Identification of Tuberculosis Bacilli in ZN-Stained Sputum Smear Images: A Deep Learning Approach

Moumen El-Melegy[1], Doaa Mohamed[1], Tarek ElMelegy[2], and Mostafa Abdelrahman[1]
[1]Electrical Engineering Department, School of Engineering
[2]Clinical Pathology Department, School of Medicine
Assiut University, Assiut 71516, Egypt
moumen@aun.edu.eg

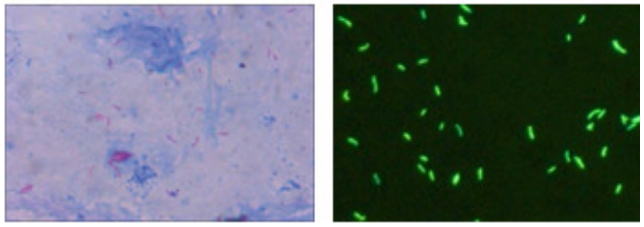
Abstract

Tuberculosis (TB) is a serious infectious disease that remains a global health problem with an enormous burden of disease. TB spreads widely in low and middle income countries, which depend primarily on ZN-stained sputum smear test using conventional light microscopy in disease diagnosis. In this paper we propose a new deep-learning approach for bacilli localization and classification in conventional ZN-stained microscopic images. The new approach is based on the state of the art Faster Region-based Convolutional Neural Network (RCNN) framework, followed by a CNN to reduce false positive rate. This is the first time to apply this framework to this problem. Our experimental results show significant improvement by the proposed approach compared to existing methods, which will help in accurate disease diagnosis.

1. Introduction

Tuberculosis (TB) is one of the leading causes of death in the world [1] that has millions of victims and patients. Over 95% of TB deaths occur in low- and middle-income countries. However TB can be treated successfully if it is diagnosed correctly at the appropriate time. There are several methods for TB diagnosis, such as chest X-ray test, culture test, interferon-γ release assay (IGRA), GeneXpert, skin test, and microscopy test. Yet diagnosis remains a challenging task especially in low and middle-income countries that depend primarily on manual diagnosis of TB with visually screening stained smears prepared from sputum.

While sputum smear test is a simple, inexpensive test and results can be available within hours, manual TB screening is a tedious work and prone to error due to work load and a dearth of properly trained technicians. Technicians view the smears slides with microscopes, looking for rod-shaped objects that may be Mycobacterium tuberculosis, the bacteria responsible for TB disease. However they may diagnose a positive TB slide as smear negative because of sparseness of acid-fast bacilli, or because too few fields have been examined. This often leads to low recall rates. Automatic methods are the best solution to improve low sensitivity of TB diagnosis, reduce human variability in slide analysis and speed up the screening process.



(a) (b)

Figure 1. (a) Image from bright field microscope (b) image from fluorescence microscope.

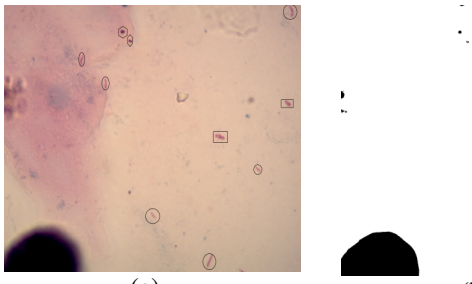
While fluorescence microscopy is more sensitive and faster to diagnose [2, 3] (see Fig. 1), conventional microscopy is mostly used in low and middle-income countries where TB prevalence rate is high because it is less expensive, easier to use and maintain [2]. As such, we will focus in this work on this kind of microscopy.

In this work we are interested in using state-of-the-art object detection methods to identify TB bacillus in bright-field microscopy images. This faces special challenges due to variations in illumination, artifacts, overstaining and lack of clear separation of bacilli from background.

