International Journal of Allied Medical Sciences and Clinical Research (IJAMSCR)

ISSN:2347-6567

IJAMSCR |Volume 8 | Issue 3 | Jul - Sep - 2020
www.ijamscr.com

Research article *Medical research*

Tuberculosis detection using convolutional neural network

Mochamad Imron, Melyana Nurul Widyawati, Siti Masrochah

Postgraduate Imaging Diagnostic Program, Health Polytechnic of Semarang, Semarang, Central Java, Indonesia

***Corresponding Author**: **Mochamad Imron**

Email id: imronjambi74@gmail.com

ABSTRACT

Background

Pulmonary tuberculosis is an infectious disease has become one of the ten leading causes of death globally. Increasing the number and variety of radiological examinations increases the workload of radiologists. This causes the radiologist to experience fatigue, and trigger an inaccurate diagnosis, missed or delayed diagnosis. Machine learning is a computational model with an algorithm that is similar to the structure and function of the biological network of the human brain. It's part of artificial intelligence that uses computer science to perform digital image processing with pattern recognition techniques. The algorithm in machine learning can calculate, recognize the pattern in the image, and make predictive diagnoses.

Objective

Generate deep learning model that can classify the chest x-rays image as tuberculosis and normal, also have the same performance with radiologists.

Methods

The deep learning model using Convolutional Neural Network (CNN) with the input image size and filter size variation has developed, then compared to the expert performance.

Results

Obtained the optimum deep learning model using an image of 200 x 200 and 5 x 5 filter size that has an accuracy, sensitivity, specificity, precision, and AUC were 0.97, 0.9667, 0.975, 0.9831, and 0.971 with CI of 0.932-1.

Conclusion

The deep learning model has 98% classification similarity with expert has obtained.

Keywords: Convolutional Neural Network, Deep learning, Tuberculosis.

INTRODUCTION

Pulmonary tuberculosis (TB) is a disease caused by bacteria (Mycobacterium tuberculosis). This disease if not treated effectively will be chronic [1]. TB is one of the 10 main causes of death and as a single infectious agent, higher than deaths due to HIV or AIDS. Every year millions of people get infected and suffer pulmonary TB. In 2017 around 1.3 million people died caused by TB. Globally the development of TB in 2017 is 10 million people [2, 3]. Diagnosis of TB is based on the patient's history, physical examination, and supporting examinations

namely laboratory and radiological examinations. Laboratory tests in the form of Acid-Resistant Bacteria Test or Xpert® MTB / RIF are the gold standard. Radiological examination for TB cases is the posterior-Anterior chest x-rays position [4]. Increasing the number and variety of radiological examinations increases the workload of the radiologist. This condition causes the radiologist to experience fatigue so that it can trigger inaccuracies in diagnosis, missed, and delayed diagnosis [5]. Besides, intra and inter-individual variability of interpretations by radiologists tends to be high [6].

In radiology, the use of film has decreasing and has been replaced by digital images. The use of digital images from Computer Radiography and Digital Radiography [7] allows processing in the form of image processing, image analysis, image understanding, and computer vision [8]. Deep learning is part of artificial intelligence that uses computer science to do digital image processing with pattern recognition techniques. The algorithm system in deep learning can calculate, recognize patterns in images, and make diagnosis predictions.

CNN is an artificial neural network consisting of several layers of computational connections such as neurons with minimal processing step by step, has experienced significant progress in the field of computer vision research. CNN architecture consists of convolutional, Relu, pooling layers, and fully connected layers [9]. The main purpose of the convolutional layer is to detect edges, lines, and visual elements such as typical local motifs. The parameters of the special filter operator called convolution [10]. The advantage of CNN is that it can learn feature representations automatically from training data. Several CNN layers aim to process imaging data with varying levels of abstraction, allowing machines to navigate and explore large data sets and discover complex structures and patterns that can be used for predictions. CNN has the high performance in classification medical image [10].

