Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X
Volume 3, Issue 10, 2025

DESIGN AND EVALUATION OF AN AI-ASSISTED MEDICAL IMAGING SYSTEM FOR PULMONARY NODULE DIAGNOSIS

Amna Arshad[*1], Haziq Hayat[2], Khadija Parwez[3]

*[*1]Department of Computer Science, Shaheed Zulfikar Ali Bhutto Institute of Science and Technology (SZABIST), Islamabad, Pakistan.*
[2]Department of Robotics and Artificial Intelligence, Shaheed Zulfikar Ali Bhutto Institute of Science and Technology (SZABIST), Islamabad, Pakistan.
[3]Department of Computing and Technology, IQRA University, H-9 Campus, Islamabad.

[*1]amnaarshad090@gmail.com, [2]hazikrizwan16@gmail.com, [3]khadijaparwez@gmail.com

DOI: https://doi.org/10.5281/zenodo.17365766

Keywords
Artificial intelligence, Medical imaging–assisted diagnosis system, Image segmentation, Pulmonary nodules

Article History
Received: 24 Agust 2025
Accepted: 02 October 2025
Published: 16 October 2025

Corresponding Author: * Amna Arshad

Abstract
The advancement of artificial intelligence (AI) technology has not only transformed social and industrial production but has also accelerated progress in the medical domain. Currently, the use of AI in healthcare is rapidly expanding, with medical imaging being a key area of application. This study focuses on integrating AI with a medical imaging–assisted diagnosis system to address the limitations and inaccuracies of conventional manual diagnosis in the detection of pulmonary nodules. By applying established principles of image segmentation, a comprehensive design and optimisation of an AI-driven diagnostic framework was developed to enhance precision in pulmonary nodule assessment. To evaluate its performance, 200 patient cases comprising 231 nodules, confirmed by pathology or with stability over more than two years of follow-up, were analysed. The AI system detected 881 true nodules with a sensitivity of 99.10% (881/889), whereas radiologists identified 385 true nodules with a sensitivity of 43.31% (385/889). Moreover, the AI system demonstrated markedly higher sensitivity in detecting non-calcified nodules compared to radiologists (99.01% vs. 43.30%, P < 0.001), with statistically significant differences. These findings highlight the potential of AI-based medical imaging systems to substantially improve diagnostic accuracy, reduce missed detections, and support radiologists in clinical decision-making.

INTRODUCTION

With the rapid progress of modern science and technology, artificial intelligence (AI) is being increasingly applied in various fields of work and everyday life. The care of human health issues is one of the most crucial fields where AI is developing rapidly. Today, the combination of AI and medical imaging has become a major research focus, and studies in this area are increasing rapidly. The AI-assisted diagnosis system functions as an intelligent tool that supports doctors by utilising deep learning to aid in the analysis of medical images. It combines the workflow of human radiologists with intelligent computer assistance, providing helpful diagnostic suggestions while the doctor reads the images. This technology enables doctors to make more accurate imaging diagnoses, reduces the likelihood of mistakes or missed findings, and enhances work efficiency.

Lung cancer is one of the most dangerous cancers to human health worldwide, ranking first in both occurrence and death rates among all cancers. Thin-slice CT scans are highly effective in detecting lung cancer early; however, they also increase the number of CT images that require review. Because of the large amount of CT data, doctors may miss or misread some lung nodules. To solve this, several companies have developed AI-based CT screening platforms. One of the most significant advantages of the AI screening system in clinical use is its speed — while doctors need several minutes to read one scan, the AI system can complete it in about three seconds. Another significant advantage is that AI does not get tired. As doctors work longer hours, their focus and reaction time decrease, which can lead to diagnostic errors. AI, however, continues to operate at the same high level of accuracy. It improves diagnostic efficiency and reduces the chance of missing small pulmonary nodules. Besides detecting nodules, AI can also notice subtle lesion changes that the human eye might miss, and some of its quantitative parameters are useful for identifying the type of lesion.

The application of AI technology is a significant step in advancing the medical field. It is a highly innovative approach for diagnosing pulmonary nodules using medical images, and many researchers have explored this topic. For example, Zheng G proposed an integrated classification method based on the INCEPTION module to diagnose lung nodules with different characteristics [1]. Zhang S applied the well-known LENET-5 model to classify lung nodules from chest CT images [2]. Zhang N aimed to study the value of CT texture analysis in distinguishing between benign and malignant solitary pulmonary nodules (SPNs). His method involved performing CT scans on 97 patients (54 with malignant and 43 with benign nodules) to measure the CT value and the maximum diameter of each nodule [3]. Vayntrub Y used PET imaging to evaluate lung nodules and verify the Herder model [4]. Wang Y pointed out that computer-aided diagnosis (CAD) based on deep learning is currently the main research direction [5]. Although these studies have advanced AI research in medicine, most remain at the theoretical level and have limited applicability in real-world clinical practice.

To achieve more accurate and reliable diagnosis of pulmonary nodules, traditional diagnostic approaches need improvement. Therefore, some researchers have proposed medical imaging–assisted diagnosis systems to enhance accuracy. Singh S developed a computer diagnostic system that can automatically generate radiological reports from medical images, thereby reducing the workload of doctors [6]. Ha W designed an automated system for breast tumour diagnosis [7]. Saygili A introduced a new three-step approach including preprocessing, segmentation, and classification [8]. Barinov L's research showed a significant accuracy difference between two diagnostic systems [9]. Loku L reviewed studies to evaluate how empirically based evidence is being applied to diagnose and treat autism spectrum disorder (ASD), including the use of different sensing technologies [10]. Although these works demonstrate progress in medical imaging–based diagnosis, most focus only on specific diseases or single applications, resulting in limited overall understanding of their full potential.

In this study, an AI-based medical imaging diagnosis assistant system was utilised for the diagnosis of pulmonary nodules. The system's clinical performance was evaluated using a subjective image quality assessment method. A test was conducted on 200 patients, covering a total of 231 nodules confirmed by pathology or stable after two years of follow-up. Results showed that the AI software detected 881 true nodules, with a sensitivity of 99.10% (881/889). The radiologist detected 385 true nodules, with a sensitivity of 43.31% (385/889). AI software demonstrated significantly higher sensitivity for continuous nodules than radiologists (99.01% vs. 43.30%, $P < 0.001$) [11]. The AI system also achieved higher sensitivity for detecting nodules smaller than 5 mm ($P = 0.000$). For nodules between 5 and 10 mm, its sensitivity remained superior to that of radiologists, with statistical significance.

The primary contribution of this paper is to demonstrate how an AI-based medical image assistance system can be effectively applied in clinical medicine, highlighting its clear advantages. It not only helps reduce the radiation dose from chest CT scans and detect more nodules, but it can also locate and describe lung nodules, predict tumour growth, estimate the pathological grade of malignant nodules, and even forecast tumour metastasis and patient prognosis [12]. Many studies now confirm that AI can significantly reduce radiologists' workloads, free up medical staff from repetitive tasks, and lower the rate of missed pulmonary nodule diagnoses by doctors.

Medical imaging diagnosis assistance system in the diagnosis of pulmonary tuberculosis

Overall scheme design

The AI-aided diagnosis system for pulmonary nodules is divided into two parts: cloud server design and user interface design. File transmission and instruction interaction are carried out through the designed transmission module. Its overall structure is shown in Fig. 1.

