# Liver Segmentation in CT Images Using Residual U-Net and 3D CRF

**Shuntaro Nagano, Tohru Kamiya**
*Department of Mechanical and Control Engineering, Kyushu Institute of Technology, 1-1 Sensui-cho, Tobata-ku, Kitakyushu-shi, Fukuoka, Japan*

**Guanguxu Li**
*Tiangong Univercity, 399 Binshuixi Road, Xiqing District, Tianjin, China*
*E-mail: nagano.shuntaro244@mail.kyutech.jp*

**Abstract**

In recent years, the development of CAD systems aimed at reducing the burden on doctors and improving diagnostic accuracy has been promoted. In this paper, we propose a segmentation method for the liver site on abdominal CT images as a pretreatment for a CAD system using dynamic CT images. The method consists of two stages. First, we segment the liver with a model based on U-net, a segmentation model using CNN. Next, the 3D CRF (Conditional Random Field) is used to make corrections that take into account the three-dimensional characteristics of the liver to improve the accuracy of segmentation. In the experiment, the accuracy was evaluated for CT images of 20 cases.

*Keywords*: Liver, Convolutional Neural Network, Segmentation, Conditional Random Field.

## 1. Introduction

According to GLOBOCAN2020, the International Agency for Research on Cancer (IARC), an external research organization of the World Health Organization (WHO), the number of people suffering from liver cancer worldwide is about 910,000. In humans [1], it ranks 7th among cancer types (4.7% of the total). In addition, the total number of deaths due to liver cancer is about 830,000, which is the second highest among cancer types (8.3% of the total), and the mortality rate for the number of affected people is as high as 92%. Therefore, early detection and appropriate treatment are important.

Imaging diagnosis by CT, MR, etc. plays an important role in early detection of liver disease. In particular, diagnostic radiographic imaging such as dynamic contrast enhanced CT provides useful information for the differential diagnosis of the liver lesions [2]. However, since image diagnosis is performed by subjective judgment based on many years of experience of doctors, there is a problem that the diagnosis results by doctors vary. In addition, since dynamic CT takes multiple images, the number of images taken increases, which raises the burden on doctors. For this reason, research is being conducted on Computer Aided Diagnosis (CAD) systems aimed at reducing the burden on doctors and improving diagnostic accuracy.

In general, the CAD system require many technical issues such as extraction of region of interest (ROI), image alignment, identification of abnormal shadows, and diagnosis. In this paper, we focus on the extraction of ROI from a CT images.

The accuracy of area extraction in medical images is important because it affects the accuracy of alignment and identification performed later. In the extraction of the liver region, there are level set method [3] and active shape model method [4]. However, in these extraction methods, omission of extraction of lesions has become a problem. On the other hand, some Convolutional Neural Network (CNN) approaches [5, 6] are introduced for the extraction of ROI. Therefore, in this paper we use the CNN, which has produced good results in the field of image segmentation, to extract the liver region including lesions on CT images. In the experiment, we performed our proposed method to 20 CT cases and evaluated our performance.

## 2. Methods

The flow of the proposed method consists of two stages: segmentation using CNN and correction by 3D Conditional Random Field (3DCRF) [7].

### 2.1. *Segmentation with CNN*

In this paper, we propose a network structure based on Residual U-Net [8], which is a segmentation model by CNN, Residual Block [9] is changed to ResNet-D [10], and added Spatial and Channel Squeeze and Excitation (scSE) [11] to the end of each block of encoder and decoder.

#### 2.1.1. *Residual U-Net*

Residual U-Net is a CNN model used in the field of segmentation that combines U-Net [12] and residual block. Similar to U-Net, the encoder extracts the image features and the decoder restores them. In addition, by connecting a shortcut from the encoder to the decoder, it is possible to restore with the decoder while retaining the local information and position information of the image.

The difference from U-Net is the introduction of residual block in the encoder part and decoder part. With the introduction of residual block, feature extraction becomes possible while suppressing deterioration of image information.

Various structures have been proposed for the residual block, however the proposed method uses residual U-Net 15 with full pre-activation [13].

#### 2.1.2. *ResNet-D*

ResNet-D is a model in which an average pooling layer is added to the shortcut connection of ResNet's residual block. ResNet's residual block causes a 3/4 loss of information in shortcut connection when down sampling. Therefore, by adding an average pooling layer to the shortcut connection, the omission of this information is prevented.

