Computer Aided Screening of Tuberculosis Using Chest Radiographs

Anandkumar A
PG Scholor
Department of Electronics and Communication
RVS college of Engineering and Technology
Coimbatore, Tamil Nadu
Anandkumar.a88@gmail.com

Kannan.R
Assistant professor
Department of electronics and communication
RVS college of Engineering and Technology
Coimbatore, Tamil Nadu
kannanvlsi@gmail.com

***Abstract* – Tuberculosis, is an infectious bacterial disease caused by Mycobacterium tuberculosis, which most commonly affects the lungs. Approximately One third of the world's population is infected with TB. The mortality rate of patients with Tuberculosis is increased if it is undiagnosed. The standard diagnosing methods of tuberculosis are slow and unreliable. In an effort to overcome the difficulty, this project presents an approach for detecting tuberculosis in conventional poster anterior chest radiograph. The lung region is extracted using a graph cut segmentation method. A set of texture and shape features are computed for the segmented lung region, which is given as input to the pre-trained SVM (Support Vector Machine) classifier. The classifier outputs its confidence in classifying the input CXR as TB positive case. The performance of the system is analyzed using MATLAB R2013a simulator.**

***Index Terms* – Computer-aided detection and diagnosis, lung pattern recognition and classification, segmentation, tuberculosis (TB), X-ray imaging.**

I. Introduction

Tuberculosis (TB) is a major global health problem. TB is an infectious disease caused by the bacillus Mycobacterium tuberculosis, which typically affects the lungs Tuberculosis (TB) is the second leading cause of death,after HIV. Around one-third of the world's population having latent TB, and an estimate of nine million new cases occurring every year. It spreads through the air when people with active TB cough, sneeze, or otherwise expel infectious bacteria. The increasing appearance of multi-drug resistant TB has further created an urgent need for a cost effective screening technology to monitor progress during treatment.

Several antibiotics exist for treating TB. While mortality rates are high when left untreated, treatment with antibiotics greatly improves the chances of survival. Diagnosing TB is still a major challenge. The definitive test for TB is the identification of tuberculosis in a clinical sputum, which is the current gold standard. However, it may take several months to identify this slow growing organism in the laboratory. Another technique is sputum smear microscopy, in which bacteria in sputum samples are observed under a microscope. This technique was developed more than 100 years ago. In addition to this several skin tests based on immune response are available for determining an individual has contracted TB. However, skin tests are not always reliable. The latest development for detection is molecular diagnostic tests that are fast and accurate, and that are highly sensitive and specific. However, more financial support is required for these tests to make in available for all.

In this project, it explains a way to discriminate between traditional and abnormal CXRs with manifestations of TB, victimization image process techniques. an automatic approach for sleuthing TB manifestations in chest X-rays (CXRs), supported respiratory organ segmentation and respiratory organ sickness classification is developed. An automatic approach to X-ray reading permits mass screening of enormous populations that might not be managed manually. A posteroanterior exposure (X-ray) of a patient's chest could be a necessary part of each analysis for TB. The chest exposure includes all pectoral anatomy and provides a high yield, given the low value and single supply. Hence, a reliable screening system for TB detection victimization radiographs would be a crucial step towards additional powerful TB medicine. Respiratory organ segmentation, feature computation, and classification (SVM classifier) measure are the stepwise method that is involved in screening of TB in chest radiographs.

II RELATED WORKS

According to a recent survey, most of the research in computer aided detection and diagnosis in chest radiography has focused on lung nodule detection. Although the target of most research attention, lung n o d u l e s are a relatively rare finding in lungs. The most common findings in chest x rays include lung infiltrates, catheters and abnormalities in the size or contour of the heart. Stefan Jaeger [1], proposed segmentation of the lung region using an optimization method based on graph cut. This method combines intensity Information with personalized lung atlas models derived from the training set. A set of shape, edge, and texture features were computed as input to a binary classifier, which then classifies the given input image into either normal or abnormal.

