Journal of Computers Vol. 32 No. 4, 2021, pp. 152-165
doi:10.53106/199115992021083204012

Lung Fields Segmentation Based on Shape Compactness in Chest X-Ray Images

Yuqin Li, Ke Zhang, Weili Shi, Zhengang Jiang*

Department of Computer Science and Technology, Changchun University of Science and Technology, Changchun 130022, China
lyqcust@outlook.com, zhangke@cust.edu.cn,shiweili@cust.edu.cn, jiangzhengang@cust.edu.cn

Received 7 January 2021; Revised 30 April 2021; Accepted 4 June 2021

Abstract. Computer-Aided Diagnosis (CAD) benefits from its early diagnosis and accurate treatment. As the preprocessing step of CAD-based chest radiograph analysis, lung segmentation affects the precision of lesion recognition and classification. With the development of artificial intelligent technologies, a lot of powerful algorithms based on machine learning, such as convolutional neural networks, are used to extract lung areas from X-ray images. However, these state-of-the-art segmentation algorithms have become inapplicable with limited training data, varied boundaries and poor contrasts. In order to overcome these problems, this paper proposes a novel lung segmentation method which integrates Graph-cut and neural network. Different from traditional methods, the proposed method is designed with an energy function which involves a shape compactness, and the conditional probabilities are calculated according to the outputs of U-Net. Furthermore, the objective function is transformed into an iterative form and decomposed into a series of easier sub-problems based on ADMM algorithm, which is used to reduce the complexity of high-order optimization. Compared with the previous methods on JSRT dataset, the segmentation results of the proposed method show a higher Dice-Coefficient. By using the proposed method, we can achieve 97.1% accuracy compared to 94.87% using the baseline U-Net model, and the segmentation accuracy of each image in JSRT dataset is improved.

Keywords: CAD, Dice-Coefficient, lung fields segmentation, shape compactness, U-Net

1 Introduction

Pulmonary diseases are a major reason for death and hospitalization around the world. The best solution for lung disease is early diagnosis and timely treatment. Chest radiograph analysis is widely used for detection and diagnosis of lung diseases, especially in the CAD technology. In CAD system, reliable lung segmentation is a prerequisite step in automatically analyzing chest X-ray images [1-3] to ensure that lung diseases detection is not confounded by regions outside lung [4].

However, accurate lung fields segmentation is a great challenge for several reasons. Firstly, the lung fields in X-Ray images include some superimposed regions and strong edges structures (e.g. clavicles, ribs), which cause blurred boundaries; Secondly, there are large variations in lung anatomical shapes due to heart dimensions or other pathologies among different patients; Thirdly, although deep learning has achieved better performance in the field of segmentation, the results are always unsatisfied due to the network structure and limited training dataset. Therefore, how to effectively further improve the accuracy of the model on the basis of deep learning is a problem that needs to be solved.

Many researchers have presented a lot of methods for lung fields segmentation. These methods can be roughly divided into some categories as follows [5-6].

Rule-based methods generally consist of a sequence of rules and steps. Some operations, such as morphologies and thresholds [1, 7-10], will be influenced greatly by the quality of medical images.

* Corresponding Author

Shape model-based methods are often used in segmentation tasks, which have also been applied to lung fields segmentation [11-13]. Ginneken optimized the method of active shape model (ASM) to segment lung fields [14]; Besbes [15] employed a graph-based shape model and Shi [16] used an adaptive shape prior; Sun got a rough initial segmentation of lung border by a 3D ASM matching [17]. However, the segmentation results rely on the initial models heavily, and is also ineffective with abnormal cases.

Atlas-based methods employ labelled databases as anatomical atlas, and use nonrigid registration methods to align the objective images with atlas [6]. However, atlas-based methods rely on the accuracy of nonrigid registration, which is time-consuming and inefficient.

Pixel-based methods regard the segmentation as a classification task and thus acquire a classifier to label each pixel as ROI or background [14]. However, these methods often make wrong classification around the boundaries.

Convolutional neural network-based methods are widely applied with its excellent ability of feature extraction and expression [18-21]. These methods don't need to extract image features manually or preprocess the images excessively. However, it requires a large number of labeled data for training to extract the ROIs of object images. The varied shapes and imaging quality also influence the results.

The traditional edge detection methods such as Canny edge detectors [22] or other derivatives-based methods are close to other irrelevant contours and not continuous. Recently, some methods were proposed for boundaries detection such as structured edge (SE) detector [23], DeepEdge [24], holistically-nested edge detection (HED) [25], and neural network-based methods [26]. However, these methods are sensitive to noises, and the amplitudes of the first and second derivatives are always larger, which will lead to wrong detection results. The results of SEDUCM [27] have outperformed other methods [28], but it needs to predefine the number of features corresponding to various segmentation tasks. Furthermore, the parameters of the dedicated chest X-Ray lung segmentation with handcrafted features are required carefully tuning, and the generalization ability and automation ability are poor.

Graph theory methods take advantages of its plasticity of the energy function and global optimality. These methods are used to obtain optimal object segmentation. In particular, the methods represented by graph cut have been developed rapidly in image segmentation. Ali [29] employed graph-cut and iterative Markov-Gibbs-random-field (MGRF) for lung fields segmentation; Hua [30] used a graph-based search of a cost function which integrated intensity, boundary smoothness, gradient, and rib information; Dai [31] proposed a new algorithm based on Gaussian mixture and an improved graph cuts for lung segmentation on CT images. In recent years, some shape constraints such as compactness [32-34], convexity [35, 36], tubularity [37] have been widely used in energy function to fulfil the segmentation task. However, generally speaking, the methods based on graph theory are more complex and always lead to complex high-order optimization problems.

Hybrid methods display efficient performance by fusing some advanced techniques [1, 14, 38-39]. Hao [40] employed fully convolutional network (FCN) and graph-cut for esophagus tumor segmentation. FCN is a deep architecture, which can improve the performance of segmentation. However, the training error rate is high in deep plain network because the gradient disappears easily in a deeper architecture [41]. U-Net has been developed based on FCN. The U-Net method combines the features from shallow and deep layers through multipath confusion effectively, which solves the spatial loss of feature map and improves the performance of semantic of segmentation. Liu [42] used improved U-Net and Graph-Cut for liver CT sequence segmentation on LiTS dataset. This method constructed the energy function based on the adjacent slices according to its types of context information. Ullah [43] used ResNet to construct boundary-level information, and then combined this information with boundary energy term of graph-cut framework. However, this method was proposed based on user interaction, which increases the burden for hospitals and inexperienced doctors. In fact, adding some shape constraints in Graph-cut is beneficial for a wide range of image segmentation application, especially when facing the problem of limited training data, varied boundaries and poor contrasts. So, this paper employs a hybrid method that integrates graph theory with neural network to segment lung fields automatically, and uses the idea of optimization method [32] to obtain the results. As a result, this paper shows an automated, high-performance and precise algorithm for lung segmentation. The main contributions of this paper can be summarized as follows:

(1) A hybrid model is developed which integrates graph-cut and U-Net network to segment the lung fields automatically.

