Identifying drug-resistant tuberculosis from chest X-ray images using a simple convolutional neural network

J Ureta¹²* and A Shrestha¹²

¹Bioinformatics Lab, Advanced Research Institute for Informatics, Computing, and Networking (AdRIC), De La Salle University-Manila, Philippines.
²Software Technology Department, College of Computer Studies, De La Salle University-Manila, Philippines.

*jennifer.ureta@dlsu.edu.ph

Abstract. Tuberculosis (TB) is one of the top 10 causes of death worldwide, and drug-resistant TB is a major public health concern especially in resource-constrained countries. In such countries, molecular diagnosis of drug-resistant TB remains a challenge; and imaging tools such as X-rays, which are cheaply and widely available, can be a valuable supplemental resource for early detection and screening. This study uses a specialized convolutional neural network to perform binary classification of chest X-ray images to classify drug-resistant and drug-sensitive TB. The models were trained and validated using the TBPortals dataset which contains 2,973 labeled X-ray images from TB patients. The classifiers were able to identify the presence or absence of drug-resistant Tuberculosis with an AUROC between 0.66–0.67, which is an improvement over previous attempts using deep learning networks.

1. Introduction

Due to technological and medical advancements, tuberculosis (TB) mortality rate is decreasing [1]. However, despite the continuous improvements in disease control, TB still remains a major public health threat — it is one of the top 10 causes of death worldwide. In 2019, an estimated 8.9–11.0 million people were diagnosed with TB, and 1.1–1.2 million people died due to the disease [2]. TB burden is high among low- and middle-income countries — the Philippines, for instance, has the third highest prevalence rate in the world, with about 1 million Filipinos having active TB, and this number continues to grow every year [3, 4, 5].

Drug-resistant TB (DR-TB) is a major public health concern, especially in resource-constrained countries, since its treatment is difficult and takes more time and money [2]. In 2019, about 3.3% of new TB cases and 18% of recurring cases were multi-drug-resistant TB (MDR-TB), and the rates of effective treatment for such cases were significantly lower [2]. MDR-TB accounts for a third of all antimicrobial resistance deaths globally [6].

Diagnosis of DR-TB remains a challenge [7]. Culture-based phenotypic drug-sensitive tests are considered the gold standard, but they require specific laboratory facilities and can take several weeks. Molecular assays that are based on detecting specific drug-resistant mutations have shown varying degrees of success (see e.g. [8] for a review), but they are expensive and
require well-equipped facilities and trained technicians which are often hard to access. Whole-genome sequencing and associated bioinformatics tools [9, 10, 11, 12], which look at the entire mutational landscape, provide a promising new direction in diagnosis and characterization of MDR-TB. However, these technologies remain out of reach for low- and middle-income countries with high TB burden.

On the other hand, conventional chest X-ray (CXR) is widely available, and has the potential to be a valuable tool for early detection and screening of DR-TB. Previous studies have suggested the presence of CXR image features that may be useful in distinguishing DR-TB from drug-sensitive TB. In a retrospective cross-sectional study, Icksan et al. compared CXR findings between drug-resistant and drug-sensitive TB and concluded that there were significant differences in the size and morphology of lesions [13]. Based on a literature survey, Wang et al. reported that common radiological signs associated with MDR-TB include “centrilobular small nodules, branching linear and nodular opacities (tree-in-bud sign), patchy or lobular areas of consolidation, cavitation, and bronchiectasis” [14], and concluded that the prevalence of thick-walled cavity lesions might be a promising feature for differential diagnosis of MDR-TB. Cha et al. [15] and Kim et al. [16] have similarly suggested the presence of radiological signatures in chest radiograph and CT images that are informative in differentiating drug-resistant and drug-sensitive TB.

Computational image processing techniques can be used to aid human interpretation and to discover novel features that might be key in differentiating drug-resistant from drug-sensitive TB. While deep convolutional neural networks have been shown to have high accuracy in diagnosing TB from CXR images [17, 18, 19], there has been limited work on applying such models to the more challenging task of classifying drug-sensitive and drug-resistant TB. Jaeger et al. employed several classifiers, including VGG-v16 pretrained on natural images and a customized convolutional neural network, to classify drug-resistant and drug-sensitive TB using CXR images of patients in Belarus. They reported that both the deep learning models did not perform well and attributed this to their small training set.

TB CXR image datasets have grown ever since. A prominent example is TB Portals, which is a web-based, open access repository of multi-domain TB data which includes linked socioeconomic, clinical, radiological, and genomic data [20]. There has also been progress in customizing neural network architectures for TB screening rather than adapting models from natural image classification tasks which have a large number of parameters and are overly data-hungry. Pasa et al. [21] describe a small and efficient but accurate neural network optimized for the task of TB diagnosis.

