Chest X-ray Image Classification for Tuberculosis using Deep Convolutional Neural Network

Aditya Nair[1], Akshay Mohan[2], Jith James Fernandez[3], Sabin Biju Seema[4], Sreena V G[5]
[1]*Pursuing eJPT*
[2]*Amritha University, Coimbatore*
[3]*Advance Data Analytics associate, Guidehouse*
[4]*Associate Software Engineer, Ospyn Technologies*
[5]*Marian Engineering College, sreenavg*

Abstract **- Lung infections are severe health conditions that seriously affect the life of human, especially those like Tuberculosis (TB) which accounts for most of deaths yearly. Therefore, proper diagnosis of TB is very much essential. Chest X-Rays (CXR) are mostly been used by medical industry for detecting tuberculosis as well as other lung diseases. Deep learning based disease identification systems are in use nowadays, one among such a deep learning architecture is Convolutional Neural Network (CNN). In this work, a prediction model using Deep Convolutional Neural Networks is proposed for detecting TB from input Chest X-ray image. Evaluation of model performance was done using confusion matrix and accuracy as metrics. The developed custom network is paired with adam optimizer, which was able to precisely differentiate between normal and TB infected CXRs. Finally an interactive web application was built to deploy the model so that it can be used to make predictions on Chest X-rays as input.**

Index Terms **– Tuberculosis classification, Chest X-Ray, Convolutional Neural network, Deep learning.**

I.INTRODUCTION

Tuberculosis (TB) is caused by Mycobacterium tuberculosis that affect the lungs mostly. It is spread from one person to another through air.A person affected with lung TB spreads the disease through coughing, spiting or sneezing. TB disease is definitely curable but timely detection is an important factor for appropriate treatment. In order to detect Tuberculosis, Chest X-Ray (CXR) has been used mostly due to its convenient efficiency as well as cost effectiveness. Along with CXR, skilled radiologists are also a necessity in order to analyse and detect anomalies from the CXR report and to provide timely treatment. Lack of medical experts is a problem faced in medical field for timely diagnosis of deadliest diseases.

To solve this issue, Deep Convolutional Neural Network (CNN) model can be used which after collecting a chest X-ray image as input can predict the disease with a very high accuracy, that aids the role of a skilled radiologist. This will be extremely useful in developing Countries where TB rate is increasing at an alarming rate along with the lack of skilled radiologists. Deep Learning has advanced drastically and with it CNN has also consistently surpassed other recognition algorithms in image-based classification to a point where it even exceeds that of humans especially in the area of computer vision. As a result, Artificial Intelligence and Deep Learning models are being widely used in the Medical field to provide better care.

There has been numerous works done in this specific feild and still the researchers are working on these type of CNN based disease classifications. In this work we used a model developed from scratch using simple CNN layers that could predict results in a better way with comparable accuracy. We also took it up a step further to actually deploy the system through various platforms using our website and android application that greatly increases the usability as well as the reachability.

The paper is organised as: Section II explains related works based on the literature review conducted before implementing the proposed system. Section III and IV discuss the detailed methodology. Section V depicts the experimental result and Section VI concludes the paper.

II. RELATED WORKS

One of the first research papers to use in-depth study techniques in medical imaging is [1]. The work details a method which utilised Local Binary Pattern along with a Gaussian Laplacian to detect presence of tuberculosis in Chest X-Ray images. The ConVnet model associating the differentiation of different TB manifestations was described in [2]. During training the model, this work focuses on uneven, slightly fragmented X-ray scanning and includes confirmation of sample shuffling. CNN has also been used [3] to exclude discriminatory and independent features on X-rays in order to differentiate various body parts. This study demonstrated the efficiency of CNN models that can surpass conventional architecture. A method based on in-depth study of the classification of CXR into normal and diseased categories is shown in [4]. There are also other methods like Support Vector Machine [5][6], Bayesian Classifier [8], K-Nearest Neighbour [7], Linear Discriminant Analysis [9], that has been utilized for TB detection and classification.

A computer-assisted diagnostic program is introduced by [11] to diagnose TB using Alexnet and VGGNet to classify a given chest X-ray as either normal or TB affected categories. This study which was based on data sets from Montgomery and Shenzhen CXR images, concluded that VGGNet defeated Alexnet due to VGGNet's deep network. The accuracy was 80.4% and 81.6% respectively for Alexnet and VGGNet. Also they added that the accuracy can further be improved as the size of data used for testing increases.

