International Journal of Engineering Trends and Technology
ISSN: 2231 – 5381 / https://doi.org/10.14445/22315381/IJETT-V70I8P233

Volume 70 Issue 8, 318-326, August 2022

Original Article

Preliminary Screening for Pulmonary Tuberculosis from Chest Radiography using Artificial Neural Network

Sucheera Phramala[1], Weeragul Pratumgul[2], Jagraphon obma[3], Worawat Sa-ngiamvibool[1*]

[1]*Faculty of Engineering, Mahasarakham University, Mahasarakham 44150, Thailand*
[2]*Faculty of Engineering, Pathumwan Institute of Technology, Bangkok 10330, Thailand*
[3]*Faculty of Engineering, Rajamangala University of Technology ISAN Khonkaen Campus, Khonkaen 40000, Thailand*

wor.nui@gmail.com

Received: 25 June 2022 Revised:11 August 2022 Accepted:13 August 2022 Published: 27 August 2022

Abstract *- This research has developed an algorithm for primary screening of pulmonary tuberculosis from chest radiographs by using the image processing principle with Artificial Neural Network (ANN) to examine the preliminary features related to the high incidence of pulmonary tuberculosis, namely Reticular Infiltration, Cavity, and Consolidation. The procedure was then used to learn 14,000 chest radiographs and tested for preliminary screening on 6,000 test images. The test result found that this method can process with an accuracy of 82.20%, a sensitivity of 86.80%, specificity of 79.59% and positive predictive values of 77.37% compared with the radiologist.*

Keywords *- Chest radiography, Artificial neural network, Image processing, Pulmonary tuberculosis.*

1. Introduction

Tuberculosis is the most common disease in Southeast Asia. It causes disabilities and the highest mortality rate. The World Health Organization (WHO) [1] indicated that Thailand was one of the 14 countries with serious problems with tuberculosis in three aspects: the incidents of tuberculosis (TB), the coincidence of tuberculosis and human immunodeficiency virus (TB/HIV) and multidrug-resistant tuberculosis (MDR(-TB)). Normally, tuberculosis can be found in all organs. However, pulmonary tuberculosis is the most common case in Thailand, with an approximate percentage of 80. The number of new patients with the disease was over a hundred thousand per year in 2018. The average number of patients with the disease in the country was 108,000. It was found that there were 156 patients per 100,000 people [2]. According to the mentioned data, pulmonary tuberculosis and the medical treatments by physicians require pulmonary images to conduct chest radiography [3 - 6], that is radiography taking the photos of chests with X-rays on films and processing the photos for physicians to analyze. It is a method providing clear images, and there are studies about various chest radiographic techniques [7-8].

Many studies used chest X-ray images to analyze pulmonary tuberculosis. For example, H.Das et al. [9] segmented the pulmonary images with the region-based active contour. The segmented parts had outstanding characteristics such as IH, GM, SD, HOG and LBP to classify the images into normal and abnormal images using the support vector machine (SVM). It took 18 minutes to search for an image. The sets of images were tested. Then, there were developments in the analyses of the chest radiographs using various artificial neural networks. For instance, Y.Cao et al. [10] used mobile phones connected to the clouds to screen the chest radiographs by improving the accuracy and the classification of the outstanding characteristics with the deep convolutional neural network and SVM. However, there was the limitation that the size of the inputs or images must be 227x227 pixels. On the other hand, S.Hwang et al. [11-12] presented the screening techniques with the artificial neural networks by using three sets of data from the Korean Institute of Tuberculosis (KIT), Montgomery College (MC) and Shenzhen College (SC). Deep learning was also suggested. [13] Tawsifur Rahman et al. [14] studied the nine layers of artificial neural networks. The accuracy, correctness, sensitivity and specificity of the diagnosis of tuberculosis with the radiographs were 96.47%, 96.62%, 96.47%, 96.47% and 96.51%. Eman Showkatian et al. [15] proposed the accepted model for classifying the artificial neural networks with an accuracy of 88.0%, the sensitivity of 87.0%, the F1 score of 87.0%, and the correctness of 87.0% and the AUC of 87.0%. The highest efficiency of the automatic classification of the types of TB was obtained. The accuracy, the sensitivity, the F1 score and the AUC were 91.0%. The correctness was 90.0%. The mentioned studies confirmed the accuracy and the correctness of the artificial neural network algorithms. Thus,

the researcher(s) presented the initial analysis of pulmonary tuberculosis with chest radiographs by applying the artificial neural network for medical benefits.

