

RESEARCH

Open Access



Application of improved Unet network in the recognition and segmentation of lung CT images in patients with pneumoconiosis

Zhongsong Zhou^{1†}, Xin Li^{2†}, Hongbo Ji¹, Xuanhan Xu¹, Zongqi Chang², Keda Wu¹, Yangyang Song¹, Mingkun Kao¹, Hongjun Chen¹, Dongsheng Wu^{2*} and Tao Zhang^{2*}

Abstract

Background Pneumoconiosis has a significant impact on the quality of patient survival. This study aims to evaluate the performance and application value of improved Unet network technology in the recognition and segmentation of lesion areas of lung CT images in patients with pneumoconiosis.

Methods A total of 1212 lung CT images of patients with pneumoconiosis were retrospectively included. The improved Unet network was used to identify and segment the CT image regions of the patients' lungs, and the image data of the granular regions of the lungs were processed by the watershed and region growing algorithms. After random sorting, 848 data were selected into the training set and 364 data into the validation set. The experimental dataset underwent data augmentation and were used for model training and validation to evaluate segmentation performance. The segmentation results were compared with FCN-8s, Unet network (Base), Unet (Squeeze-and-Excitation, SE + Rectified Linear Unit, ReLU), and Unet ++ networks.

Results In the segmentation of lung CT granular region with the improved Unet network, the four evaluation indexes of Dice similarity coefficient, positive prediction value (PPV), sensitivity coefficient (SC) and mean intersection over union (MIoU) reached 0.848, 0.884, 0.895 and 0.885, respectively, increasing by 7.6%, 13.3%, 3.9% and 6.4%, respectively, compared with those of Unet network (Base), and increasing by 187.5%, 249.4%, 131.9% and 51.0%, respectively, compared with those of FCN-8s, and increasing by 14.0%, 31.2%, 4.7% and 9.7%, respectively, compared with those of Unet network (SE + ReLU), while the segmentation performance was also not inferior to that of the Unet ++ network.

Conclusions The improved Unet network proposed shows good performance in the recognition and segmentation of abnormal regions in lung CT images in patients with pneumoconiosis, showing potential application value for assisting clinical decision-making.

[†]Zhongsong Zhou and Xin Li contributed equally to this work.

*Correspondence:

Dongsheng Wu

wdshxsy@163.com

Tao Zhang

scdxzhangtao@163.com

Full list of author information is available at the end of the article



Keywords Pneumoconiosis, Abnormal region, CT image, Improved Unet, Gaussian error linear unit

Introduction

Pneumoconiosis is a common occupational lung disease, which is mainly caused by long-term work in environments containing dust particles. It is widespread in industrialized countries, especially among workers in the mining, construction and metallurgical industries. The incidence of pneumoconiosis varies from region to region, but there is a general trend of high incidence. According to statistics, pneumoconiosis is one of the main causes of occupational disease death in many countries [1]. Therefore, the prevention and treatment of pneumoconiosis are crucial, including strengthening dust control in the working environment, regular physical examination and rational treatment [2].

The pathological changes of pneumoconiosis include fibrosis and scarring of lung tissue and damage to the alveolar walls. These changes lead to decreased lung function, dyspnea and other associated symptoms [3]. The diagnosis of pneumoconiosis usually relies on a comprehensive assessment of clinical symptoms, occupational history, X-ray examination and lung function tests. Chest X-rays and computed tomography (CT) scans are commonly used imaging techniques in the detection and diagnosis of pneumoconiosis.

Currently, image analysis of pneumoconiosis mainly relies on physicians' visual judgment and manual marking segmentation [4], which is subjective, time-consuming and error-prone [5]. In order to improve accuracy and efficiency, researchers began to explore automatic identification and segmentation methods based on computer algorithms [6]. The segmentation methods are mainly based on feature extraction and pattern recognition of chest CT images. Computer algorithms are able to automatically detect and locate abnormal areas of the lungs, such as nodules, plaques and fibrotic lesions, and perform quantitative analysis [7]. These methods reduce human error and improve the consistency and reliability of diagnosis [8].

In recent years, deep learning techniques have also made important breakthroughs in image recognition, localization and segmentation [9–13]. Fully Convolutional Network (FCN) [14] is a widely used method. By replacing the fully connected layers of traditional Convolutional Neural Network (CNN) with convolutional layers, it is able to accomplish the accurate segmentation of contents and objects. For example, in the field of medical imaging, FCN network can be used to accurately segment the tumor regions in medical images such as chest CT or chest MRI through training the network. There are three structures of FCN networks, FCN-32s, FCN-16s, and FCN-8s, among which FCN-8s shows the best

segmentation effect [15]. Ronneberger [16] et al. proposes Unet after improving and expanding on the basis of FCN, which mainly contains the encoding-decoding structure and skip connection, and it has been widely used in the field of image segmentation. However, due to the simple structure of the Unet network(Base), the feature extraction ability of the lesion region is insufficient, resulting in its model training is prone to overfitting phenomenon and relatively poor segmentation results. The Unet++ network proposed by Zhou et al. [17] in 2018 is partially adapting the direct hopping connection between the Unet encoder and decoder to a nested dense hopping connection, but it still does not allow for exploring enough information from multiscale features.

Therefore, in response to the current “difficult diagnosis” of pneumoconiosis, we proposed an improved Unet network, which aimed to fully automate the identification and segmentation of suspected abnormal regions of different sizes and locations of lung CT images in patients with pneumoconiosis, to provide a powerful support for clinical diagnosis and treatment decisions and to improve the prognostic outcomes of patients. In practical application, this model can be used in rural areas to solve the problem of diagnosis and treatment, promote the development of primary medical care, thus alleviating the problem of “difficult to diagnose” pneumoconiosis in rural areas, promoting the development of “Internet+Intelligent Medical Care”, and promoting the implementation of the strategy of Healthy China.

Materials and methods

Patients

The study retrospectively selected patients with pneumoconiosis who were treated at the West China Fourth Hospital of Sichuan University. The inclusion criteria covered the following requirements: (1) individuals with a history of dust exposure; (2) patients whose lung radiographs met or exceed the acceptable quality criteria set out in the GBZ70-2015 guidelines for the diagnosis of occupational pneumoconiosis; and (3) positive cases who had been formally diagnosed with pneumoconiosis and who had obtained diagnostic certificates from qualified units. On the other hand, the exclusion criteria included subjects with pre-existing pulmonary or pleural diseases that would interfere with the diagnosis or grading of pneumoconiosis. These may include but are not limited to, pneumothorax, pleural effusion, or incomplete resection of lung tissue on one side.

A total of 20 patients were included in the study, of which 19 were male and 1 was female. The age distribution was mainly concentrated between 30 and 60 years

old, with only 4 patients younger than 30 years old. 13 patients were between 30 and 60 years old, and there were also 3 patients older than 60 years old. The mean age of the study population was 45 years and the median age was 38 years. The age and gender distribution was shown in Fig. 1.

The study was approved by the Medical Ethics Review Committee of West China Fourth Hospital of Sichuan University, with the review committee approval number HXSU-EC-2,023,042.

Data source and processing

This study used a set of CT image dataset provided by West China Fourth Hospital of Sichuan University and developed a system Based on an innovative algorithm for marker extraction. The algorithm integrated a variety of advanced digital image processing techniques, such as watershed algorithm, region growing algorithm, flood processing, morphology and edge detection [18], which were improved and combined according to the image features in order to achieve segmentation and extraction of CT images for pneumoconiosis. The algorithm had high accuracy, strong reliability and completeness, and had been verified by expert doctors with an accuracy rate of 93.6%. It can help doctors diagnose and treat pneumoconiosis more quickly, shorten the treatment time and improve the treatment effect.

Network structure

8S network

FCN-8s replaces the traditional fully connected layers with fully convolutional layers, allowing the network to accept input images of arbitrary size and output pixel-level predictions of the same size. The FCN model can use FCN-32s, FCN-16s, and FCN-8s structures depending on the segmentation granularity, and 8s is the feature map recovered from 8-fold downsampling to the input size. The FCN-8s network structure was shown in Fig. 2.

