Deep Learning Based Automatic Tuberculosis Screening System

Chirath Dasanayaka[1], Maheshi B. Dissanayake[1]

[2]Department Electrical and Electronic Engineering, Faculty of Engineering, University of Peradeniya, Peradeniya 20400, Sri Lanka
chirath.dasanayaka@pdn.ac.lk, maheshi@ee.pdn.ac.lk

Given that, Tuberculosis (TB) is completely curable when diagnosed at an early stage, TB screening plays an important role in preventive care. According to the World Health Organization TB is on the rise worldwide, and the largest cause behind this increase is inadequate control in developing nations. Even though several laboratory based testing methods are available for the diagnosis of TB, they are time consuming, expensive and very limited in rural areas. As a solution to this problem, we present an accurate automated TB screening system using frontal chest X-rays, which can be utilized with negligible human interaction. The two datasets of chest radiography that were used in this research are available in US National Library of Medicine, and they were acquired from the Department of Health and Human Service, Montgomery Country, Maryland, USA, and Shenzhen No 3 Peoples Hospital in China. Our proposed method consists of two sub models, namely: lung segmentation and classification sub-systems. In order to exclude the area that are not pertinent to lungs, we used UNET convolutional neural network (CNN) for segmentation model and for feature extraction and classification we have used ensemble of two deep convolutional neural networks (DCNN), VGG16 and InceptionV3. By using image preprocessing techniques such as contrast limited adaptive histogram equalization (CLAHE), median filtering based noise removal with heavy data augmentation, hyper-parameter tuning, and model ensembling, we were able to achieve classification accuracy of 95% with Youden's index of 0.906, sensitivity of 97.1% and specificity of 93.5%, which is a considerable improvement compared to the existing work in the literature. The tedious task of hyper parameter tuning was done by using a genetic algorithm. Due to the hyper-parameter tuning to match the task in hand, we were able to reduce the potential of over-fitting. Overall, the model presented could be quite helpful for the countries with low human experts in this field of medicine. Furthermore, since the solution presented is inexpensive, easily accessible, highly accurate, and can be used to screen a large population instantly, it can be adopted in low income countries.

Key words: Tuberculosis, deep learning, lung segmentation, transfer learning, CAM visualization.