

Received June 22, 2020, accepted July 13, 2020, date of publication July 30, 2020, date of current version August 20, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3012990

# DCSegNet: Deep Learning Framework Based on Divide-and-Conquer Method for Liver Segmentation

CONGSHENG LI<sup>1</sup>, GUORONG YAO<sup>2</sup>, XU XU<sup>1</sup>, LEI YANG<sup>1</sup>, YI ZHANG<sup>1</sup>,  
TONGNING WU<sup>1</sup>, (Member, IEEE), AND JUNHUI SUN<sup>2</sup>

<sup>1</sup>China Academy of Information and Communications Technology, Beijing 100191, China

<sup>2</sup>The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310003, China

Corresponding authors: Tongning Wu (wutongning@caict.ac.cn) and Junhui Sun (1307005@zju.edu.cn)

This work was supported by the grants from the National Science and Technology Major Project under Grant 2018ZX10301201.

**ABSTRACT** Image segmentation plays a vital role in the medical diagnosis and intervention field. The segmentation methods can be classified as fully automated, semiautomated or manual. Among them, manual segmentation can best improve the quality of the results, but it is time-consuming and tedious, and it may lead to operator bias. A continuity-aware probabilistic network based on the divide-and-conquer method was proposed in the current work. The proposed network comprised backbone network, local segmentation and a weight network. The backbone network extracts the features from image. The local segmentation divides the data space, whereas the weight network provides the continuity-aware weights. Therefore, combining those results of the weighted segments can eventually yield precise estimations. In this study, the proposed model was evaluated against several recent methods on the three datasets, and a several performance indexes of segmentation were evaluated for liver segmentation, the results showing that it is the most advanced liver segmentation approach. The source code of this work is publicly shared at <https://github.com/licongsheng/DCSegNet> for others to easily reproduce the work and build their own models with the introduced mechanisms.

**INDEX TERMS** Liver segmentation, deep learning, divide-and-conquer.

## I. INTRODUCTION

A basic task in planning liver operations is to detect and evaluate the liver shape using abdominal CT images and computer assistance, such as radiotherapy [1]. There are a lot of researches focused on improving segmentation accuracy and efficiency [2]. Nonetheless, numerous issues should be managed in liver segmentation, including the great variations in liver shapes, their different appearances and their ambiguous boundaries [3].

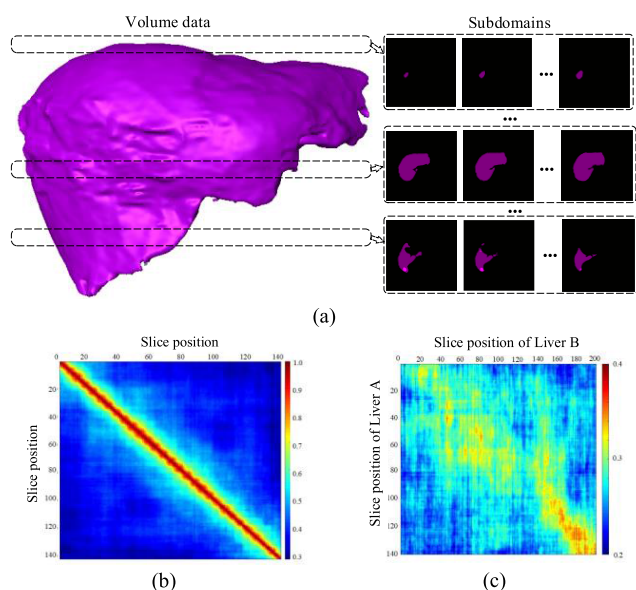
The liver segmentation algorithms are mainly classified into the following categories, including regional growth [4], graph cut [5], level set [6], and deep learning [7]. Typically, the deep learning method have become the hot research topics recently as a result of the accumulated data and the increased computing power, especially in machine vision tasks, such as the image classification [8] and segmentation [9].

The associate editor coordinating the review of this manuscript and approving it for publication was Seok-Bum Ko<sup>1</sup>.

Generally, the deep convolutional neural networks (DCNNs) designed for image segmentation are categorized into two types, namely, 2D and 3D networks [10]. Among them, the 2D DCNNs have achieved favorable performance in many 2D medical image segmentation scenario [11]–[13]. Nonetheless, the interslice correlations of the liver tissue's 3D structure is not considered in the 2D neural network [14], which together with the varying liver data structure makes it unsuitable to apply a 2D CNN to volumetric liver segmentation. It is worth noting that a volumetric liver segmentation algorithm should be able to take into account the interslice and intraslice characteristics. To address this issue, 3D DCNNs have been constructed [15], [16], such as V-net [17], denseVNet [16], Z-Net [18] and other 3D neural networks [19]. However, they will greatly increase the model complexity and the number of hyperparameters in the model [17], [20]. To this end, it is difficult to train a 3D DCNN with limited training data and hardware resources. In addition, some studies have been carried out to improve the

segmentation accuracy by adding loss function constraints, such as the boundary loss [20], the Hausdorff distance [21], the Mumford-Shah loss function [22] and signed distance map [23]. These methods contribute to improving the tissue edge segmentation accuracy of 2D slices, but they can hardly consider the continuity between layers.

A 2D DCNN segmentation network attempts to extract the features of all liver cross-sections through convolutions on the plane, which accounts for a key problem leading to the poor continuity between layers [24]. However, the structural changes in cross-sectional liver tissues are very significant; as a result, it is difficult to consider the differences between liver tissue and other tissues and the structural changes in vertical liver tissues. Based on the above considerations, this study aimed to analyze the structural similarity [25] in the cross-sectional liver tissues of different human bodies, as shown in Fig. 1. Such a continuity-induced similarity relationship was predominant in neighboring layers, indicating that if a region with high structural similarity was segmented with a local model, then a better segmentation result should theoretically be obtained.