(2) A shape compactness prior is adopted to the energy function, which is beneficial to the varied boundaries and poor contrasts in medical images. Then the ADMM algorithm is employed to cope with the problem of high-order optimization.

(3) A baseline U-Net model is designed in this work, and its outputs are regarded as the conditional probabilities. It is different from the conventional semi-automated ways when calculating data (or regional) term. The proposed method not only overcomes the problem of limited data, but also improves the segmentation performance of neural network.

The structure of the rest of this paper is organized as follows: Section 2 provides the overall proposed framework and its calculation process. Section 3 presents the experimental studies and discussions. Finally, the paper gives a concise conclusion in Section 4.

2 Overall Framework

Segmenting an object from background can be expressed as a binary labelling problem, which is always solved by energy minimization. In this case, every pixel in images needs to be assigned a label, and the binary labelling problem corresponding to the minimum energy is selected as the solution. This paper constrains the segmentation energy function with a shape compactness, which can be solved efficiently with alternating direction method of multipliers (ADMM). This work employs the outputs of U-Net to calculate conditional probabilities.

2.1 U-Net

In general, the U-Net model is regarded as a kind of FCN. It is called U-Net because of its elegant U type symmetrical structure. The U-Net network concatenates the features extracted from shallow and deep layers by skipping connections. It consists of down-sampling and up-sampling. The down-sampling structure extracts the feature information of input images, and transmits the information layer by layer through continuous operations. The up-sampling structure expands the resolution of the feature maps through deconvolution operations until the resolution of the images is fully restored. The down-sampling and up-sampling also can be regarded as encoder and decoder structures respectively. The U-Net model is suitable for image segmentation. Currently, most of the popular segmentation methods are based on U-Net and its extensions. This work adopts the U-Net as a representative to calculate the conditional probability.

The U-Net network designed in this paper is shown in Fig. 1, this architecture includes 22 convolutional layers, 4 max-pooling layers, 4 up-sample layers and 4 catenation connections. The orange boxes correspond to multi-channel feature maps, and blue boxes represent copied feature maps. The numbers at the top of the boxes represent the number of channels. Arrows with different colors indicate different operations, and the arrow with purple indicates that this operation consists of 1 up-sampling layer and 2 convolution layers.

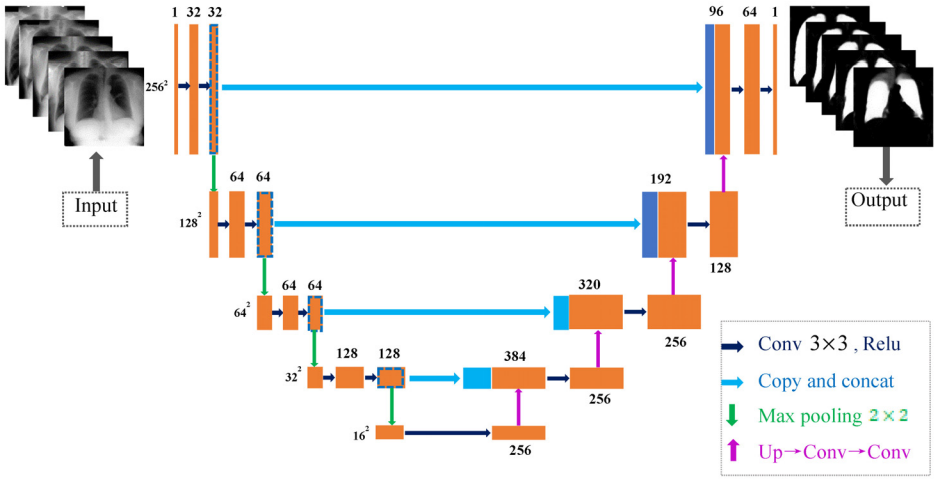


Fig. 1. U-Net architecture used in this paper. Arrows with different colors indicate different operations.

2.2 Construction of Energy Function

Let $I(p):\Omega\subset\mathbb{R}^2\to\mathbb{R}$ be an image function which is defined over an image domain Ω. The $\boldsymbol{x}_p\in\mathbb{R}^K$ is the input-feature vector of pixel p. In order to have a clearer presentation, we concentrate on a binary statement which the label set $\ell=\{0,1\}$, where 0 and 1 represent background and object in the image, respectively. The task of segmentation is that each pixel p should be assigned a label $u_p\in\ell$. As a result, an optimal region in Ω with a point set will be found. Meanwhile, it is necessary to ensured that the optimal region must obey some appearance and appearance priors. For this research, we construct the energy function as follows. Then this energy function needs to be minimized to find the optimal solution.

$$\mathrm{E}(\boldsymbol{u})=A(\boldsymbol{u})+\lambda B(\boldsymbol{u}). \tag{1}$$

2.3 Data (or Regional) Term Formulation

The first term $A(\boldsymbol{u})$ in Equation 1 is called data (or regional) term because it incorporates regional constraints. The data (or regional) term measures the performance of the model that how well the pixels fit into the object or background. In Equation 1, $\boldsymbol{u}:\Omega\to\{0,1\}$ is a binary function, $A(\boldsymbol{u})$ can be formulated as the log-likelihoods appearance:

$$A(\boldsymbol{u})=\int_\Omega u_p\log\frac{P(u_p=0\,|\,\boldsymbol{x}_p)}{P(u_p=1\,|\,\boldsymbol{x}_p)}. \tag{2}$$

In order to facilitate the subsequent calculation, let the term of log as v_p. Therefore the $A(\boldsymbol{u})$ can be written in discrete form: $A(\boldsymbol{u})=\sum_p u_p v_p$. It is well known that any techniques can be used to obtain conditional probabilities. Different from conventional semi-automatic methods, this paper designs a baseline U-Net network as a representative to calculate conditional probabilities. Then the outputs of U-Net are applied to learn $P(u_p\,|\,\boldsymbol{x}_p)$, the architecture of U-Net is shown in section 2.1.

