

Research Article

Interpretable Optimization Training Strategy-Based DCNN and Its Application on CT Image Recognition

Ronghan Wang ¹, Tao Liu,² Junwei Lu,³ and Yuwei Zhou⁴

¹Hunan Institute of Scientific and Technical Information, Changsha, Hunan 410001, China

²The Second People's Hospital of Hunan Province, Changsha, Hunan 410021, China

³Oriental Science and Technology College of Hunan Agricultural University, Changsha, Hunan 410128, China

⁴Changsha Hospital of Traditional Chinese Medicine, Changsha, Hunan 410100, China

Correspondence should be addressed to Ronghan Wang; sixifeiyue@163.com

Received 24 December 2021; Revised 18 February 2022; Accepted 24 February 2022; Published 16 March 2022

Academic Editor: Akemi Gálvez

Copyright © 2022 Ronghan Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The theoretical basis of the discrete random sample batch classification is not clear and the sample division is not scientific during the process of Deep Convolutional Neural Network (DCNN) model training. Aiming at the problems above, starting from the DCNN detection recognition mechanism, the theory of random discrete samples is given and proved, and a scientific quantitative batch of sample input method is proposed. Combined with image preprocessing, based on the strategy of random dispersion of samples, and scientifically quantified sample input batches, the DCNN model is trained with limited label samples, and then the CT image recognition of pulmonary nodules is carried out. Experimental results based on the LIDC-ID-RI public dataset show that the sensitivity, specificity, and accuracy of the proposed method have reached 96.40%, 95.60%, and 96.00%, respectively. Compared with the multiscale convolutional neural network method and the multiscale multimode image fusion method, the recognition accuracy of the proposed method is improved by 1.6 and 3.49 percentage points, respectively.

1. Introduction

Computed tomography (CT) images play an important role in the detection of pulmonary nodules, and it has been an important auxiliary means for doctors to diagnose and treat lung cancer [1]. Modern imaging techniques can help doctors better detect nodules in the lung parenchyma and improve patients' chances of survival. How to rely on the development of science and technology to achieve intelligent image detection and recognition has become a hot issue of concern [2].

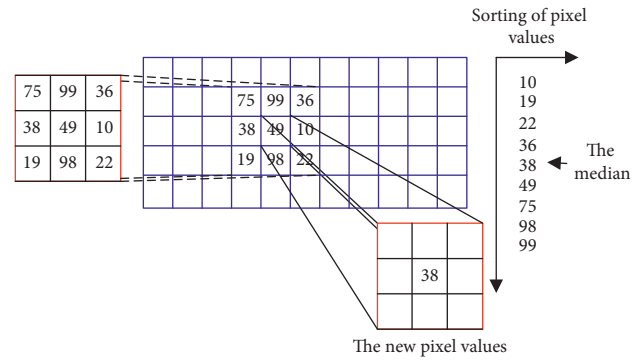
Based on the above requirements for the detection and recognition of pulmonary nodules, many researchers are committed to the realization of adaptive detection and recognition of pulmonary nodules. Zhu and Liu [2] proposed a computer-aided detection algorithm for the automatic detection of pulmonary nodules in CT images. Shi et al. [3] proposed a low-dose CT image pulmonary nodules

detection method based on the convolutional neural network. Sun et al. [4] proposed an automatic detection algorithm for pulmonary nodules based on deep learning using the threshold method, region growth algorithm, and morphological processing. Li et al. [5] proposed an automatic detection method of pulmonary nodules based on a target detection algorithm and proposed a set of pulmonary parenchyma CT image processing processes combining the threshold segmentation algorithm and digital morphology processing. Xi and Liu et al. [6], based on the deep convolutional neural network model, discussed the influence of lung nodule images of different scales and modes on model classification performance and proposed a 2D multiview fusion lung image processing method. The above methods or algorithms can achieve a certain detection and recognition effect of pulmonary nodules based on their respective datasets, but there are still problems of relying on expert experience or poor interpretation. Khan et al. [7] proposed s

a deep learning framework to support the automated detection of lung nodules in computed tomography (CT) images. In their work, a deep learning framework named VGG-SegNet has been used to mine the deep features, and then these features were serially concatenated with the handcrafted features, such as the Grey Level Co-Occurrence Matrix (GLCM), Local-Binary-Pattern (LBP), and Pyramid Histogram of Oriented Gradients (PHOG) to enhance the disease detection accuracy. Sahlol et al. [8] proposed a novel method for detecting tuberculosis in chest radiographs using artificial ecosystem-based optimization of deep neural network features. Połap et al. [9] presented research results on the application of the heuristic method for the detection of over aggregated X-ray images that come from implemented segmentation. The above methods can obtain good verification results on their respective experimental datasets, but there are still some improvements to be made. Most of the above methods still optimize the feature extraction process to obtain better detection and recognition results, but often ignore the influence of the input sample quality and model training strategy on recognition results. In most engineering practical scenarios, the quality and quantity of the labeled samples used for training models are limited, and hence how to complete the detection and recognition of unknown samples in the case of limited effective samples should also become the focus of attention of most researchers. To obtain a higher detection and recognition rate further based on intelligent detection methods has become the focus of this paper.

Affected by instrument sensitivity and transmission channel, CT images will inevitably be interfered with by noise, which will greatly affect doctors' judgment of lesions and may lead to wrong or missed detection [10]. According to the literature research, CT images of pulmonary nodules are often affected by impulse noise [11]. At present, the mainstream noise reduction algorithms mainly include wavelet noise reduction, Gaussian filtering, and other filtering algorithms [12]. As far as the whole process is concerned, the wavelet denoising algorithm is complicated and cumbersome, which is not conducive to online diagnosis. Although the Gaussian filter can have a relatively ideal filtering effect on Gaussian noise, the filtering effect on impulse noise is not very good. However, other filtering algorithms, such as mean filtering and Wiener filtering, are also limited by the computational cost and noise reduction ability; thus, they cannot become a strong universal noise reduction algorithm. Different from the above denoising algorithms, the median filter algorithm can overcome the image blurriness caused by the linear filter algorithm and the high computational cost of the nonlinear filter, and the median filter algorithm has a relatively ideal denoising effect on the impulse noise in CT images [13].

With the significant improvement of computer performance, the deep learning algorithm [14], as an intelligent detection and recognition algorithm, shines brilliantly in more and more scenes. As one of the typical representatives of deep learning algorithms, the Deep Convolutional Neural Network (DCNN) can adaptively extract features from images and perform pattern recognition based on its



classifier [15]. The DCNN method avoids manual extraction of feature indexes and classifier design to the maximum extent, thus effectively improving detection efficiency and intelligence [16]. However, like most methods requiring model training, the performance of DCNN model training has a great influence on the detection and recognition effect [17]. To obtain an ideal training model, the existing label training samples are generally discretized and input into the constructed DCNN model in multiple batches for training [18]. Han et al. [19] obtained a recognition model with better detection performance by training the model in multiple stages. Zhang et al. [20] realized pattern recognition tasks under limited samples by discrete training samples and reducing the training capacity of single batch samples. However, discrete samples and multiple batches of input into the training model are not highly explanatory and lack a rigorous theoretical derivation process. Therefore, the existing DCNN algorithm presents two major problems in the model training process:

First, what is the theoretical basis for the random dispersion of samples? Second, how to scientifically quantify the number of sample input batches?

The deep learning algorithm has been faced with the problem of the "black box" effect since it was put forward. Just because of the "black box" effect, the model training problems encountered by the DCNN algorithm in engineering practice have troubled many researchers. Therefore, aiming at the problems of the lack of theoretical guidance in the DCNN model training strategy and how to scientifically quantify training batches, this paper theoretically deduced why samples are randomly discrete, and further proposed a method of scientifically quantifying the sample input batches. Combined with the above sample pretreatment methods, this paper preprocessed images based on the median filtering algorithm, and then used DCNN under the guidance of interpretable training to detect and identify the pulmonary nodules in CT images so as to comprehensively verify the effectiveness of the proposed method and training strategy.