The deep learning method proposed is automatic detection of pulmonary TB using a Convolutional Neural Network. The Results are classification of normal and TB chest x-rays. Using laboratory results in the form of examination of Xpert® MTB / RIF and or Acid Resistant Bacteria as the gold standard [11, 12], Deep learning with CNN is expected to be able to automatically detect pulmonary TB on digital images of the chest x-rays with high performance. The results of the classification of deep learning models are compared with radiologists (expert) to get a classification similarity.

MATERIALS AND METHODS

Dataset preparation

The study involved 2,026 digital chest x-rays images taken retrospectively from computer radiographs at two hospitals for the period 2018 to 2019. Obtaining chest x-rays images, 450 were confirmed as Tuberculosis and 360 were normal so that the total sample was 810. TB samples were confirmed with Xpert® MTB / RIF or Acid Resistant Bacteria test while the normal chest x-rays was validated by two experts.

The digital chest x-rays images imported according to medical standards are DICOM. DICOM is not only an image format but also a standard for data transfer, storage, and communication protocols between medical devices. DICOM image file consists of headers containing raw data and metadata [13]. Besides DICOM images on radiographic computers have a matrix of 1024 x 1024 with a data capacity of 7 to 10 MB. (6) If processed in one CNN hidden layer node, there are 1024 x 1024 x 3 = 3,145,728 parameters.

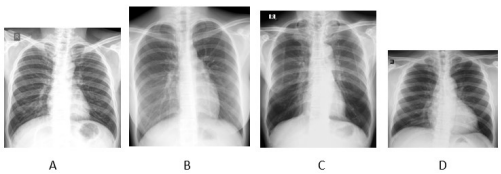


Figure 1 A. Stenic, B. Hyposthenic, C. Asthenic, D. Hyper- Sthenic

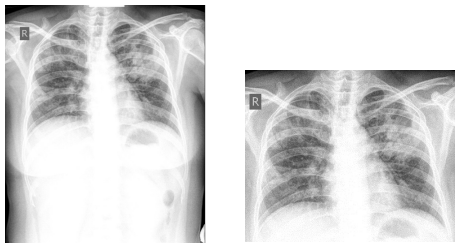
This is certainly very burdensome in the learning process. Therefore in this study downscale was carried out in the image pre-processing process. For

encoding, images in the DICOM format it is converted to JPEG.

The digital chest x-ray images obtained in the form of DICOM and have different irradiation areas. Also in general, the shape of the human body has several variations, known as body habitus. In the collected digital images of the chest x-rays image, four types of body habitus are obtained, namely the stenic, hyposthenic, asthenic, and hypersthenic types. Body habitus affects the size of the shape, position, and movement of internal organs [14].

Image pre-processing

The ImageJ application was used to crop the image according to the lung field area and convert it to JPEG format. This is done to improve readability in the deep learning model. The image size was standardized to 400 x 400 to further lighten the work of the model. The images are labelled in two different files, TB and Normal.



A B

Figure 2. A. Original image, B. Image of the cropping result

Development of the deep learning model

Developed a deep learning CNN model using Visual Geometry Group (VGG) architecture [9] with Python programming language version 3.7. The library learning machine used by Tensor Flow 2.0 is supported by Keras [15].

Deep learning architecture

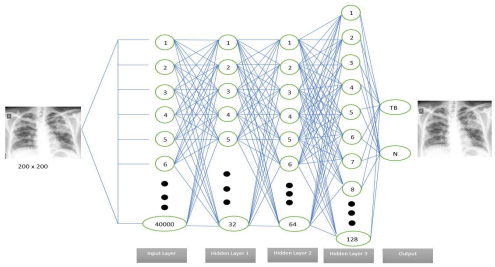


Figure 3. Deep learning model with 3 hidden layers

The Deep Learning model was composed of three hidden layers, consisting of 32 convolutions filters with the desired filter size (3 x 3 or 5 x 5) and 32 max pool, relu activation, kernel initializer "he-uniform", padding "same" in the first layer, 64 convolutions filters, and 64 max pool, relu activation, kernel initializer "he-uniform", padding "same" in the second layer, 128 convolutions filters, and 128 max pool, activation, kernel initializer "he-uniform", padding "same" in the third layer. One Flatten and two dense in the last layer. Batch size 20 and epoch 50.