Unet network

The layer structure of the encoding area of the Unet network mainly consists two 3×3 convolutional layers represented by light blue, and a max pooling layer of size 2×2 represented by light gray, and the layer structure of the decoding area mainly includes 2×2 up-sampling represented by light gray and 2×2 convolutional layers represented by light green in transposed convolutional layer, and two 3×3 convolutional layers represented by light blue, and all of the convolutions use the Rectified Linear Unit (ReLU) [19] as the activation function. The Unet network structure was shown in Fig. 3.

Unet++ network

Unet++ splits the encoder and decoder parts of the original Unet model into submodules, each of which contains an encoder and a decoder. In the encoder, each submodule downsamples the input image twice consecutively and extracts features using a convolutional layer; in the decoder, each submodule up-samples and fuses the

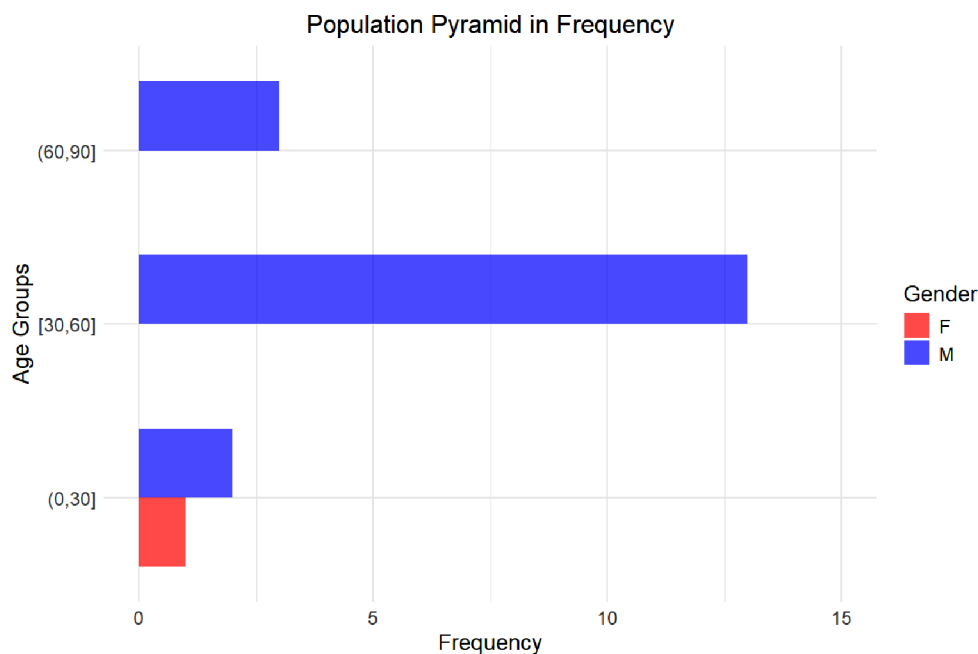


Fig. 1 Population pyramid in frequency

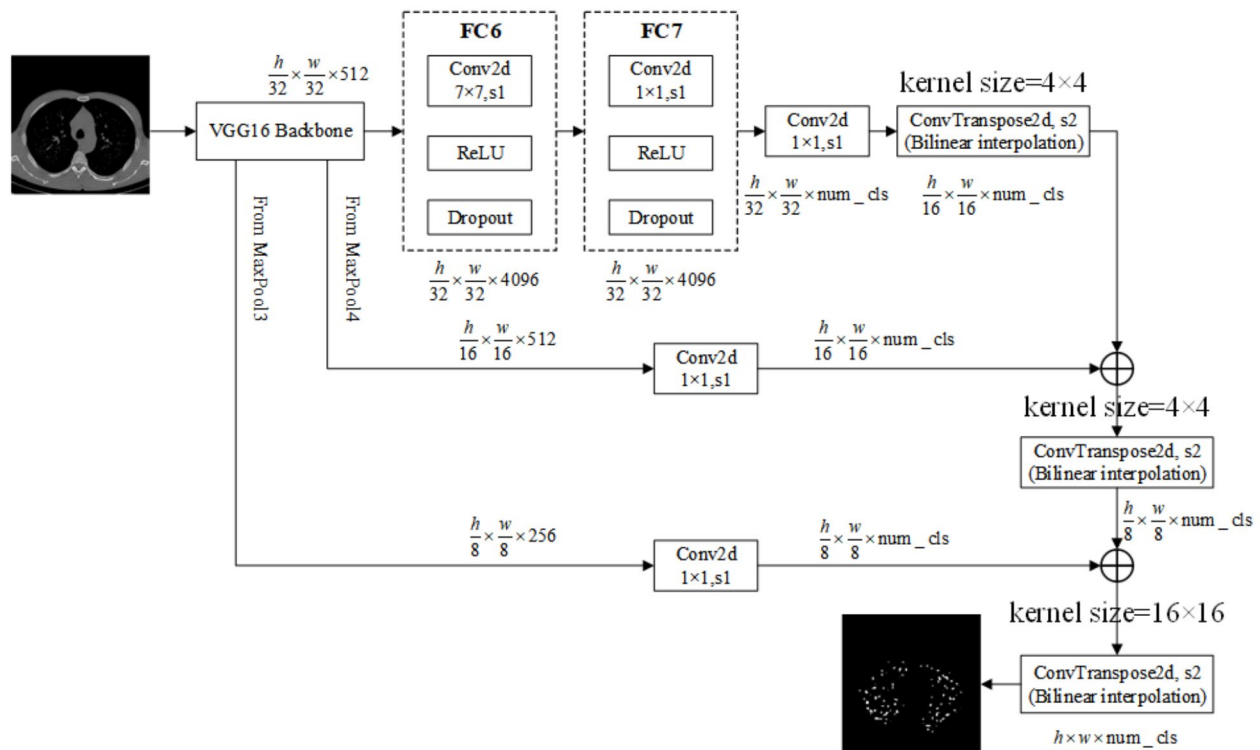


Fig. 2 Schematic diagram of FCN-8s network

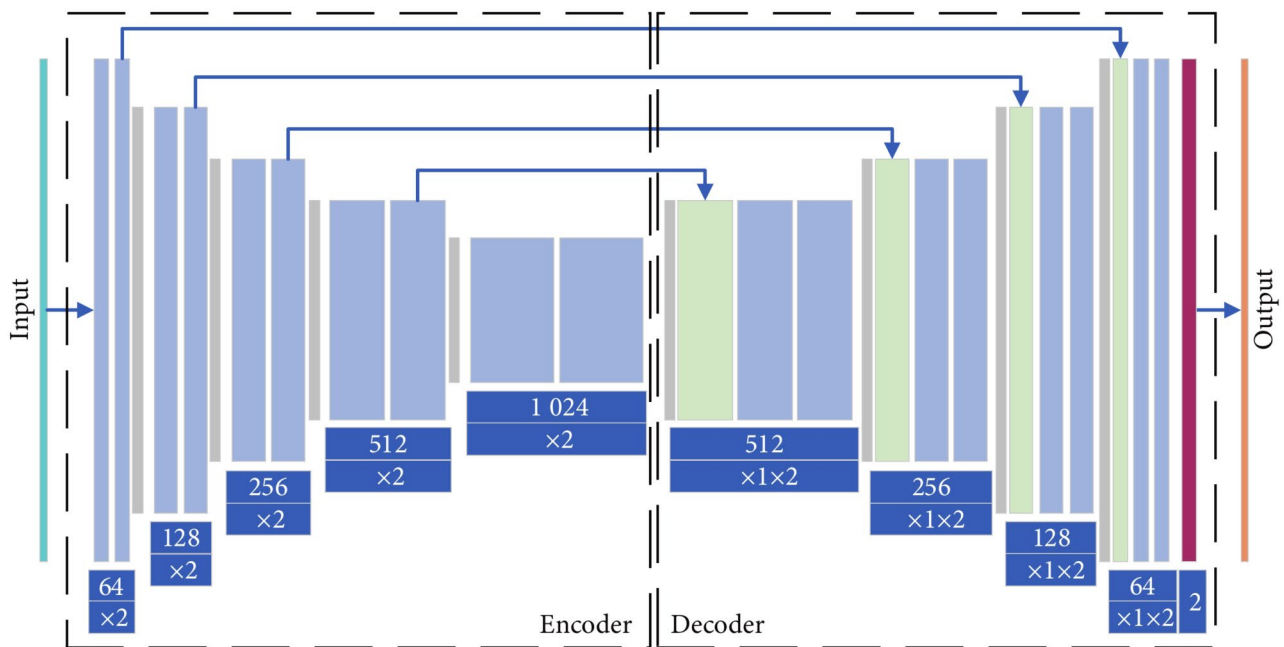


Fig. 3 Schematic diagram of Unet network

output of the previous level with the encoder features, and then performs the decoding operation at the next level. Ultimately, the output of Unet++ is formed by combining the outputs of all the submodules. The structure of Unet++ network was shown in Fig. 4.

