Lung Tumor Segmentation by Fusing 2D and 3D Models

Minh-Trieu Tran, Soo-Hyung Kim, Hyung-Jeong Yang and Guee-Sang Lee

April 16, 2020

Lung tumor segmentation by fusing 2D and 3D models

Minh-Trieu Tran[0000-0002-5015-5604] , Soo-Hyung Kim[0000-0003-3575-5035] , Hyung-Jeong Yang[0000-0003-3024-5060] and Guee-Sang Lee[0000-0002-8756-1382]

Chonnam National University, South Korea
tmtvaa@gmail.com

Abstract. In this study, an automatic lung tumor segmentation model is proposed. Over the past decade, deep neural networks have shown many strengths in the field of image segmentation. Especially in the field of medical imaging, many network models have been developed in recent years. These networks mostly use data for training as 2-dimensional (2D) data because they do not require a lot of hardware resource. However the 3-dimensional (3D) structure information of the tumor are not fully utilized, so there are also number of studies using 3D models. Our proposed method is a fusion model from 2D and 3D model. The results of proposed model are evaluated on the data set of task 4 of the StructSeg competition with very positive results compared to other methods.

Keywords: U-net, Residual Network, Lung Tumor Segmentation.

1 Introduction

Today with the development of globalization, the industry in many places flourished. The progress of science and technology has been applied in daily life, increasing the living standard of people. However, accompanied by the development of industrialization are pollution problems, as well as many dangerous incurable diseases. Prominent among them is lung cancer. This is one of the most dangerous diseases in the world with a very high mortality rate, more than 50%. The early detection and diagnosis of lung tumors has been interested by many scientists over the years. There are two types of tumors are benign and malignant. If a patient has a malignant tumor, they have cancer. Also for benign tumors, if it is detected and diagnosed late, the sizes of tumor is increased as well as the ability to turn to malignant tumors. From these problems, many scientists have studied methods of detecting and diagnosing lung tumors. Most diagnostic methods are based on computed tomography (CT) images. Doctors can diagnose easier patient's health condition when the tumor images are segmented automatically. Most of the medical image processing competition focus on segmentation [1]. Also, if the diagnosis is done manually, it greatly depend on the doctor's experience, which take a long time to identify the tumor. In this paper, a deep learning model is proposed based on the using of a combination of both 2D and 3D data for training.

2 Related Works

In recent years, the widespread development of convolutional neural networks have been applied in variety of tasks such as in computer vision and medical imaging. With the potential of a deep learning approach, many recently developed models serve the purposes of image analysis as well as image segmentation. Today, image data can be classified into two categories: 2D and 3D images. Each has its own advantages in training the model. Some studies focus on the using of 2D images because it has the advantage of rapid training, requiring low computational hardware. J. Cai et al. [2] used regression neural networks for pancreas segmentation from CT images. Ronneberger [3] and colleagues proposed a U-net model in the biomedical images segmentation. Aganj et al. [4] proposed a method of unsupervised learning in medical CT images segmentation. Besides the aforementioned benefits, the drawback of 2D data training is the omission of the object's 3D structure information. A number of studies using 3D data for network training have been proposed and developed. Dolz et al. [5] proposed a 3D convolution model for brain magnetic resonance imaging (MRI) segmentation. Andermatt and his colleagues [6] developed a 3D regression neural network model for segmenting MRI images.

