Physica Medica 63 (2019) 112–121

Contents lists available at ScienceDirect

Physica Medica

journal homepage: www.elsevier.com/locate/ejmp

Original paper

A cascaded dual-pathway residual network for lung nodule segmentation in CT images

Hong Liu, Haichao Cao, Enmin Song*, Guangzhi Ma, Xiangyang Xu, Renchao Jin, Yong Jin, Chih-Cheng Hung

Huazhong University of Science and Technology, School of Computer Science & Technology, Wuhan 430074, China

ARTICLE INFO

Keywords:
Lung nodule segmentation
Residual neural networks
Cascaded dual-pathway architecture
Deep learning

ABSTRACT

It is difficult to obtain an accurate segmentation due to the variety of lung nodules in computed tomography (CT) images. In this study, we propose a data-driven model, called the Cascaded Dual-Pathway Residual Network (CDP-ResNet) to improve the segmentation of lung nodules in the CT images. Our approach incorporates the multi-view and multi-scale features of different nodules from CT images. The proposed residual block based dual-path network extracts local features and rich contextual information of lung nodules. In addition, we designed an improved weighted sampling strategy to select training samples based on the edge. The proposed method was extensively evaluated on an LIDC dataset, which contains 986 nodules. Experimental results show that the CDP-ResNet achieves superior segmentation performance with an average DICE score (standard deviation) of 81.58% (11.05) on the LIDC dataset. Moreover, we compared our results with those of four radiologists on the same dataset. The comparison shows that the CDP-ResNet is slightly better than human experts in terms of segmentation accuracy. Meanwhile, the proposed segmentation method outperforms existing methods.

1. Introduction

Lung cancer is a relatively common and fatal cancer. At present, the mortality rate of all cancers is 19.5% [1,2], of which the incidence of lung cancer accounts for 66.67%, and the 5-year survival rate is 18% [3]. The use of computed tomography (CT) images for analysis and diagnosis is an essential strategy for early lung cancer diagnosis and survival time improvement [4]. It is worth noting that the nodule size is an important factor in the volumetric analysis of lung nodules [5]. In the clinics, the degree of malignancy of the nodules is related to their size. Non-calcified nodules larger than 20 mm in diameter are more likely to be cancerous (more malignant) than the smaller nodules [6]. In addition, with the increasing number of CT images of lung nodules, the development of a robust automatic segmentation model has outstanding clinical significance.

Specifically, due to the similar visual characteristics between nodules and their surroundings, this is very disadvantageous for the development of a superior segmentation model. As shown in Fig. 1, we divided the lung nodules into six types. Through the analysis of Fig. 1, we found that the contrast between the juxtapleural nodules (Fig. 1(b)) and the small nodules (Fig. 1(f)) and the surrounding background is low, especially the juxtapleural nodules, which are almost the same as the intensity of the lung wall; the intensity of each part of the cavitary nodule (Fig. 1(c)) has a large difference, and the intensity of the central region is generally small; the intensity of the ground-glass opacity nodules (Fig. 1(d)) and calcified nodules (Fig. 1(e)) exhibit a state of polarization. These challenges have made the traditional intensity-based segmentation method unable to adapt to all types of nodules.

Due to the heterogeneity of lung nodules on CT images, it has been difficult to obtain an attractive segmentation performance. These methods can be generally divided into non-machine learning methods and machine learning methods.

Non-machine learning methods for lung nodule segmentation are generally based on morphological operations [7–9], region growing [10,11], and energy optimization [12–18]. Among them, Kubota et al. use morphological and convex models to segment different types of lung nodules, the problem of which is that it is difficult to find morphological templates suitable for nodules of various sizes and types. For the region growing method, the common challenge is the selection of seed points and the setting of convergence conditions. Especially for various irregularly shaped nodules, even if shape constraints are used, it is still difficult to handle such nodules well. In addition, energy-optimized methods convert nodule-segmentation tasks into energy-minimization tasks, which differs significantly from the principles of the

* Corresponding author.
E-mail address: Enmin_Song@126.com (E. Song).

https://doi.org/10.1016/j.ejmp.2019.06.003
Received 29 December 2018; Received in revised form 4 June 2019; Accepted 7 June 2019
Available online 12 June 2019

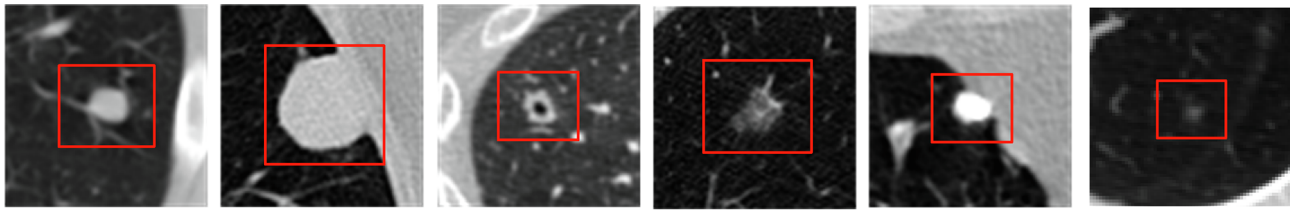

Fig. 1. Example image of a heterogeneous lung nodule in CT image. Note that (d) GGO in sub-figure (d) represents a ground-glass opacity nodule, and sub-figure (f) is a small nodule having a diameter of less than 4.4 mm.

region growing methods. However, such algorithms are often very sensitive to low-contrast nodules, particularly ground-glass nodules and juxtapleural nodules, which is similar to the region growing method [11].

Typically, in machine learning methods, researchers combine classification models with high-level features to segment lung nodules [19–22]. For example, Wu et al. proposed a stratified statistical learning approach to characterize voxels. Then, based on the probability co-occurrence maps, a conditional random field (CRF) model was trained for voxel classification [23]. Moreover, to segment lung nodule, Hu et al. first segmented the lungs, then performed vascular feature extraction based on the Hessian matrix to obtain the mask of the lung blood vessels which can be removed from the lung mask. The artificial neural networks were then used for classification [24]. Wang et al. proposed a semi-automatic central focused convolutional neural network [25] and a multi-view deep convolutional neural network [26]. They can better segment juxtapleural nodules, but there is still room for improvement in performance, especially for small nodule segmentation performance. For another, the fully convolutional networks (FCN) [27] is another trend of biomedical image segmentation. For example, the U-Net model [28] proposed by Ronneberger et al. for segmentation of the neuron structure. As well as, the V-Net model [29] proposed by Milletari et al. for the prostate segmentation in MRI.

To address the problems faced by the above six types of nodules (Fig. 1), we propose a cascaded dual-pathway residual network (CDP-ResNet), which is suitable for various types of lung nodule segmentation.

The lung nodule segmentation technique proposed in this paper has the following three innovations:

(1) For small nodules and juxtapleural nodules, the proposed CDP-ResNet model can achieve attractive segmentation performance (Table 6).
(2) The multi-view and multi-scale features of lung nodules were extracted by cascading two different-scale image patches containing multiple contiguous slices (Fig. 2).
(3) A dual-pathway architecture based on the residual block was proposed, which can extract local texture feature and richer context information of lung nodules (Section 2.1.2).
(4) Based on the sampling strategy proposed by Wang et al. [23], we further propose an improved weighted sampling strategy that can adequately sample small nodules according to the number of voxels located at the nodule edge (Section 2.2).

