed.)*. USA: CRC Press.
- Scott, W.J. (2001). *Lung Cancer: A Guide to Diagnosis and Treatment (1st ed.)*. USA: Addicus Books.
- Serpen, G., Acharya, V., Woldenberg, L.S., Coombs, R.J., & Parsai, E.I. (2000). *FIS for PLOPED compliant diagnosis of pulmonary embolism: Proceedings of Artificial Neural Networks in Engineering (ANNIE), 2000*, USA.
- Seul, M., O'Gorman, L., & Sammon, M. (2000). *Practical Algorithms for Image Analysis-Description, Examples and Code (1st ed.)*. United Kingdom: Cambridge University Press.
- Shen, L., Ranggayan, R.M., & Desautels, J.E.L.D. (1994). Application of shape analysis to mammographic calcifications. *IEEE Transactions on Medical Imaging*, Vol. 13(2), pp. 263-274.
- Sklansky, J., & Ballard, D. (1973). Tumor detection in radiographs. *Comput. Biomed. Res.*, Vol. 6(4), pp. 299-321.
- Sumner, W. (1993). A review of Iliad and Quick Medical Reference for primary care providers. Two diagnostic computer programs. *Journal of Archives of Family Medicine* Vol. 2, pp. 87-95.

- Sutton, D. (1997). *Textbook of Radiology and Imaging (2nd ed.)*. Edinburgh: Churchill-Livingstone.
- Umbough, S. (1998). *Computer Vision and Image Processing (International Edition)*. New Jersey: Prentice Hall.
- Verdaguer, A., Patak, A., Sancho, J.J., Sierra, C., & Sanz, F. (1992). Validation of the medical expert system PNEUMON-IA. *Computer on Biomedical Research, Vol.5(6)*,pp. 511-526.
- Vittitoe, N., Baker, J., & Floyd, C. (1997). Fractal texture analysis in computer- aided diagnosis of solitary pulmonary nodules. *Journal of Academic Radiology, Vol. 4*, pp. 96–101.
- Ward, J. (2002). *The Respiratory System at a Glance (1st ed.)*. United Kingdom : Blackwell Science.
- What are the radiation risks from CT?* (n.d). U.S Food and Drug Administration, USA, Center for devices and radiological health. Retrieved July 14, 2005 from <http://www.fda.gov/cdrh/ct/risks.html>.
- World Health Organization. (2005). *Practical approach to lung health : A primary health care strategy for the integrated management of respiratory conditions in people five years of age and over*.
- Wróblewska, A., Boniński, P., Przelaskowski, A., & Kazubek, M. (2003). Segmentation and feature extraction for reliable classification of microcalcifications in digital mammograms. *Opto-Electronics Review. Vol. 2(3)*, pp. 227-235.

- Xu, X., MacMahon, H., Giger, M., & Doi, K. (1997). *Adaptive feature analysis of false positives for computerized detection of lung nodules in digital chest radiographs: Proceedings of SPIE, 1997, Vol. 3034*, pp. 428–436.
- Yongbum, L., & Du-Yih, T. (2004). *Computerized classification of microcalcifications on mammograms using fuzzy logic and genetic algorithm: Proceedings of SPIE Medical Imaging 2004, Vol. 5370*, pp. 952-959.
- Yoshida, H., Xu, X., Doi, K., & Giger, M. (1995). *Computer-aided diagnosis (CAD) scheme for detecting pulmonary nodules using wavelet transforms: Proceedings of SPIE, 1995, Vol. 2434*, pp. 621-626.
- Yue, Z., Shou-nan, B., & Feng-ohi, L. (1986). Coal worker's pneumconiosis x-ray diagnosis expert system, PXDES. *Bulletin for Studies and Exchanges on Fuzziness and its Applications BUSEFAL, Vol. 29*. pp 104-108. Retrieved August 15, 2005 from <http://www.listic.univ-savoie.fr/>.
- Zhao H., Lo, S., Freedman, M., & Mun, S. (2001). *On automatic temporal subtraction of chest radiographs and its enhancement for lung cancers: Proceedings of SPIE, 2001, Vol. 4322*, pp. 1867–1872.

PUBLICATIONS

1. **Kavitha Shaga Devan, P.A Venkatachalam and Ahmad Fadzil Mohd. Hani.** “Expert System with an Embedded Imaging Module for Diagnosing Lung Diseases”. *Proceedings of 7th International Workshop on Enterprise Networking and Computing in Healthcare Industry (HEALTHCOM) 2005*, Korea. pp. 229-233.

2. **Kavitha Shaga Devan, P.A Venkatachalam and Ahmad Fadzil Mohd. Hani.** “Lung Nodule Diagnosis and Determination of Malignancy Factor”. *Proceedings of International Conference on Intelligence Systems 2005*, Malaysia.

APPENDICES

Appendix A

In this section, the decision tables for four categories of medical data are given. These categories are as listed below.

- Table 1: Patient's personal and medical background
- Table 2: Clinical symptoms
- Table 3: Clinical tests results
- Table 4: Radiological findings

For all the decision tables, the columns indicate the medical conditions and the rows indicate the list of diseases. In the tables, the condition that coincides with a disease is marked with the symbol \surd . In Table 3, some of the clinical tests are listed with the corresponding output of the tests.

