

Improving Tuberculosis Diagnostics using Deep Learning and Mobile Health Technologies among Resource-poor Communities in Perú

Marlon F. Alcantara, Yu Cao, Chang Liu, Benyuan Liu, Maria Brunette, Ning Zhang, Tong Sun, Peifeng Zhang, Qilei Chen, Ying Li, Cesar Morocho Albarracin, Jesus Peinado, Epifanio Sanchez Garavito, Leonid Lecca Garcia, Walter H. Curioso



PII: S2352-6483(17)30018-1
DOI: <http://dx.doi.org/10.1016/j.smhl.2017.04.003>
Reference: SMHL5

To appear in: *Smart Health*

Received date: 23 March 2017
Accepted date: 2 April 2017

Cite this article as: Marlon F. Alcantara, Yu Cao, Chang Liu, Benyuan Liu, Maria Brunette, Ning Zhang, Tong Sun, Peifeng Zhang, Qilei Chen, Ying Li, Cesar Morocho Albarracin, Jesus Peinado, Epifanio Sanchez Garavito, Leonid Lecca Garcia and Walter H. Curioso, Improving Tuberculosis Diagnostics using Deep Learning and Mobile Health Technologies among Resource-poor Communities in Perú, *Smart Health*, <http://dx.doi.org/10.1016/j.smhl.2017.04.003>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting galley proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Improving Tuberculosis Diagnostics using Deep Learning and Mobile Health Technologies among Resource-poor Communities in Perú

Marlon F. Alcantara^a, Yu Cao^a, Chang Liu^a, Benyuan Liu^a, Maria Brunette^b, Ning Zhang^a, Tong Sun^a, Peifeng Zhang^a, Qilei Chen^a, Ying Li^a, Cesar Morocho Albarracin^b, Jesus Peinado^c, Epifanio Sanchez Garavito^d, Leonid Lecca Garcia^c, Walter H. Curioso^e

^aDepartment of Computer Science. University of Massachusetts, Lowell, MA, USA

^bDepartment of Public Health. University of Massachusetts, Lowell, MA, USA

^cPartners in Health Perú. Carabayllo, Lima, Perú

^dNational Hospital Sergio E. Bernales Perú

^eBiomedical and Health Informatics. University of Washington, Seattle, Washington, USA

* Corresponding author. Tel.: +1-978-934-3628. E-mail address: ycao@cs.uml.edu

Abstract

Tuberculosis (TB) an infectious disease and remains a major cause of death globally. The World Health Organization (WHO) estimates that there were 10.4 million new TB cases worldwide in 2015. The majority of the infected populations come from resource-poor and marginalized communities with poor healthcare infrastructure. It is critical to reduce TB diagnosis delay in mitigating disease transmission and minimizing the reproductive rate of the tuberculosis epidemic. To combine machine learning and mobile computing techniques may help to accelerate the TB diagnosis among these communities. The goal of our research is to reduce TB patient wait times for being diagnosed by developing new machine learning techniques and mobile health technologies. In this paper, major technique barriers and proposed system architecture are first introduced. Then two major progresses are reported: (1) To develop an X-ray image database and annotation software dedicated for automated TB screening. The annotation software can help to highlight the TB manifestations, which are very useful for machine learning algorithms; (2) To develop effective and efficient computational models to classify the image into different category of TB manifestations. The model we proposed is a deep convolutional neural networks (CNN)-based models. We have conducted substantial experiments and the results have demonstrated that our approach is promising. We envision our future work includes two research activities. First, we plan to improve the performance of the algorithms with deeper neural networks. Second, we plan to implement our algorithms on mobile device and deploy our system in the city of Carabayllo, a high-burden TB area in Lima, the capital of Perú.

Keywords: tuberculosis; diagnosis; deep learning; deep convolutional neural networks, mHealth; mobile computing.

1. Introduction

Tuberculosis (TB) an infectious disease and remains a major cause of death globally. It affects the most disadvantaged populations and involves complex treatment regimes. There were more than 10.4 million

estimated new cases and 1.8 million deaths in 2015, worldwide [1]. Of the estimated 10.4 million people who developed TB in 2015, 60% occurred in six countries (India, Indonesia, China, Nigeria, Pakistan and South Africa), 95% of the deaths occurred in countries with low and middle income. [2]. Peru has the highest incidence per capita of tuberculosis in the Americas, including virulent multi-drug resistant and extensively drug resistant strains. According to the World Health Organisation (WHO), its reduction is slow at around 1.5% a year, just short of the global reduction of 1.65% annually. The rate of reduction would need to increase to a 4-5% annually, in order to reach sustainable development goal target 3.3 – ending epidemics of tuberculosis and other infectious diseases by 2030. In Low and Middle Income Countries (LMICs), efforts to eliminate the Tuberculosis (TB) epidemic are challenged by the persistent social inequalities in health, the limited number of local healthcare professionals, and the weak healthcare infrastructure found in resource-poor settings. The global health community has confronted the situation by focusing on developing and testing effective vaccines, improving the diagnosis process, and promoting patient adherence to the medical treatment.

It is critical to reduce the TB diagnosis delay in mitigating disease transmission and minimizing the reproductive rate of the TB epidemic. The ultimate goal of our research is to reduce patient wait times for being diagnosed by developing a socio-technical solution to the TB diagnostics problem. Specifically, we aim to design a user-centered, mobile device-based computing system to expedite the TB diagnosis process by developing and implementing novel image processing and machine learning techniques to analyse patients' chest X-ray images. Our study will be conducted in the city of Carabayllo, a densely occupied urban community and high-burden TB area in Lima, the capital of Perú.

Mobile computing techniques offer a unique opportunity to accelerate the TB diagnosis among resource-poor, marginalized communities with weak healthcare infrastructure and systems. However, real-world mobile computing tools and applications in TB-related clinical practice with the capacity of accurate TB screening using mobile devices are rare. A wide gap between the technological advancements and the real-world clinical practices is caused by two major barriers: (1) *the first barrier* is the lack of large-scale, real-world, well-annotated, and public available X-ray image database dedicated for automated TB screening. The majority of existing X-ray image databases, such as ImageCLEF [3], JSRT Digital Image Database [4], and ANODE Grand Challenge Database [5], were created mainly for one or two specific TB manifestations (e.g., pulmonary nodule). To the best of our knowledge, there is no large-scale, real-world, and public available chest X-ray dedicated for TB diagnosis with high-quality annotation; (2) *the second barrier* is the lack of mobile devices-based computing system that can offer accurate diagnosis by analysing chest X-ray images. The application of computer-aided chest radiography for TB screening and diagnosis [6-14] has been limited due to the modest sensitivity and specificity, and high inter- and intra-observer differences in reporting of radiographs [8]. Hence, the automatic screening for TB in chest radiographs is still a challenging task and an open research problem [8]. Furthermore, there is very few reported research on using mobile device to capture and analyse the chest radiograph images for computer-aided TB diagnosis [8].

By working closely with a local community stakeholder, Partners in Health Perú (Socios en Salud) and with a truly interdisciplinary team of public health, medicine, systems engineering, health informatics, and computers science, we aim to develop and successfully implement a mobile device-based computing solution to overcome the aforementioned barriers. Our research endeavour, anchored in the Community-Based Participatory Research (CBPR) platform [15-17], would ultimately broaden our understanding about how a socio-technical solution can serve as a platform for healthcare process optimization that is responsive and reflective of local healthcare systems and social constraints and resources.