2. Methods and materials

We will describe our theoretical method in detail below. The method is roughly divided into four sections. Section 2.1 describes network architecture. Section 2.2 describes the proposed dual-pathway architecture. Section 2.3 introduces an improved weighted sampling strategy. Section 2.4 presents the initialization and post-processing approach. Section 2.5 describes our analysis of the data. The last subsection describes the evaluation criteria.

2.1. Network architecture

The cascaded dual-pathway residual network utilizes two different-scale image patches (each image patch contains three contiguous slices) for segmentation of lung nodules. Given a voxel in the CT image, we extract two different-scale image patches containing multiple slices as the input to the CDP-ResNet model, and then output the probability that this voxel belongs to the nodule. Fig. 2 shows the architecture of the proposed cascaded dual-pathway residual network. Table 1 shows the corresponding network parameters.

The network contains two dual-pathway residual networks that share the same structure, but the image patches used for training are different. The two dual-pathway residual networks contain 57 convolutional layers (Table 1) and two max pooling layers (Fig. 2, MP1 and MP2). It should be noted that all convolution operations and pooling operations have a stride of one and the padding mode are all “valid”. The 57 convolution layers in the CNN are divided into four categories, the first is a residual block cluster consisting of three residual blocks [30], the second is a residual block cluster consisting of one residual block, the last two residual block clusters consisting of six and nine residual blocks respectively. To speed up the convergence of the model, each convolutional layer is batch-normalized [31]. After each convolution layer, a parametric rectified linear unit (PReLU) is used as an activation function [32].

We first feed the larger-scale image patch ($65 \times 65 \times 3$) to the first dual-pathway residual network that is labeled in Fig. 2. Then, its output ($35 \times 35 \times 5$) and the smaller-scale image patch ($35 \times 35 \times 3$) are concatenated and submitted to the second dual-pathway residual network (labeled as in Fig. 2) to obtain multi-scale and multi-view features of a lung nodule. The average pooling is then applied to the output of the second dual-pathway residual network ($5 \times 5 \times 5$). Finally, a fully-connected layer is used to capture the correlation of the features produced by these two dual-pathway residual networks.

The goal of network training is to maximize the probability of the correct class. We achieve this by minimizing the cross-entropy loss of each training sample. For a given input patch belonging to {0, 1} Section 2.2 describes the proposed dual path architecture., assuming that y_n is a true label, then the loss function is defined as shown in Eq. (1):

$$L = -\frac{1}{N}\sum_{n=1}^{N}[y_n log(y_n') + (1 - y_n) log(1 - y_n')] \quad (1)$$

Among them, y_n' represents the prediction probability of CDP-ResNet, and N is the number of samples.

In the experiment, we used the Stochastic Gradient Descent (SGD) algorithm [33] as a model update method. The SGD optimizer has several parameter settings: the initial learning rate is 0.001, and then the learning rate is decreased by ten percent in every five epochs. In addition, the momentum setting is 0.9. However, due to the limitations of GPU memory, only a batch size of 32 samples are used. In addition, to avoid overfitting during the training process, we adopted the early stopping training strategy [34]: if the performance is not further improved, it will stop with an extra training of 10 epochs. The number of

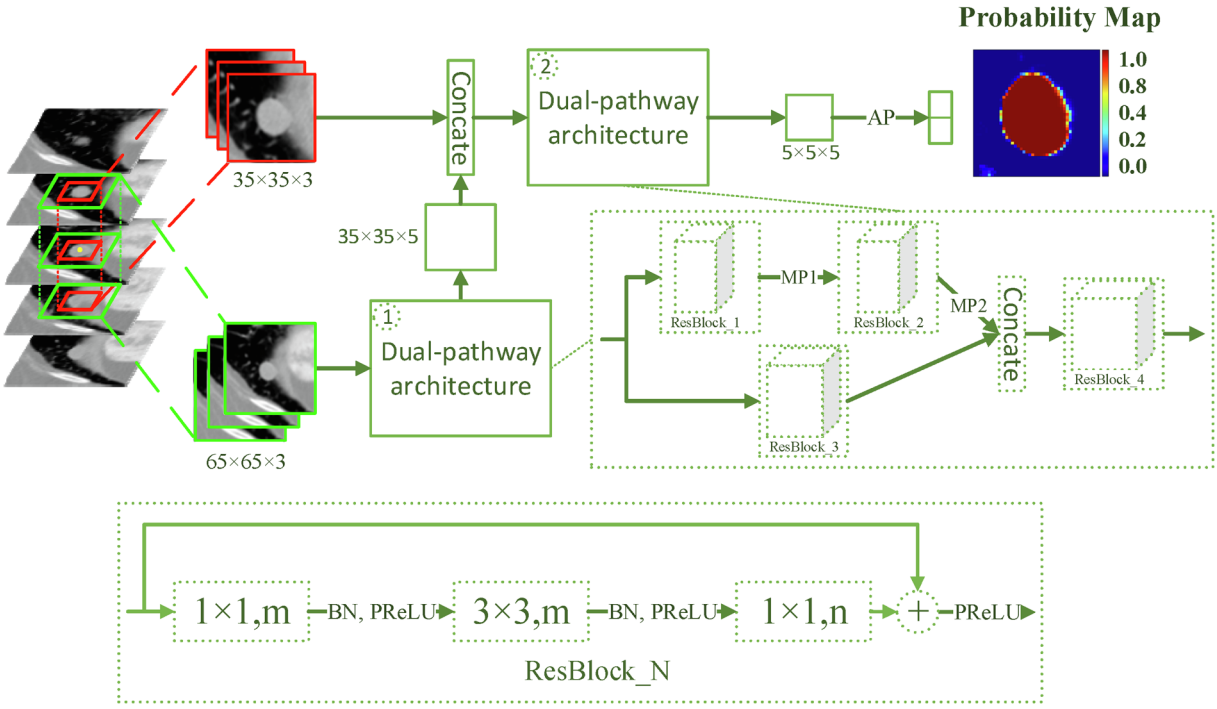

Fig. 2. Illustration of the proposed CDP-ResNet architecture. Among them, Concate, BN, AP and MP represent the concatenate operation, batch normalization operation, average pooling operation, and max pooling operation, respectively. The architecture diagrams of the residual block correspond to the bottom (ResBlock_N) of Fig. 2. The parameters m and n indicate the number of channels, and the detailed parameter information is shown in Table 1.

Table 1
Network parameters of the Dual-pathway architecture. Building blocks are shown in brackets with the numbers of blocks stacked. Downsampling is performed by ResBlock_1 and ResBlock_2 with a kernel size of 4×4, 2×2 respectively. Please note that 57 convolutional layers are calculated as $3*3+3*1+3*6+3*9=3*(3+1+6+9)=3*19=57$.

Layer name	ResBlock_1	ResBlock_2	ResBlock_3	ResBlock_4
Parameters	$\begin{bmatrix} 1\times1, & 32 \\ 3\times3, & 32 \\ 1\times1, & 64 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, & 64 \\ 3\times3, & 64 \\ 1\times1, & 64 \end{bmatrix}\times1$	$\begin{bmatrix} 1\times1, & 80 \\ 3\times3, & 80 \\ 1\times1, & 160 \end{bmatrix}\times6$	$\begin{bmatrix} 1\times1, & 10 \\ 3\times3, & 10 \\ 1\times1, & 5 \end{bmatrix}\times9$

training epochs of CDP-ResNet model is 20.

2.2. Dual-pathway architecture

The proposed dual-pathway residual network architecture consists of upper and lower pathways: one is a 7×7 pathway used in ResBlock_1 in Fig. 2 with a smaller receptive field, and the other is a 13×13 pathway used in ResBlock_3 in Fig. 2 with a larger receptive field. We refer to these two pathways as the local pathway and the global pathway. The motivation for designing this architecture is that we try to predict each voxel by two types of features: 1) the detailed features around the voxel, and 2) the richer contextual information.