In the decision tables, symbols are given to name the medical conditions. The tables below act as legends for the medical conditions in Table 1, 2 and 4. Figure 1 is the legend for Table 1, Figure 2 is the legend for Table 2 and Figure 3 is the legend for Table 4.

For Table 3, the clinical tests are named in the decision table itself, thus a legend is not needed.

Figure 1: Legend for Table 1.

H1	Age	H11	Infections
H2	Gender	H12	Choking/Inhaling foreign objects
H3	History of lung diseases	H13	Obesity
H4	History of other diseases	H14	Exposure to nature
H5	Family history of disease	H15	Activities
H6	Prolonged exposure to smoke	H16	Drug Consumption
H7	Smoking	H17	Malnutrition
H8	Exposure to chemicals	H18	Exposure to people with diseases
H9	Trauma/Injury	H19	Diseases triggered by some elements
H10	Organ Failure	H20	Allergies

Figure 2: Legend for Table 2.

S1	Coughing	S22	Dizziness
S2	Coughing blood	S23	Sweating
S3	Dyspnea	S24	Abdominal problems
S4	Wheezing	S25	Anxiety
S5	Weight Loss	S26	Chills
S6	Phlegm expectoration	S27	Itchiness
S7	Fatigue	S28	Appetite loss
S8	Cyanosis	S29	Sleeping problems
S9	Clubbing	S30	Eye problems
S10	Bad breath	S31	Skin problems
S11	Chest pain	S32	Enlarged veins
S12	Headaches	S33	Enlarged lymph nodes
S13	Rapid heart beat	S34	Disease trigger effect
S14	Leg and ankle swelling	S35	Fainting
S15	Runny nose	S36	Chest tightness
S16	Sore throat	S37	Facial swelling
S17	Chest retraction	S38	Nose bleeds
S18	Pain/aches	S39	Easily bruise
S19	Stuffy nose	S40	Speaking difficulties
S20	Sneezing	S41	Poor coordination
S21	Nasal discharge		

Figure 3: Legend for Table 4.

R1	Hyper-radiolucency	R26	Consolidation
R2	Parenchymal infiltration	R27	Cavitation
R3	Homogenous density	R28	Lobe/segment retraction
R4	Patches of density	R29	Patchy parenchymal density
R5	Diffuse reticular or nodular density	R30	Miliary pattern
R6	Radiolucent area	R31	Fungus ball
R7	Parenchymal lesion	R32	Abscess formation
R8	Cardiac silhouette	R33	Air build up
R9	Hilum enlargement	R34	Fluid build up
R10	Hilum displacement	R35	Atelectasis
R11	Flattening of diaphragm	R36	Volume loss
R12	Elevation of diaphragm	R37	AP chest diameter increase
R13	Mediastinal lymph node	R38	Hyperinflation
R14	Mediastinal shift	R39	Tramline shadows
R15	Bronchovascular marking	R40	Transient shadows
R16	Lung markings	R41	Sternum space
R17	Honeycombing	R42	Mucoid impaction
R18	Hazy lung	R43	Migratory infiltrates
R19	Calcifications	R44	Bilateral diffuse mottling
R20	Lesion/mass	R45	Lung fissures
R21	Nodular shadows	R46	Prominent Artery
R22	Bullous lesions	R47	Lung detachment
R23	Pleural effusion	R48	Heart deviation
R24	Pleural thickening	R49	Malignancy factor
R25	Fibrosis		

Table 1: Patients' personal and medical history decision table.

	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12	H13	H14	H15	H16	H17	H18	H19	H20
Acute Bronchitis			√																	
Adult Respiratory Distress Syndrome (ARDS)			√	√		√	√		√	√	√	√								
Allergic Bronchopulmonary Aspergillosis																				
Alpha-1 related Emphysema		√	√			√	√													
Asbestosis							√	√												
Aspergilloma											√									
Aspiration Pneumonitis																				
Asthma			√		√									√					√	√
Atelectasis												√								
Blastomycosis			√								√									
Bronchiectasis	√		√	√			√					√								
Bronchiolitis	√		√																	
Bronchitis Obliterans	√		√																	
Chronic Bronchitis	√	√				√	√				√		√							
Chronic Cor Pulmonale		√	√	√						√										
Chronic Necrotizing Pulmonary Aspergillosis	√			√														√		
Croup			√								√									
Coccidioidomycosis											√									
Common Cold																			√	
Cystic Fibrosis					√															
Diphtheria																				
Emphysema		√	√			√	√													
Eosinophilic Pneumonias																				
Gaucher Disease					√															
Goodpasture Syndrome	√	√																		
Hantavirus Pulmonary Syndrome														√	√					
Hayfever					√										√				√	√
Hemothorax			√	√					√											
Histoplasmosis				√											√					
Idiopathic Pulmonary Fibrosis	√	√																		
Influenza	√		√												√					
Kaposi's Sarcoma				√																

Table 1: Patients' personal and medical history decision table (contd.).