As the first step of developing the proposed sociotechnical system, we will introduce the two major progresses we have recently achieved. The first relates to the development of large-scale, real-world and well-

annotated X-ray image database dedicated for automated TB screening with the collaboration of an annotator team and a pulmonologist. The second one focuses on developing effective and efficient computational models to classify the image into different categories of TB manifestations.

2. Background and Related Work

This section include three parts: mobile computing in healthcare (sub-section A), developing chest X-ray image database (sub-section B), and computer-aided system to screen the chest radiography image for TB diagnosis (sub-section C).

2.1. Mobile Computing in Healthcare (mHealth)

Point of care delivery is critical for the success of any application in the clinical healthcare environment. In Perú, as in many developing countries, a mobile device-based computing solution is very suitable within the context of resource-poor communities in Lima, Perú. The unique characteristics of the mobile devices such as its pervasiveness and low cost provide them the opportunity to support and enable smart care decision making in a connected health scenario for automatic health scenario and tuberculosis screening.

- *mHealth in Perú:* In a recent review of the mHealth literature published in Perú, Ruiz et al. [18] showed that mobile health interventions have enormous potential to improve access and the quality of health services in Perú, increasing the effectiveness of public health programs and reducing healthcare costs [18]. Out of 19 papers selected, most of them showed a positive impact, and four were about tuberculosis. It is important to notice that most of them were implemented as pilot projects ([18]). However, the majority of the papers demonstrated that mobile health interventions are well accepted by the population and well-developed projects might contribute to reduce the gap in public health, reducing limitations such as lack of resources (human and logistic) in health care centers, high dispersion of the population and lack of infrastructure (roads, transportation and Internet connectivity).
- *mHealth for TB Diagnostics:* During the last few years, mobile phones have been successfully used for diagnosis of tuberculosis [19]. In Perú, Zimic et al. [20] proposed a relatively minimal investment with mobile phones to facilitate the diagnosis of tuberculosis using a low cost Microscopic Observation Drug Susceptibility (MODS) in remote settings where a lack of trained personnel may otherwise be a limitation [20]. Nowadays, with the advances in mobile processors, images taken by a cell-phone can be immediately processed and analyzed with the help of smart algorithms. Today's global wireless infrastructure also allows transmission of a wide variety of tuberculosis images (such as X-rays) to remote locations for telemedicine diagnosis. Therefore ubiquitous cell-phone based applications can provide unique opportunities to combat tuberculosis, especially in developing countries [21]. Recently, Schwartz et al. [22] assessed the diagnostic accuracy of digital photographs of plain film chest X-rays obtained using a mobile phone in Botswana. The authors concluded that digital photographs of chest X-rays obtained via a mobile phone equipped with a digital camera are comparable to plain film chest X-rays [22].
- *The need of a timely tuberculosis diagnosis in Perú:* tuberculosis remains as a serious public health problem. A successful treatment plan requires a proper diagnosis, in addition to good knowledge about drug susceptibility [23]. Reducing the tuberculosis diagnosis delay is critical in mitigating disease transmission and minimizing the reproductive rate of the tuberculosis epidemic. Different factors impact delays in tuberculosis diagnosis [24], such as: patient health seeking behavior, healthcare

centers with poor infrastructure and equipment, inadequate resources and information systems (mostly paper-based), lack of (or inexistent) documented processes, and lack of human resources as part of a multidisciplinary tuberculosis team.

2.2. Developing Chest X-ray Image Database

While there are some evaluation efforts in TB screening tests on developing countries [25, 26], to the best of our knowledge, there is no large-scale, real-world, well-annotated, and public available X-ray image database dedicated for TB screening diagnosis. Most of the existing research [6-14] in the area of computer-aided TB screening employed small data sets for evaluation and validation. Most of the datasets have less than 200 images. There are a few large data sets, such as ImageCLEF [3], JSRT Digital Image Database [27], and ANODE Grand Challenge Database [5], have over tens of thousands images. However, they only include one or two aspects of TB manifestations (e.g., pulmonary nodule). Without a large-scale data sets with high quality annotation, it will be very difficult to determine the efficacy of existing and proposed approach when applied to real-world clinic data. Furthermore, the new database demands an accurate annotation to make possible for the machine learning algorithms be able to differentiate among a large number of TB manifestations, the only way to provide it, is with the assistance of specialized Pulmonologists to lead and review the annotations and a software where the annotations can be done.

2.3. Computer-aided System to Screen the Chest Radiography Image for TB Diagnosis

The research activities in the area of computer-aided image analysis for tuberculosis (TB) screening from X-ray image can be broadly divided into two categories: (1) the first category is the computer-aided screening and scoring algorithms using chest radiographic features for the TB diagnosis [6-13]. Research activities in this category focus on developing different types of visual features and classification algorithms to score and screen different types of TB manifestations. Most of the papers employ texture features (e.g., Local binary patterns (LBP) [28, 29], Daubechies wavelets [30]) or geometry features (e.g., circularity, Hessian shape features). The classification algorithms employed in these papers range from simple threshold-based approach or k-nearest neighbors (K-NN) algorithm to more complicated methods, such as Decision tree and Support Vector Machine (SVM); (2) the second category of related work is focusing on X-ray image categorization on the organ and pathology level [14]. The main stream methodology in this area is based on local patch representation of the image content (e.g., visual bag of words (Visual BoW) approach). This type of dense sampling of simple features are then feed to non-linear kernel-based classifier, such as SVM classifier. The goal is to discriminate between healthy and pathological cases. It is also shown that this type of methods can successfully identify specific pathologies in a set of chest radiographs.

3. Proposed Approach

Figure 1 illustrates our proposed approach for improving TB diagnosis. We aim to to design and deploy a reliable, safe and secure, simple to use, and power efficient mobile phone-based cloud computing system to screen the chest radiography image with improved accuracy and reader consistency.

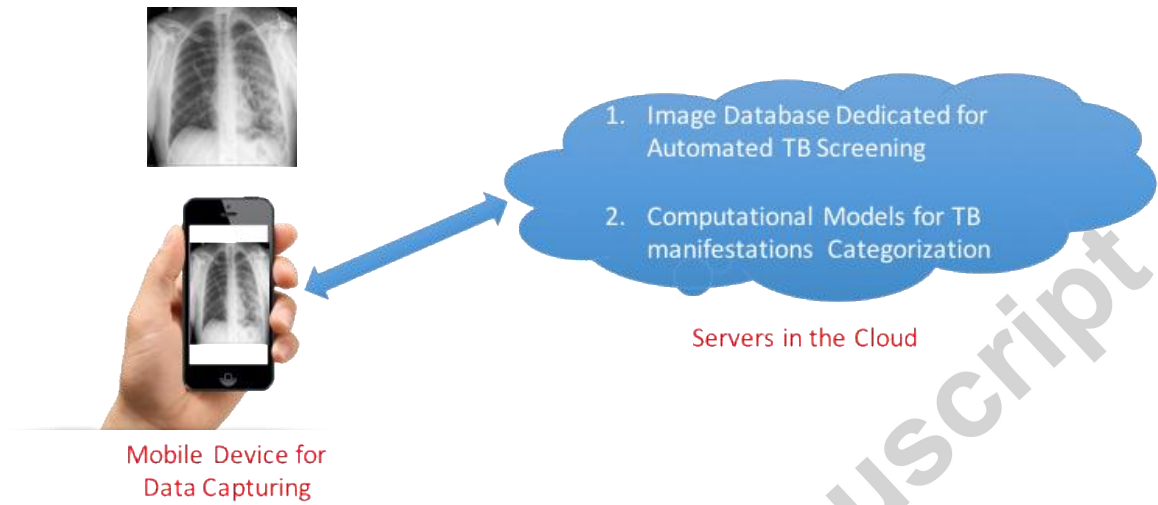


Fig. 1. Overview of the proposed mobile phone-based system for improving TB diagnosis.