2. Image Preprocessing

Based on the fact that the median filtering algorithm can take into account both the efficiency of the linear filter and the

accuracy of the nonlinear filter, the median filtering algorithm is used to preprocess the sample before the sample is input into the DCNN detection model for training or testing so as to effectively reduce the sample noise and enhance the details. The median filtering process is shown in Figure 1. Firstly, the median value of local pixels in the image is obtained, and then the local correlation is used to reset the pixels so as to effectively remove the scattered salt and pepper noise. The specific steps are as follows:

- (1) Reorder local pixels according to their size
- (2) Select the median value of pixel sequence as the new pixel value
- (3) Move the median filter scale window and carry out the contents contained in the first and second steps again until the whole image completes traversal denoising

3. Deep Convolutional Neural Network (DCNN)

DCNN implements image feature enhancement with the help of convolution operation and carries out down-sampling operation relying on the local correlation of images to achieve fast dimension reduction of samples. Figure 2 shows the basic internal calculation process of DCNN. Firstly, convolution of the denoised image ergodic formula is obtained using convolution kernels f_x of different scales. Secondly, the local correlation principle is used for noise reduction and feature enhancement, and the enhanced image c_x is obtained by fine-tuning and summing b_x . Thirdly, the lower sampling layer is fine-tuned by weighting ω_x and bias b_{x+1} . Finally, the image features are obtained by activating the Sigmoid function S_{x+1} . A typical DCNN iterates the following convolution and down-sampling processes several times, and then deeply excavates the input samples to obtain the image feature information adaptively, thus providing detection and recognition results based on its classifier.

Similar to the neural network, the training process of the DCNN model includes the determination of forwarding parameters and the adjustment of reverse parameters, and the forward and reverse processes are integrated to determine the minimum error of model reconstruction. Let the sample set composed of m samples be $\{(\mathbf{x}^{(1)}, \mathbf{y}^{(1)}), \dots, (\mathbf{x}^{(m)}, \mathbf{y}^{(m)})\}$, which belong to n categories. $\mathbf{y}^{(i)}$ are the corresponding category labels of the samples $\mathbf{x}^{(i)}$; the training objective function of the DCNN model can be expressed as

$$J(\omega, \mathbf{b}) = \frac{1}{m} \sum_{i=1}^m J(\omega, \mathbf{b}; \mathbf{x}^i, \mathbf{y}^i) \quad (1)$$

$$= \frac{1}{m} \sum_{i=1}^m \frac{1}{2} \cdot \|h_{\omega, \mathbf{b}}(\mathbf{x}^{(i)}) - \mathbf{y}^{(i)}\|^2$$

In the formula, the meanings of ω and \mathbf{b} are the same as those in Figure 2, and $h_{\omega, \mathbf{b}}(\mathbf{x}^{(i)})$ indicates the detection and recognition results. The gradient descent method is used to

minimize the objective function $J(\omega, \mathbf{b})$, and the iterative formula in the process is

$$\omega_{ij}^{(l)} = \omega_{ij}^{(l)} - \alpha \frac{\partial}{\partial \mathbf{b}_i^{(l)}} J(\omega, \mathbf{b}), \quad (2)$$

$$\mathbf{b}_i^{(l)} = \mathbf{b}_i^{(l)} - \alpha \frac{\partial}{\partial \mathbf{b}_i^{(l)}} J(\omega, \mathbf{b}). \quad (3)$$

in which α is the learning rate. Partial derivatives of (2) and (3) are obtained based on the BP algorithm. Firstly, the forward propagation is carried out to calculate the output value $h_{\omega, \mathbf{b}}(\mathbf{x}^{(i)})$ of the last layer of the network, and then the direct gap between the predicted value of the sample tag $h_{\omega, \mathbf{b}}(\mathbf{x}^{(i)})$ and the actual tag is calculated, which is defined as $\delta_i^{(nl)}$ (nl represents the output layer). Then, the residual of each layer is obtained by calculating the residual of the final output layer so as to calculate the partial derivatives of (2) and (3).

The residual calculation formula of the last layer of the traditional neural network is

$$\delta_i^{(nl)} = \frac{\partial J(\omega, \mathbf{b})}{\partial \mathbf{Z}_i^{(nl)}} \quad (4)$$

$$= \frac{1}{2} \frac{\partial}{\partial \mathbf{Z}_i^{(nl)}} \|h_{\omega, \mathbf{b}}(\mathbf{x}^{(i)}) - \mathbf{y}^{(i)}\|^2.$$

in which $\mathbf{Z}_i^{(l)}$ is the weighted sum of the input of unit i at layer l and $\mathbf{Z}_i^{(nl)}$ is the sum of the weighted input for unit i of the last layer.

4. Explainable Training Strategies for DCNN

As an intelligent image detection algorithm, the image recognition performance of DCNN depends on whether the DCNN model acquired after training has a strong enough generalization ability, i.e., whether it can achieve ideal recognition results for images with different features. As mentioned above, in order to obtain the DCNN model with good detection and recognition ability, researchers often adopt the model training strategy of random dispersion of training samples and small-capacity, multibatch input. But there are still two thorny problems: First, the random dispersion of samples lacks the theoretical basis; second, the setting method of the sample training batch is not clear. Next, this paper will focus on the detection and recognition mechanism of DCNN and explain the first problem through mathematical derivation. At the same time, aiming at the second problem, a scientific quantitative method of batch division is proposed. For the convenience of elaboration, each batch training acquisition model is defined as a single-batch Deep Convolutional Neural Network (BDCNN).

The DCNN model obtained by image sample training can be expressed as $\{p(\mathbf{x}_k, \theta_k), k = 1, 2, \dots, K\}$, where $p(\cdot)$ is the detection and recognition model; $\mathbf{x}_k \in \mathbf{X}$; \mathbf{X} is the total image training sample set; $\text{size}(\mathbf{X}) = K * n$; K is the number of input batches or the number of BDCNN obtained by image samples; n is the sample capacity of a single batch; θ_k

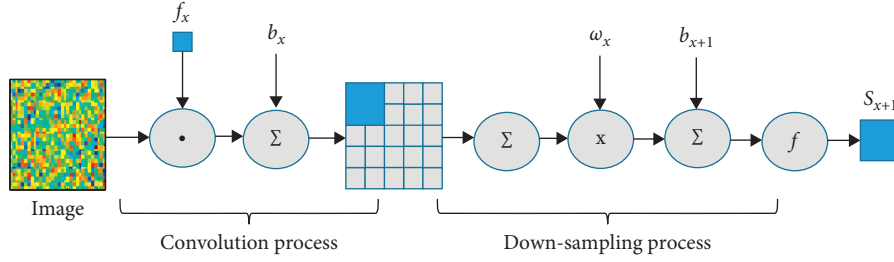


FIGURE 2: Convolution and down-sampling process of the DCNN.

is the set of model parameters obtained by training a single batch. To define the generalization error of the DCNN model:

$$\mathbf{PE} \stackrel{\text{def}}{=} P_{X,Y}(av_k I(p(\mathbf{X}, \boldsymbol{\theta}_k) = \mathbf{Y}) - \max_{J \neq Y} av_k I(p(\mathbf{X}, \boldsymbol{\theta}_k) = \mathbf{J}) < 0), \quad (5)$$

in which $P_{X,Y}(\cdot)$ is the probability function, $I(\cdot)$ is the discriminant function, $av_k(\cdot)$ is the mean function, Y is the correct label set, and J is the misclassified label set. The generalization error can be used to measure the image detection and recognition performance of the DCNN model.