	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12	H13	H14	H15	H16	H17	H18	H19	H20
Khyphoscoliosis																				
Laryngitis						√	√	√			√									
Lung Absess			√	√																
Lung Cancer			√			√	√	√												
Lymphangioliomyomatosis	√	√																		
Lymphatic Interstitial Pneumonitis	√																			
Mesothelioma							√								√					
Mountain Sickness															√					
Pertussis											√									
Pleural Effusion			√	√																
Pleurisy			√					√	√		√									
Pneumoconiosis								√												
Pneumonia	√		√	√							√									
Pneumonitis						√										√				
Pneumothorax			√						√						√					
Primary Cilia Dyskinesia					√															
Progressive Systemic Sclerosis		√																		
Pulmonary Edema				√		√			√				√			√				
Pulmonary Embolism	√			√			√		√		√				√	√				
Pulmonary Fat Embolism				√																
Pulmonary Hypertension			√	√																
Pulmonary Parenchymal Injury									√											
Respiratory Diptheria																				
Sarcoidosis	√																			
Silicosis		√					√													
Systemic Lupus Erythematosis											√									
Tuberculosis	√			√											√	√				√
Wegener's Granulomatosis																				

Table 2: Clinical symptoms decision table.

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16	S17	S18	S19	S20
Acute Bronchitis	√			√		√	√				√				√	√			√	√
Adult Respiratory Distress Syndrome (ARDS)			√					√												
Allergic Bronchopulmonary Aspergillosis	√					√	√													
Alpha-1 related Emphysema	√		√	√	√	√	√	√	√					√						
Asbestosis	√	√	√				√				√									
Aspergilloma		√																		
Aspiration Pneumonitis	√		√					√												
Asthma	√		√	√			√	√					√			√				
Atelectasis	√		√				√	√			√									
Blastomycosis	√	√			√						√									
Bronchiectasis	√		√	√	√	√	√	√	√	√	√									
Bronchiolitis	√		√	√			√								√	√	√		√	√
Bronchitis Obliterans	√		√		√		√													
Chronic Bronchitis	√		√	√		√		√	√		√	√		√						
Chronic Cor Pulmonale																				
Chronic Necrotizing Pulmonary Aspergillosis		√																		
Croup	√			√			√	√								√			√	
Coccidioidomycosis	√											√			√			√	√	
Common Cold	√						√					√			√	√		√	√	√
Cystic Fibrosis	√				√	√														
Diphtheria																				
Emphysema	√		√	√	√	√	√	√	√					√						
Eosinophilic Pneumonias	√		√	√																
Gaucher Disease																				
Goodpasture Syndrome		√	√									√								
Hantavirus Pulmonary Syndrome	√		√				√					√						√		
Hayfever	√		√												√				√	√
Hemothorax			√								√									
Histoplasmosis	√		√				√				√	√			√	√		√	√	
Idiopathic Pulmonary Fibrosis	√		√						√											
Influenza	√						√					√			√	√		√		√
Kaposi's Sarcoma																				

Table 2: Clinical symptoms decision table (contd.).

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16	S17	S18	S19	S20
Khyphoscoliosis			√																	
Laryngitis	√															√				
Lung Absess	√	√				√	√													
Lung Cancer	√	√	√	√			√				√					√				
Lymphangioleiomyomatosis		√	√					√			√		√							
Lymphatic Interstitial Pneumonitis	√		√		√						√									
Mesothelioma	√		√								√									
Mountain Sickness			√			√	√	√				√	√							
Pertussis	√		√												√					
Pleural Effusion			√																	
Pleurisy	√																			
Pneumoconiosis	√		√					√			√									
Pneumonia	√		√			√		√												
Pneumonitis																				
Pneumothorax	√		√								√									
Primary Cilia Dyskinesia			√	√																√
Progressive Systemic Sclerosis	√		√																	
Pulmonary Edema	√		√	√		√		√					√	√						
Pulmonary Embolism		√	√					√			√									
Pulmonary Fat Embolism		√						√												
Pulmonary Hypertension			√								√			√						
Pulmonary Parenchymal Injury	√	√	√					√												
Respiratory Diptheria																				√
Sarcoidosis	√		√		√		√				√						√			
Silicosis	√		√				√				√									
Systemic Lupus Erythematosis			√																	
Tuberculosis	√		√	√	√		√				√	√			√	√		√	√	√
Wegener's Granulomatosis	√	√	√	√											√	√		√		

Table 2: Clinical symptoms decision table (contd.).

	S21	S22	S23	S24	S25	S26	S27	S28	S29	S30	S31	S32	S33	S34	S35	S36	S37	S38	S39	S40	S41	
Acute Bronchitis	√																					
Adult Respiratory Distress Syndrome (ARDS)				√																		
Allergic Bronchopulmonary Aspergillosis																						
Alpha-1 related Emphysema					√																	
Asbestosis									√													
Aspergilloma																						
Aspiration Pneumonitis				√																		
Asthma														√		√					√	
Atelectasis																√						
Blastomycosis			√																			
Bronchiectasis																						
Bronchiolitis	√																					
Bronchitis Obliterans																						
Chronic Bronchitis	√																					
Chronic Cor Pulmonale																						
Chronic Necrotizing Pulmonary Aspergillosis																						
Croup																						
Coccidioidomycosis																						
Common Cold	√									√						√						
Cystic Fibrosis	√		√	√							√											
Diphtheria				√																		
Emphysema					√																	
Eosinophilic Pneumonias																						
Gaucher Disease																					√	
Goodpasture Syndrome		√																				
Hantavirus Pulmonary Syndrome						√																
Hayfever							√			√												
Hemothorax																						
Histoplasmosis	√																					
Idiopathic Pulmonary Fibrosis																						
Influenza	√																					
Kaposi's Sarcoma											√											

Table 2: Clinical symptoms decision table (contd.).