2.4 Shape Compactness Term Formulation

The second term $B(\boldsymbol{u})$ in Equation 1 is called shape compactness, which can be formulated as the ratio of length-squared to area with a discrete form:

$$B(\boldsymbol{u})=\frac{l(\boldsymbol{u})^2}{a(\boldsymbol{u})^2}. \tag{3}$$

The area can be represented as $a(\boldsymbol{u})=\sum_p u_p$, and the length is proportional to the number of the adjacent pixels with different labels assigned, that is: $l(\boldsymbol{u})\propto\sum_{p,q}w_{pq}(u_p-u_q)^2$, where $w_{pq}=1$ if p and q are neighbors, otherwise $w_{pq}=0$. Based on the theory of Laplacian, it has already been proved that $\sum_{i,j=1}^{m}w_{i,j}(y_i-y_j)^2=2y^TLy$, where L is the Laplacian matrix corresponding to the weights $w_{i,j}$. In this study, the theory is applied in Equation (3). Then we redefine the length term as $l(\boldsymbol{u})=\boldsymbol{u}^TL\boldsymbol{u}$. The energy function becomes:

$$\mathrm{E}(\boldsymbol{u})=\boldsymbol{v}^T\boldsymbol{u}+\lambda\frac{(\boldsymbol{u}^TL\boldsymbol{u})^2}{\mathbf{1}^T\boldsymbol{u}}. \tag{4}$$

where $\mathbf{1}$ is a vector with each element = 1. And the segmentation model becomes:

$$\arg\min_{\boldsymbol{u}\in\{0,1\}^{|\Omega|}}\mathrm{E}(\boldsymbol{u})=\boldsymbol{v}^T\boldsymbol{u}+\lambda\frac{(\boldsymbol{u}^TL\boldsymbol{u})^2}{\mathbf{1}^T\boldsymbol{u}}. \tag{5}$$

2.5 Energy Function Optimization

In the next step, this paper employs the ADMM algorithm to our model. The benefit of this form is that we can solving each variable by fixing all other variables.

General ADMM algorithm. Generally speaking, ADMM is suitable for the problem of convex optimization, the standard ADMM algorithm is formulated as Equation 6:

$$\min f(x)+g(x),\ \text{ s.t. } Ax+By=c. \tag{6}$$

where f and g are convex functions.

The augmented Lagrangian of this convex problem can be formulated as follows:

$$L_p(x,y,\lambda)=f(x)+g(y)+\lambda^T(Ax+By-c)+\frac{\rho}{2}\left\|Ax+By-c\right\|_2^2,\ \rho>0. \tag{7}$$

The ADMM splits the problem into some smaller subproblems which have to be solved alternatingly. The optimization processes are given by following steps. For k = 0,⋯, repeat until a convergence is reached.

$$\begin{aligned} x_{k+1}&=\arg\min_x L_p(x,y_k,\lambda_k),\\ y_{k+1}&=\arg\min_y L_p(x_{k+1},y,\lambda_k),\\ \lambda_{k+1}&=\lambda_k+\rho(Ax_{k+1}+By_{k+1}-c). \end{aligned} \tag{8}$$

ADMM for our problem. To apply the ADMM method to our work. Firstly, we also formulate the Equation 5 as an augmented Lagrangian form:

$$\arg\min_{\boldsymbol{u},\boldsymbol{r},t}\boldsymbol{v}^T\boldsymbol{u}+\frac{\lambda}{t}(\boldsymbol{u}^TL\boldsymbol{u})(\boldsymbol{r}^TL\boldsymbol{r}),\ \text{s.t. } \boldsymbol{u}=\boldsymbol{r}\ \text{ and }\ t=\mathbf{1}^T. \tag{9}$$

where $\boldsymbol{r}\in\mathbb{R}^{|\Omega|}$ and $t\in\mathbb{R}_+$ are two auxiliary variables. Then, the model can be constructed as an iterable form via augmented Lagrangian:

$$\arg\min_{\substack{\boldsymbol{u},\boldsymbol{r},t\\ \theta_1,\theta_2}}\boldsymbol{v}^T\boldsymbol{u}+\frac{\lambda}{t}(\boldsymbol{u}^TL\boldsymbol{u})(\boldsymbol{r}^TL\boldsymbol{r})+\frac{\rho_1}{2}\left\|\boldsymbol{u}-\boldsymbol{r}+\theta_1\right\|^2+\frac{\rho_2}{2}\left\|t-\mathbf{1}^T\boldsymbol{r}+\theta_2\right\|^2. \tag{10}$$

where θ_1 and θ_2 are dual variables, ρ_1 and ρ_2 are penalty parameters. Then we employ an iterative optimization method to solve this problem, with each variable is updated orderly until convergence.

Updating r. Fix the parameters $\boldsymbol{u}$ and t, only the terms with variable $\boldsymbol{r}$ are considered. Let $\alpha=\frac{\lambda}{t}(\boldsymbol{u}^TL\boldsymbol{u})$. The problem can be formulated as:

$$\arg\min_{\boldsymbol{r}\in\mathbb{R}^{|\Omega|}}\alpha\boldsymbol{r}^TL\boldsymbol{r}+\frac{\rho_1}{2}\left\|\boldsymbol{r}-(\boldsymbol{u}+\theta_1)\right\|^2+\frac{\rho_2}{2}(\mathbf{1}^T\boldsymbol{r}-(t+\theta_2))^2. \tag{11}$$

To minimize this convex quadratic problem, we compute the partial derivative about Equation 11 with respect to $\boldsymbol{r}$ and make the result = 0. Then apply the Woodbury matrix identity: $(A+UCU)^{-1}=A^{-1}-A^{-1}U(C^{-1}+VA^{-1}U)^{-1}VA^{-1}$ to Equation 11. For further derivation, let $\beta=\alpha L+\rho_1 I$, and in fact, $\beta^{-1}\mathbf{1}=\frac{1}{\rho_1}\mathbf{1}$.

$$(\alpha L+\rho_1 I+\rho_2\mathbf{1}\mathbf{1}^T)^{-1}=\beta^{-1}-\frac{1}{\rho_1}(\frac{\rho_1}{\rho_2}+|\Omega|)^{-1}\mathbf{1}\mathbf{1}^T. \tag{12}$$

Because β is very sparse, we can use the method of preconditioned conjugate gradient to solve the result of Equation 12.