Related work

In this study, Gaussian Error Linear Unit (GeLU) [20] was introduced as an activation function in improved Unet network. Although ReLU and GeLU activation functions are widely used in the field of deep learning,

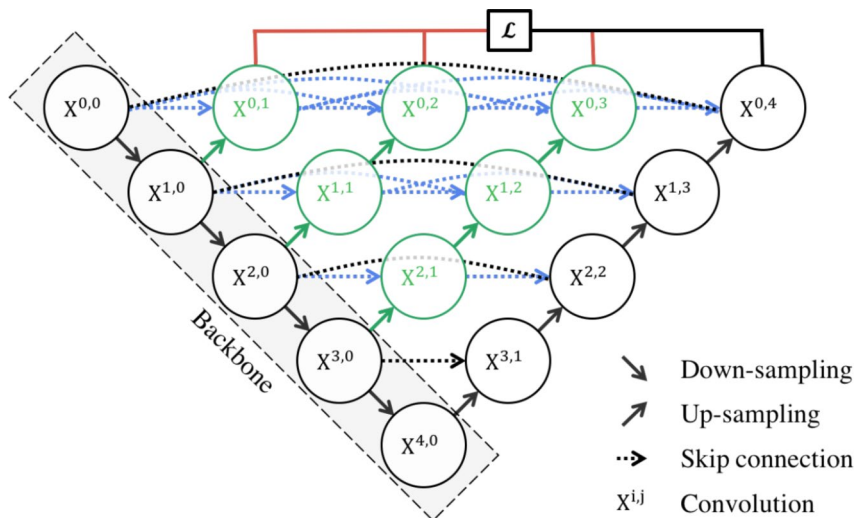


Fig. 4 Schematic diagram of Unet++ network

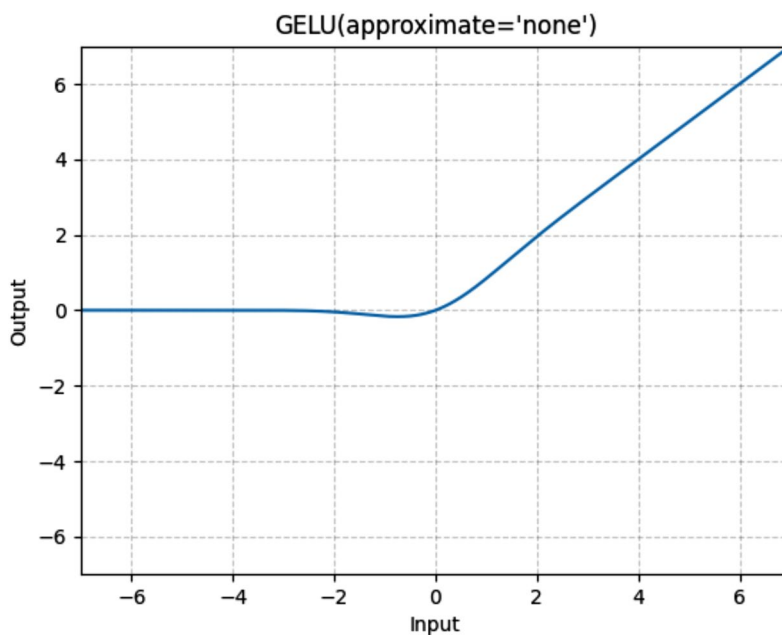


Fig. 5 GeLU activation function

the choice of using GeLU in some specific cases may lead to better performance [21]. The new activation function GeLU solved the problem that ReLU can lead to the disappearance of non-positive gradient and uneven data distribution. GeLU introduced the Sigmoid function on the basis of ReLU, which made its output smoother, and thus it could reduce the problem of oscillations occurring in the neural network [22], and made the model have better generalization ability and accuracy to improve the performance of lung particle segmentation. The formula and function image (Fig. 5) were expressed as follows:

$$GeLU(x) = 0.5 \times x(1 + \tanh \left[\sqrt{\frac{2}{\pi}}(x + 0.0447x^3) \right]) \quad (1)$$

This study incorporated the Squeeze-and-Excitation(SE) attention mechanism into the Unet network, which mainly increased the model’s attention to important features by adaptively adjusting the weights of each channel in the input feature map. Specifically, the SE attention mechanism first compressed the input feature map into a vector using a global average pooling operation, and then learned a weight vector through two fully connected layers, which contained the weight information of each

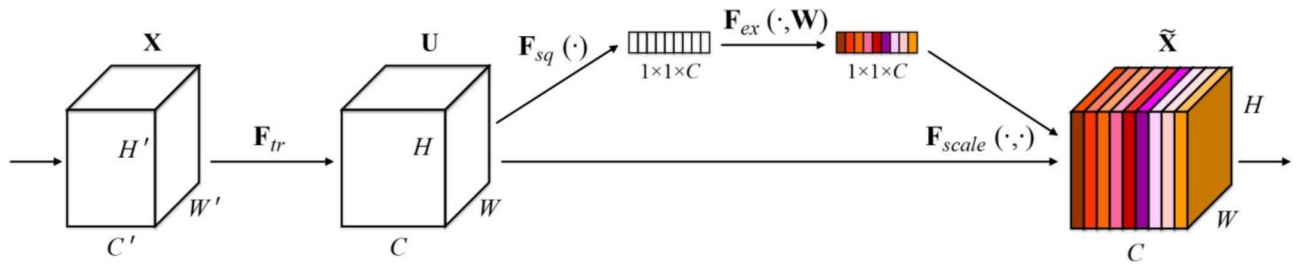


Fig. 6 Structure of SE module

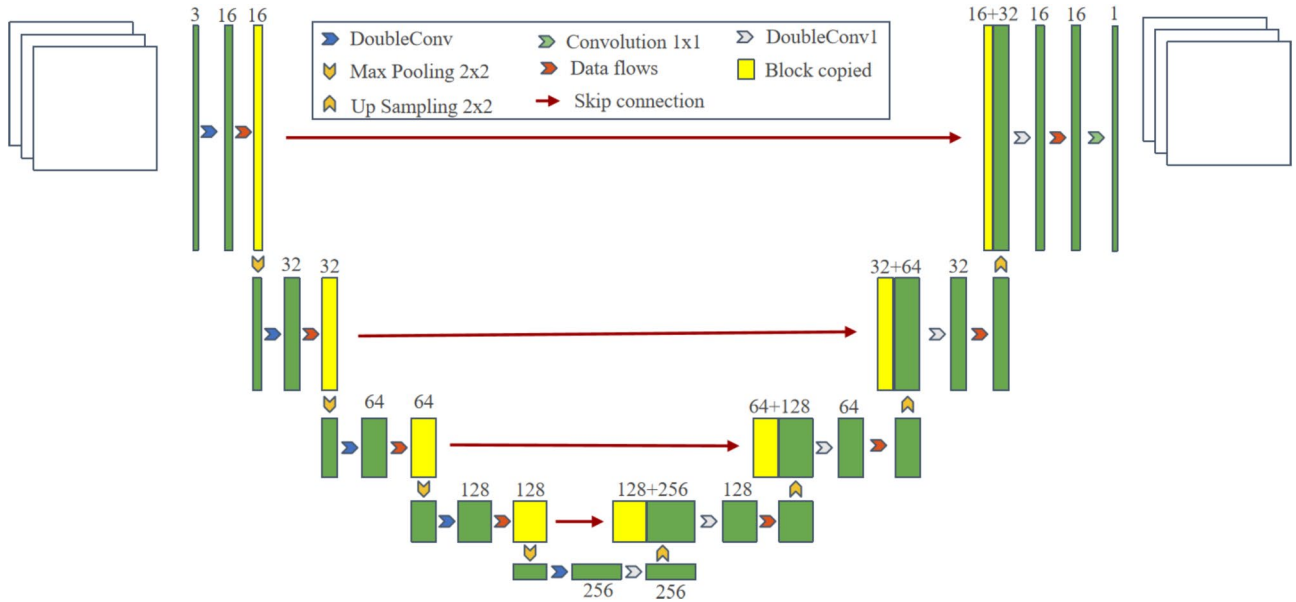


Fig. 7 Schematic diagram of improved Unet network

channel. Finally, this weight vector was multiplied with the original input feature map to obtain the enhanced feature map. In the Unet model, the SE attention mechanism can be used to enhance the attention during feature extraction and reconstruction. The structure of the SE module was shown in Fig. 6.

