For decades, tuberculosis (TB), a potentially serious infectious lung disease, continues to be a leading cause of worldwide death. Proven to be conveniently efficient and cost-effective, chest X-ray (CXR) has become the preliminary medical imaging tool for detecting TB. Arguably, the quality of TB diagnosis will improve vastly with automated CXRs for TB detection and the localization of suspected areas, which may manifest TB. The current line of research aims to develop an efficient computer-aided detection system that will support doctors (and radiologists) to become well-informed when making TB diagnosis from patients' CXRs. Here, an integrated process to improve TB diagnostics via convolutional neural networks (CNNs) and localization in CXRs via deep-learning models is proposed. Three key steps in the TB diagnostics process include (a) modifying CNN model structures, (b) model fine-tuning via artificial bee colony algorithm, and (c) the implementation of linear average–based ensemble method. Comparisons of the overall performance are made across all three steps among the experimented deep CNN models on two publicly available CXR datasets, namely, the Shenzhen Hospital CXR dataset and the National Institutes of Health CXR dataset. Validated performance includes detecting CXR abnormalities and differentiating among seven TB-related manifestations (consolidation, effusion, fibrosis, infiltration, mass, nodule, and pleural thickening). Importantly, class activation mapping is employed to inform a visual interpretation of the diagnostic result by localizing the detected lung abnormality manifestation on CXR. Compared to the state-of-the-art, the resulting approach showcases an outstanding performance both in the lung abnormality detection and the specific TB-related manifestation diagnosis vis-à-vis the localization in CXRs.

Introduction

Tuberculosis (TB), a highly contagious lung disease, is the leading cause of worldwide death followed by malaria and HIV/AIDS. The World Health Organization ([World Health Organization, 2018](https://www.frontiersin.org/articles/10.3389/frai.2020.583427/full#B45)) alludes that more than 95% of TB patients live in developing countries that lack adequate healthcare funding and supporting medical infrastructure. In descending order, two-thirds or 67% of newly TB-infected cases occur in eight developing nations beginning with India, followed by China, Indonesia, the Philippines, Pakistan, Nigeria, Bangladesh (formerly, East Bengal of British India), and South Africa. Statistics from 2000 to 2018 have projected a saving of 58 million lives via early TB diagnosis and timely treatment. Thus, timeliness in TB diagnosis is critical when mitigating its spread, improving TB preventive efforts and/or minimizing the TB death rate.

Currently, computed tomography (CT) offers the best-known TB detection method. For most earlier cases, however, TB diagnosis is confirmed via chest X-rays (CXRs) given the radiation dose, cost, availability, and the ability to reveal the unsuspected pathologic alterations among TB detection methods. For decades, researchers have focused on developing a computer-aided detection (CAD) system for the preliminary diagnosis of TB-related diseases via medical imaging. In the early stages, CAD depends on rule-based algorithms to select and extract useful pathogenic features within images to yield meaningful quantitative insight; yet, such methods are time-consuming, having to rely chiefly on the artificial extraction of patterns with useful information. As the manifestation of many diseases typically covers an extremely small region of the entire image, the challenge of the feature recognition process quickly becomes compounded. Moreover, with cumulative medical image data and evolving mutations of the disease, problems such as poor transferability among different datasets and unstable performance vis-à-vis newly generated data have stopped the CAD system from formulating a well-grounded decision with high accuracy.

With advances in deep learning, the convolutional neural networks (CNNs) have consistently surpassed other traditional recognition algorithms in achieving superordinate performance for image-based classification and recognition problems. The superlative ability to automatically extract useful features from the inherent characteristics of data makes CNN the first choice for complex medical problem solving. To date, CAD systems embedded with deep-learning algorithms have worked efficiently for medical disease detection by effectively generating a range of high-quality diagnostic solutions while spotlighting

## Materials and Methods

### **Datasets and Preprocessing**

Two public CXR datasets, the Shenzhen Hospital CXR dataset and the NIH CXR dataset, have been used in our study to test the performance of deep CNN models processed via our proposed methodology.