Conclusion to be proved: The generalization error of the DCNN training model is positively correlated with the correlation between BDCNN, and negatively correlated with the detection and recognition performance of BDCNN.

The proof process is as follows. Before that, five definitions or property theorems are stated:

- (1) The BDCNN correlation refers to the correlation between the DCNN models acquired by a single batch of training, and the correlation expression is described in detail in the subsequent proof part of this paper.
- (2) The detection and recognition performance of BDCNN is the sample recognition ability of the DCNN model acquired by a single batch of training, and there will be a detailed mathematical expression in the subsequent derivation process of this paper.
- (3) Convergence almost everywhere [21]: Let ξ and $\{\xi_n\}$ be a sequence of random variables defined in the probability space (Ω, F, P) . If there is a zero-test set Ω_0 , i.e., $\Omega_0 \in F$, $P(\Omega_0) = 0$, and $\forall \omega \in \Omega/\Omega_0$, there is $\xi_n(\omega) \rightarrow \xi(\omega)$, and ξ_n converge almost everywhere to ξ , i.e., $\xi_n \xrightarrow{a.s.} \xi$.
- (4) Borel's law of strong numbers [22]: Suppose $\{\xi_n\}$ is a sequence of independent and identically distributed random variables in the probability space (Ω, F, P) , $P(\xi_n = 1) = p$, $P(\xi_n = 0) = 1 - p$, $0 < p < 1$, $S_n = \sum_{k=1}^n \xi_k$, and $(S_n/n) \xrightarrow{a.s.} p$.
- (5) Chebyshev's inequality [23]: For a random variable x , if expectation Ex and variance Dx exist, then $\forall \xi > 0$, and $P\{|x - Ex| \geq \xi\} \leq Dx/\xi^2$.

It can be proved that the following conclusions can be obtained: When the number of BDCNN $K \rightarrow \infty$, the

following expressions have a convergence relation at every place:

$$\mathbf{PE}^* \xrightarrow{a.s.} P_{X,Y}(P_\theta(p(\mathbf{X}, \boldsymbol{\theta}) = \mathbf{Y}) - \max_{J \neq Y} P_\theta(p(\mathbf{X}, \boldsymbol{\theta}) = \mathbf{J}) < 0). \quad (6)$$

Proof process: based on the property theorem (3), comparing Equation (5) and Equation (6), it can be seen that, to prove Equation (6), it only needs to prove that for any j , there exists a null test set C in the value space of $(\theta_1, \theta_2, \dots, \theta_k, \dots)$, so that for all samples x except C , the following expression is true:

$$\frac{1}{K} \sum_{k=1}^K I(p(\mathbf{x}, \boldsymbol{\theta}_k) = j) \xrightarrow{K \rightarrow \infty} P_\theta(p(\mathbf{x}, \boldsymbol{\theta}_k) = j). \quad (7)$$

For the limited training samples and test sample set, the training batches corresponding to x of $p(\mathbf{x}, \theta) = j$ are limited. Let the training batches be S_1, S_2, \dots, S_R , where R is a finite number. For the set with a total number of N samples, $R \leq 2^N$. Define $\varphi(\theta_k) = r$ when $\{x: h(\mathbf{x}, \theta) = j\} = S_r$, and let K_r represent the times $\varphi(\theta_k) = r$ in K BDCNN, where $k = 1, 2, \dots, K$; then:

$$\frac{1}{K} \sum_{k=1}^K I(p(\mathbf{x}, \boldsymbol{\theta}_k) = j) = \frac{1}{K} \sum_{r=1}^R K_r I(\mathbf{x} \in S_r). \quad (8)$$

According to (8), the left and right sides of the equation represent the number of BDCNN misclassified. When $K \rightarrow \infty$, according to Borel's law of strong numbers, there will be

$$\frac{K_r}{K} = \frac{1}{K} \sum_{k=1}^K I(\varphi(\boldsymbol{\theta}_k) = r) \xrightarrow{a.s.} P_\theta(\varphi(\boldsymbol{\theta}) = r). \quad (9)$$

Therefore, for any j , there exists a null test set C in the value space of $(\theta_1, \theta_2, \dots, \theta_k, \dots)$, such that for all samples x except C , the following expression holds:

$$\frac{1}{K} \sum_{k=1}^K I(p(\mathbf{x}, \boldsymbol{\theta}_k) = j) \rightarrow \sum_{r=1}^R P_\theta(\varphi(\boldsymbol{\theta}) = r) I(\mathbf{x} \in S_r) = P(p(\mathbf{x}, \boldsymbol{\theta}) = j)_\theta. \quad (10)$$

Thus, (6) is deduced.

Next, the margin function $mr(\mathbf{X}, \mathbf{Y})$ of DCNN is defined as

$$mr(\mathbf{X}, \mathbf{Y}) \stackrel{\text{def}}{=} P_{\theta}(p(\mathbf{X}, \theta) = \mathbf{Y}) - \max_{j \neq \mathbf{Y}} P_{\theta}(p(\mathbf{X}, \theta) = j). \quad (11)$$

The margin function represents the degree to which the number of correctly classified BDCNN exceeds the number of misclassified BDCNN during model training. The larger the value of $mr(\mathbf{X}, \mathbf{Y})$, the higher the confidence of the DCNN model obtained by training.

Note

$$J(\mathbf{X}, \mathbf{Y}) = \operatorname{argmax}_{j \neq \mathbf{Y}} P_{\theta}(p(\mathbf{X}, \theta) = j). \quad (12)$$

$mr(\mathbf{X}, \mathbf{Y})$ can be rewritten as

$$\begin{aligned} mr(\mathbf{X}, \mathbf{Y}) &= P_{\theta}(p(\mathbf{X}, \theta) = \mathbf{Y}) - P_{\theta}(p(\mathbf{X}, \theta) = J(\mathbf{X}, \mathbf{Y})) \\ &= E_{\theta}(I(p(\mathbf{X}, \theta) = \mathbf{Y}) - I(p(\mathbf{X}, \theta) = J(\mathbf{X}, \mathbf{Y}))). \end{aligned} \quad (13)$$

Note

$$rmg(\theta, \mathbf{X}, \mathbf{Y}) = I(p(\mathbf{X}, \theta) = \mathbf{Y}) - I(p(\mathbf{X}, \theta) = J(\mathbf{X}, \mathbf{Y})). \quad (14)$$

And then

$$mr(\mathbf{X}, \mathbf{Y}) = E_{\theta} rmg(\theta, \mathbf{X}, \mathbf{Y}). \quad (15)$$

It has been proved that $\mathbf{PE}^* \xrightarrow{a.s.} P_{X,Y}(mr(\mathbf{X}, \mathbf{Y}) < 0)$, and the upper bound of the generalization error \mathbf{PE}^* can be obtained based on the analysis of $P_{X,Y}(mr(\mathbf{X}, \mathbf{Y}) < 0)$. In order to show that the detection and recognition results of the DCNN model are credible, introducing the $E_{X,Y}mr(\mathbf{X}, \mathbf{Y})$ and $E_{X,Y}mr(\mathbf{X}, \mathbf{Y}) > 0$ represents the expected degree of DCNN on the classification results of each sample. According to Chebyshev's inequality,

$$\begin{aligned} P_{X,Y}(mr(\mathbf{X}, \mathbf{Y}) < 0) &= P_{X,Y}(mr(\mathbf{X}, \mathbf{Y}) - E_{X,Y}mr(\mathbf{X}, \mathbf{Y}) < E_{X,Y}mr(\mathbf{X}, \mathbf{Y})) \\ &< P_{X,Y}(|mr(\mathbf{X}, \mathbf{Y}) - E_{X,Y}mr(\mathbf{X}, \mathbf{Y})| > E_{X,Y}mr(\mathbf{X}, \mathbf{Y})) \\ &\leq \frac{\operatorname{var}_{X,Y}(mr(\mathbf{X}, \mathbf{Y}))}{E_{X,Y}mr(\mathbf{X}, \mathbf{Y})^2}. \end{aligned} \quad (16)$$