Here, we employ the convolutional neural network architecture proposed by Pasa et al. to classify DR-TB vs. drug-sensitive TB using CXR images from the radiological dataset of TB portals. We report an AUROC of 0.66 on a test set consisting of images prior to the start of treatment, and an AUROC of 0.67 in a set including follow-up images. This result is an improvement over previous attempts by Jaeger et al. [1] to use deep learning models to detect drug-sensitive TB.

2. Methodology
2.1. Data
We used the data of TB Portals[20] for this study which was extracted on July 2020. It includes 3,051 cases classified into five classes of drug resistance: multidrug resistance(MDR), extensive drug resistance(XDR), monoresistance, polydrug resistance, and sensitive. The resistance classes are defined as follows: “monoresistance is resistance to one first-line anti-TB drug only, polydrug resistance is resistance to more than one first-line anti-TB drug (other than both isoniazid and rifampin), multidrug resistance is resistance to at least both isoniazid and rifampin, and extensive drug resistance is resistance to any fluoroquinolone and to at least one of 3 second-line injectable drugs (capreomycin, kanamycin, and amikacin) in addition to multidrug resistance” [20].
Figure 1. This shows the percentages of cases for each drug resistance type among all the cases in the TB Portal (a) and for each country (b).

We note that the TB Portals contains data that was heavily selected for MDR-TB, with almost half of them (47%) being MDR, which does not reflect the actual prevalence of anti-microbial resistant TB in the TB patient population. The higher proportion of DR-TB cases in the dataset is because TB portals has collected data from the most virulent, drug-resistant and deadly cases in order to distinguish them from available reference strains[20]. Figure 1a shows the percentage of cases in the TB Portals data for each type of drug resistance, and Figure 1b shows the cases per country and resistance type.

For this work, we used the CXR dataset of TB Portals. This radiological dataset contains 2,973 labeled CXR images. We combined all cases that belong to the resistant classes (MDR, Monoresistant, PDR and XDR) into a single class. We considered 2 datasets: a before-treatment dataset containing only those images that were taken before treatment was started (899 images with 240 drug-sensitive cases and 659 drug-resistant cases), and the full dataset (2,973 images with 636 drug-sensitive cases and 2,337 drug-resistant cases). Both datasets were randomly partitioned into a training set with 70% of the images and a testing set with 30%.

2.2. Preprocessing
We used U-Net [22] to segment the lungs in the CXR images. We trained the network using the publicly available Montgomery dataset [23] which is maintained by the National Institutes of Health (NIH). It contains 138 CXR images, in which 80 are normal while 58 have manifestations of TB. Figure 2a shows a sample image from the Montgomery dataset, and Figures 2b and 2c show the left and right mask for this image.

After segmentation, the images were further preprocessed using the same steps used by Pasa et al. [21]. Black bands or borders were cropped from the edges of the images, images were resized and the central 512 x 512 region was extracted. Lastly, the mean pixel value, which was calculated for all images in the dataset, was subtracted from each pixel and were divided by their standard deviation [21].

2.3. Model Architecture and Training
We used the convolutional neural network proposed by Pasa et al. [21] which is tailored for diagnosing TB. Figure 1 of the paper of Pasa et al. [21] shows the architecture of the network. The network is composed of 5 convolutional blocks, with each block containing two 3x3 convolutional layers with ReLU as the class activation function followed by a max pooling.
layer. Each block also has a 1x1 convolution layer that acts as a shortcut connection. The convolutional blocks are then followed by a global average pooling layer and a fully-connected softmax layer. While there are several other more complex deep learning models, most of them have been designed to be trained on large amounts of data making them unsuitable for our task which has a limited amount of data.

The model was implemented using Tensorflow. The Adam optimizer was used to train the models with the following parameters: $\beta_1 = 0.9$ and $\beta_2 = 0.999$ which are the exponential decay rates for the moment estimates, $E = 1 \times 10^{-8}$ which is a small constant for numerical stability, a learning rate of $8 \times 10^{-5}$, and a batch size of 8. Categorical cross-entropy was used as the error function.

Five-fold cross validation was performed using the training set in order to determine and validate the hyperparameters to be used for the final training. Afterwards, we trained the model on the full training set and evaluated it on the testing set. We used early-stopping to determine the number of epochs to avoid overfitting. Table 1 shows the number of epochs used for the final training on the entire training set. The models were trained on a CPU with an RTX 2070 GPU.

