ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/332033098

Liver Segmentation in CT based on ResUNet with 3D Probabilistic and Geometric Post Process

Conference Paper · March 2019

DOI: 10.1109/SIPROCESS.2019.8868690

CITATIONS
22

READS
1,544

4 authors, including:

Wendong Xu
The University of Hong Kong
4 PUBLICATIONS **33** CITATIONS

SEE PROFILE

Hong Liu
Chinese Academy of Sciences
89 PUBLICATIONS **812** CITATIONS

SEE PROFILE

Xiangdong Wang
Chinese Academy of Sciences
85 PUBLICATIONS **593** CITATIONS

SEE PROFILE

All content following this page was uploaded by Hong Liu on 28 March 2019.

The user has requested enhancement of the downloaded file.

Liver Segmentation in CT based on ResUNet with 3D Probabilistic and Geometric Post Process

Wendong Xu, Hong Liu, Xiangdong Wang, Yueliang Qian
Beijing Key Laboratory of Mobile Computing and Pervasive Device
Institute of Computing Technology, Chinese Academy of Sciences
Beijing, China
Email: xuwendong18g@ict.ac.cn, hliu@ict.ac.cn, xdwang@ict.ac.cn, ylqian@ict.ac.cn

***Abstract*—This paper proposes a novel liver segmentation framework using ResUNet with 3D probabilistic and geometric post process. Our semantic segmentation model ResUNet adds residual unit and batch normalization layer to up sampling and down sampling part of U-Net to construct a deeper network. To quick converge, we propose a new loss function DCE, which is linearly combined by Dice loss and cross entropy loss. We use continuous several CT images as input for training and testing to explore more context information. Based on initial segmentation of ResUNet, fully connected 3D conditional random field is used to refine segmentation results by exploring 2D neighbor regions and 3D volume information. Finally, 3D connected components analyzing is used to remain some large components and reduce segmentation noise. The experimental results on public dataset LiTS show our proposed framework achieve the state of the art performance for liver segmentation.**

Keywords-CT; liver segmentation; post processing, deep learning

I. INTRODUCTION

Liver cancer is one common cancer diseases and causes massive deaths every year in the world. Automatic segmentation of liver and lesion is important for deriving quantitative biomarkers for accurate clinical diagnosis and computer-aided decision systems. Liver segmentation in CT images is the basic step for liver tumor detection and recognition, which can avoid too much relying on doctors' prior knowledge. While automatic liver segmentation in CT is a challenging task due to complex liver anatomy and disturbance by other organs. And in CT sequences, the shape and size of the liver are quite different due to different image positions and different person.

Some previous work used image segmentation methods to segment liver in CT images. J. Huang et al. [1] used 3D regional growth method to segment liver region. Baâzaoui et al. [2] proposed an entropy-based fuzzy region growing method to semi-automated segment liver tumor regions. Above traditional image segmentation methods need to choose regional growth seeds, which is empirical and difficult to apply for various and complex CT sequences.

With great success of deep learning in the field of natural images, deep convolutional neural networks have performed well in image segmentation [3,4]. Researchers began to segment the liver in CT images using semantic segmentation methods. Olaf et al. [5] firstly used an end-to-end fully convolutional neural network to segment medical images. Christ et al. [6] proposed a cascaded U-Net to segment liver and liver tumor simultaneously. U-Net [5] has an encoder-decoder structure that incorporating multi-scale features by using size-corresponded down sampling and up sampling layers. They also used fully connected 3D conditional random field to improve segmentation results on public dataset 3DircaDB [7]. But in U-Net segmentation stage, Christ's method does not fully make use of adjacent images of CT volume for model training. X Li et al. [7] proposed H-DenseNet model for liver segmentation and got the state-of-the-art performance on public LiTS dataset [8]. But the two DenseNet structures may result in high computing complexity and using much time to converge model.

This paper proposes a framework for liver segmentation in CT sequences using context information, much deeper network and quicker convergence strategy. The main contributions of this paper are as follows.

