Accepted for oral presentation at CARS-2013, the 13th International Congress and Exhibition on Computer Assisted Radiology and Surgery, June 21-25, 2016, Heidelberg, Germany

Towards a lung tuberculosis CAD: selecting the CT image analysis method

Vassili Kovalev [1], *Alexander Kalinovsky* [1], *Alena Skrahina* [2], *Alex Rosenthal* [3], *Andrei Gabrielian* [3]

[1] United Institute of Informatics, Minsk, Belarus
[2] Scientific and Practical Center for Pulmonology and Tuberculosis, Minsk, Belarus
[3] National Institute of Allergy and Infectious Diseases, Bethesda, USA

Keywords:
Lung tuberculosis, CT, Image descriptors, CAD

Purpose

The lung tuberculosis (TB) remains a serious problem of public health in a number of regions [1]. This is because of a permanently growing incidence of drug resistant cases, the increasing rate of migration from the countries with low-level incomes to the developed ones, and due to some other factors.

This work is part of a large project, which involves several countries and which supposes creating free information resources on lung TB [2] as well as development of web-based computer assisted diagnosis services (see experimental web site [3]).

Thus, the purpose of this study was to compare the relevant CT image analysis methods and selecting a technique, which is suitable for describing and classification of CT images in the lung tuberculosis CAD.

Methods

Performance of different methods was assessed on the rather hard problem of automatic classification of five TB classes distinguished in real clinical practice. Thus, CT images of 500 TB patients representing 5 TB classes, 100 CTs for each class were sampled from CT archive containing about 9000 TB cases. The classes include: Infiltrative TB (TB-1), Focal TB (TB-2) Tuberculoma (TB-3), Miliary TB (TB-4), and Fibro-cavernous TB (TB-5).

Imaging was performed using LightSpeed Pro 16 scanner with typical slice thickness of 2.5 mm. The archive images can be freely downloaded from the TB portal [2] and viewed/explored using content-based image retrieval engine available in [3]. Lung segmentation was performed by method reported in [4].

The three methods compared were texture analysis methods including:

(a) The extended co-occurrence matrices [5] with voxel pairs considered in 3D and spacing of 1.4 mm (CO-OCC).

(b) The method capitalizing on commonly known Local Binary Patterns, which was adapted for 3D (LBP-3D).

(c) An original method introduced by authors, which employs binary descriptors produced by PCA-generated filters (BF-PCA).

The use of BF-PCA method supposes performing three main steps given below.

Step-1: Generation of a bank of adaptive differential filters:

(a) Random sampling of cube-shaped CT image patches from a subset of input images.

(b) Voxel-wise multiplication of patches by 3D Gaussian.

(b) Supplying the patches to the PCA method and selecting the leading components as filters.

Step-2: Creating binary filtered CT images by convolving with selected set of filters.

Step-3: Generating image descriptors which are essentially the histograms of values computed by voxel-wise concatenation of binary images.

Results

At the experimentation stage, CT images were categorized into 5 TB classed using the above three kinds of image descriptors with the help of SVM, K-nearest neighbors (K=5), and logistic regression classifiers. All the experiments were performed using v-fold validation with v=10.

The best results for all 3 methods were obtained with K-nearest neighbors classifier. Corresponding confusion matrices together with the values of general classification accuracy (i.e. the accuracy of simultaneous classification into 5 classes) are presented in Fig. 1.

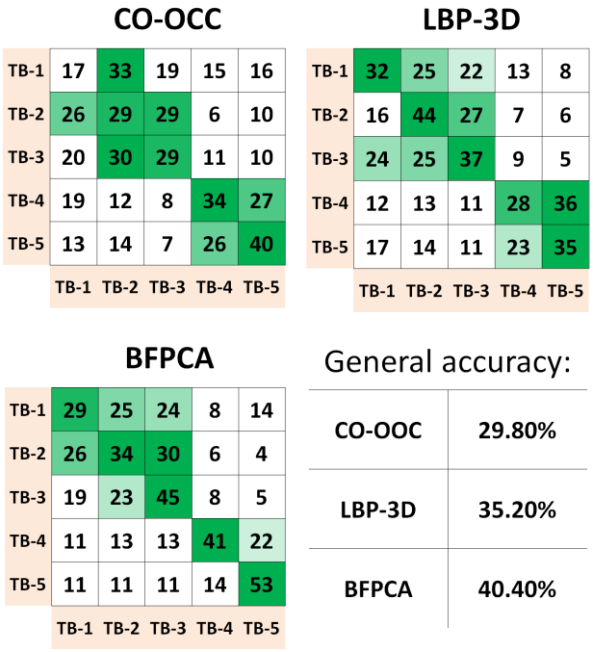

Fig. 1. Confusion matrices and the general accuracy scores for 3 different methods of classification of CT images of lung tuberculosis.

As it can be seen from Fig. 1, the BF-PCA method which employs binary descriptors produced by PCA-generated banks of filters outperforms both the 3D co-occurrence and 3D LBP methods.

Example of adaptive differential filters automatically generated with the help of PCA method based on 3D patches of CT lung images of TB patients are shown in Fig. 2.

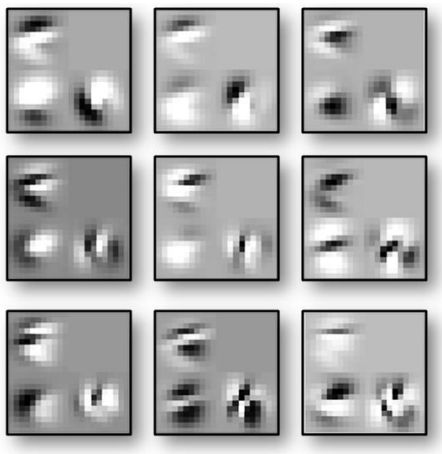

Fig. 2. Example of the first 9 leading filters automatically created by the PCA method. For each 3D filter, the 3 orthogonal sections are shown.

As it can be immediately inferred from the confusion matrices of Fig. 1, the five TB types tend cluster into two groups consisting of (TB-1, TB-2, TB-3) and (TB-4, TB-5) clusters respectively. This would allow introducing a more efficient hierarchical classification procedure on the further steps of this work.

Conclusion

The results of present study suggest that the CT image analysis method which capitalizes on binary descriptors produced by PCA-generated filters can potentially be employed as a basic tool for prospective CAD on lung tuberculosis.

Acknowledgements. This work was funded by the National Institute of Allergy and Infectious Diseases, National Institutes of Health, U.S. Department of Health and Human Services, USA through the CRDF project OISE-15-61772-1.

References

[1] Global tuberculosis report 2014. (2014) World Health Organization, WF 300, WHO Press, ISBN 978-92-4-156480-9, 170 p.

[2] http://tuberculosis.by/ Last visited 12.01.2016.

[3] http://imlab.grid.by/ Last visited 12.01.2016.

[4] Hu S, Hoffman E, and Reinhardt JM (2001). Automatic lung segmentation for accurate quantitation of volumetric X-ray CT images. IEEE Transactions on Medical Imaging, 20(6):490-498.

[5] Kovalev VA, Kruggel F, Gertz H-J, and von Cramon DY. (2001) Three-dimensional Texture Analysis of MRI Brain Datasets, IEEE Transactions on Medical Imaging, 20(5):424-433.