The Shenzhen Hospital CXR dataset ([Candemir et al., 2013](https://www.frontiersin.org/articles/10.3389/frai.2020.583427/full" \l "B5); [Jaeger et al., 2014](https://www.frontiersin.org/articles/10.3389/frai.2020.583427/full#B15); [Wang et al., 2017](https://www.frontiersin.org/articles/10.3389/frai.2020.583427/full#B43)) is compiled by the Shenzhen No. 3 People's Hospital and Guangdong Medical College in China. The dataset comprises 662 frontal posteroanterior CXR images in various sizes, among which 326 have been diagnosed as normal cases, whereas the other 336 as having TB manifestations. The NIH CXR dataset ([Shin et al., 2016b](https://www.frontiersin.org/articles/10.3389/frai.2020.583427/full#B40)) is by far one of the largest public CXR datasets. This dataset is extracted from the clinical PACS database at the NIH Clinical Center, comprising 112,120 frontal view (posteroanterior and anteroposterior) CXR images with 14 thoracic pathologies (atelectasis, consolidation, infiltration, pneumothorax, edema, emphysema, fibrosis, effusion, pneumonia, pleural thickening, cardiomegaly, nodule, mass, and hernia). As it is not anticipated for the original radiology report to be shared publicly, the disease information and labels for CXRs had to be text-mined via natural language processing techniques with accuracy of >90%. Owing to the massive amount of data, detailed annotations, and wide range of thorax diseases covered by this dataset, many researchers studying thorax disease detection vis-à-vis the deep-learning area have used it.

To improve the general quality of the NIH CXR dataset, all CXR images have been enhanced using contrast limited adaptive histogram equalization (CLAHE) ([Pizer et al., 1987](https://www.frontiersin.org/articles/10.3389/frai.2020.583427/full" \l "B35)) with the clip limit number equals 1.25; similarly, to improve the processing speed, all CXR images have also been resized from their original size to 512 × 512. Also, in diagnosing TB-related manifestations on the NIH CXR dataset, as the distribution of CXRs under each TB-related disease class presents a strongly biased trend, models trained on this dataset will tend to perform with a strong preference for their conforming predictions. As such, data augmentation techniques such as horizontal flip, rotate, contrast adjustment, and position translation have been perspicaciously implemented to increase the number of images under the classes with fewer CXRs, thereby propagating an evenly distributed data to eliminate the interference.

### **Methodology**

In this section, a CAD system driven by deep CNN models for TB diagnostics and localization from CXR images with the use of unified approaches to improve the accuracy-stability of the disease detection process is proposed. As shown in [Figure 1](https://www.frontiersin.org/articles/10.3389/frai.2020.583427/full#F1), we divide the TB diagnostic task into four subprocesses: (i) CXR image preprocessing; (ii) preliminary detection of the suspected TB patients via abnormality checking; (iii) identification of the specific TB manifestation [consolidation ([Adler and Richards, 1953](https://www.frontiersin.org/articles/10.3389/frai.2020.583427/full#B1)), effusion ([Vorster et al., 2015](https://www.frontiersin.org/articles/10.3389/frai.2020.583427/full#B42)), fibrosis ([Chung et al., 2004](https://www.frontiersin.org/articles/10.3389/frai.2020.583427/full#B7)), infiltration ([Mishin et al., 2006](https://www.frontiersin.org/articles/10.3389/frai.2020.583427/full#B28)), mass ([Cherian et al., 1998](https://www.frontiersin.org/articles/10.3389/frai.2020.583427/full#B6)), nodule ([Kant et al., 2007](https://www.frontiersin.org/articles/10.3389/frai.2020.583427/full#B16)), and pleural thickening ([Gil et al., 1994](https://www.frontiersin.org/articles/10.3389/frai.2020.583427/full#B10))]; and (iv) localization of the suspicious diseased area on CXRs.

For each process, all deep CNN models used herein have been improved by structural modification, the implementation of artificial bee colony (ABC) algorithm during fine-tuning, and the amalgamation of an ensemble model. Unlike standard object detection task encapsulating bounding box information, the localization of TB manifestations is achieved primarily via class activation mapping; essentially, this entails the production of an attention map over the image to spotlight all the detected suspicious areas instead of predicting edges of bounding boxes.

Specifically, we analyze the overall performance of the experimented deep CNN models on two publicly available CXR datasets, namely, the Shenzhen Hospital (CXR) Dataset and the NIH CXR dataset. For all datasets, our analysis emphasizes the binary classification of CXR images to differentiate among the TB abnormalities, while performing a further diagnosis and localization of specific TB-related manifestation on the NIH CXR dataset.

Compared to typical performance results from just a plain application of the original deep CNN models, our quantitative results of the proposed methodology achieve a staggering improvement of >30% points during the identification of TB manifestation, aside from outstanding prediction accuracy with an average improvement of over >8% points on the abnormality detection. For qualitative results, not only can our model provide the attention map that fully encapsulates the suspicious diseased regions vis-à-vis the diagnostic results, but the model also can successfully distinguish among diseases caused by similar reason(s). Additionally, for spatially spread out TB-related manifestations, regardless of their sizes, the model can often localize the abnormalities successfully.