#### 2.1.3. *Spatial and Channel Squeeze and Excitation (scSE)*

Spatial and channel squeeze and excitation is a method proposed to improve the performance of CNN models that perform segmentation. This is an extension of the Squeeze-and-Excitation block (cSE) [14] used in the field of image classification. In cSE block, the entire

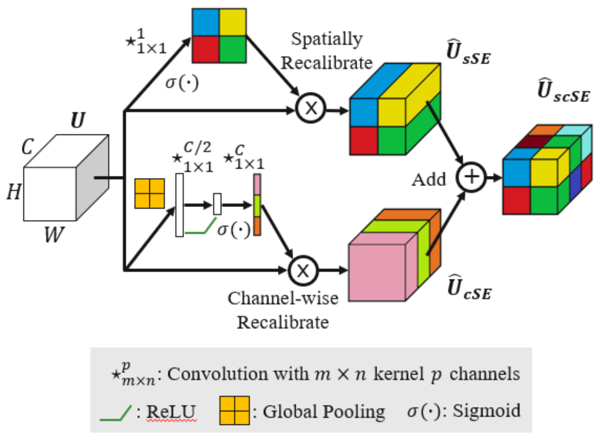



Fig. 1. Outline of scSE Block

image is squeeze and the excitation for each channel is calculated, but in segmentation, the information between pixels cannot be enhanced well. Therefore, channel squeeze and spatial excitation was proposed. Contrary to cSE block, sSE block is squeeze in the channel direction and excitation is calculated for each pixel, so that the output takes into account the relationship between pixels. scSE is a combination of both sSE and cSE. The details of scSE are shown in Fig. 1.

### 2.2. *3D Conditional random Field (3DCRF)*

The CT image is 3D information, but the segmentation model of the proposed method is performed for 2D. In other words, the liver region is extracted for each slice of the CT image. This is because there is a problem due to the number of data sets and GPU specifications for learning in 3D. Therefore, in order to make corrections that take into account the three-dimensional characteristics of the liver, 3DCRF is used in this paper.

The CRF is considered on a complete graph $G = (V, E)$ with the vertices $i \in V$ of each pixel in the image and the edges $e_{ij} \in E = \{(i,j)\forall i,j \in V \text{s.t.} i < j\}$ between the vertices. In addition, the variable vector $x \in L^N$ is used as the label for each vertex, and the labeling is optimized for each vertex by minimizing the following energy function.

$$E(x) = \sum_{i \in V} \phi_i(x_i) + \sum_{(i,j)\in E} \phi_{ij}(x_i, x_j) \qquad (1)$$

$$\phi_i(x_i) = -logP(x_i|I) \qquad (2)$$

$$\phi_{ij}(x_i, x_j) = \mu\left(w_{pos}\exp\left(-\frac{|p_i - p_j|^2}{2\sigma_{pos}^2}\right)\right.$$
$$\left. + w_{bil}exp\left(-\frac{\left|p_i - p_j\right|^2}{2\sigma_{bil}^2} - \frac{|I_i - I_j|^2}{2\sigma_{int}^2}\right)\right) \quad (3)$$

Here, $P(x_i|I)$ is the likelihood of each pixel to the class obtained from the segmentation model of the proposed method. In addition, $\mu(x_i, x_j) = 1(x_i \neq x_j)$ , and $\phi_{ij}(x_i, x_j)$ represents the strength of the bond between pixels, and represents the cost considering the color information and distance information of the image. Also, $|p_i - p_j|$ is the spatial distance between pixels $i$ and $j$, and $|I_i - I_j|$ is the difference in the color information of the image. Each element in this equation can be adjusted by the weights $w_{pos}$, $w_{bil}$ and the variances $\sigma_{pos}$, $\sigma_{bil}$, $\sigma_{int}$.

In this paper, the following optimization problem is considered with the segmentation output as the optimum labeling for each pixel.

$$x^* = \underset{x \in L^N}{\operatorname{argmin}}\, E(x) \quad (4)$$

By performing the above optimization, the accuracy of segmentation is improved.

## 3. Experimental Result

As the data set for the experiment, 20 cases of 3DIRCADb data set [15], 2542 sheets were used.

Intersection over Union (IoU) is used as the evaluation method. The formula for IoU is shown below.

$$IoU = \frac{|A \cap B|}{|A \cup B|} \quad (5)$$

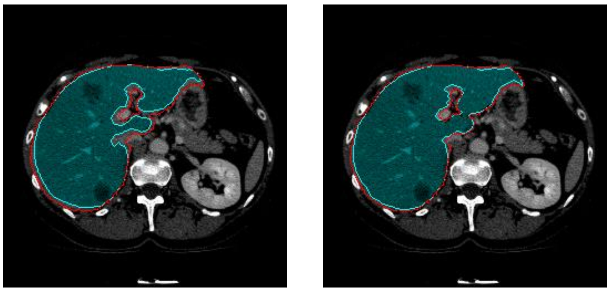
In order to verify the effectiveness of the proposed method, 5-fold cross-validation was performed. The experimental results are shown in Table 1, Fig. 2 and 3.

Table 1. Experimental result

| Approach | IoU |
|---|---|
| U-Net + 3DCRF [6] | 0.83 |
| Proposed method | |
| Residual U-Net + ResNet-D + seSE +3DCRF | 0.87 |

## 4. Discussion

From Table 1, the proposed method has higher accuracy than U-Net + 3DCRF.