**FIGURE 1. Structural similarity of liver. (a). Adjacent transverse plane of the liver (b). Similarity analysis of different transverse planes of the liver. (c). Similarity analysis of different liver transverse planes.**

A continuity-aware probabilistic network was proposed in the current work to address the abovementioned difficulties. First, a data volume is divided into several sub domains (Fig. 1 (a)). Then, a DCNN model was proposed that was constituted by a backbone network, local segmentation and weighted network. In the model, the backbone network was employed for feature extraction. All the local segmenters segment the target area according to the features extracted from the backbone network. The weighted network determines which local segmenter should be used for each subspace. In other words, the weighted network is similar to telling

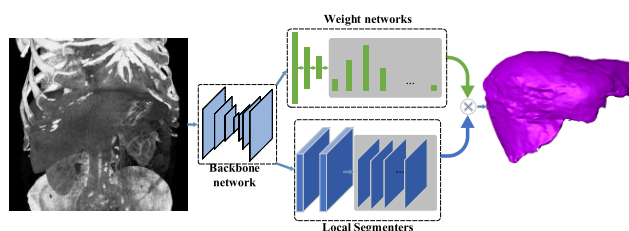
us which expert has better segmentation performance for the current sub domain. Our proposed model has several advantages. First, the local segmentation network was used for the explicit modeling of heterogeneous data according to the divide-and-conquer method [26]. Second, the weighted networks recognized the continuity of any two local segmentation networks. Third, the probabilistic soft decision, rather than the hard decision, was utilized in the weighted networks [27]; in this way, all local segmentations provided robust and accurate estimations. Fourth, the local segmentation networks were trained simultaneously with weighted networks, and the network was easily integrated with a deep neural network to form the end-to-end model. Our specific findings are as follows.

- 1) The DCsegNet model was proposed, which introduced the divide-and-conquer method to 3D data segmentation, and improved the segmentation accuracy through local segmentation.
- 2) A new data segmentation approach was proposed for training the DCsegNet model, and training data were used to train the adjacent local sequences with overlapping layers to ensure the continuity between adjacent sequences.
- 3) The proposed DCsegNet model attained higher advanced performance for three datasets.

## II. PROPOSED APPROACH

### A. OVERALL FRAMEWORK

The proposed liver segmentation framework based on the divide-and-conquer method is shown in Fig. 2. It consists of three parts, including a backbone network, a local segmentation network and weighted networks. The backbone network conducted feature extraction for any input image  $x \in X$ , and each slice image was fed into the deep convolution neural network for feature extraction. In the second part, a weighted network and a local segmentation network were employed to obtain the segmentation results from the extracted features. Notably, the final segmentation results were estimated as the weighted combination of all local segmentation networks.



**FIGURE 2. Liver segmentation framework put forward in this study.**

The local segmentation network was used to process the different cross-sectional data and divide the training data into  $L$  overlapping subdomains. Then, every data subset was adopted to train the corresponding local segmenter. We denote  $y \in Y$  as the output target of input sample  $x \in X$ , where  $X$  and  $Y$  was denoted the set of input image and

segmentation results, and  $x$  and  $y$  was denoted a sample  $X$  and  $Y$ . Therefore, the segmentation for the  $l^{th}$  subdomain ( $l = 1, 2, \dots, k$ ) was described below:

$$y(y|x, z = l) = N(y|\mu_l(x), \sigma_l^2) \tag{1}$$

where  $z$  stands for the latent variable denoting the  $\{x, y\}$  affiliation to one subdomain, whereas  $\mu_l(x)$  represents the segmentation results for the  $l^{th}$  local segmentation network of the input sample  $x$ . In addition, the Gaussian distribution  $N(y)$ , which has a mean of  $\mu_l(x)$  and a variance of  $\sigma_l^2$ , is adopted for modeling the segmentation error.

To efficiently combine the above segmentation results, a weighted network that had the novel convolution neural network structure was put forward, and it generated the weight of every local segmenter. Typically, the weighted that corresponded to the  $l^{th}$  local segmentation network was deemed to be  $\pi_l(x)$ . It was clear that for all  $x \in X$ ,  $\pi_l(x)$  was positive, and  $\sum_l \pi_l(x) = 1$ . Thereafter, liver segmentation was carried out through constructing the model for the conditional probability function:

$$P(y|x) = \sum_l \pi_l(x) N(y|\mu_l(x), \sigma_l^2) \tag{2}$$

Notably, segmentation aims to identify the mapping of  $g : x \rightarrow y$ . Subsequently, the expected distribution of the conditional probability was calculated to estimate the output  $y$  for the input sample  $x$ :

$$\begin{aligned} y &= E[P(y|x)] \\ &= E\left[\sum_l \pi_l(x) N(y|\mu_l(x), \sigma_l^2)\right] \\ &= \sum_l \pi_l(x) \mu_l(x) \end{aligned} \tag{3}$$

Therefore, the pooled results of the segmentation that were weighted based on the weight functions provided the eventual estimation. In the following sections, we will provide a detailed description of backbone network, local segmenter network and weighted networks respectively.

**B. BACKBONE NETWORK**

DeepLabv3+ extends DeepLabv3 by adding a simple yet effective decoder module to refine the segmentation results, especially along object boundaries [28]. It is able to encode multiscale contextual information by probing the incoming features with filters or pooling operations at multiple rates and multiple effective fields-of-view; in addition, it can also capture sharper object boundaries by gradually recovering the spatial information.

The backbone network for DCsegNet is shown in Fig. 3. The last convolution layer and the upsampling operation in DeepLabv3+ are removed. The input of the backbone network is a slice of a volume CT, and the output contains 256 features with  $64 \times 64$  dimensions. These output features are used as the direct input of local segmentation network and weight network respectively.

