Satyavratan Govindarajan*1 and Manuskandan SR2 and Ramakrishnan Swaminathan

Diagnostics of Multi Drug Resistant Tuberculosis in Chest Radiographs using Local Textures & Extreme Gradient Boosting

https://doi.org/10.1515/cdbme-2023-1181

Abstract. This study attempts to detect and differentiate Multi Drug Resistant (MDR) – Tuberculosis (TB) and Drug Sensitive (DS)-TB Chest Radiographs (CXR) using local texture descriptors and Ensemble Learning method. Studies report that CXR images contain likelihood information of the drug resistance which can be utilized computationally. Initially, CXR images are subjected to lung fields segmentation using Reaction Diffusion Level Set method. Further, Local Directional Texture Pattern (LDTP) features are extracted from the segmented lungs to characterize the localized textural variations. Extreme Gradient Boosting (XGBoost) classifier is employed to differentiate DS-TB and MDR-TB images. The obtained results demonstrate the ability of extracted LDTP features to characterize non-specific textural inhomogeneities in images by operating on its principal directions. XGBoost algorithm provides maximum accuracy of 93% and true positive rate of 94.6% in detecting MDR-TB. As the proposed study differentiates the MDR-TB condition using CXR images, its computerized diagnostics could be used in the early screening and follow-up of TB ridden patients for public health infection control in any setting.

Keywords: Tuberculosis, Chest Radiograph, XGBoost, Drug Sensitive, Drug Resistive

1 Introduction

Eruptions of Tuberculosis (TB) co-infections and resistance to antimicrobial drugs has reinstated this disease as a major health problem. TB can be categorized into two forms based on its drug responsiveness, namely Drug Sensitive (DS) and Drug Resistant (DR) conditions. Multi Drug Resistant (MDR), a sub-category of DR-TB, is considered insensitive to at least one of the most effective first line drugs [1]. This condition is difficult to diagnose and costs to treat patients with success rate of only 48% [1]. The prompt diagnosis of MDR-TB is essential for public health infection control and to ensure appropriate therapy for infected patients [2]. Diagnosis of Drug susceptibility is generally performed using sputum sample and genome analysis. However, the former requires several weeks to obtain results and the latter produces indecisive results [2]. Digital chest radiography can be preferred to overcome the difficulty in the timely diagnoses of MDR-TB patients due to its non-invasive procedure, high sensitivity and portability in the screening and triaging of infectious lung diseases.

To predict the resistance status of a patient, attempts have been made to utilise Chest Radiographs (CXR) and Computed Tomography (CT) image features in combination with clinical data [3]. In these studies, abnormal findings such as maximum lung involvement, consolidation, multiple cavities are commonly reported for MDR-TB [4]. Significant differences in chest X-ray findings between the two groups in terms of lesion size and morphology have also been investigated [5]. It is still highly challenging to detect these suspicious lesions in CXRs and discriminate them with normal reflections [5]. In addition to that, these manifestations depict a combination of patterns in terms of geometry and texture. Hence, there is a necessity to develop a robust computer diagnostic system to extract reliable image features that represent significant CXR characteristics [6].

Local feature descriptors operate at dense image patches to provide regional information that characterizes subtle textural variations. These descriptors are invariant to monotonic grayscale variations and possess discriminative ability. Studies have utilized pattern-based texture descriptors in different applications such as applied in facial expression detection and medical image texture classification tasks [7,8].

Automated systems using deep learning networks have been developed for the classification of DS-TB and DR-TB

***Corresponding author: Satyavratan Govindarajan:** Indian Institute of Technology Madras, Chennai, India, e-mail: s.vratan@gmail.com
Manuskandan SR: Karuvee Innovations Pvt. Ltd., Chennai, India
Ramakrishnan Swaminathan: Indian Institute of Technology Madras, Chennai, India

using CT images [9]. Cid et al. constructed a 3-dimensional (3D) graph model based on image textures and further used Support Vector Machine (SVM) to differentiate the classes [10]. However, these systems fail to detect the sub- class MDR-TB images. Since CXRs are standard practice in healthcare settings, studies have shown statistically significant differences between drug sensitive and resistance status of TB ridden patients using computerized radiological image features [3]. Jaegar et al. [2] investigated the possibility of discriminating the two classes using Convolutional neural nets combined with a set of texture and shape features from CXRs. Studies continued to employ different deep networks such as Inception-V3 with large set of CXR images in order to achieve high performance [11]. Recently, a study reported 72-79% Area Under the Curve (AUC) in classifying DR-TB using chest X-rays alone and CXRs with text data [6]. However, there have been very few studies that differentiate MDR-TB from DS-TB CXRs. Extreme Gradient Boosting (XGBoost) classifier has shown superior performance over deep learning for real-life tabulated datasets [12] and widely used for variety of applications [13].

