A Cascaded Two-step Approach For Segmentation of Thoracic Organs

S. Kim[(a)], Y. Jang[(b)], K. Han[(b):
H. Shim[(c)], and H. J. Chang[(c,d)*]*

Graduate program in Biomedical Engineering The Graduate School Yonsei University [(a)]
Brain Korea 21 PLUS Project for Medical Science, Yonsei University [(b)]
Integrative Cardiovascular Imaging Research Center, Yonsei University College of Medicine [(c)]
Division of Cardiology, Severance Cardiovascular Hospital, Yonsei University College of Medicine [(d)]

ABSTRACT

Segmentation of thoracic organ is a challenging task as demonstrated by the SegTHOR challenge. In this paper, we present an efficient yet simple framework for automatic thoracic organ segmentation. Two steps are included: first, we designed a simple network to define the ROI of the input volume. Second, we propose a network based on the encoder and decoder model. Three orthogonal views are fed into the network, respectively. To do the final segmentation, we used ensemble by majority voting. The experiments show that our approach is comparable with other segmentation networks.

Index Terms — Thoracic organs, segmentation, CT, Deep learning

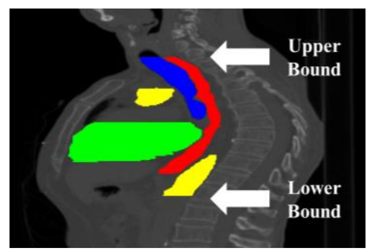
1. INTRODUCTION

Segmentation of thoracic organs in Computed Tomography (CT), which is widely used for planning treatment, is a necessary step for the computer aided diagnosis and computer assisted treatment for extracting an anatomical structure. The automatic segmentation of esophagus which has a small size and variant shape, is especially challenging compare to organs which have a large size. It is a difficult task both for the physician and for automatic algorithms to obtain an accurate and consistent segmentation result.

Early studies of organ segmentation based on atlas-based segmentation, region growing method have been proposed for organ segmentation. [1, 2] But the atlas-based segmentation approach easily affected by the atlas and registration method and the region growing segmentation is a semi-automatic method which depends on the proper location of seed points to obtain robust performance.

In recent years, Convolutional Neural Network (CNN) method has shown its efficiency in segmentation, detection, and reconstruction. Many studies of the organ segmentation based on CNN are proposed. P.Hu et al. [3] proposed a 3D CNN based method incorporating with level-set algorithm to efficiently optimize the energy function. G.Chlebus et al. [4] proposed 2D CNN based method with a shape-based post processing for liver segmentation. H.R.Roth et al. [5] employed both image patches and regions by using a P-ConvNET for patch classification and R-ConvNET for region classification to segment pancreas.

Fig. 1. Examples of SegTHOR dataset which show that CT image with corresponding ground-truth. The ground-truth has an upper and lower bound which gives ambiguity in z-axis in delineating a ground-truth.



In this work, we propose CNN based cascaded two-step approach for automatic multi-organ segmentation on CT images. It is not only training the anatomic-based clinical guideline but also segmenting the interesting slices to boost the performance. Experimental result on the dataset from the 2019 Segmentation Thoracic organ challenge shows our method is comparable to other segmentation methods.

2. MATERIALS

2.1 Data

Let $I \in R^{512\times512\times1}$ denote the input slice with corresponding ground-truth labels $Y \in R^{512\times512\times c}$, where c denotes the number of class (esophagus, heart, trachea, and aorta). The CT scans consist of anisotropic dimensions with high variation in z direction. The in-plane resolution range varies from 0.90mm and 1.37mm, z-resolution from 2mm and 3.7mm. All the imaging datasets are delineated manually following the guidelines established by the Radiation Therapy Oncology Group (RTOG). [6]

