Lung Boundary Detection in Pediatric Chest X-rays

Sema Candemir*[a], Sameer Antani[a], Stefan Jaeger[a], Renee Browning[b], George Thoma[a]
[a] Lister Hill National Center for Biomedical Communications, U.S. National Library of Medicine, National Institutes of Health, Bethesda, MD, 20894, USA.
[b] National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, 20894, USA.

ABSTRACT

Tuberculosis (TB) is a major public health problem worldwide, and highly prevalent in developing countries. According to the World Health Organization (WHO), over 95% of TB deaths occur in low- and middle- income countries that often have under-resourced health care systems. In an effort to aid population screening in such resource challenged settings, the U.S. National Library of Medicine has developed a chest X-ray (CXR) screening system that provides a pre-decision on pulmonary abnormalities. When the system is presented with a digital CXR image from the Picture Archive and Communication Systems (PACS) or an imaging source, it automatically identifies the lung regions in the image, extracts image features, and classifies the image as normal or abnormal using trained machine-learning algorithms. The system has been trained on adult CXR images, and this article presents enhancements toward including pediatric CXR images. Our adult lung boundary detection algorithm is model-based. We note the lung shape differences during pediatric developmental stages, and adulthood, and propose building new lung models suitable for pediatric developmental stages. In this study, we quantify changes in lung shape from infancy to adulthood toward enhancing our lung segmentation algorithm. Our initial findings suggest pediatric age groupings of 0 - 23 months, 2 - 10 years, and 11 - 18 years. We present justification for our groupings. We report on the quality of boundary detection algorithm with the pediatric lung models.

Keywords: Pediatric chest X-rays, Lung boundary detection, Model-based segmentation, Tuberculosis.

1. INTRODUCTION

Tuberculosis (TB) is a major public health problem worldwide, and is highly prevalent in developing countries. According to the World Health Organization (WHO) estimates, there were 8.6 million new TB cases in 2012, and as many as 1 in 10 TB cases globally are among the less than 15 year old age group, but the number may be higher because many children are simply undiagnosed [1]. It is estimated that there are around 1 million cases of TB in children younger than 15 years living in developing countries, with approximately 74,000 deaths annually [2]. Over 95% of the TB cases and deaths occur in low- and middle-income countries that often have under-resourced health care systems. In an effort to aid population screening in such resource challenged settings, the U.S. National Library of Medicine (NLM) has developed a chest X-ray (CXR) screening system that provides a pre-decision on pulmonary abnormalities and could be used to minimize the workload. The system has been trained on adult CXR images, and we published our findings in [3] [4]. This article presents enhancements toward including pediatric CXR images.

In our current system, when a digital CXR image is presented from the PACS or an imaging source, it automatically identifies the lung regions, extracts image features, and classifies it as normal or abnormal using trained machine-learning algorithms. In expanding the system capability to include pediatric CXRs, we need to be aware of change in the lung shape during the pediatric developmental stages, as shown in Figure 1. For example, in an infant, the lungs are smaller, have a triangular shape, and the cardiac contour is relatively large such that the horizontal diameter of the heart may approach 60% of thoracic horizontal diameter [5].

Our lung boundary detection algorithm is model-based [3]; therefore, we built up a pediatric lung model set. A schematic of the segmentation stage of the screening system is shown in Figure 2. As background, the segmentation algorithm registers the patient CXR to images from an expert-marked atlas by computing a correspondence mapping

*sema.candemir@nih.com; candemirsema@gmail.com; office: (301)-496-7273;

from the target to the model. Then an inverse transform is applied to the model lung boundary to obtain the patient specific lung boundary. In order to decrease the registration error, and increase the boundary accuracy, this process is repeated on several models, and we average the lung boundary results. The details of the full system can be found in [3].