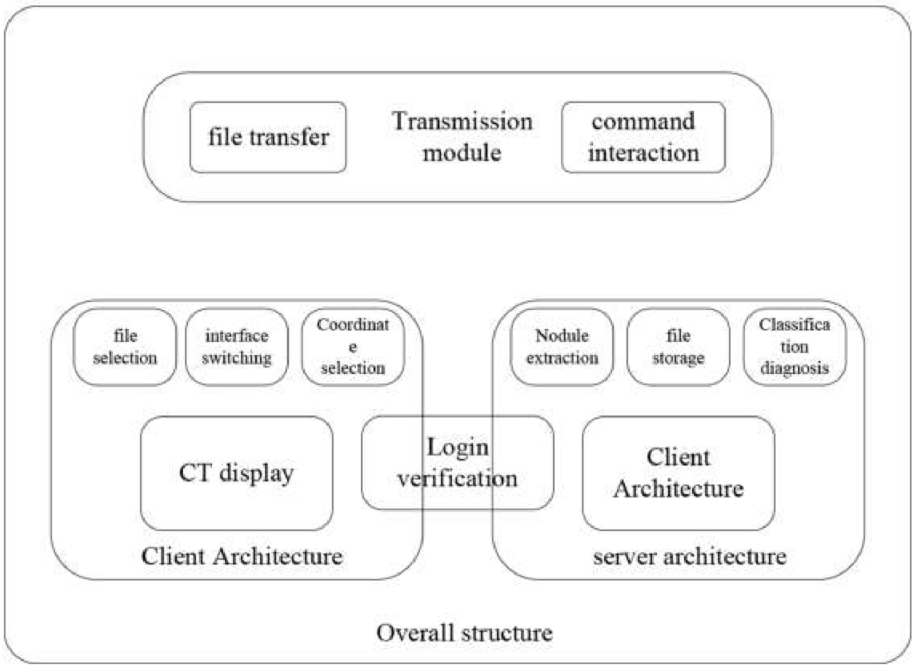

Figure 1: Overall system architecture

As shown in Fig. 1, the client architecture comprises four main components: CT display, file selection, interface switching, and coordinate selection. Its main functions are described below:

(1) Login verification

A user login verification interface is created to ensure secure access. Three types of login prompt boxes are designed to handle different situations: when the username does not exist, when the password is incorrect, and when both the username and password match successfully. In the first two cases, the system displays an error message and keeps the login page open for re-entry. Upon successful login, the interface transitions to the main operation window.

(2) File operation

The file operation function allows users to select CT image files for processing and upload them to the cloud server through the file transfer option in the transmission module. During system operation, temporary files and result data generated by the server are sent back to the client. The received results are then displayed visually on the user interface, helping users clearly see the analysis output [13].

(3) CT image display

A key feature of the system is the display of CT images. Since the CT images used are in DICOM format, users first select the folder containing the images through the file operation function. The Visual Toolkit (VTK) is then used to import and display the DICOM files. The VTK window is

embedded into the interface, allowing users to perform operations such as zooming in, zooming out, and flipping CT images. In addition, a slider is included to adjust image brightness and contrast by modifying its position, offering
better visualisation and clarity during image analysis.

(4) Instruction interaction

The interaction between the client and server allows the execution of Linux commands on the remote server through the command interaction function in the transmission module. This enables cloud-side operations, such as nodule extraction, classification, diagnosis, and result storage, all of which are handled automatically by the server.

(5) Determination of pulmonary nodule location information

When the user clicks on a CT image displayed in the interface, mouse and drawing events are triggered to record the position where the click occurred. The system calculates the exact coordinates of the lung nodule by combining the image's pixel values with the interface size. Using the nodule size extraction box provided in the interface, the user can select an appropriate region based on the nodule's size. Once selected, the data are uploaded to the cloud server via the transmission module, supplying accurate location information for subsequent lung nodule volume extraction.

Cloud Server-Side Functions

(1) MySQL database

A MySQL database is built on the cloud server to manage user-related data. It supports login name and password verification, as well as query functions. The database also stores user login times, uploaded files, and computed results in a structured way, allowing users to retrieve and review their data whenever needed.

(2) Extraction and storage of lung nodule volume data

The files uploaded by the client are stored and processed on the server to obtain CT image data and corresponding lung nodule location information. The nodule area is then identified through calculation, and its features are extracted. The processed data are stored in a predefined directory and hierarchy, ensuring that extracted nodules are ready for later classification and diagnosis.

(3) Classification diagnosis

Based on the concepts of deep residual networks, dense connection networks, and multi-resolution analysis, a multi-resolution 3D dual-path network model is constructed on the cloud server. This model is used to classify and diagnose the extracted 3D lung nodule data. The diagnostic results are saved in designated folders and categorised by level, enabling users to retrieve them when needed.

The overall workflow of the AI-assisted diagnosis system for pulmonary nodules is illustrated in **Fig. 2**.

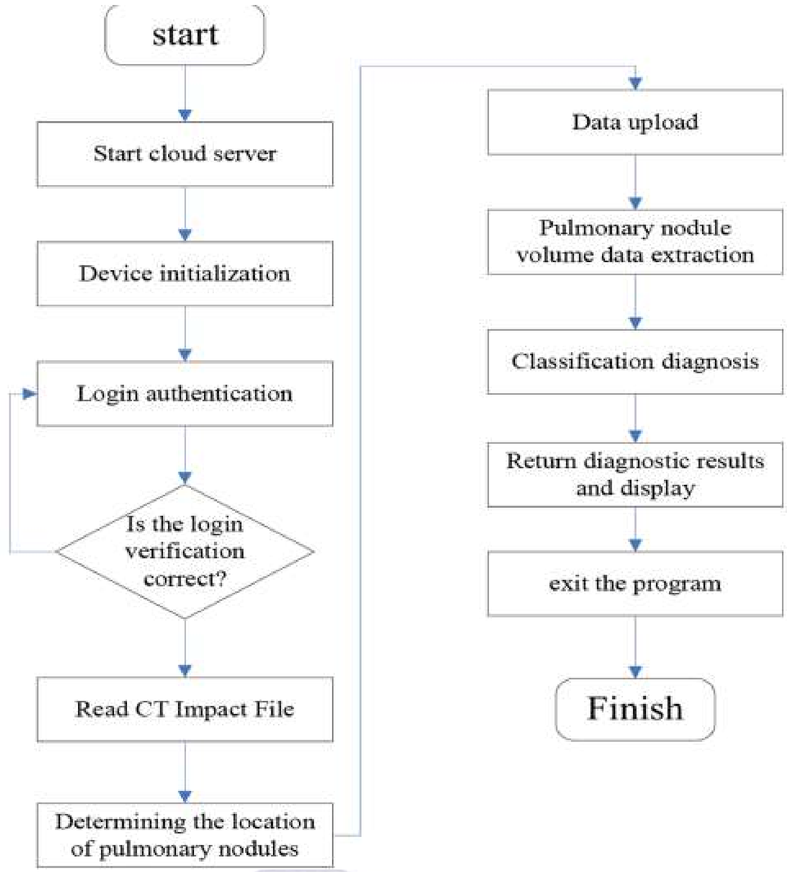

Figure 2 Schematic diagram of the overall flow of the artificial aided diagnosis system for pulmonary nodules

As shown in Fig. 2, the AI-assisted diagnosis system for pulmonary nodules is developed using the Python programming language, and the deep learning network for diagnosis is constructed on the TensorFlow platform [14]. The main component of the computer-aided diagnosis algorithm is the design of a multi-resolution 3D dual-path network model, which combines the principles of the deep residual network (ResNet) and dense connection network (DenseNet). In particular, the algorithm first extracts the 3D volume data of lung nodules using multi-resolution processing techniques, and then applies a dual-path network for feature extraction and classification. The network comprises multiple paths, each processing data at different scales and resolutions, thereby enhancing the diversity and richness of the extracted features. By training the deep learning model on the TensorFlow platform, the system can automatically learn various visual and structural features of lung nodules and generate diagnostic results. These results are saved and made accessible to users through predefined directories and hierarchical paths, ensuring easy retrieval and management of data.