The complete architecture is shown in Fig. 2 and Table 1. To fuse the features extracted by two pathways, we add a residual block (ResBlock_2) after the first residual block cluster (ResBlock_1) of the local pathway. Although this makes the effective perception field of the top layer features of each pathway to be the same, the global pathway can more directly and flexibly simulate the features of the same region. The feature map of these two pathways is then stacked and fed to the fourth residual block cluster (ResBlock_4).

There are two types of inputs for the dual-pathway residual network architecture. One is three large-scale image patches taken from three image slices, and the other is eight smaller-scale image patches. Specifically, given a voxel, we extend the current, previous, and subsequent slices centered on this voxel to extract image patches of two different scales (see Fig. 2).

In addition, we integrate the residual learning structure into our proposed segmentation network to further improve the segmentation performance of the model. Moreover, the original residual learning structure is discarded, the residual mechanism of the bottleneck structure is used, its head and end are 1×1 convolutions (to reduce and restore size), and the middle is 3×3 convolution, this structure can increase network depth while reducing network parameters [30].

2.3. Improved weighted sampling

In general, for segmentation problems, the edge voxel points of the target are critical because they usually reflect the overall characteristics of the target. If the edge voxels of the lung nodule are well recognized, even if the internal segmentation of the nodule is incomplete, a complete nodule mask can be obtained by simple morphological operations. Based on this idea, when sampling the voxels of the P class (a set of voxel points belonging to the lung nodules), we consciously pay attention to the voxel points on the edge of the lung nodules. That is, the corresponding sampling weight (the probability that the voxel point is sampled) is set based on the minimum distance between each voxel point and the edge voxel point of the lung nodule, and the smaller the distance, the larger the weight [25]. In addition, to make the voxel points near the center of the nodule have a chance to be sampled, we use the radius of the lung nodule as an adjustment factor to suppress the polarization of the sampling weight.

The formula for calculating the voxel point sampling weight of the P

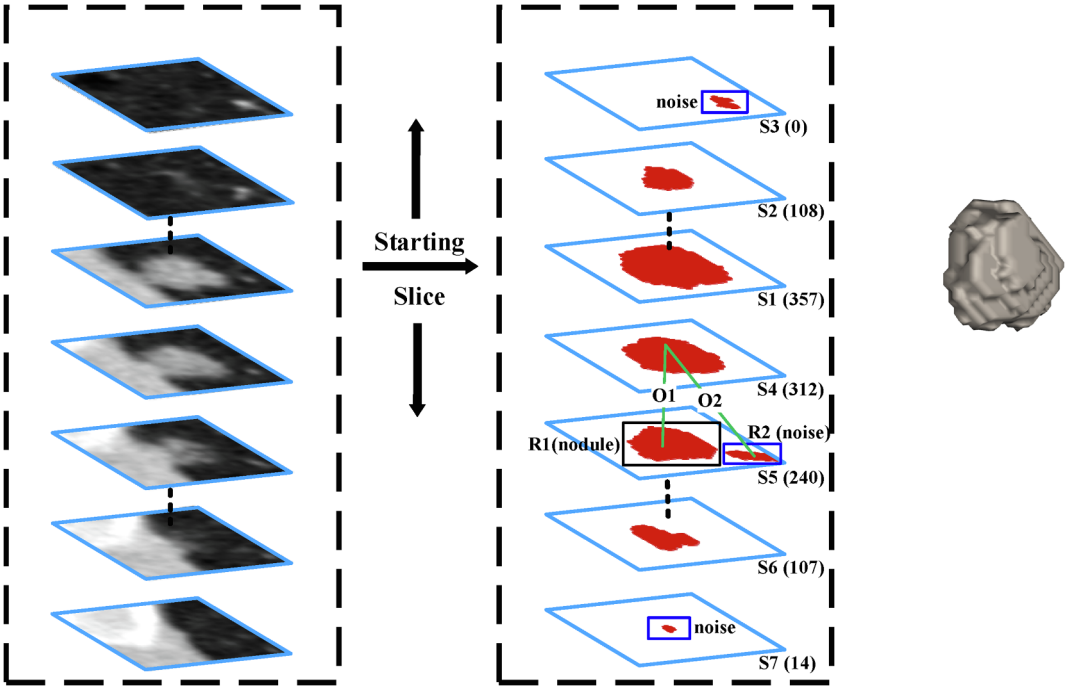

Fig. 3. Post-processing segmentation procedure. The bounding box only needs to be specified in one of the slices and then iteratively applied to the previous and next slices. The number on the right side of each slice is the intersection area of the current slice and the upper slice nodule mask, counted in voxels. The left column shows the original CT slice and the middle column displays the results of the CDP-ResNet model, where the red regions represent nodules, the region marked by the dark blue box represents the noise region. The right column shows the 3D visualization of the final segmentation result.

Table 2
Characteristic distributions of the LIDC training, validation, and testing sets. Values are shown in mean ± standard deviation.

Characteristics	Training set (n = 391)	Validation Set (n = 56)	Testing set (n = 539)
Diameter (mm)	8.30 ± 4.72	8.12 ± 4.63	7.93 ± 4.15
Sphericity	3.80 ± 0.58	3.84 ± 0.62	3.85 ± 0.58
Margin	4.07 ± 0.73	4.08 ± 0.82	4.10 ± 0.79
Spiculation	1.60 ± 0.78	1.53 ± 0.70	1.57 ± 0.74
Texture	4.56 ± 0.83	4.46 ± 0.98	4.56 ± 0.80
Calcification	5.65 ± 0.79	5.68 ± 0.78	5.66 ± 0.81
Internal structure	1.01 ± 0.16	1.03 ± 0.20	1.01 ± 0.08
Lobulation	1.73 ± 0.72	1.74 ± 0.74	1.70 ± 0.71
Subtlety	3.99 ± 0.78	3.89 ± 0.74	3.95 ± 0.75
Malignancy	2.94 ± 0.91	2.87 ± 0.77	2.92 ± 0.91

Note: The range of all characteristic values except diameter, internal structure, and calcification is 1–5, wherein the internal structure and calcification range from 1 to 4, 1 to 6, respectively. Margin indicates the clarity of the nodule edge. Spiculation and lobulation represent the amount of these shapes. Texture is a statistic of the distribution properties of the local gray information of nodules. Internal structural represents the internal composition of the nodule. Malignancy, calcification, and Sphericity indicate the possibility that the nodule is such a feature. Subtlety describes the contrast of the nodule region and its surrounding region. The characteristics of the three sets are without significant statistical difference.

class is shown in Eq. (2).

$$PW_k = \frac{\exp^{-\frac{\min d(k,t)}{r}}}{\sum \exp^{-\frac{\min d(k,t)}{r}}}, \quad k \subseteq P \tag{2}$$

where PW_k represents the sampling weight of the k-th voxel in the P class; E represents the set of voxels belonging to the edge of the lung nodule; d(k,t) represents the Euclidean distance between the k-th voxel in P and the t-th voxel in E, and r represents the radius of the lung nodule.

