Accuracy Study on Deep Learning-Based CT Image Analysis for Lung Nodule Detection and Classification

Xiyue Cheng[1], Jinyu Li[2], Mengqi Mi[1], Hao Wang[1], Jianjun Wang[2], Peng Su[3*]

[1] College of Clinical Medicine, North China University of Science and Technology, Tangshan 063000, China
[2] North China University of Science and Technology Affiliated Hospital, Tangshan 063000, China
[3] Ji Tang College, North China University of Science and Technology, Tangshan 063000, China

Corresponding Author Email: Supeng@ncst.edu.cn

ABSTRACT

Early detection and accurate classification of lung nodules are crucial for the treatment of lung cancer. With the widespread application of deep learning technologies in medical imaging analysis, significant progress has been made in the automatic detection and classification of lung nodules from computed tomography (CT) images. However, existing deep learning approaches often face challenges with limited annotated data and generalization across diverse datasets. To address these challenges, this study introduces two innovative methods: a domain-adaptive adversarial network for joint segmentation of lung CT images to enhance model generalization, and an improved deep propagation generation network (DPGN) for few-shot classification of lung CT images to reduce reliance on extensive annotated data. Through these methods, this research aims to improve the accuracy of lung nodule detection and classification, providing more reliable support for clinical diagnosis.

1. INTRODUCTION

In the past few decades, with the rapid development of computer vision and medical imaging technologies, deep learning-based analysis of CT images has shown tremendous potential in the detection and classification of lung nodules [1-4]. Lung nodules are one of the earliest signs of lung cancer, which is one of the most lethal cancers worldwide. Early diagnosis and accurate classification of nodules are crucial for devising treatment plans and improving patient prognosis. Therefore, enhancing the accuracy of lung nodule detection and classification is significant for clinical diagnosis and patient care [5, 6].

Despite some progress made by traditional medical imaging analysis methods in recent years, they usually rely on manually crafted features and expert knowledge, which limits their accuracy and generalization capability [7-9]. Deep learning technologies, especially convolutional neural networks (CNNs), have become powerful tools for improving lung nodule detection and classification performance due to their ability to automatically learn complex and abstract features from large datasets. However, the performance of deep learning models largely depends on the availability of large amounts of annotated data, which is often difficult to obtain in medical imaging analysis [10-12].

While current research methods have improved the accuracy of lung nodule detection and classification to some extent, there are still significant flaws and shortcomings. For example, most existing deep learning methods require large amounts of annotated data to train models, which are often difficult to acquire in practical applications [13, 14]. Additionally, many models exhibit poor generalization when dealing with datasets from different distributions, which is particularly evident in multicenter studies. These issues limit the application and effectiveness of deep learning methods in actual clinical environments [15-17].

This paper aims to overcome the above challenges by focusing on two main research contents, thus enhancing the accuracy of deep learning-based CT image analysis in lung nodule detection and classification. First, we propose a domain-adaptive adversarial network for joint segmentation of lung CT images, aiming to enhance the model's generalization capability when dealing with data from various sources. Second, addressing the issue of insufficient annotated data, we develop a few-shot lung CT image classification algorithm based on an improved DPGN. Through these innovations, we not only expect to improve the accuracy of lung nodule detection and classification but also aim to reduce the dependency on large volumes of annotated data, thus providing more practical and effective solutions in clinical applications.

2. DOMAIN-ADAPTIVE ADVERSARIAL NETWORK FOR JOINT SEGMENTATION OF LUNG CT IMAGES

Lung CT images come from diverse sources, and the differences in image feature distributions caused by different equipment and scanning parameters can severely affect the model's generalization ability and accuracy. To address this,

we propose a U-Net combined with a domain adversarial network (DAUNet) model, which integrates U-Net's efficient image segmentation capabilities with the strong domain adaptability of the domain adversarial network to tackle the critical issue in lung CT image analysis—the domain shift problem in lung nodule detection and classification. DAUNet employs an adversarial learning mechanism introduced by the domain adversarial network, encouraging the model to learn more generalizable feature representations, thereby effectively reducing the domain shift impact between different datasets.

2.1 Overall architecture

The model defines the training set of lung CT images as the source domain image set $\{a_u\}^V_{u=1} \in A$, and the test set or images to be analyzed as the target domain image set $\{y_u\}^V_{u=1} \in B$. The model learns feature representations from the source domain A through a generator, and uses these learned features to perform high-accuracy lung nodule segmentation in the target domain B. The model comprises two main components: the generator and the discriminator. The generator receives labeled source domain lung CT images and unlabeled target domain images as inputs a_u, extracts features of the source domain images through deep learning, and uses these features to optimize the lung nodule segmentation of the target domain images b_u. In this process, the generator produces feature maps $H(a_u)$ and $H(y_u)$ for the source and target domain images, respectively, and feeds these feature maps into the discriminator. The task of the discriminator is to distinguish the differences between the source and target domain feature maps, to determine which domain they each belong to. As training progresses, the discriminator continually improves its accuracy, while the generator strives to enhance the segmentation accuracy of source domain images and reduce the feature differences between the source and target domains.

In the network framework proposed in this paper, the segmentation module plays a crucial role, particularly in the tasks of joint segmentation of lung CT images and lung nodule detection and classification. The segmentation module first performs precise segmentation on labeled source domain lung CT images, then, by comparing the segmentation results with the actual labels, the network is adjusted and optimized. When the generator and domain discriminator work together entering the adversarial learning phase, a dynamic adversarial relationship forms between them. Through continuous mutual optimization, they enhance the network's segmentation and discrimination capabilities, until reaching a Nash equilibrium. Ideally, the discriminator would be unable to distinguish differences between the source and target domain images, with its output approaching 0.5, indicating that the generator has effectively extracted domain-invariant features between the source and target domain images.