Augmentation was used as an image data generator, where the width shift range is 0.1, and the height shift range is 0.1. For a horizontal flip, the setting is "true" or no flipping is done.

The deep learning model was developed using a variety of filter dimensions namely 3 x 3 and 5 x 5, the input image size of 50 x50, 100 x 100 and 200 x 200. Thus there are six deep learning model namely xray50_3x3, xray50_5x5, xray100_3x3, xray100_5x5, xray200_3x3 and xray200_5x5.

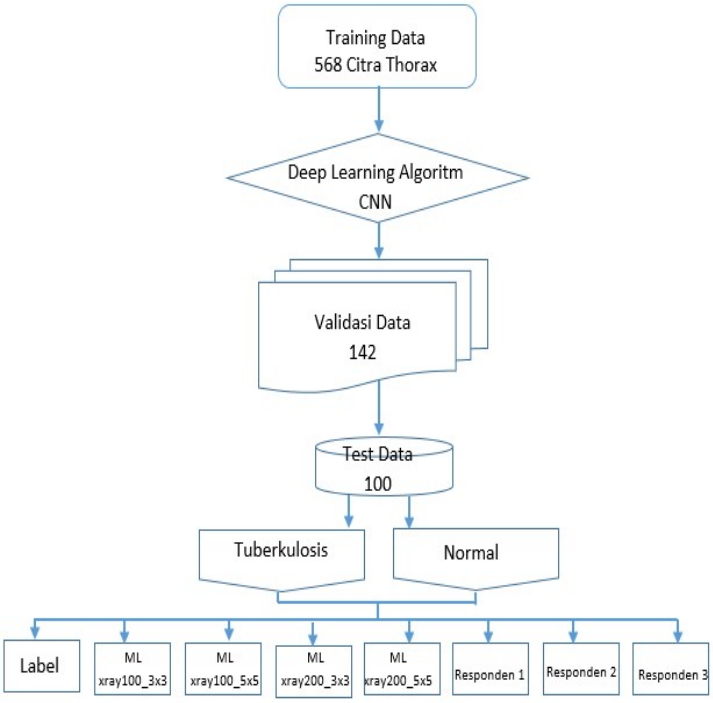


Figure 4. Data flow training, validation, and testing

Deep learning model training

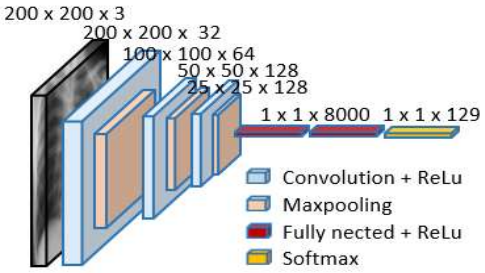


Figure 5. CNN Deep Learning Architecture with 200 x 200 input images

An input image of 200 x 200 sizes was convoluted to a 5x5 filter 32 times, activated by relu, and subjected to a 2x2 maxpool size in the first layer, resulting a 100x100 feature map. The feature map was the input for the second layer. In the second layer was convoluted to a 5x5 size filter 64 times, activated with Relu. The convolution result, Simplified by 2x2 maxpool size produce a 50x50 feature map. The feature map that produced by the second layer becomes the third layer input. In the third layer, it was convoluted with a 5x5 filter size 128 times, activated by relu. Simplified by 2x2 maxpool size produce a 25x25 feature map.

The output of the third layer is a 25x25 two-dimensional matrix and then enters the flatten layer. The flatten layer converts a two-dimensional matrix into a vector. This vector is then entered into the dense layer for the classification process.

To produce a model with high performance, training process uses the Stochastic Gradient Descent Optimizer (learning rate 0.001 and momentum 0.9). Evaluate loss models using cross-entropy. Softmax was used for classification purposes.

