BIOCYBERNETICS AND BIOMEDICAL ENGINEERING 38 (2018) 691–699

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/bbe

Original Research Article

Automatic detection of tuberculosis bacilli from microscopic sputum smear images using deep learning methods

Rani Oomman Panicker [a,c,*], *Kaushik S. Kalmady* [b], *Jeny Rajan* [b], M.K. Sabu [c]

[a] *Department of Information Technology, College of Engineering, Trikaripur, Kerala, India*
[b] *Department of Computer Science and Engineering, National Institute of Technology Karnataka, Surathkal, India*
[c] *Department of Computer Applications, Cochin University of Science and Technology, Kochi, India*

ARTICLE INFO

Article history:
Received 12 March 2018
Received in revised form
16 May 2018
Accepted 28 May 2018
Available online 13 June 2018

Keywords:
Tuberculosis
Sputum smear microscopy
Automatic TB detection
Convolutional neural network
Image processing

ABSTRACT

An automatic method for the detection of Tuberculosis (TB) bacilli from microscopic sputum smear images is presented in this paper. According to WHO, TB is the ninth leading cause of death all over the world. There are various techniques to diagnose TB, of which conventional microscopic sputum smear examination is considered to be the gold standard. However, the aforementioned method of diagnosis is time intensive and error prone, even in experienced hands. The proposed method performs detection of TB, by image binarization and subsequent classification of detected regions using a convolutional neural network. We have evaluated our algorithm using a dataset of 22 sputum smear microscopic images with different backgrounds (high density and low-density images). Experimental results show that the proposed algorithm achieves 97.13% recall, 78.4% precision and 86.76% F-score for the TB detection. The proposed method automatically detects whether the sputum smear images is infected with TB or not. This method will aid clinicians to predict the disease accurately in a short span of time, thereby helping in improving the clinical outcome.

1. Introduction

Tuberculosis (TB) is a potentially serious contagious infection and is caused by Mycobacterium tuberculosis. It is an airborne disease that spreads from one person to another through coughs, sneezes, speaks, spits etc. and it predominantly attacks the lungs (pulmonary TB). But it can also infect or damage other organs (extrapulmonary TB) such as the brain, spine, kidney etc. and may eventually lead to death, if not handled properly. These bacteria are rod-shaped, slow growing with varying curvature and have a length ranging from 1 to 10 μm [9]. Depending upon the level of infection, doctors classified TB into two forms: active TB and latent (inactive) TB. Among them, active TB is more dangerous and is contagious. Patients with latent TB do not spread infection, but sometimes it can be converted to active TB at a later stage.

* *Corresponding author at*: Department of Information Technology, College of Engineering, Trikaripur, Kerala, India.
E-mail address: oommanrani@yahoo.co.in (R.O. Panicker).
https://doi.org/10.1016/j.bbe.2018.05.007

Based on the report of World Health Organization (WHO) 2017 [1], in 2016 itself, 1.3 million people died and an estimated 10.4 million fell ill due to TB. Top TB burden countries include India, Indonesia, China, Pakistan etc., and among these countries, India (25%)is having the maximum population of TB patients [1]. Every year millions of people worldwide are affected with TB infection, leading to WHO announcing TB as a global emergency [2].

There are many established TB detection methods available such as microscopy, tuberculin skin test (TST), chest X-ray, interferon-γ release assay (IGRA), culture test, and GeneXpert etc. [3]. However, microscopic sputum smear examination using conventional microscope is a widely used technique all over the world especially in low and middle-income countries due to its low cost, ease of use and maintenance; and also provides sufficiently faster results compared to other tests [3]. Sputum smear examination can be done in two ways: one with the help of bright field or conventional microscope and the other with the help of a fluorescent microscope, and both of them vary in their power of the lens and staining. Former one uses 100× lens and Ziehl-Neelsen (ZN) acid-fast staining procedure and the latter one uses 25× and auramine-O [4].

