Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 192 (2021) 543–552

Procedia Computer Science

www.elsevier.com/locate/procedia

25th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems

MAU-Net: Multiple Attention 3D U-Net for Lung Cancer Segmentation on CT Images

Wei Chen[a], Fengchang Yang[b], Xianru Zhang[c], Xin Xu[a,*], Xu Qiao[c,*]

[a]*School and Hospital of Stomatology, Cheeloo College of Medicine, Shandong University & Shandong Key Laboratory of Oral Tissue Regeneration & Shandong Engineering Laboratory for Dental Materials and Oral Tissue Regeneration, Jinan, China*

[b]*Department of Radiology, Shandong Cancer Hospital and Institute, Cheeloo College of Medicine, Shandong University, Jinan, China*

[c]*Department of Biomedical Engineering, School of Control Science and Engineering, Shandong University, Jinan, China*

Abstract

Accurate segmentation of lung cancer from computed tomography (CT) is of great significance to constructing an automatic diagnosis system for lung cancer. This paper presents multiple attention 3D U-Net (MAU-Net), a novel deep learning-based architecture for lung cancer segmentation from CT images. In particular, we first apply a dual attention module at the bottleneck of the U-Net that models the semantic interdependencies in spatial and channel dimensions, respectively. A novel multiple attention gate module is then proposed to adaptively recalibrate and fuse multiscale features from the dual attention module, the previous decoder feature maps, and the corresponding features from the encoder. Extensive ablation studies on a clinical dataset consisting of 322 CT images demonstrate the effectiveness of our proposed method. Our model achieved an average Dice similarity coefficient, 95% Hausdorff distance and relative absolute volume difference of 0.8667, 13.0036, and 0.1552, respectively.

Keywords: Type your keywords here, separated by semicolons ;

1. Introduction

Lung cancer is one of the most common cancers. In 2020 alone, there are an estimated 2.2 million new cases of lung cancer and 1.8 million deaths. It is the second most commonly diagnosed cancer in 2020 and the leading cause of cancer deaths [1]. About 80% of lung cancer patients cannot be treated with surgery, and chemotherapy accompanied by radiotherapy is the main treatment [2]. Radiotherapy and efficacy evaluation requires accurate positioning of the tumor location. Computed tomography (CT) is the most commonly used imaging modality for diagnosis, radiation planning, and prognosis. Routinely, the treatment target is manually delineated on CT scans by experts and then transferred to

* Corresponding author.
E-mail address: xinxu@sdu.edu.cn (X.X.), qiaoxu@sdu.edu.cn (X.Q.)

1877-0509
10.1016/j.procs.2021.08.056

a treatment planning system to calculate radiation dose [3]. Besides, treatment response can usually be determined by measuring changes in tumor size, usually using one-dimensional measurement methods, such as RECIST v1.1. The recent popular radiomics research also requires tumor segmentation. Therefore, tumor segmentation has a wide range of applications in lung cancer. Manually delineating the tumor volume is labor-intensive, time-consuming and subjective, and has great variability between and within observers. Therefore, automatic segmentation methods are essential to reduce the burden of radiologists and improve the objectivity of the segmentation results.

In recent years, automatic medical image segmentation has made great progress by developing fully convolutional neural network architectures (FCNs) [4]. The representative work is U-Net [5], which is a typical encoder-decoder architecture consisting of an encoder and a decoder and has long skip connections at each resolution level. The main idea is to combine the high-resolution features of the encoder with the up-sampling features of the corresponding decoder to fuse multiscale features. Based on U-Net, recently, Isensee et al. [6] proposed nn-Unet. This deep learning-based segmentation method can automatically configure parameters such as preprocessing, network structure, training strategy, etc., and has achieved state-of-the-art results in many tasks.