2. The Principles and the Designs

2.1. The Principles of the Artificial Neural Network

To initially screen the chest radiographs with the artificial neural network implementation, the researcher presented the structures of the artificial neural network that consisted of four layers. The input layer has three neurons for receiving the values of the characteristics from the previous step. The recticular values include the infiltration cavity and the Consolidation. If the neural cells collect the signals by calculating the weighted sum of the inputs, the sum will be analyzed and interpreted with the activation function. The hidden layer has 10 neurons. The output layer displays the numeral results in the binary system. The abnormality levels of the chest radiographs consist of the normal level and the TB candidate level, as shown in Fig. 1

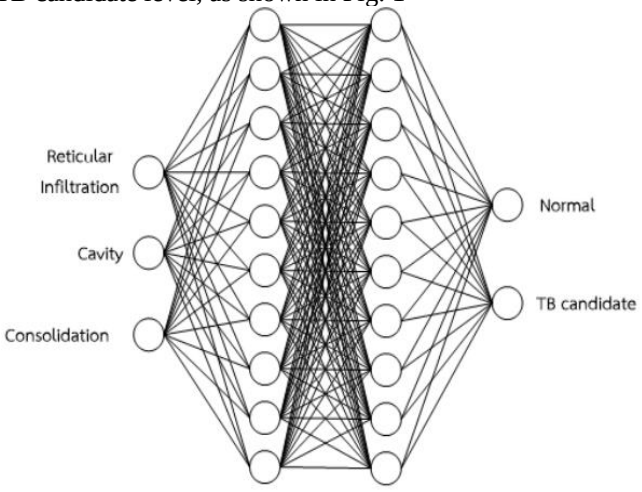


Fig. 1 The Structure of the Artificial Neural Network for Initially Screening for the Pulmonary Tuberculosis from the Chest Radiographs

The outputs from the hidden layer are evaluated with Equation 1

$$S_j^h = f(\sum_{i=1}^{10} W_{ji}^h S_i - \theta_j^h) \quad (1)$$

The outputs from the output layers are evaluated by Equation 2

$$b_k = f(\sum_{i=1}^{10} W_{kj}^0 S_j^h - \theta_k^0) \quad (2)$$

Where W_{ji}^h W_{kj}^0 is the weight between the layers, and θ_j^h θ_k^0 is the respective error.

The error vector for the hidden layer can be calculated with the equation

$$e_k = b_k(1 - b_k)(d_k - b_k)$$

The error vector for the output layer can be calculated with the equation

$$e_j = s_j^h(1 - s_j^h) \sum_{k=1}^{10} w_{kj}\ e_k$$

Where d_k is the goal of the desired outputs.

The weighted adjustments for the output and hidden layers are as follows.

$$W_{ji}^h(new) = w_{kj} + ns_j^h e_k \quad (3)$$

$$\theta_k^0(new) = \theta_k^0 + ne_k \quad (4)$$

$$w_{ji}(new) = w_{ji} + ns_j e_j \quad (5)$$

$$\theta_j^h(new) = \theta_j^h + ne_k \quad (6)$$

2.2. The user interface designs.

The structure of the artificial intelligent network must be used for initial screening for pulmonary tuberculosis from chest radiographs. The network had to perform the learning tasks, classify the forms of pulmonary tuberculosis and memorize the forms for the initial analysis. To design the user interfaces, the users must be able to conveniently and instantly test the interfaces. The users must be able to communicate with the system through the screens instead of directly typing commands. The interfaces were developed with Matlab, as shown in Fig. 2.

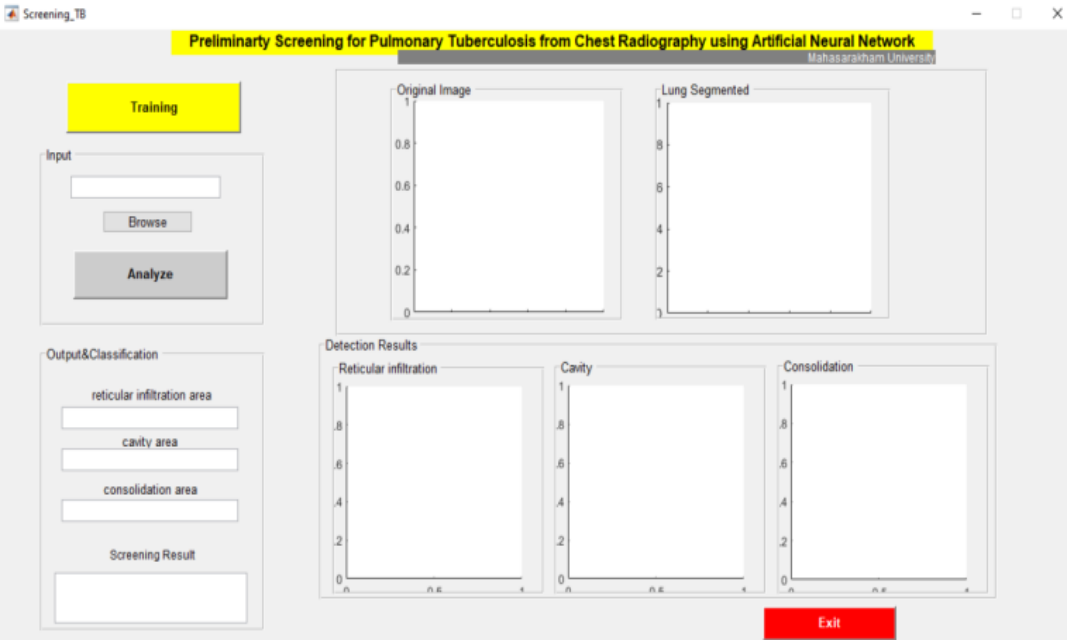


Fig. 2 The Designs of the User Interfaces with MATLAB for Initially Screening for Pulmonary Tuberculosis from the Chest Radiographs with the Artificial Neural Network