During the model training and validation stage, an extensive collection of labelled lung nodule images is required to form the training and validation datasets. In the training phase, the backpropagation algorithm in deep learning is used to optimise the model, and the classification accuracy is gradually improved by continuously adjusting the weight parameters in the network. To ensure that the model performs well across different datasets, cross-validation is used to assess its generalisation ability. During the verification phase, the model's predicted results on the validation set are compared with the ground truth labels, allowing further optimisation of the model's structure and parameters. This process ultimately produces a model with high diagnostic accuracy. The entire workflow is implemented in Python using the TensorFlow framework and integrated with the PyQt5 development toolkit to provide a user-friendly graphical interface for easy interaction.

Design of the interaction function between the client and the cloud server

(1) SSH protocol for server connection design

The system uses the Secure Shell (SSH) protocol to establish a secure connection between the client and the cloud server. The Diffie–Hellman (DH) algorithm is used for key exchange, while authentication is performed using RSA or DSA algorithms to ensure secure login and data transmission. The overall framework of the SSH protocol used in this design is shown in Fig. 3.

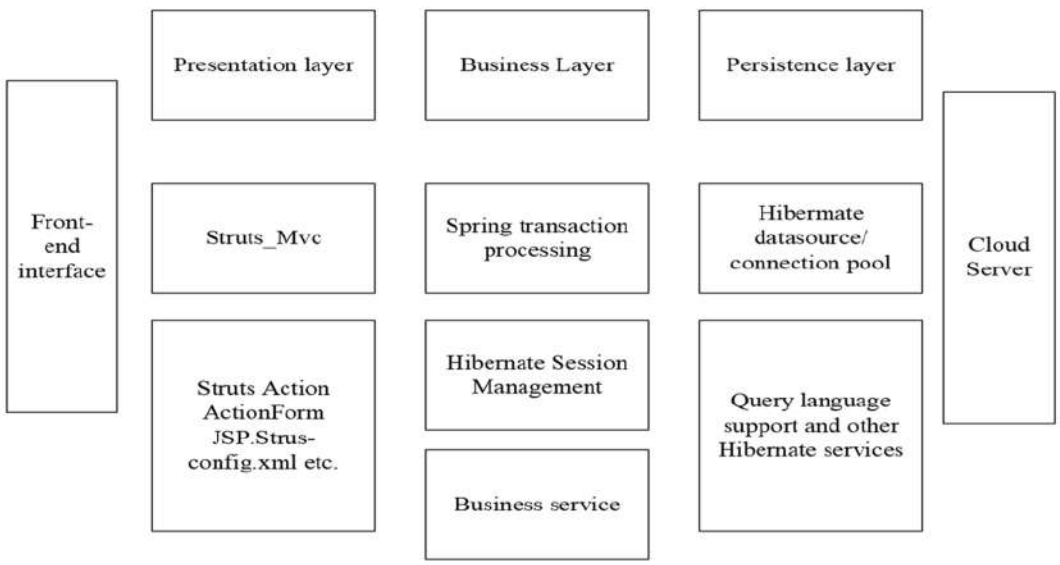

Figure 3 SSH protocol framework

The SSH protocol follows a Client–Server architecture. Paramiko is a Python library that implements and encapsulates the SSHv2 protocol, allowing developers to control remote computers through Python. The Paramiko library primarily consists of two key components: SSHClient and SFTPClient.

The SFTPClient component provides functionalities similar to standard SFTP commands and is used for remote file operations such as uploading, downloading, and managing directories. The SSHClient component is a high-level interface for establishing secure connections, functioning similarly to the Linux ssh command. It manages the creation of Transport, Channel, and SFTPClient objects.

A Channel acts as a secure communication path between the client and server, while Transport is responsible for encrypting data before it is transmitted. In most applications, Transport is used to create a remote client session and send commands from the client to the server. A Session represents the ongoing communication process between the Client and Server, maintaining the interaction between both sides [15, 16].

(2) Implementation of the Interaction Function Between Files and Commands

To achieve communication between the client and the cloud server, an SSH connection is established using the Paramiko module. File operations are performed through the SFTP class within Paramiko, enabling actions such as file uploading, downloading, directory creation, and other remote file management tasks.

Within the client interface, a QPushButton widget is used for user interaction. The button is set to trigger a clicked signal, which activates when the user presses the left mouse button. This signal is linked to the function self.xuanze(), where the actual file-selection process is implemented.

The QFileDialog class provides a graphical interface for selecting files or folders. It displays filenames and allows users to easily choose paths. Because the CT image data to be uploaded consists of multiple DICOM-format files stored in a folder, the getExistingDirectory() function of QFileDialog is used to select the appropriate directory. The selected folder path is stored in the self variable. Path.

A for loop is then used to iterate through all DICOM files within the selected folder. The upload operation is performed using sftp.put() function. The local file path is formed by combining the folder path (self.path) with the DICOM file name (dicom_name). Similarly, the remote file path on the cloud server is created using the variables self.cloud_path and dicom_name. Each file is stored in the same destination folder on the cloud server, ensuring all data is organised together in a consistent directory structure.

The complete file upload process, including path selection and transfer, is illustrated in Fig. 4.

