ISSN 2616-5775 Vol. 6, Issue 11: 26-34, DOI: 10.25236/AJCIS.2023.061104

Segmentation of nodules in CT images of the lung using an improved V-Net network model

Xiaoru Xu[1,2,a], Lingyan Du[1,2,b,*], Dongsheng Yin[1,2,c]

[1]School of Automation and Information Engineering, Sichuan University of Science and Engineering, Zigong, China
[2]Artificial Intelligence Key Laboratory of Sichuan Province, Yibin, China
[a]xiaoru1113020@163.com, [b]dulingyan927@163.com, [c]1551327511@qq.com
[*]Corresponding author

Abstract: *Lung cancer has traditionally exhibited high incidence and mortality rates, making the early detection and treatment of the disease essential in reducing mortality rates. We present a new network design to enhance the separation of lung nodules in CT lung images. The design reduces the problem of missed or incorrect segmentation due to unclear nodule morphology, varying shapes, and attachment to the pleura. In this paper, the V-Net serves as the foundation network, paired with a multi-scale feature network to enhance the original's leap connections. Two distinct attention mechanism modules are integrated into the network encoding and decoding for increased feature extraction of lung nodules. The Log-Cosh Dice Loss replaces the original loss function to address the issue of non-convexity in the Dice loss function. Additionally, the lung 3D images are cropped to resolve the problem of imbalanced distribution of positive and negative samples within lung CT images. We evaluated the performance of the model on the LUNA16 dataset. The evaluation results demonstrate the superiority of the model. We observed objective improvements compared to the initial network; 6.9% improvement in DSC values, 11.5% improvement in MIoU values, 8.5% improvement in accuracy and 99.8% pixel accuracy. This method has been found to effectively prevent missed lung nodules and produce satisfactory segmentation results.*

Keywords: *Deep learning, V-Net network, Attention mechanism, LIDC-IDRI*

1. Introduction

The lung cancer has always been one of diseases with the highest incidence and mortality among cancer disease, with over two million lung cancer cases (11.4% of all new cancers) and over one million deaths from lung cancer (18.0% of all cancer deaths) worldwide in 2020[1]. The key to reducing lung cancer mortality is to be able to detect and treat lung cancer in its early stages[2]. Pulmonary nodules are prominent in the early stage of lung cancer, and the size, type and location of pulmonary nodules are important predictors of the severity of lung cancer[3].

Currently, low-dose computed tomography (LDCT) of the chest is the most commonly used method for lung cancer screening, which results in a significant increase in the detection rate of lung nodules and provides a large sample of image data for lung nodule studies[4]. However, the manual detection of a large amount of CT image data is not only time-consuming, but also prone to errors due to false positives. In this context, automatic segmentation of lung nodules has significant research implications for assisting doctors in the diagnosis of lung cancer. Deep neural networks can be trained with big data and can automatically learn and extract features from the image data input to the network, eventually achieving automatic segmentation of lung nodules. Wang et al.[5] proposed a semi-automatic centrally focused CNN for voxel classification, but the model was not ideal for small nodules. Aresta et al.[6] proposed an interactive segmentation network iW-Net, which takes into account the user input and can greatly reduce the number of parameters, but the model must correct the lung nodule segmentation based on expert annotations and corresponding simulated user input. Jianshe Shi et al.[7] proposed a multi-scale U-Net-based automatic lung nodule segmentation algorithm, which could better solve vascular adhesion type and chest wall adhesion type lung nodules, but the segmentation took longer time. The above methods are based on two-dimensional image data for processing, while medical data is mostly three-dimensional. Using a three-dimensional network model for lung nodule segmentation will preserve its spatial characteristics and improve the lung nodule segmentation effect. Currently, the 3D U-Net basic network

ISSN 2616-5775 Vol. 6, Issue 11: 26-34, DOI: 10.25236/AJCIS.2023.061104

and 3D V-Net basic network used for image segmentation are often used to improve and then used for segmentation to obtain better segmentation results. Wenhao Wu et al.[8] segmented the lung nodules in CT lung images based on 3D U-Net networks combined with 3D conditional random field, and this method has a better segmentation effect on adherent lung nodules and ground glass lung nodules (GGNs). Kido S et al.[9] proposed a nested 3D fully connected convolutional network model with a residual cell structure, which is capable of robust and accurate 3D segmentation of the lung nodule region. Although the above methods can obtain better segmentation results of lung nodules, there is still much room for improvement in segmentation accuracy compared with the gold standard of segmentation by doctors.