Accurate and on time treatments are required for controlling the number of TB cases. Manual detection and counting of TB bacilli through the microscopic eyepiece is a tedious task, which requires highly skilled lab technicians, and lots of mental and physical (eye) strain [3]. The accuracy of TB detection always depends upon the technician's level of expertise and knowledge. Technicians, in general, have to analyze between 20 and 100 fields of each sputum smear slide, which can take about 40 min to 3 h depending upon the patient's infection level [6]. According to Veropoulos et al. [7], the manual TB detection and counting methods may miss 33–50% of active cases and the automated methods may help to increase the diagnostic sensitivity rate, as machines can screen a large number of fields in a short duration and can detect the TB bacillus more accurately in the early stage itself. The other advantages of automated TB detection methods are low mental and physical strain, faster and accurate decision making, patient record use, multi-head visualization and communication use (second opinion) [10]. The usage of these types of automated systems and methods will help the clinicians to provide better treatments to the needy and thereby the hospitals can meet the quality standards.

In the past few years, many automated methods have been proposed to detect TB bacilli from conventional as well as fluorescence microscopic images. Most of the existing methods use handcrafted feature vectors to discriminate bacilli pixels from non-bacilli pixels. The performance of these methods heavily depends on the bacilli features considered. Also, manually designing feature vectors for complex tasks requires a great deal of human time and effort [31]. In this work, we used a customized convolutional neural network (CNN) to identify the pixels in the image that belongs to bacilli. The advantage of CNN over other machine learning techniques is that instead of handcrafted feature vectors, the CNN automatically learns the characteristics of bacilli, provided enough samples to learn. We used a standard dataset provided by Costa et al. [27] to train and test the proposed CNN framework.

This paper is structured as follows. Section 2 describes the review of existing papers. In Section 3, the proposed methods are discussed. Experimental results, data sets, and discussions are given in Section 4. Finally, conclusions are drawn in Section 5.

2. Related work

Many researchers proposed fully and semi-automatic TB bacilli detection methods for sputum smear images acquired using a conventional or fluorescent microscope. The general steps involved in developing automated systems for TB bacilli detection are: (i) image acquisition and pre-processing, (ii) segmentation, (iii) feature extraction, and (iv) classification of TB bacilli. The first attempt to detect TB bacilli from microscopic sputum smear images using image processing techniques was done by Veropoulos et al. [8] in 1998. They demonstrated the efficacy of their method on images acquired with a fluorescent microscope. Costa et al. [5] were the first to publish bacilli detection using image processing techniques on bright field microscopic images. In the remaining part of this section, we will be discussing automatic TB detection methods using conventional microscopic images only.

Costa et al. [5] proposed an adaptive thresholding-based segmentation method for the detection of TB in images of ZN-stained sputum smear images. They used morphological filters for handling the artifacts. But the sensitivity of their method was relatively less (76.65%). In 2008 itself, Sadaphal et al. [11] used Bayesian segmentation and shape features (axis ratio and eccentricity) for the detection of TB bacilli. However, their method failed in identifying the overlapping or conglomerated bacilli. An adaptive hue based segmentation method was later proposed by Makkapati et al. [12] in 2009, for detecting TB bacilli. In [13], a thresholding-based segmentation was used by considering Cr and a plane of YCbCr and Lab color space for the identification of bacilli from sputum smear images. They got an accuracy of 85.7%. Zhai et al. [14] proposed a two-stage segmentation (coarse segmentation and fine segmentation) and a decision tree based classification for TB bacilli identification, and their experimental results showed an accuracy greater than 80%. In the same year itself, Nayak et al. [15] proposed HSI color space based segmentation and area-based classification for detecting TB bacilli with 93.5% success rate. A genetic algorithm based neural network (GA-NN) approach has been used for the detection of TB objects and an accuracy of 86.32% has been reported on ZN stained tissue slide images [19].

A pixel classifier combination is used for detecting TB bacilli in [9,16]. These methods failed to identify touching TB bacilli. In 2012, Costa Filho et al. [17] performed a neural network-based classification for detecting TB bacilli with a sensitivity of 91.53%. A Random Forest (RF) based learning method was proposed by Ayas et al. [18] in 2014 for detecting TB bacilli. Recently, in 2015 Costa Filho et al. [20] used segmentation and post-processing method for the automatic identification of TB. The authors achieved a sensitivity of 96.80%. Very recently, a fuzzy-based decision-making approach was developed