Recently, attention mechanisms are widely used in semantic segmentation. Attention mechanisms first appeared in natural language processing and quickly expanded to many computer vision fields such as classification, detection, and semantic segmentation. Attention can be viewed as using information transferred from other feature maps to select the most discriminative part of the input signal [7]. In semantic segmentation , the attention mechanism can help the model focus on salient features and amplify their influence while filtering away irrelevant features. Nie et al. [8] designed an attention model to segment prostate from MRI images with higher accuracy than baseline models, e.g., V-Net and FCN. Sinha and Dolz [9] proposed a multi-level attention-based architecture for abdominal organ segmentation from MRI images. Gu et al. [10] made extensive use of multiple attentions in a CNN architecture and proposed a comprehensive attention-based CNN (CA-Net) for more accurate segmentation.

Many algorithms have been developed for the detection and segmentation of lung nodules. Compared with lung nodules, lung cancer is usually larger and irregular in shape, presenting challenges of automatic segmentation algorithms that developed for lung nodules [11]. Several deep learning-based methods have been proposed in the literature for lung tumor segmentation. Kamal et al. [12] proposed Recurrent 3D-DenseUNet, which consists of a 3D encoder block to extract features, a recurrent convolutional block to extract Spatio-temporal information, and a 3D decoder block to get the segmentation masks. They achieved an average dice score of 0.7228 on the NSCLC-Radiomics dataset of 300 patients. Jiang et al. [13] proposed multiple resolution residually connected network to combine features at different image resolutions and achieved promising segmentation results in several datasets. Zhang et al. [14] introduced a modified version of ResNet and applied it to segment the gross target volume in CT of patients with non-small cell lung cancer. Byun et al. [15] proposed a coupling-net with shape-focused prior and discussed the segmentation results of cancers of different shape types. Although these works have achieved good segmentation results, the accurate segmentation of lung cancer is still an open problem.

In this paper, we proposed a novel method for lung cancer segmentation. Considering that the lung area in chest CT only occupies a part of it, the direct use of chest CT will increase the amount of calculation and affect the segmentation results. Therefore, we first used a public model to segment the lung regions to crop the original chest CT images. We then proposed a Multiple Attention 3D U-Net (MAU-Net) that integrates multiple attention mechanisms to segment lung cancer accurately. In the following section, we will provide a detailed description of our proposed network.

2. Method

The flowchart of our proposed segmentation pipeline is depicted in Figure 1. The flowchart is mainly composed of two steps. The first step involves lung segmentation and preparing the regions of interest (ROIs) for feeding into the lung cancer segmentation network. In the second step, we train the proposed cross-attention network and evaluate the testing set's performance.

2.1. Lung Segmentation

Segmentation of the lung from CT has been studied extensively [16, 17, 18, 19]. Recently, Hofmanninger et al. [20] proposed that the accuracy and reliability of lung segmentation algorithms mainly depend on the diversity of

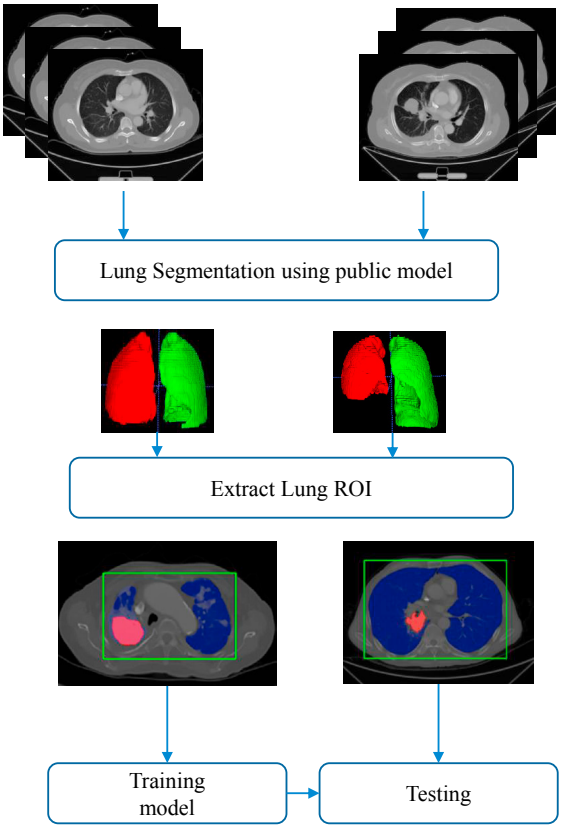

Fig. 1. The flowchart of our proposed segmentation pipeline.