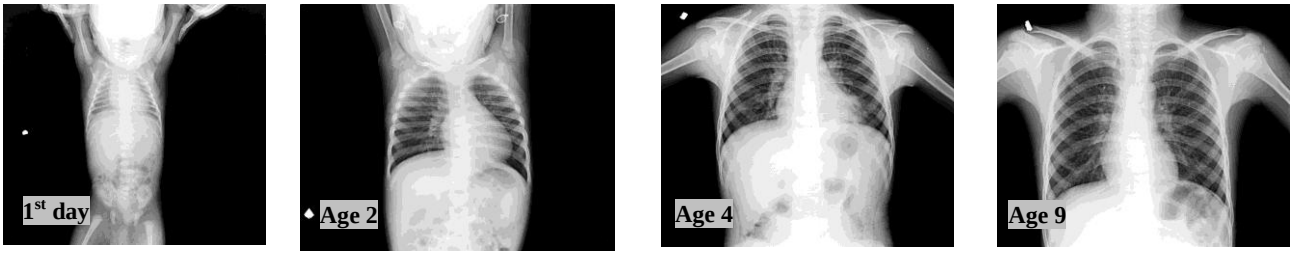


Figure 1: Example pediatric XCRs.

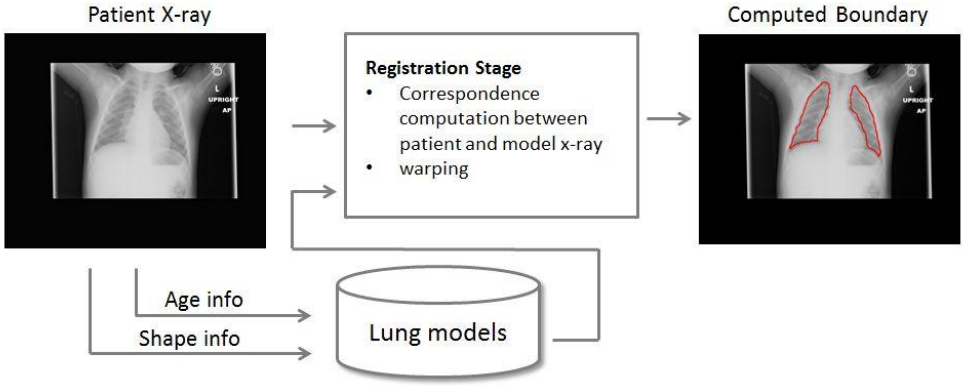


Figure 2: The segmentation approach for adult and pediatric CXR.

As reported in Table 3, we tested the system on pediatric images and we observed that adult lung models are not adequate for the pediatric cases. We note that the algorithm performance is much lower than the peak adult lung segmentation Dice score of 95.4 ± 0.015, reported in [3]. We attribute this to changes in the lung shape during the pediatric developmental stages. Consequently, in this study, we quantify changes in lung shape from infancy to adulthood and propose building new lung models suitable for various pediatric developmental stages to improve algorithm performance. Our initial findings suggest pediatric age groupings of 0 - 23 months, 2 - 10 years, and 11 - 18 years. We present justification for our groupings in the section 2, and report the quality of boundary detection algorithm with age group-based pediatric lung models in Section 3.

2. QUANTIFYING VARIATIONS IN THE PEDIATRIC LUNG SHAPE

In order to develop meaningful pediatric lung shape models, it is necessary to identify age groups within which there is minimal change in the overall lung shape. We have 161 pediatric images within a data set containing 397 CXRs from India. For comparison, we also randomly select 25 CXRs from our adult CXR image data set from the "Montgomery Dataset" maintained by NLM. This set contains 138 CXRs from the Department of Health and Human Services, Montgomery County, Maryland. Both data sets were de-identified at source and their use is exempted under NIH IRB OHSR # 5357 valid through Dec. 31, 2018. Table 1 summarizes the number of CXRs in each age group. The lung boundaries in these images have been manually delineated using the Firefly tool [6] and are used to generate binary lung mask images. As shown in Figure 3 and 4 respectively, we compute the average image within each group and the average image of the lung masks.