Previous attempts to automate the task of identifying and quantifying TB (bacilli) in the images acquired with conventional light microscopy have been reported. Costa et al. [4] were the first to propose such an automatic TB bacilli identification method based on red minus green (R-G) images. To segment the bacilli, a threshold value was computed using the histogram of the R-G image. Afterwards several filtering operations were applied to separate the bacilli from artifacts. Sadaphal et al. [5] applied a Bayesian segmentation method to segment possible TB objects, and then the segmented objects were classified based on shape and color descriptors. Osman et al. [6] presented a study on contrast enhancement using linear

stretching for bacilli detection on both RGB and HSV color spaces. The same authors [7] proposed an automatic segmentation method using hybrid multilayered neural network. Khutlang [8] proposed a combination of several pixel classifiers to segment the candidate bacillus objects, and then shape and color features were extracted for the identification of the hopefully-true bacilli. Rulaningtyas et al. [9] employed a naive Bayesian-based color segmentation for detecting TB bacilli after performing image pre-processing. CostaFilho et al. [10] classified images into low density and high density background, then a set of color features were used for segmentation and classification. Kusworo et al. [11] employed a color segmentation method on TB images, then eccentricity and compactness shape features were extracted for final classification with a support vector machine.

All the previous methods have combined workflows of image processing techniques and machine learning for TB identification. They relied on hand-crafted sets of color shape descriptors for that goal, which resulted in rather low detection accuracy. In recent years, researchers [12-15] have started applying emerging, powerful deep learning methods which allow learning discriminating features for bacilli detection and classification. The Convolutional Neural Network (CNN) is the main engine for all these methods. In [12-13, 15], the authors had to split a microscopic image into smaller patches, each containing an image object that could potentially be a TB bacillus. The CNN operated on patches (not the whole image) of the used datasets. The main drawback of these methods is how to split the larger microscopic image into such a way. The accuracy of the method largely depends on this preliminary patching step. Some authors even did not reveal the details of how this was done [13, 15], whether it is automated or even done manually. The recent work of [14] tried to overcome this drawback by using an initial stage of image binarization and pixel classification to locate foreground objects (bacilli, non-bacilli, artifacts) and then construct the required patches. Each patch presumably will contain one foreground object and is fed to the CNN stage for final classification into bacilli and non-bacilli. While this method automates the image patching, its overall accuracy depends on the success of the first binarization/pixel-classification step, which is often error-prone for the challenging conventional bright-field microscopy. It also suffers from touching foreground objects and over-stained images. Fig. 2 shows that the binarization step can cause a microscopic image lose real bacilli pixels, thus failing to classify them correctly.



(a) (b)

Figure 2. (a) Sputum smear microscopy image with true bacilli marked, (b) Corresponding binary image.

In this work we choose to use one of the state-of-the-art deep learning-based methods that has the ability for localization and classification of TB bacilli with high performance. It can avoid all previous work difficulties, such as finding proper set of features, and dividing images into patches. We propose to use a Faster Region-based Convolutional Neural Network (RCNN) framework. This framework has achieved the state-of-the-arts results on several object detection and classification challenges (e.g., Pascal VOC [16] and MS COCO [17]). *To the best of our knowledge, this framework has not been adopted before to the task at hand in this paper*. To reduce the rather high false positive rate of the Faster R-CNN we add an additional stage consisting of a CNN that classifies the bounding boxes found by the Faster R-CNN into real or false TB bacilli. This increases the detection performance of the overall approach.

Another contribution of this paper is that we assess and compare notable existing methods against ours on the same dataset. Previous works (e.g, [7-15]) used different datasets and thus it is not possible to compare between the reported results.

The paper is organized as follows. Section 2 describes the data used in this work. In Section 3 the proposed approach along its implementation details is discussed. Experimental results are reported in Section 4. Finally, our conclusions are drawn in Section 5.