3. Methodology

The methodology consists of three processes. The 1st process is the preparation of the chest radiographs. The 2nd process is the identification of the outstanding characteristics of the chest radiographs. The 3rd process is identifying the abnormality levels of the chest radiographs. The processes mentioned above can be concluded as shown in Fig. 3.

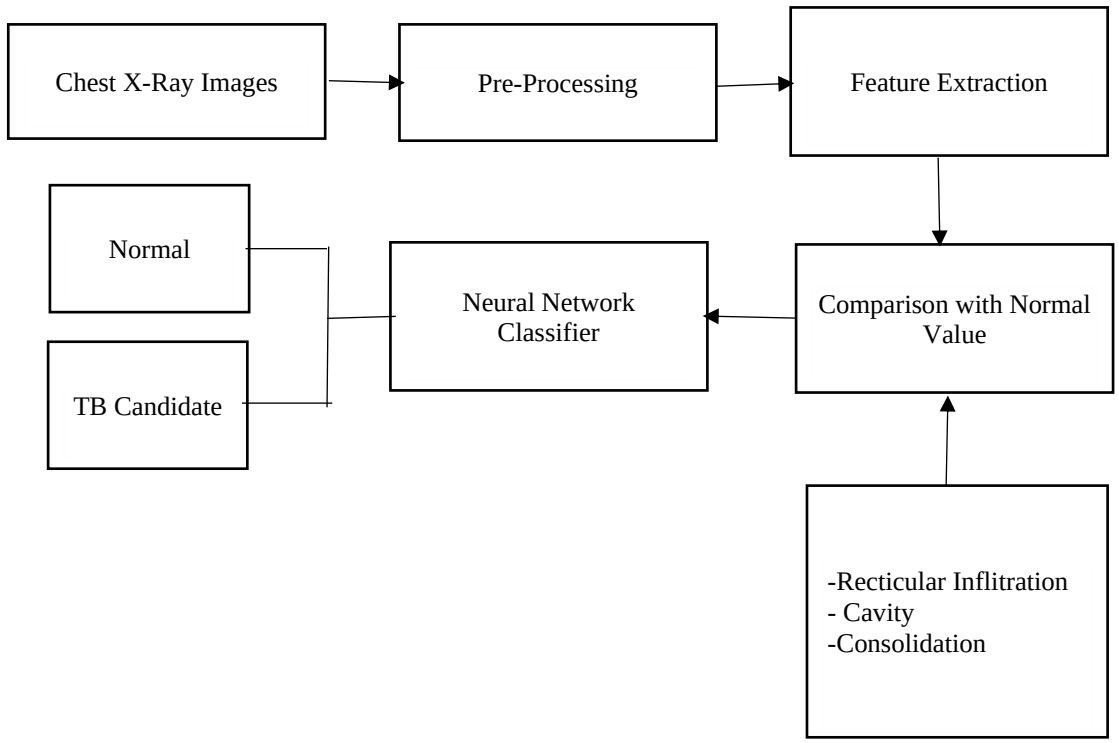


Fig. 3 The Processes for Initially Screening for Pulmonary Tuberculosis from the Chest Radiographs with the Artificial Neural Network

3.1. The 1st Process – The Preparation of the Chest Radiographs

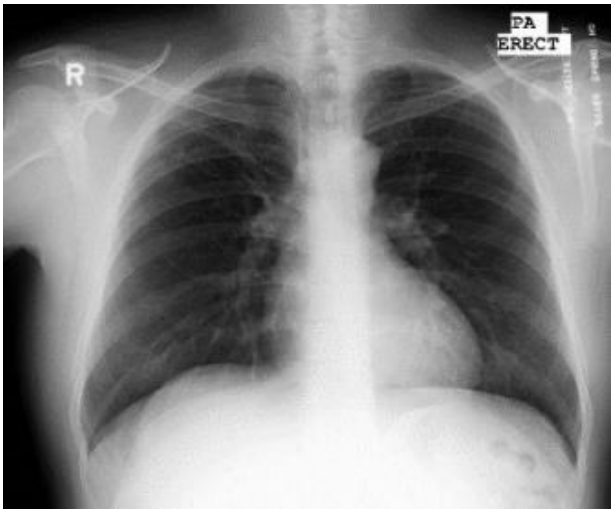
The chest radiographs were obtained from the three chest radiograph databases as follows.

