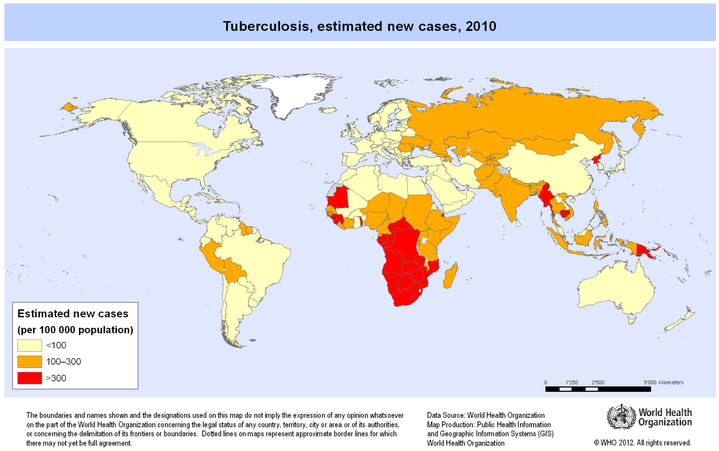
# Automatic Screening for Lung Diseases in Chest Radiographs: A Global Health Initiative

**Stefan Jaeger, Alexandros Karargyris, Sameer Antani, George Thoma**
**Lister Hill National Center for Biomedical Communications**
**U.S. National Library of Medicine, National Institutes of Health (NIH)**

## Tuberculosis

Tuberculosis (TB) is the second leading cause of death in the world killing at least 1.4 million people in 2010. Almost 95% of TB deaths are in the developing world.




It is a disease caused by germs that are spread from person to person through the air. TB usually affects the lungs, but it can also affect other parts of the body, such as the brain, the kidneys, or the spine. (CDC definition)

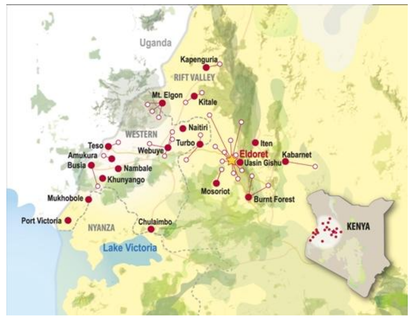
With an estimated 9 million new cases occurring every year, TB is a major global health problem. There are two important reasons contributing to this severity:

a) Opportunistic infections of HIV positive populations
b) Emergence of multi-drug resistance strains

TB is commonly diagnosed with the Mantoux skin test, the sputum test, or chest x-ray (CXR).

## TB Screening

### Our Region of Interest – Western Kenya




- ❑ 50 clinics stretched across 300 miles
- ❑ Only five clinics have x-ray machines.
- ❑ Only one clinic has direct radiologist access.

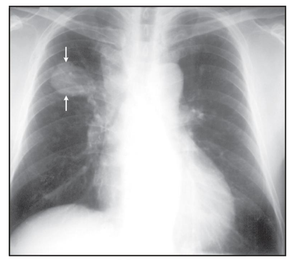
NATIONAL LIBRARY OF MEDICINE · USAID FROM THE AMERICAN PEOPLE · AMPATH ACADEMIC MODEL PROVIDING ACCESS TO HEALTHCARE PARTNERSHIP

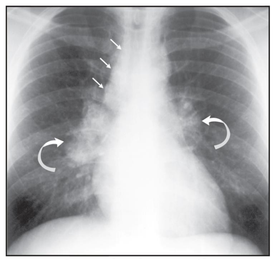
### NLM – AMPATH Collaboration

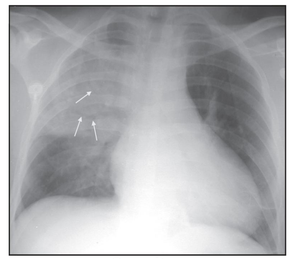
- ❑ AMPATH – The Academic Model Providing Access to Healthcare
- ❑ Partnership among USAID, five US universities and Moi University Medical School
- ❑ Runs huge AIDS treatment program (earliest and largest in Sub-Saharan Africa)
- ❑ Decided to screen everyone for TB, which is a major risk factor for AIDS patients
- ❑ Use of portable x-ray scanners

