

CXR Tuberculosis Detection Using MATLAB Image Processing

Mr. P. A. Kamble¹, Mr. V. V. Anagire², Mr. S. N. Chamtagoudar³

¹Asst. Prof., Dept. of E&TC Engineering, Walchand Institute of Technology, Solapur, Maharashtra, India

²Asst. Prof., Dept. of E&TC Engineering, Walchand Institute of Technology, Solapur, Maharashtra, India

³Asst. Prof., Dept. of E&TC Engineering, Walchand Institute of Technology, Solapur, Maharashtra, India

Abstract - Tuberculosis (TB) is very dangerous and rapidly spread disease in the world. When left undiagnosed and thus untreated, mortality rates of patients with tuberculosis are high. Standard diagnostics still rely on methods developed in the last century. They are slow and often unreliable. In the investigating cases for suspected tuberculosis (TB), chest radiography is not only the key techniques of diagnosis based on the medical imaging but also the diagnostic radiology. So, Computer aided diagnosis (CAD) has been popular and many researchers are interested in this research areas and different approaches have been proposed for the TB detection and lung disease classification. In this paper we present method for detection of Tuberculosis in CXR image by using MATLAB which includes Pre Processing of Image, Segmentation and Feature extraction from that image.

Key Words: CXR, MATLAB, tuberculosis, Segmentation, Feature extraction.

1.INTRODUCTION

Tuberculosis (TB) is one of the most important public health problems worldwide. There are 9 million new TB cases and nearly 2 million TB deaths each year [1]. Case-finding and the management of pulmonary tuberculosis is an essential target of tuberculosis control programs. However, pulmonary tuberculosis (PT) is becoming more and more of a serious problem, particularly in countries affected by epidemics of human immunodeficiency virus (HIV)-TB co-infection [2]. The diagnosis of PT using prompt and accurate methods is a crucial step in the control of the occurrence and prevalence of TB. However, the diagnosis of PT is quite complex, so there is no unified standard at present.

Image Processing is a form of signal processing where images and their properties can be used to gather and analyze information about the objects in the image. Digital image processing uses digital images and computer algorithms to enhance, manipulate or transform images to obtain the necessary information and make decisions accordingly.

Examples of digital image processing include improvements and analysis of the images of the Surveyor missions to the moon [3], magnetic resonance imaging scans of the brain and electronic face recognition packages. These techniques

can be used to assist humans with complex tasks and make them easier. A detailed analysis of an X-ray can help a radiologist.

Large number of patients with TB infections needs to be X rayed and screen for active TB to ensure a proper treatment of their infections. Taking Standard Chest X-rays (CXR) is an inexpensive way to screen for the presence of TB. The purpose of screening system is to identify everything that is or could be related to a patient having TB infections. But mass screening of a large population is a time consuming and tedious work, which require considerable effort when done manually. For this reason, Computer-aided diagnostic systems (CAD) used to detect Tuberculosis infections in chest X rayed. These systems have the potential to lessen the TB detection error risk and also depend on the radiologists [4, 5].

Computer aided detection diagnosis is also known as computer aided diagnosis which is a medical image processing for diagnosing the aid of radiology images. The goal of CAD is not only to get the better result of the diagnostic accuracy but also the consistent radiology image interpretation by using the computer output.

CAD output is able to help the radiologists can diagnosis and provide treatment based on the subjective judgment. Normally, there are two types of approaches are applied in computerized schemes for CAD which are finding the location of lesions and quantifying the features of radiography images whether normal or abnormal patterns are included in lungs. Other medical imaging techniques such as X-ray, MRI, ultrasound, etc. produce the information and aid to analyze the detection of tumor in the short time duration.

In this paper, we describe how we differentiate between normal and abnormal CXRs with manifestations of TB, using image processing techniques in MATLAB.

2.METHODOLOGY

A typical digital image processing system consists of image pre-processing, segmentation, feature selection and extraction, and classification. The image needs to be reduced to certain determining characteristics and the analysis of these characteristics gives the relevant information. Figure 1 shows a block diagram of a typical digital image processing system, showing the sequence of the operations.

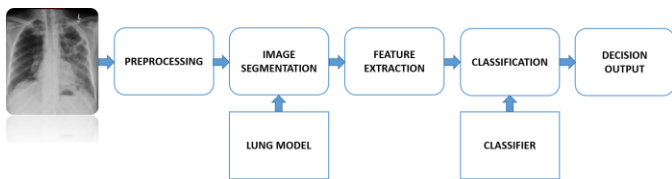


Figure 1 Block diagram of a typical digital image processing system

2.1 Preprocessing

The main purpose of the image preprocessing is not only to improve the quality of image but also to reduce the undesired portion from the background of the images. Most of preprocessing methods applied the intensity value of neighborhood pixel for obtaining the brightness intensity value of the input images.

There are dominant reasons for the requirement processes from the image preprocessing phase:

- *To meet the requirements of physician for the image quality improvement*
- *Noise reduction*
- *Contrast enhancement*
- *Correction of missing or wrong pixel values*
- *Optimal preparation of data for postprocessing*
- *Elimination of acquisition-specific artifacts*

2.2 Segmentation

The second stage of the system is detecting the lung boundary of x-ray images using image properties and the lung shape model from the previous section.

Among the preprocessing processes, image segmentation is very important to extract good quality features for classification. The key functions of image segmentation are clustering pixels having same intensity value from the whole image regions, separating regions or objects of desired part of the original image, hiding the undesired region or surfaces. Image segmentation are applied in many image preprocessing areas such as recognition of object, object occlusion, estimation of the boundary, editing, or query from image database, image security and image compression.