1. The 138 chest radiographs from the Montgomery County Chest X-ray Set (MC), including the 80 normal chest radiographs and the 58 radiographs related to tuberculosis (TB related).
2. The 662 chest radiographs from the Shenzhen Chest X-Ray Set (SC) consist of 326 normal chest radiographs and 336 TB-related radiographs.
3. The 19,200 chest radiographs from the National Institutes of Health (NIH) comprised 9,524 normal chest radiographs and 9,606 TB-related radiographs.

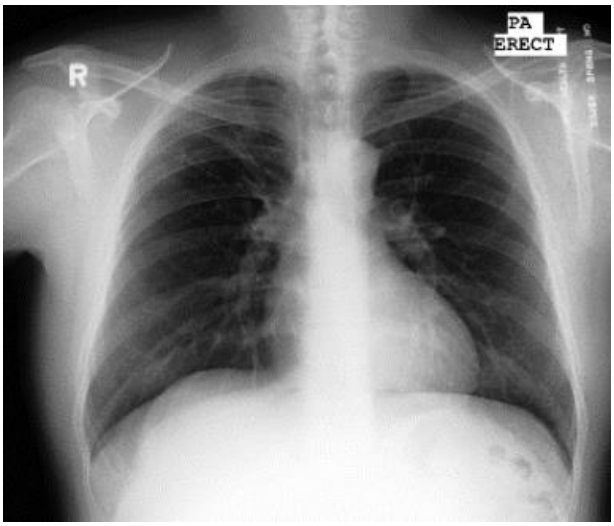
All radiographs were taken with the computer radiography (CR) system in the format of. PNG with 12 bits in the grayscale mode. There were three sizes of the images: 4,020 x 4,892 pixels, 4,892 x 4,020 pixels and 3,000 x 3,000 pixels. The images were divided into two sets for training the system to memorize the images and testing the computer's abilities to support the diagnosis of pulmonary tuberculosis: 14,000 images and 6,000 images, respectively. For the pieces of training, each group had 7,000 images; For the tests, each group had 3,000 images. The outstanding characteristics of the images were adjusted before pre-processing. The Hounsfield unit (HU) divided the grayscale into various levels. The values were from -1000HU to +1000 HU, as shown in the following details.

- Air <-500 HU (the blackest black value)
- Fat <-90 HU
- Water and CSF = 0HU
- Hematoma 50-90 HU
- Calcification >+80 HU
- Bone) >+500 HU (the whitest white value)

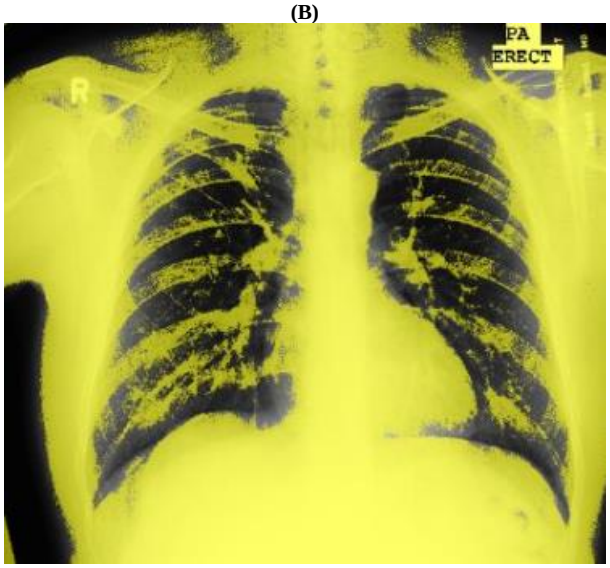
The homomorphic filter was used to increase the radiographs' brightness and reduce the multiplicative noises found in the radiographs. The rays with the inconsistent intensities resulted in the bend edges of ribs and concealed details. Then, the histogram equalization, the median filter and the contrast-limited adaptive histogram equalization (CLAHE) were used. The image segmentation was conducted using the active contour model to specify the lungs' shapes in the radiographs, as shown in Fig. 4.



(A)



(B)



(C)

Fig. 4 (A) The Original Chest Photographs (B) The Radiographs After the Contrast Adjustment (C) The Images from the Active Contour Model

3.2. The 2nd Process – The Identification of the Outstanding Characteristics in the Chest Radiographs

By adjusting the characteristics with the grayscale by separating intensity and contrast, the adjusted grayscale will be used for the reticular infiltration by using the opening function for processing mathematical shapes and applying the principle of image erosion. The image dilation follows it. The erosion function can maintain the reticular infiltration in the images while removing unwanted parts, while the dilation function can improve the visibilities of the removed parts. Then, the bones (collarbones and ribs) are removed by comparing the values and subtracting the original images before opening the images.

Then, the images are converted into binary images. The pixel with a value higher than the threshold is 1 (white), while that lower than the threshold is 0 (black). After that, the noises are removed. The AND logic is used. The pixels of the segmented images and the current images are compared. If the values are equal, the value will be 1 (white) and replace the points of the recticular infiltration. The results from using the logic are shown in Fig. 5.

