

Detection and Extraction of Tuberculosis from Chest CT Slices

Anita Titus

Assistant Professor

*Department of Electronics and Communication
Easwari Engineering College, Chennai*

P. Kannan

PG Student

*Department of Electronics and Communication
Easwari Engineering College, Chennai*

Abstract

Tuberculosis (TB) is a deadly infectious disease that mostly affects the lungs, but can also affect the other organs, including central nervous system. The basic key to control the spread of TB is the early detection and treatment. A computer aided diagnosis (CAD) system is proposed to detect and classify TB. The chest CT slices are initially preprocessed to remove the gaussian noise followed by lung segmentation and then with the features extracted from the ROIs the Bayesian classifier is been trained and finally the slices are classified. A set of 30 CT(computed tomography) slices were processed that demonstrates the accuracy and satisfactory performance of the algorithm. The CAD system achieved an accuracy of 73.3%, sensitivity of 88.6% and specificity of 78.33%.

Keywords: Computer Aided Diagnosis (CAD), Tuberculosis, Computed Tomography, Segmentation, Classification

I. INTRODUCTION

Tuberculosis is an infectious disease caused by bacillus mycobacterium that can spread through the lymph nodes and bloodstream to any organ in body. Generally the lungs are affected when TB disease becomes active. In 15-20% of active cases, the infection spreads over the outer part of the lungs, causing the other kinds of tuberculosis known as extra pulmonary TB. CAD system can pave the way for a automated diagnostic interpretation of chest radiographs that reduces the work of a radiologist and provides diagnosis result with more accuracy. The main objective of the CAD system is to make the patient's diagnosis faster, easier and more efficient. Various types of CAD schemes are being developed for detection and/or characterization of various lesions for variety of imaging modalities including computer tomography, magnetic resonance imaging (MRI). The organs currently being targeted by CAD research include breast, chest, colon, brain, liver, kidney etc. As segmentation of medical images is often a precursor to image analysis applications, its accuracy is of great concern as segmentation errors may potentially lead to misdiagnosis. Several methods have been implemented to segment the lung in pulmonary CT slices such as watershed, active contour, thresholding and region growing. The drawbacks of these techniques are described below.

Watershed algorithm gives an acceptable segmentation result with relatively low computational cost [1]. However, over-segmentation is a well-known drawback in watershed segmentation. Snake is an active contour that starts from an initial position and shape and fits itself to the shape of the desired objects [2]. The drawback of traditional snake model is that the construction of initial contour often requires human interaction and the segmentation results may be heavily sensitive to initial contour conditions [4]. Thresholding algorithm is simple in methodology and computation [5]. However, it has some drawbacks in lung segmentation that the lung density is influenced by factors such as air volume, image acquisition protocol and degree of inspiration segmentation and the scheme works well for only healthy or mildly diseased lungs. [3]. Although some work has used an adaptive threshold value, this method still has some drawbacks on its segmentation results such as missing boundaries between 2 regions when there is not a significant discontinuity between the boundaries or the borders are too close together [8]. For example, this method requires additional post-processing steps to eliminate the trachea and main stem bronchi [7]. The main drawback of region growing algorithm approach is its computational intensity as well as its performance dependence on its initial seed [6]. The limitation in a hybrid knowledge-based framework [10] is that the TB is extracted by manual operation. To overcome this problem the proposed system exhibits several innovative works to automate the detection and classification of tuberculosis.

II. CAD SYSTEM ARCHITECTURE

The CAD system involves the processing of two sets of lung CT slices – the training dataset and the testing dataset. The training dataset is the set of CT slices diagnosed by radiologists as slices affected with tuberculosis or normal. The testing dataset is the set of CT slices that have to be classified by the CAD system. The processes involved in this system are preprocessing, segmentation of lungs, ROI extraction, feature extraction and classification as shown in the Fig.1.

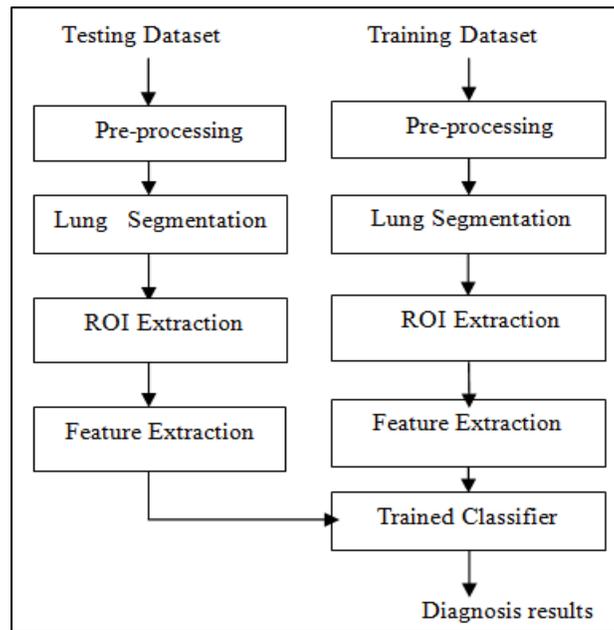


Fig. 1: Architecture of the Proposed CAD System

The chest CT slices are initially preprocessed to remove the gaussian noise by using a Weiner filter followed by lung segmentation using fuzzy clustering then with the features extracted from the ROIs the Bayesian classifier is been trained and finally the slices are classified as cavitory TB or as normal.