In this study, the residual mechanism was introduced to improve the doubleconv part of the Unet network, and each convolutional layer was followed by a Gelu activation function and a normalization layer (BatchNorm2d) to improve the feature learning ability of the model [23], and SE attention mechanism was added to Unet network. The structure of improved Unet network was shown in Fig. 7.

At the same time, in order to compare the role of the GeLU activation function, we introduced Unet(SE+ReLU) to compare with the improved Unet(SE+GeLU).

Furthermore, in order to demonstrate the impact of data augmentation techniques on model performance, we conducted ablation experiments.

Moreover, in order to explore whether data distribution and random grouping affect the performance and generalization ability of the model, we conducted a five-fold cross-validation experiment.

Evaluation indicators and methods

In this study, Dice similarity coefficient, positive prediction value (PPV), sensitivity coefficient (SC) and mean intersection over union (MIoU) were used to measure the accuracy of the improved Unet network for segmentation of abnormal regions of lung CT images [24], and to verify the superiority of the segmentation performance of the improved Unet network by comparing it with the metrics of the mainstream network models.

Dice similarity coefficient is the most common measure in medical image segmentation tasks, and it is a measure of the overlap region between labeled and modeled segmentation results. The formula was expressed as follows:

$$\begin{aligned} \textcircled{1} \text{Dice}(A, B) &= 2 \frac{|A \cap B|}{|A| + |B|} \\ \textcircled{2} \text{Dice} &= \frac{2TP}{2TP + FN + FP} \end{aligned} \quad (2)$$

In the formula (2), ① and ② are the two representations of Dice indexes, A and B in Equation ① represent the labeling map and segmentation map contour area, respectively. TP, FP, and FN in Equation ② represent the number of true positive, false positive, and false negative, respectively.

Generally, it is considered that $Dice > 0.7$ indicates that the model segmentation and the doctor's manual segmentation of the region have a high degree of repetition, and the segmentation effect is excellent [25].

The PPV evaluation index indicates the ratio of the number of true positive samples to the number of predicted positive samples. The formula was expressed as follows:

$$PPV = \frac{TP}{TP + FP} \quad (3)$$

The SC evaluation index indicates the ratio of the number of true positive samples to the number of actual positive samples. The formula was expressed as follows.

$$SC = \frac{TP}{TP + FN} \quad (4)$$

The MIoU evaluation index indicates the average of the ratios of the intersection and concatenation of the true label and the predicted value. The formula was expressed as follows.

$$MIoU = \frac{1}{k+1} \sum_{i=0}^k \frac{TP}{FN + FP + TP} \quad (5)$$

Giga Floating-point Operations Per Second (GFLOPS), Params, Frame Per Second (FPS) were used to evaluate the inference speed.

Statistical methods

R statistical software (version 4.3.3) was used for data analysis. The Shapiro-Wilk method was used to test whether the data satisfied normal distribution. Based on the results of normality and variance chi-square tests for each evaluation metric, Wilcoxon signed rank sum test was used to perform the statistical test of the performance evaluation metrics obtained from the improved Unet network and other networks, and the statistical test was corrected by the Benjamini and Hochberg FDR (BH) method, $\alpha = 0.05$ [26].

Results

Results of pre-processing and enhancement of CT image data

The CT image data were further preprocessed before applying the improved Unet network for the experiments. Watershed algorithm was introduced for dividing the labeled image pixels into foreground and background, and three-channel processing was employed to address the possible errors caused by binarization processing. Meanwhile, the accuracy of the segmented images was further improved by combining the region growing algorithm to accurately control and adjusted the segmented images by removing the misdetrcted objects and adding the missed objects.

Next, a flood processing algorithm was applied and the segmented image was regionally grown using different features according to a pre-determined regional guide to obtain more complete and detailed information about the target boundaries, which facilitated the extraction of particles and substance one by one. All image segments underwent uniform processing, and their resolution were adjusted to a size of 512×512 . Finally, the data were cleaned to exclude the effects caused by incomplete extraction of dust particles and other reasons. After selecting 1,212 patient image samples that met the requirements, 848 of them were used as the training set and the other 364 samples as the validation set by randomized sorting, and all Dicom format samples were converted to PNG format images.

Afterward, the original images underwent lung parenchyma extraction, and the corresponding gold standard images were masked. Subsequently, the processed data were resized to a resolution of 512×512 . The experimental data volume were enlarged by data cleaning and supervised single-sample data enhancement operations [27], and geometric transformations and DropBlock regularization methods [28] were used to effectively avoid overfitting problems. A fixed-size rectangular region was cropped from the center of the image using the center crop operation to resize the image to a uniform size. And the CT maps without foreground targets are removed. The data after enhancement were used for model training and validation. The image processing results were shown in Fig. 8.

The results of the ablation study were shown in Table 1.

Parameter setting and segmentation effects of the improved Unet network

The experiment used the Pytorch deep learning framework to build the experimental environment, which ran on Pycharm software. The specific experimental network code is available through the corresponding author.

In this study, the input image shapes for all networks were all $512 \times 512 \times 1$, with a depth of 8. Additionally, all

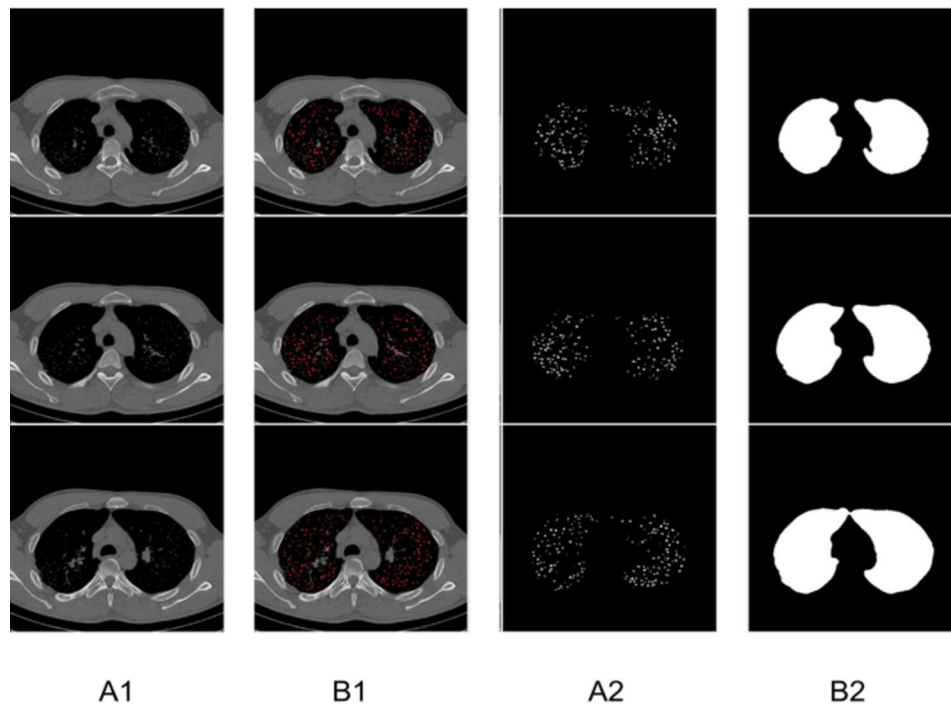


Fig. 8 Schematic diagram of processing results. A1: Original image; B1: Ground truth image; A2: Ground truth image after mask processing; B2: Original image after extracting lung parenchyma

Table 1 Summary of ablation study for the improved unet model and data augmentation

	Dice	PPV	SC	MIoU
Baseline ^a	0.7637	0.7315	0.8503	0.8219
Baseline + Prompts ^b	0.7696	0.7393	0.8502	0.8261
Baseline + Ours ^c	0.7572	0.7164	0.8466	0.8098
Baseline + Prompts + Ours ^d	0.8028	0.8006	0.8820	0.8584

a: Traditional Unet model+unaugmented data; b: Traditional Unet model+augmented data; c: Improved Unet model+unaugmented data; d: Improved Unet model+augmented data

networks utilized the Adam optimizer for optimization. According to the GPU memory and image data size, in order to improve the network training efficiency, as well as for the model to be more stable in the later stages of training, the following excellent empirical values of network parameters were obtained by comparing a large number of parameter adjustment experiments: the number of training samples batch_size for all the networks was set to 2, and the training times of network training as steps_per_epoch was 606 (i.e., the number of data set samples / batch_size = 1212 / 2). The number of training rounds for each round of network training num_classes was set to 2 according to the correlation between the parameters, and the total number of training rounds epochs was set to 100 (the total number of iterations of the network was 60,600), and the learning rate was initialized to 0.00005. The learning rate update strategy and mixed accuracy training were used to optimize the neural

network model to improve the training speed and reduce memory consumption.