Sema Candemir [2], presented a robust lung boundary detection method that is based on a simple lung model calculation and a graph cut segmentation algorithm. Using publicly available chest x-ray data set the segmentation accuracy measured around 91% which is comparable to the performance of state-of-the-art algorithms 95% and human-observer scores 94%. Y. Akgul [3], described a methodology for adaptive parameter learning to improve the segmentation performance using a multi-class classifier approach. The performance of the system demonstrated within graph cut segmentation framework. The regularization parameter using a single feature or a heuristic combination of a few features replaced by the model of characteristics of the image regions with feature vectors which include haar feature for edge, local binary patter for texture and hessian for shape information of local regions. Thus it characterizes the image regions better than using single feature.

S. Jaeger [4], developed an automated system that screens chest x-rays for manifestations of TB and other lung diseases. For a given input chest x-ray, the lung field segmented using a combination of an intensity mask, a statistical lung model mask, and a Log Gabor mask. Extract a set of features for shapes, curvatures, and textures from the segmented lung field. Using the extracted features, train a support vector machine that distinguishes between normal and abnormal x-rays.

C. Leung [5], van Ginneken et al, tells a fully automatic method to detect abnormalities.Segmentation is used t o subdivide the lung fields into over- lapping regions of various sizes. From each region the Texture features are extracted, using t h e moments of responses to a multi scale filter bank. All regions a r e classified by voting among nearest neighbors using the leave one out approach. Finally, the classification results of each region are combined.

Wei et al. [7] evaluated 210 features in search of the optimum feature set on 247 chest X-ray images. The presented CAD system consists of four processing steps: 1) detecting possible tumors by using an adaptive ring filter, 2) extraction of the boundaries of the tumor candidates, 3) extraction of feature parameters, and 4) discrimination between normal and abnormal regions. Yoshida [8] developed a method oriented to the reduction of false positives by exploring the symmetry between the right and left lungs and assuming that a nodule candidate region in one lung would correspond to a normal region.in the other. These two regions a r e matched and the difference in structure is evaluated. Using this approach the authors were able to reduce the number of false-positives. Avni et al. [9] present the visual words method, dividing the images into small patches called “visual words”. The feature extraction consists of creating a frequencies vector using the different patches, followed by classifying the images t o different pathologies using a kernel-based SVM.

Kakeda.S [12], states the complexity of developing full-fledged CAD systems for X-ray analysis, research has concentrated on developing solutions for specific sub problems. The segmentation of the lung field is a typical task that any CAD system needs to support for a proper evaluation of CXRs. Other segmentations that may be helpful include the segmentation of the ribs, heart, and clavicles. For example, van Ginneken et al.[11] compared various techniques for lung segmentation, including active shapes, rule-based methods, pixel classification, and various combinations thereof. Their conclusion was that pixel classification provided very good performance on their test data. Ginneken et al. [14] state that automated nodule detection is becoming one of the more mature applications of decision support/automation for CXR and CT. Several studies have been published evaluating the capability of commercially available CAD systems to detect lung nodules. The result is that CAD systems can successfully assist radiologists in diagnosing lung cancer. However, nodules represent only one of many manifestations of TB in radiographs.

III SYSTEM DESCRIPTION

This section presents implemented methods for lung segmentation, feature computation, and classification which is simpler compared to the previous existing methods, and has better accuracy.