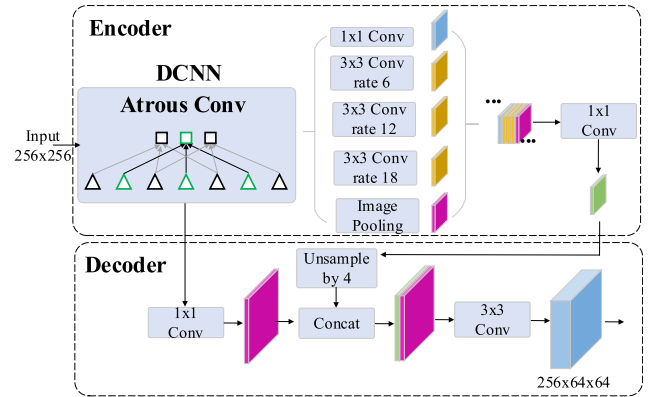


FIGURE 3. Backbone network employed in this study [28].

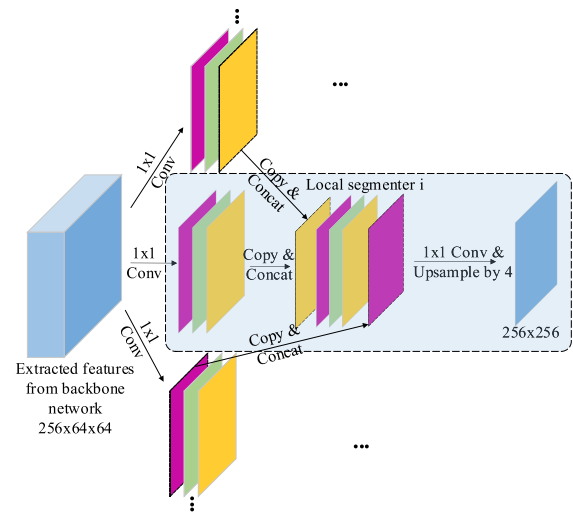


FIGURE 4. Local segmenter network architecture.

**C. LOCAL SEGMENTER**

In the divide-and-conquer method, local segmenters are adopted for the effective modeling of the subdomain data. The local segmenter network architecture is shown in Fig. 4. The number of local segmenters is same as the number of subdomains. Each local segmenter contains two convolution layers with a  $1 \times 1$  kernel. The first convolution layer extracts three feature maps from the output of the backbone network. Then, the local segmenter merges the local features with the sub features of the neighboring subdomain segmenter as the feature input of the last convolutional layer. The local segmentation networks can be deemed to be several experts, and every expert is skilled over the little segmentation subdomain, with diverse experts covering the distinct segmentation regions. Therefore, all experts are able to provide the expected outcome, regardless of the subdomain data.

To additionally model the liver label continuity, those local segmenter segmentation regions were densely overlapped. Those neighboring local segmenters shared a large degree of overlap within the specific regions, thereby rendering a great degree of similarity. As a result, multiple segmenters were

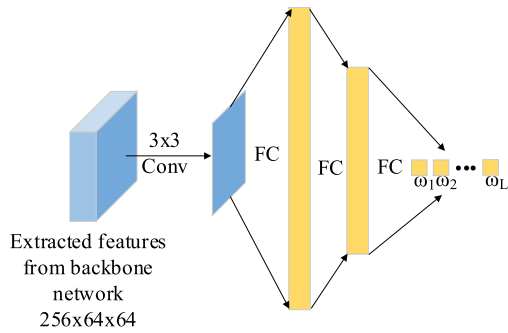


FIGURE 5. Weighted network architecture.

responsible for the liver segmentation for each subdomain, and so we were able to adopt ensemble learning to generate the precise segmentation results.

A comprehensively connected layer was employed for the dense overlapping of local segmenters. In addition, the sigmoid function was utilized to be an activation function, and the activation value of every local segmenter was mapped into the corresponding subdomain space, which was used as the expert outcome. According to the above discussion,  $\mu_l(x)$  was utilized to denote the  $l^{\text{th}}$  local segmentation network results, and the later segmentation loss was given by the following:

$$L_{seg} = \sum_{l \in L} (y_l - u_l(x))^2 \quad (4)$$

where

$$y_l = \begin{cases} \text{liver\_mask} & l = \text{index of subdomain} \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

#### D. WEIGHTED NETWORK

A weighted network was necessary in the approach proposed in this study to determine the local segmentation network weights. As a result, the use of a weighted network that had the divide-and-conquer structure boosted the cooperation between the local segmentation networks and the weighted network. With regard to the divide-and-conquer principle, the tree structure has been extensively utilized as a hierarchical structure. For instance, in the computer vision and machine learning communities, decision trees have been frequently used as classifiers in the coarse-to-fine decision-making process and as the tree structure. In this study, the weighted network contains one convolution layer and three fully connected layers. The weights  $\{\omega_i\}$ ,  $i = 1, 2, \dots, L$  are regressed from the features extracted from the backbone network. The SoftMax activation function is used to ensure that  $\{\omega_i\}$  satisfies the following requirements:

$$\begin{cases} \omega_i \in [0, 1] \\ \sum_{i=1}^L \omega_i = 1 \end{cases} \quad (6)$$

Notably, the KL divergence [29] was used to be a loss term to train the weight networks.

$$L_{weight} = - \sum_{l \in L} z_l \log(\omega_l(x)) \quad (7)$$

where  $\omega_l(x)$  is the weight of a subdomain  $x$  that was regressed by the weighted network, and  $z_l$  is one-hot vector that indicates to what subdomain  $x$  belongs.

#### E. LOSS FUNCTION

In the end, we jointly learn local segmenter and weighted networks by defining the total loss as follow:

$$Loss = L_{seg} + \lambda L_{weight} \quad (8)$$

where  $\lambda$  is utilized to balance the importance between the gating task and the regression task.  $\lambda = 1$  is used in this study.

