China Set - The Shenzhen set - Chest X-ray Database

Description: The standard digital image database for Tuberculosis is created by the National Library of Medicine, Maryland, USA in collaboration with Shenzhen No.3 People's Hospital, Guangdong Medical College, Shenzhen, China. The Chest X-rays are from out-patient clinics, and were captured as part of the daily routine using Philips DR Digital Diagnose systems.

Number of X-rays:

- 336 cases with manifestation of tuberculosis, and
- 326 normal cases.

Image parameters:

- Format: PNG
- Image size varies for each X-ray. It is approximately 3K x 3K.

Image file names are coded as `CHNCXR_#####_0/1.png`, where '0' represents the normal and '1' represents the abnormal lung.

The clinical readings of all X-rays are saved as text files following the same file format: `CHNCXR_#####_0/1.txt`. Each text file contains the patient`s age, gender, and abnormality of the lung. (PTB: pulmonary tuberculosis).

IRB Details: The dataset was de-identified by the data providers and was exempted from IRB review at their institutions. The dataset use and public release were exempted from IRB review (No. 5357) by the NIH Office of Human Research Projections Programs.

Citation and Use:

- Please do not share the dataset outside of your research group/ organization, but forward new dataset requests to us.
- Please inform us if you find errors or inconsistencies in the data.
- It is requested that publications resulting from the use of this data attribute the source (National Library of Medicine, National Institutes of Health, Bethesda, MD, USA and Shenzhen No.3 People's Hospital, Guangdong Medical College, Shenzhen, China) and cite the following publications:
    1) Jaeger S, Karargyris A, Candemir S, Folio L, Siegelman J, Callaghan F, Xue Z, Palaniappan K, Singh RK, Antani S, Thoma G, Wang YX, Lu PX, McDonald CJ. Automatic tuberculosis screening using chest radiographs. IEEE Trans Med Imaging. 2014 Feb;33(2):233-45. doi: 10.1109/TMI.2013.2284099. PMID: 24108713
    2) Candemir S, Jaeger S, Palaniappan K, Musco JP, Singh RK, Xue Z, Karargyris A, Antani S, Thoma G, McDonald CJ. Lung segmentation in chest radiographs using anatomical atlases with nonrigid registration. IEEE Trans Med Imaging. 2014 Feb;33(2):577-90. doi: 10.1109/TMI.2013.2290491. PMID: 24239990

Acknowledgement: Staff involved in the collection of this data:

NLM: Data Collection/Curation/Validation

- Sameer Antani, PhD
- Stefan Jaeger, PhD
- Sema Candemir, PhD
- Zhiyun, Xue, PhD
- George R. Thoma, PhD

Data provider:

- Pu-Xuan Lu, Department of Radiology, The Shenzhen No.3 People`s Hospital, Guangdong Medical College, Shenzhen, China.
- Yi-Xiang Wang, Department of Imaging and Interventional Radiology, Prince of Wales Hospital, The Chinese University of Hong Kong, Shatin, Hong Kong.

For additional information, please contact: Dr. Sameer Antani, Phone: +1 301-435-3218, Email: sameer *dot* antani AT nih *dot* gov