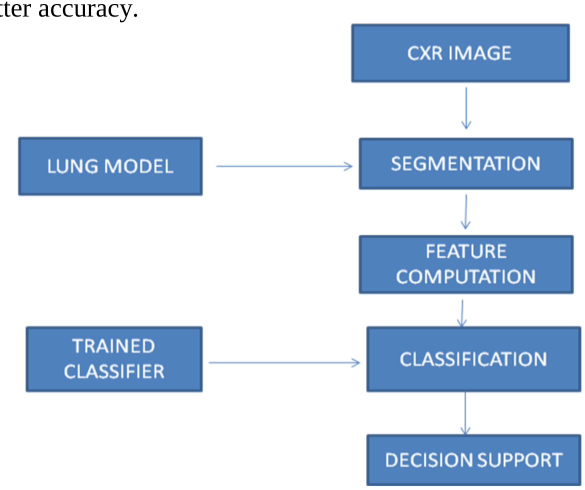


Fig.1 System overview

Fig.1 shows the architecture of the system with the different processing steps are first, segments the lung of the input CXR using a graph cut optimization method in combination with a lung model. For the segmented lung field, our system then computes a set of features as input to a pre-trained binary classifier. Finally, using decision rules and thresholds, the classifier outputs its confidence in classifying the input CXR as a TB positive case.

The method to address the chest classification problem, presenting a generic approach for the diagnosis of chest pathologies of all types. The algorithm is comprised of two main phases: The image processing phase and the classification phase. The image processing phase includes segmentation and feature extraction. The classification phase deals about the SVM classifier.

A. Lung segmentation

Lung segmentation performed by considering the properties of lung boundaries, regions, and shapes. The system has two main stages. It first computes an average shape model using training images. Then it uses a graph cut segmentation algorithm to detect the lung Regions with the help of the calculated shape model.

In general, segmentation in medical images has to cope with anatomical shape variations, poor contrast, and acquisition noise due to hardware constraints. Lung segmentation also faces theses issues. Therefore incorporate a lung model that represents the average lung shape of selected training masks. Select these masks according to their shape similarity as follows. linearly align all training masks to a given input CXR. Then compute the vertical and horizontal intensity projections of the histogram equalized images. To measure the similarity between projections of the input CXR and the training CXRs, use the Bhattacharyya coefficient.

To compute an approximate shape model, use training images which are selected according to their shape similarity. The average of all selected masks is used as an approximate shape model for the observed patient lung image. The second stage of the system detects the lung region with a segmentation algorithm.

Fig 2.CXR and its calculated lung model

Fig.2 shows a CXR and its calculated lung model. Graph cut approach and models the lung boundary detection with an objective function. To formulate the objective function, define three requirements a lung region has to satisfy: a) The lung region should be consistent with typical CXR intensities expected in a lung region, b) Neighboring pixels should have consistent labels, c) The lung region needs to be similar to the lung model computed, Fig: 4 shows a typical computed lung model.

B. Features

To describe normal and abnormal patterns in the segmented lung field, experimented with two different feature sets. Our motivation is to use features that can pick up subtle structures in an CXR.

1) Object Detection Inspired Features - Set A: As our first set, use features that have successfully applied to microscopy images of cells for which classified the cell cycle phase based on appearance patterns. It is the same set that have used in our earlier TB classification work. This set is versatile and can also be applied to object detection applications.

The first set is a combination of shape, edge, and texture descriptors. For each descriptor, compute a histogram that shows the distribution of the different descriptor values across the lung field. Each histogram bin is a feature, and all features of all descriptors put together form a feature vector that input to our classifier. Through empirical experiments, found that using 32 bins for each histogram gives us good practical results . In particular, use the following shape and texture descriptors.

- Intensity histograms (IH)
- Gradient Magnitude Histograms (GM)
- Shape descriptor histograms (SD)

$$SD=\tan^{-1}(\lambda_1/\lambda_2) \quad (5)$$

Where λ_1 and λ_2 are the eigen values of Hessian matrix with $\lambda_1 \leq \lambda_2$

Histogram of oriented gradients (HOG) is a descriptor for gradient orientations weighted according to gradient magnitude.The image is divided into small connected regions, and for each region a histogram of gradient direction or edge orientations for pixels within the region is computed. The combination of these histograms represent the descriptor.