Updating t. Fix the parameters $\boldsymbol{u}$ and $\boldsymbol{r}$, updating t is equivalent to solving the following problem:

$$\arg\min_{t\geq 0} \frac{\gamma}{t} + \frac{\rho_2}{2}(t - \mathbf{1}^T \boldsymbol{r} + \theta_2)^2 . \tag{13}$$

where $\gamma = \lambda(\boldsymbol{u}^T L\boldsymbol{u})(\boldsymbol{r}^T L\boldsymbol{r})$.

Let $g(t) = \frac{\gamma}{t} + \frac{\rho_2}{2}(t - \mathbf{1}^T \boldsymbol{r} + \theta_2)^2$. We compute the partial derivative of $g(t)$ with respect to *t* and make the result = 0.

$$\frac{\partial g(t)}{\partial t} = -\frac{\gamma}{t^2} + \rho_2(t - \mathbf{1}^T \boldsymbol{r} + \theta_2) = 0 . \tag{14}$$

Equation 14 is solved as follows:

$$t^3 - t^2(\mathbf{1}^T r - \theta_2) = \frac{\gamma}{\rho_2} . \tag{15}$$

It is well known that: $\gamma = \lambda(\boldsymbol{u}^T L\boldsymbol{u})(\boldsymbol{r}^T L\boldsymbol{r}) > 0$ and $\rho_2 > 0$. Therefore, we can derive the follows inequality relations: $t^3 - t^2(\mathbf{1}^T \boldsymbol{r} - \theta_2) > 0,\ t > \mathbf{1}^T \boldsymbol{r} - \theta_2$. Then the cube roots are analyzed and solved.

Updating ***u***. Get rid of the terms that are not related to ***u*** and regard the irrelevant variables as a constant, the Equation 10 can be rewritten as:

$$\arg\min_{\boldsymbol{u}\in\{0,1\}^{|\Omega|}} \boldsymbol{v}^T \boldsymbol{u} + \frac{\lambda}{t}(\boldsymbol{u}^T L\boldsymbol{u})(\boldsymbol{r}^T L\boldsymbol{r}) + \frac{\rho_1}{2}\left\|\boldsymbol{u} - \boldsymbol{r} + \theta_1\right\|^2 . \tag{16}$$

Let $\eta = \frac{\lambda}{t}(\boldsymbol{r}^T L\boldsymbol{r})$, $\delta = \boldsymbol{r} - \theta_1$, and the Equation 16 = $b(\boldsymbol{u})$. Hence, the problem can be expressed as following formula:

$$b(\boldsymbol{u}) = \boldsymbol{v}^T \boldsymbol{u} + \eta(\boldsymbol{u}^T L\boldsymbol{u}) + \frac{\rho_1}{2}\left\|\boldsymbol{u} - \delta\right\|^2 . \tag{17}$$

where $\boldsymbol{u}: \Omega \to \{0,1\}$ is binary function, so it's obvious that: $(u_p - u_q)^2 = |u_p - u_q|$.

It has been proven that: $f^T L f = \frac{1}{2}\sum_{i,j=1}^{n} w_{ij} \bullet (f_i - f_j)^2$, where *L* is the Laplacian matrix. Therefore, Equation 17 is derived as follows:

$$b(\boldsymbol{u}) = \sum_{p\in\Omega}(v_p + \rho_1(\frac{1}{2} - \delta_p))\ u_p + \delta_p^2 + \frac{1}{2}\eta\sum_{p,q\in\Omega} w_{p,q}\left|u_p - u_q\right| . \tag{18}$$

Equation 18 can be converted to the following problem:

$$\arg\min_{\boldsymbol{u}\in\{0,1\}^{|\Omega|}} = \sum_{p\in\Omega}(v_p + \rho_1(\frac{1}{2} - \delta_p))\ u_p + \eta\sum_{p,q\in\Omega} w_{p,q}\left|u_p - u_q\right| . \tag{19}$$

Equation 19 is a typical graph-cut problem, which can be solved by using the algorithm of Boykov-Kolmogorov. Finally, update the two dual variables on the basis of standard ADMM algorithm.

3 Experiments and Discussion

3.1 Dataset and Evaluation Indicator

A publicly available Japanese Society of Radiological Technology (JSRT) dataset [44] was adopted to test the proposed method. The dataset consists of 247 PA chest radiographs from 13 institutions in Japan and one in the United States. The images were scanned from films to the size of 2048*2048 pixels, with a spatial resolution of 0.175 mm and 12-bit gray levels. The dataset has been subdivided into two folds: fold 1 consists of 124 images, and fold 2 consists of 123 images. The manual segmentation results of lung fields for JSRT are available at the website: (http://www.isi.uu.nl/Research/Databases/SCR/).

The Dice coefficient index (DSC) has been currently considered the most popular similarity measurement. This research used DSC to evaluate the performance of the proposed method. The formular is as follows:

$$\mathrm{DSC} = \frac{2 \times |\mathrm{S} \cap \mathrm{T}|}{|\mathrm{S}| + |\mathrm{T}|}. \tag{20}$$

where S is the lung parenchyma area obtained by algorithms, and T is the ground-truth mask. DSC is the overlap between the estimated segmentation mask and the ground-truth mask. The values of the DSC are range from 0 to 1, where 1 denotes fully overlapped segmentation results.

3.2 Model Parameters Setting

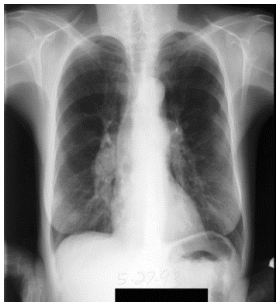
In the energy function, we set $\rho_1 = 40$ and $\rho_2 = 1500$. To facilitate the algorithm convergence, this paper increased these two parameters by 1% at each iteration. Each optimal parameter of compactness regularization λ was obtained by traversal search.

For the designed U-Net network, the batch size in this model was set at 8. To obtain the specific information for each region from the images, this architecture involved a receptive field with size of 3*3, and stride of 1. In order to solve the problem of lack of data in Deep Learning, data enhancement methods had been used to preprocess the data. This paper adopted the ImageDataGenerator module built-in the Keras framework to augment the dataset. Rotation, cropping, scaling, offset changes were applied to generate new images. By adjusting the parameters of the experiment, the network was iterated 100 times, and a training model was saved every ten times.