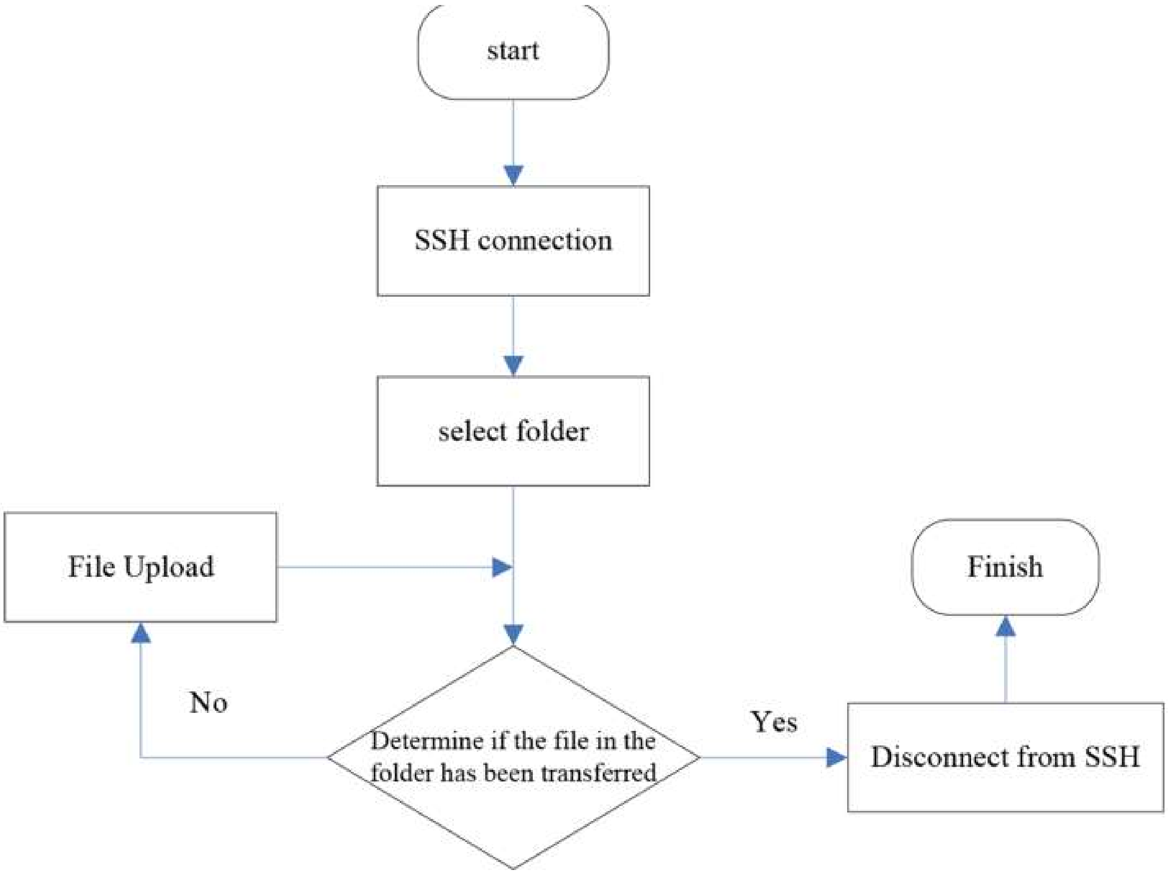

Figure 4: File upload process

After completing the previous steps, the necessary files—such as the lung CT images to be processed—are successfully transferred from the client to the cloud server. Once the data transmission is complete, the client can send Linux commands to the cloud server to perform data processing and other related operations. These commands are executed remotely on the server side.

When a command is sent, a new channel is created to carry out the instruction. The input and output streams of the command correspond to the standard input, output, and error objects that return execution results. The command parameter contains the specific instruction to be executed, while the timeout parameter defines the maximum time allowed for the command to run before it is automatically terminated [17, 18].

Subjective Image Quality Evaluation Method

The subjective evaluation method for image quality involves having multiple experienced medical professionals visually assess the enhanced images produced by the system and provide feedback based on their expertise and clinical judgment. This evaluation primarily relies on medical workers' theoretical knowledge and practical experience to assess the image quality.

Because this method does not involve a reference or standard comparison image, it is also referred to as an absolute subjective evaluation. In contrast, relative subjective evaluation is based on comparing the test image with a reference image that has relatively better quality or less distortion, allowing observers to judge differences through visual comparison. The following Equation can express the differential subjective scoring method used in this study:

$$d_{i,j}=MOS_{original}-MOS_{distorted}$$

$$d'_{i,j} = \frac{d_{i,j} - \min(d_{i,j})}{\max(d_{i,j}) - \min(d_{i,j})}$$

The MOS value here refers to a subjective score (one to five, ranging from worst to best). Di,j refers to the MOS difference between the reference image and the evaluated image, typically the MOS of the reference image. The value can be either four or five, with Di,j being the average.

Objective image quality evaluation method

However, in actual image quality evaluation, subjective evaluation is not entirely accurate or scientific. It will be limited by the self-generated experience of medical work, lack of theoretical knowledge, and emotional factors, which make the image evaluation inaccurate. Therefore, the evaluation method without reference is adopted in this paper, and the assessment is performed by calculating the relevant parameter values of the grey level co-occurrence matrix. The specific parameters are as follows:

(1) Energy: The Equation is shown in Equation 3:

$$ASM = \sum_{i=0}^{L-1}\sum_{j=0}^{L-1} P_d(i,j)^2$$

(2) Contrast: The Equation is shown in Equation 4:

$$CON = \sum_{N=0}^{L-1} n^2 \left(\sum_{j=0}^{L-1} P_d(i,j)\right)$$

(3) Correlation: The Equation is shown in Equation 5:

$$COR = \frac{\sum_{i=0}^{L-1}\sum_{j=0}^{L-1} ijP_d(i,j) - \mu_1\mu_2}{\sigma_1^2\sigma_2^2}$$

In the Equation, μ_1, μ_2, σ_1 and σ_2 are respectively the following Equations:

$$\mu_1 = \sum_{i=0}^{L-1} i \sum_{i=0}^{L-1} \hat{P}(i,j)$$

$$\mu_2 = \sum_{i=0}^{L-1} i \sum_{i=0}^{L-1} \hat{P}_d(i,j)$$

$$\sigma_1^2 = \sum_{i=0}^{L-1} i(i-\mu_1)^2 \sum_{i=0}^{L-1} \hat{P}d(i,j)$$

$$\sigma_2^2 = \sum_{i=0}^{L-1} i(i-\mu_2)^2 \sum_{i=0}^{L-1} \hat{P}d(i,j)$$

(4)Entropy: The Equation is shown in Equation 10:

$$H = \sum_{i=0}^{L-1}\sum_{i=0}^{L-1} \hat{P}_d(i,j) log \hat{P}_d(i,j)$$

(5)The Equation is shown in Equation 11:

$$l(X,Y) = \frac{2\mu_X + \mu_Y + C_1}{\mu_X^2 + \mu_Y^2 + C_1}$$

$$c(X,Y) = \frac{2\sigma_X\sigma_Y + C_2}{\sigma_X + \sigma + C_2}$$

$$S(X,Y) = \frac{\sigma XY + C_3}{\sigma_X^2\sigma_Y^2 + C_2}$$

Among them, μx, μy, σX^2, σY^2, and σXY represent the mean value and the standard deviation of x and y in the x and y directions of the evaluated image, respectively [19]. Covariance: C1, C2, *and* C3 are constant, and L is the dynamic range of the pixel grey value of the evaluation image. For eight-bit grayscale images, L = 255; for sixteen-bit grayscale images, L = 65535. The Equation for calculating the structural similarity of an image is:

$SSIM(X, Y) = [l(X, Y)]\alpha * [c(X, Y)]\beta * [S(X, Y)]\gamma$

Let $\alpha = \beta = \gamma = 1$. By improving the algorithm based on research into improvement and referring to the relevant characteristics of the human visual imaging system, a definition evaluation index for non-reference images is designed. The Equation for calculating the structural clarity is as follows:

NRSS=1-$NRSS = 1 - \frac{1}{N}\sum_{-=1}^{N} SSIM(X_i, Y)_i$

Improved OTSU lung parenchyma segmentation

The OTSU algorithm is a method that determines the optimal segmentation threshold through an exhaustive search, and it has become a relatively mature algorithm for lung parenchyma segmentation. Its process is shown in Fig. 5:

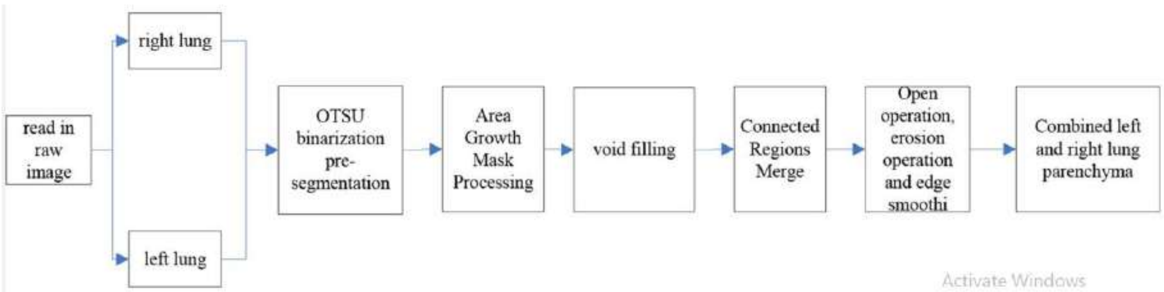

Fig. 5 Algorithm flowchart

The specific algorithm principle is as follows: In this paper, the medical image in DICOM format provided by a city people's hospital is used, and the grey level is 16 bits, so L is 65536. The number of pixels with a grey value of i is n, then the number of pixels in the image is the sum of all grey values:

$$N = \sum_{i=0}^{L-1} n_i$$

Probability of occurrence of grey level: It can be defined as the number of pixels of this grey value in the total number of pixels.

$$p_i = \frac{n_i}{N} \; i = 0, 1, 2, \ldots, L-1$$

Then, look for a valley threshold value t in the histogram of the image, and use this threshold value to divide the image into two parts: namely, Co = (0, 1,...). C1 = (t + 1,t + 2, t + 3, ... L-1), the probabilities of these two parts are:

$$w_0 = \sum_{i=0}^{t}$$

$$w_1 = \sum_{i=t+1}^{L-1} p_i = 1 - w_0$$

Through the above Equation, the grey mean value of the whole image and the grey mean value of the two parts after division can be calculated:

$$\mu = \sum_{i=0}^{L-1} i * p_i$$

$$\mu_0 = \sum_{i=0}^{t} i * p_i / w_0$$

$$\mu_1 = \sum_{L-1}^{i=t+1} i * p_i / w_1$$

By continuously adjusting the grey value and increasing or decreasing the threshold t to repeatedly calculate σt2, the best threshold for segmentation can be determined, thereby maximising the inter-class variance. This allows for the best separation of the background and the lung parenchyma.

Segmentation of the regional growing lung parenchyma based on OTSU

The image data used in the algorithm of this study are 16-bit DICOM-format lung CT medical images. The specific algorithm steps are as follows: First, the original DICOM image is read into the system. The OTSU algorithm is then applied to perform binarisation pre-segmentation on the left lung parenchyma. After obtaining the pre-segmented lung region, mask processing is carried out using the region-growing method to refine the segmented area. Once the mask processing is complete, any holes in the resulting image with an area smaller than the threshold value T are automatically filled. Larger holes are treated as connected regions, which are then processed to ensure structural consistency. After identifying the connected areas, morphological operations such as opening, erosion, and smoothing are performed to enhance the boundary clarity and eliminate noise. Through these steps, the outline of the left lung parenchyma is accurately extracted. The same procedure is then applied to the right lung parenchyma. Finally, the left and right lung parenchyma regions are merged and fused to form a complete lung parenchyma structure, which is subsequently denoised to improve image quality and accuracy [20]. The experimental outcomes of this process are presented in Fig. 6.

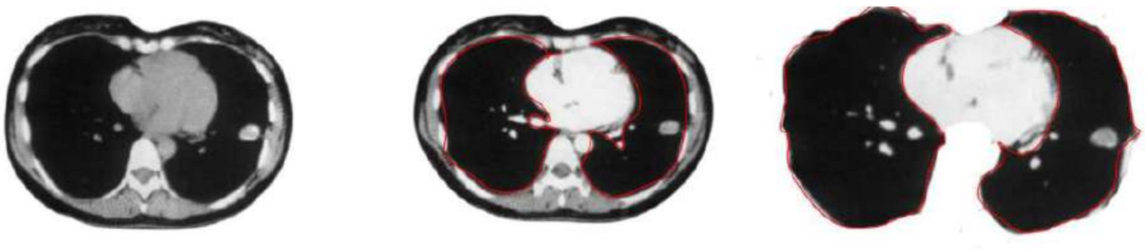

Fig. 6 OTSU lung parenchyma segmentation result

In this section, the preprocessed and segmented lung parenchyma can be reconstructed by using more reliable volume rendering and surface rendering algorithms. The azimuth observation is realised through the 3D rotation effect, which is convenient for medical workers to diagnose details and has a beneficial auxiliary impact.

Detection performance evaluation of an AI pulmonary nodule auxiliary diagnosis system

The density and size of pulmonary nodules are strongly associated with their likelihood of being benign or malignant. Research indicates that fewer than 1% of nodules with a diameter of less than 5 mm are malignant. In comparison, approximately 6% to 28% of nodules measuring 5 to 10 mm in diameter are malignant, while nodules larger than 20 mm have a 64% to 82% probability of malignancy. Studies have also demonstrated that ground-glass nodules (GGNs) are generally more likely to be malignant than solid nodules. Because nodules smaller than 5 mm are of a small size and GGNs often exhibit low density, the qualitative assessment of these nodules has historically been challenging. However, with advancements in computed tomography (CT) technology, the detection rate of pulmonary nodules has significantly improved. In this section, the study focuses on evaluating the performance of different AI-assisted diagnostic systems (AIADS) in detecting and measuring various types of pulmonary nodules. The goal is to identify the limitations and shortcomings of current AI systems and to provide valuable insights for further optimisation and enhancement of AIADS algorithms and technologies.

CT scan image data acquisition

Phantoms with 15 simulated lung nodules were scanned using SIEMENS SOMATOM Definition Flash CT, as shown in Table 1:

Table 1. CT parameter settings and radiation agent descriptive data

Sequence	Tube voltage (kVp)	Tube current (mAs)	CTDIvol (mGy)	DLP (mGy·cm)	E (mSv)
1	80	100	6.96	245	3.45
2	80	200	8.78	369	4.65
3	100	300	6.11	245	3.85
4	80	120	5.45	175	2.96
5	120	120	4.63	86	2.32
6	100	100	2.53	143	1.74
7	120	40	1.25	69	0.85

8	100	20	1.55	78	0.56
9	120	60	1.95	63	0.50
10	100	10	0.42	84	0.32
11	80	20	0.31	21	0.25
12	100	10	0.28	9	0.36

The above 12 sets of parameter images are respectively reconstructed using different convolution kernels in the filtered back-projection algorithm. To meet the AI system's needs for recognising images, the reconstructed layer thickness and layer spacing are both set to 1 mm, while other parameters remain unchanged. Finally, the 36 sets of reconstructed images are transferred to the image storage and transmission system.

Comparison of the overall performance of AIADS

The sensitivity, FP, FN and RVE results of the four AIADS and the comparison between groups are shown in Table 2.

Table 2. Descriptive data and comparison of sensitivity, false positive (FP), false negative (FN), and relative volume error (RVE) of four AI-assisted diagnostic systems (AIADS)

System	Sensitivity (%) Median (P25,P75)	FP Median (P25,P75)	FN Median (P25,P75)
A	62 (53.33, 60)	135 (5.17, 7.5)	7 (6, 7)
B	100 (100, 100)	1 (0, 25)	0 (0, 0)
C	76.34 (73.33, 73.3)	7 (3, 14)	3 (4, 4)
D	84.67 (80, 86.67)	2 (0, 1)	5 (2, 3)
Comparison	**AdSig**	**AdSig**	**AdSig**
A vs B	<0.0001*	<0.0001	<0.0001*
A vs C	<0.0001*	0.596	<0.0001*
A vs D	<0.0001*	<0.0001	<0.0001*
B vs C	<0.0001*	<0.0001 +	<0.0001*
B vs D	0.0001	0.999	<0.0001*
C vs D	0.071	<0.0001 +	0.071

It can be seen from the table that the median sensitivity of system B is the highest, up to 100%, while the median sensitivity of system A is the lowest (60%).