training data, thus emphasizing the importance of data diversity compared with model selection. Their experimental results showed that accurate lung segmentation does not require complicated methods. Once diversified training data is obtained, the segmentation architecture based on deep learning will produce state-of-the-art results. They collected four lung cancer databases, trained a semantic segmentation deep learning model and achieved state-of-the-art results in the testing set. To facilitate research on lung diseases, they released the trained model under General Public License 3.0. Given the innovation and viewpoint of this article, we did not train a new network model for lung segmentation but directly used this trained model [1].

2.2. *Lung Cancer Segmentation*

Figure 2 shows the architecture of our proposed network for lung cancer segmentation. Due to the broad applicability of U-Net [5], we use this structure as the backbone. Besides, we applied a dual attention module (DAM) that includes a spatial attention block (SAB) and a channel attention block (CAB) to make full use of spatial and channel attentions. At the end of the decoder, we proposed a new Multiple Attention Gated Module (MAGM) for the accurate segmentation of lung cancer.

2.2.1. *Base U-Net*

The U-Net is a segmentation network composed of an encoder and a decoder and has long skip connections at each resolution level. The role of the encoder is to extract high-level features on the feature maps of gradually decreasing size, while the decoder uses the encoded features to obtain the segmentation target. We set up 4 levels for the base

[1] https://github.com/JoHof/lungmask

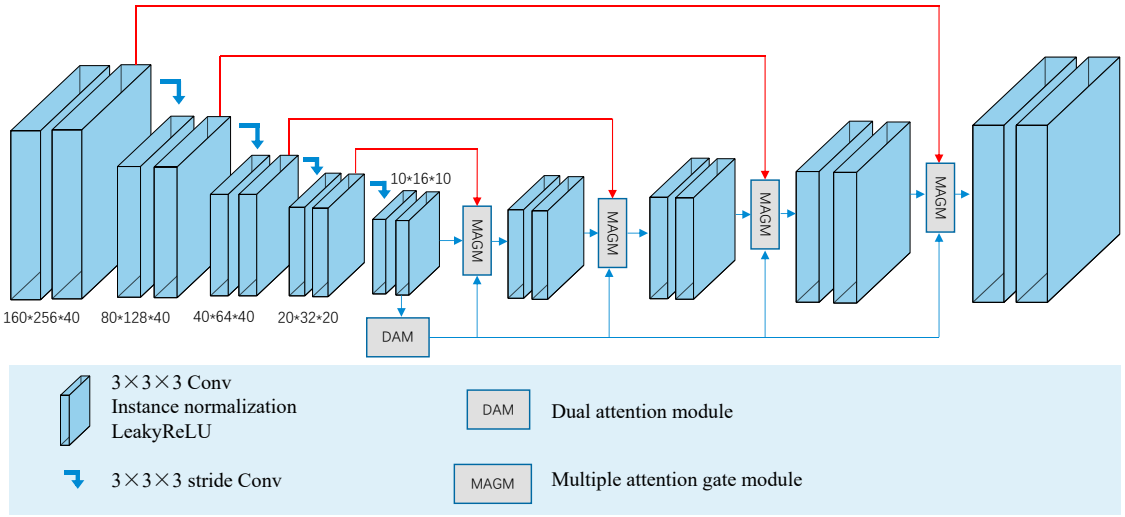

Fig. 2. The architecture of our proposed network.

U-Net. In each level, the feature map's resolution is reduced by stride convolution, and the number of channels is increased at a ratio of 2. The initial input size is $160 \times 256 \times 40$, and the initial number of channels is 32.