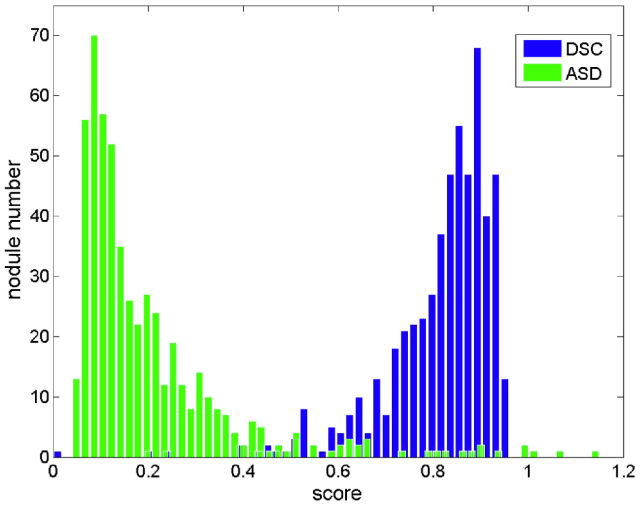

Fig. 4. DSC and ASD distributions of the LIDC testing set.

Table 3
Mean DSCs (%) of consistency comparison between CDP-ResNet and each radiologist, where R1 to R4 represent four radiologists.

	R1	R2	R3	R4	Average
R1	–	82.61	82.47	82.49	82.66 ± 0.48
R2	82.61	–	83.72	82.36	
R3	82.47	83.72	–	82.32	
R4	82.49	82.36	82.32	–	
ours	82.92	83.11	82.79	81.93	82.69 ± 0.45

For the sampling of voxel points of the N class (a set of voxel points that are not part of the lung nodules, such as blood vessels, lung walls, lung parenchyma, etc.), we consider their corresponding intensity information to calculate their sampling weights. This is mainly because

Table 4
Mean P-Value of consistency comparison between CDP-ResNet and each radiologist, where R1, R2, R3, and R4 represent four radiologists.

	R1	R2	R3	R4	Average
R1	–	0.0101	0.414	0.1454	0.1120
R2	0.0101	–	0.0754	4.93E-05	
R3	0.414	0.0754	–	0.022	
R4	0.1454	4.93E−05	0.022	–	
R*	0.3473	2.21E−05	0.03896	0.3595	0.1864

the higher intensity of the lung wall and blood vessels is more similar to the pulmonary nodules than the lung parenchyma and air. Especially for lung nodules that adhere to the lung wall, the intensity of the lung wall is almost the same as the intensity of the lung nodules. Hence, these high-related background voxels should be given a larger sampling weight. The corresponding sampling weight calculation formula is as shown in Eq. (3). It should be noted that the sampling region corresponding to the voxel of N classes is a local region other than the pulmonary nodule, and the size of this local region is determined by the circumscribed rectangle "Rect" and diameter "D" of the lung nodule.

$$NW_i = \frac{I_i \exp^{-\frac{\min_{j \subseteq E} d(i,j)}{r}}}{\sum I_i \exp^{-\frac{\min_{j \subseteq E} d(i,j)}{r}}}, \; i \subseteq N \tag{3}$$

where NW_i represents the sampling weight of the i-th voxel in the N class; d(i,j) represents the Euclidean distance between the i-th voxel in N and the j-th voxel in E; I_i represents the intensity value of the i-th voxel in the N class; r represents the radius of the lung nodule.

Finally, we also need to determine the number of voxel points sampled for each lung nodule. To adequately sample nodules of various sizes without introducing too many redundant samples, we only use the number of voxel points on the nodule edge as the reference, instead of taking all voxel points corresponding to the nodule. In the specific experiment, we set the number of sampling points to twice the number of voxel points at the edge of the nodule. To ensure that the model does not bias towards a certain class during training, we set the proportion of P class and N class samples to 1:1.

2.4. Initialization and post-processing

Since the method proposed in this paper is a semi-automatic segmentation model, it is necessary to give the volume of interest (VOI) where the nodule is located before segmentation. However, since the nodules are usually distributed over multiple CT slices, it is tedious to manually specify the region of interest (ROI) in which the nodules are located, layer-by-layer. To facilitate the doctor's operation, we performed the following post-processing operation, that is, it is only necessary to manually designate a bounding box called a starting slice on one CT slice.

Table 6
Evaluation results of the proposed method on the LIDC test set.

	LIDC testing set		LIDC testing set	
	Attached (n = 130)	Non-attached (n = 409)	Diameter < 6 mm (n = 236)	Diameter ≥ 6 mm (n = 303)
DSC (%)	78.10	82.69	79.61	83.11
ASD (mm)	0.37	0.21	0.17	0.31

The same bounding box is then applied repeatedly to the previous and next slices until the following experimental conditions are met: the nodule intersection area of the current slice and the previous slice is less than 30% of the previous slice area. For instance, slices S3 and S7 in Fig. 3 are eliminated. This is because the nodule intersection area of slice S3 and the previous slice (slice S2) is zero, and the nodule intersection area of the slice S7 and the previous slice (slice S6) is only 14 voxels, which is only 13.08% of the nodule area of the previous slice.

To eliminate the noise region, we performed the following connected region selection: 1) When noise occurs in the starting slice, we select the isolated region closest to the center of the bounding box as the nodule mask for the current slice, and 2) when the noise occurs in other slices, we select the connected region where the overlap [35] of the current slice and the previous slice nodule mask is the largest. For example, in the slice S5, it is assumed that the two candidate regions segmented by CDP-ResNet are R1 and R2 as shown in Fig. 3. The overlap ratio of these two regions with the previous slice (slice S4) is denoted as O1 and O2. Since O1 is larger than O2, we reserve the region R1 and remove the region R2 (noise).

2.5. Data and analysis

A public dataset from the Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI) was used in the experimental evaluation [36–38]. In this study, we studied 986 nodule samples annotated by four radiologists. Due to the differences in labeling between the four radiologists, the 50% consistency criterion [10] was used to generate the ground-truth boundary. All nodules are randomly divided into three subsets, corresponding to the training set, the validation set, and the testing set. The number of nodules contained in each subset was 391, 56, and 539, respectively. The statistical distribution of the three subsets is shown in Table 2, and it can be seen that they have similar data distributions.

2.6. Evaluation criteria

To evaluate the segmentation results of the CDP-ResNet model, we used the average surface distance (ASD) and dice similarity coefficient (DSC) as the primary evaluation criteria [39–41]. Moreover, to ensure the robustness of the evaluation, we also use the positive prediction value (PPV) and sensitivity (SEN) as auxiliary evaluation parameters.

Table 5
The DSC average values on different nodule groups.

Characteristics	Characteristic scores					
	1	2	3	4	5	6
Malignancy	76.67 [44]	80.33 [141]	82.98 [203]	82.75 [135]	78.51 [16]	–
Sphericity	–	81.69 [13]	79.07 [95]	81.75 [389]	85.65 [42]	–
Margin	–	73.30 [33]	81.37 [61]	81.41 [277]	83.57 [168]	–
Spiculation	82.24 [296]	80.65 [191]	83.03 [24]	79.36 [26]	84.67 [2]	–
Texture	63.49 [7]	77.03 [22]	75.95 [8]	82.10 [117]	82.13 [385]	–
Calcification	–	–	78.92 [23]	78.06 [39]	81.34 [30]	82.04 [447]
Internal structure	81.67 [536]	66.20 [3]	–	–	–	–
Lobulation	81.46 [232]	81.85 [247]	79.60 [39]	83.43 [21]	–	–
Subtlety	63.62 [1]	75.88 [28]	81.64 [87]	80.82 [304]	84.97 [119]	–

Table 7
Evaluation results corresponding to different levels of Gaussian noise added to the CT image, where m represents the mean and s represents the standard deviation.