2. Data

All images used for training and evaluation are taken from ZNSM-iDB [23] public database that consists of various categories (autofocused data, overlapping objects,single or few bacilli, views without bacilli, occluded bacilli, over-stained views with bacilli and artifacts), see Fig. 3. The database consists of three divisions, each acquired with different microscopes. Our training and test data are taken from first and second divisions, the first using Labomed Digi 3 digital microscope with an iVu 5100 digital camera 5.0 megapixel (MS-1), and the second using Motic BA210 digital microscope with a Siedentopf type binocular

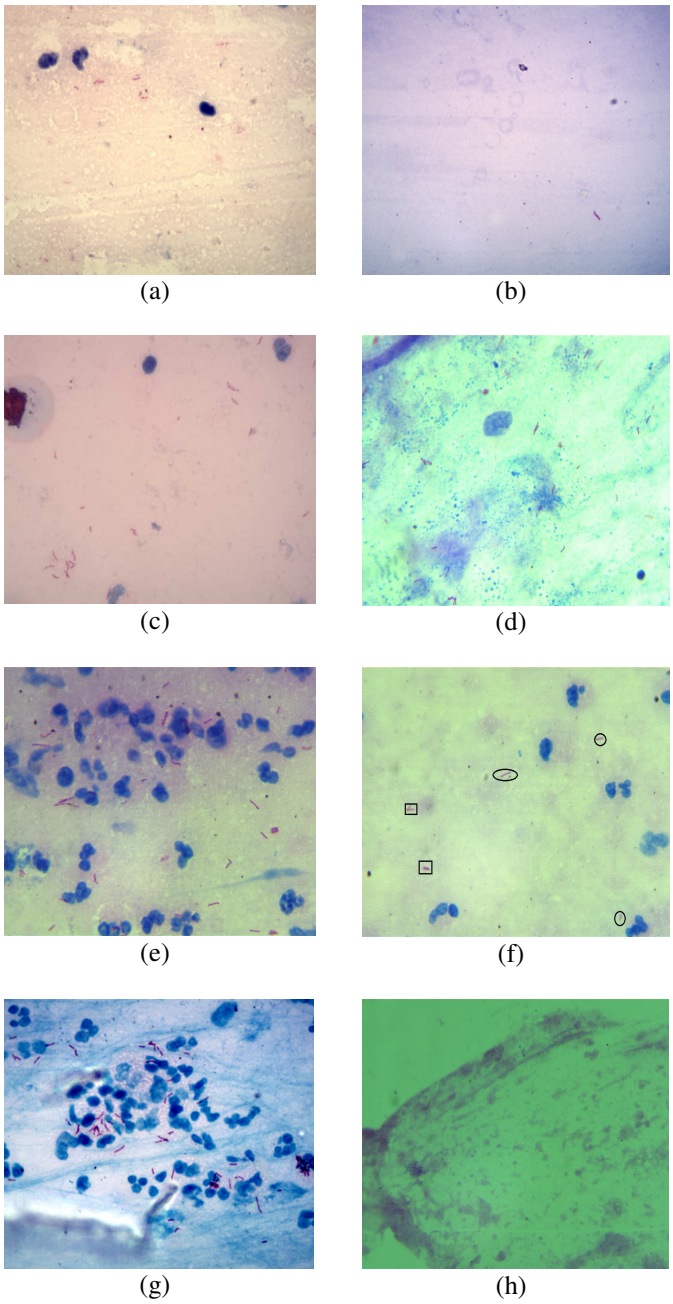


Figure 3. Sample images from the database: (a) image from Autofocus category acquired using MS-1, (b) image from single bacilli category acquired using MS-2, (c) image from overlapping bacillus category acquired using MS-2, (d) image from overlapping bacillus category acquired using MS-1, (e) image from overstaining bacillus category acquired using MS-2, (f) image from manually segmented category acquired using MS-1 circle or oval used to mark single bacilli and square used to mark occluded bacillus, (g) image from occluded category acquired using MS-2, (h) image from without bacillus category acquired using MS-1.

head and Moticam 2500 digital camera module 5.0 megapixel (MS-2). 80% of data are selected at random for training and 20% for testing.

3. Proposed Method

In this work we propose TB identification in images obtained from conventional light microscopy using a deep learning approach. The proposed approach relies on the R-CNN framework family. This framework can classify and locate objects inside images by combining CNNs and region proposal methods [19]. A region proposal method is a method that finds a set of regions, defined with bounding boxes, which might contain objects of interest. Early members of this framework used Selective Search [20] or EdgeBoxes [21]. Although it exceeded the previous best detection results on Pascal VOC 2012 by about 30% [16], it was computationally slow since each image is processed many times to detect region of interest.