So, there will be

$$\mathbf{PE}^* \leq \frac{\operatorname{var}_{X,Y}(mr(\mathbf{X}, \mathbf{Y}))}{E_{X,Y}mr(\mathbf{X}, \mathbf{Y})^2}. \quad (17)$$

The detection and recognition capability of BDCNN is defined as s , the average correlation between BDCNN is $\bar{\rho}$, and the expressions are

$$s = E_{X,Y}mr(\mathbf{X}, \mathbf{Y}), \quad (18)$$

$$\bar{\rho} = \frac{E_{\theta, \theta^*}(\rho(\theta, \theta^*)sd(\theta)sd(\theta^*))}{E_{\theta, \theta^*}(sd(\theta)sd(\theta^*))}, \quad (19)$$

in which $\rho(\theta, \theta^*)$ represents the correlation between $rmg(\theta, \mathbf{X}, \mathbf{Y})$ and $rmg(\theta^*, \mathbf{X}, \mathbf{Y})$, and $sd(\theta)$ represents the standard deviation of $rmg(\theta, \mathbf{X}, \mathbf{Y})$.

The upper bound of \mathbf{PE}^* represented by s and $\bar{\rho}$ is obtained based on the following proof procedure.

For variables θ and θ^* , if $(E_{\theta}f(\theta))^2 = E_{\theta, \theta^*}f(\theta)f(\theta^*)$, the calculation $\operatorname{var}_{X,Y}(mr(\mathbf{X}, \mathbf{Y}))$ is as follows:

$$\begin{aligned} \operatorname{var}_{X,Y}(mr(\mathbf{X}, \mathbf{Y})) &= E_{X,Y}(mr(\mathbf{X}, \mathbf{Y}))^2 - (E_{X,Y}mr(\mathbf{X}, \mathbf{Y}))^2 \\ &= E_{X,Y}(E_{\theta}rmg(\theta, \mathbf{X}, \mathbf{Y}))^2 - (E_{\theta}E_{X,Y}rmg(\theta, \mathbf{X}, \mathbf{Y}))^2 \\ &= E_{X,Y}E_{\theta, \theta^*}(rmg(\theta, \mathbf{X}, \mathbf{Y})rmg(\theta^*, \mathbf{X}, \mathbf{Y})) \\ &\quad - E_{\theta, \theta^*}(E_{X,Y}rmg(\theta, \mathbf{X}, \mathbf{Y})E_{X,Y}rmg(\theta^*, \mathbf{X}, \mathbf{Y})) \\ &= E_{\theta, \theta^*}(E_{X,Y}(rmg(\theta, \mathbf{X}, \mathbf{Y})rmg(\theta^*, \mathbf{X}, \mathbf{Y})) \\ &\quad - E_{X,Y}rmg(\theta, \mathbf{X}, \mathbf{Y})E_{X,Y}rmg(\theta^*, \mathbf{X}, \mathbf{Y})) \\ &= E_{\theta, \theta^*}(\operatorname{cov}_{X,Y}(rmg(\theta, \mathbf{X}, \mathbf{Y})E_{X,Y}rmg(\theta^*, \mathbf{X}, \mathbf{Y}))) \\ &= E_{\theta, \theta^*}(\rho(\theta, \theta^*)s d(\theta)sd(\theta^*)). \end{aligned} \quad (20)$$

Substitute (20) into (18) and (19), and $E\xi^2 \geq (E\xi)^2$; then, (21) can be obtained as

$$\begin{aligned}
 \text{var}_{X,Y}(mr(\mathbf{X}, \mathbf{Y})) &= \bar{\rho}(E_{\theta}\mathbf{sd}(\boldsymbol{\theta}))^2 \leq \bar{\rho}E_{\theta}(\mathbf{sd}(\boldsymbol{\theta}))^2 \\
 &= \bar{\rho}E_{\theta}\text{var}_{X,Y}(\text{rmg}(\boldsymbol{\theta}, \mathbf{X}, \mathbf{Y})) \\
 &= \bar{\rho}E_{\theta}\left(E_{X,Y}(\text{rmg}(\boldsymbol{\theta}, \mathbf{X}, \mathbf{Y}))^2 - (E_{X,Y}\text{rmg}(\boldsymbol{\theta}, \mathbf{X}, \mathbf{Y}))^2\right) \\
 &= \bar{\rho}\left(E_{\theta}E_{X,Y}(\text{rmg}(\boldsymbol{\theta}, \mathbf{X}, \mathbf{Y}))^2 - E_{\theta}(E_{X,Y}\text{rmg}(\boldsymbol{\theta}, \mathbf{X}, \mathbf{Y}))^2\right) \\
 &\leq \bar{\rho}\left(E_{\theta}E_{X,Y}(\text{rmg}(\boldsymbol{\theta}, \mathbf{X}, \mathbf{Y}))^2 - (E_{\theta}E_{X,Y}\text{rmg}(\boldsymbol{\theta}, \mathbf{X}, \mathbf{Y}))^2\right) \\
 &\leq \bar{\rho}(1 - s^2).
 \end{aligned} \tag{21}$$

Then, (22) can be obtained:

$$PE^* \leq \frac{\bar{\rho}(1 - s^2)}{s^2}. \tag{22}$$

Therefore, the generalization error of the DCNN training model is positively correlated with the correlation between BDCNN, and negatively correlated with the detection and recognition performance of BDCNN.

To sum up, the DCNN training model's generalization error can be reduced by reducing the correlation between BDCNN and improving the classification intensity of BDCNN so as to enhance the DCNN model's generalization ability and improve the confidence of detection and recognition results.

Therefore, during the sample input process, the correlation between BDCNN is generally reduced through sample randomization. In order to improve the detection and recognition performance of BDCNN, samples can be randomly dispersed and scientifically quantified into batches from the perspective of sample training. To a certain extent, the random dispersion of samples ensures the strong generalization ability of the trained network model and avoids the training model falling into the local optimum. At the same time, when the total number of training samples is constant, the number of sample input batches depends on the number of batch sample input and the number of iterations of the model training cycle.

In order to further illustrate the random discrete process of samples, the specific algorithm flow is given Algorithm 1.

As for how to scientifically quantify the sample input batches, as shown in Figure 3, this paper takes the final model generalization error PE^* as the cost function, and realizes the optimization of the sample input batch times through the grid optimization method based on the topographic map of the model generalization error PE^* under the different iterations obtained and the batch sample size.

In conclusion, the steps of the pulmonary nodule detection and recognition method based on image pretreatment and interpretable training-guided DCNN are shown in Figure 4. In other words, the denoising pretreatment of label samples is carried out based on the

median filtering algorithm, and then the random discrete input of the DCNN model is used for training and recognition. Based on the sample random discretization criterion mentioned above, the pre-processed training samples are randomly discretized and then input into the DCNN model to obtain a detection and recognition model with better generalization performance.