Fig. 5 The Areas of the Recticular Infiltration in the Chest Radiographs

Pulmonary tuberculosis has various chest radiographic characteristics dependent on the pathology of the disease. Some characteristics are specific to the disease, such as the wounds of pulmonary cavities. Especially for upper lobes, the patients with the wounds usually have tuberculosis diagnosed from the stained sputum, as shown in Fig. 6.

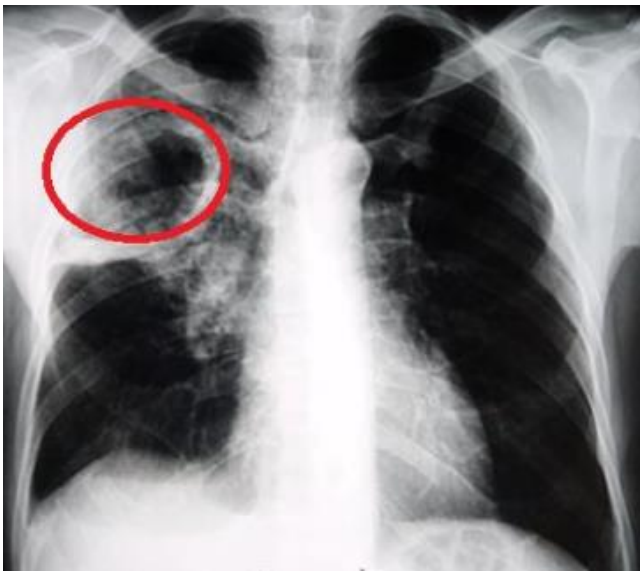


Fig. 6 The Characteristics of the Wounds of the Pulmonary Cavities on the Chest Radiographs

Source: Dr. Wannabhorn Buriwong, the Teaching documentation about Pulmonary Radiology

The grey level co-occurrence matrix (GLCM) is used to identify the pulmonary cavities' wounds. Each pixel is analyzed to examine the relationships of the two pairs of GLCM at the different greyscale levels. The relationship between the greyscale levels can be built by converting the chest Radiographs into grayscale images. The light dispersion is adjusted with the adaptive equalization technique. The relationships can be identified with the matrix to compare the grayscale levels of the lung parenchyma and the wounds of the pulmonary cavities, as shown in Figure 7 (A).

The image processing technique for identifying the dark shadows of the pathological conditions of the Consolidation starts by using the grayscale to adjust the characteristics of the images before being processed. Then, the technique is used to process the methematical shapes to remove the blood vessels in the images and identify the dark shadows. The collarbones and the ribs must be identified and removed from the images to improve the accuracy of processing the images used for the recticular infiltration because the dark shadows have the outstanding characteristics shown in the grey group(s) in the figure. To process the images, the differences in the colors of the images, dark (black) and bright (white), are used, as shown in Fig. 7 (B).

