



# Liver Tumor Segmentation of CT Image by Using Deep Fully Convolutional Network

Lingmin Jin<sup>1</sup>, Rui Ma<sup>1</sup>, Meng Zhao<sup>1</sup>, Shenghua Teng<sup>1</sup>(✉),  
and Zuoyong Li<sup>2</sup>(✉)

<sup>1</sup> College of Electronic and Information Engineering,  
Shandong University of Science and Technology, Qingdao 266590, China  
lingmin\_jin@163.com, shteng@sdust.edu.cn

<sup>2</sup> Fujian Provincial Key Laboratory of Information Processing and Intelligent  
Control, College of Computer and Control Engineering, Minjiang University,  
Fuzhou 350121, China  
fzulzytdq@126.com

**Abstract.** Accurate segmentation of liver tumors is an important guarantee for the success of liver cancer surgery, where convolutional network has been a type of popular method. However, the performance of the traditional convolutional network is limited by the network depth. To improve the accuracy of liver tumor segmentation, we propose a cascaded deep fully convolutional network (DFCN) which uses ResNet as the basis network followed by side output layer in the upsampling stage to fuse multi-scale image features. For better localizing the liver tumors, the segmentation result is further refined by a fully connected conditional random field. Experimental results show that the proposed method achieves higher segmentation accuracy than several state-of-the-art methods.

**Keywords:** Liver tumor segmentation · Deep learning · Fully convolutional network · Fully connected conditional random field

## 1 Introduction

Accurate segmentation of liver tumors is an important guarantee for the success of liver cancer surgery. CT imaging technology is a common way for doctors to diagnose liver cancer. Comparing with other medical imaging technologies, CT images have the characteristics of clear imaging and high signal-to-noise ratio, playing an important role in the diagnosis and treatment of liver diseases. In the actual clinical diagnosis, doctors are required to manually segment liver tumors on CT images. The segmentation process requires a lot of time and effort, and the segmentation results are greatly affected by human subjectivity. To alleviate this issue, researchers have developed some computer-aided methods for liver tumor segmentation. These methods can be divided into three categories, namely traditional methods [1–4], machine learning-based methods [5, 6], and deep learning-based methods [7–11].

Traditional segmentation methods mainly include thresholding [1], region-based growth [2], level set and active contours [3, 4], and so on. Traditional methods mainly use manual extraction of features, resulting in inaccurate segmentation effects,

especially the target of medical images to be segmented may not have obvious contours, etc. Machine learning based segmentation usually consist of feature extraction and classification or regression. Massoptier et al. [5] first segmented the liver in the CT image by the dynamic contour method, and then segmented the tumor on the liver using the K-means clustering method. Shi et al. [6] used the AdaBoost score Similar algorithm realizes the automatic segmentation of liver tumors. Machine learning-based methods need to manually set the features to be extracted and the features used could heavily affect the segmentation results, which are still affected by subjective experience and prior knowledge, and the segmentation efficiency needs to be improved.

Deep learning methods are widely used in the field of medical image processing. For example, Guo et al. [7] proposed ALEXNet [8] liver tumor segmentation model based on the FCN structure. Ronneberger et al. [9] proposed U-Net for medical image segmentation. Christ et al. [10] proposed a method of segmenting liver tumors with a cascaded fully convolutional network. He [11] proposed that ResNet can directly pass the input information to the next layer by introducing the identity mapping structure to form a residual network. With the advent of ResNet, more and more deep convolutional networks are used in the field of medical image processing.

After exploring existing liver tumor segmentation methods based on convolutional neural networks, we can find that they still have two limitations. Firstly, the shallow convolutional neural network is plagued by the problem of deep network degradation, and the ability of the network to extract features is limited. Secondly, the output of the high-level convolutional layer tends to lose part of the detailed information like the number of layers increases and the pooling process, resulting in a rougher feature map obtained by upsampling, which affects the accuracy of liver tumor segmentation. To alleviate these issues, we proposed a modified method for liver tumor segmentation based on DFCN. Experimental results showed that the DFCN model had better feature expression capability and more generalization performance, which improved the segmentation accuracy of the model.

## 2 Method

To solve the problem of accurate segmentation of liver tumors, we proposed an improved method of liver tumor segmentation. First of all, this method overcomes the problem of deep network degradation and improves the rough expression of the fuzzy segmentation results of the fully convolutional network, and then introduces a balanced loss function to train the network. Finally, the fully connected conditional random field is used to optimize the liver tumor segmentation results of DFCN.