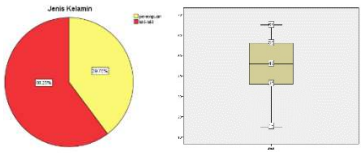
In this study, the performance of the deep learning model was assessed, namely Cross-validation, Diagnostic Test or Confusion Matrix, ROC curves, and classification speed. Three experts were taken as a comparison.

Data analysis

This study compares the CNN deep learning models and expert performance. The data obtained are categorized as a nominal scale, namely TB and normal. Variables are two research groups namely CNN deep learning and expert. Measurements were made once to 100 test data in the form of normal and TB images. Tests are categorized as unpaired comparisons. So the statistical test is Chi-square with the condition that the cell has an expected value of less than 5, a maximum of 20% of the number of cells [14].

RESULTS

Sample characteristic



A B

Figure 6. A. Gender frequency, B. Age distribution

The graph shows that the sexes of women are 39.75% (322), and men are 60.25% (488). The Minimum age was 15 and the maximum was 65 years, while the average age is 45.25 and the standard deviation is 12.09. Normality test data using the skewness value and standards error that is 0.2, the value is ≤ 2 then the sample is normally distributed.

Respondent's characteristics

To get a comparison of the results of the Deep learning classification, the test image was assessed by three respondents who have experienced working as experts for more than 5 years. The characteristics of respondents in this study are:

Table 1. Characteristic of Respondents

Respondent	Experience	Position
1	5*	Expert
2	11*	Expert
3	6*	Expert

*in year

Interrater reliability observer test

The reliability test using the percent agreement obtained the level of agreement between Respondents 1 and 2, 1 and 3, and 2, and 3 are 96%, 97%, 97%. (15) The reliability test with Cohen's kappa obtained the level of agreement between respondents 1 and 2, 1 and 3, 2, and 3 are 91.7%, 93.8%, 93.7%. Cohen's kappa test results obtained in the range between 0.90 - 1.00. Thus the level of agreement is expressed as an almost perfect agreement [16].

Classification results by the deep learning model

Softmax was used for classification purposes; the similarity value was set between 0 and 1, meaning that it is increasingly similar to a training image, the value approaches 1. The classification result that appears is the largest value, for example, if the input test image has a value of 54% similar to normal and 100 % is similar to the TB training image, the classification result that appears in the image is 100% TB.

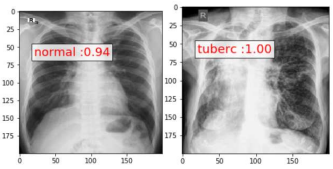


Figure 7. Classification result A. Normal, B. TB

Deep learning model performance

Cross-validation

Table 2. Accuracy and error rate deep learning model

Performance	xray50		xray100		xray200	
	3x3	5x5	3x3	5x5	3x3	5x5
Accuracy	0.90226	0,91729	0,91068	0,93115	0,93706	0,95884
Error	0,09774	0,08271	0,08932	0,06885	0,06294	0,04116

Diagnostic test

Diagnostic Test or Confusion Matrix was done after the deep learning model goes through the training and validation stages. The results of the classification by the deep learning model and the gold standard are the results of the Molecular Rapid Test (Xpert® MTB / RIF) or examination of Acid Resistant Bacteria for TB and validated by two radiologists for normal chest x-rays used as a basis for calculating the performance test. The classification results of 100 test data by the deep learning model are arranged in following table.

Table 3. Deep learning model Performance Test

Performance	xray50		xray100		xray200	
Test	3x3	5x5	3x3	5x5	3x3	5x5
Accuracy	0,90	0,88	0,92	0,91	0,96	0,97
Sensitivity	0,867	0,833	0,9167	0,85	0,95	0,9667
Specificity	0,95	0,95	0,925	1	0,975	0,975
Precision	0,963	0,961	0,9483	1	0,9827	0,9831
NPV	0,826	0,792	0,8809	0,8163	0,9286	0,9512

ROC curve [17]

