International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18 (2026) pp. 1-11

https://doi.org/10.70917/ijcisim-2026-3935
Article

Research on the Construction of an Artificial Intelligence Precision Diagnosis Model for Pulmonary Tuberculosis Based on Multi-modal Data Fusion

Haoyang Lan [1,*]

[1] College of Laboratory Medicine, Chongqing Medical University, Chongqing, 400016, China
* Correspondence author: l2812800528@163.com

Abstract: Objective: To evaluate the value of a deep learning-based CT-assisted diagnosis model for pulmonary tuberculosis in clinical applications. Methods: The case data of 2000 patients who underwent chest CT plain scans at the Chongqing Public Health Medical Treatment Center from March 2021 to July 2024 were retrospectively collected. The patients were divided into the normal lung group (1300 cases), the common pulmonary infection group (100 cases), and the pulmonary tuberculosis group (600 cases). The data set was divided into the training set (1400 cases, 70.0%) and the test set (600 cases, 30.0%) by random grouping (achieved through the sample function of R language for complete random grouping of the training set and test set). All images used the lung field automatic segmentation algorithm to obtain the lung field area. Further, the DenseNet algorithm based on the mixed-domain attention was used for classification research. The classification performance of the model was evaluated using the area under the curve (AUC), sensitivity, specificity, and accuracy. Finally, in the test data, the optimal model was compared with the diagnostic results of three radiologists of different seniority. Results: The diagnostic performance of the model in the training set and test set was AUC = 0.899, 0.831; ACC = 89.4, 84.1%. Among them, the diagnostic performance of the improved DenseNet model was superior to that of junior doctors (accuracy rates were 90.9% and 89.4%, P = 1.000, Kappa = 0.679), and was highly consistent with the diagnostic levels of middle-aged and senior doctors (accuracy rates were 90.7%, 92.4%, and 95.7%, Kappa values were 0.749 and 0.821). Conclusion: The DenseNet model can accurately identify secondary pulmonary tuberculosis and is comparable to the diagnostic level of middle-aged radiologists, and can be used as an auxiliary diagnostic tool for secondary pulmonary tuberculosis.

Keywords: Pulmonary tuberculosis; Artificial intelligence; Precision diagnostic model; DenseNet algorithm

1. Introduction

Tuberculosis is a chronic respiratory infectious disease caused by Mycobacterium tuberculosis, characterized by a long course of the disease, strong infectivity, and high recurrence rate. It is one of the major public health problems threatening human health worldwide [1-3]. In China, the prevention and control situation of tuberculosis is equally severe. Although the incidence and mortality of tuberculosis have decreased in recent years through the implementation of modern tuberculosis control strategies, due to the large population base, a large number of mobile populations, and an increase in drug-resistant tuberculosis, the prevention and control of tuberculosis still faces many challenges [4-7].

Chest imaging examinations are important means for the diagnosis and screening of tuberculosis. Among them, chest X-ray and chest CT examinations are the most commonly used imaging methods in clinical practice [8-9]. However, the traditional manual reading method has many limitations, such as low reading efficiency, strong subjectivity, and high rates of missed diagnosis and false diagnosis, especially in grassroots medical institutions, where the shortage of imaging professionals and uneven diagnostic levels have severely restricted the early detection and diagnosis of tuberculosis. With the rapid

development of artificial intelligence (AI) technology, its application in medical imaging diagnosis is becoming increasingly widespread [10]. In recent years, the application of AI technology in medical imaging diagnosis has not been a new phenomenon, and it has achieved significant improvements in the diagnostic accuracy of single-modal data [11-12]. However, medical imaging diagnosis often requires the integration of various types of information to make accurate judgments. The AI diagnosis model for tuberculosis based on multi-modal data fusion can integrate different modalities of information, improve diagnostic accuracy, and provide more comprehensive clinical decision support [13-15]. Multi-modal data refers to different types of tuberculosis data from different sources and modalities, such as medical images, text, physiological signals, etc. [16-17]. The use of multi-modal data can provide more comprehensive and accurate information, thereby helping AI models to better perform tuberculosis diagnosis [18-19].

Deep learning is a new branch of artificial intelligence, different from traditional imaging omics. Deep learning can directly complete the mapping from images to features and support automatic feature extraction. This study proposes a DenseNet algorithm based on mixed-domain attention using chest CT image data. DenseNet solves the problem of gradient disappearance by reusing features and having a closely connected network. Based on the attention mechanism of human eyes, it combines the bottom-up and top-down structures to achieve spatial attention. Based on the multi-channel characteristics of human vision, it focuses on the channel domain and ignores the information in the channel domain, studies mixed-domain attention, and introduces it into DenseNet. Therefore, this study mainly adopts the improved DenseNet algorithm to construct an auxiliary diagnosis model for secondary tuberculosis and evaluates its clinical application value, while using other network models for comparative analysis.