	S21	S22	S23	S24	S25	S26	S27	S28	S29	S30	S31	S32	S33	S34	S35	S36	S37	S38	S39	S40	S41
Khyphoscoliosis																					
Laryngitis													√								
Lung Absess			√			√															
Lung Cancer								√								√		√			
Lymphangioliomyomatosis																					
Lymphatic Interstitial Pneumonitis																					
Mesothelioma				√																	
Mountain Sickness		√			√				√												√
Pertussis				√											√						
Pleural Effusion																					
Pleurisy																√					
Pneumoconiosis																					
Pneumonia		√				√															
Pneumonitis																					
Pneumothorax				√																	
Primary Cilia Dyskinesia		√																			
Progressive Systemic Sclerosis																					
Pulmonary Edema					√							√									
Pulmonary Embolism		√			√											√					
Pulmonary Fat Embolism											√										
Pulmonary Hypertension		√		√											√	√					
Pulmonary Parenchymal Injury																					
Respiratory Diptheria													√					√			
Sarcoidosis			√						√	√			√			√					
Silicosis																					
Systemic Lupus Erythematosis																					
Tuberculosis	√		√	√	√			√													
Wegener's Granulomatosis			√							√	√					√	√				

Table C: Clinical tests decision table.

	ABG		CBC				Other Tests		
	CO ₂	O ₂	White blood cell	Red blood cell	Eosinophilic	ESR	Sweat test	Skin test	Blood Culture
Acute Bronchitis	High								
Adult Respiratory Distress Syndrome (ARDS)	Low	Low							Positive
Allergic Bronchopulmonary Aspergillosis					High			Positive	
Alpha-1 related Emphysema									
Asbestosis									
Aspergilloma					High				
Aspiration Pneumonitis	Low	Low							
Asthma		Low	High						
Atelectasis	High								
Blastomycosis									
Bronchiectasis		Low		Low					
Bronchiolitis		Low							
Bronchitis Obliterans						High			
Chronic Bronchitis		Low		Low					
Chronic Cor Pulmonale		Low							
Chronic Necrotizing Pulmonary Aspergillosis					High				
Croup									
Coccidioidomycosis									
Common Cold									
Cystic Fibrosis		Low					High		
Diphtheria									
Emphysema	High	Low	High	High					
Eosinophilic Pneumonias			High		High				
Gaucher Disease									
Goodpasture Syndrome	High		High	Low					
Hantavirus Pulmonary Syndrome		Low	High						
Hayfever					High			Positive	
Hemothorax									
Histoplasmosis								Positive	Positive
Idiopathic Pulmonary Fibrosis	Low	Low							

Table C: Clinical tests decision table (contd.).

	ABG		CBC				Other Tests		
	CO ₂	O ₂	White blood cell	Red blood cell	Eosinophilic	ESR	Sweat test	Skin test	Blood Culture
Influenza			Low						
Kaposi's Sarcoma									
Khyphoscoliosis	Low	Low							
Laryngitis									
Lung Absess									
Lung Cancer									
Lymphangioleiomyomatosis		Low	High	Low					
Lymphatic Interstitial Pneumonitis		Low							
Mesothelioma									
Mountain Sickness									
Pertussis									
Pleural Effusion									
Pleurisy		Low	High			High			
Pneumoconiosis									
Pneumonia	Low		High			High			Positive
Pneumonitis									
Pneumothorax		Low							
Primary Cilia Dyskinesia									
Progressive Systemic Sclerosis	Low	Low							
Pulmonary Edema	Low	Low							
Pulmonary Embolism	Low	Low							
Pulmonary Fat Embolism	Low	Low							
Pulmonary Hypertension	Low	Low							
Pulmonary Parenchymal Injury				Low					
Respiratory Diptheria									
Sarcoidosis	Low	Low			High	High		Negative	
Silicosis		Low							
Systemic Lupus Erythematosis	Low	Low							
Tuberculosis		High							Positive
Wegener's Granulomatosis									

Table C: Clinical tests decision table (contd.).