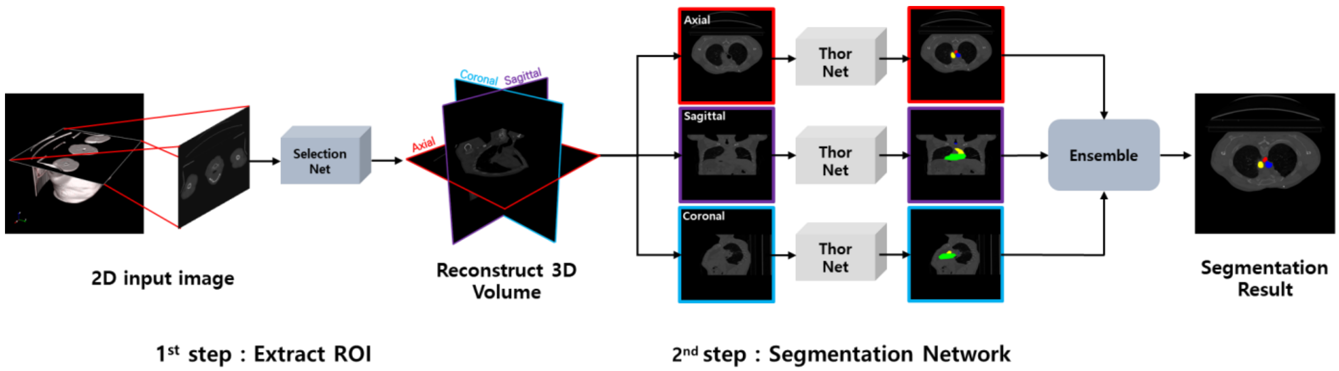


Fig.2. Framework of cascaded two-step approach

2.2 Data Preprocessing

In order to segment of four different thoracic organs, we slice the 3D CT scan along three orthogonal planes (axial, sagittal, and coronal). The intensity of each slices is limited by 1500 HU and normalized by subtracting its minimum value and divide the difference between maximum and minimum value. Each dataset consists of $512\times512\times z$ (range from 150 to 284), axial slice is 512×512 while other slices are $512\times z$ and $z\times512$. The sagittal and coronal slices are padded with a Hounsfield unit of air (-1000 HU) to isotropic images to facilitate segmentation network. Due to the significant imbalance of the class pixels, the slice which has at least one organ is included in the ground truth, is considered for training our network.

3. METHODS

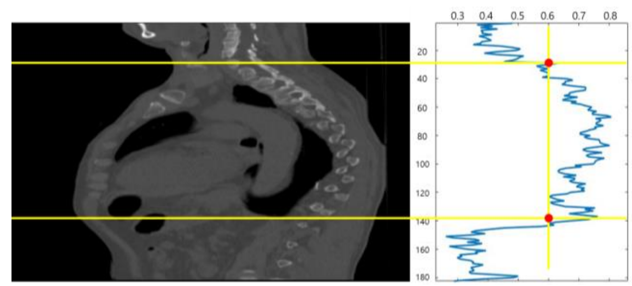
We propose a cascaded two-step approach for thoracic organ segmentation. Fig. 2 shows the framework of the proposed approach. First, the selection of slice with a simple CNN network. Second, the segmentation of thoracic organs with segmentation network with ensemble method. After the selection of slice, it is excluded or retained from the original image and is padded to isotropic 3D volume. The simple 2D CNN network was used for selection of slice which determines whether the slice to be considered or not. The segmentation network with ensemble provides a robust segmentation result with less variance. The final results are post-processed using a 3d connected component analysis.

3.1 Selection network

The purpose of selection network is to reduce the risk for segment the slice which should not be considered by guideline. The design of the selection network is as follow. It consists of 4 layers with 2D CNN and a last layer with fully connected layer. Each layer has 2D convolution with 3×3 kernel, ReLU function, Batch-normalization and Max-pooling.

The Mean square error loss was used for two classes: to be considered or not. The probability map is calculated by softmax in each slice as shown in Fig. 3.