Recent Algorithms for classifying Chest X-ray images for TB identification use transfer learning implemented on pre-trained models like AlexNet, Mobilenet, ResNet and GoogLeNet. Such models were designed, trained and tested with about one million images from Imagenet database about 1000 classes. As a result, such models require more memory, resources and processing time. Such systems may be affected from over fitting and less efficiency in producing good classification results when used for medical image classifiation with the limited number of images[12]. While in [13], the hybrid approach of applying some modifications on existing DenseNet-121 was proposed . This hybrid method was proved to be better over the existing methods for classification of 14 chest diseases as it could produce system accuracy of 84.62% in comparison to the other model with 80.97% accuracy.

While CNN achieves excellent performance on large data sets, if the dataset is of small size it may not be sufficient enough to a CNN model from bottom. Transfer learning is used in such situations. The existing pre-trained networks that are trained using a huge standard set of images, can then be used to decode specific tasks. Examples of such previously trained neural networks are VGGNet, Mobilenet, ResNet, Inception are some of the pre trained networks available. The use of CNN on chest R-ray image classification was done at [14][15], as the authors feel that the available images are closer to real-life images. The efficiency of CNN based models to classify chest infections has been demonstrated in discussed works. Here we propose a new way to expand the deeper learning structures of CNN for the purposes of the separation of medical images. This can play a major role as a computer-assisted clinical diagnostic tool on the near future.

It is pretty clear from the accompanying review that greater effort is needed to address the ongoing Tuberculosis epidemic as it is one of the most deadliest disease. For this reason, we propose an improved system that can provide better validation accuracy in classifying CXR for TB manifestation.

III. MATERIALS AND METHODS

A. Dataset

The dataset we used is a collection of Normal and Tuberculosis X-rays. A group of researchers from University of Dhaka and Qatar University associated with medical experts from Bangladesh and Hamad Medical Corporation created chest X-ray image database for Tuberculosis (TB) manifestations together with Normal images[20]. The sample dataset is shown in Fig. 1

This comprises of CXR images of patients with TB (3500) and healthy people (3500). This TB database is a collection of publicly accessible datasets such as:

1) National Library of Medicine(NLM) dataset: NLM[21] provides two publicly available Chest X-ray datasets : Shenzhen and Montgomery.

2) Belarus dataset: National Institute of Allergy and Infectious Diseases(NIAID), Ministry of

Health, Republic of Belarus formed the dataset [22] as part of a drug resistance study conducted.

3) NIAID TB dataset: This dataset [23], contains about
3000 TB affected CXR images.

4) RSNA CXR dataset: As part of RSNA pneumonia detection challenge [24], 30,000 CXRs were provided
of which 10000 are normal and remaining with lung disorders.

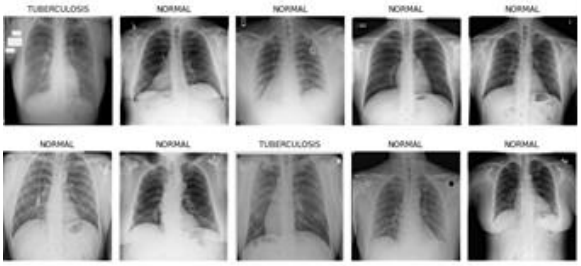

Fig. 1: Sample images from Train Set.

B. Preprocessing

Deep neural networks rely heavily on sizeable data sets to achieve high accuracy and avoid over fitting issues. Thus preprocessing on available database is necessary as a means of increasing number of samples. Total samples in the Chest X-Ray training dataset is increased by applying Image Augmentation technique. The existing training set images are modified which increases dataset variation and ultimately enhance the network model's ability to predict the category of test images.

The following types of expansions were used in the proposed model: left-to-right horizontal flip, random zoom range, shear range. Some of the performance tasks used here include resizing the image, removing the volume and measuring the histogram. These image augmentation have been done using the tensorflow.keras.preprocessing.image library. The process of adding data used in this activity is not the one that provides space for data retrieval. The batch size is 16. Batch size refers to the training examples employed in one iteration. The pixel resolution chosen is 312x312 (image height = 312, image width = 312). The image dimension and batch size is limited so as to avoid memory crash as the RAM is utilized completely. Fig. 2 shows the work flow for the proposed framework.