Influence of parameter setting on each software

(1) Sensitivity

The average sensitivity of each AIADS under different CT dose and convolution kernel conditions is shown in Fig. 7.

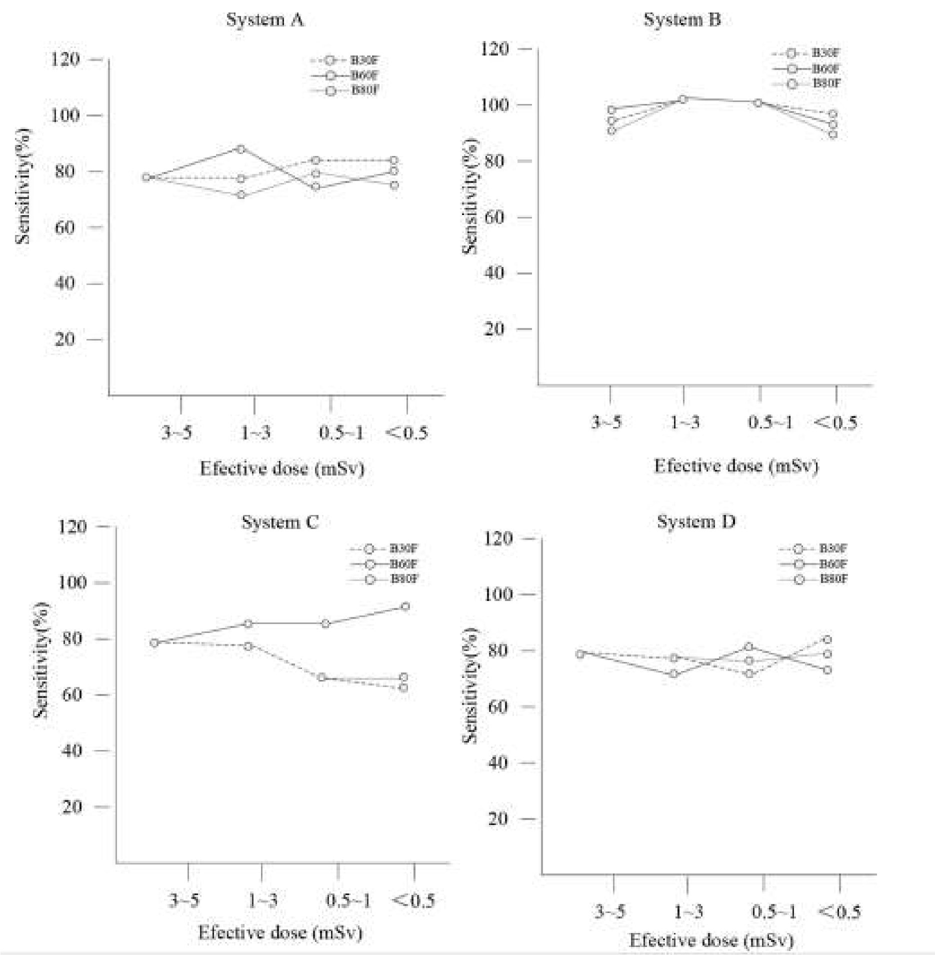

Fig. 7 Statistics of the average sensitivity of each AIADS under different

CT doses and convolution kernels

As shown in Fig. 7, for systems A, B, and C, there was no statistically significant difference in sensitivity between groups at different effective doses and convolution kernel conditions ($P > 0.05$).

(2) False positives and false negatives

A false positive refers to a test result being incorrectly displayed as positive when there is actually no disease or lesion present; a false negative refers to a test result being incorrectly displayed as negative when there is actually a disease or lesion present. Both errors will affect the accuracy of the diagnosis. False positives may lead to unnecessary treatment and examinations, while false negatives may result in the early detection of the disease being missed and treatment being delayed. Therefore, in medical testing, reducing false positives and false negatives is an important goal to improve diagnostic reliability.

This article examines false positives and False Negatives, and the results are presented in Fig. 8.

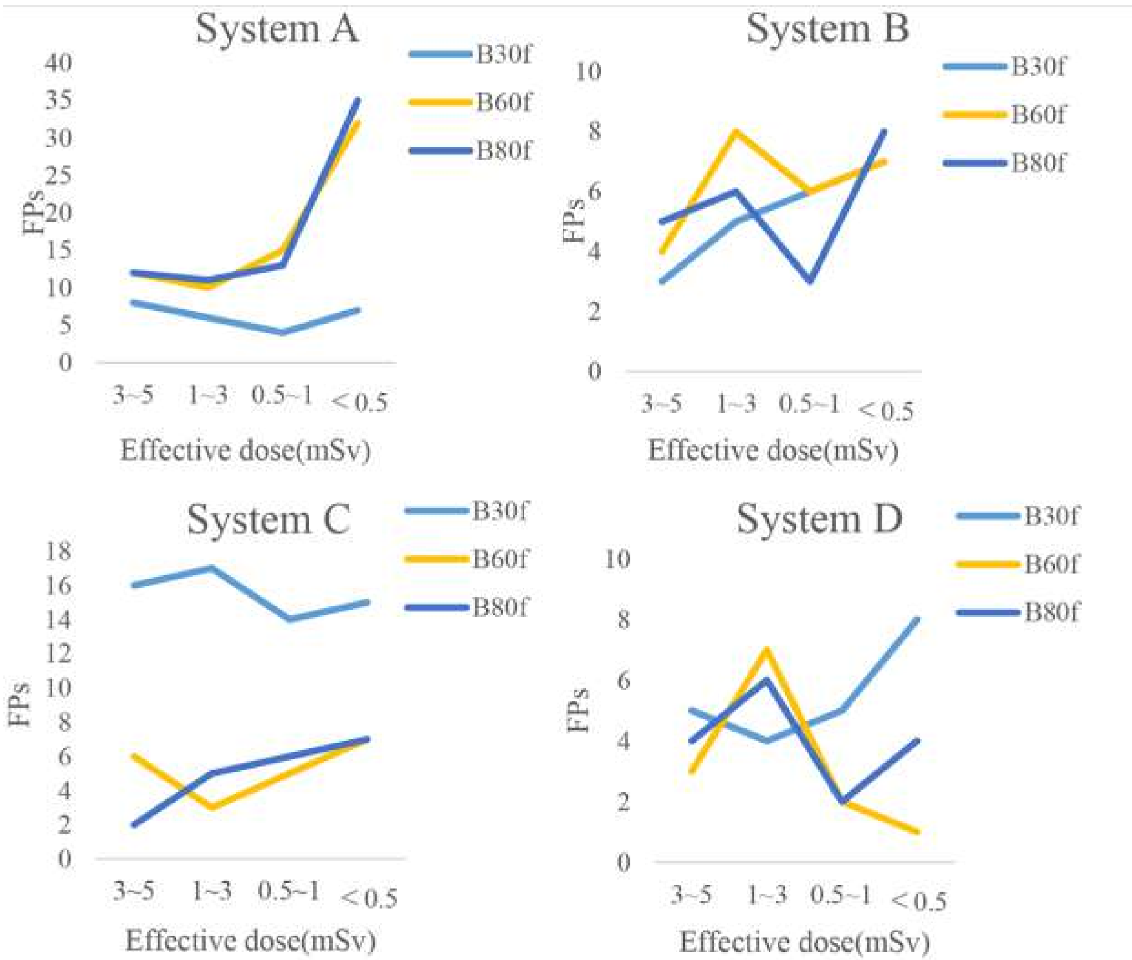

Fig. 8 Number of false positives and false negatives

As can be seen from Fig. 8, for system A, when the convolution kernels are B60f and B80f, and E decreases to its lowest value, the FPs increase to 35.67 and 39.33 (PI = 0.001), respectively. For System C, FP and FN were not significantly different between the dose groups ($P > 0.05$); however, the FPs of the B30f group were higher than those of the B80f group ($P2 < 0.0001$) under all dose conditions.