(1) This paper introduces a semantic segmentation model ResUNet with context multi-images input for liver segmentation in CT sequence. We use the U-Net structure with residual unit [9], which is named ResUNet in this paper. We add batch normalization layer [10] to make the network deeper and quick convergence.

(2) We propose a new loss function DCE to improve ResUNet training performance and obtain quick network convergence. DCE loss is combined with Dice loss and Cross-Entropy loss, which makes use of global and local segmentation performance.

(3) We use two post-processing methods including fully connected 3D conditional random field 3D_CRF and 3D connected components analysis 3D_CCA, which can use 3D context information to improve the final live segmentation performance.

Based on the proposed liver segmentation framework, we obtain state-of-the-art performance on public dataset LiTS.

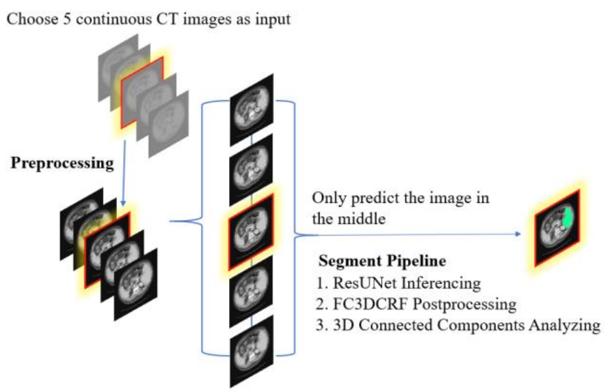
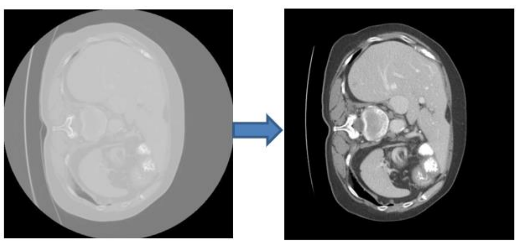


Figure 1. Our proposed framework.



(a) Original CT image (b) After pre-processing

Figure 2. Sample of CT image after preprocessing.

Figure 3. The structure of our ResUNet with BN layer.

II. OUR MAIN WORK

A. The Main Framework

This paper proposes a framework as Figure 1 shows to obtain accurate and robust liver segmentation from background in CT. For each image in CT sequence, Hounsfield adjustment and histogram equalization are used as preprocessing to strengthen image contrast and reduce noise. Then ResUNet with multi-images input is adopted to semantic segment CT image. Finally, fully connected 3D_CRF and 3D connected components analyzing 3D_CCA are used as post-processing to improve liver segmentation performance.

B. Data Preprocessing

We evaluate our method on public dataset LiTS from MICCAI2017 [8]. LiTS is more complex and bigger than 3DircaDB [9] and includes 3DircaDB. For the original CT Hounsfield values have large range and are not suited for image segmentation. We firstly adjust the original Hounsfield values of each image in CT sequence by remaining the range between -200 HU and 200 HU, which can get a clearer organ to distinguish and reduce image noise. We also use histogram equalization to enhance the contrast of the organ boundaries. Figure.2 shows CT image sample using above preprocessing method.

C. Problem Description

We define an image of CT sequence as I, and total number of pixels in the image I as N. For each pixel in a CT volume, its position is described as a triple (x, y, z), which means in the z^{th} image of the volume, the coordinate is (x, y). There are n possible classes for each pixel, which are defined by $L = \{l_1, l_2, \dots, l_n\}$. Let the class of a pixel be $v_i \in L$, then the probability that the class of pixel v_i is k can be defined as $P(v_i = k|I)$. This paper focuses on segmenting liver and background, so here k can have a value of 1 or 0 accordingly.