	Serum				PFT				
	CPK	Amylase	Calcium	LDH	TLC	RV	FEV1	FVC	FEV1/FVC
Acute Bronchitis									
Adult Respiratory Distress Syndrome (ARDS)					High	High	Low	Low	Low
Allergic Bronchopulmonary Aspergillosis									
Alpha-1 related Emphysema					Low		Low	Low	High
Asbestosis					High	High	Low	Low	Low
Aspergilloma									
Aspiration Pneumonitis					High	High	Low		Low
Asthma									
Atelectasis									
Blastomycosis							High	High	High
Bronchiectasis					Low				
Bronchiolitis						Low	Low	High	High
Bronchitis Obliterans						High	Low	Low	Low
Chronic Bronchitis									
Chronic Cor Pulmonale					High		Low	Low	Low
Chronic Necrotizing Pulmonary Aspergillosis									
Croup									
Coccidioidomycosis									
Common Cold						High	Low	Low	Low
Cystic Fibrosis		Low							
Diphtheria					High	High	Low	Low	Low
Emphysema									
Eosinophilic Pneumonias									
Gaucher Disease									
Goodpasture Syndrome									
Hantavirus Pulmonary Syndrome									
Hayfever									
Hemothorax									
Histoplasmosis				High	Low		Low	High	High
Idiopathic Pulmonary Fibrosis									
Influenza									

Table C: Clinical tests decision table (contd.).

	Serum				PFT				
	CPK	Amylase	Calcium	LDH	TLC	RV	FEV1	FVC	FEV1/FVC
Kaposi's Sarcoma					Low	Low	Low	High	High
Khyphoscoliosis									
Laryngitis									
Lung Absess									
Lung Cancer					High	High	Low	Low	Low
Lymphangioliomyomatosis					Low	Low			
Lymphatic Interstitial Pneumonitis									
Mesothelioma									
Mountain Sickness									
Pertussis									
Pleural Effusion									
Pleurisy					High	High	Low	High	High
Pneumoconiosis									
Pneumonia				High					
Pneumonitis					Low		Low	High	High
Pneumothorax									
Primary Cilia Dyskinesia					Low	Low	Low	High	High
Progressive Systemic Sclerosis									
Pulmonary Edema	High	High							
Pulmonary Embolism									
Pulmonary Fat Embolism									
Pulmonary Hypertension									
Pulmonary Parenchymal Injury									
Respiratory Diptheria					Low	Low	Low	High	High
Sarcoidosis			High			High	Low		
Silicosis					Low	Low	Low	High	High
Systemic Lupus Erythematosis									
Tuberculosis									
Wegener's Granulomatosis									

Table C: Clinical tests decision table (contd.).

	Sputum Test	
	Color	Culture
Acute Bronchitis	Dark yellow/green	Bacterial/ viral / fungal growth
Adult Respiratory Distress Syndrome (ARDS)	Pink stained/ reddish	
Allergic Bronchopulmonary Aspergillosis		Bacterial/ viral / fungal growth
Alpha-1 related Emphysema		
Asbestosis		
Aspergilloma		Bacterial/ viral / fungal growth
Aspiration Pneumonitis	Pink stained/ reddish	
Asthma	White/light yellow	Eosinophilic granules
Atelectasis		
Blastomycosis		
Bronchiectasis	White/light yellow / dark yellow/ green	Bacterial/ viral / fungal growth/ white blood cells
Bronchiolitis		
Bronchitis		
Bronchitis Obliterans		
Chronic Bronchitis	Dark yellow/green/ pink stained/ reddish	Bacterial/ viral / fungal growth
Chronic Cor Pulmonale		
Chronic Necrotizing Pulmonary Aspergillosis		Bacterial/ viral / fungal growth
Croup		
Coccidioidomycosis		Bacterial/ viral / fungal growth
Common Cold		
Cystic Fibrosis		
Diphtheria		Bacterial/ viral / fungal growth
Emphysema		
Eosinophilic Pneumonias		Bacterial/ viral / fungal growth
Gaucher Disease		
Goodpasture Syndrome		
Hantavirus Pulmonary Syndrome		
Hayfever		
Hemothorax		
Histoplasmosis		Bacterial/ viral / fungal growth
Idiopathic Pulmonary Fibrosis		

Table C: Clinical tests decision table (contd.).

	Sputum Test	
	Color	Culture
Influenza		Bacterial/ viral / fungal growth
Kaposi's Sarcoma		
Khyphoscoliosis		
Laryngitis		
Lung Absess	Dark yellow/green	Pus cells
Lung Cancer		Cancer cells
Lymphangioleiomyomatosis		
Lymphatic Interstitial Pneumonitis		
Mesothelioma		
Mountain Sickness		
Pertussis		
Pleural Effusion		
Pleurisy		
Pneumoconiosis		
Pneumonia	Dark yellow/green/ pink stained/reddish	Bacterial/ viral / fungal growth
Pneumonitis		
Pneumothorax		
Primary Cilia Dyskinesia		
Progressive Systemic Sclerosis		
Pulmonary Edema	Pink stained/ reddish	
Pulmonary Embolism		
Pulmonary Fat Embolism		
Pulmonary Hypertension		
Pulmonary Parenchymal Injury		
Respiratory Diptheria		
Sarcoidosis		
Silicosis		Black materials
Systemic Lupus Erythematosis		
Tuberculosis		
Wegener's Granulomatosis		

Table C: Clinical tests decision table (contd.).