Performance evaluation of AI-assisted detection software in identifying and measuring four types of pulmonary nodules

The research in this part analyses the performance of different AIADS software in measuring and detecting various types of pulmonary nodules, identifying the deficiencies of the AI system, and provides information that can contribute to the improvement of AIADS algorithms and techniques. According to the density and diameter of simulated lung nodules, they were divided into: small ground glass nodules with a diameter of 5.33 ± 2.18 mm and a density of - 800 HU or - 630 HU. Small solid nodules were 5.33 ± 2.18 mm in diameter and + 100 HU in density. GGN (diameter 11 ± 1.1 mm, density – 800 HU or - 630 HU) and SN (diameter 11 ± 1.1 mm, density + 100HU). A partial CT image obtained by scanning the phantom is shown in Fig. 9, showing four types of lung nodules in different scan layers.

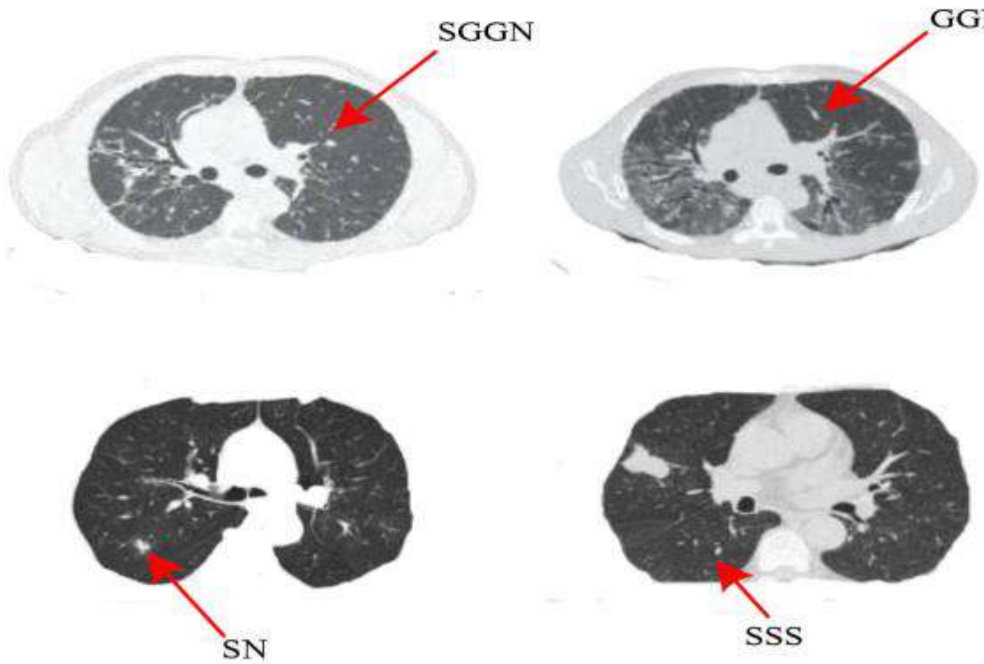

Fig. 9 Different types of simulated lung nodules CT images in the phantom

The MDR (Missed Detection Rate) of the four AIADS for different types of pulmonary nodules is shown in Table 3.

Table 3. MDR of the four AIADS for different types of pulmonary nodules

System	SGGN	SSN	GGN	SN
A	77.3	56.93	50.22	0
B	5.16	0	0	0
C	65.32	10.56	16.17	0
D	20.98	45.33	0.69	0
Comparison	Statistics	Statistics	Statistics	Statistics
A vs B	211.30	85.17	91.02	–
A vs C	5.62	76.9	75.91	–
A vs D	87	1.35	65.56	–
B vs C	145.3	Fisher	Fisher	–
B vs D	46	44.8	Fisher	–
C vs D	85.7	28.61	Fisher	–

All four systems were able to identify solid nodules (SN) correctly. Among them, System B demonstrated the lowest missed detection rate (MDR) across all types of nodules, with an MDR of 4.17% for subsolid ground-glass nodules (SGGN) and 0% for the other nodule types. For SGGN detection, statistical analysis revealed significant differences in MDR among the four systems, with the ranking order being B < D < C < A, indicating that System B performed the best, while System A performed the worst. For subsolid nodules (SSN), all systems except A and D showed significant differences, and their MDRs followed the order B < C < A < D. When detecting ground-glass nodules (GGN), System A had a noticeably higher MDR (47.22%) than the other systems, indicating poorer diagnostic accuracy in this category. The ROC curve analysis demonstrated a "one-versus-all" comparison among the four pulmonary nodule types — SGGN vs. other nodules, SSN vs. other nodules, GGN vs. other nodules, and SN vs. other nodules. Overall, Systems B, C, and D demonstrated superior diagnostic performance for all nodule types, except SSN, with AUC values exceeding 0.80. In contrast, System A performed poorly in distinguishing SGGN, SSN, and GGN, with AUCs of 0.79, 0.57, and 0.56, respectively, as illustrated in Fig. 10.