2. Materials and Methods

2.1. General Information

This study retrospectively collected the imaging and clinical data of patients who visited the Imaging Center of Chongqing First People's Hospital from March 2021 to July 2024.

Inclusion criteria: ① Patients diagnosed with pulmonary tuberculosis by sputum culture, patients with non-tuberculous pulmonary abnormalities, and normal control patients; ② Complete X-ray chest imaging data; ③ Clear X-ray chest imaging images; ④ Patients obtained anteroposterior chest X-rays in a standing position at the time of diagnosis. Exclusion criteria: ① Incomplete clinical diagnosis records; ② Patients with diseases that significantly affect the results of chest X-ray examinations (such as lung cancer, pacemaker or defibrillator placement, history of heart or lung surgery); ③ Patients with extensive pleural effusion.

Based on the above criteria, a total of 2000 patients were included, including 1170 males and 830 females, with an average age of 56 years. Among them, there were 600 tuberculosis cases, 1300 normal cases, and 100 cases in the general pulmonary infection group. The patients were randomly divided into two datasets, and were trained and tested on the improved DenseNet network for pulmonary tuberculosis lesion detection model in a ratio of 7:3.

2.2. Research Methods

(1) CT Scan

The scan was performed using the SOMATOM Definition AS 64-slice spiral CT scanner. The parameters were set as follows: slice thickness and slice interval were both 5 mm, thin slice was 0.625 mm, voltage was 120 kV, and current was 180 - 240 mA. The patient was in a supine position, with both arms raised, and the entire lung was scanned. The reconstruction slice thickness was 0.625 mm, and the lung and mediastinum windows were observed.

(2) Image Preprocessing

The original CT image data was converted to a unified format before inputting the algorithm and data cleaning was performed. The window width and window level values were adjusted within the range of -1200 to 600 HU, and the image intensity was normalized to $[0,1]$ to standardize the CT image.

Lung field segmentation grid. Using 3D nested UNet, all CT data was standardized to form an image of $256\times256\times64$ pixels (pixel). Two radiologists with more than 5 years of experience independently evaluated the image segmentation effect after pre-training. The 3D nested UNet network structure is shown in Figure 1. Residual block is the residual block, dropout is the dropout layer, pooling is the pooling layer, conv block + batch normalization + activation is the convolution block + batch normalization + activation function, conv block 3×3×3 is a 3×3×3 convolution block, full connect is the fully connected layer, down-sampling is the down-sampling, up-sampling is the up-sampling, and skip

connection is the skip connection.

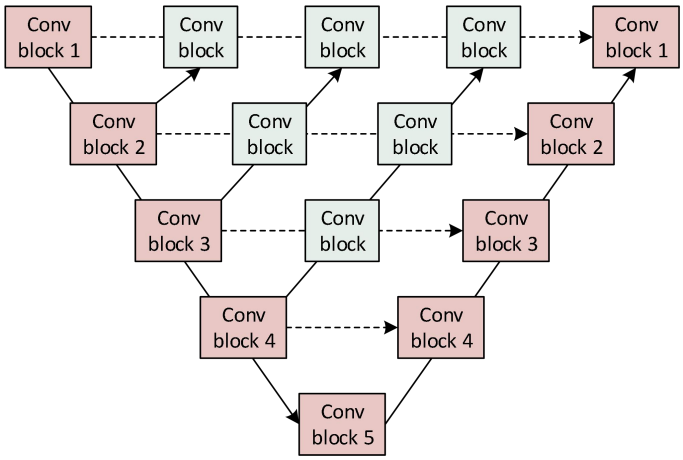

Figure 1. 3D nested net network structure

(3) Precise Diagnosis Model Based on DenseNet and Hybrid Domain Attention

(a) DenseNet

As the number of layers in CNN networks increases, problems such as gradient vanishing and model degradation occur. DenseNet breaks away from the conventional thinking of deepening network layers (ResNet) and widening network structures (Inception) to improve network performance. From the perspective of features, by reusing features and setting bypasses, it significantly reduces the number of network parameters and alleviates the problem of gradient vanishing [20]. Based on the assumption of information flow and feature reuse, it has fewer parameters, closely connected with regularization effects, and thus has a certain inhibitory effect on overfitting. DenseNet is mainly composed of multiple Dense Blocks, and its framework is shown in Figure 2.