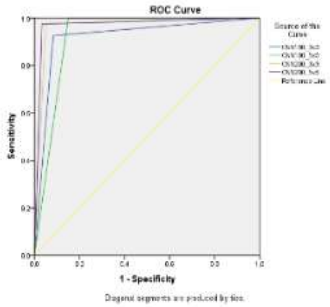


Figure 8. ROC curves of the deep learning model

Table 4. Performance of AUC and CI 95% deep learning models

Performance	xray50		xray100		xray200	
	3 X 3	5 X 5	3 X 3	5 X 5	3 X 3	5 X 5
AUC	0,908	0,892	0,921	0,925	0,963	0,971
CI95%	0,844-0,973	0,823-0,960	0,858-0,983	0,870-0,980	0,920-1,00	0,932-1,00

Classification speed

At the learning stage it was evaluated to obtain optimal performance. The learning time required for input image size of 50x50, 100x100 and 200x200 was 5 to 10, 15 to 30 and 60 to 90 minutes.

For the test used randomized 40 normal and 60 TB x-rays image. The duration of the classification process by all deep learning models takes 30 to 60 seconds. The classification results are in the form of normal or TB information with the percentage of matches.

Respondent's performance

The performance evaluation of respondents was carried out with 100 test data consisting of 40 normal and 60 TB chest x-rays (20 minor TB, 20 moderate TB, 20 extends TB). Test data are arranged randomly, which is the same data used to test the deep learning model. Chest x-rays images are in DICOM format, read using the RadiAnt application. The classification results are compared with the results of the Molecular Rapid Test (Xpert® MTB / RIF) or examination of Acid Resistant Bacteria for TB and validated by two radiologists for normal chest x-rays used as a basis for calculating the performance test.

Diagnostic test

The results of the classification by respondents were compared with the gold standard and arranged in a 2x2 table. Accuracy, sensitivity, precision, and negative predictive value were calculated. (18) Results like the following table:

Table 5 Respondent Performance Test

Performance	Respondents		
	1	2	3
Accuracy	0,97	0,99	0,98
Sensitivity	0,95	1	0,9833
Specificity	1	0,975	0,975
Precision	1	0,9836	0,9833
NPV	0,9302	1	0,975

ROC curve [18, 19]

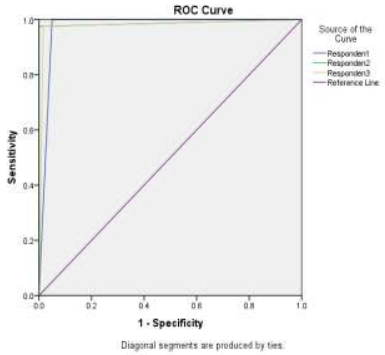


Figure 9. ROC curves of respondents

Table *6.* Performance of AUC and CI 95% respondents

Performance	Respondents		
	1	2	3
AUC	0,975	0,987	0,979
CI95%	0,942 - 1,00	0,959 - 1,00	0,945- 1,00

Classification speed

The duration of the classification process by respondents varies. Respondents 1, 2 and 3 are 5, 4, 4 minutes. The second respondent has highest performance, therefore used as comparison to the CNN deep learning model.

DISCUSSION

Comparison of the deep learning model and the second respondent performance

This comparative study was comparing the performance between CNN deep learning models with expert. The data obtained are categorized as a nominal scale, TB and normal. Variables are two research groups, CNN deep learning model and expert. Measurements were made once to 100 test data. Tests are categorized as unpaired comparisons. So the statistical test is chi-square with the condition that cells have an expected value of less than 5, a maximum of 20% of the number of cells. If the chi-square requirements are not met then the Fisher test is used as an alternative [14].

The second respondent and Xray50_3x3 deep learning model

Table 7. Table of Chi-Square test results between second respondent and xray50_3x3 deep learning model

		Second Respondent				
		TB		Normal		ρ- Value
		O	E	O	E	
xray50_3x3	TB	53	32,9	1	21,1	0
	Normal	8	28,1	38	17,9	
	Total	61	61	39	39	

Statistical test results show that the table 2x2 is worth testing with Chi-Square because there is no expected value of less than 5 with a minimum expected count of 17.9.