We will introduce two major processes we have made recently. As shown in the right side of Figure 1, the first progress is to develop a X-ray image database and annotation software dedicated for automated TB screening. The second progress focus on developing effective and efficient computational models to classify the image into different categories of TB manifestations. We will introduce these two progresses in the following two sub-sections.

3.1. To investigate, develop, and evaluate a chest X-ray image database and related software tools

- 1) *Technique Challenges:* The main challenge in this component includes: (1) where and how to gain access to the real-world, large scale TB screening images with detailed diagnostic descriptions; (2) to determine the types of TB manifestations we should target and how we can use these manifestations to annotate each X-ray images; and (3) to develop dedicated annotation software and database management software package for reviewing the chest radiography, locating important contents, annotate them, and extract the annotated contents for research, teaching, and training purposes.
- 2) *Proposed Approach:* To address the first challenge, we established an international research team which include scientists from both U.S. and Peru. One of the core team members is Dr. Jesus Peinado, header of Informatics at Partners In Health at Peru. In the past three years, his team in Peru has collected around 5,000 chest X-ray radiography images captured from real-world TB patients with detailed TB screening descriptions. In addition to that, we also explored the second source of image is the X-ray images from the 2004-2013 ImageCLEF collection, which include over 400,000 medical images, diagnostic annotations, search topics and relevance judgments. The 2004-2007 collection [31-33] contains over 66,000 images from a variety of teaching files annotated in English, French or German. The 2008-2010 collection [34-36] contains over 77,000 images and captions from the medical literature. These images were published in Radiology and Radiographics, two of the journals published by the Radiological Society of North America. The 2011-2013 collection [37-40] includes

more than 300,000 image and related text annotation from the biomedical literature (e.g., PubMed). To address the TB manifestation issue, we worked very closely with our clinical and research collaborators, Dr. Jesus Peinado from Peru and Dr. John Bernardo at Boston Medical Center (BMC) and Boston University School of Medicine (BUSM), to generate a scientific categorization of TB manifestations. As shown in Fig. 2, we have identified five TB manifestations. There are some important discoveries (which will serve as important motivations and rational for the proposed approach for image analysis and machine learning techniques) from these images. First, the variety of the TB manifestations is large. Second, each category of TB manifestation actually indicates the severity of the TB disease. Therefore, algorithms that can recognize and classify the X-ray image into different type of TB manifestation can serve the purpose of screen the X-ray image to understanding the severity of the TB disease.

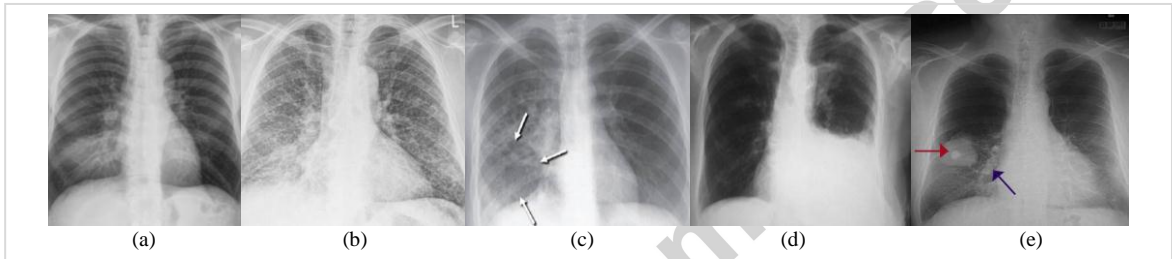


Fig. 2: (a) Air space consolidation which showing glass opacity with consolidation in the right middle lobe; (b) Miliary pattern with seed-like appearance; (c) Cavity located at the lower lobe (annotated by arrows); (d) Pleural effusion, which is excess fluid that accumulates in the pleural cavity; (e) Calcified granulomata: The red arrow indicates a large 5 cm diameter squamous cell carcinoma of the right lower lobe and there is 1.5 cm bright opacity in the middle of the mass (which is a calcified granuloma). Additional calcified granulomatous areas are medial to the mass, as indicated by blue arrow.

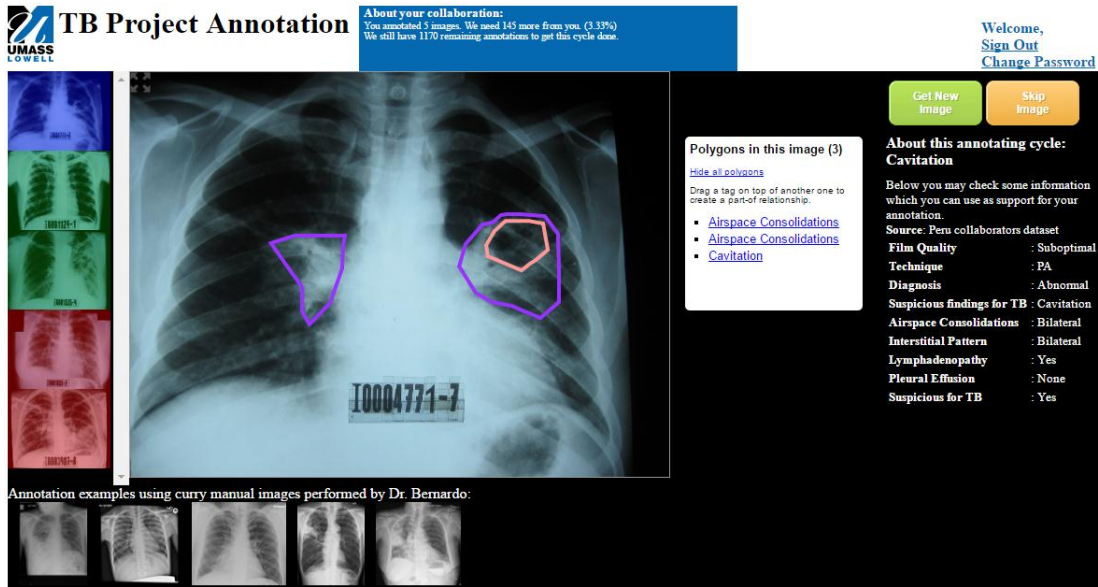


Fig. 3: Annotation software interface.

3.2. To provide an accurate and reliable annotation on the medial images.

In addition to developing the real-world database, we also developed annotation software for reviewing the chest radiography, locating important contents, annotate them, and extract the annotated contents for research, teaching, and training purposes. While there are many existing efforts in medical image annotations, [41, 42], there is few open source annotation software dedicated to annotating X-ray image to support automatic screening. Based several existing open source annotation software projects [43], we developed a web-based annotation software.