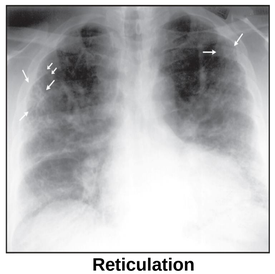
## TB Manifestations

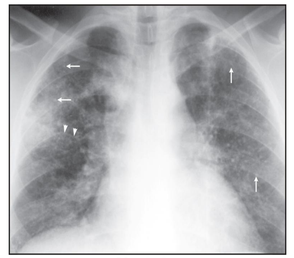
Different TB manifestations in CXRs. They vary in intensity, texture, and shape (Daley, Gotway, Jasmer).

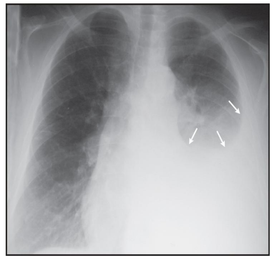

Mass

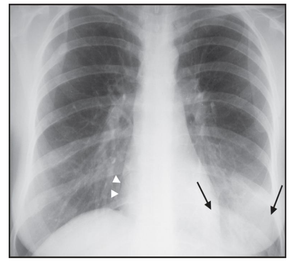

Lymphadenopathy

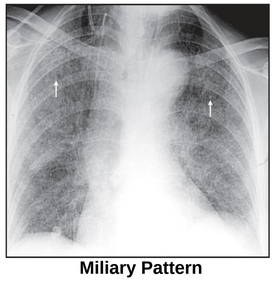

Opacity


Reticulation


Nodules


Pleural Effusion


Silhouette


Miliary Pattern

# Automatic Screening for Lung Diseases in Chest Radiographs: A Global Health Initiative


**Stefan Jaeger, Alexandros Karargyris, Sameer Antani, George Thoma**

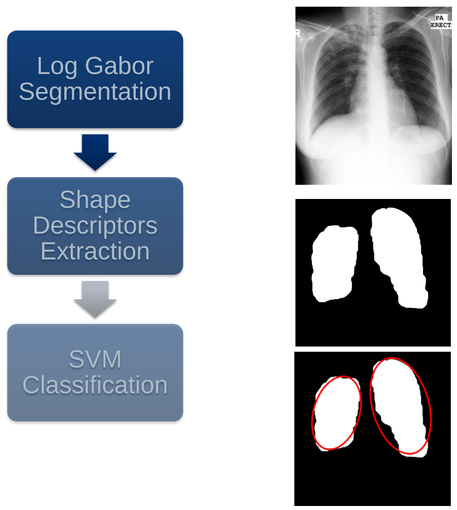
**Lister Hill National Center for Biomedical Communications**

**U.S. National Library of Medicine, National Institutes of Health (NIH)**

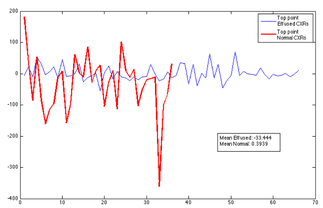

## Lung Shape Detection

Each of the TB manifestations depicts various textural and shape characteristics that may affect the shape of the lung (Pleural Effusion, Silhouette, Miliary Pattern) or may not (Mass, Lymphadenopathy, Opacity, Reticulation, Nodules).

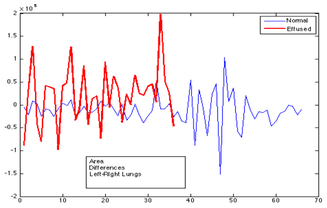
We have developed a module that addresses the detection of lung-deformed TB manifestations.