Then new members of the family were proposed to solve this drawback. Fast R-CNN [22] introduces a more effective method for training the CNN and adopts a bounding box regressor. The bounding box regressor is a layer that outputs the locations of bounding boxes where objects of interest might be located. Faster R-CNN [16] combines the Fast R-CNN with a RPN (Region Proposal Network). RPN is a network that uses the feature map, to generate regions of interest each with score, then fast R-CNN classifies the proposed regions and refines their locations. The time cost of generating region proposals is much smaller in RPN than selective search.

In order to share convolutional layers between RPN and Faster R-CNN model, an algorithm of four steps was presented in [16]. In the first step, input image is passed to a conventional network which returns feature map for that image. Second step, RPN is applied to this feature map and returns object proposals with scores. Third step, Region Of Interest (ROI) pooling layer is applied to these proposals to down sample it to the same size. In the fourth step, object proposals are passed to fully connected layer to classify and output the bounding boxes for objects. Figure 4 illustrates the four steps of the Faster R-CNN framework.

3.1 Second classification stage

Our earlier experiments have revealed that the Faster R-CNN suffers from high false positive rate, which has a negative impact on the overall accuracy. As such, we propose to detect TB bacilli in two stages; we propose to use a CNN as a second stage on top of the Faster R-CNN stage, see Fig. 5 and Fig. 6.

The bounding boxes found by the first stage of the Faster R-CNN are fed to the second stage CNN for final

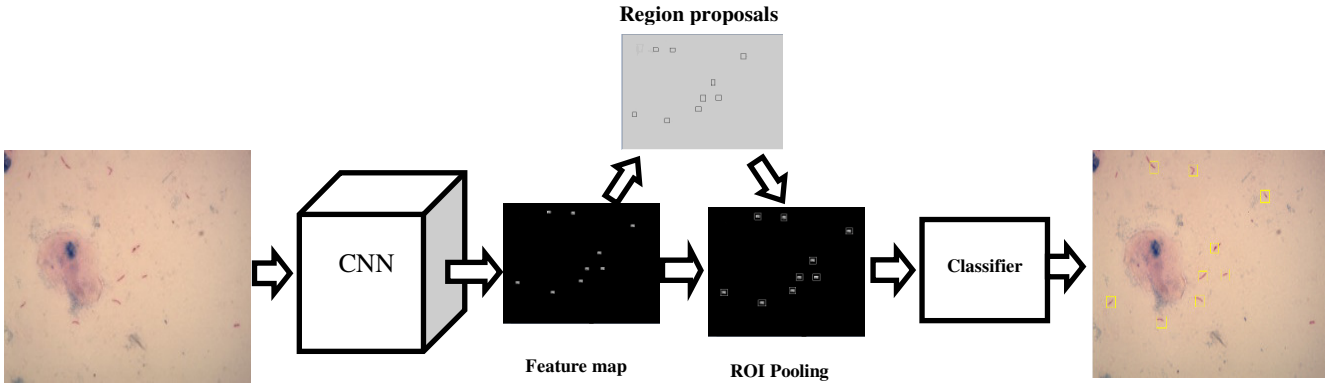


Figure 4. Faster R-CNN is a network that combines a Convolutional Neural Network, a Region Proposal Network, a Region of Interest, Pooling layer, and a classifier.

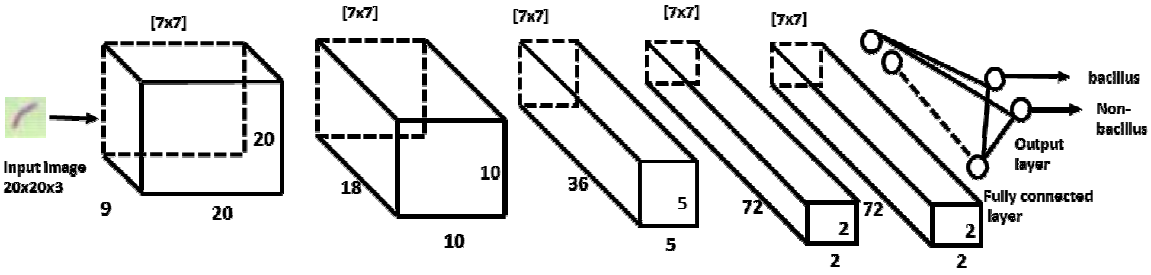


Figure 5. Structure of the second stage classifier.

classification into TB or non-TB bacillus. Figure 6 illustrates the complete proposed approach.