In general, algorithms for image segmentation are based on two fundamental properties of intensity values which are discontinuity and similarity. Discontinuity property does a partition the image based on the sharp intensity value changes. Similarity property does a partition the image into regions that are like in order to the specified criteria.

2.3 Feature Extraction

Features are the representatives of the images. When the input data to an algorithm is too large to be processed then

the input data will be transformed into a reduced representation set of features. Transforming the input data into the set of features is called feature extraction. If the features extracted are carefully chosen it is expected that the features set will extract the relevant information from the input data in order to perform the desired task using this reduced representation instead of the full size input [6].

2.4 Classification

After the feature extraction stage, the features have to be analyzed, classifiers like Bayesian classifiers and Support Vector Machines (SVM) are used to classify features and find the best values for them.

An SVM in its original form is a supervised non probabilistic classifier that generates hyperplanes to separate samples from two different classes in a space with possibly infinite dimension [7, 8].

The unique characteristic of an SVM is that it does so by computing the hyperplane with the largest margin; i.e. the hyperplane with the largest distance to the nearest training data point of any class.

3. IMPLEMENTATION

The block diagram of the proposed setup is as shown in Figure 1. Initially the system is trained with features with more number of X-ray images (data base).

In figure 1, we observe that the data i.e. the image is read and are preprocessed. Preprocessing includes histogram equalization and filtering of image using Gaussian Filter which smooth's image by removing noise. It will prepare the x-ray for the next two-process segmentation and feature extraction.

Active contour (AC) models (or snakes) [9, 10] are commonly used as segmentation techniques for medical images. Active Contour Models are used to extract femur contours in X-ray images, after doing edge detection on the image using a modified Canny filter.

Active Contour Model is susceptible to noise and other undesired edges. This method successfully extracts the femur contour with a small restriction on shape, size and orientation of the image.

To describe normal and abnormal patterns in the segmented lung field, we use a group of low-level features motivated by content-based image retrieval (CBIR) [11, 12]. Statistical feature of image histograms is used by calculating some features which includes standard deviation (std), mean, kurtosis, skewness and entropy.

To detect abnormal CXRs with TB, we use a support vector machine (SVM), which classifies the computed feature vectors into either normal or abnormal.

4.RESULT

Table 1 Feature values obtained for right lungs of both normal and Abnormal

Features	Normal	Abnormal
Mean	110.36	104.77
Variance	1237.3	1247
Third Moment	847.34	11353
Entropy	0.8	0.98

Table 2 Feature values obtained for left lungs of both normal and Abnormal

Features	Normal	Abnormal
Mean	129	102
Variance	1028.6	1140.7
Third Moment	-3394	16180
Entropy	0.78	0.99

5.CONCLUSION

We have developed an automated system that screens CXRs for manifestations of TB. When given a CXR as input, our system first segments the lung region using Active contour method. We compute statistical features as input to a classifier, which then classifies the given input image into either normal or abnormal.

In this paper, we conclude that for normal chest radiographs offers lower values in features such as variance, third moment and entropy, but higher mean. And in case of TB chest radiograph's offers lower mean value and other feature values were higher. Usually in TB case both the lungs are affected, spread of the infection on both sides is asymmetrical. TB chest radiographs offers greater disparities between the left and right on every feature values.

6.REFERENCES

[1] R.P. Tripathi, N. Tewari, N. Dwivedi, et al. "Fighting Tuberculosis: An Old Disease with New Challenges". Med Res Rev, 2005, 25(1), 93-131.

[2] R. Colebunders, WE. Bastian. "A Review of Diagnosis and Treatment of Smear-negative Pulmonary Tuberculosis". Int. J. Tuberc Lung Dis, 2000, 4, 97-107.

[3] R. Gonzalez and R. E. Woods. Digital Image Processing. Prentice Hall, New Jersey, ISBN 0201180758, 2 edition, 2002.

[4] "Implementing the WHO Stop TB Strategy", A Handbook for National TB Control Programmes, 2008.

[5] "WHO Policy on Collaborative TB/HIV activities", Guidelines for national programmes and other stakeholders, Update Version of 2004.

[6] Tony Lindeberg, KTH Royal Institute of Technology, Stockholm, Sweden, Scale Invariant Feature Transform.

[7] V. Vapnik, The nature of statistical learning theory. Springer Verlag, 2000.

[8] B. Schölkopf, C. Burges, and A. Smola, Advances in kernel methods: support vector learning. The MIT press, 1999.

[9] K. Sum and P. Cheung, Boundary vector field for parametric active contours, Pattern Recognit., vol. 40, no. 6, pp. 1635-1645, 2007.

[10] T. Wang, I. Cheng, and A. Basu, Fluid vector flow and applications in brain tumor segmentation, IEEE Trans. Biomed. Eng., vol. 56, no. 3, Mar. 2009.

[11] M. Simpson, D. You, M. Rahman, D. Demner-Fushman, S. Antani, and G. Thoma, "ITI's participation in the ImageCLEF 2012 medical retrieval and classification tasks," in CLEF 2012 Working Notes, 2012.

[12] C.-R. Shyu, M. Klaric, G. Scott, A. Barb, C. Davis, and K. Palaniappan, "GeoIRIS: Geospatial information retrieval and indexing system—Content mining, semantics modeling, complex queries," IEEE Trans. Geosci. Remote Sens., vol. 45, no. 4, pp. 839–852, Apr. 2007