D. Initial Segmentation by ResUNet and DCE Loss

This paper uses ResUNet model to obtain the probability matrix $P(x_i\,|I)$ of background and liver pixels in each CT image. For ResUNet, we add residual unit and Batch Normalization BN layer [10] to each down sampling layer and up sampling layer based on U-Net as Figure 3 shows. Residual unit is used to ensure the gradient exponential decay does not occur in the deep network part and train adequately. Bath normalization layer is used to avoid gradient disappear problem when the model goes deeper. Above strategy can make network deeper and fast convergence. We increase network layers for more precise positioning. We construct ResUNet with 5 blocks of down sampling and up sampling units instead of 3 in [11].

Our structure of ResUNet is shown in Figure 3. In our paper, all convolutional layers' kernel size is 3×3. After each pooling layer, the length and width of feature map become half of the former input. And after each un-pooling layer, the length and width of feature map become two times of the former input. So we can get the semantic segmentation results as original input image size in final. We represent the residual operation with a red plus sign (+) in Figure 3 and long connection is used to make use of former layer information.

To combine context information between adjacent images, we select two images front and after each current image as shown in Figure 1.

E. Proposed DCE Loss Function

To train above ResUNet model, we propose a new loss function called DCE combined with Dice loss [12] and Cross-Entropy loss [5].

F Milletari et al. [12] firstly applied Dice value as a loss function in their V-Net model training to segment MRI images. Dice value can describe the global segmentation performance with the ground-truth mask as follow.

$$D = \frac{2\sum_i^N p_i g_i}{\sum_i^N p_i^2 + \sum_i^N g_i^2}. \quad (1)$$

Where N denotes the sum of pixels, p_i is the predicted label and $g_i$is the ground truth label, which are in range [0, 1].

The Cross-Entropy loss [5] is widely used in various model training by measuring the degree of similarity between predict distribution and ground-truth distribution. This loss is sensitivity to unbalanced data, so we give different weights for background and liver. Let Ω stands for the image, x_k stands for pixel x belongs to k_{th} class. The weight of the k_{th} class be ω_k and the Cross-Entropy is defined as follows.

$$E = -\sum_{k\in\Omega} \omega_k \log(P(x_k \mid I)). \quad (2)$$

We calculate the softmax layer's Cross-Entropy using (2). Dice loss only uses the label of the highest probability, which is not probability level. Cross-entropy directly uses predict probability to calculate the loss. To make full use of above two losses, we combine Dice loss and Cross-Entropy loss as DCE loss, which is defined as follow.

$$DCE = \epsilon D + (1-\epsilon)E. \quad (3)$$

Where ϵ denotes a fixed factor get in range of 0 and 1, which is set as 0.5 in our experiment.

F. Refined Segmentation by 3D_CRF

After above process, we can get the initial segmentation result. But semantic segmentation by ResUNet model only considers the individual pixel recognition, which does not make use of the context and correlation information of neighbor region in 2D CT image and 3D volume. In order to explore the local information between pixels and images, we use fully connected 3D conditional random field 3D_CRF as a post-processing step. We connected all pixels in a CT volume and consider every pair of pixels relationships of position and Hounsfield value. We calculate potentials for every two pixels in (x_i, y_i, z_i) and (x_j, y_j, z_j). We define pairwise potentials function using two Gaussian kernel functions as Christ et al. [6] suggested:

$$k(f_i, f_j) = \omega^{(1)} exp\left(-\frac{|p_i - p_j|^2}{2\theta_\alpha^2} - \frac{|I_i - I_j|^2}{2\theta_\beta^2}\right) + \omega^{(2)} exp\left(-\frac{|p_i - p_j|^2}{2\theta_\gamma^2}\right) \quad (4)$$

Where $k(f_i, f_j)$ stands for the Gaussian potentials, p_i and p_j are the position of pixel i and j, I_i, I_j are the Hounsfield value of pixel i and j. $\theta_\alpha, \theta_{beta}, \theta_\gamma$ are super parameters to control the degrees of nearness and similarity. ω_1, ω_2 are linear combination weights. The first Gaussian kernel uses Hounsfield value and position of pixel i and j, and the second Gaussian kernel considers positional information of pixel i and j.