3.2 Implementation

In our implementation of this approach, the Faster R-CNN is structured as follows. The network input image layer has an input size of 400×400×3, middle layers containing 3 convolutional layers with 32 (3×3) kernels, each followed with a Rectified Linear Units (ReLUs) activation layer and a max pooling layer, and final layers consisting of a fully connected layer, a softmax layer and a classification layer. We have used 500 images with 2500 (bacilli) object, selected from all database categories. 80% of dataset used for training and 20% for testing. For proper Faster R-CNN training and to get more accuracy, multiple data augmentation methods have been used, such as image rotation, reflection and translation. Several key hyper-parameters have been tuned in the Faster R-CNN architecture, such as minimum anchor box sizes (we use 3 box sizes 10^2, 15^2, and 20^2). The batch size is 128. The learning rate is 1e-5 for first two stages and 1e-6 for third and fourth stages. The network is trained for 40 epochs.

As shown in Fig. 5, the second stage CNN has an input size of 20×20×3, and has five convolutional layers (each layer has 7×7 filter size, and Rectified Linear Units (ReLUs) activation functions, and is followed by max pooling layer). The number of filter for each layer is taken as 9, 18, 36, 72, and 72, respectively. The last layer of the CNN consists of a fully connected network employing a Softmax activation function and a classification layer with two outputs. The CNN has been trained with the same Faster R-CNN training data. The found boxes of the Faster R-CNN are re-sized to 20×20 and used as input. False positive boxes are used as training samples for the non-TB class, while true positive boxes are used as samples of the TB class. Data augmentation on these samples (random rotations and reflections) is also employed. The batch size is 128, and the network is trained for 40 epochs.

4. Experimental Results

In this section we report the results obtained from the proposed approach. To determine the efficiency of our approach, precision, recall and F-score metrics are calculated. To calculate these metrics, True positive (TP) which is total number of TB bacilli that are detected correctly by the approach, False Positive (FP) which is total number of non-bacilli that are detected as bacilli by the approach, and False Negative (FN) which is total number of true bacilli that are not detected by the approach are calculated.