5. Experiment and Result Analysis

In order to demonstrate the effectiveness of the proposed method and training strategy, this paper carried out method validation based on the online public Lung Image dataset, LDC-ID-RI (Lung Image Database Consortium) [24], and conducted a comparative analysis with the methods of related literature under the same dataset. The LDC-ID-RI dataset is provided by the National Cancer Institute of the United States, and each sample is in the standard DICOM format of 512×512 pixels. Taking dataset included in the 1018 research dataset as an example, the samples have been carried out by four experienced breast radiologists the first physician to separate samples diagnosis and the diagnosis is given, then the diagnosis of three other doctors to visit. Finally, the sample is annotated according to majority principle. 250 groups of negative and positive cases in the dataset are selected, respectively, and then the CT images are preprocessed. Meanwhile, in order to increase the training sample size of the model, the processed images are rotated and inverted, and then randomly discrete input into the DCNN model for training and recognition.

5.1. Image Preprocessing. In order to eliminate noise interference introduced by equipment and better detect and identify pulmonary nodules, the median filtering algorithm was introduced to de-noise the original CT images. In order to compare and verify the effect of median filtering, a comparison experiment of Gaussian filtering is added, and the obtained results are shown in Figure 5. Figure 5(a) shows the CT image before noise reduction, and Figure 5(b) shows the CT image after noise reduction. A specific example after median filtering is shown in Figure 6.

```

Calculate and obtain the number of input training samples:
m = size(x, 3);
Obtain the sample number after random dispersion:
kk = randperm(m);
Get batch sample size:
numbatches = m/opt.batchsize;
for l = 1: numbatches
Take out the batchsize samples and corresponding labels after disordering
batch_x = x(:, :, kk((l - 1) * opt.batchsize + 1: l * opt.batchsize));
batch_y = y(:, :, kk((l - 1) * opt.batchsize + 1: l * opt.batchsize));
Computes the network output under the current network weights and network inputs:
net = cnnff(net, batch_x);
After the above network output is obtained, bp algorithm is used to get the error pair network weight through the corresponding
sample label:
net = cnnbp(net, batch_y);
After obtaining the derivative of the error to the weight, the weight is updated by the weight update method:
net = cnnapplygrads(net, opt);
if isempty(net.rL)
net.rL(1) = net.L;
end
net.rL(end + 1) = 0.99 * net.rL(end) + 0.01 * net.L;
end

```

ALGORITHM 1: The samples are randomly discrete and then input into the model for training.

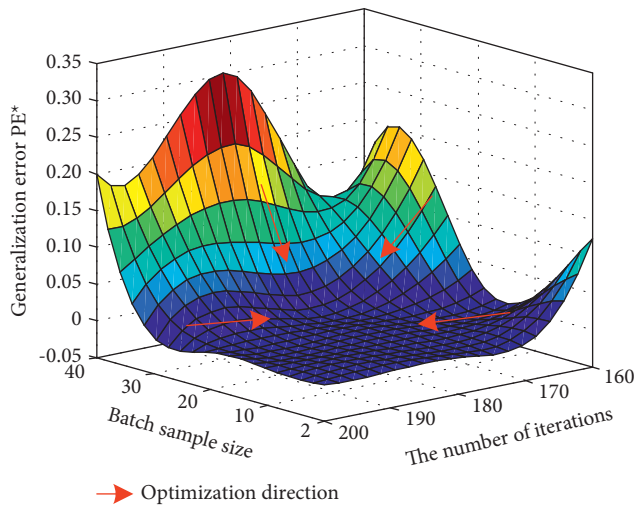


FIGURE 3: Process of sample batch optimization.

The dark spots with a large number of discrete distributions are impulse noises generated by the influence of sensors or channels. After median filtering, the results are shown in Figure 6(b), which can achieve an ideal noise reduction effect. As shown in Figure 5, although Gaussian filtering also has a certain noise reduction effect, compared with median filtering, the performance of Gaussian filtering in dealing with the noise reduction of salt and pepper noise is not obvious. It can be found that the median filtering method can not only suppress the impulse noise but also retain the image edge details, which can provide better conditions for the subsequent identification of pulmonary nodules.

From the above examples, it can be seen that the median filtering algorithm can achieve a better noise reduction effect for common impulse noise reduction problems. In addition to noise reduction, the median filtering algorithm can realize the enhancement of image details, which lays a solid foundation for subsequent DCNN model training and testing.

5.2. DCNN Framework Construction. Figure 7 shows the DCNN basic framework adopted by the method in this paper. Based on the training error and calculation cost considerations, the training error and calculation duration of the DCNN model at different depths are obtained, and the results are shown in Table 1. It can be seen that the optimal structure of the DCNN model consists of three convolution layers and three down-sampling layers. The size of the input image is 512×512 , and the down-sampling form is mean down-sampling. The size of the convolution kernel used by the first two convolution layers is 5×5 , and the size of the convolution kernel used by the third convolution layer is 6×6 . The number of convolution kernels used by each convolution layer is 6, 12, and 12 respectively. After passing through the full connection layer, pattern recognition is carried out and detection results are output.

The samples are input into the training model according to different batch numbers and different iterations are set to obtain the generalization error topographic map of the DCNN training model, as shown in Figure 8. Search to obtain the optimal batch sample number and the optimal number of iterations, and then obtain the optimal number of sample input batches, and then obtain the DCNN model completed by training.