A. Preprocessing:

The input chest CT slice is preprocessed to remove the Gaussian noise using a wiener filter of window size 3x3 and the performance is evaluated in terms of peak signal to noise ratio (PSNR). The PSNR estimates the quality of the reconstructed image with respect to the original image. The reconstructed images with higher PSNR are better.

Input: Chest CT slice (in JPEG format).

Process

- 1) Step 1: Resize the input chest CT slice to 256 x 256.
- 2) Step 2: Convert the resized slice into a gray scale slice.
- 3) Step3: Apply the wiener filter of window size 3x3 suppress the Gaussian noise in the gray scale slice [5].

Output: Preprocessed chest CT slice.

B. Segmentation:

Input: Preprocessed chest CT slice.

Process

- 1) Step 1: Threshold the preprocessed slice to a binary slice.
- 2) Step 2: Remove the background region of the input CT slice.
- 3) Step 3: Remove the border artifacts.
- 4) Step 4: Remove the unwanted objects (<100 pixel size).
- 5) Step 5: Fill the holes to obtain the mask image.
- 6) Step 6: Multiply the mask image with the preprocessed chest CT slice. Output: Segmented lung slice.

C. Roi Extraction:

Input : Segmented lung slice.

Process:

- 1) Step 1: Create a mask from the input gray scale
- 2) Step 2: Apply the mask to the input slice to retain ROI and set all else to 0.
- 3) Step 3: Convert the masked image into a black and white with grayscale threshold.
- 4) Step 4: Mark and trace the boundary of the region. Output: Extracted region of interest (ROI)

D. Feature Extraction:

The features are computed from the extracted region of interest (ROI) by using scale invariant feature transform (SIFT) descriptor.

Input: Extracted region of interest (ROI)

Process

- 1) Step 1: Compute the 'area' feature that is the number of pixels in the ROI region extracted.
- 2) Step 2: Compute the 'Mean' feature that is the measure of intensity level of a pixel.

$$\text{Mean} = \frac{\sum_{i=0}^{n-1} x_i p(x_i)}{n} \quad (1)$$

Where,

- x_i - a random variable that specifies the intensity.
- $p(x_i)$ - the histogram of the intensity level in the region.
- n - the number of possible intensity levels.

- 3) Step 3: Compute the 'Entropy' feature that is the measure of uncertainty.
- 4) Step 4: Compute the 'Energy' feature which is also known as uniformity of ASM (angular second order moment) that is the sum of mean squared signal value.

Output: Extracted features arranged as a feature vector.

E. Classification:

Texture features are computed from the extracted region of interest (ROI) by scale invariant feature transform(SIFT) descriptor . With the features extracted from the ROIs the Bayesian classifier is been trained and finally the slices are classified. The output of the classifier indicates whether the slice has been classified as normal lung or as cavitory TB.

III. RESULTS AND DISCUSSIONS

The proposed CAD system was tested on a dataset consisting of 30 CT slices that comprised of both normal and cavitory TB slices. The training data set consisted of 18 TB slices and 12 normal slices. The input chest CT slices are converted into a grayscale slice. Weiner filter is applied to the gray scale slices to remove the Gaussian noise. Then the Gaussian filtered gray scale slice is converted into a binary image and the lung borders are cleared, holes are filled and small connected components are removed as shown in the Fig. 2.

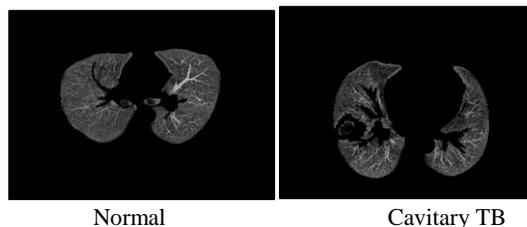


Fig. 2: Lung segmentation

The region of interest (ROI) is extracted from the segmented lung as shown in the Fig. 3 by using boundary tracing algorithm and the features extracted from the ROI are fed as input to the Bayesian network classifier. The outputs of the classifier will indicate whether the slice has been classified as normal lung or cavitory TB.