The objective of this work is to develop a CAD system to distinguish DS-TB and MDR-TB chest radiographic images using local directional texture pattern features and XGBoost classification. For this, the RDLS segmented lung fields are subjected to LDTP feature extraction. Significant LDTP features are computed and fed to XGBoost. The hyper parameters of XGBoost are tuned using Random Search (RS) technique. Performance of XGBoost is further compared with different machine learning algorithms. The flow diagram of the proposed methodology is shown in Figure 1. The contributions of this work aim to characterise the subtle abnormal intensity differences in DS-TB and MDR-TB CXR images by localized textural alterations using local directional texture patterns (LDTP). Implementation of RS-XGBoost to differentiate the two classes with high accuracy, precision and AUC.

2 Methods

2.1 Chest X-ray image database

The CXR images are obtained from the NIAID TB Portals repository [14]. The TB portals contain clinical and image data obtained prospectively and retrospectively from many countries. All the data from TB Portals Consortium cases have been curated and validated by expert physicians. 401 initial CXR images (199 DS-TB and 202 MDR-TB) belonging to five countries, with no comorbidity and no extrapulmonary cases are considered for analysis. The images used in this study are in grayscale format and, a fixed resolution of 1024x1024 pixels.

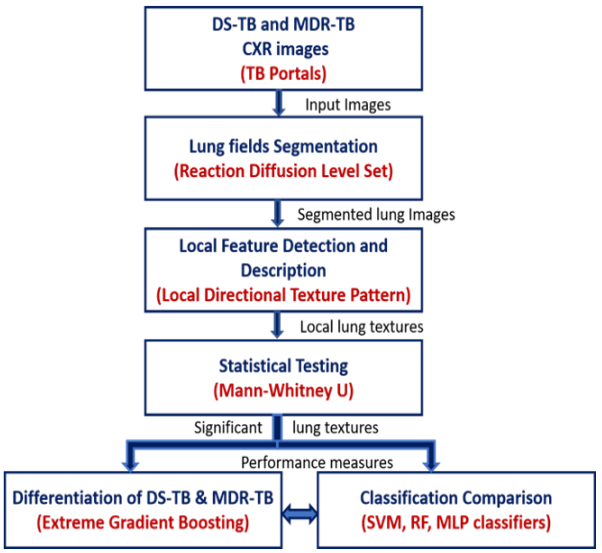

Figure 1: Pipeline of the proposed work.

2.2 Segmentation of lung fields

In chest X-ray images, segmentation of lungs is performed by means of Reaction Diffusion Level Set (RDLS) technique. The method involves dynamic motion of a curve/surface through either shrinkage or expansion. RDLS presents a rate of diffusion expression in the level set equation to attain a stable output and to avoid boundary leakage as given by [8]. Further, Anisotropic diffusion is used as an edge criterion in the level set in order to produce smooth contours. In this work, alpha parameter is fixed as 0.7 with 800 iterations in segmentation algorithm. The edge indicator with K = 30 with 15 iterations.

2.3 Extraction of LDTP features

The local directional texture pattern (LDTP) features are extracted from the mask multiplied segmented lungs to characterize DS and MDR-TB CXR images. LDTP generates a code from the combination of principal directions and the intensity differences of the two principal directions of the local image neighbourhood. This represents the structural and contrast information as directional texture numbers [15]. The principal directional numbers of the local region are found from the Kirsch compass masks in eight different directions. The difference in intensity of pixels in an image neighborhood is calculated at each of the two principal directions. In LDTP, the differences are encoded using three levels as negative, equal, and positive. Each encoded difference D_f is given in equation (1) as,

$$D_f(d) = \begin{cases} 0, & if\ \varepsilon \le d \le \varepsilon \\ 1, & if\ d < -\varepsilon \\ 2, & if\ d > \varepsilon \end{cases} \quad (1)$$

Here, d is the actual intensity difference, ε is a scalar threshold and is heuristically selected as 15. Final LDTP code is obtained by concatenating the principal directional number and the intensity differences, as given by,

$$\text{LDTP}\ (x,y) = 16P_{dir}^{1(x,y)} + 4D_f\left(d_1^{(x,y)}\right) + D_f\left(d_2^{(x,y)}\right) \quad (2)$$

Where, LDTP(x,y) represents the code for the image pixel (x,y), $P_{dir}^{1(x,y)}$ is the principal directional number of its neighbourhood and $D_f\left(d_1^{(x,y)}\right) + D_f\left(d_2^{(x,y)}\right)$ are the first two differences of its neighbourhood respectively. Therefore, the total length of LDTP code becomes 72.