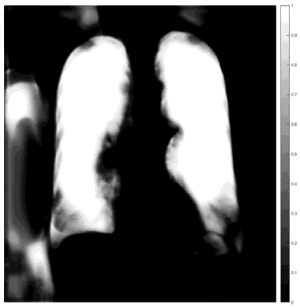
3.3 Experimental and Discussion

Training. The segmentation performance of the proposed method was based on the two-fold cross-validation, in which 20% of the train images were reserved for validating. To further verify the effectiveness of the proposed model, we adjusted the resolution without any compression to 256*256 and 512*512, and conducted our experiment respectively. A technique of data augmentation was employed to generate new images to extend the dataset, the parameters setting was as described in 3.2. The epoch was set at 100. In the U-Net baseline model, we used the Rectified Linear Unit (ReLU) as the activate function in the intermediate layers, and the sigmoid was employed in the last convolutional layer. Besides, the Adam optimizer was included to optimize the loss function.

Experiment results with its baseline U-Net model. For this research, a representative baseline U-Net network was applied to calculate the conditional probability. Fig. 2(a) presented an image selected randomly from the JSRT dataset, and the corresponding estimated lung field outputted by U-Net network was shown in Fig. 2(b). Moreover, Fig. 3 displayed some visualization results of the intermediate layer of our U-Net network in this research.



(a) An image is randomly selected from JSRT dataset



(b) the conditional probabilities map by U-Net network. Each pixel's intensity represents the probability of the pixel being part of the lung field

Fig. 2.

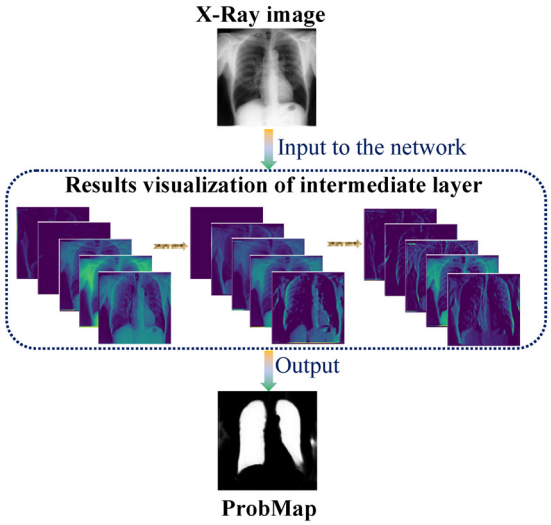
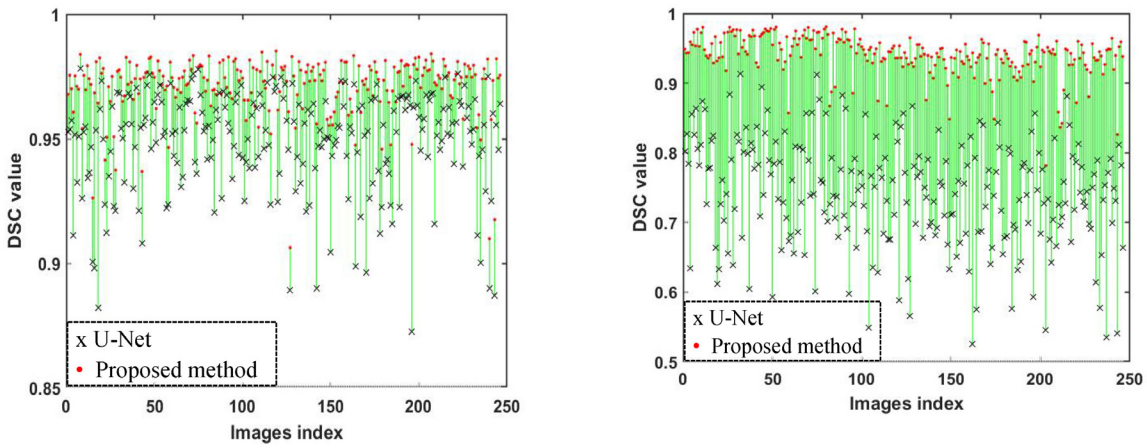


Fig. 3. Some visualization results of intermediate layer of U-Net network

To calculate the conditional probability in Equation 1, a baseline U-Net model was utilized. As a result, a sequence of training models was saved in the stage of training. In order to have a clear comparison, the highest accuracy and a lower verification model were selected to test the other subset images. At this point, two groups of test results were obtained based on the test subset, which corresponded to two different training models. Next, we swapped training and test subsets. By using the same process as above, we also got two groups of test results of another subset. Therefore, we obtained two groups of segmentation results on the whole dataset, corresponding to different training models. In this paper, we adjusted the size of images without compressing the pixel value of the images to 256*256 and 512*512. Training with the 1024*1024 imaging resolution is also possible but it requires multiple GPUs or needs to reduce the number of feature maps in the Original U-Net network.

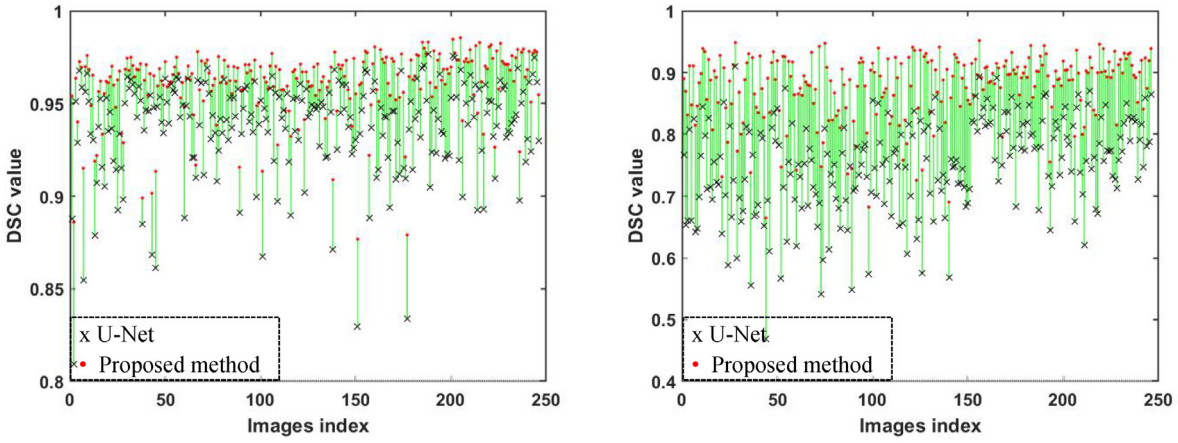
Firstly, the images with 256*256 were input in the experiment. After training, the highest accuracy training model was selected to test the other images. Then swapped the subsets and the segmentation results were obtained in the same way, in which the mean DSC of U-Net was 0.9487. These outputs were used to calculate the conditional probabilities in the proposed method, and the mean DSC was 0.9710. Fig. 4 showed the performance of U-Net and the proposed method in JSRT dataset. For comparison, a training model with lower accuracy was chosen to test the other images. The mean DSC of U-Net was 0.7412. These results were used as the conditional probabilities for the proposed method, and the proposed method achieved 0.9400.



