ISSN: 2321-7782 (Online)
Volume 3, Issue 5, May 2015
International Journal of Advance Research in Computer Science and Management Studies
Research Article / Survey Paper / Case Study
Available online at: www.ijarcsms.com

An Efficient Detection of Tuberculosis from Chest X-rays

Hrudya Das[1]
Signal Processing
Collage of Engineering Cherthala
Alappuzha, Kerala – India

Ajay Nath[2]
Signal Processing
Collage of Engineering Cherthala
Alappuzha, Kerala – India

Abstract: Tuberculosis (TB) is a major health problem in all over the world. Chest radiographs is becoming an important tool for fighting against TB .Existing Methods are less reliable in high population. So a computer aided system for detecting TB is becoming more needful for the mass screening of TB .Detecting cavities from chest x-ray is an efficient method for diagnosing the TB. So here, an automatic method is explained for detecting the TB from CXR with less effort. Region based active contour segmentation is used for segmenting lung field and the extracted features are classified using supported vector machine as normal and abnormal. The Montgomery County (MC) Data set contains 138 posterioanterior cxrs, among which 80 cxrs are normal and 58 cxrs are abnormal with manifestations of TB are used. All images of the MC set are in 12-bit grayscale, captured with an Eureka stationary x-ray machine (CR). The abnormal cxrs cover a wide range of TB-related abnormalities, including effusions and miliary patterns.the full system explanation is discussed.

Keywords: Region based segmentation, feature extraction, classification, SVM, HOG.

I. INTRODUCTION

Despite the existence of an effective and affordable cure, tuberculosis (TB) remains one of the world's major health care challenges. Mortality and morbidity rates are only slightly lower than those of the well-known HIV/AIDS epidemic [1], but TB has received less attention of the media and public. One of the reasons for this has been the decline of TB in high-income countries [2]. Tuberculosis is a major health threat in many regions of the world. While diagnosing tuberculosis still remains a challenge. Although TB can affect almost any part of the body, its main site of infection is the lungs because of the bacillus preference for high oxygen environments. The pathogenesis of TB is complex, and some of its features are not fully understood yet. Bacilli enter the lungs through the airways and end up in the alveoli where they invoke the innate immune response. In clinical practice, TB is diagnosed using a combination of clinical symptoms, chest radiography, and sputum examination .The typical symptoms associated with TB are fever, weight loss, night sweats, and coughing. Standard diagnostics still rely on methods developed in the last century. They are slow and often unreliable. Manually the detection of TB cavities is done by just looking at the X-rays/CT images by the doctors/technicians. So by means of looking at the images by the naked eye there is more chance for wrong prediction of the intensity of the cavities. Hence, because of this wrong prediction of the cavities, the physicians may not prescribe correct dosage of medicine. They may prescribe high or Low dosage of medicine. If the dosage is too high it will lead to various harmful effects such as causing other diseases. If the dosage is too low the patient cannot easily recover from the disease soon. So the accurate detection of the cavities must be done for the accurate prescription of medicine with the correct dosage to get rid of the disease completely.

So the automatic detection of tuberculosis from x-ray may helpful in the rural area where an expert Radiologist is not always available. So developing a CAD system for the diagnosing TB is a challenging task. The basic method includes the segmentation, feature extraction and classification. 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 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. In the segmentation stage proper method were used to segment the lung field correctly. In general, segmentation in medical images has to cope with poor contrast, acquisition noise due to hardware constraints, and anatomical shape variations. Depending on the lung segmentation, different feature were extracted for the further analysis. The extracted features are input to the classifier, which then classifies a given input image into either normal or abnormal. Here first extract the lung region using a region based active contour segmentation method. For this lung region, compute a set of texture and shape features, which enable the x-rays to be classified as normal or abnormal using a binary classifier.