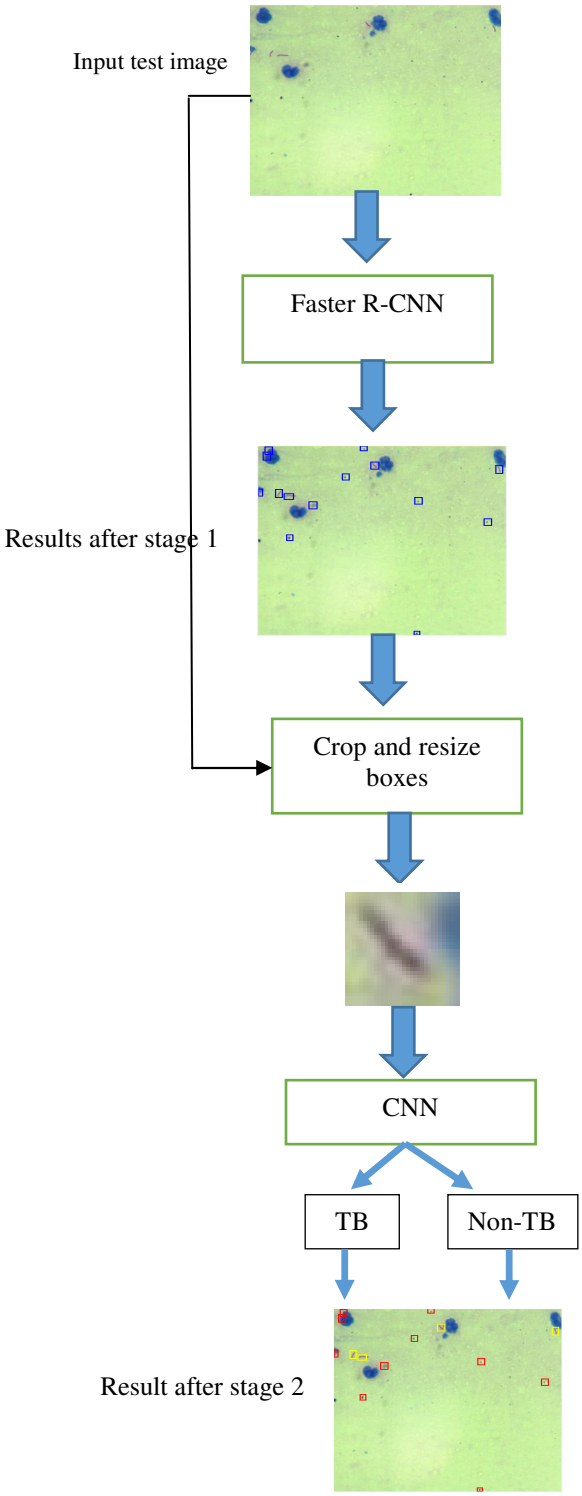


Figure 6. The complete proposed approach consisting of two stages: a Faster R-CNN and then a CNN. In the final result image, yellow boxes denote the finally-obtained TB bacilli, while red boxes denote boxes found by the Faster R-CNN but rejected by the second stage CNN.

For the sake of comparison, we have implemented the method proposed in [8]. The reason of selecting this reference is that it uses a comprehensive set of features and several machine learning classifiers for TB bacilli identification. As such, it is considered a plethora of methods altogether. It consists of four stages: first stage uses a combination of pixel classifiers to segment the candidate bacillus objects based on color intensity features. In stage two, shape and color features, such as color moment features, eccentricity, compactness, Fourier features, were extracted. In stage three, feature selection was carried out. Classic machine learning classifiers (SVMs, kNNs, linear (LDA) and quadratic (QDA) classifiers) have been applied for object classification in stage four.

Table 1 reports the accuracy of the proposed approach on the test dataset in comparisons to the other methods. It is clear that the proposed Faster R-CNN indeed demonstrates a better performance against all other existing methods. However it suffers from rather high false positive rate, which is rectified by the second stage CNN. This stage has succeeded in reducing the false positive rate by 20%. The overall 2-stage approach has presented the best overall performance in terms of the three metrics. Sample results obtained by the complete approach are shown in Fig. 7.

To probe further in the analysis of the proposed approach, we have conducted an experiment to assess the performance of the 2nd stage CNN alone. We have compared its performance against the CNNs proposed in [13] operating on about 2000 patches of the used data, each manually-segmented patch may contain one or no bacillus. The authors of [13] proposed a model (Model A) consisting of 2 convolutional layers and another (Model B) composed of 3 convolutional layers. Table 2 reports the accuracy of our CNN model in comparison to these two CNN models. Clearly it confirms the higher performance of our proposed CNN.

Table 1. Results of proposed approach vs. other methods

Classifier	Recall	Precision	F-score
SVM	94.6%	81.6%	88.0%
*K*NN	93.9%	80.7%	86.8%
QDA	96.3%	79.2%	86.9%
LDA	94.4%	79.4%	86.2%
Faster R-CNN	**98.3%**	**82.6%**	**89.7%**
Faster R-CNN +CNN	**98.4%**	**85.1%**	**91.2%**

5. Conclusions

In this paper we have proposed a deep-learning approach to the localization and classification of TB bacilli

in ZN-stained microscopic images. The new approach is based on the Faster Region-based Convolutional Neural Network (RCNN) framework. To the best of our knowledge, this is the first time to apply this framework to this problem. To reduce the rather high false positive rate of the Faster R-CNN, we have proposed a second stage consisting of a CNN that classifies the bounding boxes found by the Faster R-CNN into real or false TB bacilli. Our experimental results show significant improvement by the proposed approach compared to existing methods.