According to our results, our proposed network was easily carried out through the use of the accessible comprehensively connected, sigmoid, and softmax layers within those available deep learning frameworks. Furthermore, the comprehensively differentiable network proposed in this study was embedded in each deep convolutional neural network, making it possible to conduct end-to-end training while obtaining a superior representation of the future.

### III. EXPERIMENTS

#### A. DATASET

The data used in the current work includes three datasets. They are the Liver Tumor Segmentation Challenge (LiTS) dataset, the 3D Image Reconstruction for Comparison of Algorithm Database-01 (3D-IRCADb-01) and custom CT scans from the First Affiliated Hospital, Zhejiang University School of Medicine, China. Fig. 6 shows the volume rendered abdominal CT images of three datasets.

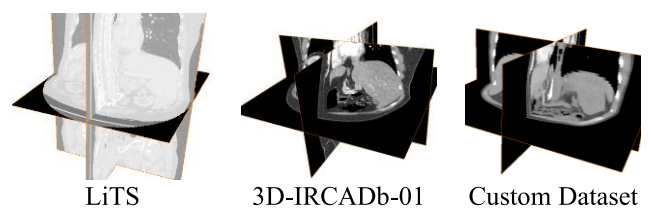


FIGURE 6. Three dataset examples used in the study.

#### 1) LiTS DATASET

A total of 200 3D contrast-enhanced abdominal CT scans, together with the segmentation labels for liver and tumor regions, are included in the LiTS dataset, and the resolution in every axial slice is  $512 \times 512$  pixels. One hundred thirty scans had the ground truth labels, including 10 as testing data, 20 as validation data, and 100 as training data. The scanning layer thickness in the data set is 0.8 mm-5 mm, and the number of liver scanning layers is 301-31. In training phase, Intensity values were clipped to the  $[-10, 200]$  HU range to ignore those uncorrelated details and normalize those



images to  $[0, 1]$ . Additionally, the images were downsampled to  $256 \times 256$ . The random cropping and normal distribution noise ( $N(0, 0.01)$ ) method was employed to enhance the training data [30].

## 2) 3D-IRCADb-01 DATASET

3D-IRCADb-01 is a database that includes several sets of anonymized medical images of patients and the manual segmentation of the various structures of interest was performed by clinical experts. The 3D medical images and masks of the segmented structures of interest are available as DICOM files. The dataset offers a higher variety and complexity of livers and their lesions and is publicly available. The 3DIRCADb includes 20 venous phase enhanced CT volumes from various European hospitals with different CT scanners. The scanning layer thickness in the data set is 1.0 mm-4 mm, and the number of liver scanning layers is 163-45. For our study, all volumes were used to validate the performance of the proposed model.

## 3) CUSTOM DATASET

The custom CT scans of ten patients from the First Affiliated Hospital, Zhejiang University School of Medicine were taken during the period from 2018 to 2019. All personal information about the patients was removed before the data left the hospital. The scanning layer thickness in the data set is 0.7 mm-1 mm, and the number of liver scanning layers is 162-134. These ten volumes were manually segmented by 3D Slicer (<https://www.slicer.org/>). In addition, all these data were used to validate the performance of the proposed model.

## B. EXPERIMENTAL SETTINGS

Our model was implemented with Torch [31] and optimized using the Adam algorithm on the NVIDIA Tesla K40c GPU. The original learning rate was set to  $1e-4$ , and it decayed according to the poly schedule learning rate. In addition, 10 volumes from the LiTS validation data were used to monitor the performance of our model. During training, each CT scan was resampled to  $256 \times 256$  pixels as the model input. 20 epochs were adopted to train the proposed model.

In the training data of DCsegNet, the number of subdomains ( $L$ ) was set as 3, 5, and 7. In addition, there were 0, 3, and 5 overlapping layers in adjacent subdomains that were used to find the best configuration for liver segmentation (as shown in Fig. 7). The labels of the weighted network's outputs were  $1, 2, \dots, L$ .

When the number of subdomains was set as 1, the proposed model could be considered to be the original DeepLabv3+. Table 1 shows the variation in the number of the training dataset of each local segmenter when the number of local segmenters and the size of overlap between adjacent regions changed.

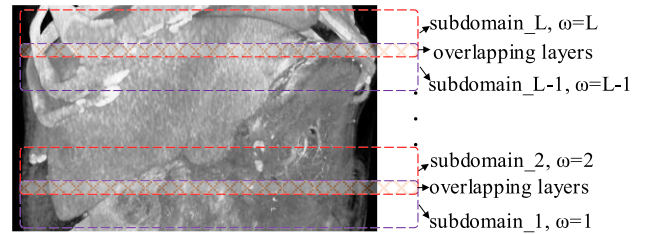


FIGURE 7. A diagram of the subspace segmentation.

TABLE 1. The number of training data for local segmenters.

No. of subdomains ( $L$ )	Overlapping layers	Number of training data per local segmenter
1	0	13624
3	3	4945
	5	5214
	7	5484
5	3	3209
	5	3940
	7	4285
7	3	2466
	5	2811
	7	3158

## C. EVALUATION METRICS

In this section,  $|TP|$ ,  $|TN|$ ,  $|FP|$ , and  $|FN|$  represent the true positive, true negative, false positive and false negative values, respectively.

### 1) DICE COEFFICIENT

The Dice coefficient ( $DICE$ ) is an approach to determine the spatial overlap between the gold standard image and the segmented one, with the value ranging from 0 (no overlap) to 1 (perfect matching), as shown below [33]:

$$DICE = \frac{2|TP|}{2|TP| + |FN| + |FP|} \quad (9)$$

### 2) RAND INDEX

The rand index ( $RI$ ) assesses the pixel consistency within the ground truth and segmented images, with a value that ranges from 0-1. Here, a value of '0' suggests that the segmentation results are completely different from the ground truth image whereas '1' represents that the segmentation result is identical to the ground truth image.