### 3. Results and Discussion

Overall the U-NET segmentation performed well on the TB Portals dataset even if it was trained on a different dataset. Figures 3a and 3b show a sample result of the segmentation process when applied to the TB Portals dataset. However there were a few failed segmentations that were produced as can be seen in the example Figure 3c and 3d.

The results of the 5-fold cross validation for the before-treatment dataset and the full dataset are shown in Table 2. After re-training the models on the full training set, we applied them to the test sets. The ROC curves and the AUC scores are shown in Figures 4a and 4b.

Previously, Jaeger et al. [1] employed a customized convolutional neural network and the VGG-v16 network for the same classification task and obtained AUC of 0.56 and 0.52, respectively. While the performance of our classifier is modest, it is a significant improvement over these previous results. We posit that the improvement is a consequence of the network
Figure 3. (a) and (b) are sample images from TB Portals (b) shows a good segmentation of (a) while (c) shows a failed segmentation of (d).

| Dataset                      | Fold 1 | Fold 2 | Fold 3 | Fold 4 | Fold 5 |
|------------------------------|--------|--------|--------|--------|--------|
| Before-treatment dataset     | 0.721  | 0.657  | 0.734  | 0.705  | 0.721  |
| Full dataset                 | 0.708  | 0.657  | 0.713  | 0.689  | 0.701  |

Table 2. AUC scores for the 5-fold cross validation

Figure 4. ROC curve of the before-treatment dataset (a) and ROC curve of the full dataset (b) where the positive class is the drug-resistant class.

architecture we used, which is customized for TB CXR image analysis, or of the larger dataset that we trained our model on, only part of which was available earlier.

Interestingly, the results of using just the before-treatment dataset was comparable to using the full dataset despite the significantly fewer number of images that were used for training the model. That the inclusion of follow-up images of the patients do not significantly improve classification accuracy was also observed by Jaegar et al [1]. Since the before-treatment images likely correspond to drug-resistance due to transmission as opposed to acquired resistance due to drug-selective pressures, our results suggest the model is learning mostly about transmitted resistance. It is estimated that in some countries, the percentage of MDR-TB resulting from transmission is higher than 90% [24]. Given that a few countries are over-represented in the TB Portals database (see Figure 1b), it is likely that the database is biased towards primary transmission cases. On the other hand, our results suggest that CXR images taken prior to treatment contain information about likely drug-resistance and thus might have potential as a cheap early screening tool for drug-resistance and to guide drug regimens.
To partly address the issue of data imbalance which we mentioned in Section 2.1, we evaluated the precision of our model accounting for the actual prevalence of the drug-resistant class which one might be encounter in practical settings, instead of the one that our model is trained and tested on. Let $\eta \in [0, 1]$ denote the real-world positive class prevalence. Brabec et al. [25] show that the empirical precision of a binary classification model can be expressed as:

$$\text{Prec}(\eta) = \frac{\text{TPR} \cdot \eta}{\text{TPR} \cdot \eta + \text{FPR} \cdot (1-\eta)}$$

where $\text{TPR}$ and $\text{FPR}$ are the empirical True Positive Rate and False Positive Rate, respectively. We used this relationship to plot the Prevalence-Precision curves of Figures 5a and 5b. We plotted the curves for the different operating points of our classifier with TPR values ranging from 40% up to 80%. For the positive class prevalence ($\eta$) we used a range of values from 3.4% to 18% based on the prevalence of multi-drug resistance TB in 2018 [26]. We can see from Figures 5a and 5b that the precision value is quite low for any prevalence value and operating point of the model. Further improvements are needed for practical adoption of the model.

A major limitation of our current approach is the interpretability of the classification results. Since the goal is to assist medical experts in identifying signs of drug resistance, it would be helpful if the areas and features indicative of drug-resistance could be identified. One way to do this is through the use of class activation maps (CAM) [27] which generate heat maps which can be used to identify image regions used by the CNN for class discrimination.

4. Conclusion

While molecular diagnosis of drug-resistant TB remains hard to access for resource-constrained countries, clinical imaging, supplemented by computer-assisted detection methods, is a promising tool for early detection and screening. While we showed significant performance improvement compared to previous work in applying deep learning models to classify drug-resistant and drug-sensitive TB from chest X-rays, further improvements are required for practical adoption. For future work it would be interesting to address the issue of imbalanced dataset at the training phase by using sampling and augmentation techniques. Adding visualization techniques and consulting with radiologists may also provide more insights and help validate the performance of the models.
References

[1] Jaeger S, Juarez-Espinosa O H, Candemir S, Poostchi M, Yang F, Kim L, Ding M, Folio L R, Antani S, Gabrielian A, Hurt D, Rosenthal A, and Thoma G 2018 Detecting drug-resistant tuberculosis in chest radiographs. *Int J Comput Assist Radiol Surg* 13 1915–1925.