	Auscultation					
	Ronchi	Bronchial sound	Rales	Friction rub	Egophony	Decreased breath sounds
Acute Bronchitis	√		√			
Adult Respiratory Distress Syndrome (ARDS)			√			
Allergic Bronchopulmonary Aspergillosis		√	√		√	
Alpha-1 related Emphysema	√					
Asbestosis						√
Aspergilloma		√	√			
Aspiration Pneumonitis	√		√			
Asthma	√					√
Atelectasis			√			√
Blastomycosis			√			√
Bronchiectasis	√		√			√
Bronchiolitis	√					√
Bronchitis						
Bronchitis Obliterans			√			
Chronic Bronchitis	√		√			
Chronic Cor Pulmonale						
Chronic Necrotizing Pulmonary Aspergillosis		√	√			
Croup						
Coccidioidomycosis						
Common Cold						
Cystic Fibrosis	√		√			
Diphtheria						
Emphysema	√					
Eosinophilic Pneumonias						
Gaucher Disease						
Goodpasture Syndrome						
Hantavirus Pulmonary Syndrome						
Hayfever	√					
Hemothorax						
Histoplasmosis						

Table C: Clinical tests decision table (contd.).

	Auscultation					
	Ronchi	Bronchial sound	Rales	Friction rub	Egophony	Decreased breath sounds
Idiopathic Pulmonary Fibrosis			√			
Influenza						
Kaposi's Sarcoma						
Khyphoscoliosis						
Laryngitis						
Lung Absess			√			√
Lung Cancer	√		√			√
Lymphangioleiomyomatosis			√			
Lymphatic Interstitial Pneumonitis			√			
Mesothelioma						√
Mountain Sickness			√			
Pertussis						
Pleural Effusion				√		√
Pleurisy				√		√
Pneumoconiosis						
Pneumonia		√		√		√
Pneumonitis						
Pneumothorax						√
Primary Cilia Dyskinesia	√		√			
Progressive Systemic Sclerosis						
Pulmonary Edema	√		√			
Pulmonary Embolism	√		√	√		√
Pulmonary Fat Embolism	√		√			
Pulmonary Hypertension						
Pulmonary Parenchymal Injury			√			
Respiratory Diptheria						
Sarcoidosis	√		√			
Silicosis						
Systemic Lupus Erythematosis						
Tuberculosis			√			√
Wegener's Granulomatosis						

Table C: Clinical tests decision table (contd.).

	Auscultation			Blood Pressure	Percussion
	Increased breath sounds	Abnormal heart sounds	Vocal femitus		
Acute Bronchitis					
Adult Respiratory Distress Syndrome (ARDS)				Low	
Allergic Bronchopulmonary Aspergillosis					
Alpha-1 related Emphysema					
Asbestosis					Dullness
Aspergilloma					
Aspiration Pneumonitis					
Asthma					Hyperresonance
Atelectasis				Low	Dullness
Blastomycosis					Dullness
Bronchiectasis					
Bronchiolitis					Hyperresonance
Bronchitis					
Bronchitis Obliterans					
Chronic Bronchitis					
Chronic Cor Pulmonale					
Chronic Necrotizing Pulmonary Aspergillosis					
Croup					
Coccidioidomycosis					
Common Cold					
Cystic Fibrosis					Hyperresonance
Diphtheria					
Emphysema					Hyperresonance
Eosinophilic Pneumonias					
Gaucher Disease					
Goodpasture Syndrome					
Hantavirus Pulmonary Syndrome					
Hayfever					
Hemothorax					
Histoplasmosis					

Table C: Clinical tests decision table (contd.).

	Auscultation			Blood Pressure	Percussion
	Increased	Abnormal	Vocal femitus		
	breath sounds	heart sounds			
Idiopathic Pulmonary Fibrosis	√				
Influenza					
Kaposi's Sarcoma					
Khyphoscoliosis					
Laryngitis					
Lung Absess					
Lung Cancer					
Lymphangioleiomyomatosis					
Lymphatic Interstitial Pneumonitis					
Mesothelioma					Dullness
Mountain Sickness					
Pertussis					
Pleural Effusion			√		Dullness
Pleurisy					
Pneumoconiosis					
Pneumonia					Dullness
Pneumonitis					
Pneumothorax					Hyperresonance
Primary Cilia Dyskinesia					
Progressive Systemic Sclerosis					
Pulmonary Edema		√		Low	
Pulmonary Embolism		√		Low	
Pulmonary Fat Embolism					
Pulmonary Hypertension				High	
Pulmonary Parenchymal Injury					
Respiratory Diptheria					
Sarcoidosis					
Silicosis					
Systemic Lupus Erythematosis					
Tuberculosis					Dullness
Wegener's Granulomatosis					

Table 4: Radiological findings decision table.

	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	R14	R15	R16	R17
Acute Bronchitis																	
Adult Respiratory Distress Syndrome (ARDS)				√	√												
Allergic Bronchopulmonary Aspergillosis																	
Alpha-1 related Emphysema	√										√				√		
Asbestosis					√			√							√	√	√
Aspergilloma																	
Aspiration Pneumonitis															√	√	
Asthma	√	√	√	√	√											√	
Atelectasis			√	√	√					√		√		√			
Blastomycosis			√	√													
Bronchiectasis																√	√
Bronchiolitis	√																
Bronchitis Obliterans				√	√												
Chronic Bronchitis															√		
Chronic Cor Pulmonale																	
Chronic Necrotizing Pulmonary Aspergillosis																	
Croup																	
Coccidioidomycosis									√				√				
Common Cold																	
Cystic Fibrosis				√	√											√	
Diphtheria																	
Emphysema	√										√				√		
Eosinophilic Pneumonias																	
Gaucher Disease																	
Goodpasture Syndrome				√													
Hantavirus Pulmonary Syndrome																	
Hayfever																	
Hemothorax																	
Histoplasmosis									√								
Idiopathic Pulmonary Fibrosis				√	√												√
Influenza																	
Kaposi's Sarcoma					√												
Khyphoscoliosis																	

Table 4: Radiological findings decision table (contd.).