Specifically, in this framework, the generator's goal is twofold: on the one hand, when the source domain image a_u is input, the generator's output $H(a_u)$ should be as close as possible to the actual lung nodule labels, achieving high-accuracy lung nodule segmentation; on the other hand, when the unlabeled target domain image b_u is input, the generator also needs to ensure that the output $H(b_u)$'s error is minimized, to adapt and optimize the segmentation performance of the target domain images. For the domain discriminator, its goal is also twofold: firstly, when the source domain image a_u is input, the discriminator $F(a_u)$'s output should be close to 0, indicating it has correctly identified the source domain image;

secondly, when the target domain image b_u is input, the discriminator $F(b_u)$'s output should be close to 1, accurately distinguishing the target domain image. Through such a training mechanism, the network aims to achieve two main objectives: to improve the accuracy of lung nodule segmentation and to effectively alleviate the domain shift problem between the source and target domains, thereby achieving accurate lung nodule detection and classification across different lung CT image datasets. The ultimate goal expression of the network is:

$$\underset{H}{MIN}\,\underset{F}{MIN}\left\{\begin{array}{l}R_{a\sim A,b\sim B}\left[\log F\left(a,b\right)\right]+\\R_{a\sim A}\left[\log\left(1-F\left(a,H\left(a\right)\right)\right)\right]\end{array}\right\} \quad (1)$$

2.2 Generator

In the proposed network structure, the generator is designed to accept two types of inputs: one is labeled source domain lung CT images, used for learning segmentation accuracy; the other is unlabeled target domain lung CT images, used to adapt and improve generalization capabilities across different datasets. This dual input mode allows the network not only to learn precise segmentation of lung nodules but also to adapt to image features under different scanning conditions, thereby enhancing model performance on new, unseen datasets. Assuming the image height and width are represented by G and Q, the network's input size is designed as $2 \times G \times Q \times 1$, considering that CT images are generally grayscale, where "1" represents a single grayscale channel.

The structure of the generator consists of two main parts: an encoder and a decoder. The encoder compresses the input images into deep feature representations through a sequence of convolution and pooling operations, carrying the high-level semantic information of the input images. Each convolution layer is followed by a ReLU activation function and DropBlock regularization, which randomly drops areas on the feature map during training to mitigate overfitting, with a dropout rate also set at 0.1. After processing by the encoder, the deep feature maps dimension becomes $2 \times G/8 \times Q/8 \times 128$, preparing for subsequent feature fusion and upsampling. The decoder is responsible for gradually restoring the compressed feature maps back to the original image size through deconvolution and convolution operations, restoring image details and structure, and ultimately outputs the probability of each pixel belonging to a lung nodule through a Sigmoid activation function, completing the segmentation task.

Moreover, a Gradient Reversal Layer (GRL) is specifically incorporated into the generator's design, located at the penultimate decoder layer, to fuse the feature maps of the source and target domains and send these fused feature maps into the domain discriminator for domain classification. Specifically, the GRL introduces an adjustable parameter η, which controls the extent of gradient reversal: during the model's backward propagation, the gradients passing through the GRL are multiplied by a negative value ($-\eta$), effecting the reversal of gradient direction. This operation forces the domain discriminator to learn to distinguish between source and target domain features while promoting the generator to produce more domain-invariant features, thereby reducing the differences between the source and target domains. Assuming the loss function for the source domain is represented by M_t and the loss function for the target domain by M_s, the specific reversal operation is expressed as follows:

$$M_{TO} = M_t - \eta M_s \tag{2}$$

Assuming the current epoch number is represented by r, and the total number of epochs by v_{EP}. The first stage is set as r_0, and the formula for η's change is given by the following equation:

$$\eta(r) = MAX\left(0, \eta_{MAX}\frac{r-r_0}{v_{EP}-r_0}\right) \tag{3}$$

2.3 Discriminator

In the proposed model, the design and function of the domain discriminator play a critical role, achieving effective recognition and adaptation of differences between the source and target domains through a refined network structure. Figure 1 displays the architecture of the domain discriminator. The domain discriminator first receives the processed source domain feature maps $H(a_u)$ and target domain feature maps $H(b_u)$ from the GRL, a step that reduces domain differences through gradient reversal, thereby prompting the model to produce more generalizable feature representations. These feature maps are then fed into a sequence of four convolutional encoding layers for deep feature extraction, each convolutional layer consisting of 3×3 convolutions, ReLU activation functions, and DropBlock to enhance the network's learning capacity and resistance to overfitting. Through a sequence of 2×2 maximum pooling operations, the feature maps are gradually reduced to 1/16 of the original size, while increasing the model's translational invariance to input features. Finally, a fully connected layer flattens the extracted feature maps and calculates a final discrimination result used to differentiate between the source and target domains. The accuracy of this discrimination result directly relates to the adversarial training effect between the generator and discriminator. We adjust the model parameters using stochastic gradient descent by calculating the error between the discrimination results and actual labels, aiming to achieve mutual opposition between the two and ultimately reach Nash equilibrium.
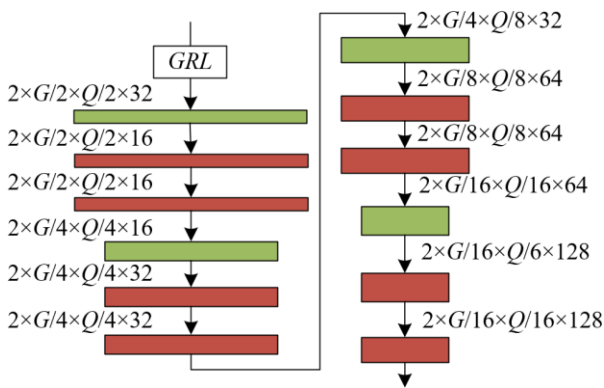