$RI$  is calculated using the following equation:

$$RI = \frac{|TP| + |TN|}{|TP| + |FN| + |TN| + |FP|} \quad (10)$$

### 3) SENSITIVITY

The sensitivity measures the percentage of actual positives values that are correctly identified whereas the specificity

measures the percentage of negative values that are correctly identified. The sensitivity (TPR) is the ratio of the pixels that are classified as positive and all the pixels that are actually positive.

$$TPR = \frac{|TP|}{|TP| + |FN|} \quad (11)$$

#### 4) HAUSDORFF DISTANCE

Hausdorff Distance (HD) is widely used in evaluating medical image segmentation methods. It is one of the most informative and useful criteria because it is an indicator of the largest segmentation error. In some applications, segmentation is one step in a more complicated multi-step process. For example, some multimodal medical image registration methods rely on segmentation of an organ of interest in one or several images. In such applications, the largest segmentation error as quantified by HD can be a good measure of the usefulness of the segmentations for the intended task. For two point sets  $X$  and  $Y$ , the one-sided HD from  $X$  and  $Y$  is defined as [21], [34]:

$$hd(X, Y) = \max_{x \in X} \min_{y \in Y} \|x - y\|_2 \quad (12)$$

and,

$$hd(Y, X) = \max_{y \in Y} \min_{x \in X} \|x - y\|_2 \quad (13)$$

The bidirectional HD between these two sets,

$$HD(X, Y) = \max(hd(X, Y), hd(Y, X)) \quad (14)$$

In the above definitions we have used the Euclidean distance, but other metrics can be used instead. Intuitively,  $HD(X, Y)$  is the longest distance one has to travel from a point in one of the two sets to its closest point in the other set. In image segmentation, HD is computed between boundaries of the estimated and ground-truth segmentations, which consist of curves in 2D and surfaces in 3D.

## IV. RESULTS AND DISCUSSION

### A. VERIFICATION OF THE WORKING PRINCIPLE OF LOCAL SEGMENTERS

In this study, a liver segmentation model was proposed in line with the divide-and-conquer principle based on the structural similarity of the adjacent liver cross-sections. Typically, high liver segmentation accuracy was achieved through training the local segmenters and the weights of the corresponding segmenter. Fig. 8 shows an example of the segmentation result for local segmenters ( $L = 5$ ) for the cross-sections at different positions.

According to the outputs of DCsegNet, the model selected the corresponding segmenter and calculated the weight of each segmenter according to the slice features extracted from the backbone network. In this study, the data of adjacent subdomains were overlapped, and the target region was segmented by multiple local segmenters for some areas of the sliced data; therefore, the weighted integration of those output

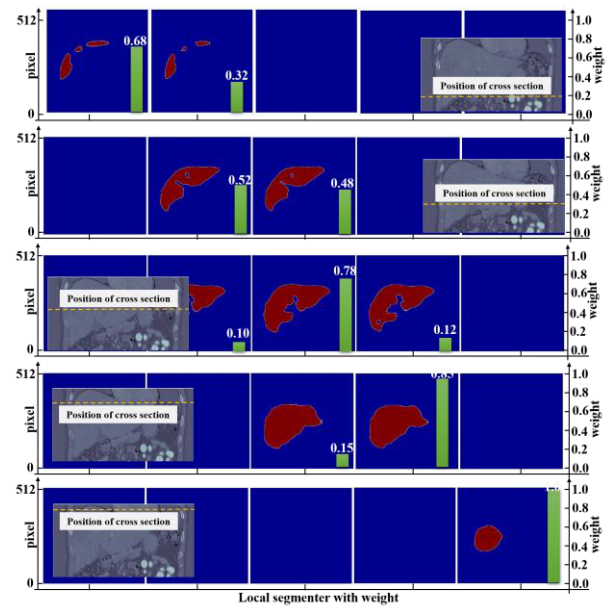


FIGURE 8. An example of the segmentation result for local segmenters on the transverse planes at different positions.

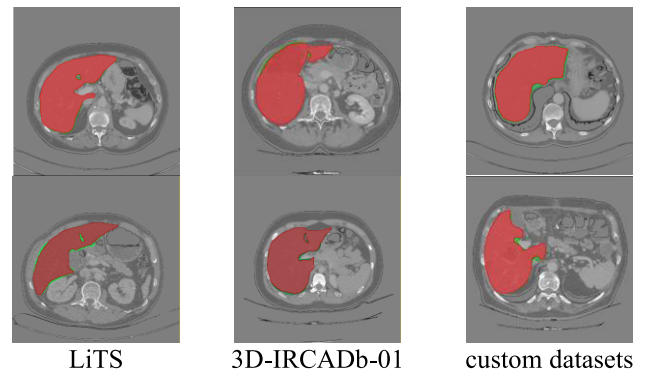


FIGURE 9. Comparison of segmentation accuracy of different segmentation algorithms. The green mask is the ground truth. The red mask is the segmentation result based on DCsegNet.

results of local segmenters by the weighted network helped to obtain the final segmentation result.

### B. COMPARISON WITH THE LATEST MODELS

Fig. 9 shows some 2D slices extracted from the LiTS, 3D-IRCADb-01 and custom datasets, as well as the ground truth and segmentation results obtained from the DCsegNet model.

The means and standard deviations of DICE, RI, TPR and 90<sup>th</sup> HD are listed in Table 2 to Table 5, respectively. These tables indicated the DCsegNet model performance with different numbers of local segmenters using three different datasets. As shown from these tables, when the number of local segmenters is 5, for the different datasets, the three segmentation indicators have obtained the best results.