[2] WHO 2020 Global tuberculosis report 2020 URL https://www.who.int/tb/publications/global report/en/

[3] WHO 2018 What is multidrug-resistant tuberculosis (mdr-tb) and how do we control it? URL: https://www.who.int/news-room/q-a-detail/what-is-multidrug-resistant-tuberculosis-(mdr-tb)-and-how-do-we-control-it.

[4] Snow K, Yadav R, Denholm J, Sawyer S, and Graham S 2018 Tuberculosis among children, adolescents and young adults in the philippines: a surveillance report URL https://ojs.wpro.who.int/ojs/index.php/wpsar/article/view/586/855.

[5] Weiler A 2019 It’s time to end tb in the philippines URL: https://www.who.int/philippines/news/comments/detail/it-s-time-to-end-tb-in-the-philippines.

[6] Stop tb http://www.stoptb.org/news/stories/2017/11 014.asp accessed: 2021-04-30.

[7] Migliori G B, Tiberi S, Zumla A, Petersen E, Chakaya J M, Wejse C, Torrico M M, Duarte R, Aflifenaar J W, Schaaf H S, Marais B J, Cirillo D M, Alagna R, Rendon A, Pontal E, Piubello A, Figueroa J, Ferlazzo G, Garc’a-Basteiro A, Centis R, Visca D, D’Ambrosio L, Sotgiu G, Arkub T A, Akkerman O, Aleksa A, Belilovski E, Bernal E, Blanc F X, Boeree M, Borisov S, Bruchfeld J, Loidi J C, Caminero J, Carvalho A, Gallardo J C, Charalampos, Danila E, Forsman L D, Denholm J, Dheda K, Diel R, Diktanas S, Dobler C, Enwerem M, Esposito S, Salinas N E, Filippov A, Formenti B, Garc’a J G, Goletti D, Rosso R G, Gualano G, Isaakidis P, Kaluzhenina A, Koiral A, Kuksa S, Kunst H, Li Y, Magis-Escurre C, Manfrin V, Manga S, Manika K, Marchese V, Robles E M, Maryandyshev A, Matteelli A, Mariani A, Mazza-Stalder J, Mello F, Mendoza L, Mesia A, Miliauskas S, Hamdan H M, Ndjeka N, Marcos M N, Ottenhoff T, Palmero D, Palmieri F, Papavasileiou A, Payen M, Pontarelli A, Dalcolmo M P, Fernandez S Q, Romero R, Silva D R, Santos A, Seaworth B, Sinitsyn M, Skrahina A, Solovic I, Spayneello A, Talodini M, Torres C, Udwaadia Z, van den Boom M, Volchenkov G, Yedilbayev A, Zaleskis R, and Zellweger J 2020 MDR/XDR-TB management of patients and contacts: Challenges facing the new decade. The 2020 clinical update by the Global Tuberculosis Network. *International Journal of Infectious Diseases* 92 S15–S25.

[8] Nguyen T N A, Berre V A L, Ban’uls A L, and Nguyen T V A 2019 Molecular Diagnosis of Drug-Resistant Tuberculosis: A Literature Review. *Frontiers in Microbiology* 10 794.

[9] Chen M L, Doddi A, Royer J, Freschi L, Schito M, Ezewudo M, Kohane I S, Beam A, and Farhat M 2019 Beyond multidrug resistance: Leveraging rare variants with machine and statistical learning models in Mycobacterium tuberculosis resistance prediction. *EBioMedicine* 43 356 – 369.

[10] Sergeev R S, Kavaliou I S, Satanesuski U V, Gabrielian A, Rosenthal A, Tartakovsky M, and Tuzikov A V 2019 Genome-Wide Analysis of MDR and XDR Tuberculosis from Belarus: Machine-Learning Approach. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* 16 1398–1408.

[11] Hunt M, Bradley P, Lapierre S G, Heys S, Thomsit M, Hall M B, Malone K M, Wintringer P, Walker T M, Cirillo D M, Comas I, Farhat M R, Fowler P, Gardy J, Ismail N, Kohl T A, Mathys V, Merker M, Niemann S, Omar S V, Sintchenko V, Smith G, van Soolingen D, Supply P, Taseen S, Wilcox M, Arandjelovic I, Peto T E A, Crook D W, and Iqbal Z 2019 Antibiotic resistance prediction for Mycobacterium tuberculosis from genome sequence data with Mykrobe. *Wellcome Open Research* 4 191 ISSN 2398-502X.