2.2.2. *Dual Attention Module*

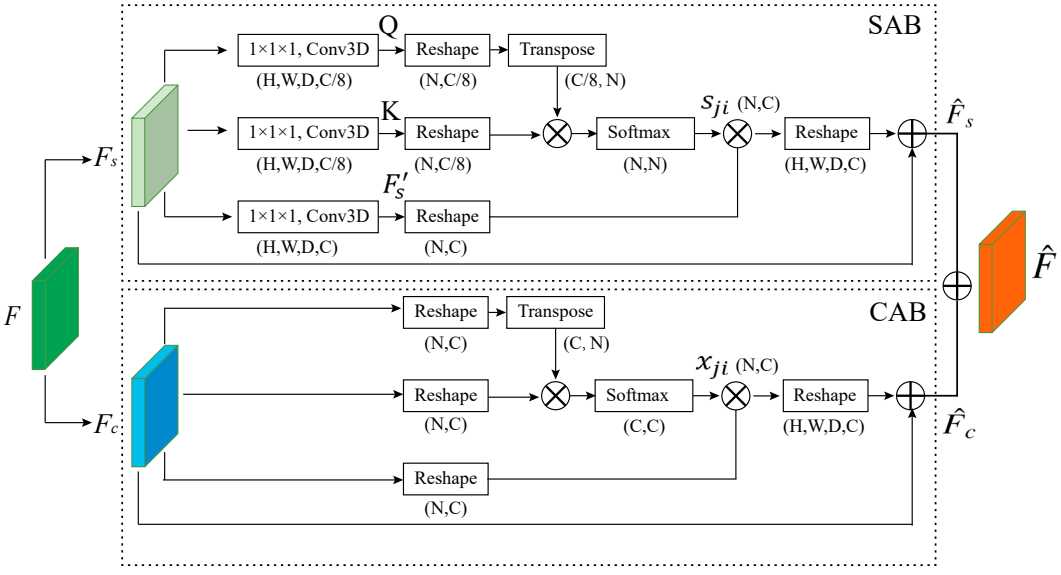

Fig. 3. The architecture of DAM.

At the end of the encoder, a DAM with position and channel attention blocks is integrated to the U-Net. Let $F \in \mathbb{R}^{512\times H\times W\times D}$ be the feature map at the end of the encoder, where 512 is the channel number, and H, W, D represent the height, width and depth, respectively. We first use two parallel $3 \times 3 \times 3$ convolutional layers with 128 channels to obtain two compressed feature maps $F_p \in \mathbb{R}^{128\times H\times W\times D}$ and $F_c \in \mathbb{R}^{128\times H\times W\times D}$. F_p is fed into a spatial attention block (SAB) and obtained $\hat{F}_s$, and F_c is fed into a channel attention block (CAB) and obtained $\hat{F}_c$. $\hat{F}_s$ and $\hat{F}_c$ are then added to get the final output $\hat{F}$. The details of DAM is shown in Figure 3.

Spatial attention block (SAB) For the input feature map F_s, two $1 \times 1 \times 1$ convolutional layers are used to get two feature maps $Q \in \mathbb{R}^{C\times H\times W\times D}$ and $K \in \mathbb{R}^{C\times H\times W\times D}$, respectively. Q and K are then reshaped to $\mathbb{R}^{C\times N}$, where $N = H \times W \times D$. The transpose of Q and K is further fused by a matrix multiplication followed by a softmax layer to generate the attention map:

$$s_{ji} = \frac{\exp(K_i \cdot Q_j^T)}{\sum_{x=1}^{N} \exp(K_i \cdot Q_j^T)} \tag{1}$$

where s_{ji} represents the i^{th} position's impact on j_{th} position. Meanwhile, we apply a convolutional layer to F_s to generate a new feature map $F_s^{'}$ and reshape it to $\mathbb{R}^{C\times N}$. Then, $F_s^{'}$ and the transpose of S are fused by a matrix multiplication and the result is reshaped to $\mathbb{R}^{C\times H\times W\times D}$. Finally, we use a scale learnable parameter α to weight the above result and perform a element-wise sum operation with F_s to obtain the final output:

$$\hat{F}_{sj} = \alpha \sum_{i=1}^{N} (s_{ji} F_{si}^{'}) + F_s \tag{2}$$

Therefore, $\hat{F}_s$ is a weighted sum of the features across all positions and original features. It can selectively highlight useful features in the global contextual view.