II. RELATED WORK

Chest X-rays (CXR) can be used for automatic detection of TB. Previous methods related to this work are also considered in this study. The purpose of the segmentation is to find corresponding regions within the lung fields. Segmentation of lung fields on PA chest radiographs has received considerable attention in the literature. ASMs have been developed Cootes and Taylor [6] and have been applied to various segmentation tasks in medical imaging [7].An object is described by points, referred to as landmark points. The landmark points are (manually) determined in a set of training images. From these collections of landmark points, a point distribution model is constructed. This method make use of active shape mode segmentation method [5] for segmenting the lung field. Since the presence of cavities in the upper thoracic area often suggests typical TB, automatic segmentation and classification of these cavities taking spatial, geometric and demographic information into consideration is useful for a CAD system. Active contour (AC) models (or snakes) [9][10] are commonly used as segmentation techniques for medical images, though prior knowledge of the region of interest (ROI) is often needed in order to guide the execution and successful convergence of these algorithms. AC models can be categorized into two types: level set and parametric. The convergence of level set techniques is usually slower than parametric methods because the deformation of a higher dimensional function is required [11].Besides, level set is very sensitive to noise resulting in the extraction of too many false objects. Given an initial contour, the external and internal forces of parametric snakes drive the evolution and converge to the final contour much faster than level set snakes. However, as said

In earlier research [10], traditional parametric AC models have limited capture range. It cannot converge accurately unless an initial contour is specified close to the region of interest. Boundary vector flow (BVF) [9], an enhanced version of the traditional parametric AC models, has successfully increased the capture range but is not suitable to extract acute concave angles. . Magnetostatic active contour (MAC) model [11] is a level set snake. Although slower than parametric methods, it is able to extract multiple objects starting from a single initial contour. The major challenge in the cavity detection from CXRs is the complicated texture and varied intensity distribution in the lung fields caused by TB infection.

Mean shift [13] is a feature space analysis technique that clusters neighboring data points with similar characteristics using a neighborhood search procedure, which locates the local maxima in a probability density function (PDF), based on the kernel density estimation (or Parzen window method). The mean shift segmentation is based on a recursive mean shift procedure If a cavity appears near the clavicles, it is very likely that the cavity is partially occluded, which makes the visible part of the cavity violate the circularity criterion. In addition, the radio-dense clavicles make the intensity distribution of lung fields near them quite different from other portions of the lung fields which may require a different GICOV threshold. Therefore, if no cavity is confirmed in the clavicle regions in second stage, those regions go through segmentation and classification stage (third stage), using new GICOV and circularity thresholds calculated for them [15]. In the hybrid scheme, GICOV algorithm for inner boundary features and circular measure for shape feature together for classification are used. A region based active contour method for x-ray lung segmentation [16] work, a level set energy for segmenting the lungs from digital Posterior-Anterior (PA) chest x-ray images is presented. Segmentation using combined lung mask [19] use an average of three different masks for lung segmentation: the intensity mask, the lung model mask, and the Log Gabor mask. A method called segmentation using serial chest radiographs [20] used two kinds of shape statistics, the population based and the patient-specific shape statistics. The former is trained from the manually marked lung field contours from a population using PCA method, and the latter is obtained from the segmentation results of a specific patient available during serial lung field segmentation procedure. A graph cut method is also used for the detection of TB[25] .SIFT[20], texture[5],Local binary pattern and shape descriptors are used as features .k-nn classifier, baysein classifier are the commonly used classifier in all of this works.

III. PROPOSED METHOD

The method for the detection of tuberculosis includes the main steps like segmentation, feature extraction and classification. Section A describes segmentation, section B describe feature extraction and classification and section C describes the implementation details.

A. LOCALIZED REGION BASED SEGMENTATION

Two main categories exist for active contours: edge-based and region-based. Edge-based active contour models utilize image gradients in order to identify object boundaries. This type of highly localized image information is adequate in some situations, but has been found to be very sensitive to image noise. One benefit of this type of flow is the fact that no global constraints are placed on the image. Thus, the foreground and background can be heterogeneous and a correct segmentation can still be achieved in certain cases. More recently, work in active contours has been focused on region-based flows inspired by the region-competition work of Zhu and Yuille.[34].These approaches model the foreground and background regions statistically and find an energy optimum where the model best fits the image. Some of the most well-known and widely used region-based active contour models assume the various image regions to be of constant intensity.

Analysis of local regions leads to the construction of a family of local energies at each point along the curve. In order to optimize these local energies, each point is considered separately, and moves to Minimize (or maximize) the energy computed in its own local region. To compute these local energies, Local neighborhoods are split into local interior and local exterior by the evolving curve. The energy Optimization is then done by fitting a model to each local region. Let *I* denote a given image defined on the domain Ω, and let ***C*** be a closed contour represented as the zero level set of a signed distance function ϕ i.e. C={x=x | ϕ(x)=0}.