Table 1: CXRs groups according to the patient age

Pediatric Group 1	52 CXRs	1 Day ≤ Age < 2Years
Pediatric Group 2	67 CXRs	2 Years ≤ Age < 11 Years
Pediatric Group 3	42 CXRs	11 Years ≤ Age < 18 Years
Adult	25 CXRs	Age ≥ 18 Years

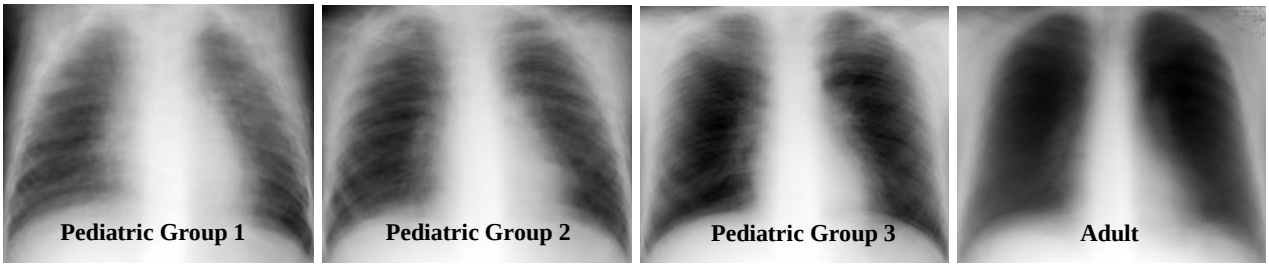


Figure 3: Average image of pediatric CXRs according to groups listed in Table 1.

Qualitatively, we note from the images in these figures that the lung shapes within-class are similar, but there are noticeable visual differences between the groups. In order to quantify this observation, we compute the mean boundaries within each group, shown as colored contours in Figure 5. Next, as reported in Table 3, the average contour distance (ACD) of each lung boundary to the mean boundary of each group is computed. ACD is the average distance between the mean boundary M and each lung boundary L. Let a_i and b_j be the points on the boundary L and M, respectively. The minimum distance of point a_i to the boundary M is defined as follows:

$$d(a_i, M) = \min \| b_j - a_i \| \tag{1}$$

For ACD, the minimum distance for each point on the boundary L to the contour M is computed. Then, the distances are averaged over all points of boundary L. In order to make the similarity measure symmetric, the computation is repeated from contour M to contour L. Following is the formulation for ACD,

$$ACD(L, M) = \frac{1}{2}\left(\frac{\sum_i d(a_i, M)}{|\{a_i\}|} + \frac{\sum_j d(b_j, L)}{|\{b_j\}|} \right) \tag{2}$$

where $|.|$ is the cardinality of the set.

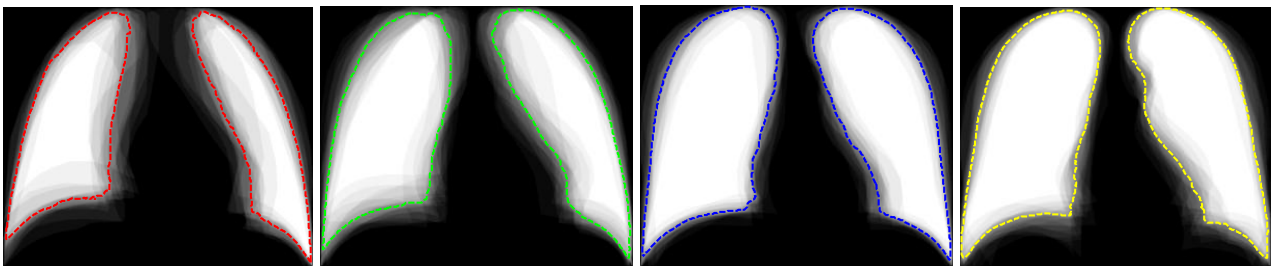


Figure 4: Average lung mask images in each pediatric age group. The dashed contours indicate the average contour.