Channel attention block(CAB). The role of CAB is to model the relationship between different channels.We first reshape F_c to $\mathbb{R}^{C\times N}$, and perform a matrix multiplication between F_c and the transpose of F_c. Then, a softmax layer is used to obtain the channel attention map $X \in \mathbb{R}^{C\times C}$:

$$x_{ji} = \frac{\exp(F_{ci} \cdot F_{cj})}{\sum_{i=1}^{C} \exp(F_{ci} \cdot F_{cj})} \tag{3}$$

where x_{ji} represents the i_{th} channel's impact on the j_{th} channel. The transpose of X and F_c are then fused by a matrix multiplication and the result is reshaped to $\mathbb{R}^{C\times H\times W\times D}$. Finally, we use a scale learnable parameter β to weight the above result and perform an element-wise sum operation with F_c to obtain the final output:

$$\hat{F}_{cj} = \beta \sum_{i=1}^{C} (X_{ji} F_{ci}) + F_{cj} \tag{4}$$

Therefore, $\hat{F}_c$ is a weighted sum of the features of all channels and can model the long-range semantic dependencies between different feature maps.

2.2.3. Multiple Attention Gated Module

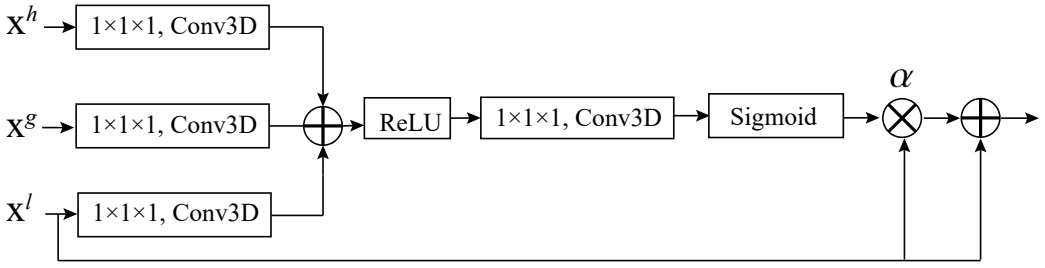

Fig. 4. The architecture of MAGM.

Due to GPU memory limitation and to reduce the amount of calculation, we only added DAM at the end of the encoder. However, in the decoder process, cascading the encoder's corresponding layer directly with the decoder will introduce additional noise, which will affect the final segmentation result. To alleviate this problem, Oktay et al. [21] proposed an attention U-Net based on grid-attention, which performs an attention gating before the cascade operation by combining image features from multiple imaging scales. The attention gate can suppress features in

irrelevant regions and highlight salient features for dense label prediction. Besides, this attention does not introduce many parameters and calculations and can achieve better performance than gating based on global feature vectors. Inspired by the attention U-Net, we proposed a Multiple-attention Gated Module (MAGM), which combines three input information: the features of the corresponding decoder layer, the features from the corresponding encoder layer, and the feature of dual attention, as shown in Figure 4.

Let x^l be the feature map from skip connection, x^g be the feature map from decoder, and x^d be the feature map from DAM. x^l, x^g and x^d are compressed by a $1 \times 1 \times 1$ convolution respectively, and the results are summed and followed by a ReLU activation function. Feature map obtained by the ReLU is then fed into another $1 \times 1 \times 1$ convolution with one output channel followed by a Sigmoid function to obtain a pixel-wise attention coefficient $\alpha \in [0, 1]^{H \times W \times D}$. x^l is then multiplied with α followed by a sum operation to obtain the final output feature maps.