(a) the highest accuracy training model (b) a lower accuracy training model

Fig. 4. Segmentation performance of U-Net and the proposed method for each image whose size is 256*256 in JSRT dataset with (a) the highest accuracy training model and (b) a lower accuracy training model.

Secondly, the images with 512*512 were input in another experiment, and the same experimental setup is as the first experiment. The highest accuracy training model and a lower were also selected. The segmentation performance for each image was shown in Fig. 5. Table 1 to Table 2 showed the mean DSC of segmentation results from different methods. Table 1 was the mean performance of U-Net and the proposed method with 256*256 image resolution. Result 1 gave the highest accuracy training model for experiment; Result 2 was a lower training model. Table 2 exhibited the results for 512*512 image resolution. From Fig. 4 to Fig. 5, we can conclude that the accuracy of the segmentation was increased by the proposed method for every X-ray image.



(a) (b)

Fig. 5. Segmentation performance of U-Net and the proposed method for each image whose size is 512*512 in JSRT dataset with the highest accuracy training model (a) and a lower accuracy training model

As mentioned earlier, we did four comparative experiments with different image resolutions. By using the baseline U-Net model, it predicted a probability score for each pixel in an image. This paper used the outputs of the U-Net to calculate the conditional probability in objective energy function. Fig. 4 to Fig. 5 showed the segmentation performance for each sample in the JSRT dataset. Each subfigure displayed the change from the baseline U-Net to the proposed model. As we can see, all these experiments produced higher performance for every X-Ray image compared with the baseline U-Net. Besides, it is worth mentioning that even if a lower training model was selected, the final results can still maintain better performance. But it has to be said that the segmentation performance will be improved when the accuracy of the conditional probability increases. The U-Net model is superior to others due to its simplicity and fewer training parameters. The modern segmentation methods are mostly based on U-Net and its extensions. In these experiments, we only used the U-Net as a representative to calculate the conditional probability.

Experiments with a smaller training dataset. We did an experiment to verify the effect of the size of training data for segmentation model. Considering the completeness of the JSRT dataset, only the number of images in training subset was reduced to half of the original U-Net network, and keeps other parameters or architectures unchanged. In the same way, a two-fold cross-validation method was employed to evaluate the results of segmentation. By swapping training and test subsets, the segmentation results of whole JSRT dataset were obtained. Next, the two groups of U-Net outputs were used as conditional probabilities directly. Therefore, two groups of different results corresponding to different size of training data were acquired. Fig. 6 showed the segmentation performance with different methods.

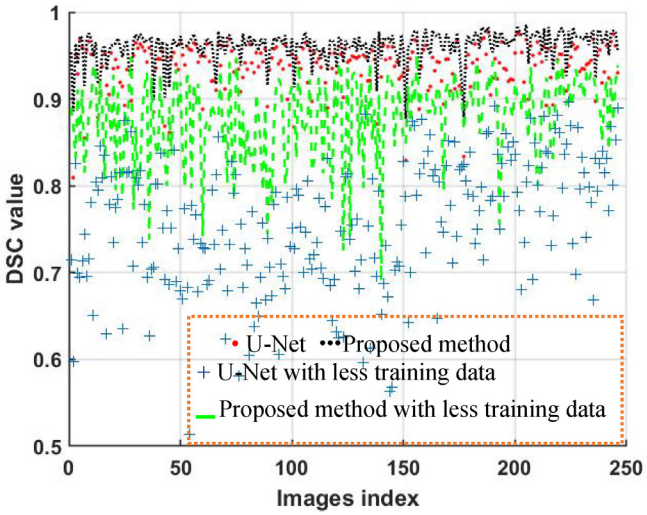


Fig. 6. Segmentation performance with different methods for each image in JSRT dataset

From Fig. 4 to Fig. 6, we can conclude that the amount of training data had a crucial impact on the model of deep neural network. However, the proposed method not only overcame the problem of limited data, but also improved the segmentation performance.

Table 1. Mean DSC of segmentation Results of 256*256 image resolution. The result 1 denotes the highest training model which is inputted to the experiment, and the Result 2 is a lower model

Method	Result1 (Mean DSC)	Result2 (Mean DSC)
U-Net	0.9487	0.7412
Proposed method	0.9710	0.9400

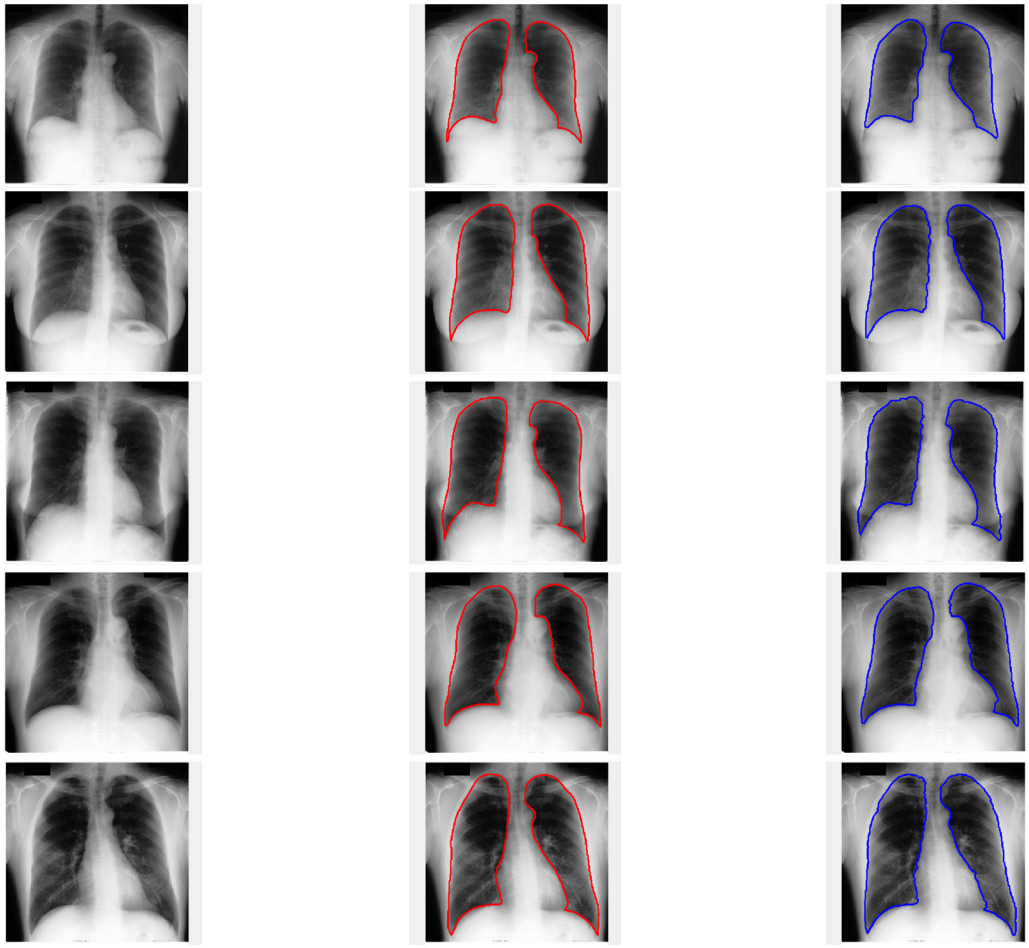
Table 2. Mean DSC of segmentation Results of 512*512 image resolution. The Result1 denotes the highest training model which is inputted to the experiment, and the Result2 is a lower model

