

# Automatic x-ray screening for tuberculosis

Sameer Antani

*Automatic analysis of x-ray images enables on-site screening for tuberculosis infection, addressing the shortage of trained radiologists in rural Kenya.*

Tuberculosis (TB) is widespread among immunocompromised HIV-positive patients in developing nations that have scarce medical resources. Tackling the disease, while at the same time avoiding complications of antiretroviral (HIV) treatment, requires rapid-result tests for detecting infection. Previously, healthcare practitioners determined TB exposure using immune response skin tests. However, if the patient had previously been vaccinated, this could compromise the accuracy of the result, and the test was therefore not a true indicator of active disease. A recent definitive test for TB diagnosis based on DNA analysis<sup>1</sup> speeds up earlier sputum culture microscopy-based methods, but is expensive and inappropriate for population screening. Consequently, analyzing chest x-ray (CXR) images remains a relatively inexpensive and mandatory part of every patient evaluation. In developing regions, the challenge is to address the imbalance between the size of the affected population and the available radiology services. To achieve this aim—and to fulfill the objectives of the National Institutes of Health global health initiative—we have conducted research and development in advanced biomedical image diagnostics, and translated it into low-cost automated tools. This article highlights our work in developing and deploying a reliable, low-cost, automatic screening system for pulmonary TB in CXR images.

We have conducted population screening in western Kenya in collaboration with Academic Model Providing Access To Healthcare (AMPATH), an NGO that provides treatment to some 200,000 HIV-positive patients. For TB infection screening, we provide lightweight digital x-ray units that are mounted in custom-designed trucks (see Figures 1 and 2). Using a low-power computer, the system analyzes a digital CXR image immediately after acquisition (see Figure 3) using advanced image processing and machine learning algorithms, and



*Figure 1. An HIV-positive patient being led to the truck-mounted x-ray unit for tuberculosis (TB) screening.*



*Figure 2. Radiographer preparing a patient for x-ray imaging of the lungs.*

presents the decision in a simplified format to regional medical officers, who may have little or no background in radiology.

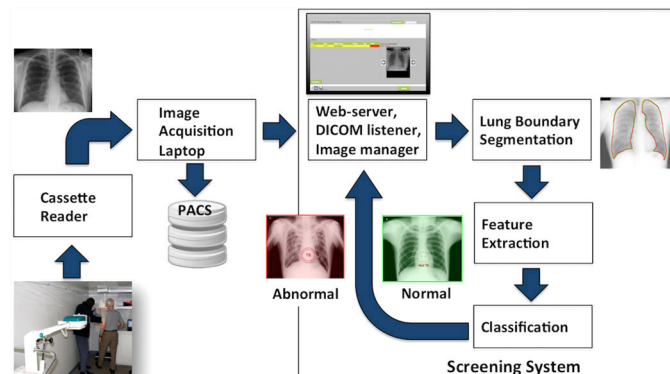
The screening process comprises a set of cascaded modules that communicate with the CXR imaging system (see Figure 4).

*Continued on next page*

A key first step is the automatic detection of the lung regions, for which we have developed a novel automated lung boundary detection algorithm for CXR images. We evaluated the system on publicly available datasets,<sup>2</sup> where it demonstrated accuracy in the range 92–95%, although the method is adversely affected by severe pleural effusion and pediatric cases<sup>3</sup> (and therefore requires further development). After detection, image features are extracted from within the localized lung, leading to a two-class normal/abnormal decision for the input CXR image. We analyzed several features with different classifiers and determined that multiscale image texture and edge curvature features yield the best performance when using a neural network (NN) classifier (a model based on biological neural networks) with one hidden layer.

CXR images can serve as both screening and diagnostic tools. In a population screening program, all participants who have CXR images that are either ‘suspicious for TB’ or ‘minimally abnormal’ are generally ‘intentionally over-read’ (that is, they would be flagged for follow-up testing, such as sputum culture or DNA tests).’ Similarly, we tuned the classifier to favor sensitivity over specificity, while keeping the overall accuracy (measured as the area under the receiver operating characteristic curve<sup>4</sup>) as high as possible. The NN classifier resulted in an overall accuracy of 99.3%, with a false positive rate of less than 7%.