The variation of the Loss curve of the improved Unet network was shown in Fig. 9, where it can be observed that the Loss value was higher at the beginning of the training, but with the increase of the number of iterations, the Loss gradually decreased and stabilized. This indicated that the improved Unet network was able to gradually learn a better feature representation during the training process and improve the segmentation performance. The variation of Accuracy curves for the improved Unet network and other networks was shown in Fig. 10.

Comparison of the segmentation effect of the improved Unet network and the base network

Figure 11 illustrated the segmentation effect of the improved Unet network, FCN-8s network, Unet(Base) network, Unet(SE+ReLU) network and Unet++ network.

Results of five-fold cross-validation

Figure 12 illustrated the Accuracy curves of five-fold cross-validation.

Results of multiple comparative experiments

The experimental networks used in this study included the FCN-8s network, the Unet network (Base), the Unet(SE+ReLU) network, the improved

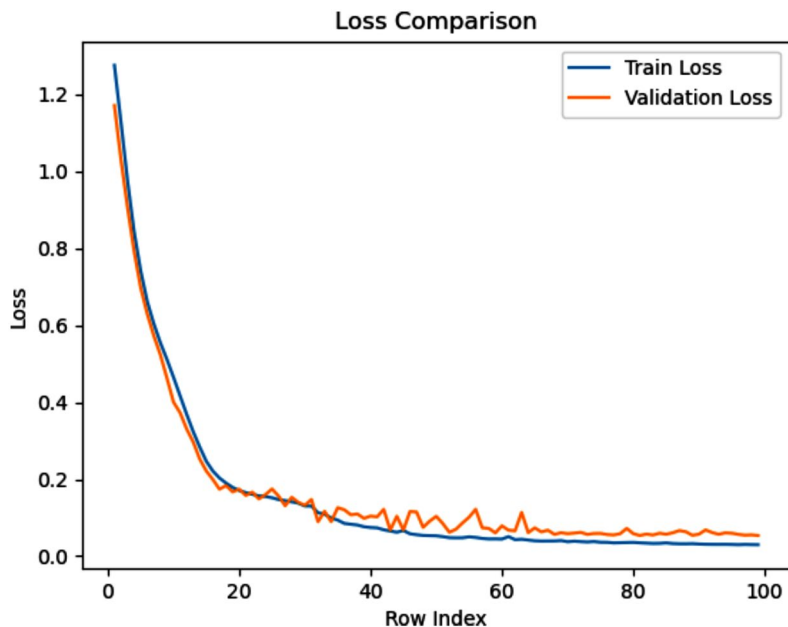


Fig. 9 Loss curve variation diagram

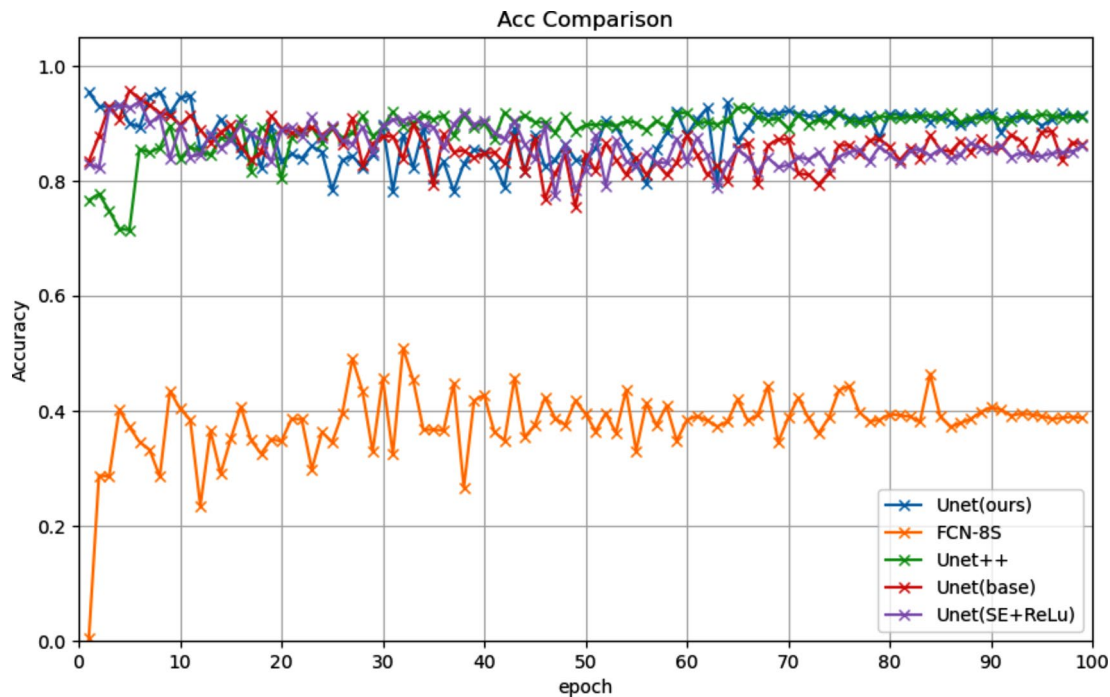


Fig. 10 Accuracy comparison diagram

Unet(SE+GeLU) network and the Unet++ network. And the training parameters of all networks were consistent.

Figure 13 demonstrated the segmentation performance of the improved Unet network and Unet (Base) network on the CT image dataset.

Table 2 demonstrated the segmentation performance of each network on a dataset of lung CT images of pneumoconiosis patients. The median values of Dice, PPV, SC and

MIOU for the improved Unet network in this study were 0.848, 0.884, 0.895 and 0.885, respectively, increasing by 7.6%, 13.3%, 3.9% and 6.4%, respectively, compared with those of Unet network (Base), and increasing by 187.5%, 249.4%, 131.9% and 51.0%, respectively, compared with those of FCN-8s, and increasing by 14.0%, 31.2%, 4.7% and 9.7%, respectively, compared with