	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	R14	R15	R16	R17
Laryngitis																	
Lung Abscess																	
Lung Cancer		√			√	√	√										
Lymphangioleiomyomatosis																	
Lymphatic Interstitial Pneumonitis					√												√
Mesothelioma																	
Mountain Sickness																	
Pertussis																	
Pleural Effusion			√	√										√			
Pleurisy																	
Pneumoconiosis				√	√		√										
Pneumonia			√	√	√												
Pneumonitis																	
Pneumothorax			√	√										√			
Primary Cilia Dyskinesia																	
Progressive Systemic Sclerosis		√			√												√
Pulmonary Edema				√	√			√							√	√	
Pulmonary Embolism		√										√			√		
Pulmonary Fat Embolism				√													
Pulmonary Hypertension																	
Pulmonary Parenchymal Injury		√	√		√		√										
Respiratory Diphtheria																	
Sarcoidosis		√			√				√				√				
Silicosis			√														
Systemic Lupus Erythematosus												√					
Tuberculosis				√	√		√		√				√				
Wegener's Granulomatosis																	

Table 4: Radiological findings decision table (contd.).

	R18	R19	R20	R21	R22	R23	R24	R25	R26	R27	R28	R29	R30	R31	R32	R33	R34
Acute Bronchitis																	
Adult Respiratory Distress Syndrome (ARDS)																	
Allergic Bronchopulmonary Aspergillosis														√			
Alpha-1 related Emphysema					√												
Asbestosis		√					√										
Aspergilloma														√			
Aspiration Pneumonitis																	
Asthma																	
Atelectasis											√						
Blastomycosis																	
Bronchiectasis								√								√	
Bronchiolitis																	
Bronchitis Obliterans																	
Chronic Bronchitis																	
Chronic Cor Pulmonale																	
Chronic Necrotizing Pulmonary Aspergillosis								√		√				√			
Croup																	
Coccidioidomycosis						√			√	√							
Common Cold																	
Cystic Fibrosis					√										√		
Diphtheria																	
Emphysema					√												
Eosinophilic Pneumonias																	
Gaucher Disease																	
Goodpasture Syndrome												√					
Hantavirus Pulmonary Syndrome																	
Hayfever																	
Hemothorax																	
Histoplasmosis												√	√				
Idiopathic Pulmonary Fibrosis	√																
Influenza																	
Kaposi's Sarcoma						√											
Khyphoscoliosis																	

Table 4: Radiological findings decision table (contd.).

	R18	R19	R20	R21	R22	R23	R24	R25	R26	R27	R28	R29	R30	R31	R32	R33	R34
Laryngitis									√						√		
Lung Abscess										√							
Lung Cancer		√	√			√				√							
Lymphangioliomyomatosis						√											
Lymphatic Interstitial Pneumonitis								√									
Mesothelioma			√				√										√
Mountain Sickness																	√
Pertussis																	
Pleural Effusion																	√
Pleurisy						√											√
Pneumoconiosis				√					√	√							
Pneumonia						√											
Pneumonitis																	
Pneumothorax																	
Primary Cilia Dyskinesia																	
Progressive Systemic Sclerosis																	
Pulmonary Edema						√											
Pulmonary Embolism						√											
Pulmonary Fat Embolism																	
Pulmonary Hypertension																	
Pulmonary Parenchymal Injury																	
Respiratory Diptheria																	
Sarcoidosis																	
Silicosis		√		√													
Systemic Lupus Erythematosus						√											
Tuberculosis		√				√				√	√						
Wegener's Granulomatosis																	

Table 4: Radiological findings decision table (contd.).

	R35	R36	R37	R38	R39	R40	R41	R42	R43	R44	R45	R46	R47	R48	R49
Acute Bronchitis															
Adult Respiratory Distress Syndrome (ARDS)															
Allergic Bronchopulmonary Aspergillosis						√									
Alpha-1 related Emphysema			√	√			√								
Asbestosis		√													
Aspergilloma															
Aspiration Pneumonitis										√					
Asthma				√											
Atelectasis	√	√									√				
Blastomycosis															
Bronchiectasis	√				√										
Bronchiolitis															
Bronchitis Obliterans									√						
Chronic Bronchitis					√										
Chronic Cor Pulmonale												√			
Chronic Necrotizing Pulmonary Aspergillosis															
Croup															
Coccidioidomycosis															
Common Cold															
Cystic Fibrosis	√							√							
Diphtheria															
Emphysema			√	√			√								
Eosinophilic Pneumonias									√						
Gaucher Disease															
Goodpasture Syndrome															
Hantavirus Pulmonary Syndrome															
Hayfever															
Hemothorax															
Histoplasmosis															
Idiopathic Pulmonary Fibrosis		√													
Influenza															
Kaposi's Sarcoma															
Khyphoscoliosis															

Table 4: Radiological findings decision table (contd.).