Figure 1. Domain discriminator architecture

2.4 Loss functions

In the proposed model, the design of the loss functions includes two key parts: segmentation loss and adversarial loss. The segmentation loss is primarily used to evaluate and optimize the model's performance in segmenting lung nodules on source domain images, ensuring the segmentation model

can accurately predict the location and scope of lung nodules. This is crucial for enhancing the model's accuracy and reliability in actual clinical applications. At the same time, adversarial loss is responsible for adjusting the adversarial relationship between the domain classifier and the feature extractor. By optimizing this loss, the model can better identify and adapt to the feature distribution of target domain images, thereby significantly improving the model's adaptability to differences between various data sources. By simultaneously optimizing these two loss functions, our model not only achieves precise lung nodule segmentation on source domain data but also maintains high segmentation performance and good domain adaptability when facing target domain data. Assuming the segmentation result for the u-th sample is represented by $B^{\hat{}u}$, its corresponding label M_{SE} is the Dice loss function, represented by B^u, elements in image U belonging to structure j are represented by U^j, the probability of the u-th sample belonging to the target domain is represented by $f^{\hat{}u}$, and the binary cross-entropy loss function is represented by M_{DO}, then the loss function expressions are:

$$M_{TO} = \sum_{u=1}^{v} M_{SE}\left(\hat{B}^u, B^u\right) + \sum_{u=1}^{v} M_{DO}\left(\hat{f}^u, f^u\right) \tag{4}$$

$$M_{SE}\left(U^j, K^j\right) = 1 - \frac{2\left|U^j \cap K^j\right|}{\left|U^j\right| + \left|K^j\right|} \tag{5}$$

$$M_{DO}\left(\hat{f}^u, f^u\right) = -\left(f\log\left(\hat{f}\right) + (1-f)\log\left(1-\hat{f}\right)\right) \tag{6}$$

3. FEW-SAMPLE LUNG CT IMAGE CLASSIFICATION BASED ON IMPROVED DPGN

This paper further employs an improved DPGN model to address the few-sample lung CT image classification problem, which is designed for specific applications in lung nodule detection and classification. This model adopts a 2way-1shot learning setup, where the training process utilizes two categories of support sets and one category of query sets, simulating the scenario of sample scarcity in actual applications. To better handle the complex features in lung CT images and improve classification accuracy, significant improvements have been made to the feature extraction module, introducing attention mechanisms and optimizing normalization layers and activation functions, resulting in an improved backbone network named SF_ResNet12. Through SF_ResNet12, the model effectively extracts key features from lung CT images. In the classifier module, a Point-to-Distribution (P2D) propagation strategy is adopted, and dynamic convolution is introduced in the embedding space, using Mahalanobis distance (MD) and earth mover's distance (EMD) as metrics.
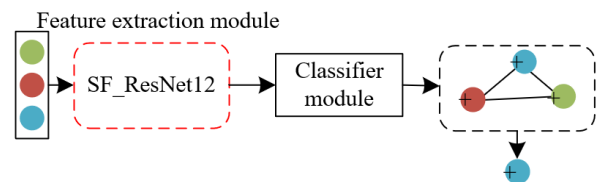

Figure 2. Few-sample lung CT image classification model architecture

These improvements enable the model to perform effective iterative updates between the nodes of the point and distribution graphs based on distance, thereby accurately classifying lung nodule types in a small number of samples. Figure 2 displays the architecture of the few-sample lung CT image classification model.

3.1 Feature extraction module optimization

This paper specifically employs the DPGN as the foundational framework and integrates the Simultaneous Attention Module (SimAM) into the residual blocks of its feature extraction module, aiming to significantly enhance the network's ability to capture features of lung CT images without introducing additional parameters. The optimized residual block architecture is shown in Figure 3.
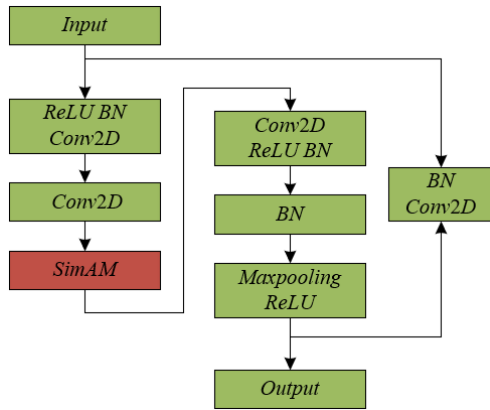
Figure 3. Residual block architecture incorporating SimAM attention mechanism

This improvement involves introducing SimAM after the second convolution operation on the main path of each residual block, thereby generating three-dimensional weights for image features across spatial and channel dimensions, which are effectively aggregated with the feature maps. Consequently, each pixel receives weights of varying sizes based on its importance, calculated using a simple and effective energy function derived using binary labels. The introduction of this attention mechanism, particularly for the specific application of lung CT images, enables the model to focus more on the key areas of lung nodules, thereby improving the accuracy of lung nodule classification under few-sample conditions. The definition formula for the energy function R_s is given as follows:

$$R_s = \left[1 - \left(Q_u u + y_u\right)\right]^2 + \frac{1}{l-1}\sum_{x=1}^{l-1}\left[-1 - \left(Q_u a_x + y_u\right)\right]^2 + \eta Q_u^2 \tag{7}$$

Theoretically, each channel has l energy functions. Let $\omega_u = 1/l - 1 \Sigma^{l-1}_{x=1} a_x$, $\delta^2_u = 1/l - 1 \Sigma^{l-1}_{x=1} (a_x - \omega_u)^2$, where:

$$Q_u = -\frac{2\left(u - \omega_u\right)}{\left(u - \omega_u\right)^2 + 2\delta_u^2 + 2\eta} \tag{8}$$

$$y_u = \frac{u^2 - \omega_u^2}{\left(u - \omega_u\right)^2 + 2\delta_u^2 + 2\eta} \tag{9}$$