To specify the area just around the curve, use the derivative of heviside function. Here use B(x, y) to mask local region. This function will be 1 when the point is within a ball of radius centred at, and 0 otherwise.

Energy function is given as

$$E(x) = \int_{\Omega_x} \delta\phi(x) \int_{\Omega_y} B(x,y).F(I(y), \Phi(y)) d_y d_x$$

The function, F is a generic internal energy measure used to represent local adherence to a given model at each point along the contour. The various internal energies are uniform modelling energy, the means separation energy, and the histogram separation energy. The uniform modelling flow finds its minimum energy when the interior and exterior are best approximated. Mean Separation (MS) Energy relies on the assumption that foreground and background regions should have maximally separate mean intensities. Optimizing the energy causes the curve to move so that interior and exterior means have the largest difference possible. Histogram Separation (HS) Energy, more complex energy that looks past simple means and compares the full histograms of the foreground and background

Mean seperation Enerrgy

Another important global region-based energy that uses mean Intensities is the one proposed by Yezzi *et al.* [16] which Refer to as *means separation energy.*

$$E_{MS} = \int_{\Omega_y} (u - v)^2.$$

This energy works on the assumption that foreground and Background regions should have maximally separate mean intensities. Optimizing the energy causes the curve to move so that interior and exterior means have the largest possible difference. There is no restriction on how well the regions are modeled by u and v. A corresponding ***F*** is formed by localizing the global energy with local mean equivalents as shown here

$$F_{\text{MS}} = (u_x - v_x)^2.$$

Local region based flow is

$$\begin{aligned}\frac{\partial \phi}{\partial t}(x) = \delta\phi(x) \int_{\Omega_y} &\mathcal{B}(x,y)\delta\phi(y) \\ &\cdot \left(\frac{(I(y) - u_x)^2}{\mathrm{A}_u} - \frac{(I(y) - v_x)^2}{\mathrm{A}_v} \right) dy \\ &+ \lambda\delta\phi(x)\, \mathrm{div} \left(\frac{\nabla\phi(x)}{|\nabla\phi(x)|} \right)\end{aligned}$$

Where A_u and A_v the areas of the local interior and local Exterior regions respectively given by

$$\begin{aligned}\mathrm{A}_u &= \int_{\Omega_y} \mathcal{B}(x,y) \cdot \mathcal{H}\phi(y)\, dy \\ \mathrm{A}_v &= \int_{\Omega_y} \mathcal{B}(x,y) \cdot (1 - \mathcal{H}\phi(y))\, dy.\end{aligned}$$

The optimum of this energy is obtained when u_x and v_x are the most different at every *x* along the contour. In some cases, this is more desirable than attempting to fit a constant model. This method encouraging local foreground and background means to be different. that means it is not constant. This allows this energy to find image edges clearly and easily without being distracted when interior or exterior regions are not uniform.

B. FEATURES AND CLASSIFICATION

To describe normal and abnormal patterns in the segmented lung field,following features are extracted

- Intensity histograms (IH).This histogram is a graph showing the number of pixels in an image at each different intensity value found in that image. For an 8-bit greyscale image there are 256 different possible intensities, and so the histogram will graphically display 256 numbers showing the distribution of pixels amongst those greyscale values.
- Gradient Magnitude Histograms (GM).
- Shape descriptor histograms (SD).

  Where λ_1 and λ_2 are the eigenvalues of the Hessian matrix, with $\lambda_1 \leq \lambda_2$
- Histogram of oriented gradients (HOG) is a descriptor for gradient orientations weighted according to gradient magnitude [24].The image is divided into small connected regions, and for each region a histogram of gradient directions or edge orientations for the pixels within the region is computed. The combination of these histograms represents the descriptor.
- Local binary patterns (LBP) is a texture descriptor that codes the intensity differences between neighbouring Pixels by a histogram of binary patterns [27], LBP is thus a histogram method in itself. The Binary patterns are generated by thresolding the Relative intensity between the central pixel and its neighbouring pixels. Because of its computation