**TABLE 2.** Dice coefficient validation against different methods on different datasets (dice %).

Architecture	LiTS	3D-IRCADb-01	Custom Dataset
DeepLabV3+	95.13 ± 1.36	95.91 ± 1.46	92.58 ± 2.13
DCSegNet-3(3)	95.51 ± 1.13	96.68 ± 1.32	92.97 ± 1.13
DCSegNet-5(3)	<b>95.97 ± 1.03</b>	<b>97.51 ± 1.13</b>	<b>93.37 ± 1.01</b>
DCSegNet-7(3)	95.07 ± 1.21	95.20 ± 1.45	91.21 ± 1.21

**TABLE 3.** Rand index validation against different methods on different datasets (RI %).

Architecture	LiTS	3D-IRCADb-01	Custom Dataset
DeepLabV3+	99.26 ± 0.18	99.55 ± 0.15	99.12 ± 0.68
DCSegNet-3(3)	99.38 ± 0.19	99.73 ± 0.54	99.25 ± 0.54
DCSegNet-5(3)	<b>99.42 ± 0.14</b>	<b>99.97 ± 0.05</b>	<b>99.29 ± 0.35</b>
DCSegNet-7(3)	99.12 ± 0.25	99.43 ± 0.14	99.14 ± 0.44

**TABLE 4.** Sensitivity validation against different methods on different datasets (TPR %).

Architecture	LiTS	3D-IRCADb-01	Custom Dataset
DeepLabV3+	97.42 ± 0.48	97.77 ± 0.51	95.03 ± 1.51
DCSegNet-3(3)	98.28 ± 0.43	98.01 ± 0.51	95.41 ± 1.34
DCSegNet-5(3)	<b>98.84 ± 0.47</b>	<b>98.64 ± 0.54</b>	<b>96.14 ± 1.11</b>
DCSegNet-7(3)	86.16 ± 2.02	97.84 ± 1.29	96.12 ± 1.36

**TABLE 5.** 90<sup>th</sup> Hausdorff distance validation against different methods on different datasets (HD in mm).

Architecture	LiTS	3D-IRCADb-01	Custom Dataset
DeepLabV3+	3.60	3.25	4.38
DCSegNet-3(3)	3.41	3.12	4.21
DCSegNet-5(3)	<b>3.18</b>	<b>2.97</b>	<b>4.16</b>
DCSegNet-7(3)	3.81	3.57	4.87

### C. ABLATION INVESTIGATION AND PARAMETER ANALYSIS

To evaluate the effect of the number of different local subdomains and the overlap between adjacent subdomains on the segmentation results, the number of subdomains in the LiTS dataset is set a 3, 5, and 7 and the overlap between adjacent subdomains is 0, 3, and 5. We trained different network models, and used the models to verify the data of the three data sets. The comparison of the means and standard deviations of the DICE are shown in Table 6.

#### 1) THE NUMBER OF LOCAL SEGMENTERS

Table 6 displays some specific conclusions. First, after comparing the impact on the segmentation results of using 3, 5, and 7 local segmenters to construct the segmentation network, it can be seen that the best results are obtained using 5 local segmenters. In theory, highly accurate segmentation was acquired from those abovementioned structures through the use of the set number of segmenters, which was reasonable since a higher number of local segmenters suggested a greater quantity of local segmentation networks, whereas

**TABLE 6.** Validation against different on three datasets (dice %).

Architecture (NO. of hyper parameters)	Over lap	LiTS	3D-IRCADb- 01	Custom Dataset
DeepLab V3+ (75.79 M)	/	95.13 ± 1.36	95.91 ± 1.46	92.58 ± 2.13
DCSegNet-3 (96.13 MB)	0	95.48 ± 1.21	96.64 ± 1.37	92.92 ± 1.20
	3	95.51 ± 1.13	96.68 ± 1.32	92.97 ± 1.13
	5	95.58 ± 1.12	96.68 ± 1.25	92.97 ± 1.12
DCSegNet-5 (96.14 M)	0	95.91 ± 1.12	97.42 ± 1.17	93.32 ± 1.03
	3	95.97 ± 1.02	97.51 ± 1.13	93.37 ± 1.01
	5	96.01 ± 1.00	97.52 ± 1.12	93.41 ± 1.00
DCSegNet-7 (96.15 M)	0	95.05 ± 1.22	95.14 ± 1.57	91.20 ± 1.29
	3	95.07 ± 1.21	95.20 ± 1.45	91.21 ± 1.21
	5	95.07 ± 1.18	95.19 ± 1.45	91.29 ± 1.14

a greater number of local segmentation networks indicated higher expert intelligence. Furthermore, the researchers continued to increase the number of local segmenters to analyze its impact on the segmentation results. For example, when the number of local segmenters increases to 7, the segmentation accuracy decreases by different amounts. Therefore, we can draw a conclusion that the performance tended to be saturated when the number of local segmenters was large enough since certain neighboring local segmentation networks corresponded to identical training data in the presence of an excessive number of local segmenters, making it impossible to boost the actual number of experts. In addition, as seen from Table 1, as the number of local segmenters increases, the amount of training data for each local divider will be significantly reduced, which may result in the insufficient training of the hyperparameters for each local divider. That is, the model is underfit. In addition, due to the individual differences in the liver disease structures of the patients and the differences in the CT imaging layer thickness, the similarity of the CT image structure in the same subdomain is reduced, which will further increase the difficulty of model fitting. These factors need to be further verified by standardizing the imaging process (such as selecting the same layer thickness) and increasing the number of training sets.

#### 2) THE SIZE OF THE OVERLAP LAYERS FOR SUBDOMAINS

In this study, different sized overlapping region for adjacent subdomains of the training data were selected to study their impact on the segmentation results. It can be seen from table 6 that increasing the number of overlapping layers of adjacent subdomains will not significantly improve the segmentation accuracy, but the standard deviation of the segmentation will decrease. Table 1 shows that when increasing the number of overlapping layers of adjacent subdomains, for each local segmenter, the training data for the optimizer network parameters will increase accordingly, which is beneficial to improving the model performance. In addition, as the number of

overlaps between adjacent subdomains increases, the data of the adjacent subdomains of adjacent local segmenters has better compatibility. In other words, each expert is not only good at the data segmentation of their respective subdomains but also the data segmentation of adjacent subdomains will be more professional. This results in better stability of the segmentation results, and a lower standard deviation in the DICE coefficient.

