Level Set Segmentation Method for Automatic Tuberculosis Screening Using Chest Radiographs

Sajmi Salam

P.G Scholar, Department of Computer Science, Caarmel Engineering College, Pathanamthitta, India

Abstract: *Tuberculosis is a major global health problem worldwide, after HIV. Diagnosing tuberculosis is still remains a challenge because of the multi-drug-resistant bacteria which cause this disease. Tuberculosis is an infectious disease caused by the bacillus Mycobacterium tuberculosis which mainly affects the lungs. Mortality rates of the patients with tuberculosis are high when left undiagnosed and thus untreated. Standard diagnosis still remains slow and often unreliable. In an effort to reduce the disease, this paper presents an automated approach for detecting tuberculosis in chest radiographs. In this paper, we first extract the lung region using a level set segmentation method. We compute a set of texture and shape features, which enable the X-rays to be classified as normal or abnormal using a binary classifier for this lung region. Many antibiotics exists for TB. Unfortunately, diagnosing TB is still a major problem. For this problem, we are using a cost effective screening technology to monitor progress during treatment. Level Sets are an important category of modern image segmentation techniques are based on partial differential equations (PDE), i.e. progressive evaluation of the differences among neighboring pixels to find object boundaries. The level set method was initially proposed to track moving interfaces. It can be used to efficiently address the problem of curve/surface/etc. Initial contour is taken over the object that time the level will be zero then this contour is moving towards the object boundary then it settle down in the boundary of the object well. At first we take the outside contour. Then move towards the object. If the gray level changing, move it else not. Stop when gray level changes.*

Keywords: Computer-aided detection and diagnosis, lung pattern recognition and classification, segmentation, tuberculosis (TB), X-ray imaging.

1. Introduction

Tuberculosis (TB) is the second leading cause of death from an infectious disease worldwide, after HIV. With about one-third of the world's population having latent TB, and an estimated nine million new cases occurring every year, Tuberculosis is a major global health problem. Tuberculosis is an infectious disease caused by the bacillus Mycobacterium tuberculosis, which affects the lungs. It spreads through the air when people with active TB cough, sneeze, or otherwise expel this infectious bacteria. TB is mostly found in sub-Saharan Africa and Southeast Asia, where widespread poverty and malnutrition reduce resistance to this disease. Moreover, opportunistic infections in immune-compromised HIV/AIDS patients have exacerbated this problem. The increasing appearance of multi-drug resistant TB has created an urgent need for a cost effective screening technology to monitor progress during treatment.

For treating TB there are many existing antibiotics. While mortality rates are so high when left untreated, treatment with antibiotics greatly improves the chances of survival. Unfortunately, diagnosing TB is still a major problem. The definitive test for Tuberculosis is the identification of Mycobacterium tuberculosis in a clinical sputum or pus sample, which is the current gold standard. However, it may take many months to identify this slow-growing organism in the lab. Another method is sputum smear microscopy, in which bacteria in sputum samples are observed under a microscope. This method was developed more than 100 years ago. In addition, several skin tests based on immune response are available for identifying whether an individual has contracted TB. However, skin tests are not reliable always. The latest developments for detection are molecular diagnostic tests that are very fast and accurate, and that are highly sensitive. However, further economical support is required for these tests to become common place. In this paper, I present an automated approach for detecting Tuberculosis manifestations in chest X-rays (CXRs) in lung segmentation and lung disease classification. An automated approach to X-ray reading allows mass screening of very large populations that could not be managed manually. A postero-anterior radiograph (X-ray) of a patient's chest is a major part of every evaluation for Tuberculosis. The chest radiograph includes all thoracic anatomy and provides a very high yield, given the low cost and single source. Therefore, a most reliable screening system for Tuberculosis detection using radiographs would be a major critical step towards more powerful Tuberculosis diagnostics.

Figure 1: Examples of normal CXRs in the MC dataset
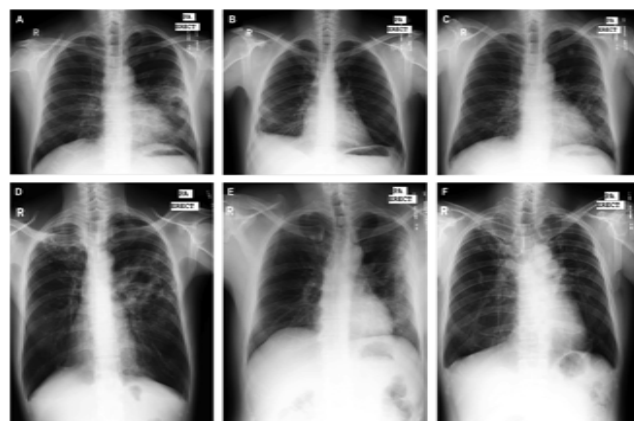

Figure 2: Examples of abnormal CXRs in the MC dataset.