### 2.1 DFCN Segmentation Model

ResNet with the fully connected layer removed is used as the basic network of DFCN, where ResNet is formed by stacking residual units and has 24 layers in total. Each residual unit is composed of two convolutional layers with BN layers. The basic network is divided into 5 convolutional stages with the pooling layer as the demarcation point. The scale of the feature map generated at each stage is different: from

shallow to deep, the original image size, 1/2 original image size, 1/4 original image size, 1/8 original image size, and 1/16 original image size. A side output layer is connected at the end of each convolution stage, and each side output layer is responsible for supervising the feature map generated by the convolution stage. The side output layer is composed of a convolution layer with a convolution kernel size of  $3 \times 3$  and an output channel of 16 and a deconvolution layer. The deconvolution layer is responsible for upsampling feature maps of different scales to the original size. The feature maps with different scale information generated by each side output layer are superimposed and input into the fusion layer. The fusion layer linearly fuses the features of each scale through a convolution layer with a convolution kernel size of  $1 \times 1$ . Finally, the fused result is sent to the classifier as the output of DFCN for classification.

In this paper, when using DFCN for CT image liver tumor segmentation, we find that the receptive field in the first convolution stage is small, and it is easy to extract local image noise, which affects the entire tumor segmentation, so we only use feature maps for the last four convolution stages. In this paper, inspired by the structure of the cascaded full convolution network proposed by Christ et al. [10], the DFCN of the cascaded structure is designed. As shown in Fig. 1, two DFCNs with the same structure are trained to segment the liver and tumor respectively. The first DFCN focuses on segmenting the liver from the CT slices of the abdomen, and then the liver ROI was cut out from the original image through the liver segmentation results, and the second DFCN focuses on segmenting liver tumors from the liver ROI.

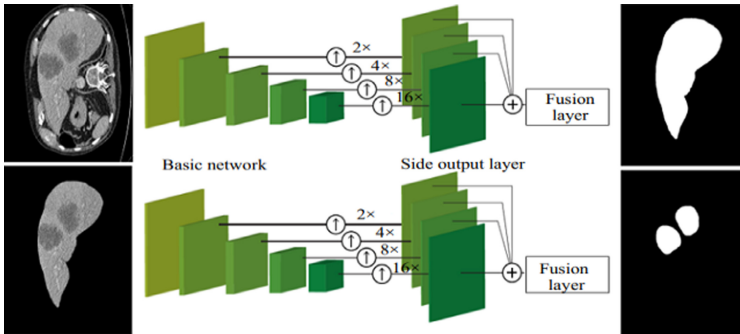


Fig. 1. Cascaded liver tumor segmentation network.

## 2.2 Training of DFCN Network

This article introduces the cost-sensitive loss objective function [12] in the process of calculating the network loss function, the loss function generated by all the side output layers in the DFCN network is:

$$L_{side}(W, w) = \sum_{m=1}^M \alpha_m l_{side}^{(m)}(W, w^{(m)}). \quad (1)$$

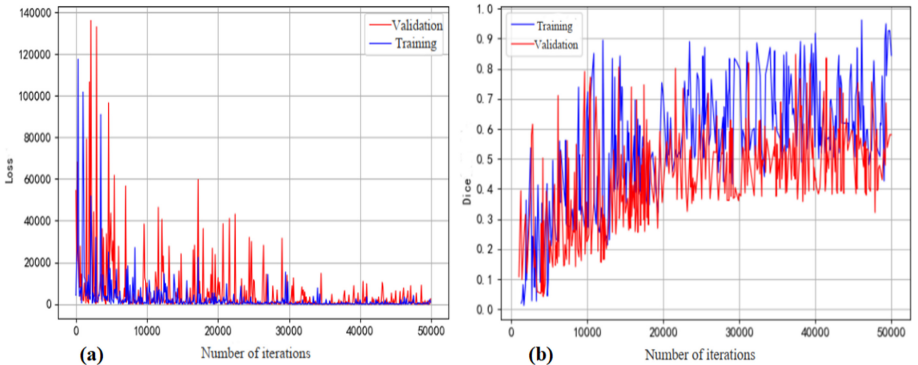
Because of the imbalance of the number of positive and negative sample pixels, this paper introduces the balance parameter  $\beta$  according to the cost-sensitive method. For each side output layer, the specific result of the loss function is:

$$L_{side}(W, w^{(m)}) = -\beta \sum_{j \in Y_+} \log \Pr(y_i = 1 | X; W, w^{(m)}) - (1 - \beta) \sum_{j \in Y_-} \log \Pr(y_j = 0 | X; W, w^{(m)}), \quad (2)$$

where  $\beta = |Y_-|/Y$ ,  $1 - \beta = |Y_+|/|Y|$ ,  $|Y_-|$  represents the set of positive sample label pixels, and  $|Y_+|$  represents the set of negative sample label pixels. The loss function of the network includes two parts: the loss function  $L_{side}(W, w)$  generated by the output layer on all sides and the loss function  $L_{side}(W, w, h)$  generated when the fusion layer predicts the final segmentation result, where  $h$  represents the weight parameter of the fusion layer.

In this paper, a stochastic gradient descent algorithm with momentum parameters is used to optimize the loss function of the network model. During the training process: the learning rate is set to  $10^{-7}$  and the momentum parameter is 0.9. In order to prevent overfitting, the regular term coefficient is set to 0.0002, a total of 50000 iterations. In order to visualize the training process, this paper records the Loss generated by the network segmentation tumor every 100 iterations and draws a line graph.

The Loss line chart is shown in Fig. 2(a). In the training set and verification set, not all CT slices contain tumors, the Loss generated when there is no tumor is small, so it can be seen from the figure that the Loss fluctuates locally. But as the number of iterations increases, the overall trend of the Loss gradually declined and eventually stabilized in a lower range. The Dice line chart is shown in Fig. 2(b). Similarly, when recording the Dice similarity coefficient, the Dice similarity coefficient of CT slices that do not contain tumors will be 0. When drawing a line chart, this article discards the item whose Dice similarity coefficient is 0. As the number of iterations increases, the Dice similarity coefficient gradually increases, and finally stabilizes in the training set in the range of  $70\% \pm 20\%$ , and in the validation set in the range of  $55\% \pm 20\%$ .



**Fig. 2.** Line charts during training: (a) Loss, (b) Dice.

### 2.3 FC-CRF Optimization Process

DFCN improved the roughness of segmentation results. However, it does not fully consider the relationship between pixels and lacks a priori constraints on context information, resulting in a lack of spatial consistency in segmentation results. To resolve this issue, we use fully FC-CRF [13] to further optimize the segmentation results.

The energy function  $E(x)$  in the fully connected conditional random field is:

$$E(x) = \sum_i \varphi_u(x_i) + \sum_{i \neq j} \varphi_p(x_i, x_j), \quad (3)$$

where  $\varphi_u(x_i)$  indicates the unary energy term, which represents the probability that the  $i$ -th pixel belongs to the category label  $x_i$ , and  $\varphi_p(x_i, x_j)$  represents the binary energy term, which represents the probability that the pixel points  $i$  and  $j$  belong to the labels  $x_i$  and  $x_j$  at the same time. The binary energy term considers the interaction between adjacent pixels and uses spatial context information. Its expression is:

$$\varphi_p(x_i, x_j) = \mu(x_i, x_j) \left( w^{(1)} \exp \left( -\frac{\|p_i - p_j^2\|}{2\sigma_\alpha^2} - \frac{\|I_i - I_j\|}{2\sigma_\beta^2} \right) + w^{(2)} \exp \left( -\frac{\|p_i - p_j^2\|}{2\sigma_\gamma^2} \right) \right), \quad (4)$$

$\mu(x_i, x_j) = [x_i \neq x_j]$  is the label compatibility function. When adjacent pixels are assigned different category labels,  $\mu(x_i, x_j)$  is a penalty term, and it can be understood that similar pixels tend to be classified into the same category. The parameters  $\sigma_\alpha$ ,  $\sigma_\beta$  and  $\sigma_\gamma$  are used to control the scale of the Gaussian kernel function.

The solution of FC-CRF can be transformed into the energy function minimization problem. The average field approximation algorithm proposed by Krähenbühl [13] et al. is used to reduce computational complexity. First, the pre-processed abdominal CT image is input into the DFCN to predict the probability of each pixel being classified as a tumor and output a probability map, and then connect an FC-CRF to optimize the DFCN segmentation results. The input of FC-CRF includes two parts, which are the probability map and the pre-processed CT image. The probability map provides unary potential energy, and the color and spatial position information between pixels provided by the CT image after preprocessing is used as binary potential energy. Finally, it continuously iterates through the average field approximation algorithm until the energy function value is minimum, and then outputs the liver tumor segmentation results.