3. Experimental Results

3.1. Dataset

Our institutional review board has approved the study and does not require the patient's informed consent. A total of 322 patients diagnosed with lung cancer from June 2014 to June 2019 were collected retrospectively. All lung CT examinations were performed using four CT scanners; the tube voltage was 120 kV, the tube current was 220 mAs, and the thickness of the interlayer slices was 4–5 mm. For each patient, an experienced radiologist (with eight years of experience) used the Medical Imaging Interaction Toolkit (MITK) software (version 2013.12.0; http://www.mitk.org/) to contour the tumor area. Another experienced radiologist (with 15 years of experience) then confirmed the segmentation. If there were different explanations, the final consensus would be reached through discussion. We divided the dataset into a training set, validation set, and testing set at a ratio of 7:1:2.

3.2. Implementation and Evaluation Methods

All methods were implemented in the Pytorch framework. A hybrid loss that combines Dice loss and cross-entropy loss was used. We used stochastic gradient descent with Nesterov momentum ($\mu = 0.99$) with an initial learning rate of 0.01 for training. We used a batch size of 2 and trained for a total of 300 epochs. We randomly extracted a patch of size $160 \times 256 \times 40$ from each CT image at each training epoch. The learning rate was decayed according to the 'poly' learning rate policy with $(1 - epoch/epoch_{max})^{0.9}$. A variety of data augmentation techniques, including rotation, Gaussian blur, mirroring and so on, were applied on the fly during training. The training was implemented on one NVIDIA Tesla V100 GPU. The best model on the validation set was used for the testing set.

We utilized three evaluation metrics to conduct quantitative evaluation. (i) Dice similarity coefficient (DSC). DSC measures the overlap between a segmentation and the ground truth, which is defined as

$$DSC = \frac{2|S \cap G|}{|S| + |G|} \tag{5}$$

where S and G denote the region segmented by the network and the ground truth, respectively (the higher, the better). (ii) Hausdorff distance (Haus). Haus measures the surface distance between two segmentation boundaries. In specific, given two volumes S and G and their corresponding surface ∂S and ∂G , Haus calculates the shortest least-squares distance $d(s, g)$ between the points s on the ∂S and the points g on the ∂G, and vice versa, finally returns the maximum value over all d

$$\text{Haus}(S, G) = \max\{\sup_{s \in \partial S} \inf_{g \in \partial G} d(s, g), \sup_{g \in \partial G} \inf_{s \in \partial S} d(g, s)\} \tag{6}$$

Directly calculating the maximum over all surface distances will makes the Haus very susceptible to small outlying subregions. Therefore, we used Haus95, which calculates the 95% quantile surface distance between S and G instead of the maximum distance (the smaller, the better). (iii) Relative absolute volume difference (RAVD). RAVD measures the absolute size difference of the segmentation and the ground truth, divides it by the ground truth to compare two volumes in percent (the smaller, the better).

$$RAVD = \left(\frac{|S| - |G|}{|G|}\right) \times 100 \tag{7}$$

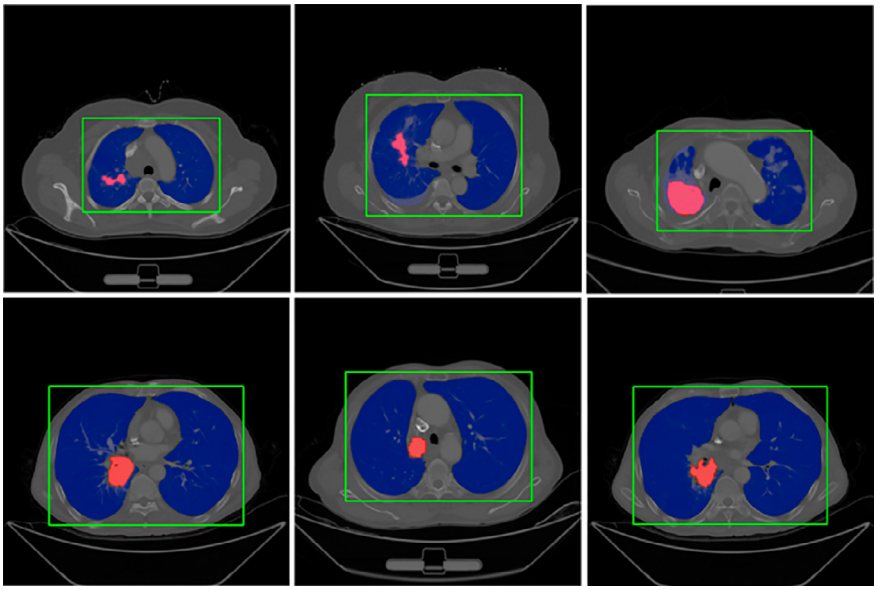

Fig. 5. The results of lung segmentation.