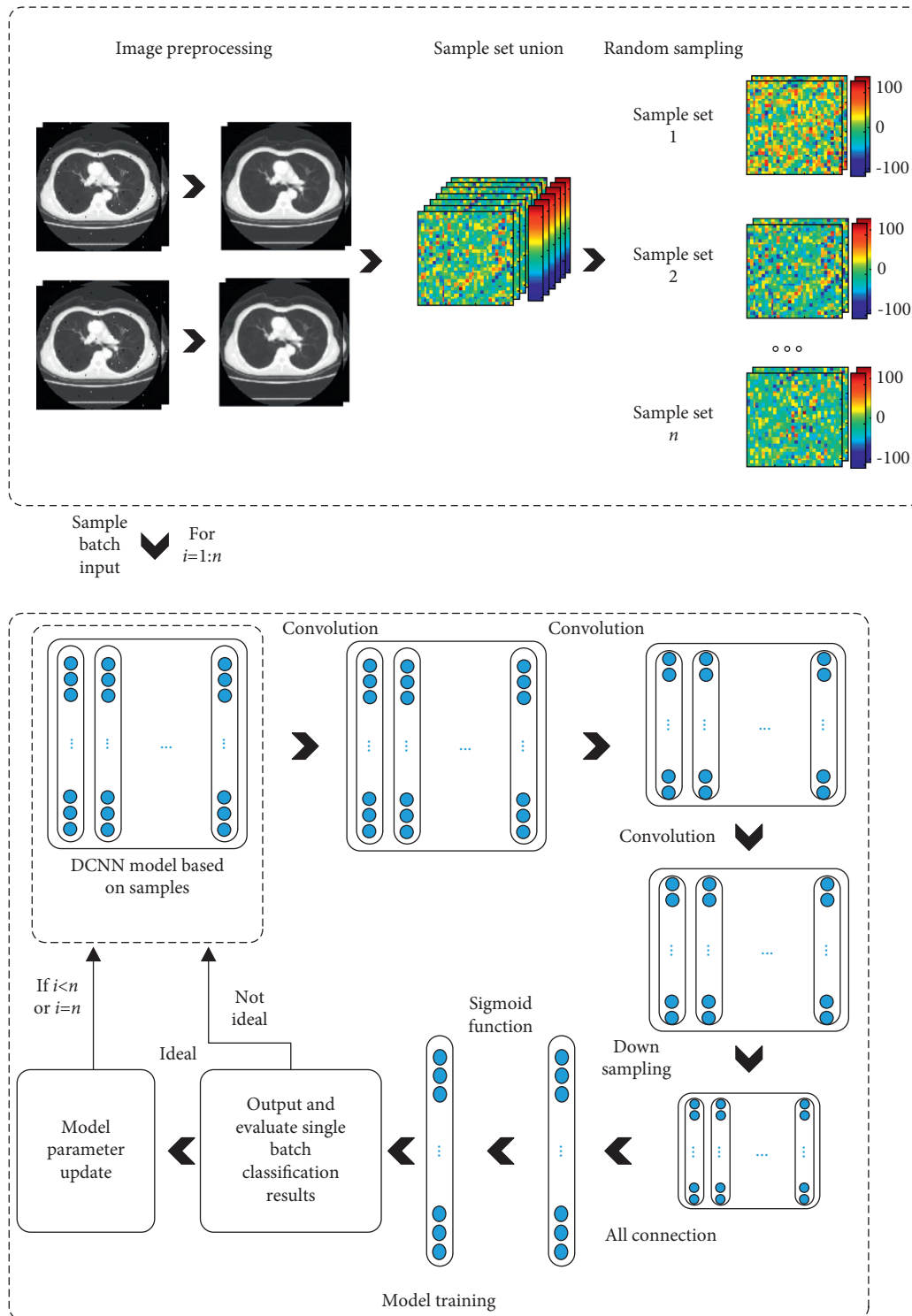


FIGURE 4: Steps of the lung nodule detection and recognition method based on the DCNN algorithm guided by interpretable training.

In order to express the influence of sample discretization on image features of in-depth model mining more intuitively, the features mined from the test samples are visualized. The obtained results are shown in Figures 9(a) and 9(b). Sample Nos. 1–10 are healthy, and sample nos. 11–20 are pulmonary nodules. It can be seen that in the case of

nondiscrete samples, the features mined by the detection model are chaotic, while the features mined by the model trained by the method proposed in this paper can distinguish the two types of samples well, indicating that sample discretization can effectively improve the feature mining capability of the model.

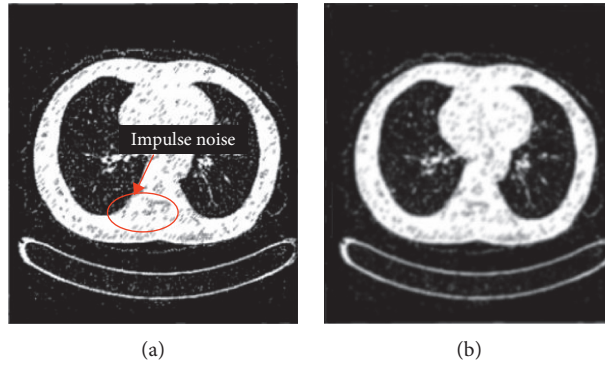


FIGURE 5: Gaussian filtering example. (a) Original LUNG CT images. (b) Gaussian filtering CT images.

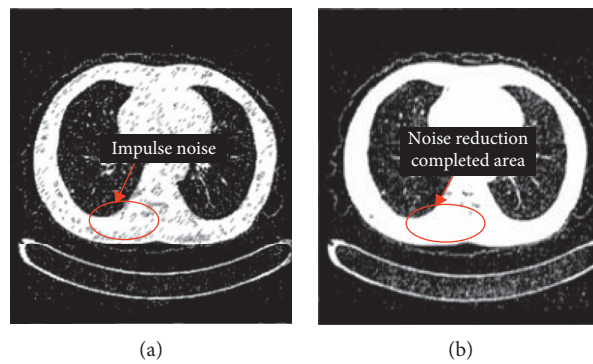


FIGURE 6: Median filtering example. (a) Original LUNG CT images. (b) Median filtering CT images.

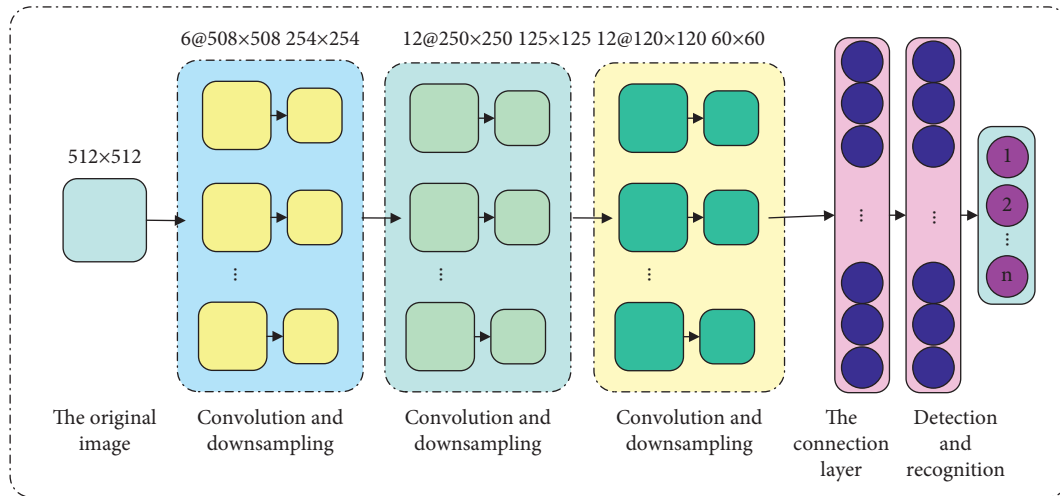


FIGURE 7: Schematic diagram of the DCNN structure and related parameters.

TABLE 1: Structure optimization results of DCNN.

No.	The training error	Calculation cost (s)
1	0.01524	12.53
2	0.01203	17.45
3	0.01025	18.26
4	0.01145	25.79

5.3. Comparative Analysis of Evaluation Indicators and Detection and Identification Results. In order to better measure the effectiveness of this method, based on the same dataset, a comparative analysis has been conducted with the detection and recognition methods of pulmonary nodules proposed in other literature. In order to quantify the comparative

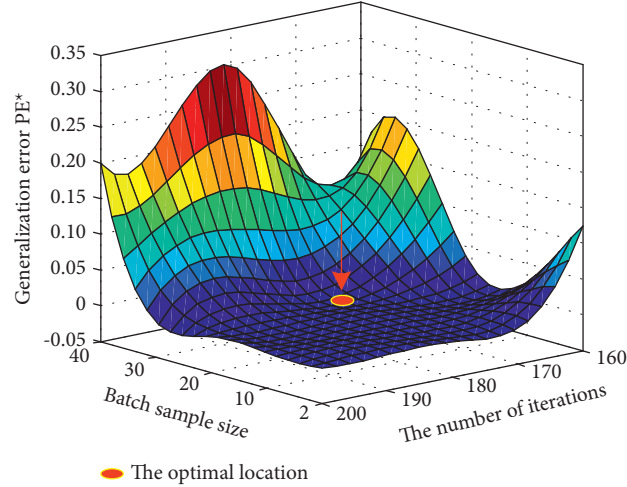


FIGURE 8: Optimal training batch optimization process.