[12] Phelan J E, O’Sullivan D M, Machado D, Ramos J, Oppong Y E A, Campino S, O’Grady J, McNerney R, Hibberd M L, Viveiros M, Huggett J F, and Clark T G 2019 Integrating
informatics tools and portable sequencing technology for rapid detection of resistance to anti-tuberculous drugs. *Genome Medicine* **11**.

[13] Ikcsan A, Napitupulu M, Nawas M, and Nurvidiya F 2018 Chest X-ray Findings Comparison between Multi-Drug-Resistant Tuberculosis and Drug-Sensitive Tuberculosis. *The Journal of Natural Science, Biology and Medicine* **9**.

[14] Wang Y X, Chung M, Skrahin A, Rosenthal A, Gabrielian A, and Tartakovsky M 2018 Radiological signs associated with pulmonary multi-drug resistant tuberculosis: an analysis of published evidences. *Quantitative Imaging in Medicine and Surgery* **8** 161–173.

[15] Cha J, Lee H Y, Lee K S, Koh W J, Kwon O J, Yi C A, Kim T S, and Chung M J 2009 Radiological Findings of Extensively Drug-Resistant Pulmonary Tuberculosis in Non-AIDS Adults: Comparisons with Findings of Multidrug-Resistant and Drug-Sensitive Tuberculosis. *Korean Journal of Radiology* **10** 207-216.

[16] Kim H C, Goo J M, Lee H J, Park S H, Park C M, Kim T J, and Im J G 2004 Multidrug-resistant tuberculosis versus drug-sensitive tuberculosis in human immunodeficiency virus-negative patients: computed tomography features. *Journal of Computer Assisted Tomography* **28** 366–371.

[17] Murphy K, Habib S S, Zaidi S M A, Khowaja S, Khan A, Melendez E T, Amad F, Schalekamp S, Verhagen M, Philipson R H H M, Meijers A, and van Ginneken B 2020 Scientific Reports **10**.

[18] Nash M, Kadavigere R, Andrade J, Sukumar C A, Chawla K, Shenoy V P, Pande T, Huddart S, Pai M, and Saravu K 2020 Scientific Reports **10**.

[19] Pande T, Cohen C, Pai M, and Ahmad Khan F 2016 Computer-aided detection of pulmonary tuberculosis on digital chest radiographs: a systematic review. *Int J Tuberc Lung Dis* **20** 1226–1230.

[20] Rosenthal A, Gabrielian A, Engle E, Hurt D E, Alexandru S, Crudu V, Sergueev E, Kirichenko V, Lapitskii V, Snezhko E, Kovalev A, Astrovko A, Skrahina A, Taaffe J, Harris M, Long A, Wollenberg K, Akhundova I, Ismayilova S, Skrahin A, Mammadabayov E, Gadirova H, Abuzarov R, Seyfaddinova M, Avaliani Z, Strambu I, Zaharia D, Muntean A, Ghita E, Bogdan M, Mindru R, Spiniu V, Sora A, Ene C, Vashakidze S, Shubladze N, Nanava U, Tuzikov A, and Tartakovsky M 2017 The TB Portals: an Open-Access, Web-Based Platform for Global Drug-Resistant-Tuberculosis Data Sharing and Analysis. *Journal of Clinical Microbiology* **55** 3267–3282.

[21] Pasa F, Golkov V, Pfeiffer F, Cremers D, and Pfeiffer D 2019 Scientific Reports **9**.

[22] Ronneberger O, Fischer P, and Brox T 2015 U-Net: Convolutional Networks for Biomedical Image Segmentation. *International Conference on Medical Image Computing and Computer-Assisted Intervention* 234–241.

[23] Jaeger S, Candemir S, Antani S, Wang Y X J, Lu P X, and Thoma G 2014 Two public chest X-ray datasets for computer-aided screening of pulmonary diseases. *Quantitative imaging in medicine and surgery* **4** 475–477.

[24] Kendall E A, Fofana M O, and Dowdy D W 2015 Burden of transmitted multidrug resistance in epidemics of tuberculosis: a transmission modelling analysis. *The Lancet Respiratory Medicine* **3** 963–972.

[25] Brabec J, Komárek T, Frane V, and Machlica L 2020 On Model Evaluation under Non-constant Class Imbalance. *Computational Science – ICCS 2020 (Cham: Springer International Publishing)* 74–87.

[26] Intimicrobial resistance URL: https://www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance.

[27] Zhou B, Khosla A, Lapedriza A, Oliva A, and Torralba A 2016 2016 IEEE Learning Deep Features for Discriminative Localization. *Conference on Computer Vision and Pattern Recognition (CVPR)* 2921–2929.