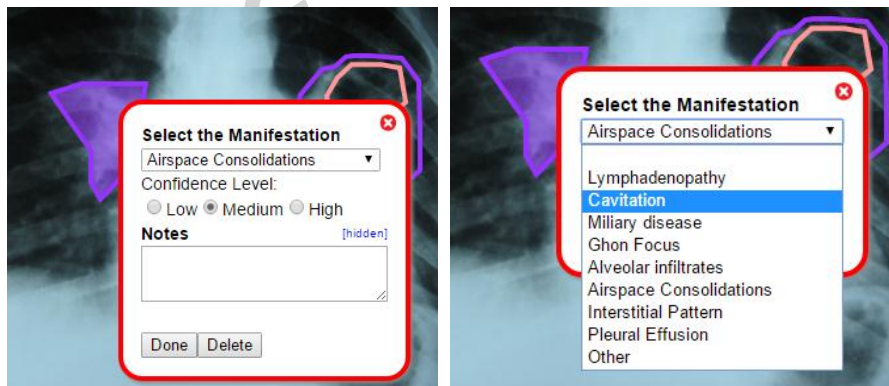


Fig. 4: Pop up to inform the manifestation and details.

In general lines, the annotation software is a tool in which someone capable can highlight the location at the image in which some TB manifestations is occurring. However, this task is not possible without the guidance of a specialist. We trained a team with around 10 members who were trained directly by the Pulmonologist Dr. John J. Bernardo. The train were provided for each TB manifestation isolated, and some extra training to verify complex cases. The annotation is performed using our own software developed exclusively for this purpose. The common interface of the annotation software can be seen on Figure 3. In the left panel are all the images in which the annotator is working on.

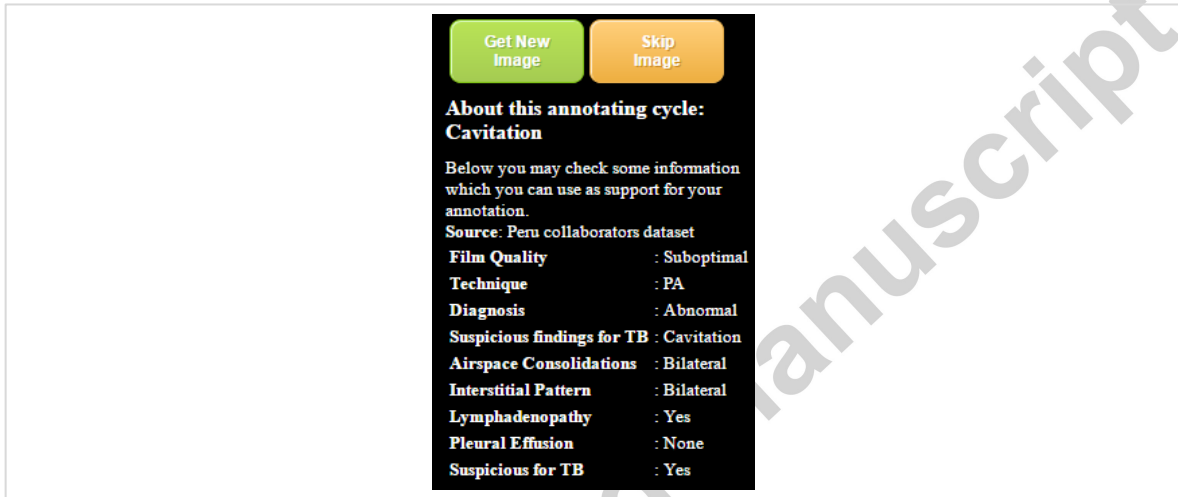


Fig. 5: Right Panel, buttons and golden information.

The images have a color to identify the current annotation status: blue for the image shown in the central panel, green for the images already annotated and red for images with no annotation. On the middle, we have the main panel which contains the current image with every annotation on it, the system pick one specific color for each different manifestation, on the right side of the annotation panel the annotator can see all the annotated polygons listed, also, the annotator can hide the polygons to see a clean image and can highlight individually each polygon as well. When the annotator starts an annotation, he/she clicks in a point to be the initial vertex of a polygon, after, he/she keeps indicating each subsequent vertex of the polygon, and finally close the polygon with a click in the initial vertex, in the moment in that the polygon is finished, a popup window shows up asking the details about the regions highlighted, it includes the TB manifestation, the confidence level, and possibly some notes (Fig. 4 - left). The manifestation may be informed using some default options in a select box (Fig. 4 - right), or the annotators may choose to pick the option other and write the manifestation themselves.

The right panel (Fig. 5) shows important information provided by the team in Peru coordinated by Dr. Peinado that helps the annotator to provide the correct manifestation, the items cover the main TB manifestations possible to be found in x-ray images (seen in Fig. 5) may be verified with some details in this panel, including sometimes the side (right, left or bilateral) in which a specific manifestation is.

Also, the right panel provide two buttons, the first one is a “Get New Image” that adds to the annotator’s personal collaboration one more image to be annotated. This button, randomly select an image lacking annotation and gives it to the annotator, so, there is no pre-selected image to a specific annotator. The

annotator also may skip the image using the button “Skip Image”, it makes the current image returns to the set of non-annotated images, any annotation made is discarded and the image are available to any annotator again.

The information shown in this panel is not the only support the annotators have, on the bottom there are some images annotated by a pulmonologist using the same software to work as a sample to the annotators, when any image is clicked, a new window opens showing the sample as in Fig. 6.

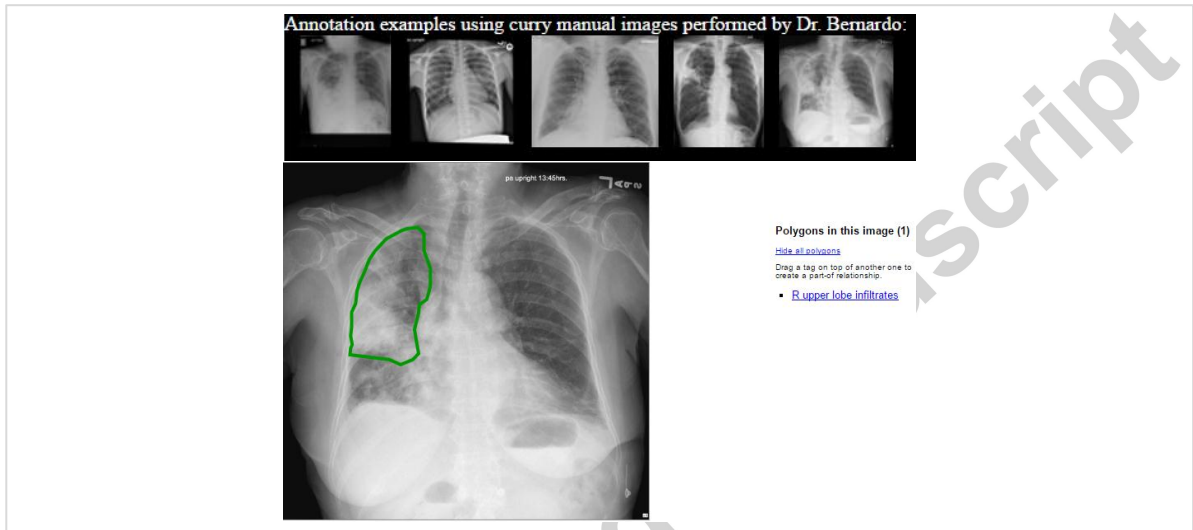


Fig. 6: Images annotated by a pulmonologist (top) to be choose, one specific image open as a sample (bottom).