Method	Result1 (Mean DSC)	Result2 (Mean DSC)
U-Net	0.9374	0.7501
Proposed method	0.9610	0.8770

Comparison with state-of-the-art methods. To compare the performance of the proposed model with some representative segmentation method. Table 3 summarized the corresponding results on JSRT dataset. The results of traditional methods such as thresholding and clustering [10] were unsatisfactory. These approaches were rule-based methods, which were influenced greatly by the quality of imaging protocol. SIFT-Flow [5] achieved a better result, but its nature of nonrigid registration methods to align the objective images was unsatisfactory. The single U-Net network achieved 0.9487. Fig. 7 shows some segmentation results with proposed method on the JSRT dataset.

Table 3. Performance of different methods for lung field segmentation in term of Dice coefficient index (DSC)

Method	Mean Dice coefficient index (DSC) (%)
K_Means Clustering [10]	67.5
Otsu's Thresholding [10]	55.56
SIFT-Flow [5]	96.7±0.8
Single U-Net	94.87
CRF as post-processing	95.3
Proposed method	97.1±1.4



(a) Initial images (b) Ground-truth (c) The proposed method

Fig. 7. Some segmentation results with proposed method on the JSRT dataset. Red and blue contours indicate ground truth and automatic segmentation results, respectively

4 Conclusions

In this paper, a robust algorithm was presented for lung field segmentation. An energy function model with shape compactness was constructed for the task of lung fields segmentation. In the energy function, the data term can be represented as a log-likelihoods appearance, and the shape compactness term can be formulated as the ratio of length-squared to area with a discrete form. Because a squared sum of pairwise potentials was involved in the energy function, it will lead to a complex high-order optimization problem. In order to meet this challenge, the objective energy function was transformed into an iterative form and split it into several sub-problems which can be solved easily based on the theory of the ADMM. In this research, the outputs of the U-Net network were employed to learn conditional probabilities, which improve the robustness and automation of the model.

To verify the effectiveness of the proposed method, a publicly available JSRT dataset with different image resolution was employed to evaluate the performance. For the sake of proving the robustness of the proposed method, the training models with the highest and a lower verification accuracy in U-Net were used to calculate conditional probability respectively. From the segmentation and evaluation, the proposed method achieved a higher accuracy and outperformed the state-of-the-art on JSRT dataset. A limitation of the study is the selection of optimal hyperparameters. Besides, further experiments are needed on some larger datasets for a more reliable estimation about the performance of the method. We intend to make our model more robust, universal, automatic and accurate in the future.

Acknowledgements

This work is supported by the local science and technology development funds guided by central government: Basic Research Project of Jilin Province (No. 202002045JC).

References

[1] B.v. Ginneken, B.M.t.H. Romeny, Automatic segmentation of lung fields in chest radiographs, Medical Physics 27(10)(2000) 2445-2455.

[2] T. Xu, M. Mandal, R. Long et al., An edge-region force guided active shape approach for automatic lung field detection in chest radiographs, Computerized Medical Imaging and Graphics 36(6)(2012) 452-463.

[3] M.S. Brown, L.S. Wilson, B.D. Doust et al., Knowledge-based method for segmentation and analysis of lung boundaries in chest X-ray images, Computerized Medical Imaging and Graphics 22(6)(1998) 463-477.

[4] A.M.R. Schilham, B.V. Ginneken, M. Loog, A computer-aided diagnosis system for detection of lung nodules in chest radiographs with an evaluation on a public database, Medical Image Analysis 10(2)(2006) 247-258.

[5] S. Candemir, S. Jaeger, K. Palaniappan et al., Lung segmentation in chest radiographs using anatomical atlases with nonrigid registration, IEEE Transaction on Medical Imaging 33(2)(2014) 577-90.

[6] S. Candemir, S. Antani, A review on lung boundary detection in chest X-rays, International Journal of Computer Assisted Radiology and Surgery 14(4)(2019) 563-576.

[7] W.J. Choi, T.S. Choi, Automated pulmonary nodule detection based on three-dimensional shape-based feature descriptor, Computer Methods and Programs in Biomedicine 113(1)(2014) 37-54.

[8] J. Dehmeshki, X. Ye, X. Lin et al., Automated detection of lung nodules in CT images using shape-based genetic algorithm, Computerized Medical Imaging and Graphics 31(6)(2007) 408-417.

[9] Q. Gao, S. Wang, D. Zhao et al., Accurate lung segmentation for X-ray CT images, The Third International Conference on Natural Computation (ICNC 2007), 2007.

[10] N.R. Kasu, C. Saravanan, Segmentation on Chest Radiographs Using Otsu's and K-Means Clustering Methods, 2018 International Conference on Inventive Research in Computing Applications (ICIRCA), 2018.

[11] B.V. Ginneken, S. Katsuragawa, B.M. ter Haar Romeny et al., Automatic detection of abnormalities in chest radiographs using local texture analysis, IEEE Transactions on Medical Imaging 2(2)(2002) 139-149.