Fig. 3. The probability maps for each slice in z-direction. The threshold value of 0.6 is considered as criteria to select the slice. The probability map is shown parallel to z axis (Right)



We threshold at 60% of probability map to extract a ROI in z direction. After the reconstruction of 3D volume, it is fed to the segmentation steps based on the result.

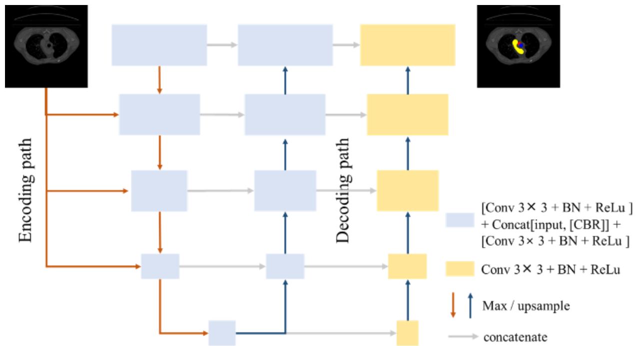
3.2 Segmentation network

We designed the encoder-decoder based CNN named Thor network which is designed for pixel-wise segmentation of the four thoracic organs along three orthogonal planes including axial, sagittal, and coronal at the 2D slice level.

Thor network consists of two types of block; type-1 block contains a 3×3 convolution, batch-normalization and ReLU activation. Type-2 block has concatenation of the feature from the type-1 block and input feature then feed to another type-1 block. In encoding path, we use type-1 block to minimize the loss of local information by concatenating pooled feature. The feature dimensions are decreasing by max pooling layer which leads a large receptive field. In contrary, the network use both type-1 and type-2 blocks and has upsampling layer to get the probability of each pixel in decoding path. We incorporate a shortcut path to concatenate the features from the encoder to the decoder path to compensate localization features and to recover detail feature from organ.

We use a Softmax Cross-Entropy loss to train our Thor-network. An ensemble of same models which is trained on different orthogonal views can improve the segmentation performance by a majority voting. Thor networks in each path are trained respectively.

Fig. 4. Architecture of Thor network. (Note that CBR means Convolution layer + Batch normalization + ReLU function.)



4. EXPERIMENTS AND RESULTS

4.1 Experiments

We evaluate our method using testing data provided from SegTHOR 2019 challenge. The training set composed of CT images from 40 patients and test set comprises of images from 20 patients. We found that the images are highly imbalanced problem. We only trained each network with the slice paired with the labelled mask which contains one organ at least. The total amount of slice is 4,650 in axial, 5,700 in sagittal and 6,205 in coronal, respectively. Our network was trained from scratch with a random weight initialization.

The experimental setup for segmentation of thoracic organs is as follows. We trained our network for 200 epochs with the batch size of 15. We used an Adaptive Moment Estimation (ADAM) which combines RMSProp and Moment to stabilize the gradient descent process with epsilon = 10^{-10}. Other parameters are as follows. Learning rate = 10^{-3} and decay = 0.9.

4.2 Results and Quantitative analysis

We show a sample results from the SegTHOR 2019 validation set in Fig. 5. The results are uploaded to the SegTHOR website to evaluate the performance of our method. The SegTHOR competition has two metrics as the accuracy indexes of the validation results. The *Dice similarity coefficient (DSC)* and *Hausdorff distance* between the two surfaces of A and B which is represented as following expressions:

$$DSC(G,S) = \frac{2|G \cap S|}{|G|+|S|} \quad (1)$$

$$HD(G,S) = \max(h(G,S), h(S,G)) \quad (2)$$

Table 1. shows that the DSC and Hausdorff distance for each class. It can be observed that the segmentation result of our framework seems to be underestimate the esophagus which have various shape and size.

Fig. 5. Qualitative results on sample dataset. The first column presents the input images of CT. The second column shows its corresponding segmentation results.