C. Transfer learning

To train the neural network from scratch, a lot of data along with enough processing power and time is required, which is not exactly practical. Therefore, in transfer learning approach, the last few layers of the existing pre-trained models are tuned to be suitable for the problem of interest. The effectiveness

Fig. 2: Wok flow of the proposed method of two widely used models, namely MobilenetV2 [16] and ResNet50 [18] are examined in this study.

D. Proposed Model

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 310, 310, 32)	320
conv2d_1 (Conv2D)	(None, 308, 308, 32)	9248
max_pooling2d (MaxPooling2D)	(None, 154, 154, 32)	0
conv2d_2 (Conv2D)	(None, 152, 152, 32)	9248
conv2d_3 (Conv2D)	(None, 150, 150, 32)	9248
max_pooling2d_1 (MaxPooling2	(None, 75, 75, 32)	0
conv2d_4 (Conv2D)	(None, 73, 73, 64)	18496
max_pooling2d_2 (MaxPooling2	(None, 36, 36, 64)	0
conv2d_5 (Conv2D)	(None, 34, 34, 64)	36928
max_pooling2d_3 (MaxPooling2	(None, 17, 17, 64)	0
flatten (Flatten)	(None, 18496)	0
dense (Dense)	(None, 128)	2367616
dense_1 (Dense)	(None, 64)	8256
dense_2 (Dense)	(None, 1)	65

Fig. 3: Proposed model architecture

This work proposes a Deep Convolutional Neural Network(CNN) based model to detect and classify TB. Feed-forward

network architecture is used in CNN models to automatically select and extract layers but the Convolution layer, Max-Pooling, and fully integrated layers are considered the primary layers [3]. Various parameters are adjusted in the convolution layers during training process, for educing the descriptive elements before being transferred to a fully connected layer. The extracts are subdivided into targeted categories as "normal and Tuberculosis" by the fully connected layer.

The proposed CNN model is built with feature extraction and feature classification phases. The former phase includes Convolution with ReLU activation function and max pooling operation. The latter phase consists of a fully integrated layer, flatten layer, dense layers and softmax activation. To extract definite features, six convolution layers are there in the network. The first, second, third and fourth convolutional layers use 32, 3×3 filters. Fifth and sixth convolutional layers employ 64, 3×3 filters. All these convolution layers have default hyperparameters such as strides= (1, 1), padding as "valid". All the convolutional layers are employed with ReLU activation function. Max pooling layers with a pooling size of 2×2 are used in between. The flatten layer converts the 3-dimensional output from previous layer to a 1-dimensional vector of size 18496. It is then fed to two dense layers of 128 and 64 neurons respectively which is densely mapped to 1 neuron as needed by the sigmoid classifier. Finally the images are categorized into normal and tuberculosis classes. The proposed TB detection model is shown in Fig 3.

Inorder to apply multiple filters to the same image, each variation can be done separately and stack the results on top of each other and combining them all together. If a convolutional layer accepts d1 feature maps of size (w1 x h1), it produces d2 feature maps of size (w2 x h2), which is calculated by equation(1). features. The deepness of the CNN model influence the extracted features. CNN variants have multiple

$$[w_2, h_2, d_2] = \left[\frac{w_1+2P-F}{s} + 1 \;,\; \frac{h_1+2P-F}{s} + 1 \;, n_f \right] \quad (1)$$

Here F represents filter size, S: stride, P : padding and nf : no. of filters in the corresponding layer. In First convolution layer, convolution of input image of size (312,312) with a filter of (3,3) size and additional parameters like stride and dilation set to 1 and padding 'valid', the number of filters as 32, it produces output size as (310,310,32). Each convolution layer is associated with ReLU activation function to modify feature release maps. The relation between output and input maps corresponding to each convolution layer can be represented as:

$$y_n^k = f\left(\sum_{m \in N_n} x_m^{k-1} * w_{mn}^k + b_n^k \right) \quad (2)$$

where y_n^k denotes the n^{th} output feature of the k^{th} layer, $f(.)$ is a nonlinear ReLU function, x_m^{k-1} refer to the m^{th} input map of $k\text{-}1^{th}$ layer, w_{mn}^k represents the weights for the input and output map in the n^{th} layer, and b^k is the bias associated with the n^{th} map output.

For the first Max Pooling layer, with stride (2,2) and filter size (2,2) applied to a feature map of size (308,308) gives an output feature map of (154,154) size. The Max Pooling Layer performs the down sampling function, thus reducing the spatial dimension of the input by computing the maximum value among the pixels across which the filter is applied.