### 3) THE LOCAL SEGMENTATION ARCHITECTURE

To further improve the segmentation performance of the local subregions on adjacent subdomains, the middle layer of each local partition in the study introduces part of the middle CNN layer features from the adjacent local partitions, as shown in Fig. 4. In the study, we compared the effect of introducing the middle layer feature of the adjacent local splitter and not introducing this feature on the segmentation results. The results are shown in table 7.

**TABLE 7. Influence of local segmentation architecture differences on segmentation results (DICE %).**

DCSegNet-5(3)	LiTS	3D-IRCADb-01	Custom Dataset
with cross features	95.97 ± 1.03	97.51 ± 1.13	93.37 ± 1.01
without cross features	95.96 ± 1.10	97.51 ± 1.16	93.35 ± 1.09

It can be seen from the calculation results that the use of different splitter architectures does not have a significant impact on the segmentation accuracy. This is mainly due to the overlapping of the subdomains in the training data. In addition, the training set data in LiTS use different layer thicknesses for the scanning procedures, which results in a reduction in the similarity of the data structure of the same subdomain and the overlap in the data structures of adjacent subdomains. Therefore, by introducing the local features of the middle layer of the adjacent local dividers in this study, the segmentation performance is not significantly improved.

## V. CONCLUSION

The current work proposes the DCsegNet model, which is the continuity-aware probabilistic network used in liver segmentation. DCsegNet can definitely model the established continuity relationship across different components using the local segmentation networks based on the probabilistic network. According to the experimental results from using the three datasets, this model shows higher accuracy than the other advanced models when the appropriate parameters are selected.

Due to the different layer thicknesses of the data in the training set, the individual differences in the abdominal features, and the limited amount of data in the training set, the local segmenter is difficult to achieve the best expert performance. Therefore, it can be seen from the results that this study still does not solve the problem of continuity between

adjacent layers. It is necessary to standardize the imaging parameters and collect more data to establish training in the future datasets to further optimize and improve the local segmenter to achieve more accurate segmentation results.

## ACKNOWLEDGMENT

Thanks to the creators of the LiTS (<https://competitions.codalab.org/competitions/17094>) and 3D-IRCADb-01 (<https://www.ircad.fr/research/3d-ircadb-01/>) datasets, and thank VainF for providing the deeplabv3+ code on Github (<https://github.com/VainF/DeepLabV3Plus-Pytorch>).

## REFERENCES

- [1] M. D. A. van Gastel, M. E. Edwards, and V. E. Torres, "Automatic measurement of kidney and liver volumes from MR images of patients affected by autosomal dominant polycystic kidney disease," *J. Amer. Soc. Nephrology*, vol. 30, no. 8, pp. 1514–1522, 2019.
- [2] M. Moghbel, S. Mashohor, R. Mahmud, and M. I. B. Saripan, "Review of liver segmentation and computer assisted detection/diagnosis methods in computed tomography," *Artif. Intell. Rev.*, vol. 50, no. 4, pp. 497–537, Dec. 2018.
- [3] R. L. Araujo and M. M. Linhares, "Pushing the limits of liver surgery for colorectal liver metastases: Current state and future directions," *World J. Gastrointestinal Surg.*, vol. 11, no. 2, p. 34, 2019.
- [4] W. Xu, H. Liu, X. Wang, and Y. Qian, "Liver segmentation in CT based on ResUNet with 3D probabilistic and geometric post process," in *Proc. IEEE 4th Int. Conf. Signal Image Process. (ICSIP)*, Jul. 2019, pp. 685–689.
- [5] W. Wu, Z. Zhou, S. Wu, and Y. Zhang, "Automatic liver segmentation on volumetric CT images using supervoxel-based graph cuts," *Comput. Math. Methods Med.*, vol. 2016, Mar. 2016, Art. no. 9093721.
- [6] J. Wang, H. Zu, H. Guo, R. Bi, Y. Cheng, and S. Tamura, "Patient-specific probabilistic atlas combining modified distance regularized level set for automatic liver segmentation in CT," *Comput. Assist. Surg.*, vol. 24, no. 2, pp. 20–26, Oct. 2019.
- [7] M. Ahmad, Y. Ding, S. F. Qadri, and J. Yang, "Convolutional-neural-network-based feature extraction for liver segmentation from CT images," in *Proc. 11th Int. Conf. Digit. Image Process. (ICDIP)*, vol. 11179, Aug. 2019, Art. no. 1117934.
- [8] S. Cao, Z. Chen, C. Li, C. Lv, T. Wu, and B. Lv, "Landmark-based multi-region ensemble convolutional neural networks for bone age assessment," *Int. J. Imag. Syst. Technol.*, vol. 29, no. 4, pp. 457–464, Dec. 2019.
- [9] L. Saba, M. Biswas, V. Kupilli, E. C. Godia, H. S. Suri, D. R. Edla, T. Omerzu, J. R. Laird, N. N. Khanna, S. Mavrogeni, A. Protergerou, P. P. Sfikakis, V. Viswanathan, G. D. Kitas, A. Nicolaidis, A. Gupta, and J. S. Suriq, "The present and future of deep learning in radiology," *Eur. J. Radiol.*, vol. 114, pp. 14–24, May 2019.
- [10] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. W. M. van der Laak, B. van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, Dec. 2017.
- [11] A. E. Kavur, N. S. Gezer, M. Barış, Y. Şahin, S. Özkan, B. Baydar, U. Yüksel, Ç. Kılıkçier, Ş. Olut, G. B. Akar, and G. Ünal, "Comparison of semi-automatic and deep learning-based automatic methods for liver segmentation in living liver transplant donors," *Diagnostic Interventional Radiol.*, vol. 26, no. 1, p. 11, 2020.
- [12] X. Li, H. Chen, X. Qi, Q. Dou, C.-W. Fu, and P.-A. Heng, "H-DenseUNet: Hybrid densely connected UNet for liver and tumor segmentation from CT volumes," *IEEE Trans. Med. Imag.*, vol. 37, no. 12, pp. 2663–2674, Dec. 2018.
- [13] W. Qin, J. Wu, F. Han, Y. Yuan, W. Zhao, B. Ibragimov, J. Gu, and L. Xing, "Superpixel-based and boundary-sensitive convolutional neural network for automated liver segmentation," *Phys. Med. Biol.*, vol. 63, no. 9, May 2018, Art. no. 095017.
- [14] X. Han, "Automatic liver lesion segmentation using a deep convolutional neural network method," 2017, *arXiv:1704.07239*. [Online]. Available: <http://arxiv.org/abs/1704.07239>
- [15] Q. Dou, H. Chen, Y. Jin, L. Yu, J. Qin, and P.-A. Heng, "3D deeply supervised network for automatic liver segmentation from CT volumes," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, 2016, pp. 149–157.