Our current focus is directed towards improving the performance of the proposed approach. One underway direction to do this is to increase the size of the data used for the deep network training. Since the ZNSM-iDB database [23] was the only one available to us in the public domain, we plan to collect our own samples from Assiut University Hospital and local centers, and use them along with the available ZNSM-iDB data. This is expected to increase the performance of the proposed approach.

Acknowledgement. This research is supported by Information Technology Industry Development Agency of Egypt (grant # CFP139).

Table 2. Comparison of our 2nd stage CNN versus other CNN models.

Model	**F-score**
Model A	80.76%
Model B	84.55%
Ours	**95.27%**

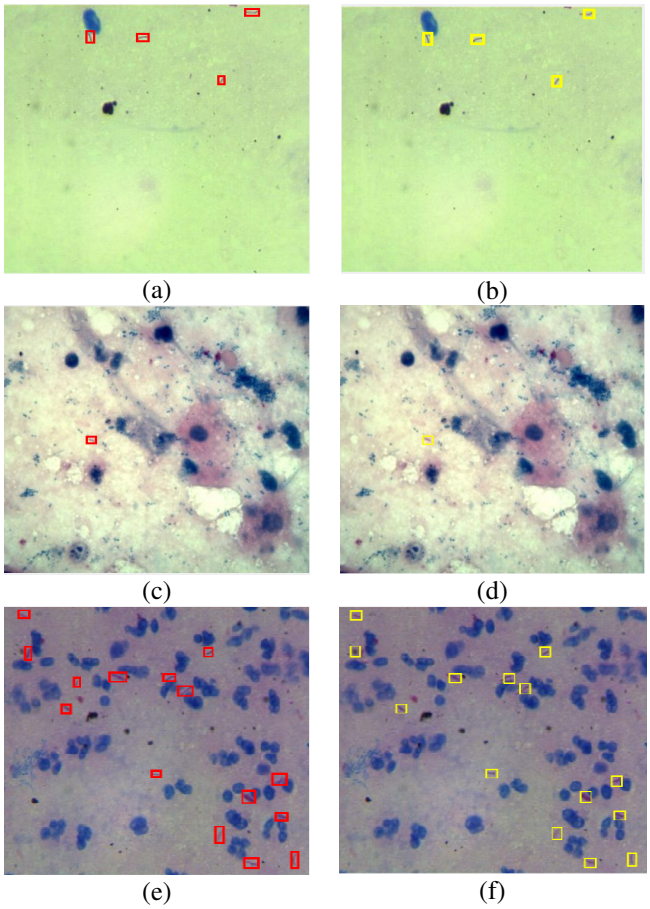


Figure 7. Sample results by the proposed approach: left coloumn shows ground truth images (ground-truth TB bacilli are marked in red boxed), right coloumn shows indentified TB bacilli as marked in yellow boxes.

References

[1] WHO Global tuberculosis report 2018, available online, https://www.who.int/tb/publications/global_report/en/

[2] K. Steingart, M. Henry, V. Ng, P. Hopewell, A. Ramsay, J. Cunningham, R. Urbanczik, M. Perkins, M. Aziz, and M. Pai, Fluorescence versus conventional sputum smear microscopy for tuberculosis: A systematic review. Lancet Infect. Dis., pp. 570–581, Sep. 2006.

[3] Manpreet B., Zeeshan S., P.P.Sharma, RituSinghal, V.P.Myneedu, Rohitsarin, Performance of light-emitting diode fluorescence microscope for diagnosis of tuberculosis, International journal of Mycobacteriology. pp.174-178, September 2013.

[4] M. G. F. Costa, C. F. F. Costa Filho, J. F. Sena, J. Salem, and M. O. de Lima, Automatic identification of mycobacterium tuberculosis with conventional light microscopy. 30th Ann. Int. IEEE EMBS Conf., pp. 382–385, 2008.

[5] Sadaphal, P., Rao, J., Comstock, G.W., and Beg, M.F., Image Processing techniques for identifying Mycobacterium tuberculosis in ziehl-neelsen stains, International Journal of Tuberculosis and Lung Disease. 12(5):579–582, 2008.