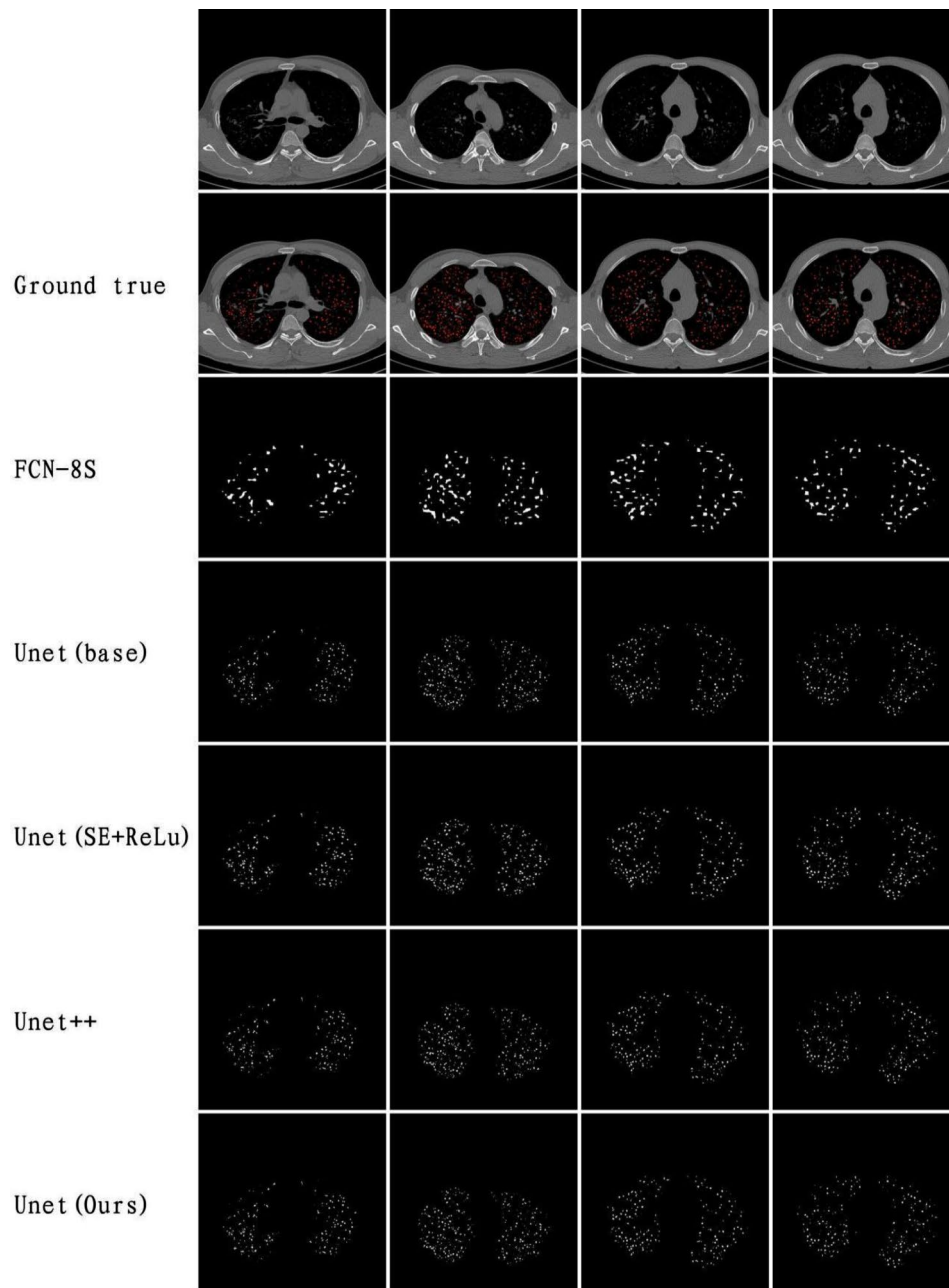


Fig. 11 Schematic diagram of FCN-8s, Unet(Base), improved Unet(Ours), Unet(SE+ReLU) and U-net++ network segmentation result

those of Unet network (SE+ReLU), while the segmentation performance was also not inferior to that of the Unet++ network.

It could be seen that the performance of the improved Unet network was significantly better than that of the Unet network (Base) and FCN-8s. As can be seen from the P-values in Table 1, all of the above performance enhancements were statistically significant except for the Unet++ network, suggesting that the improved Unet network was effective in segmenting the particulate region of the lung.

Table 3 demonstrated the inference speed of different experimental networks.

Discussion

When working with lung CT images of the patients with pneumoconiosis, accurate identification and segmentation of the abnormal areas is crucial for making therapeutic decisions. It helps to prevent the early spread of pulmonary particles, improve patient prognosis, and reduce mortality and disability.

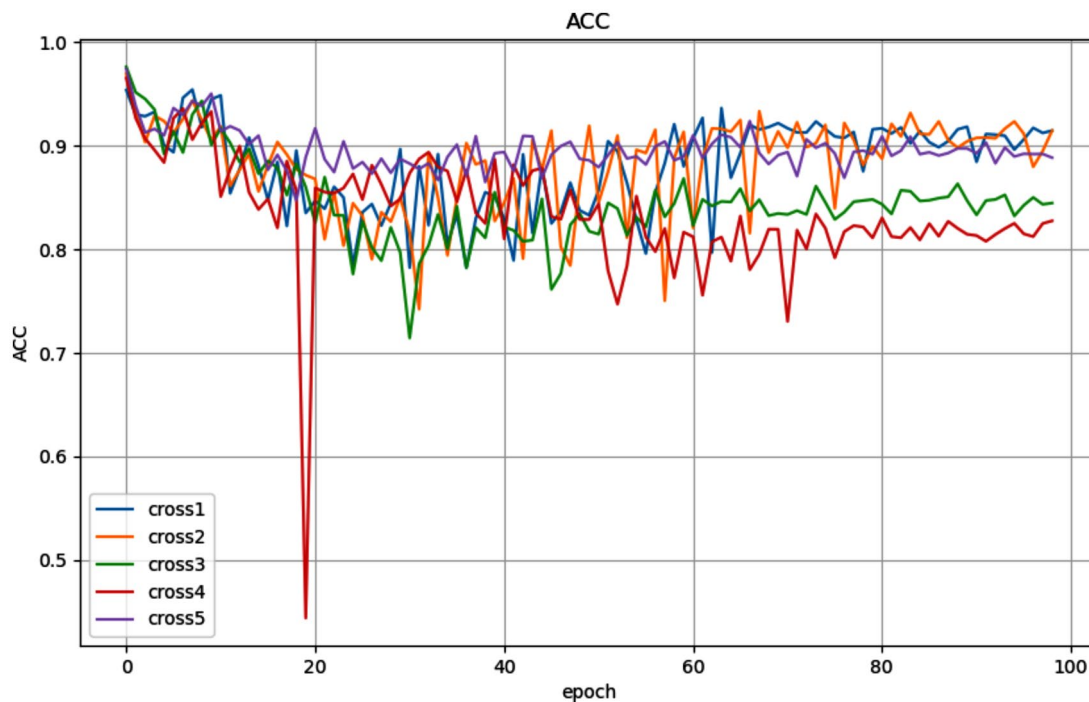


Fig. 12 Results of five-fold cross-validation

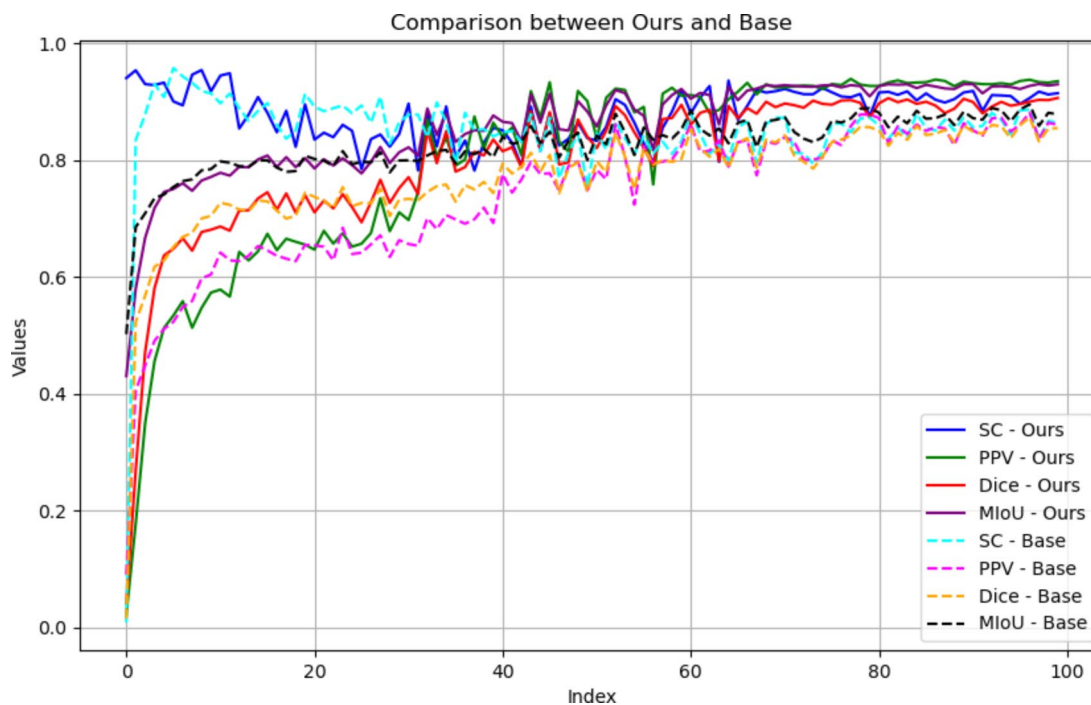


Fig. 13 Comparison of segmentation metrics of improved Unet network and Unet (base) network

The hospital where the patients were diagnosed and treated is the only Tertiary Specialized Hospital for occupational diseases among the 44 hospitals under the auspices of the National Health Commission of China, and it diagnoses and treats patients with pneumoconiosis from all over the southwest region of China. The academic

level of pneumoconiosis diagnosis is in the leading position in China. The epidemiologic characteristics (sex, age) of the patients in this study were similar to those of pneumoconiotics from other regions [29–31].