Obtained a continuity correction value of 64.743, the chi-square table value for df 1 with α = 0.05 is 3.84146. Asymp Sig. Value (2 sided) or a significance of 0,000. Contingency coefficient 0.636. The difference in normal and tuberculosis classification between the second respondent and the xray50_3x3 deep learning model is 9%.

From the data it's known that $x^2_{count} > x^2_{table}$ (64.743> 3.84146) or significance value <α (0.000 <0.05) so that the alternative hypothesis applies. Thus it can be concluded that there is a similarity between the classification by experts and deep learning models with the level of closeness of a relationship of 0.636.

The second respondent and xray50_5x5 deep learning model

Table 8. Table of Chi-Square second respondent test results and the deep xray50_5x5 model

		Second Respondent				
		TB		Normal		ρ- Value
		O	E	O	E	
xray50_5x5	TB	51	31,7	1	20,3	0
	Normal	10	29,3	38	18,7	
	Total	61	61	39	39	

Statistical test results show that the table 2x2is worth testing with Chi-square because there is no expected value of less than 5 with a minimum expected count of 18.7.

Obtained a continuity correction value of 59.395, the chi-square table value for df 1 with α = 0.05 is 3.84146. Asymp Sig. Value (2 sided) or a significance of 0,000. Contingency coefficient 0,620. The difference in the normal and tuberculosis classification between the second respondent and the xray50_5x5 deep learning model is 11%.

From the data it's known that $x^2_{count} > x^2_{table}$ (59.395> 3.84146) or significance value <α (0.000 <0.05) so that the alternative hypothesis applies. Thus it can be concluded that there is a similarity between the classification by experts and deep learning models with the level of closeness of a relationship of 0.620.

The second respondent and xray100_3x3 deep learning model

Table 9Table of Chi-Square test results between second respondent and xray100_3x3 deep learning model

		Second Respondent				
		TB		Normal		*ρ*- Value
		O	E	O	E	
xray100_3x3	TB	56	35,4	2	22,6	0
	Normal	5	25,6	37	16,4	
	Total	61	61	39	39	

Statistical test results show that the 2x2 table is worth testing with Chi-Square because there is no expected value of less than 5 with a minimum expected count of 16.4.

Obtained a continuity correction value of 69.853, the chi-square table value for df 1 with α = 0.05 is 3.84146. Asymp Sig. Value (2 sided) or a significance of 0,000. Contingency coefficient 0.651. The Difference in the normal and tuberculosis classification between the second respondent and the xray100_3x3 deep learning model is 7%.

From the data, it's known that $x^2_{count} > x^2_{table}$ (69.885> 3.84146) or significance value <α (0.000 <0.05) so that the alternative hypothesis applies. Thus it can be concluded that there is a similarity between the classification by experts and deep learning models with the level of closeness of a relationship of 0.651.

The second respondent and xray100_5x5 deep learning model

Table 10. Table of Chi-Square test results between second respondent and xray100_5x5 deep learning model

		Second Respondent				
		TB		Normal		*ρ*- Value
		O	E	O	E	
xray100_5x5	TB	51	31,1	0	19,9	0,000
	Normal	10	29,9	39	19,1	
	Total	61	61	39	39	

Statistical test results show that the table 2x2 is worth testing with Chi-Square because there is no expected value of less than 5 with a minimum expected count of 19.11.

Obtained correction value 63,240, the chi-square table values for df 1 with α = 0.05 is 3.84146. Asymp Sig. Value (2 sided) or a significance of 0.000. Contingency coefficient 0,632. The difference in the normal and tuberculosis classification between the second respondent and the xray100_5x5 deep learning model is 10%.

From the above data, it is known that $x^2_{count} > x^2_{table}$ (63.240> 3.84146) or significance value <α (0.000 <0.05) so that alternative hypotheses applies. It can be concluded that there are similarities between the classification by experts and models deep learning with a relationship level of 0.632.