Fig. 1 shows examples of normal CXRs without signs of TB. These examples are from Montgomery County (MC) dataset. Fig. 2 shows positive examples with manifestations of Tuberculosis, which are from the same dataset. Typical manifestations of TB in chest X-rays are, infiltrations, cavitations, effusions, or miliary patterns. For instance, CXR A and C in Fig. 2 have infiltrates in both lungs. CXR B is a good example of pleural TB, which is indicated by the abnormal shape of costophrenic angle of the right lung..

In CXRD the irregular infiltrates in the left lung with a large area of cavitation. There is scarring in the right apical region. CXRE shows the peripheral infiltrates in the left lung. Finally, CXRF shows TB scars resulting from an older Tuberculosis infection. In this paper describes how to discriminate between normal and abnormal CXRs with manifestations of Tuberculosis, using image processing techniques. This paper presents the level set segmentation method for the lung segmentation Level Sets are an important category of modern image segmentation techniques are based on partial differential equations (PDE), i.e. progressive evaluation of the differences among neighboring pixels to find object boundaries. The level set method was initially proposed to track moving interfaces. It can be used to efficiently address the problem of curve/surface/etc. Initial contour is taken over the object that time the level will be zero then this contour is moving towards the object boundary then it settle down in the boundary of the object well. First take the outside contour. Then move towards the object. If the gray level changing, move it else not. Stop when gray level changes.

2. Related Works

The invention of the possibility of digital image processing and digital chest radiography has given new impetus to computer aided screening and diagnosis. Still the standard CXR is a very complex imaging tool despite of its omnipresence in medical practice. In the last 10 years, several ground-breaking papers have been published on computer-aided diagnosis (CAD) in CXRs. However, there is no doubt that more research is needed to meet the practical performance requirements for deployable diagnostic systems. In a recent survey, van Ginneken *et al.* state that 45 years after the initial work on computer-aided diagnosis in chest radiology, there are still no systems that can accurately read chest radiographs. 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 radiologists in diagnosing lung cancer. However only one of many manifestations of TB in radiographs represents the nodules. In recent years, due to the complexity of developing full-fledged CAD systems for X-ray analysis, research has concentrated on developing solutions for specific subproblems, 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.* 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. Dawoud presented an iterative segmentation approach that combines intensity information with shape priors trained on the publicly available JSRT database.
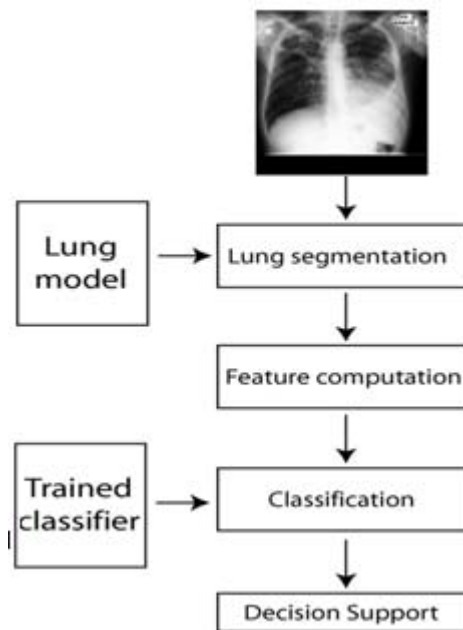

Figure 3: CXR and its calculated lung model

Different feature types and ways to aggregate them have been reported in the literature depending on the lung segmentation. For example, van Ginneken *et al.* subdivide the lung into overlapping regions of various sizes and extract features from each region. The use the moments of responses to a multiscale filter bank is used to detect abnormal signs of diffuse textural nature. In addition, they use the difference between corresponding regions in the left and right lung fields as features. A separate training set is constructed for each region and final classification is done by voting and a weighted integration.

3. Method

This section presents the implemented lung segmentation methods, feature computation, and classification. Fig. 3 shows the architecture of our system with the different processing steps, which the following sections will discuss in more detail. First, our system segments the lung of the input CXR using a graph cut optimization method in combination with a lung model. Our system then computes a set of features as input to a pre-trained binary classifier for the segmented lung field,. Finally, using decision rules and thresholds, the classifier outputs its confidence in classifying the input CXR as a TB positive case, for example.
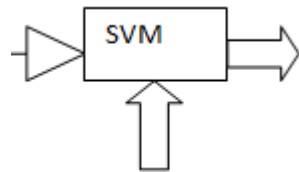
3.1 Level Set Segmentation Method

Lung segmentation is modeled as an optimization problem that takes properties of lung boundaries, regions, and shapes into account. In general it can be explained as segmentation in medical images has to cope with poor contrast, acquisition noise due to hardware constraints, and anatomical shape variations. Lung segmentation is no exception in this regard.