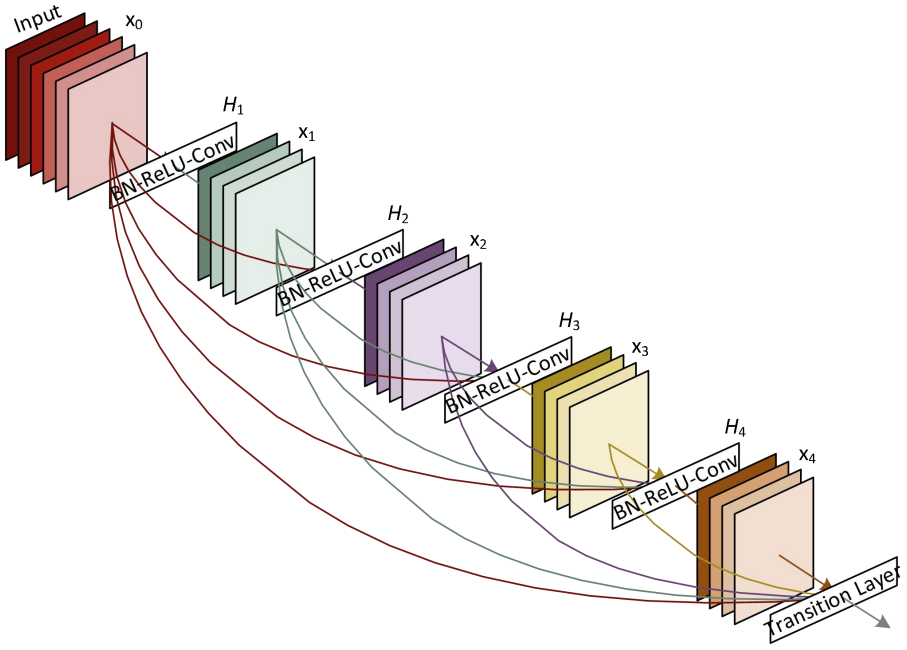

Figure 2. Five-layer DenseNet network

As shown in Figure 2, in the traditional convolutional neural network, if there are L layers, then there will be L connections. However, in DenseNet, there will be $L(L+1)/2$ connections. To put it simply, the input of each layer comes from the outputs of all previous layers. In the figure, x_0 is the input, x_1 is the output of H_1, the input of H_1 is x_0, and the input of H_2 is x_0 and x_1, that is:

$$x_l = H_l\left(\left[x_0, x_1, \cdots, x_{l-1}\right]\right) \tag{1}$$

Among them, l represents the layer, x_l represents the output of the l th layer, and $\left[x_0, x_1, \cdots, x_{l-1}\right]$ represents the channel merging of the output feature maps from layers 0 to $l-1$, similar to Inception. $H_l(\cdot)$ represents the combined function of three operations, which are batch normalization (BN), rectified linear unit (ReLU), and convolution operation.

The connection method of DenseNet makes the transmission of features and gradients more effective, and the network is easier to train. The deeper the network, the more likely the problem of gradient vanishing will occur, which is caused by the transmission of input information and gradient information between many layers. The dense connection of DenseNet is equivalent to each layer directly connecting to the input and the loss function, so this phenomenon can be alleviated when the network is deeper.

(b) Attention Module

The human visual system has multi-channel characteristics. Spatial domain attention ignores the information in the channel domain and treats the image features in each channel equally. This is not very interpretable when applied to other layers of neural networks. Channel domain attention directly performs global average pooling on the information within a channel and ignores the local information within each channel. To address this issue, this paper designs a mixed-domain attention mechanism model.

The spatial attention mechanism is implemented using a bottom-up-top-down structure. The attention network is used to learn the regions activated by high-level features, simulating the attention characteristics of human visual perception, without the need to design an additional attention mechanism for the network. Other networks that use attention often require adding a branch to extract attention and conducting separate training on it. However, this structure can extract the attention of the model during the forward process, making the model training simpler. First, through a series of convolutions and downsampling, the high-level features are gradually extracted and the model's receptive field is enlarged. The activated pixels in the high-level features can reflect the region where the attention is located, and then the attention map is enlarged to the same size as the original input by the same number of upsampling operations. The upsampling is achieved through bilinear interpolation. This maps the attention region to each pixel of the input, obtaining the attention map.

This paper introduces a channel attention mechanism for generating feature maps of different scales used to generate candidate regions. Instead of averaging the feature information of all channels in the feature map, different weights are assigned to each channel of the feature map.

The number of non-zero elements in the feature plane of the k th channel can be expressed as:

$$M_k = \sum\nolimits_{ij} 1\left[\lambda_{ij}^{(k)} > 0\right] \tag{2}$$

Among them, M_k represents the number of non-zero elements in the feature plane of the k th channel.

In channel weighting, the weight of the k th channel is defined as:

$$B_k = \log\left(\frac{\sum_{l=1}^{L} M_l}{\varepsilon + M_k}\right) \tag{3}$$

Here, b_k represents the weight of the k th channel; M_l represents the total sum of non-zero elements in the L th layer; ε represents a normal quantity to maintain the stability of the fraction and ensure that the denominator is not zero, with $\varepsilon = 1$.