### 3 Experimental Results

#### 3.1 Data Preprocessing

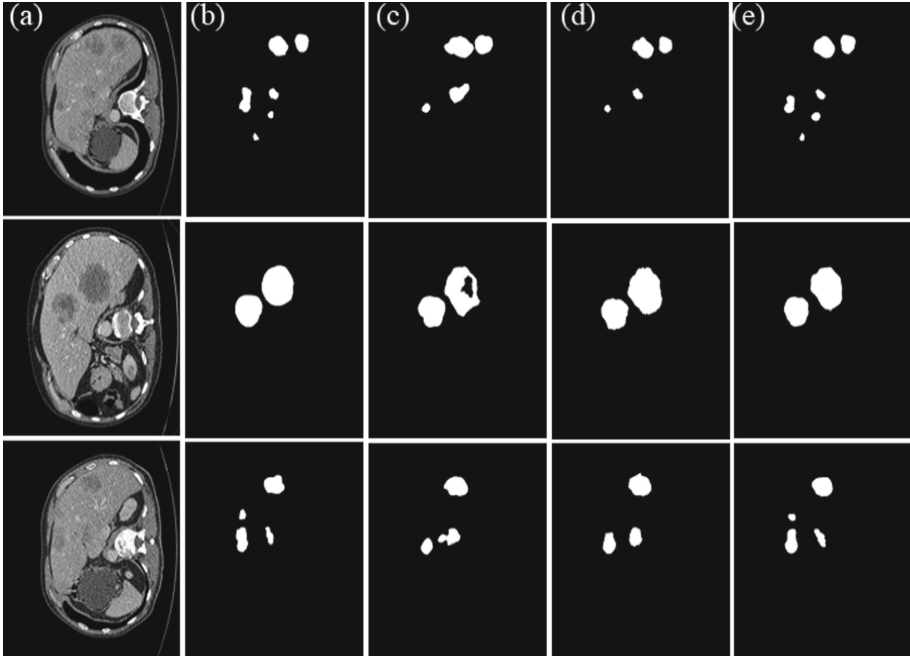
The experimental data use the data set officially provided by the CT Image Live Tumor Segmentation Challenge Competition (LiTS) [14]. Since the sponsor of LiTS does not disclose the label information of the liver and tumor of 70 patients in the data set, the data of 130 patients are used in the experiment. Among them, 100 patients' data are used for network training, 10 patients' data are used for verification, and 20 patients' data are used to test the trained network. Abdominal CT images need to be pre-processed before segmentation.

The pre-processing mainly includes: window technology processing [15], data enhancement, and normalization processing. K. Sahi et al. [16] have given that the window width of the liver is  $[-62,238]$ . To enhance the contrast of the liver in the abdominal CT image, the window of the abdominal CT is set to  $[-150,250]$  through the window technique. Since the density of liver lesions will decrease compared with normal liver tissue, the lower limit of the window is set to  $-150$  to ensure that the liver lesions are not removed. Due to the lack of CT image data, this paper uses the way of data enhancement to expand the data. To solve the comparability between different data features, this paper uses the minimum-maximum standardized processing method. It processes the pixel matrix of the image, finds the minimum value  $X_{min}$  and maximum value  $X_{max}$  in the entire pixel matrix, and then we normalize these data by using Eq. (5).  $f$  is a coefficient in Eq. (5) that can control the range of normalization. If normalized to  $[0,1]$ , the factor value is 1, normalized to  $[0,255]$ , the factor value is 255.

$$X_{norm} = f * \frac{X - X_{min}}{X_{max} - X_{min}}. \quad (5)$$

#### 3.2 Segmentation Results

To verify the superiority of the proposed liver tumor segmentation method over the counterparts including FCN [17] and DRIU [18], this article selected 30 photos in the test set for comparison. Figure 3 compares the similarities and differences between doctor marks and the results of DFCN, FCN, and DRIU. It can be seen that the segmentation result of FCN is relatively rough, the tumor contour is quite different from the tumor label image marked by the doctor, and when the tumor size is small and the grayscale is uneven, FCN cannot segment those tumors. The segmentation results of DRIU is more accurate than that of FCN, and the segmentation result is closer to the label map. However, when the tumor size is smaller and the grayscale is uneven, DRIU cannot also segment those tumors. The segmentation result of DFCN is more accurate than DRIU, and the segmentation result is closest to the tumor label map marked by the doctor. It can also be segmented if it encounters tumors with smaller grayscale unevenness and smaller size. The experimental environment is Ubuntu 16.04 + python2.7 + TensorFlow, the experimental equipment is a Dell computer with GPU, and its GPU model is TITAN X.