After using the fast inference algorithm [13], we can get class labels for all pixels in CT after 3D_CRF processing.

G. 3D Connected Component Analyzing

Since CT images may be affected by spatial noise during the capture process. After using 3D_CRF processing, some small background pixels may still be recognized as liver. We use 3D connected component analysis 3D_CCA to reduce such noise. We calculate the 3D connected regions in the above segmented CT sequence and keep the largest n 3D connected regions as the final liver segmentation result.

The computation complexity of original connected components algorithm is $O(VE)$, where V is the number of vertex and E is the total number of edges. To reduce computing complexity, we adopt the unite-find set [14], which can get $O(\mathrm{E}\alpha(x))$, where x is the height of the unite-find tree and $\alpha(x)$ is the inverse Ackermann function. It is guaranteed in [14] that using path compression techniques, the computing complexity can reduce to $O(E)$, which is much faster than original connected components algorithm.

III. Experiments

A. Dataset and Experiment Setting

We evaluate our proposed framework on public dataset LiTS from MICCAI2017 Liver Tumor Segmentation Challenge [8]. LiTS dataset includes 200 CT scans, which are provided by various clinical sites in the world and the liver and lesion region is provided by mask for each image of CT sequence. As Table I shows, the training set contains 50,622 abdominal scan images of 130 CT scans from 91 patients.

The test set includes 26,608 images of 70 CT scans from 40 patients. For CT sequences, about 35% images contain liver region.

TABLE I. The train set and test set of LiTS Dataset

Dataset Part	Volume number	Image number
Train set	130	50,622
Test set	91	26,608

We design three sets of experiments to evaluate our framework and methods for liver segmentation in CT scans. 1) We validate the effectiveness of ResUNet with DCE loss by comparing Cross-Entropy loss. 2) We evaluate post-processing performance of 3D_CRF by comparing ResUNet and ResUNet-2D_CRF. 3) We evaluate 3D connected component analyzing 3D_CCA by remaining different numbers the connected components and comparing with other state-of-the-art methods.

We use Dice value from LiTS competition as our merit to evaluate the accuracy of liver segmentation. Dice value is similar to MIOU, which is an effective measure value for semantic segmentation in natural images. The formula to calculate Dice is defined as equation (1). The higher the Dice score is, the better the segmentation performance is.

B. *Results of ResUNet_DCE*

This paper introduces a semantic segmentation model ResUNet with context multi-images input for liver segmentation in a CT sequence. We adopt U-Net structure with residual unit and batch normalization layer to make the network deeper and quick convergence. For quick training and converging, we propose a new loss function DCE which combines Dice loss and Cross-Entropy loss.

To evaluate the performance of ResUNet model with DCE loss, we compare the semantic segmentation results of ResUNet with DCE loss and Cross-Entropy loss. And we also compare different training iterations of DCE loss and Cross-Entropy loss. We use Adam optimizer from [15] to minimize the loss and mini-batch strategy to train model. In our experiment, we use five images as ResUNet input and set the background and liver class weight for cross entropy by 0.2 and 1.2 in equation (2) to reduce the influence of imbalance of background and liver pixels, and set $\epsilon = 0.5$ for DCE loss in equation (3).

TABLE II. Segmentation results with ResUNet and different loss

Model	Iteration Times	Dice
ResUNet_CrossEntropy	≈400,000	0.906
ResUNet_DCE	**≈200,000**	**0.919**
ResUNet_DCE	≈400,000	0.918

TABLE III. Segmentation results using 2D_CRF and 3D_CRF

Model	Dice
ResUNet_DCE	0.919
ResUNet_DCE +2D_CRF	0.943
ResUNet_DCE +3D_CRF	**0.948**