(c) DenseNet with mixed-domain attention

To sum up, based on the spatial domain attention map, each channel's feature map is multiplied by the corresponding weight to achieve mixed-domain attention. Then, mixed-domain attention is introduced into the Densenet Block, implemented by point-wise multiplication with the feature map of each layer's output of the Densenet Block, forming a mixed-domain attention DenseNet, abbreviated as MATDenseNet. The mixed attention module is shown in Figure 3. Here, $\otimes$ represents the point-wise multiplication operation.

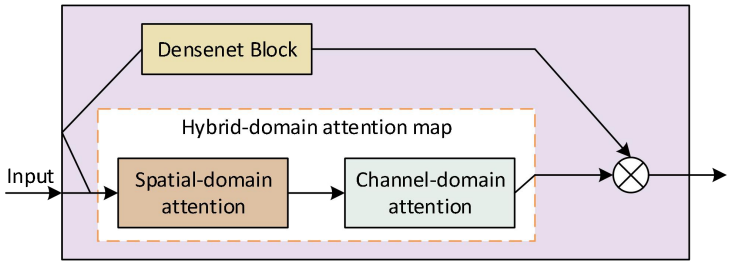

Figure 3. Mixed domain attention module

2.3. Evaluation Indicators

Since the data in the used dataset is divided according to patients, re-dividing may result in CT images from the same patient being present in both the training set and the test set, which would affect the accuracy. To ensure the objective and accurate nature of the experiment, the original dataset's division was used for the classification experiment. To evaluate the method proposed in this paper, accuracy (Accuracy), specificity (Specificity), sensitivity, and the AUC curve were selected as performance indicators. The true examples (TP), false positives (FP), true negatives (TN), and false negatives (FN) in the confusion matrix are the main components for calculating performance indicators. TP represents the number of cases where the true value is positive and the model considers it positive; FN represents the number of cases where the true value is positive but the model considers it negative; FP represents the number of cases where the true value is negative and the model considers it positive; TN represents the number of cases where the true value is negative and the model considers it negative. Accuracy and specificity are defined in Equations (4) to (5):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

$$Specificity = \frac{TN}{TN + FP} \tag{5}$$

2.4. Statistical Analysis

The clinical data and CT sampling conditions of the patients were analyzed descriptively using SPSS 26.0 software. Count data were described as "number, percentage/relative frequency (%)", and the differences between groups were compared using the χ^2 test or Fisher's exact probability method; for skewed distribution of measurement data, they were described as "median (quartiles) [M (Q1, Q3)]", and the differences between groups were compared using the rank sum test. The criterion for statistical significance was a two-sided P value < 0.05. The model efficacy was evaluated by calculating the area under the ROC curve (AUC), sensitivity, and specificity of the test set and independent validation set models.

3. Results and Analysis

3.1. Comparison of General Information

The demographic characteristics of the patients are shown in Table 1. The age of patients with tuberculosis was higher than that of non-tuberculosis patients (P < 0.001). The incidence of cough in patients with tuberculosis was higher (p = 0.025), while the incidence of chest pain in non-tuberculosis patients was higher (p = 0.009), and the differences were statistically significant. In addition, the age and gender distributions of the tuberculosis group and the non-tuberculosis group in the training set and the test set were studied. In both the training set and the test set, the average age of the same tuberculosis patients was greater than that of the non-tuberculosis group, and the difference was statistically significant.

Table 1. Demographic characteristics of tuberculosis and non-tuberculosis patients

Clinical feature	Total (n=1900)	Non-tuberculosis	Tuberculosis	P
Age	56.22±18.11	48.15±19.22	45.71±18.22	<0.001
Gender				0.081
Female	830	430	400	
Male	1070	420	650	
Hemoptysis	255	127	128	0.775
Cough	1455	520	935	0.025
Sputum	1550	770	780	0.705
Fever	1000	450	550	0.099
Gasp	275	175	100	0.078
Chest pain	20	15	5	0.009
Chest stuffy	300	168	132	0.633
Dyspnea	50	15	35	0.478
Fatigue	460	340	120	0.135
Night sweat	240	105	135	0.622
Physical reduction	460	250	210	0.508

3.2. Evaluation of Model Diagnostic Performance

The classification performance of each model on the training set and the test set is shown in Table 2. The ROC curves of each model on the training set are shown in Figure 4, and the ROC curves of each model on the test set are shown in Figure 5. The AUC values of each model in the training set are as follows: MATDenseNet model is 0.899, ResNet model is 0.841, Transformer model is 0.866, SENet model is 0.822, ShuffleNet model is 0.843, Swin-transformer model is 0.833, and DenseNet model is 0.831. The AUC values of each model in the test set are as follows: MATDenseNet model is 0.831, ResNet model is 0.818, Transformer model is 0.820, SENet model is 0.813, ShuffleNet model is 0.788, Swin-transformer model is 0.775, and DenseNet model is 0.739. It can be seen that the MATDenseNet model has the optimal AUC value.