**Fig. 3.** Qualitative comparison of segmentation results obtained by different methods: (a) original CT image, (b) tumor label, (c) FCN [17], (d) DRIU [18], (e) DFCN.

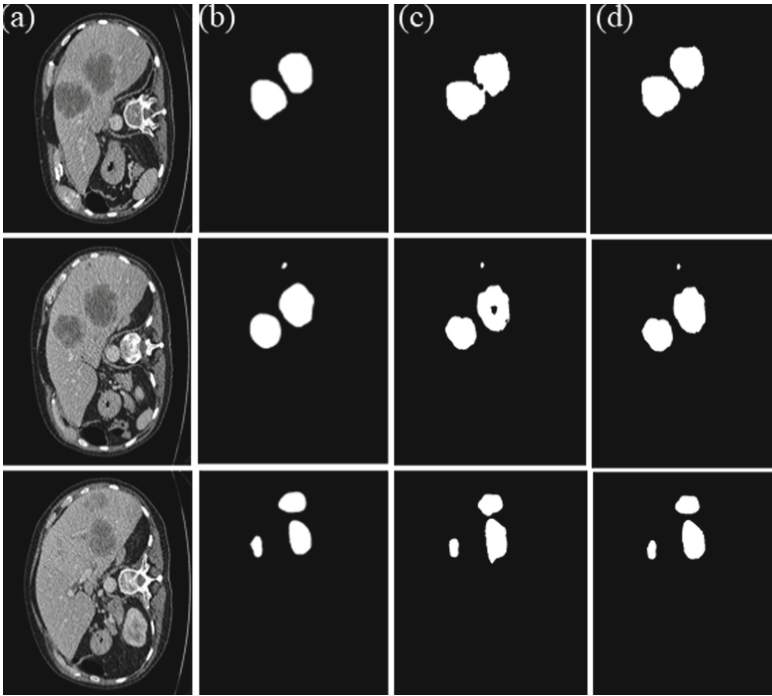
**Table 1.** Quantitative comparison on liver tumor segmentation.

| Method    | Dice   | Recall | Precision | F-measure |
|-----------|--------|--------|-----------|-----------|
| FCN [17]  | 0.7229 | 0.6919 | 0.7886    | 0.7653    |
| DRIU [18] | 0.7938 | 0.7243 | 0.8226    | 0.8057    |
| DFCN      | 0.8234 | 0.7961 | 0.8705    | 0.8520    |

Table 1 lists the quantitative comparison of segmentation results. It can be seen from Table 1 that the liver tumor segmentation effect of the DFCN model is superior to the other two deep learning segmentation methods in the Dice similarity coefficient, Recall, Precision, and F-measure [19]. Regardless of FCN or DRIU, their network layers are relatively shallow, and they cannot learn the deep semantic features of liver tumors in CT images.

### 3.3 Optimization Results

To verify the optimization effect of the FC-CRF model on the DFCN tumor segmentation results, 100 CT images are selected in the test data set for comparative experiments. Part of tumor segmentation results is shown in Fig. 4, followed by the abdominal CT map, tumor labeling map, DFCN segmentation results and FC-CRF



**Fig. 4.** FC-CRF optimization results: (a) original CT image, (b) tumor label, (c) DFCN, (d) DFCN + FC-CRF [13].

optimization results. It can be seen from Fig. 4 that the segmentation results of liver tumors optimized by FC-CRF add more detailed expression than the segmentation results before optimization, which is closer to the label diagram marked by the doctor. The experiment is conducted on the Ubuntu system, where FC-CRF is implemented using python's pydensecrf package. The densecrf interface provided by the program can solve FC-CRF using mean-field approximation algorithm.