As mentioned, the annotation software is designed for capable annotators, in our context, capable annotators are people who received a training to identify and localize TB manifestations. To make annotations in an x-ray is a tough task even to experienced pulmonologists. Because of this, the annotators received one training for each TB manifestation, during the training also is collected some images to be used as a sample as shown in Figure 6, after the training all the annotators work in the same manifestation until all images from that specific manifestation is covered. However, images may contain more than one manifestation at the same time, in these cases, only the current manifestation is being annotated, and the remaining wait for the next training.

We call the training and the following annotation as cycle as we can see in Figure 7, the cycle repeats for each new manifestation. In each cycle the annotation software also provides a feedback to the annotator about your current collaboration, including how many images left in your work or in the entire cycle. The cycles starts with the training following to the annotation and, after finished the annotation for each collaborator, a new meeting with the pulmonologist is made to work on cases annotated with low confidence, for this review the pulmonologist has a special page that loads only the images with low confidence and without any button, the pulmonologist can edit the annotation or just change the confidence level if it is well annotated, the pulmonologist has access to the collaborator responsible by that one to further clarification.

After this closing meeting, all the annotations are stored and no more are going to appear to annotators. From the second cycle onwards, the same images may appear again asking for the annotation of different manifestations, in these cases, all the previous annotations are shown to help the annotator to differentiate the

current manifestation from the previous ones, the annotator also can correct the previous annotation for a more accurate one.

After all the cycles, except by the images containing health lungs, all the other images contain the annotations for each manifestation so, the images are ready for be used in a region based classifier.

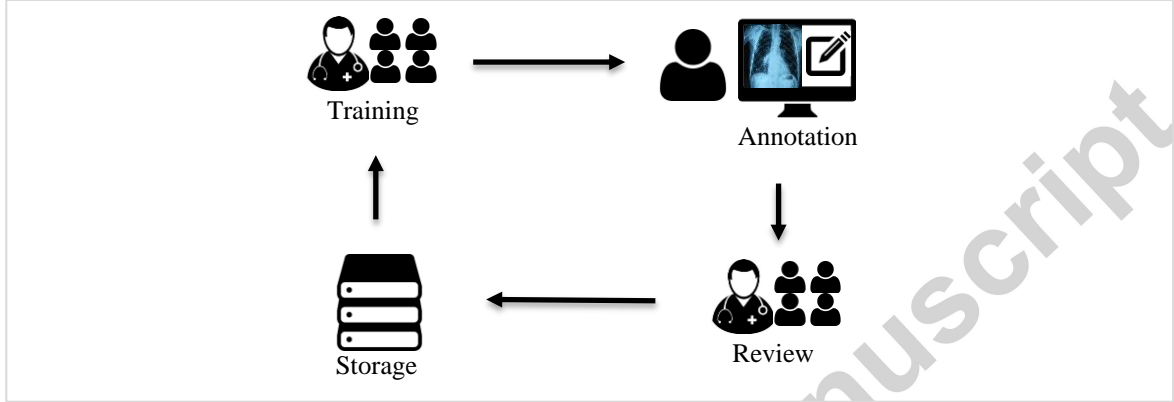


Fig. 7: Cycle for annotation process that repeats for each TB manifestation.

3.3. To Research and Develop Effective and Efficient Computational Model for TB Manifestation Analysis.

- 1) **Technique Challenges:** The main challenges of automatic TB screening come from the extremely complexity and large variety of the TB manifestations. This is true in clinic practice. As we have shown in Figure 2, the variations of TB manifestations can range from subtle military patters to apparent effusions. Via close collaborations and discussions with domain experts, we have discovered that unusual or abnormal TB manifestations affect the texture and geometry of the anatomy. Therefore, most of the existing techniques employ texture and/or geometry features. Usually, different features are useful for different manifestation. For example, texture features, such as Mean, Variance, Entropy, and Third moment, can be employed for detecting infiltration and dense. Local binary patterns (LBP), another type of texture spectrum feature, can be used for cavity detection. Template matching on Fourier domain, a method for geometry feature extraction, maybe useful for detecting the miliary pattern. Hessian shape features, another type of geometry features, could help to detect the nodules. Recently, researchers [6, 9] have shown that combine multiple features can improve the performance of abnormal TB image detection. For example, in paper [9], LBP and histogram of oriented gradients (HOG) are combined for cavity detection. In paper [6], a mixture of Intensity, LBP, and Hessian shape features are employed to measure normal and abnormal patterns in the X-ray image. The literatures have shown that the choice of features play the key role for system performance. Hence, the key issue is how to improve these hand-tuned features. To address these challenges, we plan to explore new solutions based on recent advances in deep learning [44, 45].

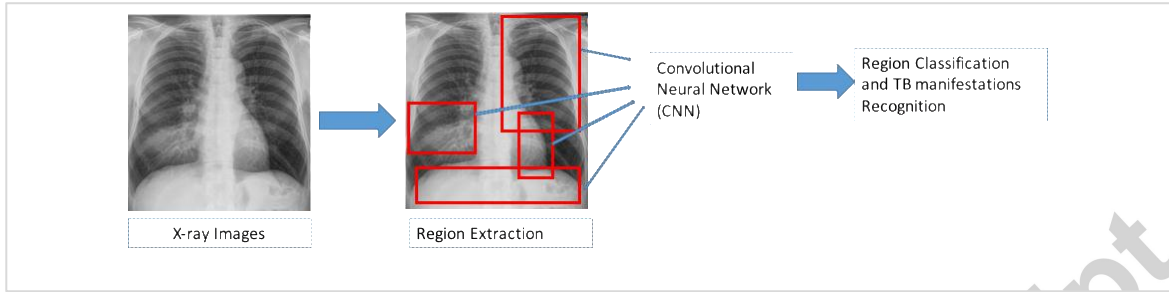


Fig. 8: Proposed approach for X-ray image analytics.

Deep learning [44, 45] aims to learn multiple levels of representation and abstraction that help infer knowledge from data such as images, videos, audio, and text, is making astonishing gains in computer vision, speech recognition, multimedia analysis, and drug designing. The impact of deep learning is far reaching on applications in medical, social and commercial domains [46-48]. Briefly speaking, there are two main classes of deep learning techniques: purely supervised learning algorithms (e.g., Deep Convolutional Network [49, 50]), unsupervised and semi-supervised learning algorithms (e.g., Denoising Autoencoders [51, 52], Restricted Boltzmann Machines [53, 54], Deep Boltzmann Machines [55]). Our proposed approach is rooted from Deep Convolutional Neural Network (CNN) [56] and Region-based CNN [57].

2) *Proposed Approach*: Our main objective is to analyse the X-ray images and to screen the chest radiography image with improved accuracy and reader consistency. We convert the screening problem into a classification problem. More specifically, we will investigate effective and efficient computational models to segment the image into smaller region and classify the image region into different category of TB manifestations (e.g., Air space consolidation, Miliary pattern, Cavity, Bronchiectasis, Opaque, etc.). As shown in Figure 8, our proposed approach includes the following three steps.

