International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-7 Issue-4, September 2018

Segmentation of the Lungs from Chest X-Rays: A Simplified Computer Aided Approach

Prashant A. Athavale, H. D. Kattimani, P. S. Puttaswamy

Abstract*: The lungs reflect the health condition of a person, and hence it has been imaged and analysed by diagnosticians for over a century, and it requires knowledge and experience. The human observer's time and effort could be used productively if the lung image analysis is automated. This is especially true in case of screening of the lung chest x-rays. The lung segmentation is by default the first of a series of steps to analyse and interpret the images using a computer. One of the traditional approaches to segmentation of the lungs has been the use of statistical models, and the other is the rule based approach. This paper proposes a fusion method to segment the lungs on chest x-rays, as this modality of imaging is low cost, easy to operate, and gives first-hand information required for diagnosis. The results that are obtained are fast and promising accuracy has been documented. The entire approach can be extended to any organ on a medical image, or any object of interest in a general segmentation problem

Keywords: Lungs Segmentation, Active Shape Models, Computer Aided Diagnosis, Morphological Operation

I. INTRODUCTION

An unknown genetic change happening in the basic building block of the human body, the cell, leads to a change in the normal life cycle of it. The birth of a new cell, its growth, division and finally its removal from the system - all happen in a well-established normal cyclic format guided by the body's control system. This cycle could be affected by various natural and external factors leading to the uncontrollable growth of the cells. The growth of these cells forms a mass or a nodule commonly known as tumour. Such tumours are classified as benign or malignant; the latter is cancerous and can develop in any part of the human body. The medical field classifies the cancer based on the origin of such tumour. Carcinomas are the cancer prevalent in organs like lungs, breast, prostate etc. and form solid masses. The affected cells may be transferred to other parts of the body via blood and eventually develop in other healthy parts. This process is known as metastasis. A diagnostician can conduct several tests to diagnose the condition of the patient. Medical imaging is one of such way to know more about the body and its condition.

Of the 200 varieties, lung cancer is a common type of cancer in the adults and affects millions of people around the world. A timely diagnosis and detection together with proper treatment can cure the lung cancer. But the detection of the cancer in the lungs could be accidental in many cases. A low cost non-invasive method of the diagnosis of the lung cancer as against the biopsy is the analysis of digital chest x-rays (CXR). The Computed Tomography (CT) is the expensive imaging alternative, seldom found in rural areas. The location, size and other details of the nodules are visible on a CXR, and the diagnostician is trained in interpreting the images and classifying them. But during routine check-up, the nodules could be missed by a human observer, and there is inter-observer variability in the interpretation, Qin et al [1].

This gives rise to a requirement of a computer aided detection of such nodules on the CXR. The computer can be trained to detect the nodules and classify the same. Early research work aimed at replacing the doctor with the computers, which could mostly be attributed to the early hype created by the artificial intelligence.

But the later works have been focusing on making the computer generated information available to the doctor and assist him in his decision making. This could reduce the actual nodules being missed by the doctor. This use of computers in the diagnosis of the medical condition is known as Computer Aided Diagnosis (CAD).

The CXR database and the challenges posed by the dataset:

In computer aided diagnostic systems, the computer analyses the given image, and returns its findings, which are in-turn given to the diagnostician. If the findings are negative, the case is closed. If there are any candidate nodules or lesions, then the output of the CAD is further analysed by the diagnostician. This reduces the burden on the human being, and can enhance the productivity, thereby improving the patient care. An accurate determination of even the subtle nodules on the image at an early stage could be lifesaving.

A typical CAD system has the following sequence of operations for detecting the nodules on the lung images: Pre-processing, Segmentation of the Lungs, removal of the rib shadows, detection of the nodules, characterization and classification of the nodules as shown in Fig 1.