	R35	R36	R37	R38	R39	R40	R41	R42	R43	R44	R45	R46	R47	R48	R49
Laryngitis															
Lung Abscess															
Lung Cancer															√
Lymphangioliomyomatosis															
Lymphatic Interstitial Pneumonitis		√													
Mesothelioma											√				
Mountain Sickness															
Pertussis															
Pleural Effusion															
Pleurisy															
Pneumoconiosis															
Pneumonia															
Pneumonitis															
Pneumothorax		√											√	√	
Primary Cilia Dyskinesia															
Progressive Systemic Sclerosis															
Pulmonary Edema															
Pulmonary Embolism	√	√													
Pulmonary Fat Embolism															
Pulmonary Hypertension												√			
Pulmonary Parenchymal Injury	√														
Respiratory Diphtheria															
Sarcoidosis															
Silicosis		√													
Systemic Lupus Erythematosus	√	√													
Tuberculosis															
Wegener's Granulomatosis															

Appendix B

The LUNEX system was tested with 60 test cases which were obtained from hospitals and internet and is listed in Table 5. The findings of the doctors and LUNEX are given for comparison. The LUNEX is capable of diagnosing more than one disease. However, here, only the first one or two diseases with the highest confidence factor are given.

Table 5: LUNEX diagnoses.

Case Studies	Doctor's diagnosis	LUNEX diagnosis
Case study 1	Tuberculosis	Tuberculosis
Case study 2	Pneumonia	Pneumonia
Case study 3	Lung Abscess	Lung Abscess
Case study 4	Aspergilloma	Aspergilloma, Blastomycosis
Case study 5	Pulmonary parenchymal injury	Pulmonary parenchymal injury , ARDS
Case study 6	Pneumothorax	Hemothorax
Case study 7	Pneumonia	Pneumonia
Case study 8	Lung cancer	Lung cancer
Case study 9	Lung cancer	Lung cancer
Case study 10	Lung cancer, Tuberculosis	Lung cancer, Tuberculosis
Case study 11	Lung cancer	Lung cancer
Case study 12	Lung cancer	Lung cancer
Case study 13	Pleural effusion	Pleural effusion
Case study 14	Chronic bronchitis	Acute bronchitis
Case study 15	Emphysema	Emphysema
Case study 16	Asthma	Asthma
Case study 17	Acute Bronchitis	Acute Bronchitis
Case study 18	Pneumonia	Pneumonia
Case study 19	Pulmonary hypertension	Pulmonary hypertension, Chronic cor pulmonale
Case study 20	Lung Abscess	Lung Abscess
Case study 21	Emphysema	Emphysema
Case study 22	Aspergilloma	Aspergilloma, Blastomycosis
Case study 23	Atelectasis	Atelectasis
Case study 24	ARDS	ARDS, Pulmonary parenchymal injury
Case study 25	Pleural effusion	Pleural effusion
Case study 26	Chronic Bronchitis	Chronic Bronchitis
Case study 27	Chronic cor pulmonale	Chronic cor pulmonale
Case study 28	Bronchiectasis	Chronic bronchitis
Case study 29	Pulmonary edema	Pulmonary edema
Case study 30	Pulmonary embolism	Pulmonary embolism

Table 5: LUNEX diagnoses (contd.).

Case Studies	Doctor's diagnosis	LUNEX diagnosis
Case study 31	Tuberculosis	Tuberculosis
Case study 32	Asthma	Asthma
Case study 33	Pneumothorax	Pneumothorax
Case study 34	Hemothorax	Pneumothorax
Case study 35	Laryngitis	Laryngitis
Case study 36	Influenza	Common cold
Case study 37	Bronchitis Obliterans	Bronchitis Obliterans
Case study 38	Pneumonoconiosis	Pneumonoconiosis
Case study 39	Sarcoidosis	Sarcoidosis
Case study 40	Pneumonitis	Pneumonitis
Case study 41	Pertussis	Pertussis
Case study 42	Kyhphoscoliosis	Kyhphoscoliosis
Case study 43	Silicosis	Silicosis
Case study 44	Chronic cor pulmonale	Chronic cor pulmonale, Pulmonary hypertension
Case study 45	Emphysema	Emphysema
Case study 46	Diphtheria	Diphtheria
Case study 47	Hayfever	Hayfever, Histoplasmosis
Case study 48	Idiopathic pulmonary fibrosis	Idiopathic pulmonary fibrosis
Case study 49	Atelectasis	Atelectasis
Case study 50	Blastomycosis	Aspergilloma, Blastomycosis
Case study 51	Pulmonary embolism	Pulmonary edema
Case study 52	Pleural effusion	Pleural effusion
Case study 53	Lung Abscess	Lung Abscess
Case study 54	Common cold	Influenza, Common cold
Case study 55	Acute bronchitis	Acute bronchitis
Case study 56	Asbestosis	Asbestosis
Case study 57	Silicosis	Silicosis
Case study 58	Influenza	Common cold
Case study 59	Lung cancer	Asbestosis
Case study 60	Lung cancer	Lung cancer