Category	DSC	ASD	SEN	PPV	Parameter
Gaussian	81.69 ± 10.74	0.23 ± 0.31	87.32 ± 14.13	79.63 ± 13.01	m = 0, s = 0.02
Gaussian	81.56 ± 10.63	0.23 ± 0.27	87.20 ± 14.33	79.63 ± 12.97	m = 0, s = 0.05
Gaussian	81.34 ± 11.13	0.24 ± 0.26	86.89 ± 15.49	78.84 ± 12.69	m = 0.05, s = 0.05
Ours	81.58 ± 11.05	0.25 ± 0.40	87.30 ± 14.30	79.71 ± 13.59	*

Table 8
Evaluation results corresponding to different types of noise added to the CT image. Where a, b represent the parameters for determining the shape and rate of the gamma distribution, respectively, and the parameter p represents the probability of setting the intensity of voxel to 255.

Category	DSC	ASD	SEN	PPV	Parameter
Gamma	81.56 ± 10.63	0.23 ± 0.27	87.20 ± 14.33	79.63 ± 12.97	a = 25, b = 3
Uniform	81.30 ± 11.47	0.24 ± 0.35	86.98 ± 15.01	79.57 ± 13.02	U(−0.05, 0.05)
impulse	81.04 ± 11.80	0.24 ± 0.34	86.12 ± 15.56	79.77 ± 12.50	p = 0.01
Ours	81.58 ± 11.05	0.25 ± 0.40	87.30 ± 14.30	79.71 ± 13.59	*

Table 9
Mean ± standard deviation of the results for various segmentation methods. Note that, “RS” represents the Random Sampling strategy; “WS” represents the Weighted Sampling strategy proposed in [25]; “IWS” represents the Improved Weighted Sampling strategy proposed in this paper. The performance of our method is marked in bold. Note that CDP-ResNet elsewhere in the paper represents CDP-ResNet + IWS in Table 9.

Network Architecture	DSC (%)	ASD (mm)	SEN (%)	PPV (%)
CF-CNN [25]	78.55 ± 12.49	0.27 ± 0.35	86.01 ± 15.22	75.79 ± 14.73
MC-CNN [42]	77.51 ± 11.40	0.29 ± 0.31	88.83 ± 12.34	71.42 ± 14.78
MV-CNN [43]	75.89 ± 12.99	0.31 ± 0.39	87.16 ± 12.91	70.81 ± 17.57
MV-DCNN [26]	77.85 ± 12.94	0.33 ± 0.36	86.96 ± 15.73	77.33 ± 13.26
MCROI-CNN [44]	77.01 ± 12.93	0.30 ± 0.35	85.45 ± 15.97	73.52 ± 14.62
FCN-UNet [28]	77.84 ± 21.74	1.79 ± 7.52	77.98 ± 24.52	82.52 ± 21.55
FCN-VNet [29]	79.59 ± 23.00	1.55 ± 7.56	79.37 ± 24.20	83.04 ± 22.38
CDP-ResNet + RS	79.72 ± 12.10	0.27 ± 0.38	85.13 ± 14.93	77.52 ± 13.62
CDP-ResNet + WS	80.33 ± 11.61	025 ± 0.34	86.05 ± 14.93	78.20 ± 13.58
CDP-ResNet + IWS	**81.58 ± 11.05**	**0.25 ± 0.40**	**87.30 ± 14.30**	**79.71 ± 13.59**

The four formulas are defined as shown in Eqs. (4)–(7).

$$DSC = \frac{2 \times V(Gt \cap Seg)}{V(Gt) + V(Seg)} \tag{4}$$

$$ASD = \frac{1}{2}(mean_{i \in Gt} min_{j \in Seg} d(i,j) + mean_{i \in Seg} min_{j \in Gt} d(i,j)) \tag{5}$$

$$SEN = \frac{V(Gt \cap Seg)}{V(Gt)} \tag{6}$$

$$PPV = \frac{V(Gt \cap Seg)}{V(Seg)} \tag{7}$$

Among them, “Gt” represents the result of expert labeling, “Seg” represents the segmentation result of CDP-ResNet model, V represents the volume size calculated in voxel units and d (i, j) represents the Euclidean distance between the voxel i and voxel j measured in millimeters.

3. Results

This section is divided into two subsections, Section 3.1 describes the overall performance of our method, and Section 3.2 introduces the robustness of the proposed segmentation model.

3.1. Overall performance

To better observe the performance of the proposed method in the testing set, we plot the histogram between the DSC value (blue mark) and the ASD value (green mark) and the number of nodules, based on all samples in the testing set, as shown in Fig. 4. As can be seen from Fig. 4, most nodules have DSC values above 0.8.

To see if the segmentation results of our proposed method are comparable to those hand-labeled by human experts, we performed a consistency comparison between CDP-ResNet and four radiologists, as shown in Table 3. Our results show that the stability of CDP-ResNet is slightly better than that of four different radiologists. Meanwhile, the DSC between CDP-ResNet and each radiologist is 82.69% on average, which is slightly higher than the average inter-radiologist variability of 82.66%.

In addition, to quantify the difference between our segmentation result R^* and the radiologists' labeled results, we calculate the P-Value. Specifically, we use the mask generated by 50% criterion as the gold standard GT and use it as a benchmark for calculating the P-Value. In other words, we first need to obtain five sets of data by calculating the DSC between GT and R1, R2, R3, R4 and R^*. Then based on these five sets of data, we calculate the P value between them to reflect the statistical difference between our segmentation results and the radiologists' labeled results. Details on the P-Value are shown in Table 4.

We know that the smaller the p-value, the greater the statistical difference. From the analysis of Table 4, the difference between our segmentation results and the fourth radiologist is significantly smaller than those of other radiologists, and the P value between R^* and R1 is ranked as the second highest among all the four P values. Although the P value between R2 and R^* is the smallest, it is also very different from the other three radiologists, especially the difference between R2 and R4. In addition, as can be seen from the last column of Table 4, the average P value between the four radiologists is 0.1120; the average P value between our segmentation method and the four radiologists is 0.1864, which reflects in a certain extent that our segmentation results are closer to the GT than those of the radiologists' results.