We could find that data augmentation significantly enhanced the segmentation performance of the model,

Table 2 Data comparison in the experimental network

	Dice			PPV			SC			MIoU		
	Median (P ₂₅ ,P ₇₅)	$\bar{x} \pm s$	P	Median (P ₂₅ ,P ₇₅)	$\bar{x} \pm s$	P	Median (P ₂₅ ,P ₇₅)	$\bar{x} \pm s$	P	Median (P ₂₅ ,P ₇₅)	$\bar{x} \pm s$	P
Unet (Ours)	0.85 (0.74,0.89)	0.80±0.13		0.88 (0.67,0.93)	0.80±0.17		0.89 (0.85,0.92)	0.88±0.04		0.88 (0.80,0.93)	0.86±0.08	
FCN- 8s	0.30 (0.28,0.30)	0.28±0.04	<0.001	0.25 (0.24,0.26)	0.24±0.03	<0.001	0.39 (0.36,0.40)	0.37±0.07	<0.001	0.59 (0.58,0.59)	0.58±0.01	<0.001
Unet (Base)	0.79 (0.73,0.83)	0.77±0.10	<0.001	0.78 (0.65,0.83)	0.74±0.13	<0.001	0.86 (0.84,0.88)	0.85±0.09	<0.001	0.83 (0.80,0.87)	0.82±0.05	<0.001
Unet (SE+ ReLU)	0.74 (0.73,0.82)	0.75±0.10	<0.001	0.67 (0.65,0.82)	0.71±0.13	<0.001	0.86 (0.84,0.89)	0.85±0.09	<0.001	0.81 (0.79,0.86)	0.81±0.05	<0.001
Unet ++	0.87 (0.85,0.88)	0.85±0.05	0.277	0.90 (0.87,0.90)	0.87±0.07	0.871	0.90 (0.87,0.91)	0.88±0.04	0.463	0.90 (0.89,0.91)	0.89±0.04	0.458

Table 3 Inference speed in the experimental network

Model	GFLOPS(G)	Params(M)	FPS(Hz)
Unet(Ours)	82.186	4.506	69.70
FCN-8s	160.604	138.358	51.26
Unet(Base)	81.036	4.318	85.41
Unet(SE+ReLU)	82.186	4.506	71.27
Unet++	301.077	26.905	26.15

especially in our model. We hypothesize that this may be due to the fact that removing redundant information and noise from the edges of the image and retaining the main feature regions of the image improves the performance of the model. By cropping different central regions, multiple different samples can be generated, thus expanding the dataset and improving the generalization ability of the model.

Each of the several networks mentioned in this paper has its own characteristics in processing medical images, and they differ in capturing specific structures and regions: the FCN-8s is a fully convolutional network mainly used for semantic segmentation tasks, and it has certain advantages for segmenting organs and structures in large regions. In medical image segmentation, Unet can effectively capture fine structures and edges in medical images, and has better results for segmenting cells, organs, and lesion regions, etc. Unet++ introduces a deeper network structure and a more complex feature fusion mechanism, which is advantageous for processing complex medical images. However, these methods have a common problem that manual labeling of data has a great influence on the segmentation results and is time-consuming.

To address this issue, this study applied the improved Unet network to the recognition and segmentation of granular regions in lung CT images. By comparing with the other networks, the improved Unet network showed a better segmentation performance. Compared to the FCN-8s network, the improved Unet network had better performance in capturing detail and edge parts, and segmented the target object more accurately. Compared to the Unet(Base) network, the improved Unet network had higher segmentation accuracy and better segmentation results for more accurate localization and segmentation of the target region. Compared with the Unet (SE+ReLU) network, the segmentation results of the improved Unet network were better, showing the superiority of the GeLU activation function. Compared with the Unet++ network, the improved Unet network performed equally in terms of stability and clarity of segmentation results. Therefore, the experimental results showed that the improved Unet network had great advantages in image segmentation tasks and achieved good segmentation results in different scenarios, showing potential application value for clinical decision aid.

This research method improved the doubleconv structure of the Unet network and introduces the residual mechanism, which can effectively avoid the problem of degradation of model performance due to training overfitting caused by the disparity of lung region sizes and the disappearance of lung region detail information after multi-layer convolution. The introduction of the SE attention mechanism can make the neural network pay more attention to the useful feature information, so as to improve the model's accuracy, ppv and sc value and other indicators, to better distinguish the effective features and reduce their impact on the model.

The method in this study improved the activation function of the Unet network, and the GeLU activation function enhances the nonlinear representation and generalization ability of the network, allowing the model to maintain accuracy while being lightweight, and reduces the computational bottleneck and memory cost for model training. The adoption can better accommodate the unbalanced scenario of background and lung particle regions. The method adopted mixed-accuracy training to speed up computation and reduce memory consumption, thus making the training of large-scale models more efficient. Based on this, the use of appropriate learning rate updating strategy can better control the parameters such as gradient size and direction in the optimization process. The joint use of mixed-accuracy training and learning rate updating strategies can better optimize the neural network, thus improving the convergence speed and stability of the whole network.

The proposed method achieved a processing speed of 69.70 FPS, indicating its potential for real-time applications. A higher FPS value signified faster model execution, which was crucial for applications that require quick response times. At the same time, it should be pointed out that the improved Unet network proposed in this study further improves the segmentation performance of the Unet model on the suspected particulate region of the patient's lungs, but the segmentation accuracy still has a certain amount of room for improvement. The lung parenchyma extraction operation was performed in the experimental data processing of this study, and the generalization ability of the model might be further improved if CT image data without extracted lung parenchyma were used to train the network. Currently, the model has been applied to pneumoconiosis and pulmonary nodule. However, to extend this study to cancer, it is essential first to explore the datasets used for such applications. The primary difference between tumors and pneumoconiosis nodules lies in their morphology: tumors tend to be larger in size, whereas pneumoconiosis nodules are smaller, typically measuring only 0.3–0.5 cm. Tumors are generally more isolated, often appearing as single entities, whereas pneumoconiosis nodules are more numerous

and diffusely distributed. Given that the model has successfully identified the smaller and more numerous pneumoconiosis nodules, it is reasonable to expect that it could also identify the larger and less numerous tumors. Thus, this model has potential applications for lung cancer, brain tumors, and other cancers. These are elements that need to be studied and addressed in future research.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12880-024-01377-3>.

Supplementary Material 1

Acknowledgements

We would like to thank all participants of this study.

Author contributions

Conception and design: Tao Zhang, Dongsheng Wu and Zhengsong Zhou; (II) Provision of study materials or patients: Tao Zhang, Dongsheng Wu, Xin Li; (III) Collection and assembly of data: Zhengsong Zhou, Xin Li, Hongbo Ji; (IV) Run of experiments: Hongbo Ji, Xuanhan Xu, Keda Wu, Yangyang Song; (V) Data analysis and interpretation: Zongqi Chang, Mingkun Kao and Hongjun Chen; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Funding

This study was supported by the National Key R&D Program of China (No. 2022YFC3600600), Sichuan Science and Technology Program (No. 2022YFS0293, 2022YFS0109, 2023NSFSC0651).

Data availability

The raw data that support the conclusions of this article will be provided by the corresponding author on reasonable request.