The pixels along all image channels are converted to a one dimensional vector by the flatten layer. Thus input of (17, 17, 64) is flattened to 18496 values. The number of parameters in each layer is calculated by

No. of parameters=filter-height * filter-width * input-channels * output-channels) + (output-channels)

For compiling the model, Adam optimizer is used. Adam or Adaptive Moment Estimation is an algorithm for optimization technique for gradient descent. For calculating the loss function binary cross entropy is used, since the model is built for a binary classification problem.

An important part of this model is the use of Callback list while training the model. Early Stopping method is used during training to stop the epoch iterations based on some measures. For our model, val loss metric is used as the measure of early stopping, as it needs to be minimum always. After a minimum val loss is achieved, if no further improvement occurs in the val loss for the next few iterations then the training stops at that epoch itself.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

A. Model Parameters

Out of the 7000 images in database, 70% of the total samples were used for network training. This accounts for a total of 4900 images including both Normal and TB manifested cases. 1400 images(20% of total samples) were used for validation and remaining 700(10%) for testing. The initial learning rate was chosen as .000001 which changes after some epochs, depending on some metrics as discussed in section III. Other parameters include batch size - 16, Adam optimizer for model optimization and binary cross entropy to determine loss function. Initially epoch was set to 25, but it stops iteration when early stopping condition is met.

B. Evaluating parameters

Accuracy, Sensitivity(Recall), Specificity and Precision are some of the parameters used for testing the accuracy of the developed model.

$$\text{Accuracy} = t_p + t_n / t_p + t_n + f_p + f_n \quad (3)$$

$$\text{Sensitivity(Recall)} = t_p / t_p + f_n \quad (4)$$

$$\text{Specif icity} = t_n / t_n + f_p \quad (5)$$

$$\text{P recision} = t_p / t_p + f_p \quad (6)$$

where t_p or true positive and t_n or true negative denote the number of correctly labeled TB and Non-TB samples respectively. f_p(false positive) denote the count of TB samples incorrectly labeled as non TB, and f_n(false negative) is the count of non TB samples wrongly classified as TB. The values for t_p, f_p, t_n and f_n can be obtained by plotting the Confusion matrix. This is the best method used to evaluate the performance of any given model by comparing the actual values with the predicted ones. The left diagonal represents total number of Tuberculosis and Normal cases correctly predicted by the model, and the right diagonal shows the incorrect classifications by the model.

C. Performance

We measure the performance of our proposed model using the parameters training loss, validation loss, training accuracy and validation accuracy. Our model stopped training at 20th epoch with a training loss: 0.0439, training accuracy: 98.41%, validation loss: 0.0639 and validation accuracy: 98.18%. Fig.5 illustrates these parameters along with learning rate.

The confusion matrices of the pretrained and proposed models for the given dataset is shown in Fig. 4. We further evaluate our model with the help of parameters such as precision, recall and specificity. Table 1 shows the detailed comparison of pre-trained models with proposed one.

D. Implementation Environment

The network was trained and tested in Python 3 on i5 core CPU with 8 GB RAM support along with the help of Google Colaboratory (Colab), an online 12GB NVIDIA Tesla K80 GPU [25]. The model was built using Keras library [26] supported with Tensorflow backend [27]. After testing phase, the CNN model is integrated to an Android application as well as a Web application. A web page preview is shown in Fig. 6

Since the model is built using Python language and its libraries, Flask web framework was used for creating API handlings in Python which acts as a backend for android and web applications. The input image is uploaded to the backend server through API calls. The android app uses PUT as a request method which in turn produces a JSON response from the server after predicting the X-ray. The result will then be shown to the user through a pop-up message inside the application. Web application makes use of POST request with the X-ray image to the server. The server then renders another web page with the result in it. Both android and web applications are implemented in such a way so as to hide user identity and provide data security. The CXR image that is to be predicted will be renamed as soon as it reaches the server in order to ensure user privacy.