**Flowchart of process for detecting lung-deformed TB manifestations**




**Difference between the top points of the left lung and right lung. Normal lung fields tend to have their top points at the same height.**




**Area difference between left and right lung. Clearly the area disparity is more unstable for effusion cases compared to normal cases which tend to be more stable.**

## Texture Detection

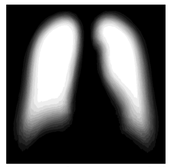
**Ground-truth segmentation for JSRT CXR**

**CXR image (Source: JSRT)**

**System Overview**

**Histogram-equalized CXR**

Lung Shape Models

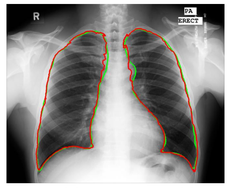


**Lung shape model for JSRT data.**

**Preprocessing:**
Intensity normalization, noise removal, contrast enhancement

Non-rigid Lung Registration Method

**Lung Segmentation:**
We use intensity information to align the registered lung shapes, for which we apply methods such as graph-cuts and level-sets.



**CXR with segmented lung**

**Feature Extraction:**
We compute the following histogram feature types: intensity, gradient magnitude, shape descriptors, curvature descriptors, oriented gradients (HOG), local binary patterns (LBP).

Trained Classifier

**Classification:**
A linear SVM classifier discriminates between normal and abnormal CXRs and outputs confidence values for its decisions.

**Normal**

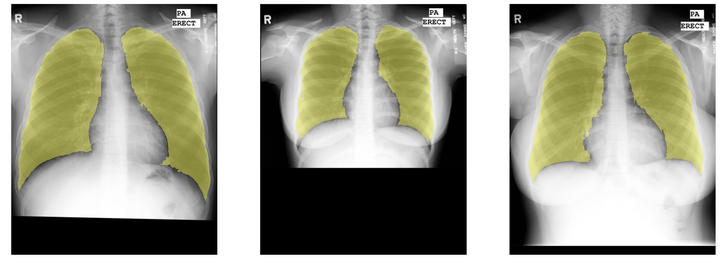
**Abnormal**

## Acknowledgment


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).


## Results

Qualitative segmentation results:

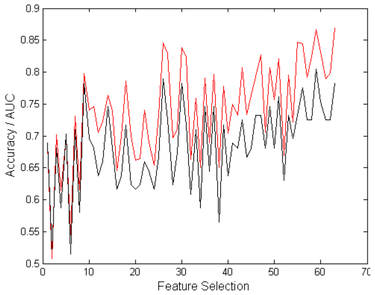


Quantitative segmentation results:

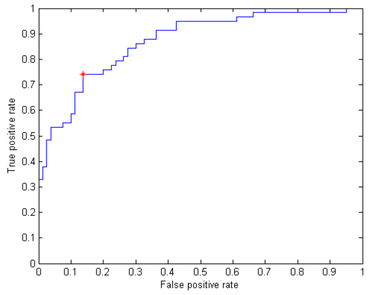
**Dice similarity scores between the ground-truth and calculated segmentation masks, computed on three different CXR sets from Japan, U.S., and India. The scores for the JSRT set are state-of-the-art (computed by S. Candemir).**

| Data Set | Avg ± std | Min | Median | Max |
|---|---|---|---|---|
| JSRT | 96.6 ± 1.4 | 88.5 | 97.0 | 98.3 |
| Montgomery County TB CXR | 96.0 ± 1.4 | 90.0 | 96.4 | 98.1 |
| India CXR set | 94.4 ± 2.1 | 85.4 | 94.9 | 97.1 |

Classification Performance:




**Feature subset evaluation of the accuracy (black curve) and the area under the ROC curve (AUC, red curve). The set with all features provides the best performance.**




**ROC curve for classification with all features. The area under the ROC curve is ~87%, with ~78% accuracy.**

## Conclusion

We have developed a TB screening software that is within reach of the performance of human experts when tested on TB CXRs from a local TB clinic.