The second respondent and Xray200_3x3 deep learning model

Table 11. Table Chi-Square test results between second respondent and xray200_3x3 deep learning model

		Second Respondent				
		TB		Normal		ρ- Value
		O	E	O	E	
Xray200_3x3	TB	58	35,4	0	22,6	0,000
	Normal	3	25,6	39	16,4	
	Total	61	61	39	39	

Statistical test results show that the 2x2 table is worth testing with Chi-Square because there is no expected value of less than 5 with a minimum expected count of 16.38.

Obtained a continuity correction value 84.430. The chi-square table values for df 1 with α = 0.05 is 3.84146. Asymp Sig. Value (2 sided) or a significance of 0.000. Contingency coefficient 0.685. The difference in normal and tuberculosis classification between the second respondent and xray200_3x3 deep learning model is 3%.

From the data, it is known that $x^2_{count} > x^2_{table}$ (84.430> 3.84146) or significance value <α (0,000 <0.05) so that the alternative hypotheses applies, and it can be concluded that there is a similarity between the classification by radiologists and models deep learning with a relationship level of 0.685.

The second respondent and Xray200_5x5 deep learning model

Table 12. Table of Chi-Square test results between second respondent and xray200_5x5 deep learning model

		Second Respondent				
		TB		Normal		ρ- Value
		O	E	O	E	
Xray200_5x5	TB	59	36,0	0	23,0	0,000
	Normal	2	25,0	39	19,1	
	Total	61	61	39	39	

Statistical test results show that the 2x2 table is worth testing with Chi-Square because there is no expected value of less than 5 with a minimum expected count of 18.3.

Obtained a continuity correction value of 80.981. The chi-square table value for df 1 with α = 0.05 is 3.84146. Asymp Sig. Value (2 sided) or a significance of 0.000. Contingency coefficient 0.677. The difference in normal and tuberculosis classification between the second respondent and the xray200_5x5 deep learning model is 2 %.

From the above data, it is known that $x^2_{count} > x^2_{table}$ (80.981> 3.84146) or significance value <α (0,000 <0.05) so that alternative hypotheses apply and it can be concluded that there are similarities between the classification by radiologists and models deep learning with a relationship level of 0.677.

All deep learning model has the same classification as experts. The deep learning model xray200_5x5 model with an input image size of 200x200 and a filter size 5x5 has the highest level of similarity 98%.

In the deep learning model, the normal image classification results obtained 0.925 - 1 following the gold standard. The average similarity was between 0.88 and 0.97 on training data. Minor, moderate and extends TB classification were 0.5- 0.9, 0.85-1, 1 following the gold standard. The average similarity were 0.851 - 0.93, 0.885 - 0.985, 0.98 – 1 on training data.

The lowest similarity was minor TB and the highest similarity was extends TB. This shows that the more features of TB images, the more easily recognized. The input image size affects the performance of the deep learning model. The greater resolution, the accuracy tends to increase. The greater of image resolution, less image information is lost, conversely, the smaller image size more image information is lost.

Image compression with lossless, data may be compressed into half or a quarter. Compression that exceeds that is called lossy. In deep learning or machine learning the use of lossy compressed images decreases the amount of information from the image so that there is the potential for error reading or prediction (12), however, in machine learning or deep learning models the use of lighter image sizes in learning uses smaller image sizes. To prevent reading or prediction errors, the deep learning model is controlled by the gold standard, in this case, the Molecular Rapid Test (Xpert® MTB / RIF) or examination of Acid Resistant Bacteria.

The limitation of this study is the ability of deep learning model classification is only one diagnosis on the chest x-rays image, namely tuberculosis. The deep learning model can be applied clinically but it is better if the development of multi-class classification capabilities or diagnoses such as pneumonia, asthma, bronchitis, emphysema, and COVID-19 is done.

CONCLUSION

Filter and image size affect the performance of the deep learning model. The resulting deep learning model with an image size of 200 x 200 and a filter 5 x 5 has a sensitivity and specificity of 96.67% and 97.5% and has a 98% classification similarity to expert and faster in classifying.