In this paper, the V-Net network with residual structure is the basic network, and the following improvements are made to reduce the phenomenon of missing segmentation and mis-segmentation during lung nodule segmentation: 1. multi-scale feature fusion for each layer using a new jump connection to exploit the multi-scale features in image segmentation; 2. adding different attention mechanisms to the encoding-decoding part of V-Net to increase the feature extraction of spatial information of lung nodules; 3. Using Log-Cosh Dice Loss to replace the original loss function and solve the problem of non-convexity of the Dice loss function. Through the further processing and fusion of features, the improved network model can achieve efficient and accurate segmentation of various types of pulmonary nodules, and improve the evaluation index of pulmonary nodule segmentation.

2. Materials and Methods

2.1. Network structure in this paper

The V-Net network has a typical encoding-decoding structure complex, which can be trained end-to-end on 3D images, and its jump connections fuse shallow simple features with deep abstract features. In this paper, based on the V-Net network, and the basic framework of the network structure in this paper is shown in Figure 1.

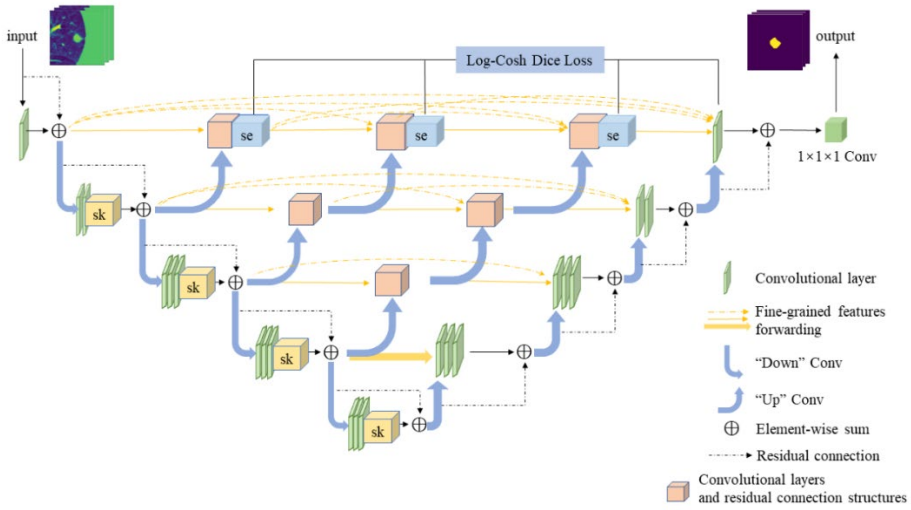

Figure 1: Network structure of this paper.

We use the V-Net network as the base network and makes the following improvements on the basis of the original V-Net network: (i) using the new jump connection to perform multi-scale feature fusion for each layer; (ii) SE attention mechanism module in up-sampling to assign different weight values to channels, and add SK attention mechanism module in down-sampling to strengthen the channel feature information of the input feature map; (iii) using Log-Cosh Dice Loss to replace the original loss function, which solves the non-convexity problem of Dice loss function. In addition, the convolution kernel uses two 3×3×3 voxel convolution to replace the 5×5×5 voxel combination in the original V-Net network to increase the perceived field of view of the network, and adds the Batch-Normalization (BN) module to each part of the network and a Dropout value of random depth to the residual network to reduce the overfitting of the network.

2.2. Principle of jump connection structure

The feature extraction stage of the original V-Net network has two shortcomings: (i) the optimal depth

ISSN 2616-5775 Vol. 6, Issue 11: 26-34, DOI: 10.25236/AJCIS.2023.061104

is priori unknown, and requires step-by-step experiments to reach the optimal depth; (ii) the jump connection is quite limited. The feature map can be fused only when the subnetwork channels of the encoder and decoder are the same, moreover, the network only pays attention to the complex features on a single scale and ignores the features on other scales. Lung nodules are often prone to the missed detection because of their small size. Therefore, it is necessary to make full use of the characteristics of lung nodules at different scales in order to improve the segmentation performance of the network. Consequently, in this paper, the jump connection of the V-Net network is improved by using the improved method of jump connection of the U-Net++ [10]. The improved jump connection structure is shown in Figure 2.