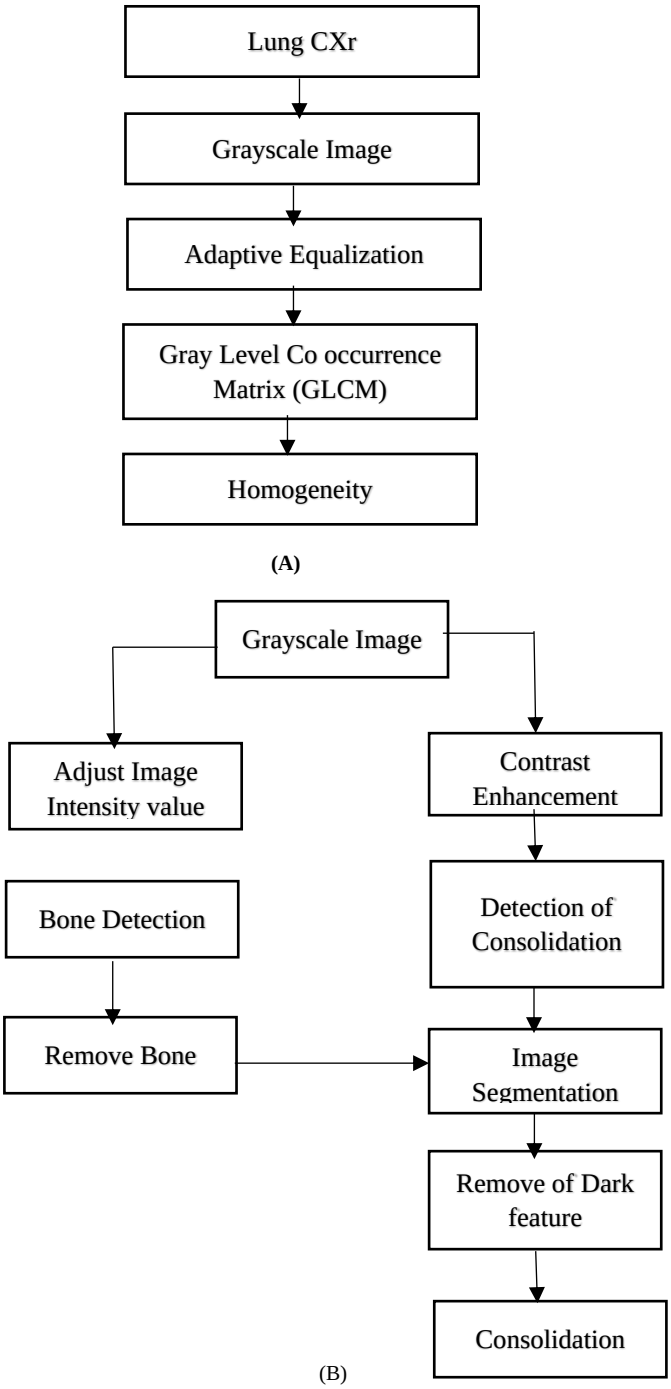


Fig. 7 (A) The Processes of the Identification of the Pulmonary Cavities and (B) the Processes of the Identification of the Dark Shadows of the Pathological Conditions of the Consolidation

3.3. The 3rd Process - The Identification of the Abnormal Values of the Chest Radiographs

To identify the abnormal values of the chest radiographs by using the data for training and testing the abilities of the artificial neural network with the 14,000 chest radiographs, the images are divided into two sets: the set for training the system to memorize the images and the set for testing the abilities of the computer to initially diagnose pulmonary tuberculosis. Each set has 3,000 images. These include the 3,000 normal chest radiographs and the 3,000 TB-related images. After processing the images, the outputs are presented in the binary system, as shown in Table 1.

Table 1. The Substitution of the Abnormal Numbers in the Chest Radiographs

The severity of the disease	Binary
Normal	00
Abnormal (Suspect Pulmonary Tuberculosis)	10

The examples of initially screening for pulmonary tuberculosis with the artificial neural network are displayed on the user interfaces of MATLAB. After the users command the system to analyze the data, the images are displayed on the right window consisting of two parts: the upper window showing the original images being analyzed and the images from the lung segmentation as well as the lower window showing the images from the feature extraction for the three characteristics L reticular infiltration, cavity and Consolidation. The lower left window also shows the calculation results of the areas of outstanding characteristics and the abnormalities of the chest radiographs. The results are normal or TB related, as shown in Fig. 8.