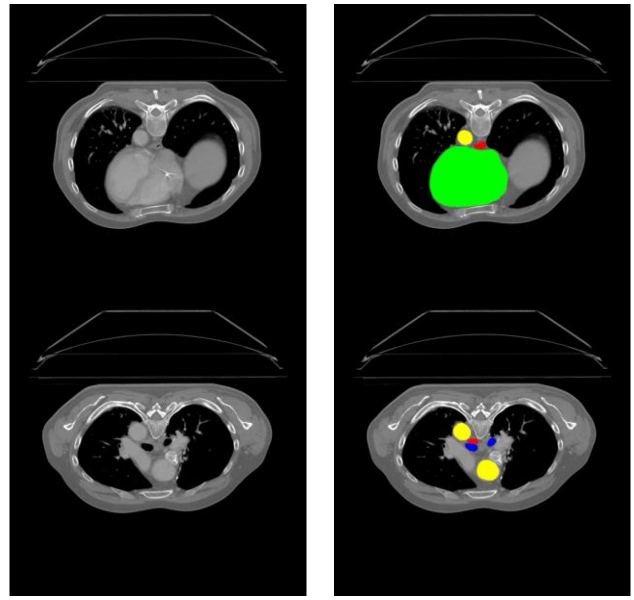


Table. 1. Summary of the results of validation dataset.

Esophagus		Heart	
DSC	**HD**	**DSC**	**HD**
0.7518	0.9267	0.9328	0.2184
Trachea		**Aorta**	
DSC	**HD**	**DSC**	**HD**
0.8885	0.6164	0.8919	1.1300

5. CONCLUSION

In this work, we have described a CNN based architecture for automatic thoracic organ segmentation. We propose a cascaded two-step approach that shows a comparable performance with other networks. We incorporate a simple classification network to choose the axial slice and ensemble method to stabilize our segmentation results. Using a majority voting based combiner, we achieve DSC of 0.7518, 0.9328, 0.8885, and 0.8919 Hausdorff of 0.9267, 0.2184, 0.6164, and 1.1300. In the future work, we will to improve the performance of small object such as esophagus to improve thoracic organ segmentation task.

References

[1] J. Wu, F. Ye, J.-L. Ma, X.-P. Sun, J. Xu, and Z.-M. Cui, "The segmentation and visualization of human organs based on adaptive region growing method," in 2008 IEEE 8th International Conference on Computer and Information Technology Workshops, 2008, pp. 439-443: IEEE.

[2] Y. Zhou and J. Bai, "Multiple abdominal organ segmentation: an atlas-based fuzzy connectedness

approach," IEEE Transactions on Information Technology in Biomedicine, vol. 11, no. 3, pp. 348-352, 2007.

[3] P. Hu, F. Wu, J. Peng, Y. Bao, F. Chen, and D. Kong, "Automatic abdominal multi-organ segmentation using deep convolutional neural network and time-implicit level sets," International journal of computer assisted radiology and surgery, vol. 12, no. 3, pp. 399-411, 2017.

[4] G. Chlebus, A. Schenk, J. H. Moltz, B. van Ginneken, H. K. Hahn, and H. Meine, "Deep learning based automatic liver tumor segmentation in CT with shape-based post-processing," in International conference on Medical Imaging with Deep Learning (MIDL), 2018.

[5] H. R. Roth et al., "Deeporgan: Multi-level deep convolutional networks for automated pancreas segmentation," in International conference on medical image computing and computer-assisted intervention, 2015, pp. 556-564: Springer.

[6] Roger Trullo, Caroline Petitjean, Su Ruan, Bernard Dubray, Dong Nie, and Dinggang Shen. Segmentation of organs at risk in thoracic CT images using a sharpmask architecture and conditional random fields. In IEEE International Symposium on Biomedical Imaging (ISBI), pp. 1003-06, 2017.