1. *Extraction of region proposals*: In this step, we will extract regions from the image using different methods like selective search [58].
2. *Initial Image feature extraction*: In this step, we will extract both global and local features from the X-ray images captured by the mobile device. While there exists a large number of global features ranging from color, texture, to edge features, we mainly choose texture and shape features because the cutting-edge research in computer-aided TB screening using X-ray imaging have shown that texture and shape features are most effective [6, 9, 14, 28-30, 59-63]. Some sample global features we used include: Gabor features and Local Binary Patterns (LBP) features. Some sample local features include SIFT features [64] and PHOG [65].
3. *Step 3: Deep Convolutional Neural Network (CNN)-based X-ray Image Analysis*: Recall in our first step, we have extracted some features from the original image. Hence, we do not need to transmit the entire image. Instead, we resize the image at mobile phone and only transmit the image with much smaller size. In our preliminary test, we reduce the size by half and the results are still acceptable. By doing so, our system will consume substantially less power, compared with transmitting the original image. Another contribution in this component is that we plan to employ the deep convolutional neural networks (CNN) [57, 66] for region classification. Our proposed techniques are rooted from recent advances on deep learning, such as region deep convolutional

neural networks [57], which take full use of region features. As shown in Figure 4, there are several steps in this approach:

4. *Extraction of region proposals*: In this step, we will extract regions from the image using different methods like selective search [58]. For each region, train a CNN model to calculate the new features for further classification. Please note, before we perform the feature extraction process, the region should be scaled to a fixed size 227x227 (in order to the same vector dimension in our further handling). After the above handling, we should generate a 4096-dimensional feature vector. The features from CNN will be combined with features originally transmitted from mobile phone. We will apply the linear classifier like Support Vector Machine (SVM) [67], the combined features for final region classification and TB manifestations recognition. For the implementation purpose, we use the open source Caffe [68] for training. The training of the proposed CNN approach could include two steps: (a) supervised training on a large dataset using CNN; (b) fine-tuning the CNN feature for detection using a smaller dataset. Our preliminary study has shown the feasibility of the proposed approach for a small group of images. In this study, we plan to extend and refine our preliminary results to large scale, real-word X-ray image datasets.

4. Experimental Results

In this section, we present our experimental results, which includes the data set used in our experiments and the reported accuracy. The dataset we used is provided by Dr. Peinado (one of the core team members from Perú). This dataset includes 4701 images. There are 453 normal images (from patients without TB) and 4248 abnormal images (from patients with different types of TB manifestations).

The first experiment we conducted is the binary categorization. This means our goal is to classify any X-ray TB image into two categories: normal and abnormal. We adapt the GoogleNet model [56] from Caffe [68] as the pre-training model, which was generated by a supervised training procedure from a very large dataset (ImageNet [69]). Then we use the 4701 images from our Perú collaborators for finetuning. We used 4/5 of the TB images for training and 1/5 of the images for testing, which is corresponding to 3760 images for training and 941 images for testing, respectively. The results (average precisions after certain number of iterations) are listed in Table 1 below. From this table, we can tell that we can achieve 89.6% of accuracy for binary classification (normal/abnormal). Please note, we achieve this performance without using any pre-processing techniques. The input to the system is the raw image pixels. Therefore, this number (89.6%) is very impressive considering the fact that we only using raw pixels.

Table 1: Average precisions for binary classification

# of Iterations	10,000	30,000	50,000	80,000	100,000
Average Precisions	82.8%	88.6%	89.0%	89.5%	89.6%

The second type of experiments we conduct is multiclass categorization. We use the same GoogleNet model as the previous experiment, and also 4/5 of images for training and 1/5 of images for testing. But in this

experiment, we add more categories named specific TB manifestation. In these experiments, our goal is to classify the TB X-ray images into multiple categories. Each category is corresponding to one of the TB manifestations. Table 2 below illustrates the characteristics of the data

Table 2: Characteristics of the Data Distribution in the Second Experiments

Category (Name of TB Manifestations)	Total Image #	Image # Used for Training	Image # Used for Testing
Cavitation	1182	946	246
Lymphadenopathy	202	162	40
Infiltration	2252	1802	450
Pleural Effusion	560	448	112

Table 3 shows the results of the multi-class categorizations. From this table, we can tell that we can achieve 62.07% of accuracy for multi-class classification. We believe there are several reasons contributing the low accuracy: (1) The input to the system is the raw image pixels and we did not use any pre-processing techniques; (2) We are performing multiclass categorizations. The intra-class differences are large while the inter-class differences are small. Due to the space limit, we did not draw the table showing the confusion matrix.

Table 3: Average precisions for multi-class classification

# of Iterations	10,000	30,000	50,000	80,000	100,000
Average Precisions	43.48%	61.68%	61.92%	62.05%	62.07%

5. Conclusions and Future Directions

Tuberculosis (TB) is one of the world's deadliest diseases, according to the World Health Organization (WHO) and Center for Disease Control and Prevention (CDC). 10.4 million people around the world became sick in 2015. Most the infected populations were from resource-poor communities with weak healthcare infrastructure. Mobile technologies have the potential to reduce the burden of TB by providing mobile computing and communication techniques and devices for better diagnosis, treatment and prevention in these marginalized communities. The goal of our research is to reduce patient wait times to be diagnosed with TB by implementing a socio-technical solution to optimize the diagnosis process. As the first step of this research project, in this paper, we introduce the two major progresses we have made. The first progress is to build a large-scale, real-world, and well-annotated chest X-ray image database dedicated for TB screening. The second progress is to develop effective and efficient computational model for TB manifestation categorization. Preliminary results have demonstrated the feasibility of the proposed approach.

Based on the proposed framework and the preliminary work reported in this paper, our final goal is to produce a scalable solution to improve healthcare system in Perú, and globally, with mobile technologies. We will continue to develop the large scale, real-world X-ray TB database with reliable content annotated by a trained team and verified by a specialized Pulmonologist, as well as continuing to improve the performance of the computing algorithms. We will also implement a scalable solution by making the mobile device-based computer-aided system available as an open source software platform. We will conduct field-testing in tuberculosis clinics in the city of Carabayllo (Lima, Perú).

Acknowledgement

This project is supported in partial by National Institutes of Health of the United States (Award #: 1R01EB021900), National Science Foundation of the United States (Award No. 1547428, 1541434, 1440737, 1229213, and 1156639). Points of view or opinions in this document are those of the authors and do not represent the official position or policies of the U.S. NIH and NSF.