Table 2. The classification properties of each model in the training set and the test set

Model	Training set (n=1400)				Test set (n=600)			
	AUC	ACC	Sensitivity	Specificity	AUC	ACC	Sensitivity	Specificity
ResNet	0.841	83.5%	83.3%	84.7%	0.818	76.6%	80.7%	75.6%
SENet	0.822	84.1%	84.5%	84.9%	0.813	77.7%	78.2%	78.3%
DenseNet	0.831	84.7%	86.2%	84.7%	0.739	76.6%	78.2%	76.6%
ShuffleNet	0.843	83.3%	82.2%	85.1%	0.788	79.1%	80.7%	78.9%
Transformer	0.866	84.7%	84.9%	85.5%	0.820	79.1%	78.6%	80.3%
Swin-trans	0.833	85.6%	85.5%	86.6%	0.775	80.3%	78.6%	82.5%
MATDenseNet	0.899	89.4%	87.9%	91.1%	0.831	84.1%	85.6%	83.7%

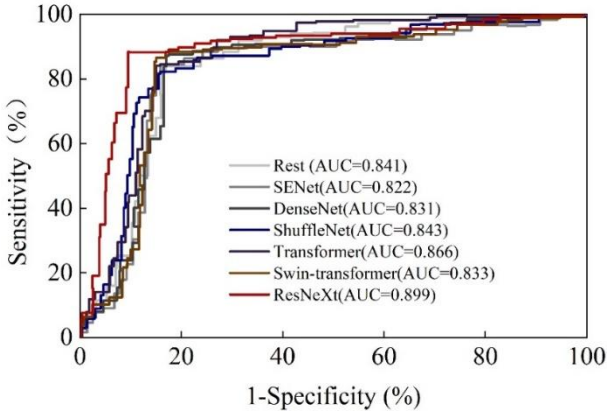

Figure 4. The model of the roc graph on the training set

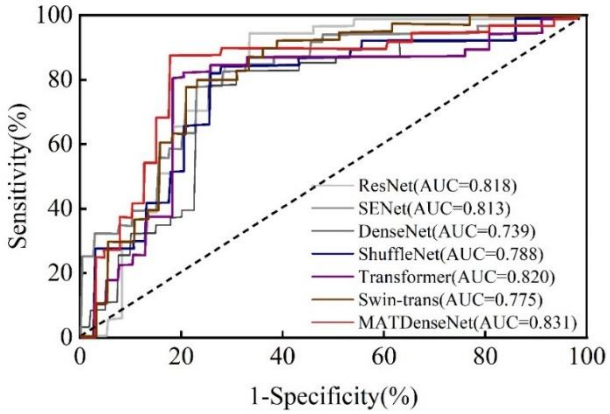

Figure 5. The model of the roc graph on the test set

3.3. Comparison of Diagnostic Performance between Artificial Intelligence Models and Radiologists

Using an independent test set, the performance of the most optimal artificial intelligence model was compared with that of three radiologists. Table 3 shows the average sensitivity, specificity, and accuracy of MATDenseNet and the three radiologists for the three-classification task. Figures 6 to 8 respectively illustrate the diagnostic performance of the MATDenseNet model and the three radiologists in the normal lung, common lung infection, and tuberculosis groups. Among these three indicators, the diagnostic performance of the MATDenseNet model was superior to that of junior (5-year) doctors, with no statistically significant difference (accuracy rates were 90.9% and 89.4% respectively, P = 1.000, Kappa = 0.679). It was comparable to that of mid-level (10-year) doctors, with a high degree of consistency (accuracy rates were 90.9% and 92.4% respectively, P = 0.375, Kappa = 0.749). There was no statistically significant difference in diagnostic performance compared to senior (15-year) doctors, and the consistency was excellent (accuracy rates were 90.9% and 95.7% respectively, P = 1.000, Kappa = 0.821). Radiologists with more experience had higher diagnostic accuracy for common lung infections and tuberculosis, but the sensitivity for identifying patients with common lung infections was only 74.9% and 83.4% respectively for mid-level and senior doctors, as shown in Figure 7.