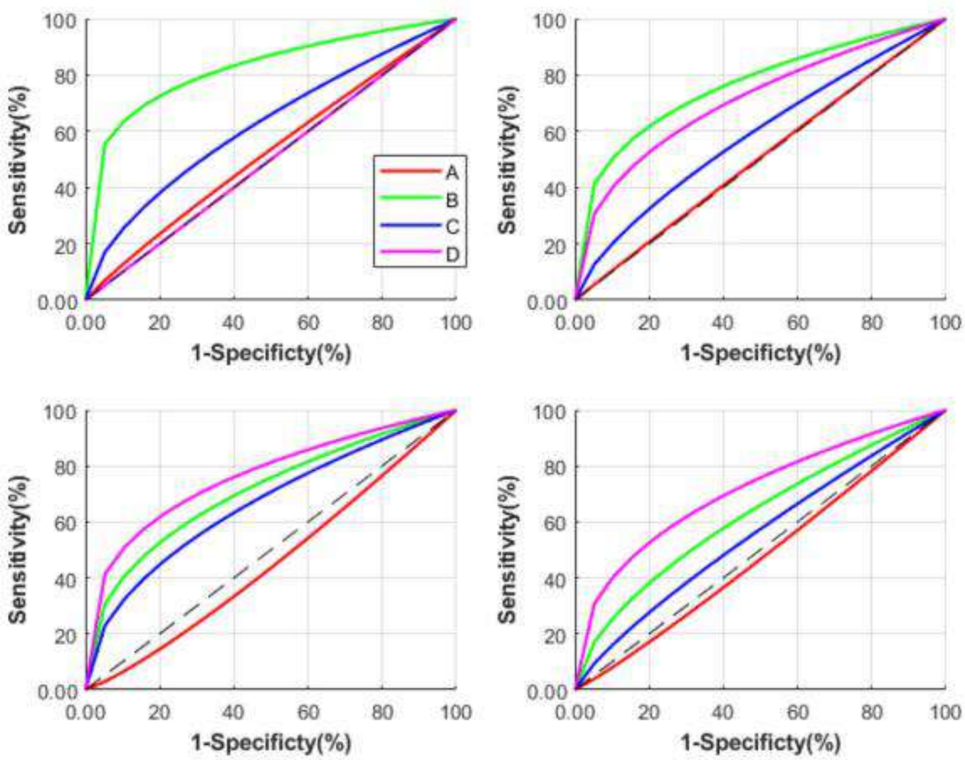

Fig. 10 One-to-many ROC curve

As shown in Fig. 10, this study analyses the ability of the AI-assisted diagnostic systems (AIADS) to classify four different types of pulmonary nodules. Among the systems, System B demonstrated a significantly lower missed detection rate (MDR) than the others and showed higher sensitivity in detecting small and low-density nodules. All systems performed well in identifying solid nodules (SN), which is likely due to their larger size and higher density, making them easier to detect. In contrast, subsolid ground-glass nodules (SGGN) were more challenging to identify. Their low density, small diameter, and poorly defined boundaries make it difficult for the software to recognise these subtle lesions accurately. The study also revealed significant differences in detection performance among the AIADS systems. While some systems excel in certain aspects, for example, System D performs best in nodule volume measurement accuracy, and System B achieves the highest accuracy in nodule type classification, no single system consistently outperforms the others in all evaluation categories.

Conclusions

The system is an AI-based medical diagnosis assistance system that combines several algorithms to increase the accuracy of the diagnosis. Three better segmentation algorithms were employed to obtain high-precision segmentation of the lung parenchyma, and the other algorithm aimed to refine the edge contours of the lung parenchyma through preprocessing and filtering edits. Additionally, point-based segmentation was used manually, allowing the boundaries to be refined to an acceptable level.

To conclude, artificial intelligence (AI) can be effectively utilised in the healthcare sector, yielding significant benefits. It has achieved an impressive performance, particularly in enhancing the rate of identifying pulmonary nodules, as well as distinguishing between benign and malignant lung lesions. The use of AI in medical imaging not only decreases the amount of work radiologists have, but also helps physicians more accurately detect and classify pulmonary nodules.

Even though AI technology is not devoid of limitations, the general potential of this technology is very promising, particularly when it comes to grassroots hospitals that lack senior medical specialists and sophisticated diagnostic equipment. AI systems can empower such hospitals to make quicker and more precise diagnoses, thereby enabling a more balanced distribution of high-quality medical services at the regional level. One way of mitigating the disparity between the morbidity and mortality of various populations due to cancer is to provide equitable access to reliable computer-aided diagnostic tools, which can identify cancer at an early stage. As AI technology is continuously developed and optimised, it is expected to enhance radiologists' capabilities to diagnose and identify early-stage lung cancer, leading to a decrease in mortality and a considerable improvement in clinical treatment and patient outcomes.

REFERENCES

1. Zheng G, Han G, Soomro NQ.An Inception Module CNN Classifiers Fusion Method on Pulmonary Nodule Diagnosis by Signs[J]. Tsinghua Sci Technol. 2020;25(3):368–83.
2. Zhang S, Sun F, Wang N, Zhang C, Yu Q, Zhang M, et al. Computer-Aided Diagnosis (CAD) of Pulmonary Nodule of Thoracic CT Image Using Transfer Learning[J]. J Digit Imaging 2019;32:995–1007.
3. Zhang N, E L, Wu S, Wu Z.CT texture analysis in differential diagnosis of benign and malignant solitary pulmonary nodules [J]. Chin J Med Imaging Technol. 2018;34(8):1211–15.
4. Vayntrub Y, Gartman E, Nici L, Jankowich MD.Diagnostic Performance of the Herder Model in Veterans Undergoing PET Scans for Pulmonary Nodule Evaluation[J]. Lung. 2021;199(6):653–57.
5. Wang Y, Wu B, Zhang N, Liu J, Zhao L.Research progress of computer aided diagnosis system for pulmonary nodules in CT images[J]. J X-Ray Sci Technol. 2019;28(1):1–16.
6. Singh S, Karimi S, Ho-Shon K, Show HL.Tell and summarise: learning to generate and summarise radiology findings from medical images[J]. Neural Comput Appl. 2021;33(13):7441–65.
7. Ha W, Vahedi Z.Automatic Breast Tumor Diagnosis in MRI Based on a Hybrid CNN and Feature-Based Method Using Improved Deer Hunting Optimisation Algorithm[J]. Comput Intell Neurosci. 2021;2021(3):1–11.
8. Saygili A, Albayrak S.An efficient and fast computer-aided method for fully automated diagnosis of meniscal tears from magnetic resonance images[J]. AI Med. 2019;97(JUN.):118–30.
9. Barinov L, Jairaj A, Becker M, Seymour S, Lee E, Schram A.Impact of Data Presentation on Physician Performance Utilizing AI-Based Computer-Aided Diagnosis and Decision Support Systems[J]. J Digit Imaging. 2019;32(3):408–16.
10. Loku L, Fetaji B, Krsteski A.Automated medical data analyses of diseases using big data[J]. Knowl Int J. 2018;28(5):1719–24.
11. Nasrallah NA, Sears CR.Biomarkers in Pulmonary Nodule Diagnosis: is It Time to Put Away the Biopsy Needle?[J]. Chest. 2018;154(3):467–68.
12. Shimomura M, Ishihara S.A case of pulmonary MALT lymphoma with preoperative definitive diagnosis based on IgH rearrangement in paraffinembeded specimen[J]. J Jpn Ass Chest Surg. 2017;31(2):221–26.
13. Havinga P, Meratnia N, Bahrepour M.AI based event detection in wireless sensor networks[J]. Univ Twente. 2017;85(6):1553–62.
14. Glauner P, Meira JA, Valtchev P, State R, Bettinger F.The Challenge of NonTechnical

Loss Detection using AI: a Survey[J]. Int J Comput Intell Syst. 2017;10(1):760–75.
15. Chen AF, Zoga AC, Vaccaro ARP.Counterpoint: AI in Healthcare[J]. Healthc Transform. 2017;2(2):84–92.
16. Hutson M.AI faces reproducibility crisis[J]. Science. 2018;359(6377):725–26.
17. Lemley J, Bazrafkan S, Corcoran P.Deep learning for consumer devices and services: pushing the limits for machine learning, AI, and computer vision[J]. IEEE Consum Electron Mag. 2017;6(2):48–56.
18. Price S, Flach PA.Computational support for academic peer review: a perspective from AI[J]. Commun ACM. 2017;60(3):70–79.
19. Agrawal A, Gans JS, Goldfarb A.What to expect from AI[J]. MIT Sloan Manage Rev. 2017;58(3):23–26.
20. Labovitz DL, Shafner L, Gil MR, Hanina A.Using AI to Reduce the Risk of Nonadherence in Patients on Anticoagulation Therapy[J]. Stroke. 2017;48(5):1416–19.