Combining the above three equations yields the following expression for minimizing the energy function:

$$R_u^* = \frac{4\hat{\delta}^2 + 4\eta}{\left(u - \hat{\omega}\right)^2 + 2\hat{\delta}^2 + 2\eta} \tag{10}$$

Traditional batch normalization combined with ReLU activation function, although performing well in most cases, can affect model performance under few-sample learning environments, especially when handling highly heterogeneous lung CT images, due to inaccurate statistical estimates caused by the small number of samples. To adapt to the specific needs of lung nodule detection and classification, this paper further optimizes the normalization method and activation function within the residual blocks of the feature extraction module. Specifically, BN and ReLU in the ResNet12 network are replaced with Filter Response Normalization (FRN) and Thresholded Linear Unit (TLU), respectively. The optimized residual block architecture is shown in Figure 4. FRN eliminates the dependency of normalization operations on batch size, and TLU effectively addresses any deviations from zero that FRN may introduce, which could inhibit the learning process when using ReLU, thus impacting the overall network performance. The combination of FRN and TLU not only enhances the robustness and adaptability of the feature extraction module but also optimizes the network's ability to capture and represent lung nodule features, making it more suitable for accurate classification of lung nodules under few-sample conditions.
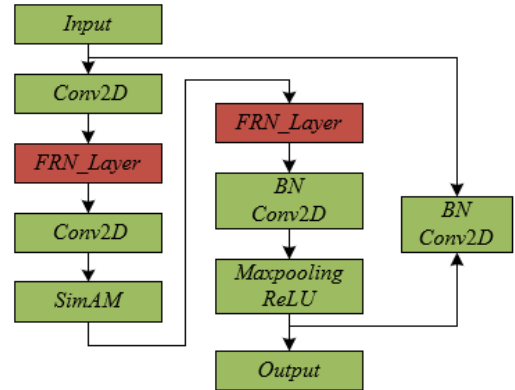
Figure 4. Residual block architecture with FRN layer

Assuming the feature vector A_x processed by FRN and TLU results in output vector C_x, where the output vector is represented by B_x, and the input feature vector is represented by A_x, with dimensions $[Y,Q,G,Z]$, mini-batch size represented by Y, spatial range mapped by Q, G, the number of filters used in convolution represented by Z, and the vector for the Y-th batch's Z-th filter represented by $v = Q \times G$. The learnable parameters are represented by e, k. The definition of FRN is given as follows:

$$B_x = e\frac{A_x}{\sqrt{\sum_x \frac{A_x^2}{v} + r}} + k \tag{11}$$

Compared to other few-sample image classification algorithms, this strategy tailored for lung nodule classification aligns more closely with the characteristics of medical image

analysis and effectively enhances the model's ability to recognize complex and subtle features of lung lesions. Assuming the output vector of FRN is represented by B_x, the output vector of TLU is represented by C_x, and the learnable parameters are denoted by π, the definition of TLU is given by the following formula:

$$
\begin{aligned}
C_x &= MAX\left(B_x, \pi\right) = MAX\left(B_x, -\pi, 0\right) \\
&= RE\left(B_x, -\pi\right) + \pi
\end{aligned}
\tag{12}
$$

3.2 Classifier module optimization

This paper focuses on the detection and classification of lung nodules, particularly in the context of few-sample lung CT image classification. To this end, key improvements have been made to the classifier module within the improved DPGN model. Figure 5 illustrates the classifier module architecture. Specifically, Omnidirectional Dynamic Convolution (ODConv) is introduced to replace traditional convolution operations, enhancing the backbone neural network's feature extraction capability when processing lung CT images. ODConv, by implementing a parallel strategy, introduces multidimensional attention mechanisms to dynamically learn across the four spatial dimensions of the convolution kernels, thereby significantly improving the accuracy of lung nodule classification.
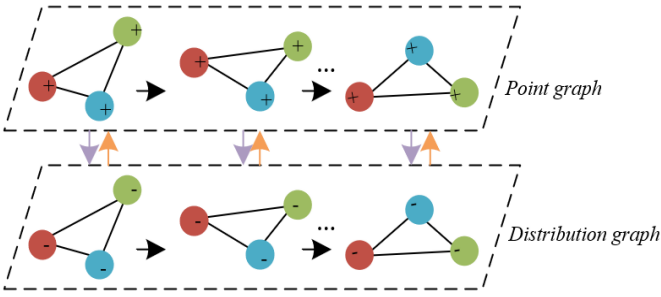

Figure 5. Classifier module architecture

Furthermore, within the P2D embedding space, ODConv is also applied, consisting of ODConv convolution, BN normalization, and leakyReLU activation functions, aimed at enhancing the representational capacity of the embedding space. The specific architecture is depicted in Figure 6. During the process of point graph distance measurement and distribution graph distance measurement, the introduction of ODConv further enhances the model's learning and representational capabilities. ODConv's core advantage lies in its use of multiple convolutional kernels' linear combination and the application of attention mechanisms for dynamic weighting, making the convolution operation dependent on the input, thus more effectively adapting to the complex features of lung nodules in lung CT images.