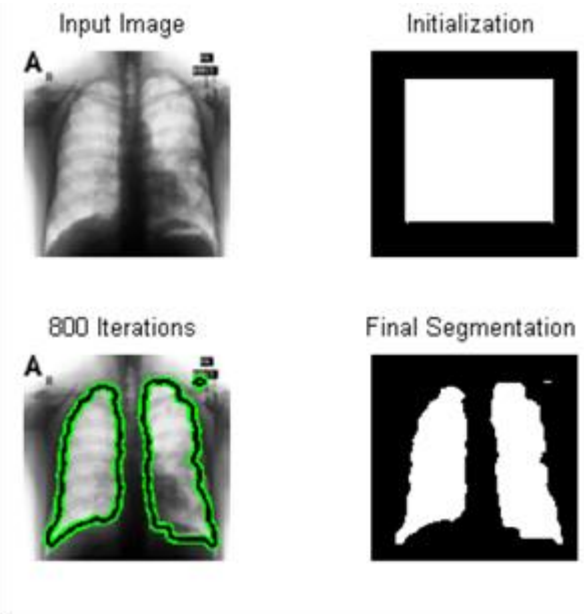
simplicity and efficiency, LBP is successfully used in various computer vision applications often in combination with HOG.

To detect abnormal CXRs with TB, here use a support vector machine (SVM), which classifies the computed feature vectors into normal and abnormal class. An SVM is a supervised non-probabilistic classifier that generates hyperplanes to separate samples from two different classes in a space with possibly infinite dimension. The important characteristic of an SVM is that it classifying the data by computing the hyperplane with the largest margin. It use the hyperplane with the largest distance to the nearest training data point of any class. 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.

C. EXPERIMENTAL RESULT

We have implemented this work on MATLAB 2015 which takes nearly 18 sec to complete the detection process.

Her 1800 iteration used. When the iteration increases better result will obtain.



IV. CONCLUSION

TB can be detected from Chest x-ray images by using image processing methods like segmentation, Feature Extraction and classification. Existing diagnostics method such as sputum staining has become less reliable in high population so this method will be helpful in rural areas. When increasing the features selected and when using another segmentation method it may get more accurate result.

References

1. World Health Organization. Global tuberculosis report 2012,
2. Lawn S. and Zumla A. Tuberculosis. Lancet, 2011.
3. Daffner R. H. Clinical radiology, the essentials. Williams and Wilkins,
4. Baliton Baltimore, 2nd edition,1999. Daffner R. H. Clinical radiology, the essentials. Williams and Wilkins,Baltimore, 2nd edition,1999.
5. Bram van Ginneken, Shigehiko Katsuragawa, Bart M. ter Haar Romeny, KunioDoi, and Max A. Viergever, Member, IEEE, Automatic Detection of Abnormalities in Chest RadiographsUsing Local Texture Analysis
6. T. F. Cootes, C. J. Taylor, D. Cooper, and J. Graham,Active shape models and their training and application,Comput.Vis.Image Understanding,1995.
7. G. Behiels, D. Vandermeulen, F. Maes, P. Suetens, and P. Dewaele, Active shape modelbased segmentation of digital X-ray images, in Lecture Notes in Computer Science. Berlin,Germany: Springer-Verlag,1999,
8. Rui Shen, Student Member, IEEE, Irene Cheng, Senior Member, IEEE, and Anup Basu, Senior Member, IEE ,EA Hybrid Knowledge-Guided Detection Technique for Screening of In- fectious Pulmonary Tuberculosis from Chest Radiographs,IEEE Transaction On Biomedical Engineering, VOL. 57,NO. 11, November,2010
9. K. Sum and P. Cheung, Boundary vector field for parametric active contours, Pattern Recognit.,vol. 40, no. 6, pp. 16351645, 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. X. Xie and M. Mirmehdi, MAC: Magnetostatic active contour model, IEEE Trans. Pattern Anal. Mach. Intell., vol. 30, no. 4,May 2008.

12. B. van Ginneken, B. ter Haar Romeny, and M. Viergever, Computeraided diagnosis in chest radiography: A survey, IEEE Trans.Med. Imag., vol. 20, no. 12, pp. 12281241, Dec. 2001.