3.3. Lung Segmentation

For all samples, we used the model released by Hofmanninger et al. [20] to segment the lung region, and six randomly selected slices from six subjects are shown in Figure 5. The blue regions represent the lung region, red regions represent the cancer region, and the green bounding boxes represent the ROIs fed to the next lung cancer segmentation stage. It can be seen that all the lung areas have been segmented correctly, and all the cancer areas are contained within the bounding boxes.

3.4. Ablation experiments

Table 1. Ablation of different methods.

Method	Metrics	3D U-Net	p	DAU-Net	p	MDA-UNet
DSC	median	0.9083	0.1030	0.9084	0.3714	0.9131
	mean± std	0.8434 ± 0.1791		0.8538 ± 0.1828		0.8667 ± 0.1316
Haus95	median	4.8665	0.0354	4.3584	0.2388	4.1240
	mean± std	21.8934 ± 40.1581		18.4078 ± 38.7084		13.0036 ± 27.9934
RAVD	median	0.0548	0.1776	0.0449	0.3909	0.0284
	mean± std	0.2240 ± 0.5616		0.1773 ± 0.6872		0.1552 ± 0.5929

In this section, we studied in detail the impact of each component in the proposed network. Specifically, we first performed the original 3D U-Net, which does not contain any attention mechanism. Second, we added the DAM to the 3D U-Net, which we called DAU-Net. Finally, we conducted our proposed MAU-Net. All of these networks were trained using the same training configuration. The segmentation results are listed in Table 1. The base U-Net achieves 0.8434, 21.8934 and 0.2240 in DSC , Haus95 and RAVD, respectively. When we add the DAM, it yields results of 0.8538, 18.4078 and 0.1773 in DSC , Haus95 and RAVD, respectively. DAU-Net significantly outperforms the base