3 Proposed Method

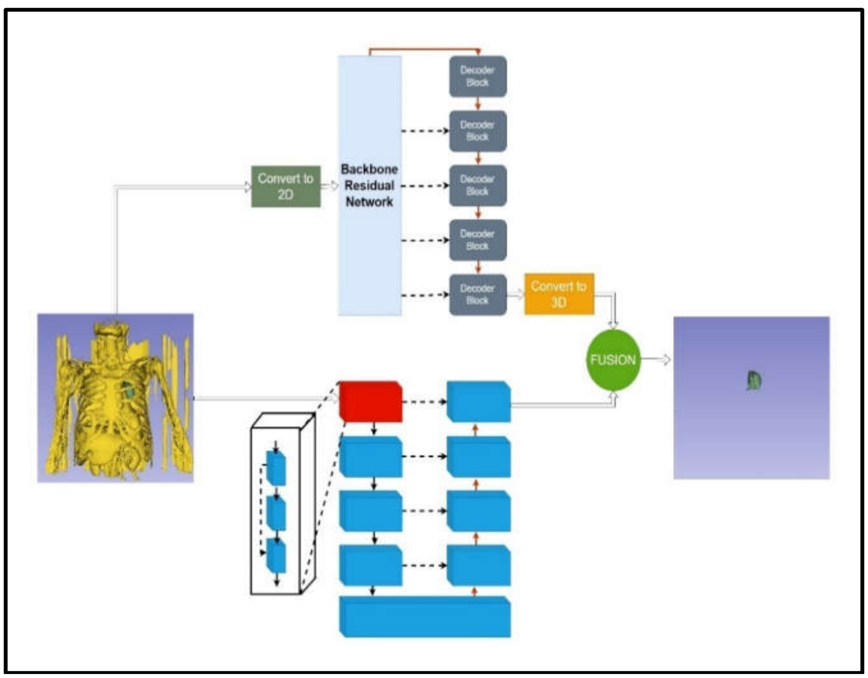

Fig. 1. The architecture of our proposed method.

Understanding the advantages of network training with 2D and 3D images, the proposed method used both 2D and 3D images from the data set for training. The network is a fusion model from separate subnets. The first is a subnet that uses 2D image data for training. With this model, the U-net network is used as the basis. The encoder part of this model used residual network backbone with pretrained weight. This pretrained weight is trained with the Imagenet [7] dataset. The model is able to learn the characteristics of data more optimally because of using pretrained weight. The second sub-net is a 3D model, we use the 3D U-net [8] model with the residual networks is added to the first block encoder for the purpose of learning features more effectively [9], while minimizing the possibility of overfitting phenomena in the model. Although the original image data in the task 4th of the StructSeg competition is the 3D images, 2D images were created using the simpleITK conversion library. This 2D image is obtained from the axial projection. The output images of the 2D subnet are reconstructed into 3D images by the simpleITK library. Then the two 3D images at the two outputs of the two subnets are fusion by taking the smallest pixel value and this produces better results. Fig. 1 shows the model of the proposed method.

4 Experimental Results

4.1 Dataset and training details

With the aim of creating 2D data for training the model, the original 3D images are converted to 2D images using the SimpleITK library. There are 50 3D images and 4775 2D images. The sizes of the 2D image are 512x512. The data set evaluated during this study is taken from task 4th of the StructSeg competition organized by MICCAI in 2019. The dataset is divided into two separate sections: training and validation. Data of the first 40 patients are used as a training dataset while the remaining 10 patients are used as a validation dataset. The model trained with 30 epochs and the Adam optimizer used to train.

4.2 Dice score coefficient

The Dice Score Coefficient (DSC) measures the overlap between label and segmentation output results. DSC function is shown below:

$$Dice = \frac{2TP}{2TP+FP+FN}, \tag{1}$$

where TP is the number of the tumor pixels and they are predicted as tumor pixels, FP is the number of background pixels but they are predicted as tumor pixels, and FN is the number of the tumor pixels but they are predicted as background pixels.

4.3 Results

The proposed method has shown superiority over other methods. Our proposed method makes the segmentation of lung tumor images more accurate, so it can be studied and applied in the medical field later. Table 1 shows the outstanding Dice score value of our model with 0.3777 while the model using 2D data only has a Dice score of 0.1576 and the model using only 3D data is only 0.3170. Fig. 2 shows lung tumor segmentation results between our proposed method with other methods

Table 1. Dice score results

Method	Dice score
2D U-net with backbone residual network	0.1576
3D U-net	0.2063
3D U-net with residual block [10]	0.3170
Proposed method	0.3777