2) CBIR-based Image Features - Set B: For second feature set, Set B, use a group of low-level features motivated by content-based image retrieval (CBIR). This feature collection includes intensity, edge, texture and shape moment features, which are typically used by CBIR systems. The entire feature vector has 594 dimensions, which is more than three times larger than the feature vector of Set A, and which allows us to evaluate the effect of high-dimensional feature spaces on classification accuracy. Extract most of the features, except for Hu moments and shape features, based on the Lucene image retrieval library. In particular, Feature Set B contains the following features

Tamura texture descriptor: The Tamura descriptor is motivated by the human visual perception. The descriptor comprises a set of six features. Only use three of these features, which have the strongest correlation with human perception: contrast, directionality, and coarseness.

• CEDD and FCTH: CEDD (color and edge direction descriptor) and FCTH (fuzzy color and texture histogram) incorporate color and texture information in one histogram. They differ in the way they capture texture information.

• Hu moments: These moments are widely used in image analysis. They are invariant under image scaling, translation, and rotation and use the DISCOVIR system (Distributed Content-based Visual Information Retrieval) to extract Hu moments.

• CLD and EHD edge direction features: CLD (color layout descriptor) and EHD (edge histogram descriptor) are MPEG-7 features. CLD captures the spatial layout of the dominant colors on an image grid consisting of 8 by 8 blocks and is represented using DCT (discrete cosine transform) coefficients. EHD represents the local edge distribution in the image, i.e. the relative frequency of occurrence of five types of edges (vertical, horizontal, 45-degree diagonal, 135-degree diagonal, and non directional) in the sub-images.

• Primitive length, edge frequency, and autocorrelation: These are well-known texture analysis methods, which use statistical rules to describe the spatial distribution and relation of gray values.

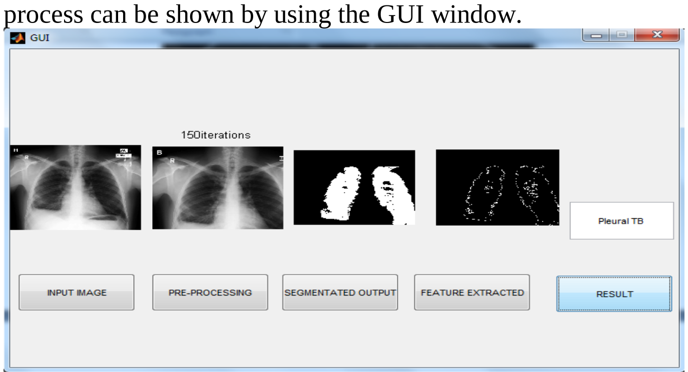
• Shape features: using a collection of shape features provided by the standard Matlab implementation (region props), such as the area or elliptical shape features of local patterns.

C. Classification

To detect abnormal CXRs with TB, use a support vector machine (SVM), which classifies the computed feature vectors into either normal or abnormal. An SVM in its original form is a supervised non-probabilistic classifier that generates hyper planes to separate samples from two different classes in a space with possibly infinite dimension. The unique characteristic of an SVM is that it does so by computing the hyper plane with the largest margin; i.e. the hyper plane with the largest distance to the nearest training data point of any class. Ideally, the feature vectors of abnormal CXRs will have a positive distance to the separating hyper plane, and feature vectors of normal CXRs will have a negative distance. The larger the distance the more confident are in the class label.Therefore use these distances as confidence values to compute the ROC curves.

IV. RESULTS

The sample images that have been taken from the available resources are filtered using Median Filtering technique. The Median filtering technique is used remove the noises which the user can't identify with naked eye. The Median filter thus improves the quality of the images. The sample images after applied the filtering technique are given to the process of lung segmentation. The lung segmentation process only segments the lung region from the sample x-ray images. The sample images after applied the segmentation technique are given to the Feature extraction. The overall process can be shown by using the GUI window.



V. CONCLUSION

An automated system that screens CXRs for manifestations of TB has been developed. The lung region is extracted using a graph cut segmentation method. A set of texture and shape features are computed for the segmented lung region, which is given as input to the pre-trained SVM (Support Vector Machine) classifier. The classifier outputs its confidence in classifying the input CXR as TB positive case. The performance of the system is analysed using MAT LAB simulator. In Future, the detection and diagnosing Tuberculosis can be done by using LBP and GLCM features and ANFIS Classifier and the overall performance analysis will be compared with present SVM classification system.

