A Knowledge Based System for Diagnosis of Lung Diseases from Chest X-Ray Images

By

Zul Waker Mohammad Al-Kabir

The thesis is submitted in fulfilment of the requirements for the degree of Master of Information Science in the school of Information Sciences and Engineering under the division of Business, Law, and Sciences

At the

University of Canberra

May 2006

COPYRIGHT

I, Zul Waker Mohammad Al-Kabir, grant permission to University of Canberra, Australia to reproduce and to redistribute publicly paper and electronic copies of this thesis in whole or in part. Any reproduction and usage will not be for commercial use or for profit.

Printed in the University of Canberra
University Drive
Bruce ACT-2617
Australia

Acknowledgements

I would like to express my sincere gratitude to my supervisors Dr. Kim Le and Associate Prof. Dharmendra Sharma of the University of Canberra for their kindness and valuable advice, suggestions, feedback and encouragement during the research. I would also like to give my heartfelt thanks to AProf. John Campbell, AProf. Craig McDonald for their suggestions and feedback and for making me feel part of the research community. I am also grateful to Dr. Dat Tran for his invaluable support in the Mathematical areas.

I would like to thank Dr. Aslam Jaman and Ms. Rejowana Majid for their prompt reply and assistance in understanding and interpreting medical images and any clinical issues. I also thank all the medical people from Bangladesh and Thailand.

I would like to thank the Department of Education, Science and Training (DEST), the University of Canberra and School of Information Sciences and Engineering for funding me with a Research Training Scheme scholarship. I would also like to thank everyone who provided support for numerous workshops and other research activities.

I would also like to take the opportunity to acknowledge Ms. Sue Prentice for editing the thesis prior to submission.

Finally, I thank my departed mother, my still going energetic father, my sisters with whom I left all my frustrations for a good night sleep. I also thank other friends for their cordial support during the research.

Abstract

The thesis develops a model (that includes a conceptual framework and an implementation) for analysing and classifying traditional X-ray images (MACXI) according to the severity of diseases as a Computer-Aided-Diagnosis tool with three initial objectives.

- The first objective was to interpret X-ray images by transferring expert knowledge into a knowledge base (CXKB): to help medical staff to concentrate only on the interest areas of the images.

- The second objective was to analyse and classify X-ray images according to the severity of diseases through the knowledge base equipped with an image processor (CXIP).

- The third objective was to demonstrate the effectiveness and limitations of several image-processing techniques for analysing traditional chest X-ray images.

A database was formed based on collection of expert diagnosis details for lung images. Five important features from lung images, as well as diagnosis rules were identified and simplified. The expert knowledge was transformed into a Knowledge base (KB) for analysing and classifying traditional X-ray images according to the severity of diseases (CXKB). Finally, an image processor named CXIP was developed to extract the features of lung images features and image classification.

CXKB contains 63 distinct lung diseases with detailed descriptions. Some 80-chest X-ray images with diagnosis details were collected for the database from different sources,

including online medical resources. A total of 61 images were used to determine the important features; 19 chest X-ray images were not used because of low visibility or the difficulty of diagnosis. Finally, only 12 images were selected after examining the diagnosis details, image clarity, image completeness, and image orientation. The most important features of lung diseases are a pattern of lesions with different levels of intensity or brightness. The other major anatomical structures of the chest are the hilum area, the rib area, the trachea area, and the heart area.

Seven different severity levels of diseases were determined. Development and simplification of rules based on the image library were analysed, developed, and tested against the 12 images. A level of severity was labelled for each image based on a personal understanding of all the image and diagnosis details. Then, MACXI processed the selected 12 images to determine the level of severity. These 12 images were fed into the CXIP for recognition of the features and classification of the images to an accurate level of severity. Currently, the processor has the ability to identify diseased lung areas with approximately 80% success rate.

A step by step demonstration of several image processing techniques that were used to build the processor is given to highlight the effectiveness and limitations of the techniques for analysing traditional chest X-ray images is also presented.

Contents

List of Figures

List of Tables

List of Abbreviations

AI	Artificial Intelligence
ANN	Artificial Neural Network
CAD	Computer Aided Diagnosis
CRT	Cathode Ray Tube
CT	Computed Tomography
CXR	Chest X-ray
FL	Fuzzy Logic
HLND	Hybrid Lung Nodule Detection system
IDL	Interactive Data Language
IS	Information Systems
KB	Knowledge Base
MATLAB	Matrix Laboratory
MACXI	Model for Analysing and Classifying Traditional Chest X-ray Images
CXIP	Image processor for Analysing and Classifying Traditional Chest X-ray Images
CXKB	Knowledge base for Analysing and Classifying Traditional Chest X-ray Images
MRI	Magnetic Resonance Imaging
MSAccess	Microsoft Access
MSVB	Microsoft Visual Basic 6.0
OS	Operating System
PA	Posterior Anterior
ROC	Receiver Operating Characteristic
ROI	Region of Interest

RSNA	Radiological Society of North America
SD	Systems Development
SDLC	Systems Development Life Cycle
SDRM	Systems Development Research Method

Glossary

Throughout the thesis, new items were highlighted when first encountered. Here, the most significant of these terms are presented with brief definitions.

Computed Tomography (CT)

Computed tomography is a tomographic image acquisition system using X-ray transmissions for gathering cross sectional slice-image from projection images: used primarily in medical imaging applications (Baxes, 1994)

Computer Aided Diagnosis (CAD)

Computer Aided Diagnosis is a computer-based automated tool to diagnose medical images to a certain extent.

Fluoroscopy

Fluoroscopy is based on the same techniques as traditional X-ray, with the photographic plate replaced by a fluorescent screen (Columbia University Press, 2005).

IDL

IDL is a programming application.

Image Analysis

The processing of an image to extract quantitative object measurements and then classify the results (Baxes, 1994).

Magnetic Resonance Imaging (MRI)

A tomographic image acquisition system using magnetic excitation for gathering cross sectional slice images from projection images; used primarily in medical imaging applications (Baxes, 1994). MRI is a non-invasive procedure that uses powerful magnets and radio waves to construct pictures of the body.

MATLAB

MATLAB stands for MATrix LABoratory. It is a software tool for a range of engineering works.

Morphological Process

A morphological process is a group of processes that evaluates each pixel in a binary or gray scale image along with its neighbouring pixels. Resulting pixel brightness is determined by looking at the input pixel brightness patterns (binary image case) or minimum and maximum values (gray-scale image case) (Baxes 1994).

Traditional Chest X-ray Image (CXR)

Traditional Chest X-ray Image is A stream of photons, that has a penetrating power, is passed through any organ of human body (e.g.-lung) to produce a photograph on a plastic photographic plate (Columbia University Press, 2005).

X-ray

X-ray is a relatively high-energy photon having a wavelength in the approximate range from 0.01 to 10 nanometres (Columbia University Press, 2005)