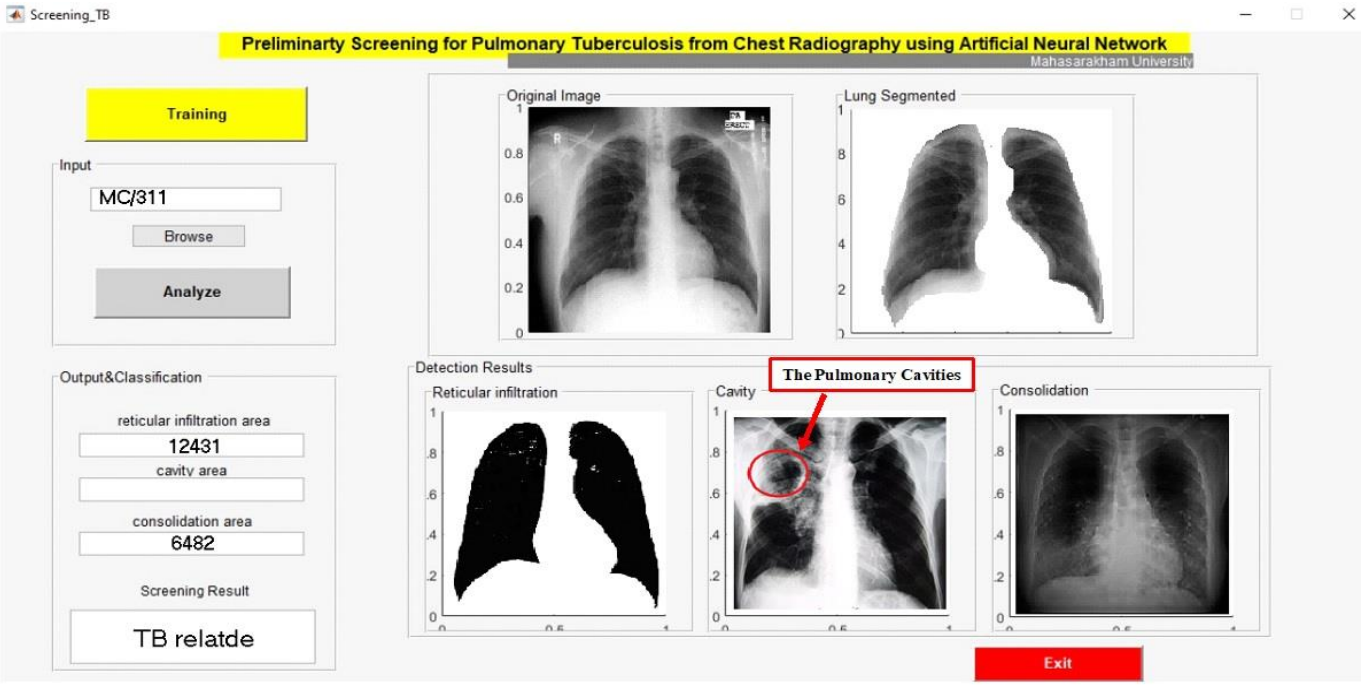


Fig. 8 The Examples of the User Interfaces of MATLAB for Initially Screening for Pulmonary Tuberculosis from the Chest Radiographs with the Artificial Neural Network

4. Results

The results from the initial screening for pulmonary tuberculosis included the 14,000 images divided into two groups: the normal group and the TB-related group. Each group had 7,000 images. The number of images that tested the system's accuracy for screening pulmonary tuberculosis was 6,000. These images were divided into the normal group and the TB-related group. Each group had 3,000 images. The sensitivity, specificity and probability of pulmonary tuberculosis were tested to check whether the values are positive predictive by using Equations 7, 8 and 9. The numbers are arranged in Table 2 [3] as follows.

$$\text{Sensitivity} = \frac{A}{A + C} \tag{7}$$

$$\text{Specificity} = \frac{D}{B + D} \tag{8}$$

$$\text{Positive Predictive Value} = \frac{D}{B + D} \tag{9}$$

Table 2. The Calculation of the Screening Efficiency

The proposed screening	**Screening and confirmation from the clinic**	
Test results	**Pathology**	
	Diseased	**Not diseased**
Positive	(True Positive) A	(False Positive) B
Negative	(False Negative) C	(True Negative) D

The results from testing the accuracy of the proposed technique were as follows.

1. The results from processing and analyzing the chest radiographs of the normal people or groups showed 2,647 correct images from the 3,000 tested images. The accuracy was 88.23%, as shown in Table 2. Then, the results from calculating the screening efficiency are shown in Table 3.

Table 3. The Accuracy of the Proposed Technique

Picture Group	**The number of images training**	**The number of images testing**	**The number of images correctly specified**	**Accuracy (%)**
Normal	7,000	3,000	2,674	88.23
with pulmonary tuberculosis	7,000	3,000	2,321	77.37
Total	14,000	6,000	4,968	82.80

2. The results from processing and analyzing the chest radiographs of the patients with pulmonary tuberculosis or the TB-related group showed 2,321 correct images of the 3,000 tested images. The accuracy was 77.37%. Generally, this study could initially check and screen pulmonary tuberculosis. There were 4,968 images of the 6,000 tested images. The accuracy was 82.80%, as summarized in Table 3

Table 4. The Results from Calculating the Screening Efficiency

True Positive	True Negative	False Negative	False Positive	Sensitivity	Specificity	Positive Predictive Value
2,321	679	353	2,647	86.80%	79.59%	77.37%

The results from calculating the screening efficiency must also be checked in the laboratory(ies) to confirm the results. Therefore, the proposed technique in this study had a sensitivity of 86.80%. It was at a high level. It can be improved actually to screen for pulmonary tuberculosis.

5. Conclusion and Discussion

The proposed technique in this study uses the characteristics for identifying the relationships with pulmonary tuberculosis, including reticular infiltration, cavities and Consolidation. The accuracy was 86.80%. The accuracy level was quite high. It shows that the three characteristics are related to the disease. Hence, this confirms the potential for improving the technique for actually screening for pulmonary tuberculosis.

Then, the errors from the technique were as follows. The chest radiographs of the patients with pulmonary tuberculosis were normal and did not have the three characteristics of the patients with both pulmonary tuberculosis and HIV. The patients with tuberculosis have milliary infiltration, hilar adenopathy and diffuse pulmonary involvement that are different from that without HIV.

Finally, the other errors were the normal images with an outstanding characteristics that met the threshold. The causes of the errors could be the pathological conditions of other diseases such as emphysema, pulmonary edema, pneumonia, atelectasis, pnemoconiosis, lung tumors and stage 1 lung cancer.

References

[1] World Health Organization, Global Tuberculosis Control: WHO Report, 2012, [Online]. Available: http://www.who.int/tb/publications/global_report/2012/gtbr12_full.pdf

[2] Tuberculosis Office, Department of Disease Control, Ministry of Public Health, THAILAND, "Tuberculosis Data Center.TBcm(Data Center," 2021. [Online]. Available: http://122.155.219.72/tbdc/frontend/web/index.php.

[3] Kanchana Chansung, "Critical Appraisal of Scientific Paper(Part III)," *Srinagarind Med J*, vol. 14, no. 1, pp. 62–67, 1999.