We are also developing methods that will provide a screening decision without the need for accurate lung boundary segmentation and a support vector machine classifier (a supervised learning model). Our method uses multiscale-oriented edge analysis to find spurious contour patterns that could be



**Figure 4.** Schematic showing the image acquisition and screening workflow. DICOM: Digital Imaging and Communications in Medicine. PACS: Picture archiving and communications system.

indicative of disease. This method examines only the chest region bounded by the shoulders at the top, the sides of chest, and the diaphragm at the bottom, and uses an NN classifier. Using this method, we have achieved results comparable with the lung boundary segmentation approach, but 25 times faster.

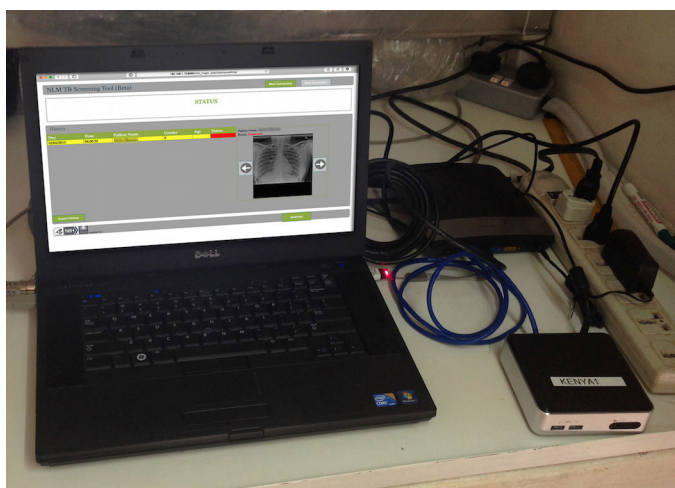
In summary, we have designed and developed a low-cost, reliable, and robust automatic CXR screening system for cardiopulmonary diseases, with a particular focus on TB detection because of its significance with HIV treatment. In collaboration with AMPATH, our future work will focus on advancing the system’s capability based on results from its use in the field.

*The author acknowledges contributions from the following researchers at the National Library of Medicine: Clement McDonald (MD), George Thoma, Sema Candemir, Stefan Jaeger, Santosh KC, Alexandros Karargyris, Zhiyun Xue, and Szilard Vajda (all PhD); and Marc Kohli (MD) Radiologist, Indiana University School of Medicine. All work reported here is conducted as part of the Intramural Research Programs of the Lister Hill National Center for Biomedical Communications, the US National Library of Medicine, and the National Institutes of Health.*

### Author Information

**Sameer Antani**  
National Library of Medicine  
Bethesda, MD

Sameer Antani is a scientist with expertise in biomedical imaging, informatics, computer science, and engineering technology.



**Figure 3.** The TB screening system (small black box on the right).

*Continued on next page*

His research advances biomedical sciences, global health, and the dissemination of scientific information. He is a senior member of SPIE.

#### References

1. [http://www.who.int/tb/features\\_archive/xpert\\_factsheet.pdf](http://www.who.int/tb/features_archive/xpert_factsheet.pdf) Details of the WHO DNA test for TB. Accessed 24 September 2015.
2. S. Jaeger, S. Candemir, S. K. Antani, Y. Wang, P. Lu, and G. R. Thoma, *Two public chest x-ray datasets for computer-aided screening of pulmonary diseases*, **Quant. Imag. Med. Surg.** **4**, pp. 475–477, 2014. doi:10.3978/j.issn.2223-4292.2014.11.20
3. S. Candemir, S. Antani, S. Jaeger, R. Browning, and G. R. Thoma, *Lung boundary detection in pediatric chest x-rays*, **Proc. SPIE** **9418**, p. 94180Q, 2015.
4. [https://en.wikipedia.org/wiki/Receiver\\_operating\\_characteristic](https://en.wikipedia.org/wiki/Receiver_operating_characteristic) Explanation of receiver operating characteristic curve. Accessed 18 September 2015.