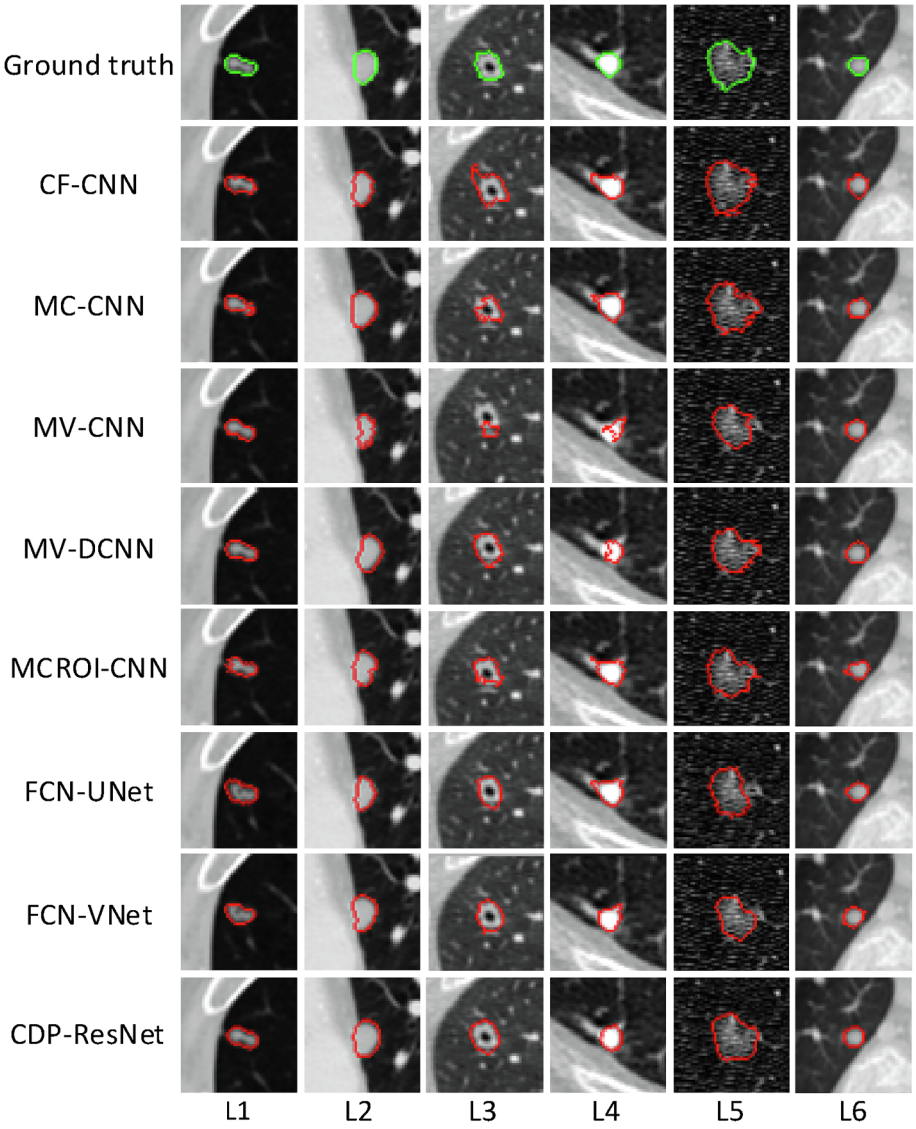

Fig. 5. Segmentation results visualization. From top to bottom: nodule with ground truth, CF-CNN segmentation, MC-CNN segmentation, MV-CNN segmentation, MV-DCNN segmentation, MCROI-CNN segmentation, FCN-UNet segmentation, FCN-VNet segmentation, and CDP-ResNet segmentation. L1-L6 are nodules of different types from the LIDC testing set.

3.2. Robustness of segmentation

To prove the robustness of the proposed method, we base the nine characteristics corresponding to each nodule as the benchmark, and divide the testing set into different groups according to the characteristic scores of the nodule. Table 5 lists the DSC average values of different nodule groups. As can be seen from Table 5, our method can handle all types of nodules with similar performance, which reflects the segmentation robustness of CDP-ResNet.

Meanwhile, we have collated the evaluation results of challenging small nodules and attached nodules. The relevant results are shown in Table 6. According to the experimental results in Table 6, it is not difficult to see that our segmentation method is independent of the attached conditions of nodules, and the size of the nodule.

To further evaluate the robustness of the proposed method, we evaluated our model by adding different levels of noise (Table 7) and different types (Table 8) of noise to the CT image. The data and experiments are shown in Tables 7 and 8.

From the observations of Tables 7 and 8, it can be seen that the noise has a certain degree of influence on our model, but this is within an acceptable range. Specifically, after we added different levels of noise and different types of noise to CT images, although DSC, SEN, and PPV showed slightly decreasing in the accuracy, the magnitude of the decline was very small. In addition, we can see that the ASD metrics used is still good when the proposed segmentation method is applied to the CT images with noise. This shows that our proposed method is robust to noise.

4. Discussion

This chapter is divided into two subsections, Section 4.1 presents the experimental comparison with other methods below, and Section 4.2 discusses the characteristics and limitations of the proposed segmentation method.

4.1. Experimental comparison

To verify the superiority of the proposed method, we conducted the following experimental comparisons, namely comparison with various different types of segmentation methods proposed recently. According to the experimental data shown in Table 9, the proposed method is superior to the existing segmentation method.

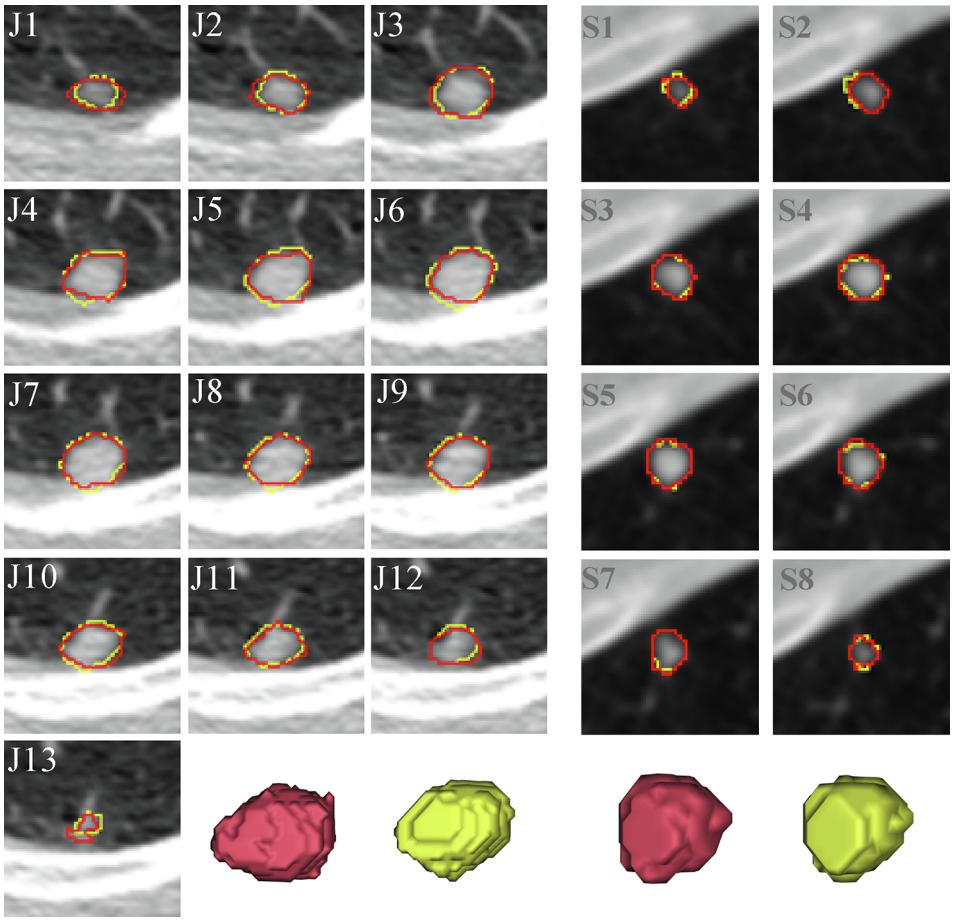

Fig. 6. Segmentation results of CDP-ResNet on juxtapleural nodule (J1–J13) and small nodule with a diameter of 4.7 mm (S1–S8) from the LIDC testing set. The red and yellow contours represent the ground truth and the segmentation results of CDP-ResNet, respectively. The red volume data and the yellow volume data corresponding to the 3-D renderings of the ground truth and the CDP-ResNet, respectively. The number in the upper left corner of each image represents the slice number where the nodule is located.