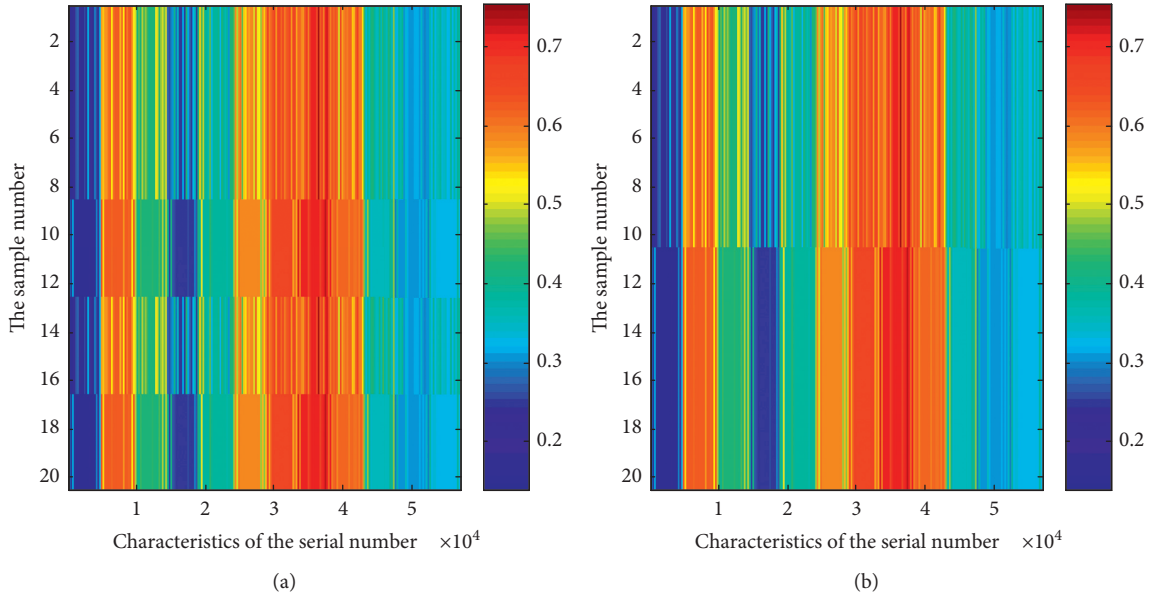


FIGURE 9: (a) Features mined by the trained model in the case of nondiscrete samples; (b) features mined by the trained model in the case of discrete samples.

analysis's results, the following 4 commonly used evaluation indicators have been introduced:

$$\begin{aligned}
 SEN &= \frac{TP}{TP + FN} \times 100\%, \\
 SPE &= \frac{TN}{TN + FP} \times 100\%, \\
 ACC &= \frac{TP + TN}{TP + TN + FP + FN} \times 100\%, \\
 FDP &= 1 - ACC.
 \end{aligned} \tag{23}$$

SEN indicates sensitivity, SPE indicates specificity, ACC indicates accuracy, and FDP indicates false diagnosis proportion. According to the actual evaluation criteria of

TABLE 2: Pulmonary nodule detection evaluation standard.

Test results	The gold standard	
Positive	TP	FP
Negative	FN	TN

pulmonary nodules in the medical field, the test criteria can be obtained as shown in Table 2. TP represents the proportion of true-positive results, TN represents the proportion of true-negative results, FP represents the proportion of false-positive results, and FN represents the proportion of false-negative results.

Based on the proposed method, the detection and recognition results of pulmonary nodules obtained from the LDC-ID-RI dataset are shown in Table 3 and Figure 10. In

TABLE 3: Comparison of pulmonary nodules' recognition based on the LIDC-ID-R1 datasets.

Method	Test results			
	SEN (%)	SPE (%)	ACC (%)	FPF (%)
1 Method of literature [3]	85.00	40.10	68.97	31.03
2 Method of literature [4]	92.37	73.24	85.20	14.80
3 Method of literature [6]	91.93	93.63	92.51	7.49
4 Method of literature [5]	96.24	92.11	94.40	5.60
5 The proposed method Scheme 1	56.80	47.20	52.00	48.00
6 The proposed method Scheme 2	83.20	80.80	82.00	18.00
7 The proposed method Scheme 3	62.40	61.60	62.00	38.00
8 The proposed method Scheme 4	96.40	95.60	96.00	4.00

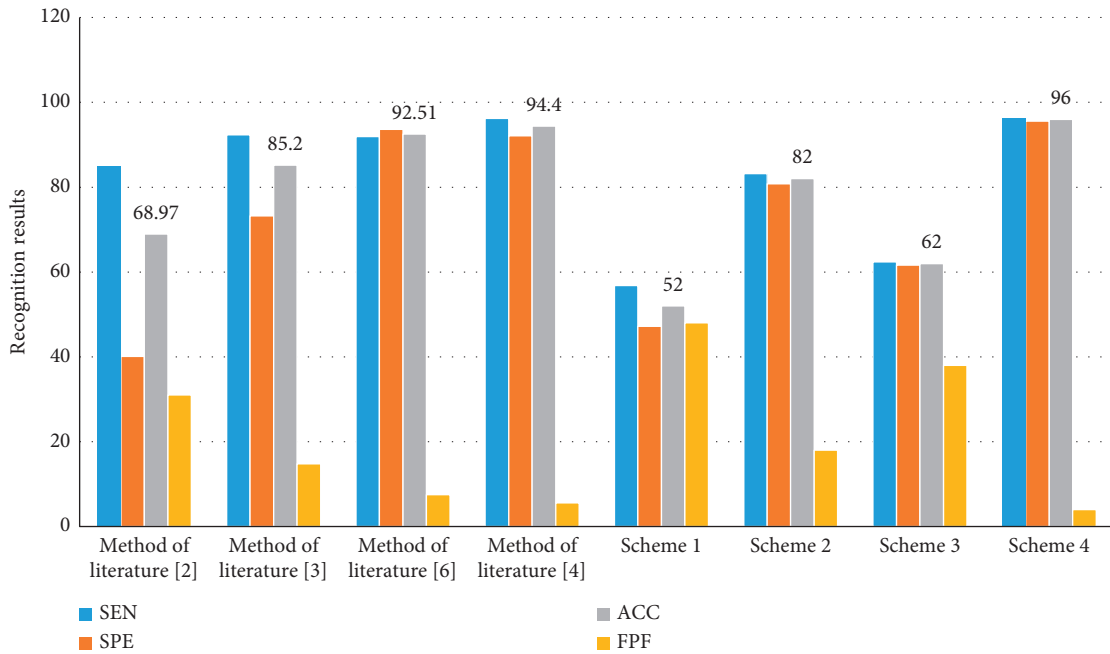


FIGURE 10: Recognition results of different methods.

order to compare and verify the superiority of sample preprocessing by median filtering and the influence of sample random dispersion on the detection model, different experimental schemes are set up for comparison and verification. In scheme 1, the original samples are directly input into the DCNN model for detection and recognition without median filtering and random discretization. Scheme 2 indicates that the original samples are not preprocessed but randomly dispersed and then input into DCNN for detection and recognition. Scheme 3 represents that the original sample is processed by median filtering but not randomly input into DCNN for detection and identification. Scheme 4 means that the original sample is firstly processed by median filtering, and then randomly discrete and then input to DCNN for detection and recognition.

Figure 10 mainly compares the advantages and disadvantages of each method and sample training strategy horizontally. In order to compare the advantages and disadvantages of each method and sample training strategy vertically again, the spider plot is shown in Figure 11.

As can be seen from the above experimental results, compared with the results obtained by the methods used in the literature [3, 4, 6], the experimental results obtained by scheme 1, scheme 2, and scheme 3 are not very ideal, but the internal comparison of schemes shows that the median filtering algorithm and the random dispersion of samples can effectively improve the detection and recognition performance of DCNN. Moreover, sample random dispersion has a greater impact on the detection and recognition performance of the DCNN model. Thanks to the intelligent detection and recognition ability of DCNN guided by image pretreatment and interpretable training, scheme 4 can achieve better detection and recognition effect of pulmonary nodules. Although the advantages of scheme 4 are not obvious compared with the algorithm in the literature [5], after image pretreatment, the samples are input into the DCNN model under the guidance of explanatory training for detection and recognition, which has achieved effective improvement in all indicators, especially in the specificity index *SPE* and accuracy index *ACC*. Thus, the validity of the proposed method is verified again.