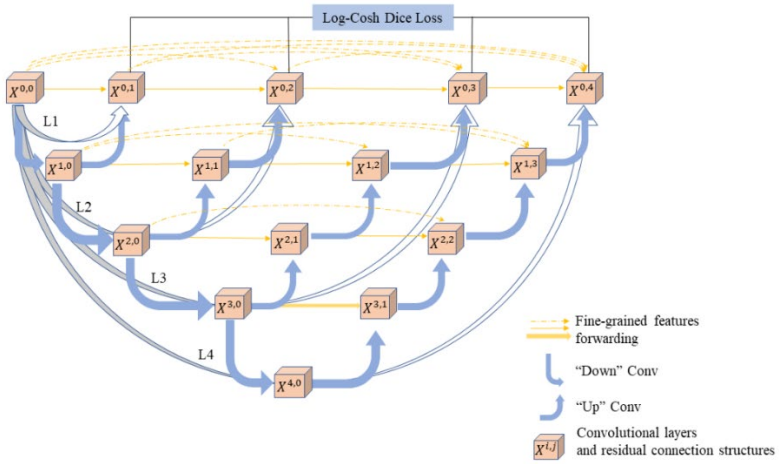

Figure 2: Multi-scale jump connection.

Where L1, L2, L3, and L4 are the four different scales of the multi-scale structure, and each scale network has a symmetric encoder-decoder structure through which the features at the four different scales can be obtained. $X^{i,j}$ denotes the nodes in the network to replace the convolutional layers and their residual connectivity structure in the V-Net network, in which the i denotes the ordinal number of the down-sampling layer in the encoder, and the j denotes the ordinal number of the nodes with the same feature resolution. The network node uses a convolution kernel with a size of 3 × 3 ×3 and a step of 1 to perform convolution operations on the feature map, and the down-sampling is completed by a convolution kernel with a size of 3 × ×3 3 and a step of 2 for convolution operations, and the up-sampling is completed by a convolution kernel with a size of 3 × 3 ×3 and a step of 2 for deconvolution operations.

This structure aggregates features of different semantic scales on the decoder subnetwork, which can form a highly flexible feature fusion scheme and can alleviate unknown network depths through V-Nets at different depths.

2.3. Principle of attention mechanism module

To prevent the network from overfitting, this paper adds the Batch-Normalization (BN) module[11] to the network. After encoding the convolutional layer, the SK attention mechanism module[12] is added to strengthen channel feature information of the input feature map,, and the SE attention mechanism module[13] is added after decoding the convolutional layer to assign different weights to the channels to improve the model performance.

The BN module effectively solves the problems such as slow convergence, unstable learning and gradient disappearance during the training process. Its core formula is:

$$y^{(i)} = \gamma^{(i)} \frac{x^{(i)} - \mu^{(i)}}{\sqrt{\left(\sigma^{(i)}\right)^2 + \epsilon}} + \beta^{(i)} \tag{1}$$

where the superscript i denotes the dimension of the data and BN is performed independently on each dimension of a batch of data; $x^{(i)}$ is the input data and $y^{(i)}$ is the output data after BN; $\mu^{(i)}$ and $\sigma^{(i)}$ are the mean and standard deviation of the current batch of input data, respectively. $\beta^{(i)}$ and $\gamma^{(i)}$ are the learnable translation and scaling parameters, respectively; ϵ is to prevent the denominator from being zero.

In V-Net network, not all features obtained by the encoder can be effectively used for segmentation,

ISSN 2616-5775 Vol. 6, Issue 11: 26-34, DOI: 10.25236/AJCIS.2023.061104

but features from different channels and spatial locations have different weights in segmentation. Therefore, this paper introduces the SE attention and SK attention mechanism module.

The structure of the SE attention mechanism and the structure of the SK attention mechanism is shown in Figure 3.