REFERENCES

[1]. BañUls AL, Sanou A, Van Anh NT, Godreuil S. Mycobacterium tuberculosis: Ecology and evolution of a human bacterium. J Med Microbiol. 64(11), 2015, 1261–9.
[2]. Magnabosco GT, Lopes LM, Andrade RL de P, Brunello MEF, Monroe AA, Villa TCS. Tuberculosis Control in People Living With HIV/AIDS. Rev Lat Am Enfermagem. 24(e2798), 2016, 1–8.
[3]. Anderson L, Baddeley A, Monica Dias H, Floyd K, Baena IG, Gebreselassei N, Global Tuberculosis Report. Geneva: World Health Organization; 2018.
[4]. Kowalczyk N. Radiologic Pathology for Technologists. Sixth Edit. Ohio: Elsevier Mosby; 2014, 472.
[5]. Reiner BI, Krupinski E. The insidious problem of fatigue in medical imaging practice. J Digit Imaging. 25(1), 2012, 3–6.
[6]. Muenzel D, Engels HP, Bruegel M, Kehl V, Rummeny EJ, Metz S. Intra- and inter-observer variability in measurement of target lesions: Implication on response evaluation according to RECIST 1.1. Radiol Oncol. 46(1), 2012, 8–18.
[7]. Norweck JT, Seibert JA, Andriole KP, Clunie DA, Curran BH, Flynn MJ, ACR-AAPM-SIIM technical standard for electronic practice of medical imaging. J Digit Imaging. 26(1), 2013, 38–52.
[8]. Soffer S, Ben-Cohen A, Shimon O, Amitai MM, Greenspan H, Klang E. Convolutional Neural Networks for Radiologic Images: A Radiologist's Guide. Radiology. 290(3), 2019, 590–606.
[9]. Sarıgül M, Ozyildirim BM, Avci M. Differential convolutional neural network. Neural Networks. 116, 2019, 279–87.
[10]. Lee JG, Jun S, Cho YW, Lee H, Kim GB, Seo JB, Deep learning in medical imaging: General overview. Korean J Radiol. 18(4), 2017, 570–84.
[11]. Steingart KR, Schiller I, Horne DJ, Pai M, Boehme CC, Dendukuri N. Xpert ® MTB / RIF assay for pulmonary tuberculosis and rifampicin resistance in adults (Review) Xpert ® MTB / RIF assay for pulmonary tuberculosis and rifampicin resistance in adults. Cochrane Libr. (1), 2014, 1–3.
[12]. Tuberculosis Coalition for Technical Assistance. Handbook for Using International Standar for Tuberculosis Care. USAID, editor. World Health Organization; 2007.
[13]. Pianykh OS. Digital Imaging and Communications in Medicine (DICOM). Second Edi. New York: Springer; 2012, 23.
[14]. Bruce W, Rollins J. Merrill ' S Atlas of Radiographic Positioning & Procedures. Thirteenth. St. Louis: Elsevier Mosby; 2016.
[15]. Raschka S. Python Machine Learning. Birmingham: Packt Publishing Ltd; 2016, 425.
[16]. McHugh ML. Lessons in biostatistics interrater reliability : the kappa statistic. Biochem Medica. 22(3), 2012, 276–82.
[17]. Huang J, Ling CX. Using AUC and accuracy in evaluating learning algorithms. IEEE Trans Knowl Data Eng.

17(3), 2005, 299–310.
[18]. Hajian K. Receiver Operating Characteristic (ROC) Curve Analysis for Medical Diagnostic Test Evaluation. Casp J Intern Med. 4(2), 2013, 627–35.
[19]. Burgess AE. Visual perception studies and observer models in medical imaging. Semin Nucl Med. 41(6), 2011, 419–36.
Available from: http://dx.doi.org/10.1053/j.semnuclmed.2011.06.005

How to cite this article: Mochamad Imron, Melyana Nurul Widyawati, Siti Masrochah. Tuberculosis detection using convolutional neural network. Int J of Allied Med Sci and Clin Res 2020; 8(3): 535-546.
Source of Support: Nil. **Conflict of Interest:** None declared.