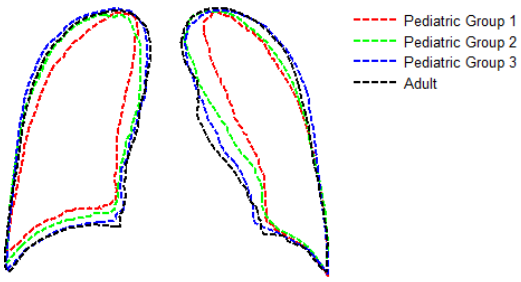


Figure 5: Superimposed mean lung boundaries from each pediatric age group demonstrate (i) lung shape development, and (ii) high similarity in lung shape between pediatric group 3 and adults (shown in black for clarity)

Table 2: Average contour distances of lung boundaries within-class and between classes

	Average Contour Distance			
	Mean Boundary of Pediatric Group 1	**Mean Boundary of Pediatric Group 2**	**Mean Boundary of Pediatric Group 3**	**Mean Boundary of Adult Lungs**
Pediatric Group 1	**6.8042**	8.9769	11.1652	10.9807
Pediatric Group 2	8.5739	**5.9532**	6.7504	6.8543
Pediatric Group 3	10.5131	**5.8211**	**4.6895**	**5.2760**

As shown in Table 2, the average contour distance of the mean boundary with the lung boundaries within each group is minimum. However, we note that the distance is higher across groups, suggesting that the pediatric age groups selected, following the initial experiment, are reasonable. Further, we observe that there is a very small difference (ACD = 5.276) between the average contour distance of the mean boundary of the adult lung and pediatric group 3. This suggests that the lungs for youth older than age 11 are reasonably mature, and very similar to the adult lung shapes. This is can be verified visually on the superimposed mean lung shape outlines shown in Figure 5. As a result we expect that it will be necessary to develop at least 2 pediatric lung models for ages 0 - 23 months and 2 - 11 years. It may be necessary to expand the adult lung model to include lung shapes from pediatric ages 11 years and older.

3. EXPERIMENTAL RESULTS

In this section, we report the lung boundary detection algorithm performance on pediatric CXRs. As we reported in Section 2, we have 161 pediatric images. We grouped them into 3 classes according to the patient age (cf. Table 1), and created pediatric models for each age group. Pediatric models are vectors which contain the horizontal and the vertical profiles of each chest X-ray.

As a first experiment, we run the segmentation algorithm on each pediatric group to see their response to adult lung models. The segmentation quality is measured with the following evaluation metrics: average contour distance (ACD), overlap measure (the Jaccard similarity coefficient) and Dice`s similarity. ACD is also used to measure the distance between the average lung mask and lung boundaries in each group, and is defined in Section 2. The overlap measure is computed with the following formulation,

$$\Omega = \frac{|S \cap GT|}{|S \cup GT|} \tag{3}$$

and, Dice`s coefficient is computed with the following formulation,

$$Dice = \frac{|S \cap GT|}{|S| + |GT|} \tag{3}$$

where S is the segmentation mask and GT is the radiologist marking as ground truth boundary. The Table 3 summarizes the system performance on each pediatric group.

Table 3: The performance of the system (which is tuned for adult lungs) on each pediatric group. Average overlap, average dice and average contour distance scores between the ground truth boundary and the computed segmentation boundary for each pediatric CXR.