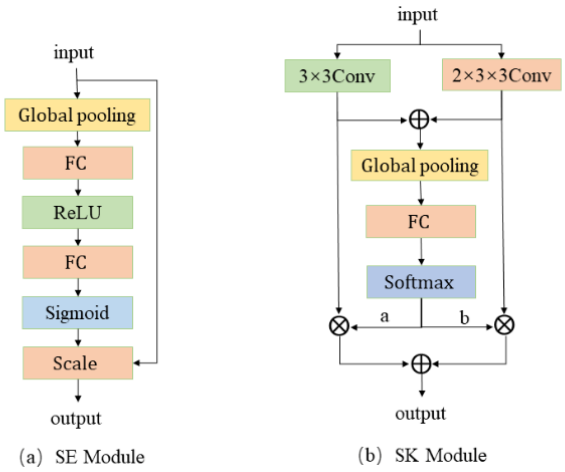

Figure 3: Structure of the SE module and SK module mechanism.

The SE attention mechanism consists of two main parts, compression and stimulation. In the compression part, the feature map undergoes a global average pooling operation to obtain a 1×1×C feature vector.

In the excitation part, the different weight values are obtained by learning through FC fully connected layer operations, where the first layer activation function is ReLu and the second layer activation function is Sigmoid. The operations involved are formulated as follows:

$$\mathcal{F}_{ex}(\mathrm{z},\mathrm{W}) = \delta\big(\mathrm{g}(\mathrm{z},\mathrm{W})\big) = \delta\big(\mathrm{W}_2\delta(\mathrm{W}_1 z)\big) \tag{2}$$

Where W_1 and W_2 are two fully connected layers.

In the SK attention mechanism, the dimension of the input feature map is C × H × W. In the Split stage, different convolution kernels can be used to convolve the original map to generate multiple paths, and in this paper, considering the small target of lung nodule segmentation, the smaller 3×3 and two 3 × 3 convolution kernels are used to convolve any input feature map to obtain feature map $\tilde{\mathrm{U}}$ and feature map $\hat{\mathrm{U}}$ respectively, then fuse to obtain a new feature map U. Involving the following equation:

$$\mathrm{U} = \tilde{U} + \hat{U} \tag{3}$$

The Fuse stage combines information from multiple pathways to obtain the global selection weights, and this step selectively filters the output of the previous layer, mainly through a gating mechanism, so that each branch carries a different stream of information to the next neuron. The equations involved are as follows:

$$\mathcal{F}_{gp}(\mathrm{U}_c) = \frac{1}{H\times W}\sum_{i=1}^{H}\sum_{j=1}^{W}\mathrm{U}_c\,(i,j) \tag{4}$$

$$\mathcal{F}_{fc}(s) = \delta\big(\mathcal{B}(Ws)\big) \tag{5}$$

The soft attention between channels can be selected with different dimensions. The channel attention information a(C×1×1) obtained by applying the softmax operation is multiplied channel by channel with the previous feature map (C×H×W) processed by the convolution kernel, and the final output is a feature map with channel attention dimension (C×H×W); the channel attention information b(C×1×1) obtained by applying the softmax operation is multiplied channel by channel with the previous feature map (C×H×W) processed by another convolution kernel, and the final output is another feature map with channel attention dimension (C×H×W). The equation involved is as follows:

$$a_c = \frac{e^{A_c z}}{e^{A_c z} + e^{B_c z}} \tag{6}$$

ISSN 2616-5775 Vol. 6, Issue 11: 26-34, DOI: 10.25236/AJCIS.2023.061104

$$\sigma(z)_j = \frac{e^{zj}}{\Sigma_{k=1}^{K} e^{zk}} \quad for\ j = 1, \ldots, K. \tag{7}$$

$$b_c = \frac{e^{B_c z}}{e^{A_c z} + e^{B_c z}} \tag{8}$$

The Select stage aggregates the feature maps of kernels of different sizes according to the selection weights, and finally fuses the feature maps of convolutional kernel A and convolutional kernel B channel attention to obtain a feature map V with channel attention dimension (C×H×W). Involving the following equation:

$$\mathrm{V} = \mathrm{A} + \mathrm{B} \tag{9}$$

2.4. Log-Cosh Dice loss function

The loss function of the original V-Net is Dice loss, which is also a common similarity coefficient loss function for medical image segmentation, and it focuses more on mining the foreground region during training. The expression is:

$$L_{DSC} = 1 - \frac{2|X \cap Y|}{|X| + |Y|} \tag{10}$$

where X is the prediction result of the network and Y is the actual value of the region. This loss function can better solve the problem of severe imbalance between positive and negative samples when segmenting lung nodules, but the training loss tends to be unstable when mining smaller targets like for lung nodules, and gradient saturation can occur in extreme cases.