Assume the convolutional kernels are represented by $\{Q_1, Q_2, ..., Q_{v-1}, Q_v\}$, the attention function is represented by $\pi_{wi}(X)$, and the attention scalars calculated based on $\pi_{wi}(X)$ are represented by $\{\beta_{q1}, \beta_{q2}, ..., \beta_{q(v-1)}, \beta_{qy}\}$. ODConv is composed of $\{Q_1, Q_2, ..., Q_{v-1}, Q_v\}$ and $\{\beta_{q1}, \beta_{q2}, ..., \beta_{q(v-1)}, \beta_{qy}\}$. Further assume that the output value is represented by B, the input value by A; the u-th convolutional kernel is represented by Q_u; the attention scalar of the convolutional kernel Q_u is represented by β_{qu}; the attention weights along the spatial dimension by β_{Tu}, along the input channel dimension by β_{Zu}, along the output

channel dimension by β_{du}, and the attention weights for the convolutional kernel dimension by β_{qu}; the step-by-step multiplication operator is represented by $\times$, and the convolution operator by $*$. The computation is then given as follows:

$$
B = \begin{pmatrix} Q_1 \times \beta_{T1} \times \beta_{Z1} \times \beta_{d1} \times \beta_{v1} \\ + \cdots + Q_v \times \beta_{Tv} \times \beta_{Zv} \times \beta_{dv} \times \beta_{qv} \end{pmatrix} * A
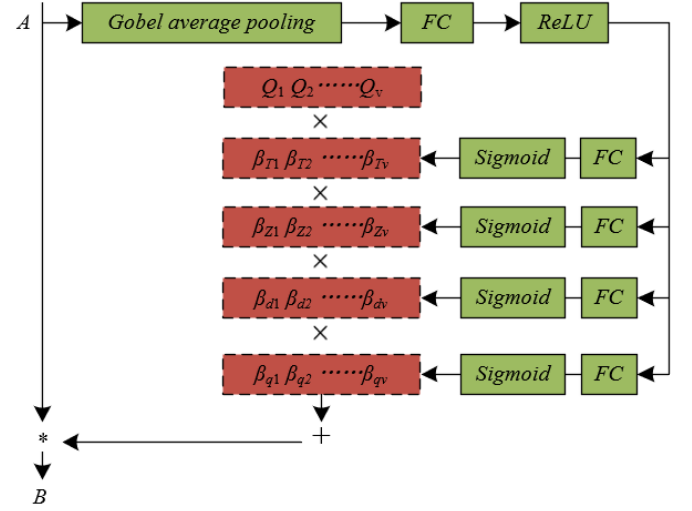\tag{13}
$$

Figure 6. ODConv architecture

The choice and optimization of distance measurement methods in the classifier module are also crucial for enhancing classification accuracy. Considering the different representations of lung nodule features in point graphs and distribution graphs, this paper uses the MD for distance calculation between point graphs because MD can account for the correlation between data features, making it more suitable for describing the similarity between individual lung nodule samples. For distance calculation between distribution graphs, the EMD is used, which can effectively measure the similarity between two sample set distributions, as EMD calculates the "cost" of transforming one distribution into another, suitable for describing overall differences between sample sets. These improvements in distance measurement methods, particularly optimized for the characteristics of lung nodule CT images, place greater emphasis on matching the uniqueness of lung nodule images. By precisely measuring the distances between samples and their distributions, more refined classification is achieved. This not only enhances the accuracy of the model in lung nodule classification tasks but also strengthens the model's adaptability to the diversity of lung nodule features.

4. EXPERIMENTAL RESULTS AND ANALYSIS

In this paper, the training process of the lung CT image joint segmentation model based on a domain-adaptive adversarial network adopted a strategy of gradually adjusting the adjustable parameter η of the GRL. From the experimental data shown in Figure 7, during the first 80 epochs, η remained at 0, indicating that the model primarily focused on learning from the source domain data without forcibly adapting to the target domain data distribution. This approach helped the model to stably learn and extract key features from lung CT images in the initial phase, providing a solid foundation for

subsequent domain adaptation. Starting from the 80th epoch, η gradually increased, reaching 0.05, 0.1, 0.15, and 0.2 at epochs 100, 120, and 140 respectively. This indicates that as training progressed, the model gradually increased its requirements for adaptability to the target domain, intensifying the adaptation from the source to the target domain through the introduction of the GRL, and promoting alignment of feature distributions between the two domains. It can be concluded that this strategy of gradually adjusting the GRL is very effective. In the early stages of training, the model could fully learn from the source domain data, avoiding potential training instability caused by prematurely introducing domain adaptation. As η gradually increased, the model was progressively guided to learn feature representations with good generalization ability for the target domain, a process that helped to gradually reduce the distribution differences between the source and target domains without sacrificing source domain performance, thereby improving the model's generalization ability in the target domain. Additionally, by dynamically adjusting the value of η, more precise control of the domain adaptation process could be achieved, avoiding premature model saturation or overfitting in specific domains.
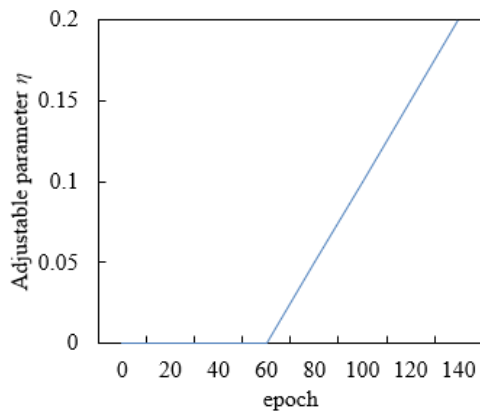
Figure 7. Training process variations of the lung CT image joint segmentation model based on domain-adaptive adversarial network
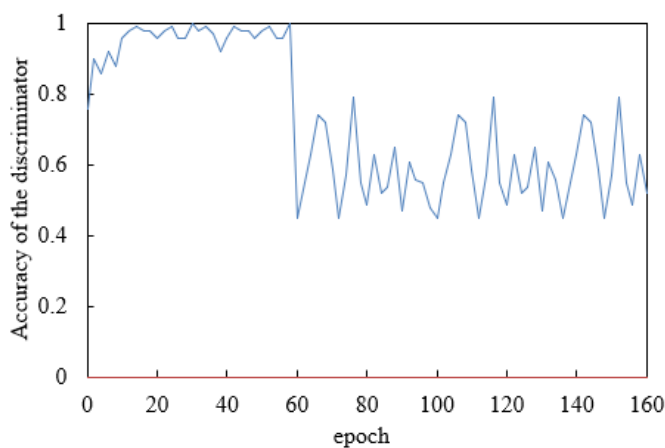
Figure 8. Accuracy changes of the discriminator in the lung CT image joint segmentation model based on domain-adaptive adversarial network