Revised Manuscript Received on 30 September 2018.
***** Correspondence Author

Prashant A. Athavale, Department of Electrical & Electronics Engineering BMS Institute of Technology & Management, Bengaluru, India

H. D. Kattimani, Department of Electrical & Electronics Engineering BMS Institute of Technology & Management, Bengaluru, India

Dr. P. S. Puttaswamy, Department of Electrical & Electronics Engineering, PES College of Engineering, Mandya, India

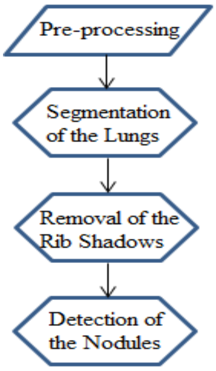

Fig. 1 A flow-chart of the typical steps in a CAD system. Pre-processing is to prepare the image for the next stages

And enhances the signal content in it. Each stage of the above mentioned crucial steps has a plethora of possible approaches [2,3]. The choice and the implementation could be guided by the problem at hand. For the study purpose, the images from the Japanese Society of Radiological Technology (JSRT) have been considered. This is a publicly available database, with around 247 CXRs each of 2048 X 2048 sizes. This set is divided in to two groups of images based on the presence of the nodules. Among them 154 images are with nodules and 93 images are without nodules, Shiraishi et al [4].

There are a multitude algorithms and different approaches to the segmentation of the lungs on the CXR. Few of them are mentioned here. Convolutional neural networks have been used to detect the nodules, and an effort to reduce false positives has been carried out by Li et al [5]. Apart from the shape model the local grey level distribution is also modelled in the statistical models [6]. The appearance model based segmentation of the lungs is reported by Dieter et al. Another work reported is of Nagata et al [7], which uses rule based algorithm. The marking of the landmark points is dealt with considerable detail in the work of Dorde et al [8]. Yan et al [9] also present a segmentation algorithm for the two dimensional CXR.

The CXRs from the database pose various challenges to the segmentation algorithms, and hence in the following discussion the same is highlighted. Each of the CXRs is having unique multimodal distribution of the grey level intensities. Hence a thresholding operation cannot produce a satisfactory result.

An example image is shown in Fig. 2a and its histogram in Fig. 2b. The Fig 2.c to 2.h in the next rows show the binary images obtained based on various threshold values, 'T'. The operation is to classify the pixels as '1' if the value of it in input image Im is above the T, else group it under '0', shown mathematically as,

$$I_o(i,j) = \begin{cases} 0, & I_m(i,j) < T \\ 1, & I_m(i,j) \geq T \end{cases} \quad (1)$$

A satisfactory result is not given by any of those chosen values. This is because of the distribution of the grey level values across the image, with great variability. The image which results from the imaging system depends on attenuation of the x-rays by various muscles internally, and also depends on the position of the patient with respect to the imaging device. Hands and neck portion are visible in some images of the database, where as it not seen in remaining. A generalised segmentation algorithm may not work here.

A. Variability in shapes:

The 3D thorax is superimposed on a 2D image, and hence there is an overlap of the rib cage, lungs, heart and other organs like vertebrate, windpipes etc. In spite of this the lungs to be segmented are having a characteristic shape which is easily recognizable by a human observer. The diagnosticians are able to observe lungs alone by mentally subtracting the overlapping structures because of the training and experience. The same operation can be done by a computer but at some compromise in quality.

Hence only one approach cannot solve the problem of segmentation and requires multiple operations either in series or parallel, to obtain a satisfactory result.

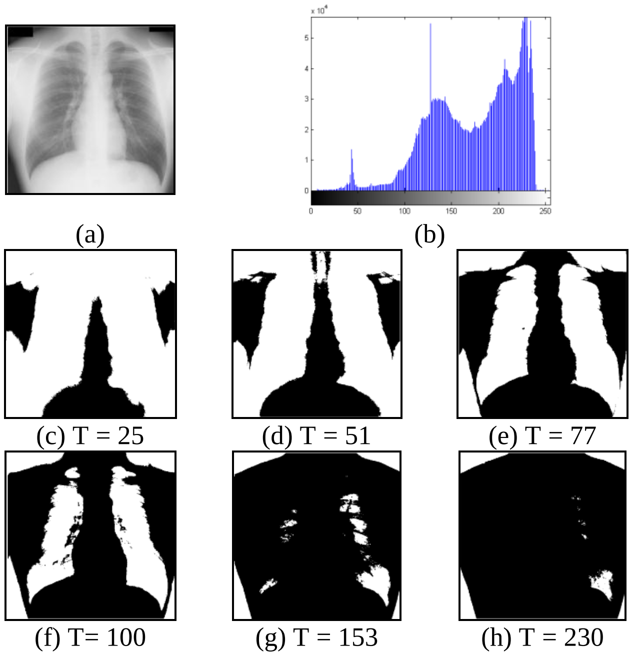

Fig. 2 (a) An input image (b) its histogram.(c) to (h) are the binary images obtained by thresholding the input image at various thresholds 'T'.