Table 3. The comparison results of the overall evaluation of the three groups

Model	Sensitivity%	Specificity%	ACC%
MATDenseNet	85.5%	93.1%	90.9%
5 years old doctor	83.6%	92.5%	89.4%
10 years old doctor	88.5%	94.4%	92.4%
15 years old doctor	93.1%	96.6%	95.7%

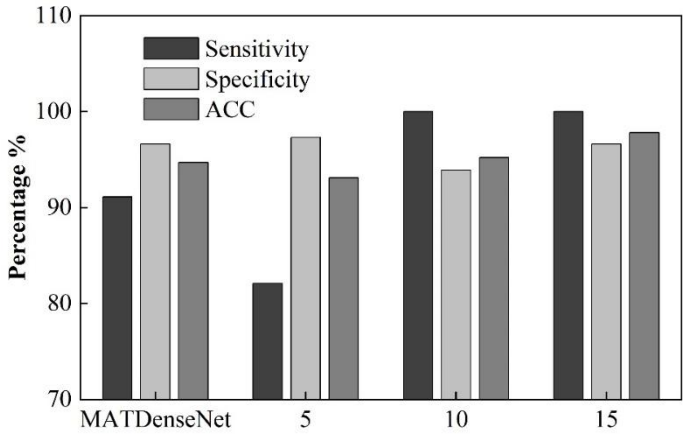

Figure 6. Comparison of the evaluation of lung normal group

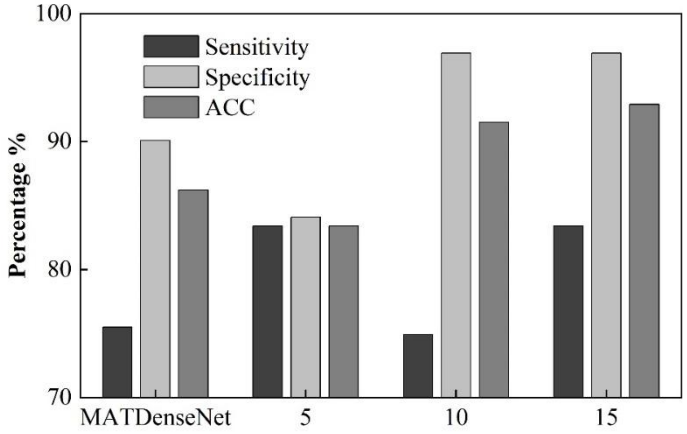

Figure 7. The comparison results of the general lung infection group evaluation

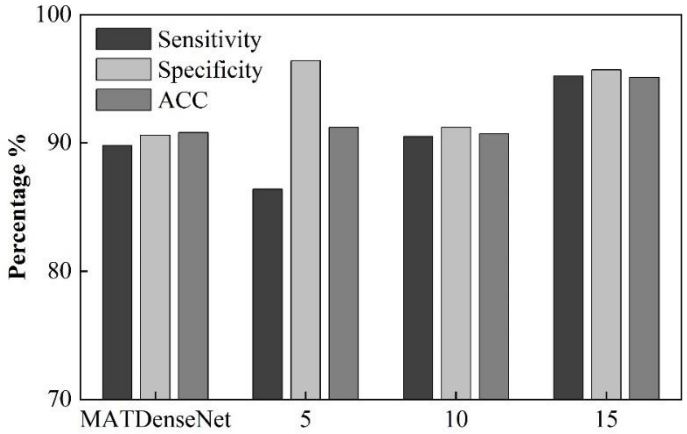

Figure 8. The comparison of the evaluation of the tuberculosis group

4. Discussion

With the development of the application of artificial intelligence technology in the field of medical images, many studies based on deep learning algorithms are dedicated to developing automatic detection systems for tuberculosis, and have achieved certain progress. Recent research indicates that deep learning methods can distinguish tuberculosis patients from normal lungs on chest X-rays, which is a major research direction for computer-assisted lung disease diagnosis in the future. This study explores the diagnosis of tuberculosis assisted by artificial intelligence and builds a classification model based on the deep learning framework (MATDenseNet) to analyze chest CT image data. The MATDenseNet model was verified for its diagnostic efficiency for tuberculosis patients, and it was concluded that the speed of artificial intelligence diagnosis is faster than that of radiologists.

Deep learning is an emerging branch of artificial intelligence. Its basic framework is a more complex and intelligent neural network model. Compared with traditional machine learning, the biggest advancement of deep learning is to use the method of automatically extracting features instead of manual feature extraction. Without manual marking, the machine can automatically identify the hidden information of the original image and automatically obtain higher-level features through a multi-layer neural network structure. In the field of pulmonary imaging, deep learning has demonstrated ideal effects in various medical scenarios such as pneumonia classification, interstitial lung disease classification, malignant risk grading of pulmonary nodules, and radiotherapy target delineation for lung cancer. There are also related studies in the diagnosis of NTM pulmonary disease. Wang et al. [21] used the ResNet network to build a binary classification model for NTM pulmonary disease and tuberculosis CT images. The image preprocessing process did not involve manual delineation of feature regions. The results showed that the AUC value of ResNet on the test dataset was 0.86, demonstrating good diagnostic performance, and the AUC value of this model was significantly higher than that of three radiologists of different seniority.