- [16] E. Gibson, F. Giganti, Y. Hu, E. Bonmati, S. Bandula, K. Gurusamy, B. Davidson, S. P. Pereira, M. J. Clarkson, and D. C. Barratt, "Automatic multi-organ segmentation on abdominal CT with dense V-networks," *IEEE Trans. Med. Imag.*, vol. 37, no. 8, pp. 1822–1834, Aug. 2018.
- [17] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-net: Fully convolutional neural networks for volumetric medical image segmentation," in *Proc. 4th Int. Conf. 3D Vis. (3DV)*, Oct. 2016, pp. 565–571.
- [18] P. Li, X.-Y. Zhou, Z.-Y. Wang, and G.-Z. Yang, "Z-net: An anisotropic 3D DCNN for medical CT volume segmentation," 2019, *arXiv:1909.07480*. [Online]. Available: <http://arxiv.org/abs/1909.07480>
- [19] H. Tang, M. Moradi, K. C. L. Wong, H. Wang, A. El Harouni, and T. Syeda-Mahmood, "Integrating deformable modeling with 3d deep neural network segmentation," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Springer, 2018, pp. 377–384.
- [20] H. Kervadec, J. Bouchtiba, C. Desrosiers, É. Granger, J. Dolz, and I. B. Ayed, "Boundary loss for highly unbalanced segmentation," 2018, *arXiv:1812.07032*. [Online]. Available: <http://arxiv.org/abs/1812.07032>
- [21] D. Karimi and S. E. Salcudean, "Reducing the hausdorff distance in medical image segmentation with convolutional neural networks," *IEEE Trans. Med. Imag.*, vol. 39, no. 2, pp. 499–513, Feb. 2020.
- [22] B. Kim and J. C. Ye, "Mumford–Shah loss functional for image segmentation with deep learning," *IEEE Trans. Image Process.*, vol. 29, pp. 1856–1866, Sep. 2020.
- [23] Y. Zhang, X. Pan, C. Li, and T. Wu, "3D liver and tumor segmentation with CNNs based on region and distance metrics," *Appl. Sci.*, vol. 10, no. 11, p. 3794, May 2020, doi: [10.3390/app10113794](https://doi.org/10.3390/app10113794).
- [24] J. Zhang, Y. Xie, P. Zhang, H. Chen, Y. Xia, and C. Shen, "Light-weight hybrid convolutional network for liver tumor segmentation," in *Proc. 28th Int. Joint Conf. Artif. Intell.*, Macao, China, Aug. 2019, pp. 10–16.
- [25] G. P. Renieblas, A. T. Nogués, A. M. González, N. Gómez-Leon, and E. G. del Castillo, "Structural similarity index family for image quality assessment in radiological images," *J. Med. Imag.*, vol. 4, no. 3, Jul. 2017, Art. no. 035501.
- [26] M. Rebsamen, U. Knecht, M. Reyes, R. Wiest, R. Meier, and R. McKinley, "Divide and conquer: Stratifying training data by tumor grade improves deep learning-based brain tumor segmentation," *Frontiers Neurosci.*, vol. 13, p. 1182, Nov. 2019.
- [27] M. Tomlinson, C. J. Tjhai, M. A. Ambroze, M. Ahmed, and M. Jibril, "Soft decision and quantised soft decision decoding," in *Error-Correction Coding and Decoding*. SpringerLink, 2017, pp. 43–58.
- [28] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," in *Proc. Comput. Vis. Pattern Recognit.*, 2018, pp. 833–851.
- [29] A. Malinin and M. Gales, "Reverse KL-divergence training of prior networks: Improved uncertainty and adversarial robustness," in *Proc. Adv. Neural Inf. Process. Syst.*, 2019, pp. 14520–14531.
- [30] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *J. Big Data*, vol. 6, no. 1, p. 60, Dec. 2019.
- [31] A. Paszke et al., "PyTorch: An imperative style, high-performance deep learning library," in *Proc. Adv. Neural Inf. Process. Syst.*, 2019, pp. 8024–8035.
- [32] L.-C. Chen, G. Papandreou, F. Schroff, and H. Adam, "Rethinking atrous convolution for semantic image segmentation," 2017, *arXiv:1706.05587*. [Online]. Available: <http://arxiv.org/abs/1706.05587>
- [33] T. Konopczyński, J. Rathore, T. Kröger, L. Zheng, C. S. Garbe, S. Carmignato, and J. Hesser, "Reference setup for quantitative comparison of segmentation techniques for short glass fiber CT data," 2019, *arXiv:1901.01210*. [Online]. Available: <http://arxiv.org/abs/1901.01210>
- [34] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*. Berlin, Germany: Springer-Verlag, 1998.

• • •