[6] Osman, M. K., Mashor, M. Y., and Jaafar, H., Detection of Mycobacterium tuberculosis in Ziehl-Neelsen stained tissue images using Zernike moments and hybrid multilayered perceptron network. In Proc. of IEEE International Conference on Systems Man and Cybernetics (SMC), 2010.

[7] M. K. Osman, M. Y. Mashor, H. Jaafar, and A. S. W. Wahab, Segmentation of tuberculosis bacilli in ziehl-neelsen tissue slide images using hibrid multilayered perceptron network. 10th International Conference on Information Science, Signal Processing and their Applications (ISSPA 2010), May 2010.

[8] R. Khutlang et al., Classification of Mycobacterium tuberculosis in Images of ZN-Stained Sputum Smears. IEEE Transactions on Information Technology in Biomdicine, July 2010.

[9] Rulaningtyas, R., Suksmono, A. B., Mengko, T. L. R., and Saptawati, P., Color segmentation using Bayesian method of Tuberculosis bacteria images in ziehl-neelsen sputum smear. In Proc. Of WiSE Health, Indonesia, 2012.

[10] C. F. F. Costa Filho, P. C. Levy, C. M. Xavier, M. G. F. Costa, L. B. M. Fujimoto, and J. Salem, Mycobacterium Tuberculosis Recognition with Conventional Microscopy. 34th Annual International Conference of the IEEE EMBS, San Diego, 2012.

[11] Kusworo A, Gernowo R, Sugiharto A, Sofjan K, Adi P, Ari B, Tuberculosis (TB) identification in the Ziehl-Neelsen sputum sample in Ntsc channel and support vector machine (SVM) classification. International Journal of Innovative Research in Science, Engineering and Technology, 2:5030–5035, 2013.

[12] John A. Quinn, Rose Nakasi, Pius K. B. Mugagga, Patrick Byanyima, William Lubega, Alfred Andama, Deep Convolutional Neural Networks for Microscopy-Based Point of Care Diagnostics. International Conference on Machine Learning for Health Care, August 2016.

[13] Y. P. López, C. F. F Costa, L. M. R. Aguilera, M. G. F. Costa, Automatic classification of light field smear microscopy patches using Convolutional Neural Networks for identifying Mycobacterium Tuberculosis. CHILEAN Conference on Electrical, Electronics Engineering, Information and Communication Technologies, Oct.2017.

[14] R.O. Panicker, K. S. Kalmady, J. Rajan, M.K. Sabu, Automatic detection of tuberculosis bacilli from microscopic sputum smear images using deep learning methods. Biocybernetics and Biomedical Engineering, pp.691-699, 2018.

[15] S. Kant,M. M. Srivastava, Towards Automated Tuberculosis detection using Deep Learning. arXiv preprint, arXiv:1801.07080, 2018.

[16] S. Ren, K. He, R. Girshick, and J. Sun, Faster R-CNN: Towards real-time object detection with region proposal networks. Proceedings of the Advances in Neural Information Processing Systems, pp. 91–99, December 2015.

[17] Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Doll´ar, P., Zitnick, C.L.: Microsoft COCO: Common Objects in Context. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) ECCV 2014. LNCS, pp. 740–755, 2014.

[18] M. I. Shah, S. Mishra, V. K. Yadav, A. Chauhan, M. Sarkar, S. K. Sharma, Ch. Rout, Ziehl–Neelsen sputum smear microscopy image database: a resource to facilitate automated bacilli detection for tuberculosis diagnosis. Journal of Medical Imaging, June 2017.

[19] R. Girshick, J. Donahue, T. Darrell, and J. Malik, Rich feature hierarchies for accurate object detection and semantic segmentation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 580–587, 2014.

[20] J. R. R. Uijlings, K. E. A. van de Sande, T. Gevers, and A. W. M. Smeulders, Selective search for object recognition. Visual Communication and Image Representation, pp. 154–171, September 2013.

[21] P. D. Larry Zitnick, Edge boxes: Locating object proposals from edges. Proceedings of the European Conference on Computer Vision, pp. 391–405, September 2014.

[22] R. Girshick, Fast R-CNN. Proceedings of the International Conference on Computer Vision, Santiago, Chile, pp. 1440–1448, December 2015.