(a) U-Net + 3DCRF (b) Proposed method

Fig. 2. Comparison of extraction accuracy of liver part

(Blue: Extracted result region, Red: Boundary of ground truth)

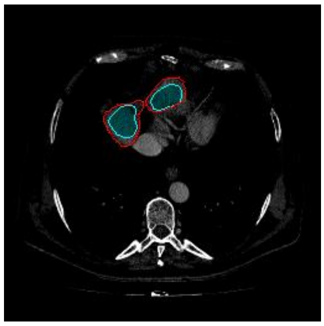


Fig. 3. Extraction result of liver terminal

From Fig. 2, it can be seen that the proposed method can extract the vicinity of the boundary with other organs, which was difficult in the past. We believe that the retention of information by the residual block and the characterization of the spatial position by the scSE block are particularly effective.

Figure 3 shows the extraction results of the terminal part of the liver, but shown in the figure, there are still some parts that have not been successfully extracted. This is because the shape of the liver often varies greatly from person to person, so it is considered that they have not been fully learned. The solution to this is the augment of data.

□

## 5. Conclusion

In this paper, we proposed segmentation of the liver region using deep learning and 3DCRF. By improving U-Net, which is a segmentation model, the experimental results showed that IoU was 0.87, which exceeded the conventional method of 0.83. Future tasks are further improvement of accuracy and introduction of alignment method and cancer detection method.

## References


1. Hyuna Sung et al., "Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries", A Cancer Journal for Clinicians, pp.209-249, 2021.
2. Jiarong Zhou et al., "Automatic detection and classification of focal liver lesions based on deep convolutional neural networks: A preliminary study", Frontiers in Oncology, pp.1-11, 2021.
3. Kentaro Saito et al., "Automatic liver segmentation from multiphase CT images by using level set method", International Conference on Control Automation and Systems, pp.1590-1592, 2017.
4. Esfandiarkhani Mina et al., "A generalized active shape model for segmentation of liver in low-contrast CT volumes", Computers in Biology and Medicine, 82, pp.59-70, 2017.
5. Christ Patrick Ferdinand et al., "Automatic liver and lesion segmentation in CT using cascaded fully convolutional neural networks and 3D conditional random fields", International Conference on Medical Image Computing and Computer-Assisted Intervention, pp.415-423, 2016.
6. Kentaro Saito et al., "ROI-based Fully Automated Liver Registration in Multi-phase CT Images", International Conference on Control, Automation and Systems, pp.645-649, 2018.
7. Krähenbühl Philipp et al., "Efficient inference in fully connected crfs with Gaussian edge potentials", NIPS, pp.109-117, 2011.
8. Zhang et al., "Road Extraction by Deep Residual U-Net", IEEE Geoscience and Remote Sensing Letters, Vol.15, Issue.5, pp.749-753, 2018.
9. He et al., "Deep Residual Learning for Image Recognition", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp.770-778. 2016.
10. T. He, "Bag of Tricks for Image Classification with Convolutional Neural Networks", Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp.558-567, 2019.
11. Abhijit Guha Roy et al., "Concurrent Spatial and Channel Squeeze & Excitation in Fully Convolutional Networks", arXiv preprint arXiv:1803.02579, 2018.
12. Olaf Ronneberger et al., "U-net: Convolutional networks for biomedical image segmentation." In: MICCAI, vol. 9351, pp. 234–241, 2015.
13. He et al., "Identify Mapping in Deep Residual Networks", European Conference on Computer Vision, pp.630-645, 2016.
14. Jie Hu et al., "Squeeze-and-Excitation Networks", arXiv preprint arXiv: 1709.01507, 2018.
15. Soler L et al., "3d image reconstruction for comparison of algorithm database: a patient-specific anatomical and medical image database", 2012.


## Authors Introduction

Mr. Shuntaro Nagano

He graduated from the Department of Mechanical and Control Engineering, Faculty of Engineering, Kyushu Institute of Technology, Japan in 2020. He is currently a Master course student in Kyushu Institute of Technology, Japan.

Dr. Tohru Kamiya

Tohru Kamiya received his B.A. degree in electrical engineering from Kyushu Institute of Technology in 1994, the Masters and Ph.D. degree from Kyushu Institute of Technology in 1996 and 2001, respectively. He is a professor in the department of control engineering at Kyushu Institute of Technology. His research interests are focused on medical application of image analysis. He is currently working on automatic segmentation of multi-organ of abdominal CT image, and temporal subtraction of thoracic MDCT images.

Dr. Guanguxu Li

Guangxu Li received his B.E. and M.E. of automation engineering from Liaoning Institute of Technology in 2006 and 2009 separately. And Ph.D. degree from Kyushu Institute of Technology in 2013. He joined Tiangong University after graduated. His research interests focus on the geometric methods applied to medical image.