Table 9 shows the quantification results for the different types of segmentation methods. The data format is "mean ± standard deviation". To ensure the fairness of the comparison, the methods comparison with CDP-ResNet in Table 9, except that the network structure is different, other conditions are consistent with CDP-ResNet, including sampling strategy and post-processing methods. For "CDP-ResNet + RS" and "CDP-ResNet + WS", only the sampling method is different, and other conditions are the same. In addition, by comparing the last three rows of Table 9, we can find the effect of the proposed improved weighted sampling strategy.

It should be noted that the reason why the performance of the CF-CNN method we implemented differs from the original paper is that the size of the dataset used is different. Wang et al. used the dataset published by LIDC before 2015 (there are 893 data, which cannot be found). In recent years, due to the continuous updating of data, the current dataset size is 986. Similarly, for the other comparison methods in Table 9, the author did not publish the source code for various reasons. Therefore, we can only re-implemented as much as we can according to the description in the paper, and then evaluate on the same training set, validation set and testing set to ensure the fairness of the comparison.

Moreover, we also visually compare the segmentation results of different methods. We selected six representative nodules from the LIDC testing set for visual comparison. L1–L6 shown in Fig. 5 correspond to the isolated nodule, juxtapleural nodule, cavitary nodule, calcific nodule, ground-glass opacity nodule, and small nodule less than 6 mm in diameter. From the visual comparison in Fig. 5, it can be seen that the overall performance of the MV-CNN method is slightly inferior to other methods, especially for cavitary nodules and GGO nodules. For isolated nodules, MC-CNN and MCROI-CNN methods performed slightly worse. MCROI-CNN method is slightly less effective for juxtapleural nodules. For central calcified nodules, the segmentation results of the MV-DCNN method are incomplete. For small nodules, CF-CNN method is less adaptable. For ground-glass opacity nodule, U-Net and V-Net performed poorly. In contrast, CDP-ResNet is still robust for segmenting these nodules, which verifies its significant feature-learning capability.

Fig. 6 shows the segmentation results of CDP-ResNet for juxtapleural nodules and small nodules cases from the LIDC testing set. This comparison shows that our segmentation results are largely overlapping with ground truth.

4.2. Characteristics and limitations

The segmentation method we proposed is unique in terms of network architecture and sampling methods. On the network architecture, we extract the local texture information and richer context information of the nodule by designing the dual path structure based on the residual mechanism. At the same time, to extract the multi-scale and multi-view features of the nodules, we cascaded the image blocks of two different scales with multiple views.

In the sampling method, to further improve the robustness of the model, we improve the weighted sampling strategy proposed by Wang et al. [25]. In other words, we use the nodule boundary voxel as the benchmark to fully sample the small nodules, which better alleviates the imbalance of training labels.

In the clinical application, lung nodule segmentation is an important practical application. Specifically, the doctor needs to perform texture analysis and morphological feature comparison based on the segmented nodules to more accurately judge the benign and malignant nodules. Moreover, if the patient has a history test record, it can be inferred whether it is in a deterioration stage by the change in the size of the

nodule. In addition, lung nodule segmentation is also an important step in the lung cancer CAD system and a prerequisite for judging its benign and malignant. In other words, the CAD system usually determines whether the texture features and morphological features meet the characteristics of malignant nodules based on the results of the nodule segmentation. Hence, the proposed algorithm can assist the doctor in the diagnosis of benign and malignant.

However, our segmentation method also has certain limitations. Specifically, our segmentation method is a semi-automatic segmentation method. Although it is not necessary to specify the location of the nodule in all layers, it is still necessary to give the ROI of a certain layer where the nodule is located. In doing so, our segmentation method can automatically segment the lung nodules based on this layer.

5. Conclusion

In this work, we proposed a CDP-ResNet model for lung nodule segmentation to extract features by cascading two dual-pathway residual networks. In addition, we compared our proposed method with the existing methods for lung nodule segmentation. As can be seen from Table 9, the CDP-ResNet model showed excellent performance in terms of segmentation accuracy with 81.58% DSC. In particular, our method can successfully segment challenging cases as shown in Table 9. Meanwhile, we compared the results of CDP-ResNet with the inter-radiologists consistency on LIDC testing set. The results show that CDP-ResNet not only performs slightly better than radiologists in terms of DICE scores but also has a slight advantage in terms of stability as shown in Table 3.

In the future work, we plan to develop an algorithm for the lung nodule detection, and then combine it with our method to implement a fully automated system for the lung nodule segmentation. Moreover, we intend to integrate the cascading architecture into the FCN network to speed up the network training and prediction process.

Acknowledgements

This work is supported by the National Key R&D Program of China (Grant No. 2017YFC0112804). The authors acknowledge the National Cancer Institute and the Foundation for the National Institutes of Health and their critical role in the creation of the free publicly available LIDC-IDRI Database used in this study.

References

[1] Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer 2015;136:E359–86. https://doi.org/10.1002/ijc.29210.

[2] Bhavanishankar K, Sudhamani MV. Techniques for detection of solitary pulmonary nodules in human lung and their classifications -a Survey. Int J Cybern Inf 2015;4:27–40. https://doi.org/10.5121/ijci.2015.4103.

[3] Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. CA Cancer J Clin 2018;68:7–30. https://doi.org/10.3322/caac.21442.

[4] Aerts HJWL, Velazquez ER, Leijenaar RTH, Parmar C, Grossmann P, Carvalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 2014;5:4006. https://doi.org/10.1038/ncomms5006.

[5] Reeves AP, Chan AB, Yankelevitz DF, Henschke CI, Kressler B, Kostis WJ. On measuring the change in size of pulmonary nodules. IEEE Trans Med Imaging 2006;25:435–50. https://doi.org/10.1109/TMI.2006.871548.

[6] MacMahon H, Austin JHM, Gamsu G, Herold CJ, Jett JR, Naidich DP, et al. Guidelines for management of small pulmonary nodules detected on CT Scans: a statement from the fleischner society. Radiology 2005;237:395–400. https://doi.org/10.1148/radiol.2372041887.

[7] Kostis WJ, Reeves AP, Yankelevitz DF, Henschke CI. Three-dimensional segmentation and growth-rate estimation of small pulmonary nodules in helical CT images. IEEE Trans Med Imaging 2003;22:1259–74. https://doi.org/10.1109/TMI.2003.817785.

[8] Gonçalves L, Novo J, Campilho A. Hessian based approaches for 3D lung nodule segmentation. Expert Syst Appl 2016;61:1–15. https://doi.org/10.1016/j.eswa.2016.05.024.

[9] Sargent D, Park SY. Semi-automatic 3D lung nodule segmentation in CT using dynamic programming. Proc SPIE 2017:10133.

[10] Kubota T, Jerebko AK, Dewan M, Salganicoff M, Krishnan A. Segmentation of pulmonary nodules of various densities with morphological approaches and convexity models. Med Image Anal 2011;15:133–54. https://doi.org/10.1016/j.media.2010.08.005.

[11] Dehmeshki J, Amin H, Valdivieso M, Ye X. Segmentation of pulmonary nodules in thoracic CT scans: a region growing approach. IEEE Trans Med Imaging 2008;27:467–80. https://doi.org/10.1109/TMI.2007.907555.

[12] Farag AA, Munim HEAE, Graham JH, Farag AA. A novel approach for lung nodules segmentation in chest CT using level sets. IEEE Trans Image Process 2013;22:5202–13. https://doi.org/10.1109/TIP.2013.2282899.

[13] Ye X, Beddoe G, Slabaugh GG. Automatic graph cut segmentation of lesions in CT using mean shift superpixels. Int J Biomed Imaging 2010;2010:983963:1–983963:14. https://doi.org/10.1155/2010/983963.