The diagnostic performance of the model developed in this study for the training set and test set was AUC = 0.899 and 0.831; ACC = 89.4 and 84.1%. Compared with related studies reported previously, the model developed in this study shows excellent performance in the diagnosis of active tuberculosis and

has good generalization ability. For example, Li et al. [22] used the most advanced three-dimensional convolutional neural network model, collecting 501 CT image datasets from 223 active tuberculosis patients and 501 datasets from healthy individuals as negative samples. The detection recall rate and accuracy were 85.8% and 89.2% respectively. Wang et al. [21] retrospectively collected 301 cases of non-tuberculous mycobacterial pulmonary disease and 804 cases of tuberculosis mycobacterial pulmonary disease with chest CT images, and developed a 3D ResNet model. Randomly extracting data for training, validation and testing, the AUC values of the model in training, validation and testing were 0.90, 0.88 and 0.86 respectively. This study differs from previous studies. Cui et al. [23] used radiology to distinguish tuberculosis and lung cancer, but the disadvantage was that it took a lot of time to delineate the interested areas in the early stage and the results were greatly affected by the delineation. In contrast, the method in this study used deep learning networks, which learn from the entire image of the patient and can fully explore the information in the image. Although Ma et al. [24] adopted the deep learning method, the lung field pre-segmentation model used traditional morphological operations, and the deep learning method was more robust than morphological operations. Additionally, the authors used a two-dimensional classification network, but the classification network in this study fully utilized the three-dimensional features of the image, confirming the stability, repeatability and adaptability of the deep learning model in multiple aspects.

To further verify the performance of the model, this study used an independent test set to compare the performance of the MATDenseNet model with the diagnostic results of three radiologists of different seniority levels. The study found that in the average sensitivity, specificity, and accuracy of the three-classification task, the diagnostic performance of the MATDenseNet model was superior to that of junior doctors, was comparable to that of mid-level doctors, and had a high consistency. There was no statistically significant difference in the diagnostic level with senior doctors, and the consistency was excellent, with Kappa = 0.821. The above results indicate that the MATDenseNet model is comparable to the diagnostic level of mid-level radiologists in identifying diseases such as secondary pulmonary tuberculosis, and is expected to be used as an auxiliary diagnostic tool for secondary pulmonary tuberculosis in clinical practice. The study also found that the more experienced the physician, the higher the diagnostic accuracy. However, even mid-level and senior-level doctors have difficulty accurately identifying patients with common pulmonary infections, with sensitivities of 74.9% and 83.4% respectively. Thus, the diagnosis of common pulmonary infections is a difficult point in the clinical work of radiologists, and medical staff still need to continuously enhance their learning and improve their diagnostic capabilities.

In conclusion, the model in this paper can effectively achieve the differential diagnosis of non-tuberculosis patients and tuberculosis, and the precise staging of pneumoconiosis. It is worthy of further promotion and application. However, this study still has certain limitations. Firstly, the data comes from the same institution, and the generalization ability of the model still needs further research. Secondly, for deep learning, the sample size still needs to be expanded. Finally, this study only refers to the diagnostic criteria for digital radiography anteroposterior chest films, and the Chinese Health Supervision Association released the "Guidelines for Chest CT Assisted Diagnosis of Pneumoconiosis" on February 27, 2023. Therefore, future research should consider combining CT, and include different data types such as demographic characteristics, dust exposure type, and dust exposure years for modeling.

References

1. Lyon, S. M., & Rossman, M. D. (2017). Pulmonary tuberculosis. Microbiology spectrum, 5(1), 10-1128.
2. Nachiappan, A. C., Rahbar, K., Shi, X., Guy, E. S., Mortani Barbosa Jr, E. J., Shroff, G. S., ... & Hammer, M. M. (2017). Pulmonary tuberculosis: role of radiology in diagnosis and management. Radiographics, 37(1), 52-72.
3. Conradie, F., Diacon, A. H., Ngubane, N., Howell, P., Everitt, D., Crook, A. M., ... & Spigelman, M. (2020). Treatment of highly drug-resistant pulmonary tuberculosis. New England Journal of Medicine, 382(10), 893-902.
4. Jiang, H., Liu, M., Zhang, Y., Yin, J., Li, Z., Zhu, C., ... & Guo, X. (2021). Changes in incidence and epidemiological characteristics of pulmonary tuberculosis in mainland China, 2005-2016. JAMA network open, 4(4), e215302.