According to the experimental data shown in Figure 8, we can observe the trend of accuracy changes of the discriminator in the lung CT image joint segmentation model based on a domain-adaptive adversarial network. During the initial

training phase (0 to 80 epochs), the discriminator's accuracy gradually increased from 0.76 to 0.96, showing significant progress in learning source domain data features and enhancing its discrimination ability for source domain data. As the training progressed to between 100 and 140 epochs, the accuracy reached a peak (0.98 to 1.00), indicating that the discriminator could very accurately distinguish between real and generated data at this stage. However, in further training (160 epochs), the accuracy declined (to the range of 0.92 to 0.79), which may suggest overfitting or a disruption in the dynamic balance between the discriminator and the generator. Notably, the early discriminator accuracy setting was always at 0, meaning that the model's training was conducted without early overfitting prevention measures. From these results, we can conclude that the lung CT image joint segmentation method based on a domain-adaptive adversarial network demonstrated its effectiveness at different stages of model training. The accuracy improvements in the initial phase reflect the model's effectiveness in learning to distinguish source domain data. The further increase and peak in accuracy during the mid-training phase indicate that the model's learning and adaptation capabilities reached an optimal level, successfully promoting knowledge transfer between the source and target domains, which is crucial for enhancing the model's generalization ability. Although a decline in accuracy occurred in the later stages of training, this also provides important clues for further optimizing the model, preventing overfitting, and maintaining the balance between the source and target domains.

Table 1. Domain adaptation results from source domain to target domain 1

Model	Task 1_*Dice* Coefficient	Task 2_*Dice* Coefficient	Absolute Error Rate
3D U-Net	0.8125	0.8745	0.1254
Dense V-Net	0.8123	0.8752	0.1147
Bayesian SegNet	0.8369	0.8956	0.1236
The Proposed Model	0.8354	0.8745	0.1147

Table 2. Domain adaptation results from source domain to target domain 2

Domain Adaptation Method	Task 1_*Dice* Coefficient	Task 2_*Dice* Coefficient	δ
3D U-Net	0.6321	0.7254	0.1124
Dense V-Net	0.6254	0.7456	0.0935
Bayesian SegNet	0.6359	0.7652	0.0854
The Proposed Model	0.6348	0.7789	0.0715

Analyzing Tables 1 and 2, we can observe the performance of the lung CT image joint segmentation model based on domain-adaptive adversarial networks across two different target domains. In the tables, the source domain is a large, well-labeled, high-quality lung CT image dataset, specifically the publicly available Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI) dataset. The target domains are chosen from two lung CT image datasets that exhibit clear differences in certain features compared to the source domain. These differences include scanning equipment (CT machines from different manufacturers), scanning parameters (such as slice thickness and image resolution), and patient demographics (variations due to age, gender, and geographical location).

Task 1 is an unsupervised domain adaptation for joint segmentation, while Task 2 is a semi-supervised or weakly supervised domain adaptation for joint segmentation. In the domain adaptation from the source domain to Target Domain 1, the model achieved a Dice coefficient of 0.8354 in Task 1 and 0.8745 in Task 2. Compared to other models (3D U-Net, Dense V-Net, and Bayesian SegNet), our model demonstrated comparable performance to the best-performing model in Task 2, and matched the absolute error rate with Dense V-Net at 0.1147. This indicates good domain adaptation effects in Target Domain 1, effectively handling lung nodule detection and classification tasks. In the adaptation results from the source domain to Target Domain 2, the model achieved a Dice coefficient of 0.6348 in Task 1 and 0.7789 in Task 2, with the best performance differential (δ) between the two tasks at 0.0715, lower than all other comparison models. This suggests that our model exhibits better stability and consistency in the more challenging Target Domain 2.

These experimental results demonstrate the significant effectiveness of the proposed lung CT image joint segmentation method based on domain-adaptive adversarial networks. Especially in scenarios with significant feature differences between the source and target domains, our model not only achieves competitive performance in both types of tasks but also excels in maintaining consistency between task performances. This highlights the model's robust adaptability and efficacy in challenging domain adaptation scenarios.

Table 3. Ablation study results of the lung CT image joint segmentation model based on domain-adaptive adversarial networks

Task	Source Domain → Target Domain	Without Adversarial Training	With Adversarial Training
Task 1	Source Domain →Target Domain 1	0.8754	0.8754
	Source Domain →Target Domain 2	0.7126	0.7985
	Target Domain 1→Source Domain	0.8695	0.8741
	Target Domain 1→Target Domain 2	0.8123	0.8123
	Target Domain 2→Source Domain	0.6895	0.7456
	Target Domain 2→Target Domain1	0.8326	0.8795
Task 2	Source Domain →Target Domain 1	0.7321	0.8324
	Source Domain →Target Domain 2	0.6124	0.6325
	Target Domain 1→Source Domain	0.7236	0.7895
	Target Domain 1→Target Domain 2	0.5248	0.5412
	Target Domain 2→Source Domain	0.6124	0.6658
	Target Domain 2→Target Domain 1	0.5269	0.6693

Table 4. Experimental results of different lung CT image classification models

Model	Year	Back Bone	Training Set acc/%	Testing Set acc/%
CBAM	2021	*Resnet*12	75.23±0.48	91.25±0.32
FPN	2021	*Resnet*12	72.14±0.64	88.25±0.36
PANet	2021	*Resnet*12	75.74±0.82	85.23±0.51
DANN	2022	*Resnet*12	75.25±0.62	88.32±0.42
ADDA	2022	*Resnet*12	85.23±0.18	89.23±0.12
MAML	2022	*Resnet*12	75.14±0.57	92.23±0.91
*GCNs with ResNet*12	2023	*Resnet*12	77.24±0.21	87.24±0.15
Contrastive Learning	2023	*Resnet*12	75.85±0.52	91.24±0.41
RotNet	2023	*Resnet*12	77.11±0.51	92.23±0.35
The proposed model	2023	*SF Resnet*12	82.36±0.45	93.58±0.38