REFERENCES

[1]. Stefan Jaeger," Automatic Tuberculosis Screening Using Chest Radiographs", IEEE TRANSACTIONS ON MEDICAL IMAGING, VOL. 33, NO. 2, FEBRUARY 2014.

[2] S. Candemir, S. Jaeger, K. Palaniappan, S. Antani, and G. Thoma, "Graph-cut based automatic lung boundary detection in chest radiographs," in Proc. IEEE Healthcare Technol. Conf.: Translat. Eng. Health Med., 2012, pp. 31–34.

[3] S. Candemir, K. Palaniappan, and Y. Akgul, "Multi-class regularization parameter learning for graph cut image segmentation," in Proc. Int. Symp. Biomed. Imag., 2013, pp. 1473–1476.

[4] S. Jaeger, A. Karargyris, S. Antani, and G. Thoma, "Detecting tuberculosis in radiographs using combined lung masks," in Proc. Int. Conf. IEEE Eng. Med. Biol. Soc., 2012, pp. 4978–4981

.

[5] C. Leung, "Reexamining the role of radiography in tuberculosis case finding," Int. J. Tuberculosis Lung Disease, vol. 15, no. 10, pp. 1279–1279, 2011.

[6] Yoshihiko Nakamura, Gentaro Fukano, Hotaka Takizawa, Shinji Mizuno, Shinji Yamamo, Tohru Matsumoto, Yukio Tateno, and Takeshi Iinuma. Eigen Nodule: View-based Recognition of Lung Nodule in Chest X-ray CT Images Using Subspace Method. International Conference on Pattern Recognition, 4:681–684, 2004.

[7] J. Wei, Y. Hagihara, A. Shimizu, and H. Kobatake. Optimal im- age feature set for detecting lung nodules on chest x-ray images. In Proc. Int. Workshop on Computer-Aided Diagnosis. Citeseer, 2002.

[8] Yoshida. Local contralateral subtraction based on bilateral symmetry of lung for reduction of false positives in comput- erized detection of pulmonary nodules. Biomedical Engineering, IEEE Transactions on, 2004

[9] Uri Avni, Hayit Greenspan, Eli Konen, Michal Sharon, and Jacob Goldbergeri. X-ray Categorization and Retrieval on the Organ and Pathology Level, Using Patch-Based Visual Words. IEEE Transactions on Medical Imaging, 30, 2011

[10] S. Jaeger,A.Karargyris, S. Candemir, J. Siegelman, L. Folio, S.Antani, and G. Thoma, "Automatic screening for tuberculosis in chest radiographs: A survey," Quant. Imag. Med. Surg, 2013

[11] B. Van Ginneken, B. ter Haar Romeny, and M. Viergever, "Computeraided diagnosis in chest radiography: A survey," IEEE Trans. Med. Imag.,Dec. 2001.

[12] S. Kakeda, J. Moriya, H. Sato, T. Aoki, H. Watanabe, H. Nakata, N. Oda, S. Katsuragawa, K. Yamamoto, and K. Doi, "Improved detection of lung nodules on chest radiographs using a commercial computer-aided diagnosis system," Am. J. Roentgenol.2004.

[13] A. Dawoud, "Fusing shape information in lung segmentation in chest radiographs," Image Anal. Recognit.2010.

[14] P. Maduskar, L. Hogeweg, H. Ayles, and B. van Ginneken, "Performance Evaluation of automatic chest radiograph reading for detection

of tuberculosis (TB): A comparative study with clinical officers and certified readers on TB suspects in sub-Saharan Africa," in Eur. Congr. Radiol., 2013.

[15] A. Hoog, H. Meme, H. van Deutekom, A. Mithika, C. Olunga, F. Onyino, and M. Borgdorff, "High sensitivity of chest radiograph reading by clinical officers in a tuberculosis prevalence survey," Int. J. Tuberculosis Lung Disease, 2011.