**Table 2.** Segmentation accuracy comparison of DFCN without and with FC-CRF optimization.

| Method             | Dice   | Recall | Precision | F-measure |
|--------------------|--------|--------|-----------|-----------|
| DFCN               | 0.8431 | 0.8158 | 0.8802    | 0.8468    |
| DFCN + FC-CRF [13] | 0.8579 | 0.8318 | 0.8933    | 0.8615    |

It can be seen from Table 2 that after the DFCN liver tumor segmentation results are optimized by FC-CRF, all four indicators are improved, and the segmentation results are closer to the tumor label. Not only the prediction probability of each pixel is considered, but also the full use of the correlation between the gray value and position of all pixels in the CT image, which increases the constraints of context information, thereby improving the detailed expression and spatial consistency of the liver tumor segmentation results.



## 4 Conclusion

In this paper, we propose a liver segmentation method based on a deep fully convolutional network. The proposed method develops a cascaded network to segment liver tumors, and uses fully connected conditional random fields to further optimize the segmentation results. We qualitatively and quantitatively evaluate the proposed method on clinical data containing 30 sets of CT images. Experimental results show that the proposed method improves the accuracy of liver tumor segmentation. However, the proposed method does not use spatial information of liver tumors for the segmentation. In the future, we will develop 3D convolutional networks to solve this problem.

**Acknowledgements.** This work is partially supported by National Natural Science Foundation of China (61972187, 61772254), Fujian Provincial Leading Project (2017H0030, 2019H0025), Government Guiding Regional Science and Technology Development (2019L3009), and Natural Science Foundation of Fujian Province (2017J01768 and 2019J01756).

## References

1. Otsu, N.: A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern.* **9**(1), 62–66 (2007)
2. Song, H., Wang, Y., Huang, X., et al.: Liver CT image tumor segmentation algorithm based on dynamic adaptive region growth. *J. Beijing Inst. Technol.* **34**(1), 72–76 (2014). (in Chinese)
3. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: active contour models. *Int. J. Comput. Vision* **1**(4), 321–331 (1988)
4. Osher, S., Sethian, J.A.: Fronts propagation with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations. *J. Comput. Phys.* **79**(1), 12–49 (1988)
5. Massoptier, L., Casciaro, S.: A new fully automatic and robust algorithm for fast segmentation of liver tissue. In: *Proceedings of the MICCAI Workshop on 3D Segmentation in the Clinic: A Grand Challenge II* (2008)
6. Shimizu, A., Narihira, T., Furukawa, D., et al.: Ensemble segmentation using AdaBoost with application to liver lesion extraction from a CT volume. In: *Proceedings of the MICCAI Workshop on 3D Segmentation in the Clinic: A Grand Challenge II* (2008)
7. Guo, S., Ma, S., Li, J., et al.: Research on liver CT image segmentation based on fully convolutional neural network. *Comput. Eng. Appl.* **53**(18), 126–131 (2017). (in Chinese)
8. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. In: *International Conference on Neural Information Processing Systems*, pp. 1097–1105. Curran Associates Inc. (2012)
9. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, William M., Frangi, Alejandro F. (eds.) *MICCAI 2015. LNCS*, vol. 9351, pp. 234–241. Springer, Cham (2015). [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28)
10. Christ, P.F., Ettlinger, F., Grün, F., et al.: Automatic liver and tumor segmentation of CT and MRI volumes using cascaded fully convolutional neural networks. *Medical Image Analysis* (2017)

11. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)
12. Xie, S., Tu, Z.: Holistically-nested edge detection. In: Proceeding of IEEE International Conference on Computer Vision, vol. 1, pp. 1395–1403 (2015)
13. Krähenbühl, P., Koltum, V.: Efficient inference in fully connected CRFs with Gaussian edge potentials. In: Advances in Neural Information Processing System, pp. 109–117 (2011)
14. WWW: Web page of the Liver Tumor Segmentation Challenge. <https://competitions.codalab.org/competitions/17094>
15. Huaijun, L., Hua, Y., Pingyong, F., et al.: The value of window technology in CT diagnosis. *J. Pract. Radiol.* **2**, 109–110 (1992). (in Chinese)
16. Sahi, K., Jackson, S., Wiebe, E., et al.: The value of liver windows settings in the detection of small renal cell carcinomas on unenhanced computed tomography. *Can. Assoc. Radiol. J.* **65**, 71–76 (2014)
17. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **39**(4), 640–651 (2014)
18. Bellver, M., Maninis, K.K., Pont-Tuset, J., et al.: Detection-aided liver lesion segmentation using deep learning (2017)
19. Dice, L.R.: Measures of the amount of ecologic association between species. *Ecology* **26**(3), 297–302 (1945)