Table 3 displays the performance comparison of the lung CT image joint segmentation model based on domain-adaptive adversarial networks across different domain adaptation tasks, with and without the strategy of adversarial training. An analysis of the ablation study results clearly shows that in most tasks, the performance of models incorporating adversarial training surpasses those without it. Notably, in the transition from the source domain to Target Domain 2, adversarial training increased the Dice coefficient from 0.7126 to 0.7985, demonstrating significant improvement; similarly, in the adaptation from Target Domain 2 to the source domain, the Dice coefficient improved from 0.6895 to 0.7456. These results confirm the effectiveness of adversarial training in promoting the model's ability to learn cross-domain generalizable features, especially in scenarios with significant differences between the source and target domains.

From the results of Task 2, a similar trend can be observed, where adversarial training significantly raised the Dice coefficient from 0.7321 to 0.8324 for the source domain to Target Domain 1 adaptation, and from 0.5269 to 0.6693 for the Target Domain 2 to Target Domain 1 adaptation. This further underscores the potential and advantage of adversarial networks in addressing issues of insufficient labeled data. Overall, these ablation study results strongly suggest that the joint segmentation strategy for lung CT images based on domain-adaptive adversarial networks proposed in this paper effectively enhances the model's generalization ability and classification accuracy across different data sources. By comparing the performance differences with and without adversarial training, the critical role of adversarial training in improving model adaptability and reducing distribution differences between the source and target domains is clearly demonstrated.

Table 4 details the experimental results of various lung CT image classification models using ResNet12 as the backbone network across different years, including the accuracy of each model on both the training and testing sets. From the table, it is observed that since 2021, the performance of these models has significantly improved with the application of various advanced techniques such as CBAM, FPN, PANet, DANN,

ADDA, MAML, GCNs with ResNet12, Contrastive Learning, and RotNet. However, the model based on SF_ResNet12 proposed in this paper demonstrated the best performance in 2023, achieving an accuracy of 82.36±0.45% on the training set and an impressive 93.58±0.38% on the testing set. This result not only surpasses other models from the same year but also exceeds the performance of all models from previous years.

These experimental results clearly prove the effectiveness of the research on few-sample lung CT image classification based on the improved DPGN. Particularly in terms of testing set accuracy, the performance of this model highlights its excellent generalization ability, which is especially crucial for medical image processing as it directly relates to the model's applicability in real medical scenarios. The enhancement in performance is attributed to the improved DPGN algorithm proposed in this paper, which, by deeply mining the rich information within a small number of samples, significantly enhances the model's ability to recognize new, unseen samples. This breakthrough underscores the potential of advanced machine learning techniques in transforming the field of medical image analysis, enabling more accurate and reliable diagnostics in clinical settings.

Table 5 displays the ablation study results for the few-sample lung CT image classification model based on the improved DPGN with various components combined. From the table, it can be observed that as the model progressively integrates components like the SimAM, FRN layer, Orthogonal Decomposition (OD) convolution, MD, and EMD, both the training and testing set accuracies gradually improve. The most basic model, without any integrated advanced components, had a testing set accuracy of 92.35%. As SimAM, FRN layer, OD convolution, and MD and EMD were sequentially introduced, the model's testing set accuracy improved to 81.24%, 92.26%, 93.58%, and finally 94.25%, respectively. This indicates that the addition of each component effectively enhanced the model's performance, especially on the testing set. These ablation study results emphasize the effectiveness of the few-sample lung CT image classification algorithm based on improved DPGN. Each step of component addition positively impacted the model's performance, particularly the increase in testing set accuracy, demonstrating the model's generalization ability to new samples. The high accuracy achieved by the final model validates the efficacy of the proposed method in enhancing the performance of lung CT image classification, especially in a few-sample learning environment. These experimental results not only prove the superiority of the improved DPGN in handling few-sample issues but also show that by introducing advanced deep learning techniques and algorithmic improvements, the performance and practical value of models in tasks like lung nodule detection and classification can be significantly enhanced.

Table 5. Ablation study results of the few-sample lung CT image classification model based on improved DPGN

SimAM	FRN_layer	OD Conv	MD +EMD	Training Set acc/%	Testing Set acc/%
				74.26	92.35
√				78.26 (±2.69)	81.24 (±1.29)
√	√			81.24 (±5.01)	92.26 (±2.19)
√	√	√		82.84 (±7.24)	93.58 (±2.39)
√	√	√	√	83.69 (±7.69)	94.25 (±2.69)

5. CONCLUSION

This paper effectively enhanced the accuracy of deep learning-based CT image analysis in lung nodule detection and classification through two main research directions. Firstly, by proposing a lung CT image joint segmentation method based on domain-adaptive adversarial networks, this study significantly enhanced the model's generalization ability when handling data from different sources. Secondly, addressing the common issue of insufficient labeled data, this study developed a few-sample lung CT image classification algorithm based on the improved DPGN, effectively improving classification performance. In specific experimental analyses, this paper detailed the adaptability of the domain-adaptive adversarial network model across multiple domains, clarified the contributions of various components to model performance through ablation studies, and compared the performance of various lung CT image classification models, highlighting the superiority of the proposed model. Particularly noteworthy is that, through adversarial training strategies and the improved DPGN algorithm, the model not only performed excellently on source domain data but also demonstrated good generalization ability and high accuracy on target domain data, which is especially important in the field of cross-domain medical image processing.