V. CONCLUSION

In this paper, we developed a customized deep CNN model which was trained from scratch by using the dataset described in section III. The proposed model can detect Tuberculosis from chest X-ray images and classify them as either Normal or Tuberculosis cases and that it could produce a comparable testing accuracy with the pretrained models. As a future work, we wish to build our own dataset in collaboration with medical

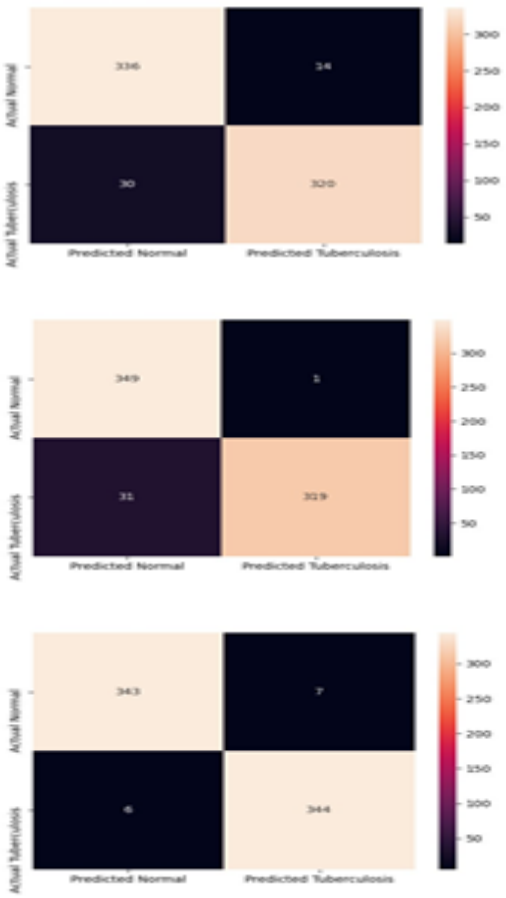

Fig. 4: Confusion matrix a) MobilenetV2 b) Resnet50 c) Proposed method

Sl.	Model	Precison	Recall(i	Specificity(in %)	Test accuracy(in
1	Mobilenet V2	91.42	95.80	91.80	93.71
2	Resnet50	91.14	99.68	91.84	95.42
3	Proposed method	98.28	98	98.28	98.14

TABLE I: Comparison of proposed model with pre-trained models

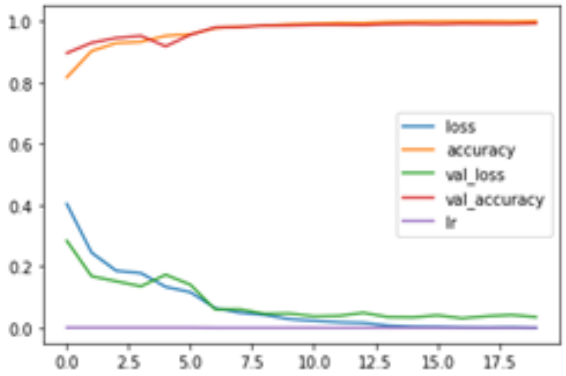

Fig. 5: Accuracy and loss plot

Fig. 6: Webpage preview

experts to increase the diversity as well as the accuracy even further. We would also explore possibilities on including more and more disease classification rather than just Tuberculosis.

REFERENCES

[1] Leibstein, J.M., Nel, A.L, "A.: Detecting tuberculosis in chest radiographs using image processing techniques", University of Johannesburg (2006).

[2] JRohilla A., Hooda R., Mittal A, "PTB detection in chest radiograph using deep learning architecture", Proceeding of 5th International Conference on Emerging Trends in Engineering, Technology, Science and Management (ICETETSM-2017), pp. 136–147 (2017).

[3] Srinivas, M., Roy, D., Mohan, C.K.: Discriminative feature extraction from X-ray images using deep convolutional neural networks. In: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 917–921 (2016).

[4] Hooda, R., Sofat, S., Kaur, S., Mittal, A., Meriaudeau, F.: Deep-learning: a potential method for tuberculosis detection using chest radiography. In: IEEE International Conference on Signal and Image Processing Applications (ICSIPA), pp. 497–502 (2017).

[5] Jaeger, S., Karargyris, A., Antani, S., Thoma, G.: Detecting tuberculosis in radiographs using combined lung masks. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 4978–4981 (2012).

[6] Karargyris, A., et al.: Combination of texture and shape features to detect pulmonary abnormalities in digital chest X-rays. Int. J. Comput. Assist. Radiol. Surg. 11(1), 99–106 (2015).

[7] Van Ginneken, B., Katsuragawa, S., ter Haar Romeny, B.M., Doi, K., Viergever, M.A.: Automatic detection of abnormalities in chest radiographs using local texture analysis. IEEE Trans. Med. Imaging 21(2), 139–149 (2002).