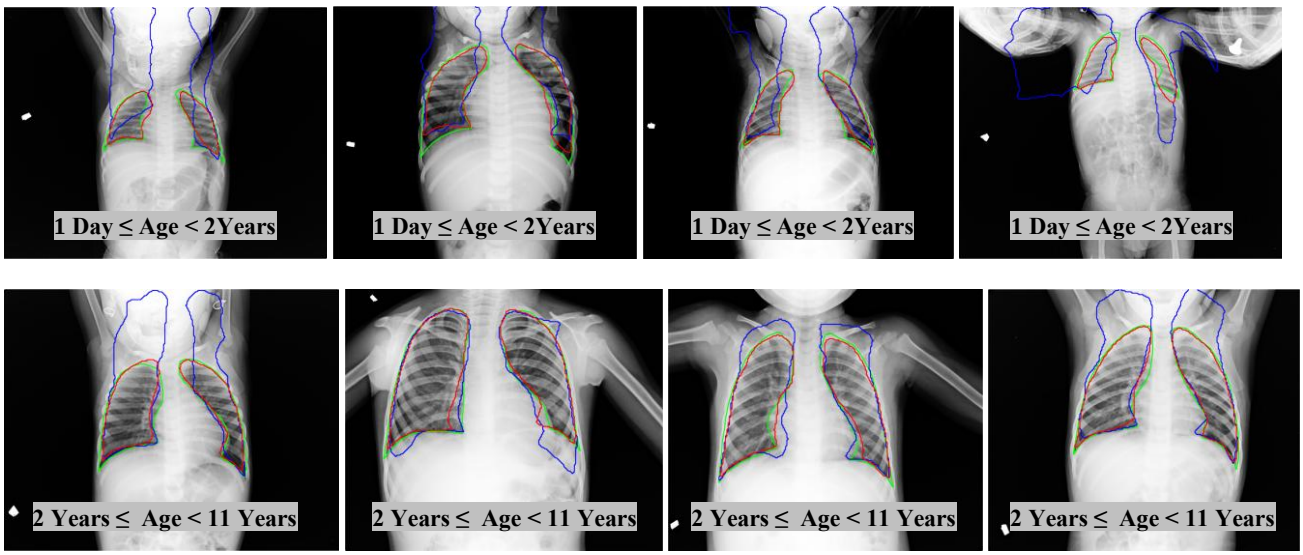
Test Set	Age Range	Model Set	Overlap Ω	Dice Score	ACD
Pediatric Group 1	1 Day ≤ Age < 2Years	Adult	26.612	36.837	148.365
Pediatric Group 2	2 Years ≤ Age < 11 Years	Adult	74.259	83.232	33.316
Pediatric Group 3	11 Years ≤ Age < 18 Years	Adult	88.460	93.761	15.307

As we observed in Section 2, the lungs for youth older than age 11 years are very similar to the adult lung shape. This observation was confirmed by the system performance on Pediatric Group 3 which contains patients older than 11 years old. However, it did not perform well on younger patients, especially those younger than 2 years old. This experiment showed that, in order to improve the system performance on youths younger than age 11 and infants, it is necessary to adapt the system to the pediatric CXRs. Therefore, we modified the module as in Figure 2. When a CXR image is presented to the system, it extracts the age information from the metadata and selects the appropriate model set. If patient is younger than 2 years old, the model set for the algorithm will be group 1; if the patient age is between 2 and 11, then the model set will be Group 2; and if the patient is older than 11, then model set will be Group 3. In order to measure the extended system performance, we repeat the first experiment with the new model sets. The Table 4 summarizes the expanded system performance on pediatric CXRs groups if lung models are selected according to the patient age.

Table 4: Expanded system performance on pediatric CXRs.

Test Set	Age Range	Model Set	Overlap Ω	Dice Score	ACD
Pediatric Group 1	1 Day ≤ Age < 2Years	Selected according to age	76.384	85.966	15.355
Pediatric Group 2	2 Years ≤ Age < 11 Years	Selected according to age	86.190	92.534	11.107
Pediatric Group 3	11 Years ≤ Age < 18 Years	Selected according to age	89.674	94.524	12.276

According to the results, the system performance is slightly increased with the pediatric models for patients older than 11. The performance increased 10% (according to the overlap measure) for second group (patient age is between 2 to 11). However the performance is drastically increased for the first group (patients younger than 2), from 26% overlap to 76% overlap. Figure 6 shows the visual results of the expanded-system. The first column contains example CXRs from group 1; the CXRs from group 2 are in column 2; and the CXRs from group 3 (patients older than 11) are in column 3. The green contour is the ground truth, the blue contour is the boundary computed with adult models, and the red contour is the boundary computed by selecting the model according to patient age. As can be seen from the figures, for Group 1 and Group 2, age-based pediatric models helped to improve the system performance.