13. D. Comaniciu and P. Meer, Mean shift: A robust approach toward feature space analysis, IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 5, May 2002

14. N. Ray and S. Acton, Active contours for cell tracking, in Proc. SSIAI, 2002, pp. 274278.

15. V. Ezhil Swanly, L. Selvam, P. Mohan Kumar, J. Arokia Renjith, M. Arunachalam, and K. L. Shunmuganathan,Smart Spotting of Pulmonary TB Cavities Using CT Images,Computational and Mathematical Methods in Medicine Volume 2013,

16. P. Annangi, S. Thiruvenkadam, A. Raja, Hao Xu, Xiwen Sun, Ling Mao,A region based active contour method for x-ray lung segmentation using prior shape and low level features,IEEE International Symposium on Biomedical Imaging , pp. 892-895, 2010

17. T.F. Chan and L.A Vese, Active contours without edges, IEEE Trans. Image Processing, vol.10, no. 2, 2001.

18. Yunmei Chen, Hemant Tagare, et al., Using prior shapes in geometric active contours in a variational framework., IJCV, vol. 50, no. 3, 2002.

19. Stefan Jaeger, Alexandros Karargyris, Sameer Antani, and George Thoma,Detecting Tuberculosis in Radiographs Using Combined Lung Masks,International Conference of the IEEE EMBS,September, 2012

20. Yonghong Shi, Feihu Qi, Zhong Xue, Liya Chen, Kyoko Ito, Hidenori Matsuo, and Dinggang Shen,Segmenting Lung Fields in Serial Chest Radiographs Using Both Population-Based and Patient-Specific Shape Statistics,IEEE Transactions on medical imaging, vol. 27, no. 4, april 2008

21. B. V. Ginneken, A. F. Frangi, J. J. Staal, B. M. T. H. Romeny, and M. A. Viergever, Active shape model segmentation with optimal features, IEEE Trans. Med. Imag., vol. 21, no. 8, pp. 924933, Aug. 2002

22. D. Lowe, Distinctive image features from scale-invariant keypoints,Int .J. Comput. Vis., vol. 60, no. 2, pp. 91110, Nov. 2004.

23. Stefan Jaeger, Alexandros Karargyris, Sema Candemir, Les Folio, Jenifer Siegelman, Fiona Callaghan, Zhiyun Xue, Kannappan Palaniappan, Rahul K. Singh, Sameer Antani, George Thoma, Yi-Xiang Wang, Pu-Xuan Lu, and Clement J. McDonald,Automatic Tuberculos Screening Using Chest Radiographs,IEEE Transactions on medical imaging, vol. 33, no. 2, february 2014

24. N. Dalal and B. Triggs, Histograms of oriented gradients for human detection, in Int. Conf. Comp. Vision Patt. Recog., vol. 1, 2005, pp. 886893.

25. S. Candemir, S. Jaeger, K. Palaniappan, S. Antani, and G. Thoma, Graph-cut based automatic lung boundary detection in chest radiographs, in IEEE Healthcare Technology Conference: Translational Engineering in Health and Medicine, 2012, pp. 3134.

26. L. Chen, R. Feris, Y. Zhai, L. Brown, and A. Hampapur, An integrated system for moving object classification in surveillance videos, in Int.Conf. Advanced Video and Signal Based Surveillance, 2008, pp. 5259

27. X. Wang, T. X. Han, and S. Yan, An HOG-LBP human detector with partial occlusion handling, in Int. Conf. Computer Vision, 2009, pp. 3239.

28. T. Ojala, M. Pietikainen, and T. Maenpaa, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 24, no. 7,pp. 971987, 2002.

29. Bram van Ginneken*, Alejandro F. Frangi, Joes J. Staal, Bart M. ter Haar Romeny, and Max A. Viergever, Active Shape Model Segmentation With Optimal Features,IEEE transactions on medical imaging, vol. 21, no. 8, august 2002

30. G. N. Srinivasan, and Shobha G.,Statistical Texture Analysis,world academy of science, engineering and technology volume 36 december 2008

31. S. Candemir, K. Palaniappan, and Y. Akgul, Multi-class regularization parameter learning for graph cut image segmentation, in International Symposium on Biomedical Imaging, 2013, pp. 14731476.

32. C. Daley, M. Gotway, and R. Jasmer, Radiographic manifestations of tuberculosis, A primer for Clinicians. San Francisco: Francis J. Curry National Tuberculosis Center, 2009.

33. Y. Boykov and G. Funka-Lea, Graph cuts and efficient n-d image segmentation, Int. J. Computer Vision, vol. 70, pp. 109131, 2006.

34. Shawn Lankton, Student Member, IEEE, and Allen Tannenbaum, Member, IEEE , Localizing Region-Based Active Contours, IEEE transactions on image processing, vol. 17, no. 11, november 2008