Non-parametric Mann-Whitney U test with a significance level at 95% confidence interval is executed to select optimal features for better differentiation of DS-TB and MDR-TB CXRs.

2.4 Classification using XGBoost

XGBoost is a scalable and faster ensemble learning that integrates individual decision tree outputs into a joint single output. XGBoost operates with parallel tree boosting technique that increases model generalization by utilizing varied regularization techniques [13]. In XGBoost, the regularized loss function in the objective function is different than gradient boosting trees. To minimise the loss function, the subsequent learner based on the greedy algorithm is applied to compute possible tree separations on the features. In this technique, tree splitting begins from the root node and subsequently the branches are added iteratively. In this work, Random search is used to find the best hyperparameters of the XGBoost classifier model.

3 Results and discussion

Figure 2 (a) shows DS-TB and MDR-TB original chest X-rays of a representative subject. Visually, the original images possess high contrast variations locally inside lung fields. The abnormal opacities are also responsible for local intensity changes by producing intricate local textural attributes. Lung contours are also found to have structural variations. These challenges pose issues in describing lung fields as region of interest for segmentation in the detection DS and MDR-TB images. From Figure 2 (b), it is observed that the complex lung borders are accurately delineated by RDLS representing smooth and continuous contours. This method retains the normal and abnormal texture patterns in lung fields. However, the differences in DS and MDR-TB characteristics from these images are still not visually perceivable.

From Figure 2(c), it is observed that the higher-order derivatives of LDTP provided local variations with more details which illustrate the non-specific alterations due to TB abnormalities. Finer patterns of localized textural changes are extracted in DS-TB. Whereas local intensity variations in terms of coarser patterns are mostly observed of MDR-TB. 18 features are evaluated with statistical significance ($p<0.05$) between DS-TB and MDR-TB images from Mann Whitney U test.

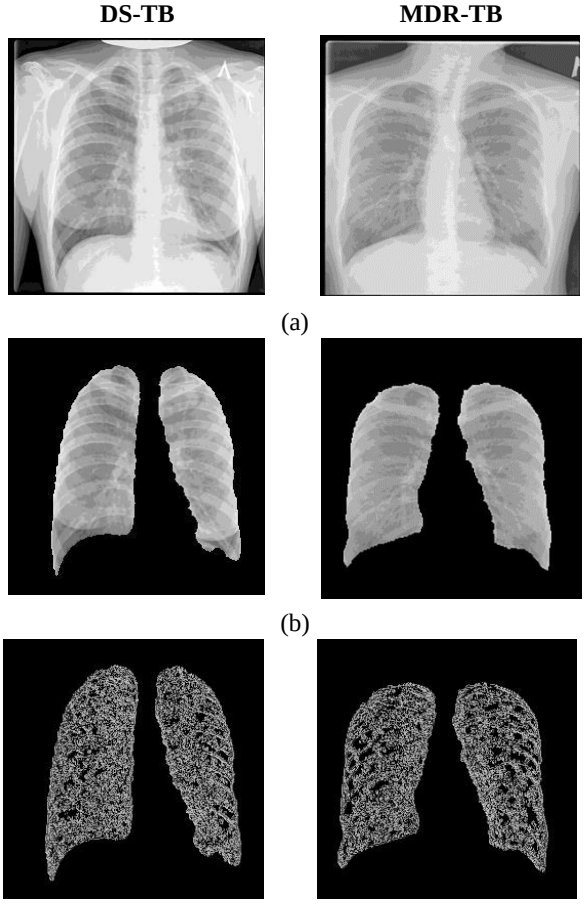

Figure 2: Representative original CXRs, (b) RDLS segmented lungs, and (c) Final LDTP features extracted

Figure 3 shows the comparison of DS and MDR-TB chest radiographic images for representative 13th and 69th LDTP significant features respectively. From Fig. 8, it is found that the mean and median values are found to be distinct for DS and MDR-TB images. Further, the interquartile range is calculated as 37 with third quartile value of 42.3 for DS-TB images, while it is 30 with third quartile value of 37.9 for MDR-TB images. This indicates that greater number of DS-TB images show inconsistent local textural variations in comparison to the other class. Although the features show distinction in their statistics, overlap in feature values are observed. Hence, machine learning algorithm to better classify

these classes is necessary.