In order to solve the above problems, this paper tries to use the Log-Cosh Dice Loss[14] as the loss function, which is derived by merging the Cosh(x) function and the Log(x) function. Cosh(x) value can range up to infinity, so it can be captured in a certain range by using the log function. The Log-Cosh Dice Loss function can solve the problem of non-convexity of the Dice loss. In this paper, a Laplace smoothing factor of 1 is added to the Dice loss avoiding the problem of division by zero and overfitting, thus improving the segmentation effect. The expression is as follows:

$$L_{lc-dce} = \log\left(\cosh\left(1 - \frac{2|X \cap Y| + 1}{|X| + |Y| + 1}\right)\right) \tag{11}$$

The formula for Cosh(x) is as follows:

$$cosh(x) = \frac{e^{x} + e^{-x}}{2} \tag{12}$$

2.5. Experimental data and environment

This paper uses the LUNA16 dataset of lung CT image slices greater than 3 mm thick from the Lung Nodule Open Reference Database: LIDC-IDRI[15] as experimental data. Generate lung nodule contour labels with the same size as the lung nodule image from an XML file of the nodule profile information annotated by the doctor.

The experimental environment for this paper is as follows:

Processor: AMD Raider 7 5700X 8-core@3400MHz; Hard disk: WD Blue SN570 1TB SSD; Graphics card: GeForce RTX 3060 (GA104); Operating system: Windows 10; Development language: Python 3.8.15; Deep learning framework: PyTorch 1.13.1.

3. Results

3.1. Evaluation indicators

For the evaluation metrics, this paper uses three evaluation metrics: Dice Similarity Coefficient (DSC), Sensitivity and Precision. DSC is a standard evaluation metric commonly used in segmentation problems. It is commonly used to calculate the similarity of two samples and is represented by a confusion matrix as:

$$DSC = \frac{2TP}{2TP + FP + FN} \tag{13}$$

Mean Intersection over Union (MIoU) represents the average ratio of the intersection and

ISSN 2616-5775 Vol. 6, Issue 11: 26-34, DOI: 10.25236/AJCIS.2023.061104

concatenation of the predicted and true outcomes, and IoU is represented by the confusion matrix as:

$$IoU = \frac{TP}{TP+FP+FN} \tag{14}$$

The Precision can be expressed in terms of the confusion matrix as:

$$P = \frac{TP+TN}{TP+TN+FP+FN} \tag{15}$$

PA is expressed as the ratio of correctly labelled pixels to the total number of correct pixels with the following formula:

$$\mathrm{PA} = \frac{\Sigma_{i=0}^{K} \mathrm{P}_{ij}}{\Sigma_{i=0}^{K} \Sigma_{j=0}^{K} \mathrm{P}_{ij}} \tag{16}$$

Where TP (true positive), indicates an actual positive sample and a positive predicted outcome; FP (false positive) means that the actual sample is negative but the predicted outcome is positive; FN (false negative) means that the actual sample is positive but the predicted outcome is negative; TN (true negative) means that the actual sample is negative and the predicted outcome is also negative. The above indicators are used for a comprehensive assessment of lung nodule segmentation results, with higher values representing better segmentation results.

3.2. Experimental process and parameter settings

The experimental process in this paper mainly includes four parts: data preprocessing, model building, model training, and model validation. The data pre-processing steps include: (i) converting the original CT image into a binary image; (ii) According to the xml annotation file in LIDC-IDRI, the binary image of the lung mask was generated, so that the original lung image data was superimposed with the lung mask image, and the two largest labels were saved to obtain the segmented lung region; (iii) The CT value range is selected, leaving the grayscale value image with the CT value at [-1200,600], which reduces the influence of other areas such as lung air and water on the extraction of target features; (iv) Resampling to ensure that the pixel pitch in the x, y, and z directions of the image is 1mm, eliminating error caused by different sampling intervals, and normalizing and deaveraging the original images. The steps of the experimental process in this paper are shown in Figure 4.