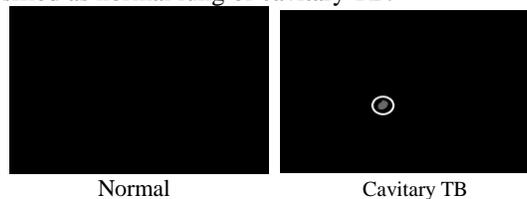


Fig. 3: ROI Extraction

Texture features are computed from the extracted region of interest (ROI) by using scale invariant feature transform (SIFT) descriptor. With the features extracted from the ROIs the Bayesian classifier is trained and finally the slices are classified. The output of the classifier indicates whether the slice has been classified as normal lung or as cavitory TB.

Tests can be positive or negative, True positive (TP) refers to the slices that test positive and have the disease. False Positive (FP) refers to the slices that test positive but do not have the disease. True negative (TN) refers to the slices that test negative and do not have the disease. False negative (FN) refers to the slices that test negative but have the disease. The performance metrics derived to check the classifier performance are mentioned below

$$\begin{aligned}\text{Accuracy} &= \frac{(\text{TN}+\text{TP})}{(\text{TN}+\text{TP}+\text{FN}+\text{FP})} \\ \text{Sensitivity} &= \frac{\text{TP}}{(\text{TP} + \text{FN})} \\ \text{Specificity} &= \frac{\text{TN}}{(\text{FP}+\text{TN})}\end{aligned}$$

IV. CONCLUSIONS

A CAD system is proposed to detect and classify tuberculosis (TB). In this work the chest CT slices were initially preprocessed to remove the gaussian noise followed by lung segmentation and then the boundary tracing algorithm is used to extract the region of interest (ROI) and with the features extracted from the ROIs the Bayesian classifier is trained and finally the images were classified. A set of 30 CT (computed tomography) slices were processed that demonstrates the accuracy and satisfactory performance of the algorithm. The proposed CAD system has achieved an accuracy of 73.3%, sensitivity of 88.6% and specificity of 78.33%.

REFERENCES

- [1] Almas Pathan, Bairu. K. saptalkar, (2010), 'Detection and Classification of Lung Cancer Using Artificial Neural Network', IEEE Trans . On Med. Image., Vol. 25, No. 4.
- [2] Bianca Lassen, Eva M. van Rikxoort, Michael Schmidt, Sjoerd Kerkstra, Bram van Ginneken, and Jan-Martin Kuhnigk, (2013), 'Automatic Segmentation of the Pulmonary Lobes From Chest CT Scans Based on Fissures, and Broanchi', IEEE Trans. On Med. Image., Vol. 32, No.4.
- [3] Deniz Aykac, Eric A. Hoffman, (2009) 'Segmentation and Analysis of the Human Airway Tree From Three-Dimensional X-Ray CT Images', IEEE transactions on medical imaging, Vol. 22, pp.no. 1434-1450.
- [4] Eva van Rikxoort, Bartjan de Hoop, Saskia van de Vors, (2010) 'Automatic Segmentation of Pulmonary Segments From Volumetric Chest CT Scans', IEEE transactions on medical imaging, Vol. 2, pp.no.22-34.
- [5] EzhilSwanly, L. Selvam, P. Mohan Kumar, J. Arokia Renjith, M. Arunachalm, K.L. Shanmuganathan, (2013) 'Smart Spotting of Pulmonary TB Cavities Using CT Images', Vol. 2, pp.no.52-67.
- [6] Ingrid Sluimer, Arnold Schilham, Mathias Prokop, Bram van Ginneken, (2011), 'Computer Analysis of Computed Tomography Scans of the Lung: A Survey' IEEE Trans. On Med. Image., Vol. 25 No. 4.
- [7] JinsaKuruville, K.Gunavathi,(2014, ' Lung Cancer Classification using Neural Networks' Elsevier-Computer methods & programs in Biomedicine., Vol.25 No. 4.
- [8] Lindsay Semler, Lucia Dettori, Jacob Furst, (2010), 'Wavelet-based Texture Classification of Tissues in computed Tomography' 18th IEEE Symposium on Computer-Based Medical Systems.
- [9] Murat Ceylan, Yuksel Ozbay, O.NuriUcan, Erkan Yildirim, (2010), 'A novel method for lung segmentation on chest CT images complex-valued artificial neural network with complex wavelet transform' Turk J Elec. Eng. & Comp Sci Vol. 18, No.4
- [10] Sudha. V, Jayashree.P. (2012, 'A hybrid knowledge-guided detection technique for screening of infectious pulmonary tuberculosis from chest radiographs', International Journal of Emerging Sci. and Eng. (IJESE) ISSN: 2319-6378, Vol.1. pp.no. 1434-1450