The Active Shape Model has been used for the segmentation of the lungs in CXRs and has given good results. But the limitation of these statistical models makes it vulnerable to local intensity features. All the points on the statistical model are annotated by an expert, and an overage of all such shapes from each training image is the model. The shape model so built is initially placed on the target image. The model has to stretch itself to fit to an instance of the lung in it, if present. The movement of the shape, in or out, is guided by local grey level model. The grey level model is a one dimensional sample of the image at that point, and then averaged.

This sample is searched along the normal to the boundary of the object, before settling at a best fit location. At each of the iterations, the model adjusts itself to the available instance of the lung and an energy minimization function decides the number of further iterations. In case the point deviates from the actual boundary, then it will never return to the correct edge of the object as there is no reset option in ASM as described in [10].

B. Refinement of the results:

The use of the ASM is restricted to only a coarse shape of the lungs, and its salient features like diaphragm, Costophrenic (CP) angle and the apex of the two lobes. Further refinement of the results to obtain an accurate segmentation of the lungs is done by the application of morphological operations. This is interspersed with rule based approaches also.

The above mentioned approach has the following advantages:

(i) lesser model points on the ASM. The statistical model can be built with only few annotated points, and hence the complexities of computations are overcome. The lateral sides of the lungs are almost linear and hence fe points are sufficient to describe them. The major curvature is at the CP angle, and a couple of landmark points are used in modelling it. (ii) The model does not settle at non-edge location: Since only first of the iteration is used in getting the location of the salient features as mentioned before, the problem of the model swayed away by a non-edge location is overcome.

(iii) The method is faster (iv) the operations are guided by the local grey level distributions only. For processing a portion of the image, a suitable region is carved out and it makes the process independent of the grey level distribution elsewhere in the image. (v) The two lungs are not modelled separately in the shape model, but are considered as a bell shaped structure. This will remove the two lungs from the rest of the image in first stage and then separate them more easily. (vi)The result accuracy is higher. Various stages of the proposed method are shown in a collage of images in Fig 3.

II. RESULTS AND DISCUSSION:

The proposed method has been tested on a select 35 images from the same database, and these were not part of the set of training images. One of the authors has segmented the right and left lung regions manually, under the guidance of an expert. This has been considered as the standard for comparison of the automated segmentation [11]. The quantification of the results has been done based on the standard evaluation metrics: Sensitivity, specificity and accuracy. These are defined in equations 2, 3, 4, in terms of True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN).

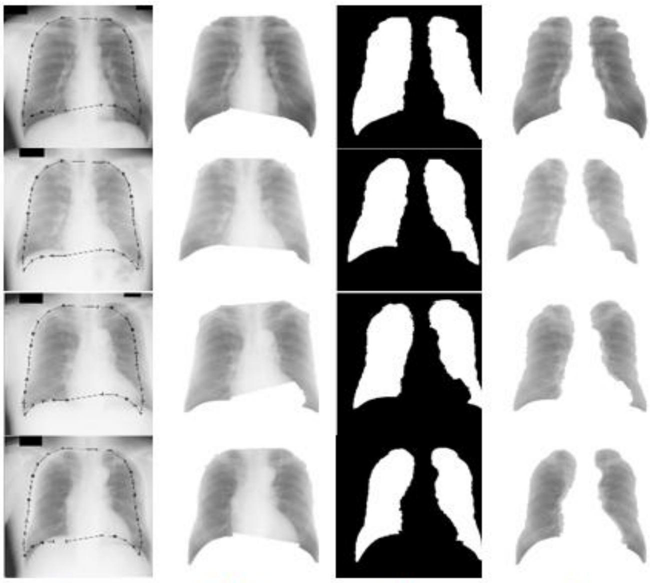

Fig. 3 The intermediate results of application of the method on some of the images. (a)The first column shows the input image from the database, superimposed with the ASM lungs model. (b) The second column is the corresponding outermost edges detected by the ASM. (c) Third column is the binary version of separation of the right and left lungs. (d) Fourth column is the detected right and left lungs.

These are defined as: TP is the set of pixels correctly classified as belonging to the lungs and TN are the set of pixels correctly classified as not belonging to the lungs. By the same token, FP is the total number of pixels falsely classified as belonging to the lungs and FN is the total number of pixels falsely marked as non-lung pixels.

$$\text{Sensitivity} = \frac{TP}{TP+FN} \quad (2)$$

$$\text{Specificity} = \frac{TN}{TN+FP} \quad (3)$$

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

These parameters and the corresponding values have been tabulated in Table No. 1. The proposed algorithm was tested on three image sizes as depicted. These images were obtained by down- sampling the original image. Correspondingly the time requirement is also less for the segmentation for lower size image.