TABLE IV. Results using 3D_CCA with different components

Model	Dice
ResUNet_DCE+3D_CRF	0.948
ResUNet_DCE+3D_CRF+3D_CCA_1	0.958
ResUNet_DCE+3D_CRF+3D_CCA_2	0.959
ResUNet_DCE+3D_CRF+3D_CCA_3	**0.960**

TABLE V. Results of liver segmentation with state of art methods

Method	Dice
Ours	0.960
Li [7]	0.960
Han [11]	**0.961**

Table II gives the liver segmentation results using ResUNet model with different loss functions. With the same iteration of 400,000, the ResUNet model with Dice loss gets 0.918 Dice score, which is higher than the ResUNet with Cross-Entropy loss with 0.906 Dice value. And at iteration of 200,000, ResUNet with DCE loss gets 0.919 Dice value, which is similar to using twice iteration. DCE loss is helpful to improve liver segmentation results and easier to converge with less time for training than Cross-Entropy loss.

C. *Results of ResUNet_DCE+3D_CRF*

After above liver segmentation by ResUNet_DCE model, we can get the initial segmentation results. To make use of adjacent and context information of 2D image and 3D volume of CT scans, we further use full connection 3D_CRF to refine the segmentation results. We test our methods with ResUNet_DCE with 2D_CRF, which only use 2D context information without exploring correlating information from adjacent CT images.

Table III gives the segmentation results. ResUNet_DCE with 2D_CRF gets 0.943 Dice value, which improves 0.024 compared with ResUNet_DCE. While ResUNet_DCE with 3D_CRF gets 0.948 Dice value, which has 0.005 promotion than using 2D_CRF. For 3D_CRF considers not only the position and Hounsfield value of the pixel pair in the same 2D CT image, but also the pixels between the adjacent images in CT volume, which can capture more context information of liver regions.

D. *Results of ResUNet_DCE+3D_CRF+3D_CCA*

For every liver region is a 3D object in CT sequence, to further reduce noise we use 3D connected component analyzing 3D_CCA after 3D_CRF. In our paper, we test the segmentation performance with remaining different numbers of largest components including 1, 2 and 3 ones.

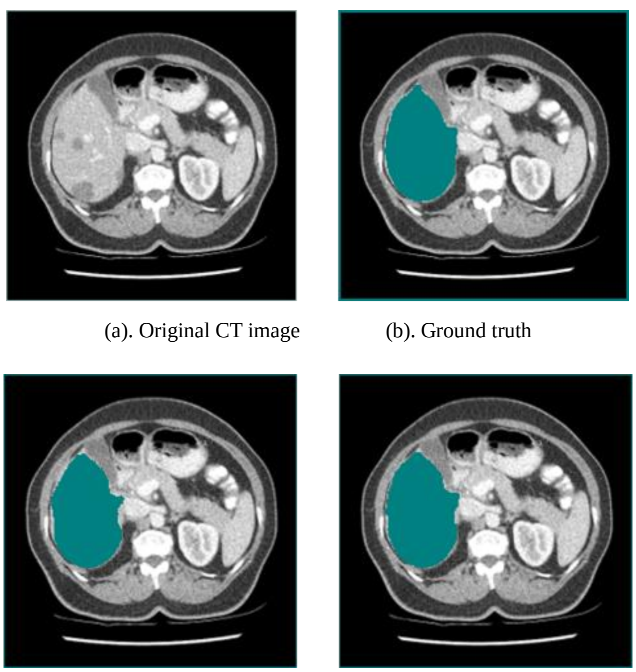
Table IV shows 3D_CCA with one largest component has 0.01 promotion compared with without 3D component analyzing. And all the results are improved by 3D_CCA post-processing and 3D_CCA with three largest components obtains the best Dice value of 0.960. This result maybe after the processing of 3D_CRF, there are still many noises components and many disconnected liver regions. Remaining three largest connected components can reduce some small noise region and reserve some independent liver regions, which can improve segmentation performance.