References

- [1] T. Ferkol and D. Schraufnagel, "The global burden of respiratory disease," *Annals of the American Thoracic Society*, vol. 11, pp. 404-406, 2014.
- [2] D. adfas. *Tuberculosis Fact Sheets*, by World Health Organization, available at <http://www.who.int/mediacentre/factsheets/fs104/en/>.
- [3] B. Caputo, T. Tommasi, H. Müller, T. M. Deserno, and J. Kalpathy-Cramer, "ImageCLEF 2009 lung nodule detection and medical annotation task, Available at <http://www.imageclef.org/2009/medanno>," 2009.
- [4] . *Japanese Society of Radiological Technology (JSRT) Digital Image Database*, available at <http://www.jsrt.or.jp/jsrt-db/eng.php>.
- [5] B. van Ginneken, S. G. Armato, B. de Hoop, S. van Amelsvoort-van de Vorst, T. Duindam, M. Niemeijer, K. Murphy, A. Schilham, A. Retico, and M. E. Fantacci, "Comparing and combining algorithms for computer-aided detection of pulmonary nodules in computed tomography scans: the ANODE09 study," *Medical image analysis*, vol. 14, pp. 707-722, 2010.
- [6] S. Jaeger, A. Karagyris, S. Antani, and G. Thoma, "Detecting tuberculosis in radiographs using combined lung masks," in *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, 2012, pp. 4978-4981.
- [7] S. Jaeger, S. Antani, and G. Thoma, "Tuberculosis screening of chest radiographs," *SPIE Newsroom*, 2011.
- [8] S. Jaeger, A. Karagyris, S. Candemir, J. Siegelman, L. Folio, S. Antani, G. Thoma, and C. J. McDonald, "Automatic screening for tuberculosis in chest radiographs: a survey," *Quantitative imaging in medicine and surgery*, vol. 3, pp. 89-99, 2013.
- [9] T. Xu, I. Cheng, R. Long, and M. Mandal, "Novel coarse-to-fine dual scale technique for tuberculosis cavity detection in chest radiographs," *EURASIP Journal on Image and Video Processing*, vol. 2013, pp. 1-18, 2013.
- [10] S. Ya-Lin and Y. Yang, "Localization algorithm and implementation for focal of pulmonary tuberculosis chest image," in *Machine Vision and Human-Machine Interface (MVHI), 2010 International Conference on*, 2010, pp. 361-364.
- [11] P. Maduskar, L. Hogeweg, R. Philipsen, S. Schalekamp, and B. van Ginneken, "Improved texture analysis for automatic detection of tuberculosis (TB) on chest radiographs with bone suppression images," in *SPIE Medical Imaging*, 2013, pp. 86700H-86700H-6.
- [12] O. Mohd Rijal, H. Ebrahimian, and N. M. Noor, "Determining features for discriminating PTB and normal lungs using phase congruency model," in *Biomedical and Health Informatics (BHI), 2012 IEEE-EMBS International Conference on*, 2012, pp. 341-344.
- [13] J. H. Tan, U. R. Acharya, C. Tan, K. T. Abraham, and C. M. Lim, "Computer-assisted diagnosis of tuberculosis: a first order statistical approach to chest radiograph," *Journal of medical systems*, vol. 36, pp. 2751-2759, 2012.
- [14] U. Avni, H. Greenspan, E. Konen, M. Sharon, and J. Goldberger, "X-ray categorization and retrieval on the organ and pathology level, using patch-based visual words," *Medical Imaging, IEEE Transactions on*, vol. 30, pp. 733-746, 2011.
- [15] B. A. Israel, A. J. Schulz, E. A. Parker, and A. B. Becker, "Review of community-based research: assessing partnership approaches to improve public health," *Annual review of public health*, vol. 19, pp. 173-202 % @ 0163-7525, 1998.
- [16] N. Wallerstein, "Empowerment to reduce health disparities," *Scandinavian journal of public health*, vol. 30, pp. 72-77 % @ 1403-4948, 2002.
- [17] N. B. Wallerstein and B. Duran, "Using community-based participatory research to address health disparities," *Health promotion practice*, vol. 7, pp. 312-323 % @ 1524-8399, 2006.
- [18] E. F. Ruiz, Á. Proaño, O. J. Ponce, and W. H. Curioso, "Mobile health for public health in Peru: lessons learned," *Revista Peruana de Medicina Experimental y Salud Pública*, vol. 32, pp. 264-272, 2015.
- [19] V. Mani, S. Wang, F. Inci, G. De Libero, A. Singhal, and U. Demirci, "Emerging technologies for monitoring drug-resistant tuberculosis at the point-of-care," *Advanced drug delivery reviews*, vol. 78, pp. 105-117, 2014.
- [20] M. Zimic, J. Coronel, R. H. Gilman, C. G. Luna, W. H. Curioso, and D. A. Moore, "Can the power of mobile phones be used to improve tuberculosis diagnosis in developing countries?," *Transactions of the Royal Society of Tropical Medicine and Hygiene*, vol. 103, pp. 638-640, 2009.