[14] Chan TF, Vese LA. Active contours without edges. IEEE Trans Image Process 2001;10:266–77. https://doi.org/10.1109/83.902291.

[15] Nithila EE, Kumar SS. Segmentation of lung nodule in CT data using active contour model and Fuzzy C-mean clustering. Alexandria Eng J 2016;55:2583–8. https://doi.org/10.1016/j.aej.2016.06.002.

[16] Wang J, Guo H. Automatic approach for lung segmentation with juxta-pleural nodules from thoracic ct based on contour tracing and correction. Comput Math Methods Med 2016;2016:2962047https://doi.org/10.1155/2016/2962047.

[17] Mukherjee S, Huang X, Bhagalia RR. Lung nodule segmentation using deep learned prior based graph cut. In 2017 IEEE 14th Int. Symp. Biomed. Imaging (ISBI 2017), 2017, p. 1205–8. doi:10.1109/ISBI.2017.7950733.

[18] Gui L, Li C, Yang X. Medical image segmentation based on level set and isoperimetric constraint. Phys Med 2017;42:162–73. https://doi.org/10.1016/j.ejmp.2017.09.123.

[19] Lu L, Barbu A, Wolf M, Liang J, Salganicoff M, Comaniciu D. Accurate polyp segmentation for 3D CT colongraphy using multi-staged probabilistic binary learning and compositional model. In 2008 IEEE Conf. Comput. Vis. Pattern Recognit., 2008, p. 1–8. doi:10.1109/CVPR.2008.4587423.

[20] Lu L, Bi J, Wolf M, Salganicoff M. Effective 3D object detection and regression using probabilistic segmentation features in CT images. CVPR 2011;2011:1049–56. https://doi.org/10.1109/CVPR.2011.5995359.

[21] Mukhopadhyay S. A segmentation framework of pulmonary nodules in lung CT images. J Digit Imaging 2016;29:86–103. https://doi.org/10.1007/s10278-015-9801-9.

[22] Shen S, Bui AAT, Cong J, Hsu W. An automated lung segmentation approach using bidirectional chain codes to improve nodule detection accuracy. Comput Biol Med 2015;57:139–49. https://doi.org/10.1016/j.compbiomed.2014.12.008.

[23] Wu D, Lu L, Bi J, Shinagawa Y, Boyer K, Krishnan A, et al. Stratified learning of local anatomical context for lung nodules in CT images. In 2010 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., 2010, p. 2791–8. doi:10.1109/CVPR.2010.5540008.

[24] Messay T, Hardie RC, Tuinstra TR. Segmentation of pulmonary nodules in computed tomography using a regression neural network approach and its application to the lung image database consortium and image database resource initiative dataset. Med Image Anal 2015;22:48–62. https://doi.org/10.1016/j.media.2015.02.002.

[25] Wang S, Zhou M, Liu Z, Liu Z, Gu D, Zang Y, et al. Central focused convolutional neural networks: Developing a data-driven model for lung nodule segmentation. Med Image Anal 2017;40:172–83. https://doi.org/10.1016/j.media.2017.06.014.

[26] Wang S, Zhou M, Gevaert O, Tang Z, Dong D, Liu Z, et al. A multi-view deep convolutional neural networks for lung nodule segmentation. In 2017 39th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., 2017, p. 1752–5. doi:10.1109/EMBC.2017.8037182.

[27] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. In 2015 IEEE Conf. Comput. Vis. Pattern Recognit., 2015, p. 3431–40. doi:10.1109/CVPR.2015.7298965.

[28] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation BT – medical image computing and computer-assisted intervention – MICCAI. Cham: Springer International Publishing; 2015. p. 234–41.

[29] Milletari F, Navab N, Ahmadi SA. V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. In 2016 Fourth Int. Conf. 3D Vis., 2016, p. 565–71. doi:10.1109/3DV.2016.79.

[30] He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. In 2016 IEEE Conf. Comput. Vis. Pattern Recognit., 2016, p. 770–8. doi:10.1109/CVPR.2016.90.

[31] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. BT – Proceedings of the 32nd International Conference on Machine Learning. Lille, France: ICML 2015; 2015. p. 448–56.

[32] He K, Zhang X, Ren S, Sun J. Delving deep into rectifiers: surpassing human-level performance on ImageNet classification. In 2015 IEEE Int. Conf. Comput. Vis., 2015, p. 1026–34. doi:10.1109/ICCV.2015.123.

[33] Wang L, Yang Y, Min R, Chakradhar S. Accelerating deep neural network training with inconsistent stochastic gradient descent. Neural Networks 2017;93:219–29. https://doi.org/10.1016/j.neunet.2017.06.003.

[34] Caruana R, Lawrence S, Giles CL. Overfitting in Neural Nets: Backpropagation, Conjugate Gradient, and Early Stopping. In BT – Advances in Neural Information Processing Systems 13, Papers from Neural Information Processing Systems (NIPS) 2000, Denver, CO, USA 2000:402–8.

[35] Uijlings JRR, van de Sande KEA, Gevers T, Smeulders AWM. Selective search for object recognition. Int J Comput Vis 2013;104:154–71. https://doi.org/10.1007/s11263-013-0620-5.

[36] Armato SG, McLennan G, Bidaut L, McNitt-Gray MF, Meyer CR, Reeves AP, et al.

The lung image database consortium (lidc) and image database resource initiative (idri): a completed reference database of lung nodules on CT scans. Med Phys 2011;38:915–31. https://doi.org/10.1118/1.3528204.
[37] Setio AAA, Ciompi F, Litjens G, Gerke P, Jacobs C, van Riel SJ, et al. Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks. IEEE Trans Med Imaging 2016;35:1160–9. https://doi.org/10.1109/TMI.2016.2536809.
[38] Setio AAA, Traverso A, de Bel T, Berens MSN, van den Bogaard C, Cerello P, et al. Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: the LUNA16 challenge. Med Image Anal 2017;42:1–13. https://doi.org/10.1016/j.media.2017.06.015.
[39] Havaei M, Davy A, Warde-Farley D, Biard A, Courville AC, Bengio Y, et al. Brain tumor segmentation with deep neural networks. Med Image Anal 2017;35:18–31. https://doi.org/10.1016/j.media.2016.05.004.
[40] Valverde S, Oliver A, Roura E, González-Villà S, Pareto D, Vilanova JC, et al. Automated tissue segmentation of MR brain images in the presence of white matter lesions. Med Image Anal 2017;35:446–57. https://doi.org/10.1016/j.media.2016.08.014.
[41] Gao Y, Shao Y, Lian J, Wang AZ, Chen RC, Shen D. Accurate segmentation of ct male pelvic organs via regression-based deformable models and multi-task random forests. IEEE Trans Med Imaging 2016;35:1532–43. https://doi.org/10.1109/TMI.2016.2519264.
[42] Shen W, Zhou M, Yang F, Yu D, Dong D, Yang C, et al. multi-crop convolutional neural networks for lung nodule malignancy suspiciousness classification. Pattern Recognit 2017;61:663–73. https://doi.org/10.1016/j.patcog.2016.05.029.
[43] Kang G, Liu K, Hou B, Zhang N. 3D multi-view convolutional neural networks for lung nodule classification. PLoS One 2017;12. https://doi.org/10.1371/journal.pone.0188290.
[44] Sun W, Zheng B, Qian W. Automatic feature learning using multichannel ROI based on deep structured algorithms for computerized lung cancer diagnosis. Comput Biol Med 2017;89:530–9. https://doi.org/10.1016/j.compbiomed.2017.04.006.