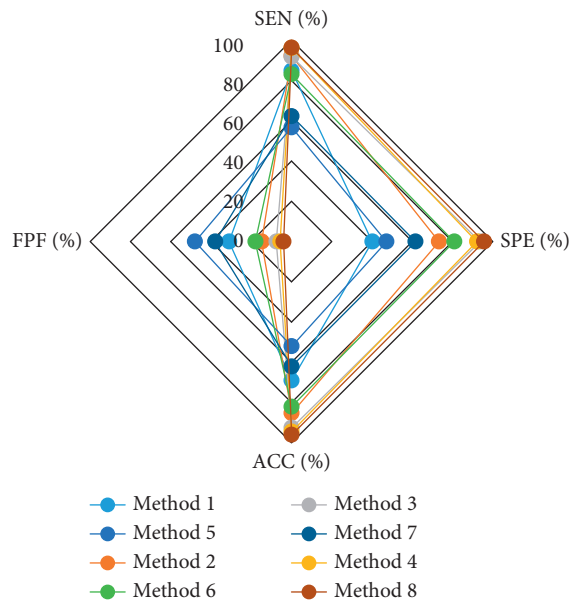


FIGURE 11: Spider plot to compare the CT image classification performance of different methods.

6. Conclusion

Aiming at the problem that CT images are interfered by noise, the median filtering algorithm is used in this paper to preprocess the detected images, which can significantly reduce the salt and pepper noise in CT images and highlight the target information, thus providing better quality training and detection and recognition samples for subsequent detection and recognition models. For the problems that DCNN training sample random discrete theoretical basis is not clear and training batch cannot be scientific and quantitative, from the perspective of DCNN training strategy, the theoretical basis of the sample derivation has been proved, and a training batch quantitative method has been proposed. Experimental verification results show that the proposed method has good performance in sensitivity, specificity, and accuracy, which has obtained 96.40%, 95.60%, and 96.00% test results, respectively. Compared with other algorithms in the literature, the proposed method has improved in all indicators. Moreover, the proposed method has strong transferability, and has a good guiding significance for the training of most deep learning network models, and can be used in more image recognition tasks. However, the method in this paper still has the following aspects that can be further studied: First, the pattern recognition problem with incomplete sample coverage; second, the adaptive construction and optimization of the network structure of the deep learning model; third, the uncertainty evaluation of the test results. In the future research work, in-depth research will focus on the above problems.

Data Availability

This paper carried out method validation based on the online public Lung Image dataset, LDC-ID-RI (Lung Image Database Consortium) [24].

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This study was supported by Research and Development Project of Hunan Institute of Science and Technology Information, China (no. 2018305).

References

- [1] Z. Gao and X. Wan, "Lung nodule recognition based on deep learning," *Journal of South-central University for Nationalities*, vol. 38, no. 3, pp. 393–396, 2019.
- [2] Y. Zhu and R. Liu, "Qiongtao huang fine-grained image recognition of weak supervisory information based on deep neural network," *Journal of Electronic Measurement and Instrument*, vol. 34, no. 2, pp. 115–122, 2020.
- [3] Y. Shi, X. Zeng, and H. Zhang, "Research on automatic detection algorithm of pulmonary nodules," *Chinese Journal of Scientific Instrument*, vol. 27, no. 6, pp. 2265–2267, 2006.
- [4] J. Sun, An Gong, and H. Shi, "Detection of pulmonary nodules in low-dose CT images based on convolutional neural network," *Computer Technology and Development*, vol. 29, no. 11, pp. 173–177, 2019.
- [5] Li Li, Lu Qiao, and H. Zhang, "Automatic lung nodule detection algorithm based on deep learning," *Computer Applications and Software*, vol. 37, no. 7, pp. 95–100, 2020.
- [6] X. Xi and W. Liu, "Pulmonary nodule assistant diagnosis system based on target detection algorithm," *Computer and Modernization*, vol. 11, pp. 1–7, 2020.
- [7] M. A. Khan, V. Rajinikanth, S. C. Satapathy et al., "VGG19 network assisted joint segmentation and classification of lung nodules in CT images," *Diagnostics*, vol. 11, no. 12, pp. 1–16, 2021.
- [8] A. T. Sahlol, M. Abd Elaziz, A. Tariq Jamal, R. Damaševičius, and O. Farouk Hassan, "A novel method for detection of tuberculosis in chest radiographs using artificial ecosystem-based optimisation of deep neural network features," *Symmetry*, vol. 12, no. 7, p. 1146, 2020.
- [9] D. Połap, M. Woźniak, R. Damasevicius, and W. Wei, "Chest radiographs segmentation by the use of nature-inspired algorithm for lung disease detection," in *Proceedings of the 2018 IEEE Symposium Series on Computational Intelligence (SSCI)*, pp. 2298–2303, IEEE, Bangalore, India, November 2018.
- [10] N. Tang, Z. Wei, and R. Zhang, "Comparative study on classification of pulmonary nodules based on multi-scale and multi-mode images," *Computer Engineering and Applications*, vol. 56, no. 3, pp. 165–175, 2020.
- [11] X. Zhang and Q. Yan, "End-to-end pulmonary nodules detection model based on dual-path network," *Computer Engineering and Design*, vol. 41, no. 12, pp. 3503–3509, 2020.
- [12] D. Zhang, F. Qin, and L. Cao, "Denoising method of pepper and salt image based on median filter," *Journal of Yibin University*, vol. 14, no. 6, pp. 108–110, 2014.
- [13] H. Guo, Y. Yue, and F. Lin, "Image denoising algorithm based on high density noise and mixed noise," *Journal of Taiyuan Normal University (Natural Science Edition)*, vol. 19, no. 2, pp. 41–47, 2020.

- [14] X. Chen and Mu Kang, "An image median filtering algorithm that can retain detailed information," *Information Technology and Information Technology*, vol. 11, pp. 128-129, 2020.
- [15] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504-507, 2006.
- [16] L. Wan, X. Lan, and H. Zhang, "A review of deep reinforcement learning theory and application," *Pattern Recognition and Artificial Intelligence*, vol. 32, no. 1, pp. 67-81, 2019.
- [17] L. Shi, H. Ma, and C. Zhang, "Pulmonary nodules detection method based on improved residual structure," *Computer Applications*, vol. 40, no. 7, pp. 2110-2116, 2020.
- [18] Y. Xiao, Q. Jiang, X. M. Wang, Y. J. Du, and Z. X. Wang, "Image recognition of convolutional neural network model driven by radius interval boundary," *Journal of Xihua University (Natural Science Edition)*, vol. 40, no. 2, pp. 71-81, 2021.
- [19] P. Han, Y. Qiang, and J. H. Liu, "A two-stage lung nodule detection method based on deep convolutional neural network," *Computer Engineering and Design*, vol. 42, no. 3, pp. 755-761, 2021.
- [20] W. Zhang, C. Li, G. Peng, Y. Chen, and Z. Zhang, "A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load," *Mechanical Systems and Signal Processing*, vol. 100, pp. 439-453, 2018.
- [21] S. S. Dong and Z. X. Huang, "A brief theoretical overview of random forests," *Journal of Integration Technology*, vol. 2, no. 012013, pp. 1-7, 2013.
- [22] Y. Sheng, G. Shi, and Z. Qin, "A stronger law of large numbers for uncertain random variables," *Soft Computing*, vol. 22, no. 17, pp. 5655-5662, 2018.
- [23] Y. H. Huo, "Chebyshev inequality and its applications," *Journal of Changchun University of Technology (Natural Science Edition)*, vol. 33, pp. 712-714, 2012.
- [24] A. Medhat, A. Shalaby, and M. S. Sayed, "A highly Parallel SAD architecture for motion estimation in HEVC encoder," in *Proceedings of the IEEE Asia Pacific Conference on Circuits & Systems*, pp. 280-283, IEEE, Ishigaki, Japan, November 2014.