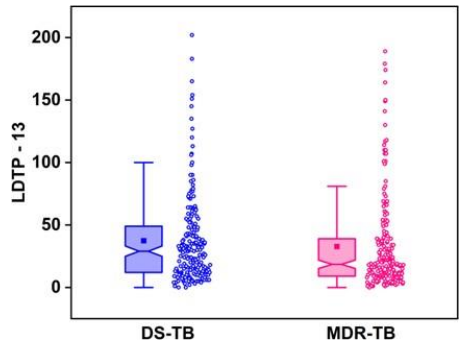

Figure 3: Comparison of DS and MDR TB CXR images using exemplar significant LDTP features

From Table 1. It is observed that RS optimized XGBoost classifier provides the maximum classification performance in the differentiation of DS-TB and MDR-TB images using significant LDTP features. Maximum accuracy of 93% with a large difference of at least 18% from the standard classifiers is obtained using RS-XGBoost. Maximum TPR and F-score values are also obtained in detecting MDR-TB subjects. MLP algorithm achieves second highest performance in correctly classifying MDR-TB. Better performance of XGBoost is attributed to its working on functional space when reducing the model cost.

Table 1: Classification performance (in %) for differentiating DS-TB and MDR-TB chest X-rays

Classifier	Accuracy	TPR	F-score	AUC
SVM	71	71.4	71.4	81
MLP	75	74.9	74.8	82
RS-XGBOOST	93	81.7	81.4	85

4 Conclusion

Early detection of multi drug-resistant cases of tuberculosis is essential for its infection control. This pilot study aims at differentiating MDR-TB images from DS-TB using a local texture descriptor and XGBoost. Results of LDTP are found to characterize the discriminative textural features of images that illustrate non-specific abnormal alterations through its measurable intensity differences. The obtained results of classification signify the efficacy of classifiers by yielding better model accuracies. RS-XGBoost with its scalable feature learning approach is able to provide superior differentiability of classes with maximum accuracy and precision. Hence, the proposed system could be used to assist physicians in the automated diagnostics of MDR-TB conditions during patient's first visit to the clinical setting.

Author Statement

Research funding: The author state no funding is involved.
Conflict of interest: Authors state no conflict of interest.

References

[1] Jang JG, Chung JH (2020) Diagnosis and treatment of multidrug-resistant tuberculosis. Yeungnam Univ J Med 2020; 37(4):277-285.
[2] Jaeger S, Juarez-Espinosa OH, Candemir S, et al. Detecting drug-resistant tuberculosis in chest radiographs. Int J Comput Assist Radiol Surg 2018;13(12):1915-25.
[3] Kovalev V, Liauchuk V, Kalinovsky A, et al. Utilizing radiological images for predicting drug resistance of lung tuberculosis. Int J Comput Assist Radiol Surg 2015;10:1-312.
[4] Wáng YX, Chung MJ, Skrahin A, et al. Radiological signs associated with pulmonary multi-drug resistant tuberculosis: an analysis of published evidences. Quant Imag Med Surg. 2018;8(2):161-173.
[5] Icksan AG, Napitupulu MR, Nawas MA, Nurwidya F. Chest X-ray findings comparison between multi-drug-resistant tuberculosis and drug-sensitive tuberculosis. J Nat Sci Biol Med. 2015;9(1):42-46.
[6] Karki M, Kantipudi K, Yang F, et al. Generalization Challenges in Drug-Resistant Tuberculosis Detection from Chest X-rays. Diagnostics. 2022;12(1):188.
[7] Turan C, Lam KM Histogram-based local descriptors for facial expression recognition (FER): A comprehensive study. J Vis Comm Image Rep 2018;55:331-41.
[8] Govindarajan S, Swaminathan R Extreme learning machine based differentiation of pulmonary tuberculosis in chest radiographs using integrated local feature descriptors. Comput Methods Programs Biomed. 2021;204:106058.
[9] Allaouzi I, Ahmed MB. A 3D-CNN and SVM for Multi-Drug Resistance Detection. InCLEF (Working Notes). 2018.
[10] Cid YD, Müller H. Texture-based Graph Model of the Lungs for Drug Resistance Detection, Tuberculosis Type Classification, and Severity Scoring: In CLEF (Working Notes) 2018.
[11] Karki M, Kantipudi K, Yu H, et al. Identifying Drug-Resistant Tuberculosis in Chest Radiographs: Evaluation of CNN Architectures and Training Strategies. In 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society 2021; 2964-2967.
[12] Shwartz-Ziv R, Armon A. Tabular data: Deep learning is not all you need. Information Fusion. 2022;81:84-90.
[13] Nasiri H, Hasani S. Automated detection of COVID-19 cases from chest X-ray images using deep neural network and XGBoost. Radiography. 2022;28(3):732-738.
[14] Rosenthal A, Gabrielian A, Engle E, et al. The TB portals: an open-access, web-based platform for global drug-resistant-tuberculosis data sharing and analysis. J Clin Microbiol. 2017;55(11):3267-82.
[15] Rivera AR, Castillo JR, Chae O (2015) Local directional texture pattern image descriptor. Pattern Recognit Lett. 51:94-100.