Table 1: Performance Indicators and Corresponding Values

Parameter	512 X 512	1024 X 1024	2048 X 2048
Sensitivity	0.862	0.87	0.88
Specificity	0.821	0.833	0.854
Accuracy	0.92	0.91	0.901

III. CONCLUSION

The problem of segmentation of the lungs can be approached in multiple ways, but each approach has its own limitations, and a hybrid approach is required to refine the results, to be of any clinical importance. All the objects of interest in a medical image are bound by a characteristic shape, but vary with in a permissible variability.

Modelling the shape and including in it the variability is a challenging task, as the model depends on the training dataset. In the proposed method, the model is dependent on the training dataset, but not its accuracy as it is a hybrid method. The simple framework explained in this paper can be extended to any shape on medical or other images.

REFERENCES

1. Chunli Qin, Demin Yao, Yonghong Shi, Zhijian Song, Computer Aided detection in chest radiography based on artificial intelligence: A survey, BioMedical Engineering On line 2018.doi:10.1186/s12938-018-0544
2. Story A, Aldridge RW, Abubakar I, Stagg HR, Lipman M, et al. (2012) Active case finding for pulmonary tuberculosis using mobile digital chest radiography: An Observational Study. International Journal of Tuberculosis and Lung Disesase 16: 1461–1467.
3. Shi, Yonghong, Feihu Qi, Zhong Xue, Kyoko Ito, Hidenori Matsuo, and Dinggang Shen. "Segmenting lung fields in serial chest radiographs using both population and patient-specific shape statistics." In International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 83-91. Springer, Berlin, Heidelberg, 2006.
4. J. Shiraishi, S. Katsuragawa, J. Ikezoe, T. Matsumoto, T. Kobayashi, K. Komatsu, M. Matsui, H. Fujita, Y. Kodera, and K. Doi, "Development of a digital image database for chest radiographs with and without a lung nodule: receiver operating characteristic analysis of radiologists' detection of pulmonary nodules", American Journal of Roentgenology, vol. 174, p. 71-74, 2000.
5. Chaofeng Li, Guoce Zhu, Xiaojun Wu, Yuanquan Wang, 'False-Positive Reduction on Lung Nodules Detection in Chest Radiographs by Ensemble of Convolutional Neural Networks' IEEE Access, Vol-6, 2018. DOI: 10.1109/ACCESS.2018.2817023
6. Seghers, Dieter 'Image segmentation using local shape and gray-level appearance models', Medical Imaging 2006: Image Processing. Vol. 6144. International Society for Optics and Photonics, 2006.
7. Nagata, Ryoichi, Tsuyoshi Kawaguchi, and Hidetoshi Miyake. "A Rule-Based Algorithm for Detection of Ribcage Boundary in Chest Radio-graphs." In ITC-CSCC: International Technical Conference on Circuits Systems, Computers and Communications, pp. 1001-1004. 2009.
8. Starcevic, Đorđe S., Vladimir S. Ostojic, and Vladimir S. Petrovic. 'An open-source digital diagnostic radiography image annotation software', In Telecommunications Forum (TELFOR), 2016 24th, pp. 1-4. IEEE, 2016.
9. Zhennan Yan, Jing Zhang, Shaoting Zhang, and Dimitris N. Metaxas, 'Automatic Rapid Segmentation of Human Lung from 2D Chest X-Ray Images', MICCAI workshop on Sparsity Techniques in Medical Imaging, 2012.
10. Prashant A. Athavale, P. S. Puttaswamy, 'Median Values of Gray Levels for Detection of Lung Contours', International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT-2015) PESCE, ISBN: 978-1-4673-9563-2
11. Muyoyeta M, Maduskar P, Moyo M, Kasese N, Milimo D, Spooner R, 'The Sensitivity and Specificity of Using a Computer Aided Diagnosis Program for Automatically Scoring Chest X-Rays of Presumptive TB Patients Compared with Xpert MTB/RIF in Lusaka Zambia'. PLoS ONE 9(4): e93757. (2014) https://doi.org/10.1371/journal.pone.0093757