Declarations

Ethical approval

The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. Studies involving human participants were reviewed and approved by the Medical Ethical Review Committee of West China Fourth Hospital, Sichuan University (Number HXSY-EC-2023042). Patients / participants provided their written informed consent to participate in this study. All methods strictly adhered to relevant guidelines and regulations.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Electronic Information Engineering, Chengdu Jincheng College, Chengdu, China

²West China School of Public Health and West China Fourth Hospital, Sichuan University, Chengdu, China

Received: 17 March 2024 / Accepted: 22 July 2024

Published online: 19 August 2024

References

- Kruger RP, Thompson WB, Turner AF. Computer diagnosis of pneumoconiosis. *IEEE Trans Syst Man Cybernetics*. 1974;140–9. <https://doi.org/10.1109/TSMC.1974.5408519>.
- Zhang KF, Liu RZ, Wang Z. Current status of imaging diagnosis of pneumoconiosis. *Wisdom Health*. 2020;6(34):22–4. <https://doi.org/10.19335/j.cnki.2096-1219.2020.34.008>.
- Deng K. Computer-aided diagnosis in pneumoconiosis based on deep learning and traditional image technology. Wuhan: Huazhong University of Science and Technology; 2017.
- Qi XM, Luo Y, Song MY, et al. Pneumoconiosis: current status and future prospects. *Chin Med J*. 2021;134(08):898–907. <https://doi.org/10.1097/cm9.0000000000001461>.
- Wang Z, Zhang JF, Qian QJ. Evolution and development of the diagnostic criteria for occupational pneumoconiosis in China. *Chin J Industrial Med*. 2017;169–71. <https://doi.org/10.13631/j.cnki.zggyyx.2017.01.028>.
- Hua M, Zhang DH, Qiu CY et al. Analysis on examination result of chest film reading at Guangdong provincial training class for pneumoconiosis diagnosis. *Chin J Industrial Med*. 2011;5:364–6. https://kns.cnki.net/kcms2/article/abstract?v=0Vs2Vpqj5wdqYbPZ_DL_CWbRzty-nwa521JBrYDPsepDpD-cBZeqeKlmEaWfM5GdaYVmcEviqQBsnyj94lnntgZnw8zqfwRbtEp-JGi01Khuv90JNY6Lf88fjHGJOKVU%26uniplatform=NZKPT%26language=CHS
- Fujimura N. Pathology and pathophysiology of pneumoconiosis. *Curr Opin Pulm Med*. 2000;6(2):140–4. <https://doi.org/10.1097/00063198-200003000-00010>.
- Wang Z. Application of computer-aided diagnosis in the diagnosis of pneumoconiosis. The world's latest doctor. *Learn Inform Abstract*. 2019;19(8):78–9. <https://doi.org/10.19613/j.cnki.1671-3141.2019.08.035>.
- Shin HC, Roth HR, GAO M, et al. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Trans Med Imaging*. 2016;35(5):1285–98. <https://doi.org/10.1109/tmi.2016.2528162>.
- Pacal I. A novel swin transformer approach utilizing residual multi-layer perceptron for diagnosing brain tumors in MRI images. *Int J Mach Learn Cyber*. 2024. <https://doi.org/10.1007/s13042-024-02110-w>.
- Pacal I, Alaftekin M, Zengul FD. Enhancing skin cancer diagnosis using swin transformer with hybrid shifted window-based multi-head self-attention and SwiGLU-based MLP. *J Digit Imaging Inf med*. 2024. <https://doi.org/10.1007/s10278-024-01140-8>.
- Pacal I, Celik O, Bayram B, et al. Enhancing EfficientNetv2 with global and efficient channel attention mechanisms for accurate MRI-Based brain tumor classification. *Cluster Comput*. 2024. <https://doi.org/10.1007/s10586-024-04532-1>.
- Pacal I, MaxCerVixT. A novel lightweight vision transformer-based approach for precise cervical cancer detection. *Knowl Based Syst*. 2024;289:111482. <https://doi.org/10.1016/j.knosys.2024.111482>.
- Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. *IEEE Trans Pattern Anal Mach Intell*. 2015;39(4):640–51. <https://doi.org/10.1109/cvpr.2015.7298965>.
- Zhou NN. Design and implementation of ultrasonic blood vessel image segmentation and recognition software based on deep learning. Chongqing: Chongqing University; 2019. <https://doi.org/10.27670/d.cnki.gcqdu.2019.001281>.
- Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: *International conference on medical image computing and computer-assisted intervention*. Cham: Springer; 2015, pp. 234–241. <https://doi.org/10.48550/arxiv.1505.04597>.
- Zhou Z, Siddiquee MMR, Tajbakhsh N, et al. Unet++: a nested u-net architecture for medical image segmentation//Deep learning in medical image analysis and multimodal learning for clinical decision support. Cham: Springer; 2018, pp. 3–11.
- Wang GQ, Zhou XH, Wei LL. Image segmentation based on watershed algorithm. *Comput Simul*. 2009;5:255–8. https://kns.cnki.net/kcms2/article/abstract?v=0Q9DRdE49fLY7vDJ0-vJmCGxicZWGIRJcPRbYMs-8bN1ne0nK05aGH4wm1Alq-B24cmiYU3RCMzVOLyiCPA9TK1juHkkrAECotO-OnXglikB1tUXWYvnnvnlqolCV6_aj76c2-NFYZDb4=&uniplatform=NZKPT&lan guage=CHS.
- Kaeas T, Aila T, Laine S et al. Progressive growing of gans for improved quality, stability, and variatio. arXiv: 1710.10196,2017. <https://doi.org/10.48550/arXiv.1710.10196>.
- Hendrycks D, Gimpell K. Gaussian error linear units(ge-lus). arXiv:1606.08415. 2016. <https://doi.org/10.48550/arXiv.1606.08415>.
- Zhao JM, Wu DK, Wang ZF. Medical image segmentation algorithm based on RT-Uet. *Radio Eng*. 2023;53(2):381–6.
- Deng H, Zhang J. Melanoma classification method based on the improved ResNet50 network. Nanjing: Nanjing University of Posts and Telecommunications; 2023. <https://doi.org/10.27251/d.cnki.gnjdc.2023.002076>.
- AlSumairi SB, Ismail MMB. X-ray image based pneumonia classification using convolutional neural network. *ACCENTS Trans Image Process Comput Vis*. 2020;6(20):54. <https://doi.org/10.19101/tpcv.2020.618050>.
- He YS, Jiang JL, Yu H, et al. Comparison of dice coefficient and Hausdorff distance in image segmentation. *Chin J Med Phys*. 2019;36(11):13071311. <https://doi.org/10.3969/j.issn.1005-202X.2019.11.012>.
- Hu J, Shen L, Sun G. Squeeze-and-excitation networks. *Proc IEEE Conf Comput Vis Pattern Recognit*. 2018;7132–41. <https://doi.org/10.1109/tpami.2019.2913372>.
- Zhou ZS, Chen XM, Zhang HY, et al. The application of modified U-Net network in regional recognition and segmentation of brain CT images. *J Sichuan Univ (Medical Edition)*. 2022;53(1):114–20. <https://doi.org/10.12182/20220160302>.
- Perez L, Wang J. The effectiveness of data augmentation in image classification using deep learning(2017). <https://doi.org/10.48550/arXiv.1712.04621>.
- Ghiasi G, Lin TY, Le QV. DropBlock: a regularization method for convolutional networks. In *Proceedings of the 32nd international conference on neural information processing systems (NIPS'18)*. Curran Associates Inc., Red Hook, NY, USA, 10750–10760. <https://doi.org/10.5555/3327546.3327732>.
- Wei F, Xue P, Zhou L, et al. Characteristics of pneumoconiosis in Zhejiang Province, China from 2006 to 2020: a descriptive study. *BMC Public Health*. 2023;23(1):378. <https://doi.org/10.1186/s12889-023-15277-8>.
- Kurth L, Casey ML, Mazurek JM, Blackley DJ. Pneumoconiosis incidence and prevalence among US Medicare beneficiaries, 1999–2019. *Am J Ind Med*. 2023;66(10):831–41. <https://doi.org/10.1002/ajim.23519>.
- Wei TT, Mei LY, Zhang H, Yao YX, Zhen Z. [Epidemiological characteristics and trend of new cases of occupational pneumoconiosis in Hubei Province from 2011 to 2020]. *Zhonghua Lao Dong Wei Sheng Zhi Ye Bing Za Zhi*. 2022;40(6):426–30. Chinese. <https://doi.org/10.3760/cma.j.cn121094-20210420-00224>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.