However, this study also has certain limitations. For example, although the model performed well in two specific domain adaptation tasks, its applicability and stability in a broader range of medical image datasets and more complex cross-domain scenarios need further verification. Additionally, the computational resource consumption and training time under specific configurations are important directions for future optimization. Future research could explore the following directions: further optimize the model structure and algorithms to reduce dependence on computational resources and enhance model efficiency; expand the model's applicability to test its generalization capability on more medical image analysis tasks; and explore new data augmentation and simulation techniques to address extreme data imbalances and the scarcity of labeled data. Through these efforts, it is hoped to further advance the application of deep learning technologies in the field of medical image analysis, providing more powerful and accurate tools for clinical diagnosis and disease research.

REFERENCES

[1] Xu, X., Wei, Y., Zheng, J., et al. (2023). Multi-scale supervised contrastive learning for benign-malignant classification of pulmonary nodules in chest CT scans. In 2023 IEEE 20th International Symposium on Biomedical

Imaging (ISBI), Cartagena, Colombia, pp. 1-4. https://doi.org/10.1109/ISBI53787.2023.10230567

[2] Liu, H., Yao, X., Xu, B., Zhang, W., Lei, Y., Chen, X. (2022). Efficacy and safety analysis of multislice spiral CT-guided transthoracic lung biopsy in the diagnosis of pulmonary nodules of different sizes. Computational and Mathematical Methods in Medicine, 2022: 8192832. https://doi.org/10.1155/2022/8192832

[3] Lu, Z., Long, F., He, X. (2022). Classification and segmentation algorithm in benign and malignant pulmonary nodules under different CT reconstruction. Computational and Mathematical Methods in Medicine, 2022: 3490463. https://doi.org/10.1155/2022/3490463

[4] Lan, Y., Xu, N., Ma, X., Jia, X. (2022). Segmentation of pulmonary nodules in lung CT images based on active contour model. In 2022 14th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC), Hangzhou, China, pp. 132-135. https://doi.org/10.1109/IHMSC55436.2022.00039

[5] Chen, Y., Hou, X., Yang, Y., Ge, Q., Zhou, Y., Nie, S. (2023). A novel deep learning model based on multi-scale and multi-view for detection of pulmonary nodules. Journal of Digital Imaging, 36(2): 688-699. https://doi.org/10.1007/s10278-022-00749-x

[6] Ji, Z., Zhao, Z., Zeng, X., Wang, J., Zhao, L., Zhang, X., Ganchev, I. (2023). ResDSda_U-Net: A novel U-Net based residual network for segmentation of pulmonary nodules in lung CT images. IEEE Access, 11: 87775-87789. https://doi.org/10.1109/ACCESS.2023.3305270

[7] Kim, D., Han, J.Y., Baek, J.W., Lee, H.Y., Cho, H.J., Heo, Y.J., Shin, G.W. (2023). Effect of the respiratory motion of pulmonary nodules on CT-guided percutaneous transthoracic needle biopsy. Acta Radiologica, 64(7): 2245-2252. https://doi.org/10.1177/02841851221144616

[8] Shi, F., Chen, B., Cao, Q., et al. (2021). Semi-supervised deep transfer learning for benign-malignant diagnosis of pulmonary nodules in chest CT images. IEEE Transactions on Medical Imaging, 41(4): 771-781. https://doi.org/10.1109/TMI.2021.3123572

[9] Forero, M.G., Santos, J.M. (2021). Evaluation of deep learning techniques for the detection of pulmonary nodules in computer tomography scans. Applications of Digital Image Processing XLIV, 11842: 335-341. https://doi.org/10.1117/12.2594562

[10] Zhang, W., Cui, L. (2021). Detection algorithm of pulmonary nodules based on deep learning. In 2021 2nd International Conference on Big Data & Artificial Intelligence & Software Engineering (ICBASE), Zhuhai, China, pp. 185-188. https://doi.org/10.1109/ICBASE53849.2021.00042

[11] Fu, Y., Xue, P., Zhao, P., et al. (2022). 3D multi-resolution deep learning model for diagnosis of multiple pathological types on pulmonary nodules. International Journal of Imaging Systems and Technology, 32(1): 74-87. https://doi.org/10.1002/ima.22642

[12] Zhou, C., Zhao, X., Zhao, L., Liu, J., Chen, Z., Fang, S. (2022). Deep learning-based CT imaging in the diagnosis of treatment effect of pulmonary nodules and radiofrequency ablation. Computational Intelligence and Neuroscience, 2022: 7326537. https://doi.org/10.1155/2022/7326537

[13] Wang, J., Chen, X., Lu, H., et al. (2020). Feature-shared adaptive-boost deep learning for invasiveness classification of pulmonary subsolid nodules in CT images. Medical Physics, 47(4): 1738-1749. https://doi.org/10.1002/mp.14068

[14] Wang, D., Zhang, T., Li, M., Bueno, R., Jayender, J. (2021). 3D deep learning based classification of pulmonary ground glass opacity nodules with automatic segmentation. Computerized Medical Imaging and Graphics, 88: 101814. https://doi.org/10.1016/j.compmedimag.2020.101814

[15] Gao, R., Li, T., Tang, Y., et al. (2022). Reducing uncertainty in cancer risk estimation for patients with indeterminate pulmonary nodules using an integrated deep learning model. Computers in Biology and Medicine, 150: 106113. https://doi.org/10.1016/j.compbiomed.2022.106113

[16] Gao, Y., Tan, J., Liang, Z., Li, L., Huo, Y. (2019). Improved computer-aided detection of pulmonary nodules via deep learning in the sinogram domain. Visual Computing for Industry, Biomedicine, and Art, 2: 15. https://doi.org/10.1186/s42492-019-0029-2

[17] Yang, J., Deng, H., Huang, X., Ni, B., Xu, Y. (2020). Relational learning between multiple pulmonary nodules via deep set attention transformers. In 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), Iowa City, IA, USA, pp. 1875-1878. https://doi.org/10.1109/ISBI45749.2020.9098722