[12] A. Dawoud, Lung segmentation in chest radiographs by fusing shape information in iterative thresholding, IET Computer Vision 5(3)(2011) 185-190.

[13] A. Soliman, F. Khalifa, A. Elnakib et al, Accurate Lungs Segmentation on CT Chest Images by Adaptive Appearance-Guided Shape Modeling, IEEE Transactions on Medical Imaging 36(1)(2017) 263-276.

[14] B.V. Ginneken, M.B. Stegmann, M. Loog, Segmentation of anatomical structures in chest radiographs using supervised methods: a comparative study on a public database, Medical Image Analysis 10(1)(2006) 19-40.

[15] A. Besbes, N. Paragios, Landmark-based segmentation of lungs while handling partial correspondences using sparse graph-based priors, 2011 IEEE International Symposium on Biomedical Imaging, 2011.

[16] Y. Shi, F. Qi, Z. Xue et al., Segmenting lung fields in serial chest radiographs using both population-based and patient-specific shape statistics, IEEE Transactions on Medical Imaging 27(4)(2008) 481-494.

[17] S. Sun, C. Bauer, R. Beichel, Automated 3-D segmentation of lungs with lung cancer in CT data using a novel robust active shape model approach, IEEE Transactions on Medical Imaging 31(2)(2012) 449-460.

[18] A.A. Novikov, D. Lenis, D. Majoret et al., Fully Convolutional Architectures for Multi-Class Segmentation in Chest Radiographs, IEEE Transactions on Medical Imaging 37(8)(2018) 1865-1876.

[19] A.P. Harrison, Z. Xu, K. George et al., Progressive and Multi-path Holistically Nested Neural Networks for Pathological Lung Segmentation from CT Images, International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2017.

[20] M. Negahdar, D. Beymer, T.S. Mahmood, Automated volumetric lung segmentation of thoracic CT images using fully convolutional neural network, SPIE Medical Imaging, 2018.

[21] B.A. Skourt, A.E. Hassani, A. Majda, Lung CT Image Segmentation Using Deep Neural Networks, The First International Conference on Intelligent Computing in Data Sciences, 2018.

[22] J. Canny, A Computational Approach to Edge Detection, IEEE Transactions on Pattern Analysis and Machine Intelligence 8(6)(1986) 679-698.

[23] P. Dollár, C. L. Zitnick, Fast Edge Detection Using Structured Forests, IEEE Transactions on Pattern Analysis and Machine Intelligence 37(8)(2015) 1558-1570.

[24] G. Bertasius, J. Shi, and L. Torresani, DeepEdge: A multi-scale bifurcated deep network for top-down contour detection, The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015.

[25] S. Xie, Z. Tu, Holistically-Nested Edge Detection, The IEEE International Conference on Computer Vision (ICCV), 2015.

[26] O. Tsujii, M.T. Freedman, S.K. Mun, Lung contour detection in chest radiographs using 1-D convolution neural networks, Journal of Electronic Imaging 8(1)(1999) 46-53.

[27] W. Yang, Y. Liu, L. Lin et al. Lung Field Segmentation in Chest Radiographs from Boundary Maps by a Structured Edge Detector, IEEE Journal of Biomedical and Health Informatics 22(3)(2017) 842-851.

[28] L. Bi, D. Feng, J. Kim, Dual-Path Adversarial Learning for Fully Convolutional Network (FCN)-Based Medical Image Segmentation, The Visual Computer 34(6-8)(2018) 1043-1052.

[29] A.M. Ali, A.E. Baz, A.A. Farag, A novel framework for accurate lung segmentation using graph cuts, 2007 4th IEEE International Symposium on Biomedical Imaging, 2007.

[30] P. Hua, Q. Song, M. Sonka et al., Segmentation of pathological and diseased lung tissue in CT images using a graph-search algorithm, 2011 IEEE International Symposium on Biomedical Imaging, 2011.

[31] S. Dai, K. Lu, J. Dong et al., A novel approach of lung segmentation on chest CT images using graph cuts, Neurocomputing 168(2015) 799-807.

[32] J. Dolz, I. B. Ayed, C. Desrosiers, Unbiased Shape Compactness for Segmentation, International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2017.

[33] I.B. Ayed, M. Wang, B. Miles et al., TRIC: Trust region for invariant compactness and its application to abdominal aorta segmentation, International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2014.

[34] P. Das, O. Veksler, V. Zavadsky et al., Semi-automatic segmentation with compact shape prior, Image and Vision Computing 27(1-2)(2009), 206-219.

[35] L. Gorelick, O. Veksler, Y. Boykov et al., Convexity shape prior for binary segmentation, IEEE Transactions on Pattern Analysis and Machine Intelligence 39(2)(2017) 258-271.

[36] O. Veksler, Star shape prior for graph-cut image segmentation, European Conference on Computer Vision (ECCV), 2008.

[37] Y. Kitamura, Y. Li, W. Ito et al., Coronary lumen and plaque segmentation from cta using higher-order shape prior, International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2014.

[38] G. Coppini, M. Miniati, S. Monti et al., A computer-aided diagnosis approach for emphysema recognition in chest radiography, Medical Engineering Physics 35(1)(2013) 63-73.

[39] S. Jaeger, A. Karargyris, S. Antani et al., Detecting tuberculosis in radiographs using combined lung masks, 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2012.

[40] Z. Hao, J. Liu, and J. Liu, Esophagus tumor segmentation using fully convolutional neural network and graph cut, Proceedings of 2017 Chinese Intelligent Systems Conference, 2017.

[41] S. Chen, H. Yang, J. Fu et al., U-Net Plus: Deep Semantic Segmentation for Esophagus and Esophageal Cancer in Computed Tomography Images, IEEE Access (2019) 82867-82877.

[42] Z. Liu, Y. Song, V.S. Sheng et al., Liver CT sequence segmentation based with improved U-Net and Graph-cut, Expert Systems with Applications 126(2019) 54-63.

[43] M. Ullah, A. Iltaf, Q. Hou et al, A foreground extraction approach using convolutional neural network with Graph cut, 2018 IEEE 3rd International Conference on Image, Vision and Computing (ICIVC), 2018.

[44] J. Shiraishi, S. Katsuragawa, J. Ikezoe et al., Development of a digital image database for chest radiographs with and without a lung nodule: receiver operating characteristic analysis of radiologists' detection of pulmonary nodules, American Journal of Roentgenology 174(1)(2000) 71-74.