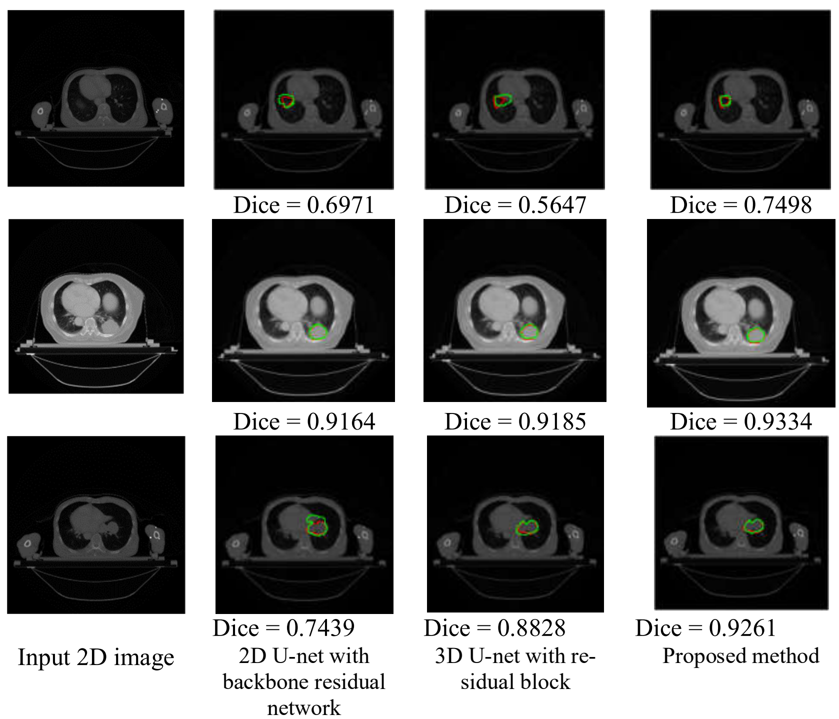

Fig. 2 Lung tumor segmentation results between our proposed method with other methods.

5 Conclusion

To summarize our work, we propose a fusion model for lung tumor segmentation which includes the 2D model and 3D model. Our model is fully utilized 3D information of the 3D image and achieves higher results than the independent 2D model. Our proposed method is evaluated on the dataset of task 4 of the StructSeg competition and produced very positive results. This study opens up many directions for medical imaging segmentation and helps doctors in early detection and diagnosis of patients who have lung tumors. Thereby increasing the survival rate of lung cancer patients.

Acknowledgment

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF - 2018R1D1A3B05049058 & NRF – 2017R1A4A1015559) and by a grant (HCRI 19 001-1*H2019-0295) Chonnam National University Hwasun Hospital Institute for Biomedical Science and by the Bio & Medical Technology Development Program of the NRF funded by the Korean government (MSIT) (NRF-2019M3E5D1A02067961). The corresponding author is Guee-Sang Lee.

References

1. Maier-Hein, L., Eisenmann, M., Reinke, A., Onogur, S., Stankovic, M., Scholz, P. , Arbel, T., Bogunovic, H., Bradley, A., P., Carass, A., Feldmann, C., Frangi, A., F., Full, P., M., van Ginneken, B., Hanbury, A., Honauer, K., Kozubek, M., Landman, B., A., Mark, K., Maier, O., Maier-Hein, K., Menze, B., H., Muller, H., Neher, P., F., Niessen, W., Rajpoot, N., Sharp, G., C., Sirinukunwattana, K., Speidel, S., Stock, C., Stoyanov, D., Taha, A., A., van der Sommen, F., Wang, C., W., Weber, M., A., Zheng, G., Jannin, P., Kopp-Schneider, A.: Why Rankings of Biomedical Image Analysis Competitions Should Be Interpreted With Care. Nature Communications (2018).
2. Cai, J., Lu, L., Xing, F., Yang, L.: Pancreas Segmentation in CT and MRI Images via Domain Specific Network Designing and Recurrent Neural Contextual Learning. In: Arxiv:1803.11303v1 (2018).
3. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional Networks for Biomedical Image Segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), 234-241 (2015).
4. Aganj, I., Harisinghani, M., G., Weissleder, R., Fischl, B.: Unsupervised Medical Image Segmentation Based on the Local Center of Mass. Scientific Reports (2018).
5. Dolz, J., Desrosiers, C., Ayed, I., B.: 3D Fully Convolutional Networks for Subcortical Segmentation in MRI: A Large-Scale Study. NeuroImage, 456-470 (2018).
6. Andermatt, S., Pezold, S., Cattin, P.: Multi-dimensional Gated Recurrent Units for the Segmentation of Biomedical 3D-Data. In: Deep Learning and Data Labeling for Medical Applications (2016).

7. Deng, J., Dong, W., Socher, R., Li, Li, L., K., Fei-fei, L.: ImageNet: A Large-scale Hierarchical Image Database. In: Conference on Computer Vision and Pattern Recognition, 248-255 (2009).
8. Cicek, O., Abdulkadir, A., Lienkamp, S., S., Brox, T., Ronneberger, O.: 3D U-net: Learning Dense Volumetric Segmentation from Sparse Annotation. In: Arxiv:1606.06650v1 (2016).
9. He, K., Zhang, X., Ren, S., Sun, J.: Deep Residual Learning for Image Recognition. In: Conference on Computer Vision and Pattern Recognition (2016).
10. Tran, M., T., Kim, S., H., Yang, H., J., Lee, G., S.: Utilizing 3D Information from Three-Dimensional Images in Lung Tumor Segmentation. In: International Conference on Smart Media and Applications (2019).