Paper ID: OCT14346   1634

Level set segmentation algorithm as follows: 1.Taking the chest image 2.Using graph gut segmentation algorithm segment only region of the Lung from the background 3. Taking the special characteristics of the image is known as Feature extraction. The following features are extracted:

- Intensity histograms (IH).
- Gradient magnitude histograms(GM)
- Shape descriptor histograms (SD), Shape features, classifier, It is used to classify the different kind of data with the help of training feature extracted data. Here we are using SVM classifier to classify the input data

Input Output Normal/Affected

Training Feature Data (Data Base)

An important category of modern image segmentation techniques that are based on partial differential equations (PDE) are Level Sets. PDE is the progressive evaluation of the differences among neighboring pixels to find object boundaries. The level set method was initially proposed to track moving interfaces. It can be used to efficiently address the problem of curve/surface/etc. Initial contour is taken over the object that time the level will be zero then this contour is moving towards the object boundary then it settle down in the boundary of the object well.

1. Take the outside contour
2. Move towards the object
3. If gray level changing move it else not
4. Stop when gray level changes

This implementation is based on following PDE update:
Φ
with
$g =$
$\Delta T=$
Φ= iso surface at current iteration i
W_a= Advection weight
F_a=Advection force
W_c=Curvature weight
F_c=Curvature force

B. Features
I experimented with two different feature sets to describe normal and abnormal patterns in the segmented lung field. The motivation is to use features that can pick up subtle structures in a CXR.

1) Object Detection Inspired Features

Set A: As the first set, I use features that I have been successfully applied to microscopy images of cells for which I classified the cell cycle phase based on appearance patterns. It is the same set that is used in our earlier TB

classification work. This set is versatile and can also be applied to object detection applications for example the first set is a combination of shape, edge, and texture descriptors. For each descriptor, first an histogram is computed 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 are input to our classifier. Through empirical experiments, it is found that using 32 bins for each histogram gives good practical results. The following shape and texture descriptors are used,

- Intensity histograms (IH).
- Gradient magnitude histograms (GM).
- Shape descriptor histograms (SD)

2) CBIR-Based Image Feature

Set B: For the second feature set, Set B, I 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 to evaluate the effect of high dimensional feature spaces on classification accuracy. Then extract most of the features, except for Hu moments and shape features, based on the Lucene image retrieval library. The Feature Set B contains the following features.

- Tamura texture descriptor: The human visual perception is the motivation of the Tamura descriptor. The descriptor comprises a set of six features. Only three of these features are used in our method, 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 mainly used in image analysis. They are invariant under image scaling, translation, and rotation. The DISCOVIR system (distributed content-based visual information retrieval) is used 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* 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 diagonal, 135 diagonal, and non-directional) in the sub-images.
- Primitive length, edge frequency, and autocorrelation: These are a well-known texture analysis method that uses statistical rules to describe the spatial distribution and relation of gray values.
- Shape features: A collection of shape is used features provided by the standard MATLAB implementation

Paper ID: OCT14346

1635

(region props), such as the area or elliptical shape features of local patterns.

C. Classification

The proposed system uses a support vector machine (SVM) to detect abnormal CXRs with TB, which classifies the computed feature vectors into either normal or abnormal. An SVM is a supervised non-probabilistic classifier in its original form that generates hyperplanes 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 hyperplane with the largest margin; i.e., the hyperplane 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 hyperplane, and feature vectors of normal CXRs will have a negative distance. The larger the distance the more confident in case of the class label. Therefore use these distances as confidence values to compute the ROC curves.

4. Conclusion

An automated system that screens CXRs for manifestations of TB has been developed. When given a CXR as input, the proposed system first segments the lung region using an optimization method based on level set method. This method combines intensity information with personalized lung atlas models derived from the training set. We compute a set of shape, edge, and texture features as input to a binary classifier, which then classifies the given input image into either normal or abnormal. In this paper, I compare two different established feature sets: one set specially used for object recognition and the other used in a system that can assist radiologists and public health providers in the screening and decision process. These comparison results have encouraged to test the system in the field under realistic conditions. In future experiments, i will evaluate our system on larger datasets that we will collect using our portable scanners.

References

[1] S. Candemir, S. Jaeger, K. Palaniappan, S. Antani, and G. Thoma ,"Automatic Tuberculosis screening using chest radiographs,"in IEEE Medical imaging,vol. 33,February 2014.

[2] World Health Org., Global tuberculosis report 2012.

[3] World Health Org., Global tuberculosis control 2011 2011.

[4] Stop TB Partnership, World Health Org. , The Global Plan to Stop TB 2011–2015 2011

[5] 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.

[6] 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.

[7] 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.

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

Paper ID: OCT14346

1636