[8] Shen, R., Cheng, I., Basu, A.: A hybrid knowledgeguided detection technique for screening of infectious pulmonary tuberculosis from chest radiographs. IEEE Trans. Biomed. Eng. 57(11), 2646–2656 (2010).

[9] Hogeweg, L., Mol, C., de Jong, P.A., Dawson, R., Ayles, H., van Ginneken, B.: Fusion of local and global detection systems to detect tuberculosis in chest radiographs. In: Jiang, T., Navab, N., Pluim, J.P.W., Viergever, M.A. (eds.) MICCAI 2010. LNCS, vol. 6363, pp. 650–657. Springer, Heidelberg (2010).

[10] Leibstein, J.M., Nel, A.L.: Detecting tuberculosis in chest radiographs using image processing techniques. University of Johannesburg (2006).

[11] Rohilla, A., Hooda, R., Mittal, A.: TB detection in chest radiograph using deep learning architecture. In: Proceeding of 5th International Conference on Emerging Trends in Engineering, Technology, Science and Management (ICETETSM-2017), pp. 136–147 (2017).

[12] Shin, H.C.; Roth, H.R.; Gao, M.; Lu, L.; Xu, Z.; Nogues, I.; Yao, J.; Mollura, D.; Summers, R.M. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. IEEE Trans. Med. Imaging 2016, 35, 1285–1298.

[13] Ho, T.K.K.; Gwak, J. Multiple feature integration for classification of thoracic disease in chest radiography. Appl. Sci. 2019, 9, 4130.

[14] Shin, H.C.; Roberts, K.; Lu, L.; Demner-Fushman, D.; Yao, J.; Summers, R.M. Learning to read chest x-rays: Recurrent neural cascade model for automated image annotation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 2497–2506.

[15] Li, Z.; Wang, C.; Han, M.; Xue, Y.; Wei, W.; Li, L.J.; Fei-Fei, L. Thoracic disease identification and localization with limited supervision. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Palace Convention Cetner, Salt Lake City, UT, USA, 28–23 June 2018; pp. 8290–8299.

[16] Mark Sandler; Andrew Howard; Menglong Zhu; Andrey Zhmoginov; Liang-Chieh Chen; MobileNetV2: Inverted Residuals and Linear Bottlenecks, June 2018, https://ieeexplore.ieee.org/document/8578572

[17] Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv e-prints [Internet]. 2014 September 01, 2014:[arXiv:1409.556 p.].

[18] Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna ZB. Rethinking the Inception Architecture for Computer Vision2016.

[19] Christian Szegedy; Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning; Aug 2016

[20] He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition2016. 770–8 p.

[21] Huang G, Liu Z, van der Maaten L, Weinberger K. Densely Connected Convolutional Networks2017.

[22] Tawsifur Rahman, Amith Khandakar, Muhammad A. Kadir, Khandaker R. Islam, Khandaker F. Islam, Zaid B. Mahbub, Mohamed Arselene Ayari, Muhammad E. H. Chowdhury. (2020) ”Reliable Tuberculosis Detection using Chest X-ray with Deep Learning, Segmentation and Visualization”. IEEE Access, Vol. 8, pp 191586-191601.DOI.10.1109/ACCESS.2020.3031384.

[23] S. Jaeger, S. Candemir, S. Antani, Y.-X. J. Wa´ng, P.-X. Lu, and G. Thoma, ”Two public chest X-ray datasets for computer-aided screening of pulmonary diseases,” Quantitative imaging in medicine and surgery, vol. 4 (6), p. 475(2014)

[24] B. P. Health. (2020). BELARUS TUBERCULOSIS PORTAL [Online]. Available: http://tuberculosis.by/. [Accessed on 09-June-2020]

[25] NIAID TB portal program dataset [Online]. Available: https://data.tbportals.niaid.nih.gov/.

[26] kaggle. RSNA Pneumonia Detection Challenge [Online].Available: https://www.kaggle.com /c/ rsna-pneumonia-detection-challenge/data. [Accessed on 09-June-2020]

[27] Bisong, E. Building Machine Learning and Deep Learning Models on Google Cloud Platform; Springer: Berlin, Germany, 2019.

[28] Chollet, F. Keras. 2015. Available online: https://github.com/fchollet/keras
[29] Abadi, M.; Agarwal, A.; Barham, P.; Brevdo, E.; Chen, Z.; Citro, C.; Corrado, G.S.; Davis, A.; Dean, J.; Devin, M.; et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems. 2015. Available online: https:// tensorflow.org