[4] Sant Anna CC, Schmidt CM, Pombo March MF, Pereira SM, Barreto ML, "Radiologic Findings of Pulmonary Tuberculosis in Adolescents," *Braz J Infect Dis*, vol. 15, no. 1, pp. 40–44, 2014.

[5] Tempon Kuamak, Ravivan Pantaverakulans Pasuporn Pho-ngennark,"Chest Radiographic Findings in Pulmonary Tuberculosis," *Buddha Chinaraj Medical Journal*, vol. 32, no. 2, pp. 134–141, 2016.

[6] Wasinan Pholphuech, "Chest Radiographic Finding in Pulmonary Tuberculosis (Banmee Hospital)," *Research and Development Health System Journal*, vol. 13, no. 2, pp. 11–19, 2020.

[7] A.Fojnica, A.Osmanovic and A. Badnjevic, "Dynamic Model of Tuberculosis-Multiple Strain Prediction Based on Artificial Neural Network," In 5th *Mediterranean Conference on Embedded Computing*, pp. 290–293, 2016.

[8] S.Candemir, S.Jaeger, K.Palaniappa, J.P.Musco, R.K.Singh, Z.Xue, A.Karargyris, S.Antani, G.Thoma, and C.J.McDonald, "Lung Segmentation in Chest Radiographs Using Anatomical Atlases with Nonrigid Registeration," *IEEE Transaction on Medical Imaging*, vol. 3, no. 2, pp. 577–590, 2014.

[9] H.Das and A.Nath, "An Efficient Detection of Tuberculosis from Chest X-Rays," *International Journal of Advance Research in Computer Science and Management Studies*, vol. 3, no. 5, pp. 149–154, 2015.

[10] Cao, Yu et al., "Improving Tuberculosis Diagnostics Using Deep Learning and Mobile Health Technologies among Resource-Poor and Marginalized Communities," Proceedings - *IEEE 1st International Conference on Connected Health: Applications, Systems and Engineering Technologies, CHASE*, vol. 2016, no. 1, pp. 274–281, 2016.

[11] S.Hwang, H.Kim, Ji Jeongb, and H.Kimc, "A Novel Approach for Tuberculosis Screening Based on Deep Learning Convolutional Neural Networks," In *SPIE Optical Engineering Press*, 2016a.

[12] "A Novel Approach for Tuberculosis Screening Based on Deep Learning Convolutional Neural Networks," In Proceedings of SPIE, *Medical Imaging: Computer-Aided Diagnosis*, 2016b.

[13] Liu, Chang et al., "TX-CNN: Detecting Tuberculosis in Chest X-Ray Images Using Convolutional Neural Network," Proceedings - *International Conference on Image Processing, ICIP*, pp. 2314–2318, 2018

[14] Tawsifur Rahman, AmithKhandakar, Muhammad Abdul Kadir, Khandaker R. Islam, Khandaker F. Islam, Rashid Mazhar, Tahir Hamid, Mohammad T. Islam, Zaid B. Mahbub, Mohamed ArseleneAyari and Muhammad E. H. Chowdhury, "Reliable Tuberculosis Detection using Chest X-ray with Deep Learning, Segmentation and Visualization," *IEEE Access*, vol. 8, pp. 191586 – 191601, 2020.

[15] EmanShowkatian, Mohammad Salehi, Hamed Ghaffari, Reza Reiazi,corresponding and Nahid Sadighi, “Deep Learning-Based Automatic Detection of Tuberculosis Disease in Chest X-Ray Images,” *Pol J Radiol*, vol. 87, pp. 118 – 124, 2022.

[16] P. Seetha Subha Priya , S. Nandhinidevi , M. Thangamani and S. Nallusamy4, “A Review on Exploring the Deep Learning Concepts and Applications for Medical Diagnosis,” *International Journal of Engineering Trends and Technology*, vol. 68, no. 10, pp. 63–66, 2020.

[17] Yasser Mohammad Al-Sharo, Amer Tahseen Abu-Jassar, SvitlanaSotnik and Vyacheslav Lyashenko, “Neural Networks as a Tool for attern Recognition of Fasteners,” *International Journal of Engineering Trends and Technology*, vol. 69, no. 10, pp. 151–160, 2021.

[18] Nafis Khuriyati, Agung Putra Pamungkas, Anggraito Agung P., "The Sorting and Grading of Red Chilli Peppers (Capsicum annuum L.) Using Digital Image Processing," *SSRG International Journal of Agriculture & Environmental Science*, vol. 6, no. 4, pp. 17-23, 2019. *Crossref*, https://doi.org/10.14445/23942568/IJAES-V6I4P104

[19] Shubham Srivastava, Himanshu Bhardwaj, Aman Dixit, Prof. Namita Kalyan Shinde, "ECG Pattern Analysis Using Artificial Neural Network," *SSRG International Journal of Electronics and Communication Engineering*, vol. 7, no. 5, pp. 1-4, 2020. *Crossref*, https://doi.org/10.14445/23488549/IJECE-V7I5P101.