5. Wang, X. N., He, T. L., Geng, M. J., Song, Y. D., Wang, J. C., Liu, M., ... & He, G. X. (2018). Prevalence of and risk factors for tuberculosis among healthcare workers in Chinese tuberculosis facilities. Infectious diseases of poverty, 7(1), 26.
6. Li, T., Zhang, B., Du, X., Pei, S., Jia, Z., & Zhao, Y. (2024). Recurrent pulmonary tuberculosis in China, 2005 to 2021. JAMA Network Open, 7(8), e2427266.
7. Long, Q., Guo, L., Jiang, W., Huan, S., & Tang, S. (2021). Ending tuberculosis in China: health system challenges. The Lancet Public Health, 6(12), e948-e953.
8. World Health Organization. (2016). Chest radiography in tuberculosis detection: summary of current WHO recommendations and guidance on programmatic approaches (No. WHO/HTM/TB/2016.20). World Health Organization.
9. Rea, G., Sperandeo, M., Lieto, R., Bocchino, M., Quarato, C. M. I., Feragalli, B., ... & Lacedonia, D. (2021). Chest imaging in the diagnosis and management of pulmonary tuberculosis: the complementary role of thoraci ultrasound. Frontiers in medicine, 8, 753821.
10. Najjar, R. (2023). Redefining radiology: a review of artificial intelligence integration in medical imaging. Diagnostics, 13(17), 2760.
11. Srivastav, S., Chandrakar, R., Gupta, S., Babhulkar, V., Agrawal, S., Jaiswal, A., ... & Wanjari, M. B. (2023). ChatGPT in radiology: the advantages and limitations of artificial intelligence for medical imaging diagnosis. Cureus, 15(7), e41435.
12. Gore, J. C. (2020). Artificial intelligence in medical imaging. Magnetic resonance imaging, 68, A1-A4.
13. Nansamba, B., Nakatumba-Nabende, J., Katumba, A., & Kateete, D. P. (2025). A systematic review on application of multimodal learning and explainable AI in tuberculosis detection. IEEE Access.
14. Shao, J., Ma, J., Yu, Y., Zhang, S., Wang, W., Li, W., & Wang, C. (2024). A multimodal integration pipeline for accurate diagnosis, pathogen identification, and prognosis prediction of pulmonary infections. The Innovation, 5(4).
15. Kumar, S., Sharma, S., & Megra, K. T. (2025). Transformer enabled multi-modal medical diagnosis for tuberculosis classification. Journal of Big Data, 12(1), 5.
16. Kumar, S., Rani, S., Sharma, S., & Hong, M. (2024). Multimodality fusion aspects of medical diagnosis: A comprehensive review. Bioengineering, 11(12), 1233.
17. Mapari, R. B., Tamboli, A. I., Tamhan, P., Shelar, Y., Bhise, S., & Kumar, C. A. (2025). Multi-Modal AI Approach for Early Tuberculosis Detection by Combining Symptom, Imaging, and Clinical Data. Indian Journal of Tuberculosis.
18. Hansun, S., Argha, A., Bakhshayeshi, I., Wicaksana, A., Alinejad-Rokny, H., Fox, G. J., ... & Marks, G. B. (2025). Diagnostic performance of artificial intelligence–based methods for tuberculosis detection: Systematic review. Journal of medical Internet research, 27, e69068.
19. Ou, Y., Zhang, Q., Meng, C., Zhou, X., Na, M., Yu, Z., ... & Huang, C. (2026). Artificial Intelligence in the Diagnosis and Management of Pulmonary Tuberculosis: A Review of Current Applications and Future Perspectives. Therapeutics and Clinical Risk Management, 596651.
20. Cunyuan Luan & Huabo Liu. (2026). AdaDenseNet-LUC: Adaptive Attention DenseNet for Laryngeal Ultrasound Image Classification. BioMedInformatics,6(1),5-5.
21. Wang Li, Ding Wenlong, Mo Yan, Shi Dejun, Zhang Shuo, Zhong Lingshan... & Xing Zhiheng. (2021). Distinguishing nontuberculous mycobacteria from Mycobacterium tuberculosis lung disease from CT images using a deep learning framework. European journal of nuclear medicine and molecular imaging,48(13),1-14.
22. Xukun Li, Yukun Zhou, Peng Du, Guanjing Lang, Min Xu & Wei Wu. (2020). A deep learning system that generates quantitative CT reports for diagnosing pulmonary Tuberculosis. Applied Intelligence,51(6),1-12.

23. Cui ENuo, Yu Tao, Shang ShengJie, Wang XiaoYu, Jin YiLin, Dong Yue... & Jiang XiRan. (2020). Radiomics model for distinguishing tuberculosis and lung cancer on computed tomography scans. World journal of clinical cases,8(21),5203-5212.
24. Ma Luyao, Wang Yun, Guo Lin, Zhang Yu, Wang Ping, Pei Xu... & Lure Fleming Y M. (2020). Developing and verifying automatic detection of active pulmonary tuberculosis from multi-slice spiral CT images based on deep learning. Journal of X-ray science and technology,28(5),939-951.