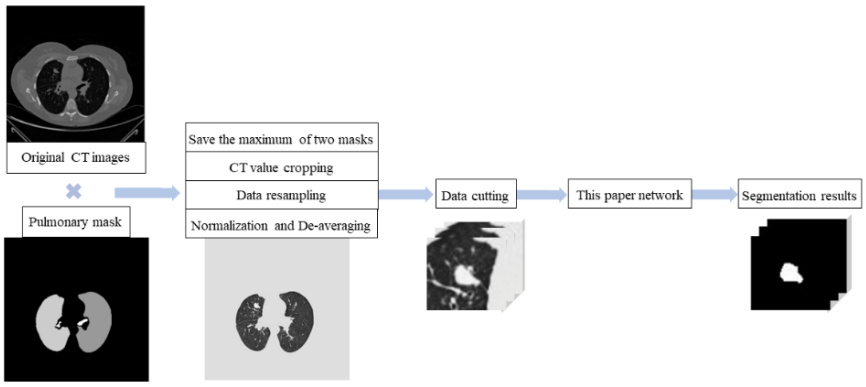

Figure 4: Step-by-step description of the data preprocessing and segmentation process.

In the training verification stage, this paper cuts the preprocessed data with a size of 96 × 96 × 16 (height ×width × depth) to reduce the negative impact of the imbalance of positive and negative samples during training. And the data is divided into training set, validation set and test set according to the ratio of 8:1:1. For parameter settings, this paper uses the validation set to adjust the model parameters, adopts the early stop mechanism, runs a total of 200 epochs before debugging, and the batch size is 4. And the method used in the model backpropagation is Stochastic Gradient Descent (SGD), in which the learning rate is set to 0.001, and every 10 iterations of training is performed, the learning rate decreases by 10%, the momentum is set to 0.9, the Dropout value is set to 0.5 and running 150 epochs, the models used for training and testing are implemented on the GPU using Python's PyTorch deep learning library.

ISSN 2616-5775 Vol. 6, Issue 11: 26-34, DOI: 10.25236/AJCIS.2023.061104

3.3. Experimental results

In this paper, according to the network parameter settings in section 3.2, Dice loss and Log-Cosh Dice Loss were used in the test set for the base V-Net network model, the multi-scale V-Net network model (MV-Net), the added attention mechanism model SK-MV-Net, the SE-MV-Net network model and the modified network model of this paper, respectively the validation was carried out and the relevant evaluation indexes derived are shown in Table 1.

Table 1: Performance comparison of different models.

Loss function	Model	DSC/%	MIoU/%	Precision/%	PA%
DSC Loss	V-Net	76.97	69.15	79.62	85.36
	MV-Net	78.52	72.48	80.23	89.21
	SK-MV-Net	79.82	73.17	81.78	98.74
	SE-MV-Net	80.23	74.48	86.90	99.38
	Our model	82.67	79.86	87.39	99.79
Log-Cosh Dice Loss	V-Net	77.95	70.23	79.71	85.64
	MV-Net	79.41	74.68	80.43	93.26
	SK-MV-Net	80.35	75.25	83.40	98.97
	SE-MV-Net	80.75	76.37	87.21	99.58
	Our model	83.87	80.65	88.19	99.83

As can be seen from Table 1, the evaluation metrics derived from the validation of the network model in this paper have significantly higher values compared to the network model before the improvement, and the addition of the dual-attention module provides a more pronounced improvement in evaluation metrics compared to the single-attention module, with better segmentation performance. Table 2 compares with three common segmentation networks and the results show that the network model in this paper also gives better segmentation results.

Table 2: Performance comparison of different model.

Model	DSC/%	MIoU/%	Precision/%	PA%
U-Net++[10]	80.67	79.56	84.73	96.29
Dense-Net[16]	82.23	78.82	80.23	97.96
CF-CNN[5]	82.15	80.02	87.83	99.31
Our model	83.87	80.65	88.19	99.83

The Figure 5 shows the 3D segmentation results of five pulmonary nodule images by four networks, where the first row represents the gold standard of pulmonary nodule contour annotation by physicians. And the original V-Net network segmentation can roughly segment the pulmonary nodules, but there are some false positive results; the MV-Net network segmentation effect is better, but there are still cases of missing segmentation; the SK/SE-MV-Net network with attention mechanism has better segmentation results than the first two networks, and there are very few missing and wrong segmentations. The SK/SE-MV-Net network with attention mechanism is superior to the first two networks in segmentation effect, and there are few false segmentation and missing segmentation, but the accuracy of segmentation of pulmonary nodule edge is not high. The combination of the two attention mechanisms and the improved jump connection improves the extraction of pulmonary nodule features in space and channels respectively, and also improves the extraction of edge features of pulmonary nodules, effectively reducing false segmentation and missing segmentation of pulmonary nodules and providing better segmentation results for the edges of pulmonary nodules.