We also compared the performance of our framework for liver segmentation with other state of the art methods on LiTS dataset. Li et. al. [7] trained a fine-tuned H-DenseNet and reach the 0.960 Dice value. Han et. al. [11] built a three-class U-Net model with residual unit and 3D connect component analyzing and save the largest connected component as the liver region which can get 0.961 Dice value. From the liver segmentation results in Table V, we can find our proposed framework can get the state-of-the-art performance on public dataset LiTS, which shows the effectiveness of proposed framework.

Figure 4 gives the liver segmentation results on one CT image and the green pixels denote the liver region. We can see the liver contour is similar to the ground truth label using our proposed framework than only using ResUNet model.

IV. CONCLUTION

This paper proposes an automatic liver segmentation framework using ResUNet with 3D_CRF and 3D_CCA as a post process. Residual unit, batch normalization layer and DEC loss for U-Net can effectively prevent the gradient disappear and get quick converge. Five continuous images for model input can make use of adjacent images. Fully connected 3D_CRF can explore the context information of adjacent pixels between 2D images and 3D CT volume. Finally, 3D connected components analysis can remove some noise prediction. The proposed framework gets the state-of-the-art performance for liver segmentation on the public dataset. In the future, we will improve our methods for liver tumor segmentation and evaluate our methods on more CT scans from the hospital.



(a). Original CT image (b). Ground truth

(c). ResUNet_DCE (d). ResUNet_DCE+3D_CRF

Figure 4. A sample of segmentation results.

ACKNOWLEDGMENT

This work is supported in part by Beijing Natural Science Foundation (4172058) and Beijing Haidian Original Innovation Joint Foundation (L182054).

REFERENCES

[1]. J. Huang, et al., “Based on Statistical Analysis and 3D Region Growing Segmentation Method of Liver”, in ICACC, 2011

[2]. Baâzaoui, A., et al. "Semi-automated segmentation of single and multiple tumors in liver CT images using entropy-based fuzzy region growing.", in IRBM, 2017

[3]. Yu, Changqian, et al. "Learning a Discriminative Feature Network for Semantic Segmentation.", in CVPR, 2018

[4]. Zhang, Hang, et al. "Context encoding for semantic segmentation", in CVPR, 2018

[5]. Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." , in MICCAI, 2015.

[6]. Christ, Patrick Ferdinand, et al. "Automatic liver and lesion segmentation in CT using cascaded fully convolutional neural networks and 3D conditional random fields.", in MICCAI, 2016.

[7]. Li, Xiaomeng, et al. "H-DenseUNet: Hybrid densely connected UNet for liver and liver tumor segmentation from CT volumes.", in arXiv preprint arXiv:1709.07330, 2017.

[8]. Christ, P. F., F. Ettlinger, and G. Kaissis. "LiTS: liver tumor segmentation challenge. 2017.", 2017.

[9]. Soler, L., et al. "3D Image reconstruction for comparison of algorithm database: A patient specific anatomical and medical image database.", in 2010.

[10]. Ioffe, Sergey, and Christian Szegedy. "Batch normalization: Accelerating deep network training by reducing internal covariate shift.", in arXiv preprint arXiv, 2015.

[11]. Han, Xiao. "Automatic Liver Lesion Segmentation Using A Deep Convolutional Neural Network Method." arXiv preprint arXiv:1704.07239, 2017.

[12]. Milletari, Fausto, Nassir Navab, and Seyed-Ahmad Ahmadi. "V-net: Fully convolutional neural networks for volumetric medical image segmentation.", in 3DV, 2016.

[13]. Krähenbühl, Philipp, and Vladlen Koltun. "Efficient inference in fully connected crfs with gaussian edge potentials.", in NIPS, 2011.

[14]. Tarjan, Robert Endre. "Efficiency of a good but not linear set union algorithm.", in JACM, 1975

[15]. Kingma, Diederik P., and Jimmy Ba. "Adam: A method for stochastic optimization.", in arXiv:1412.6980, 2014

View publication stats