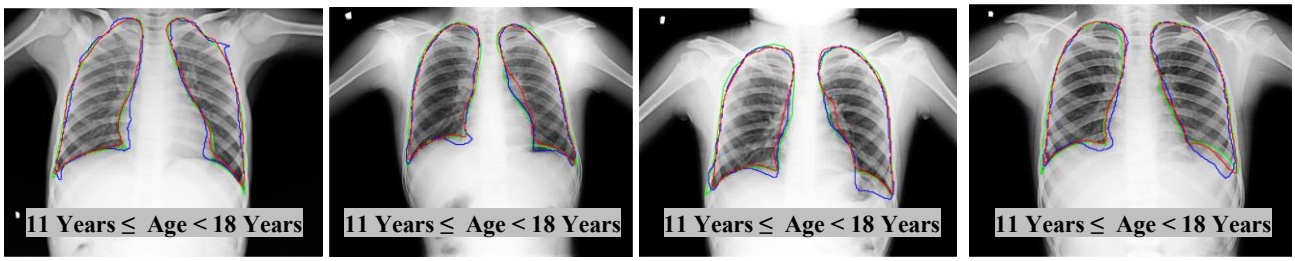


Figure 6: Segmentation results on example CXRs. Green contour is the GT, blue is the segmentation boundary with adult models, and red contours are the boundaries with pediatric models which are selected according to patient age.

4. CONCLUSIONS

Lung boundary segmentation is a key first step toward screening digital CXRs for pulmonary abnormalities. Working with an appropriate lung model increases the accuracy of the registration, and hence the accuracy of the lung boundary segmentation. Our current CXR screening system is tuned for adult lung images, and we observed that the system performed less accurately on pediatric CXRs. We also observed that there are varying lung shapes at different pediatric developmental stages. Therefore, we intended to expand the current system to include the pediatric CXRs. We modified the system and then observed the system performance with new models. The system boundary detection performance increased 10% (according to the overlap measure) for patients whose age is between 2 to 11; and increased drastically for patients younger than 2 (overlap score is increased from $\Omega = 26\%$ to $\Omega = 76\%$).

Acknowledgments

This research was supported by the Intramural Research Program of the National Institutes of Health (NIH), National Library of Medicine (NLM), and Lister Hill National Center for Biomedical Communications (LHNCBC). The authors would like to thank Dr. C.S. Durgi, Consulting Radiologist, Mediscan Diagnostic Care, and Associate Professor of Radiology, Mahadevappa Rampure Medical College, Gulbarga, India, and Dr. S. Qasba, Medical Director of Montgomery County`s TB control program, for providing CXR images. The authors thank to Dr. Devasena Gnanashanmugam, National Institute of Allergy and Infectious Diseases National Institutes of Health. The authors also thanks to Bryce Hammond for his help organizing the data.

REFERENCES

[1] World Health Organization. Global Tuberculosis report 2013. Available at http://apps.who.int/iris/bitstream/10665/91355/1/9789241564656_eng.pdf (last access is June 24, 2014).

[2] H. Jenkins, A. Tolman, C. Yuen, et al., "Incidence of multidrug-resistant tuberculosis disease in children: systemic review and global estimates", The Lancet, 383:1572-1579, (2014).

[3] Candemir, S., Jaeger, S., Palaniappan, K., Musco, J., Singh, R., Xue, Z., Karargyris, A., Antani, S., Thoma G., McDonald C., "Lung Segmentation in Chest Radiographs Using Anatomical Atlases with Non-rigid Registration." IEEE transactions on medical imaging, 33(2), 577-590, (2014).

[4] Jaeger, S, Karargyris, A., Candemir, S., Folio, L., Siegelman, J., Callaghan, F., Xue, Z., Palaniappan, K., Singh, R., Antani, S., Thoma G., Wang, Y.X., Lu, P.X., McDonald C.,"Automatic tuberculosis screening using chest radiographs", IEEE transactions on medical imaging, 33(2), 233-245, (2014).

[5] Arthur, R. "Interpretation of the paediatric chest X-ray." Paediatric respiratory reviews 1.1, 41-50, 2000.

[6] Firefly interactive labelling tool, http://firefly.cs.missouri.edu/# (last access is June 24, 2014).