ISSN 2616-5775 Vol. 6, Issue 11: 26-34, DOI: 10.25236/AJCIS.2023.061104

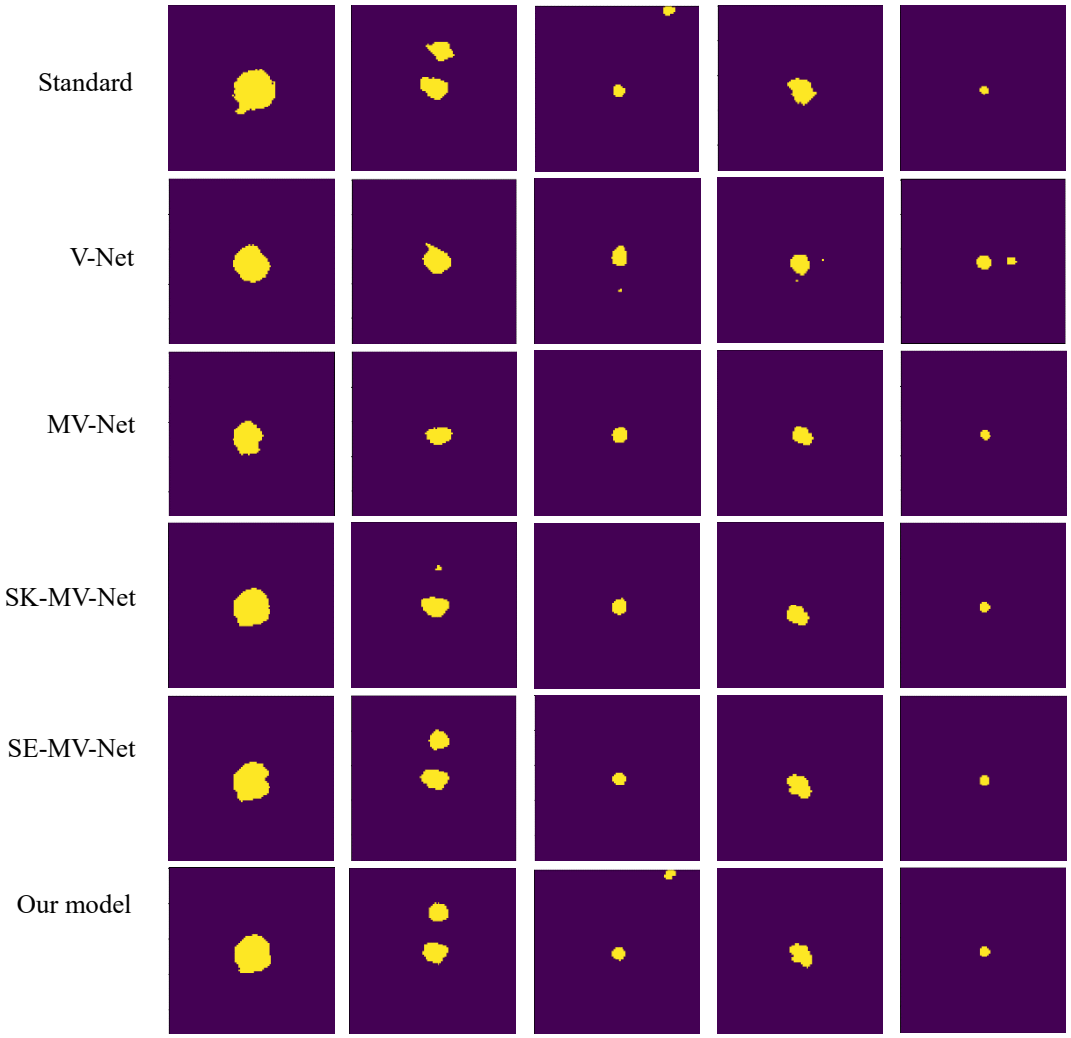

Figure 5: The 3D segmentation results.