- [21] H. Zhu, S. O. Isikman, O. Mudanyali, A. Greenbaum, and A. Ozcan, "Optical imaging techniques for point-of-care diagnostics," *Lab on a Chip*, vol. 13, pp. 51-67, 2013.
- [22] A. B. Schwartz, G. Siddiqui, J. S. Barbieri, A. L. Akhtar, W. Kim, R. Littman-Quinn, E. F. Conant, N. K. Gupta, B. A. Pukenas, and P. Ramchandani, "The accuracy of mobile teleradiology in the evaluation of chest X-rays," *Journal of telemedicine and telecare*, vol. 20, pp. 460-463, 2014.
- [23] C. J. Murray, K. F. Ortblad, C. Guinovart, S. S. Lim, T. M. Wolock, D. A. Roberts, E. A. Dansereau, N. Graetz, R. M. Barber, and J. C. Brown, "Global, regional, and national incidence and mortality for HIV, tuberculosis, and malaria during 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013," *The Lancet*, vol. 384, pp. 1005-1070, 2014.
- [24] D. G. Storla, S. Yimer, and G. A. Bjune, "A systematic review of delay in the diagnosis and treatment of tuberculosis," *BMC public health*, vol. 8, p. 15, 2008.
- [25] M. Breuninger, B. van Ginneken, R. H. Philipsen, F. Mhimira, J. J. Hella, F. Lwilla, J. van den Hombergh, A. Ross, L. Jugheli, and D. Wagner, "Diagnostic accuracy of computer-aided detection of pulmonary tuberculosis in chest radiographs: a validation study from sub-saharan Africa," 2014.
- [26] M. Muyoyeta, P. Maduskar, M. Moyo, N. Kasese, D. Milimo, R. Spooner, N. Kapata, L. Hogeweg, B. van Ginneken, and H. Ayles, "The sensitivity and specificity of using a computer aided diagnosis program for automatically scoring chest X-rays of presumptive TB patients compared with Xpert MTB/RIF in Lusaka Zambia," *PloS one*, vol. 9, p. e93757, 2014.
- [27] . ANODE.
- [28] L. Wang and D.-C. He, "Texture classification using texture spectrum," *Pattern Recognition*, vol. 23, pp. 905-910, 1990.
- [29] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, pp. 971-987, 2002.
- [30] I. Daubechies, *Ten lectures on wavelets* vol. 61: SIAM, 1992.
- [31] H. Müller, P. Clough, W. Hersh, and A. Geissbuhler, "IMAGECLEF 2004–2005: results, experiences and new ideas for image retrieval evaluation," in *CBMI 2005 - Fourth International Workshop on Content-Based Multimedia Indexing*, Riga, Latvia, 2005.
- [32] H. Müller, T. Deselaers, T. Deserno, P. Clough, E. Kim, and W. Hersh, "Overview of the ImageCLEFmed 2006 Medical Retrieval and Medical Annotation Tasks," in *7th Workshop of the Cross-Language Evaluation Forum, CLEF 2006*, 2006, pp. 595-608.
- [33] H. Müller, T. Deselaers, T. Deserno, J. Kalpathy-Cramer, E. Kim, and W. Hersh, "Overview of the ImageCLEFmed 2007 Medical Retrieval and Medical Annotation Tasks," in *8th Workshop of the Cross-Language Evaluation Forum, CLEF 2007*, 2007, pp. 1-11.
- [34] H. Muller, J. Kalpathy-Cramer, I. Eggel, S. Bedrick, J. Reisetter, C. E. K. Jr., and W. Hersh, "Overview of the CLEF 2010 medical image retrieval track," in *11th Workshop of the Cross-Language Evaluation Forum*, 2010, pp. 1-11.
- [35] H. Muller, J. Kalpathy-Cramer, I. Eggel, S. Bedrick, S. i. Radhouani, B. Bakke, C. E. K. Jr, and W. Hersh, "Overview of the CLEF 2009 medical image retrieval track," in *10th Workshop of the Cross-Language Evaluation Forum*, 2009, pp. 1-11.
- [36] H. Müller, J. Kalpathy-Cramer, C. E. K. Jr, W. Hatt, S. Bedrick, and W. Hersh, "Overview of the ImageCLEF 2008 Medical Image Retrieval Task," in *9th Workshop of the Cross-Language Evaluation Forum*, 2008, pp. 1-11.
- [37] (2011). *ImageCLEF, Medical Image Retrieval Task*. Available: <http://www.imageclef.org/2011/medical>
- [38] H. Müller, A. G. S. de Herrera, J. Kalpathy-Cramer, D. Demner-Fushman, S. Antani, and I. Eggel, "Overview of the ImageCLEF 2012 Medical Image Retrieval and Classification Tasks," in *CLEF (Online Working Notes/Labs/Workshop)*, 2012.
- [39] B. Caputo, H. Muller, B. Thomee, M. Villegas, R. Paredes, D. Zellhofer, H. Goeau, A. Joly, P. Bonnet, and J. M. Gomez, "ImageCLEF 2013: the vision, the data and the open challenges," in *Information Access Evaluation. Multilinguality, Multimodality, and Visualization*, ed: Springer, 2013, pp. 250-268.
- [40] A. G. S. de Herrera, J. Kalpathy-Cramer, D. Demner-Fushman, S. Antani, and H. Müller, "Overview of the ImageCLEF 2013 medical tasks," *Working notes of CLEF*, 2013.
- [41] B. Hu, S. Dasmahapatra, P. Lewis, and N. Shadbolt, "Ontology-based medical image annotation with description logics," in *Tools with Artificial Intelligence, 2003. Proceedings. 15th IEEE International Conference on*, 2003, pp. 77-82.
- [42] D. L. Rubin, P. Mongkolwat, V. Kleper, K. Supekar, and D. S. Channin, "Medical Imaging on the Semantic Web: Annotation and Image Markup," in *AAAI Spring Symposium: Semantic Scientific Knowledge Integration*, 2008, pp. 93-98.
- [43] B. C. Russell, A. Torralba, K. P. Murphy, and W. T. Freeman, "LabelMe: a database and web-based tool for image annotation," *International journal of computer vision*, vol. 77, pp. 157-173, 2008.
- [44] Y. Bengio, "Learning deep architectures for AI," *Foundations and trends® in Machine Learning*, vol. 2, pp. 1-127, 2009.
- [45] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, pp. 504-507, 2006.
- [46] (2012). *Scientists See Promise in Deep-Learning Programs*, by John Markoff, *New York Times*. Available: <http://www.nytimes.com/2012/11/24/science/scientists-see-advances-in-deep-learning-a-part-of-artificial-intelligence.html>

- [47] (2013). *Deep Learning makes MIT Tech Review list of top-10 breakthroughs of 2013 by MIT Tech Review*. Available: <http://www.technologyreview.com/featuredstory/513696/deep-learning/>
- [48] (2013). *NYU Deep Learning Professor LeCun Will Head Facebooks New Artificial Intelligence Lab*, by Josh Constine, *Ntechcrunch.com*. Available: <http://techcrunch.com/2013/12/09/facebook-artificialintelligence-lab-lecun/>
- [49] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *NIPS*, 2012, p. 4.
- [50] D. Eigen, J. Rolfe, R. Fergus, and Y. LeCun, "Understanding Deep Architectures using a Recursive Convolutional Network," *arXiv preprint arXiv:1312.1847*, 2013.
- [51] P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol, "Extracting and composing robust features with denoising autoencoders," in *Proceedings of the 25th international conference on Machine learning*, 2008, pp. 1096-1103.
- [52] Y. Bengio, L. Yao, G. Alain, and P. Vincent, "Generalized denoising auto-encoders as generative models," in *Advances in Neural Information Processing Systems*, 2013, pp. 899-907.
- [53] R. Salakhutdinov, A. Mnih, and G. Hinton, "Restricted Boltzmann machines for collaborative filtering," in *Proceedings of the 24th international conference on Machine learning*, 2007, pp. 791-798.
- [54] G. E. Hinton, "A practical guide to training restricted boltzmann machines," in *Neural Networks: Tricks of the Trade*, ed: Springer, 2012, pp. 599-619.
- [55] R. Salakhutdinov and G. E. Hinton, "Deep boltzmann machines," in *International Conference on Artificial Intelligence and Statistics*, 2009, pp. 448-455.
- [56] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going Deeper With Convolutions," in *Computer Vision and Pattern Recognition (CVPR), 2015 IEEE Conference on*, 2015.
- [57] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*, 2014, pp. 580-587.
- [58] J. R. Uijlings, K. E. van de Sande, T. Gevers, and A. W. Smeulders, "Selective search for object recognition," *International Journal of Computer Vision*, vol. 104, pp. 154-171, 2013.
- [59] . *CUDA Toolkit v6.5 Documentation - Profiler*, Available at <http://docs.nvidia.com/cuda/profiler-users-guide/> - abstract.
- [60] . *Intel Pin Tool*, Available at <https://software.intel.com/en-us/articles/pintool>.
- [61] . *Multiprocessor scheduling wiki page*, Available at http://en.wikipedia.org/wiki/Multiprocessor_scheduling.
- [62] P. Li, X. Huang, T. Zhao, Y. Luo, and Y. Cao, "Sparkling: Identification of Task Skew and Speculative Partition of Data for Spark Applications," presented at the Spark Summit 2014, 2014.
- [63] M. Wang, H. Li, D. Tao, K. Lu, and X. Wu, "Multimodal graph-based reranking for web image search," *Image Processing, IEEE Transactions on*, vol. 21, pp. 4649-4661, 2012.
- [64] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision (IJCV)*, vol. 60, pp. 91-110, November 2004.
- [65] A. Bosch, A. Zisserman, and X. Munoz, "Representing shape with a spatial pyramid kernel," in *Proceedings of the 6th ACM international conference on Image and video retrieval*, 2007, pp. 401-408.
- [66] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, 2012, pp. 1097-1105.
- [67] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, pp. 273-297, 1995.
- [68] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, "Caffe: Convolutional architecture for fast feature embedding," in *Proceedings of the ACM International Conference on Multimedia*, 2014, pp. 675-678.
- [69] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, 2009, pp. 248-255.