4. Conclusions

In this paper, we invoke the jump connection of multi-scale structure and improve the segmentation model by adding the attention mechanism and applying the new loss function as well as changing the size of convolutional kernel, and the experimental results show that the network has a significant improvement in segmentation effect compared with the original network, and the evaluation indexes are 6.9% improvement in DSC value, 11.5% improvement in MIoU value, and 8.57% improvement in Precision value. 8.57%, and the pixel accuracy reaches 99.8%. The multi-scale segmentation model can effectively improve the feature extraction ability of lung nodules in details and achieve accurate segmentation of lung nodules. The Log-Cosh Dice Loss function used in this paper is also proved to be effective in segmentation through experiments.

In conclusion, the method in this paper can effectively improve the phenomenon of missed segmentation of lung nodules and achieve multi-nodule segmentation, and it also has better performance in the processing of edge features of lung nodules. In addition, in deep learning, the increase of samples has a certain optimisation effect on the training model, and further research can be done to increase the training samples through 3D data enhancement to improve the segmentation of lung nodules.

References

[1] Taher F, Prakash N, Shaffie A, et al. An overview of lung cancer classification algorithms and their performances [J]. IAENG International Journal of Computer Science, 2021, 48(4).

ISSN 2616-5775 Vol. 6, Issue 11: 26-34, DOI: 10.25236/AJCIS.2023.061104

[2] Dong T, Wei L, Nie S D. Research progress of lung nodule segmentation based on CT images. Journal of Image and Graphics, 2021, 26(4): 751-765.
[3] Tammemagi M C, Mayo J R, Lam S. Cancer in pulmonary nodules detected on first screening CT[J]. The New England journal of medicine, 2013, 369(21): 2060-2061.
[4] Vachani A, Carroll N M, Simoff M J, et al. Stage migration and lung cancer incidence after initiation of low-dose computed tomography screening[J]. Journal of Thoracic Oncology, 2022, 17(12): 1355-1364.
[5] Wang S, Zhou M, Liu Z, et al. Central focused convolutional neural networks: Developing a data-driven model for lung nodule segmentation[J]. Medical image analysis, 2017, 40: 172-183.
[6] Aresta G, Jacobs C, Araújo T, et al. iW-Net: an automatic and minimalistic interactive lung nodule segmentation deep network[J]. Scientific reports, 2019, 9(1): 11591.
[7] Shi J, Ye Y, Zhu D, et al. Comparative analysis of pulmonary nodules segmentation using multiscale residual U-Net and fuzzy C-means clustering[J]. Computer Methods and Programs in Biomedicine, 2021, 209: 106332.
[8] Wu W, Gao L, Duan H, et al. Segmentation of pulmonary nodules in CT images based on 3D-UNET combined with three-dimensional conditional random field optimization[J]. Medical Physics, 2020, 47(9): 4054-4063.
[9] Kido S, Kidera S, Hirano Y, et al. Segmentation of lung nodules on ct images using a nested three-dimensional fully connected convolutional network[J]. Frontiers in artificial intelligence, 2022, 5: 782225.
[10] Zhou Z, Siddiquee M M R, Tajbakhsh N, et al. Unet++: Redesigning skip connections to exploit multiscale features in image segmentation[J]. IEEE transactions on medical imaging, 2019, 39(6): 1856-1867.
[11] Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift[C]//International conference on machine learning. pmlr, 2015: 448-456.
[12] Li X, Wang W, Hu X, et al. Selective kernel networks[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019: 510-519.
[13] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 7132-7141.
[14] Jadon S. A survey of loss functions for semantic segmentation[C]//2020 IEEE conference on computational intelligence in bioinformatics and computational biology (CIBCB). IEEE, 2020: 1-7.
[15] Armato III S G, McLennan G, Bidaut L, et al. The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans [J]. Medical physics, 2011, 38(2): 915-931.
[16] Huang G, Liu Z, Van Der Maaten L, et al. Densely connected convolutional networks[C] //Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 4700-4708.