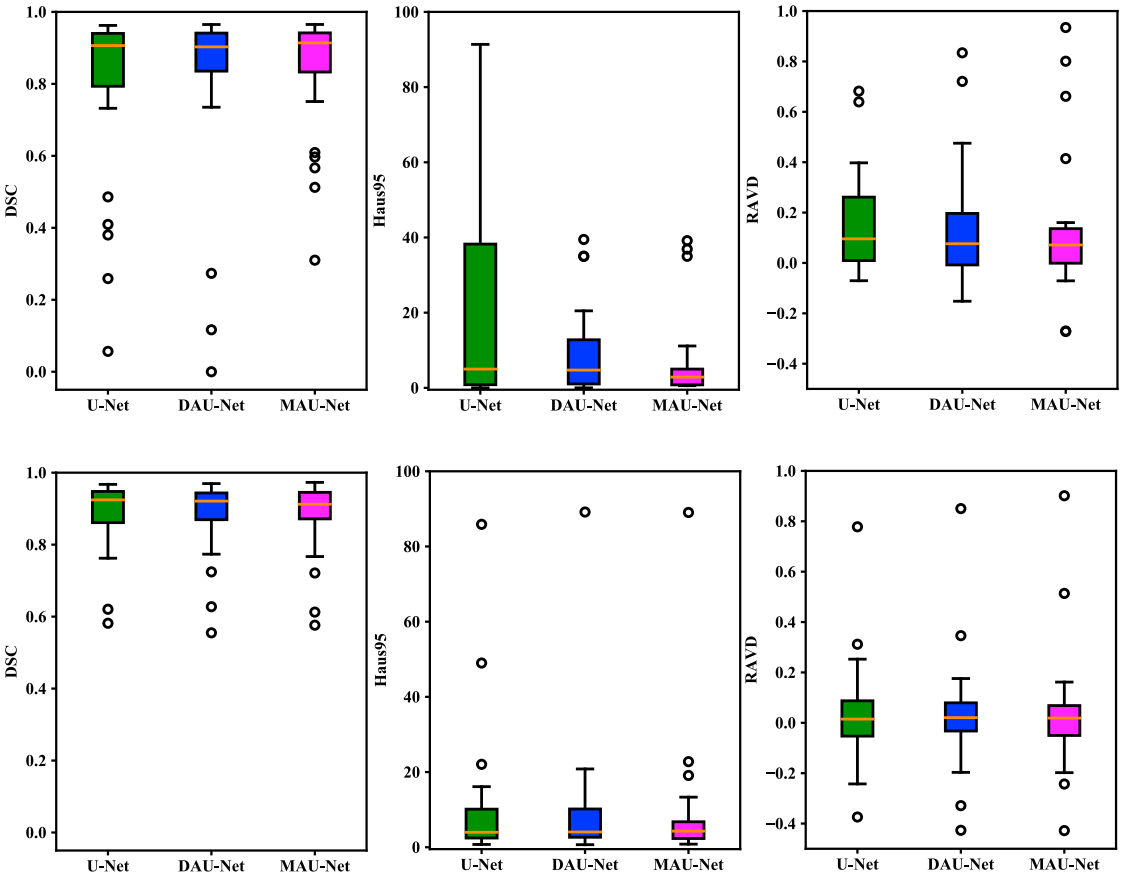

Fig. 6. Comparison of segmentation results of different tumor volume sizes. The first row represents the segmentation results for the group with a small cancer region. The second row represents the segmentation results for the group with a larger cancer region.

U-Net, proving that the attention mechanism is of vital importance for segmentation. With the same training configuration, our proposed MAU-Net achieves performances of 0.8667, 13.0036 and 0.1552 in DSC, Haus95 and RAVD, respectively, which improves the DSC by 2.3% points compared to the base 3D U-Net. The obvious performance gains for all three metrics illustrate the effectiveness of the proposed MAU-Net. Besides, we performed a paired sample t-test for the results, as illustrated in Table 1. The p-value of Haus95 is significant at $p = 0.05$ level, reveals the better performance of MAU-Net compared to the original 3D U-Net. However, DSC and RAVD are not significant, which may be due to insufficient data. In the future, we will collect more data to further validate our algorithm.

To verify whether MAU-Net can better identify smaller cancer areas, we calculated the number of voxels in the cancer area of each subject and divided all subjects into two groups based on a threshold of 10000. For each group, we computed the DSC, Haus95 and RAVD of each network architecture. The results are shown in Figure 6. For the group with a smaller cancer region, MAU-Net achieved much better segmentation results than base U-Net and DAU-Net. For the group with a larger cancer region, each network achieved comparable segmentation results.

4. Conclusion

In this paper, we proposed MAU-Net, a novel 3D encoder-decoder based CNN architecture for accurately segmenting lung cancers from volumetric CT images. Motivated by existing dual attention and attention gate methods, we make further improvements for better lung cancer segmentation. We hope to modify the model architecture further and explore its performance in other medical imaging tasks in future work.

Acknowledgements

This work was supported by National Natural Science Foundation of China (Grant No. U1806202), the 360 Department of Science and Technology of Shandong Province (Grant No. 2017CXGC1502), the Taishan Scholars 361 Project in Shandong Province (ts201511106) and the Academic promotion programme of Shandong First Medical 362 University(NO. 2019QL023).

References

[1] H. Sung, J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, F. Bray, Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA: A Cancer Journal for Clinicians n/a (n/a). doi:https://doi.org/10.3322/caac.21660.

[2] J.-J. Sonke, J. Belderbos, Adaptive radiotherapy for lung cancer, Seminars in Radiation Oncology 20 (2) (2010) 94–106, adaptive Radiotherapy. doi:https://doi.org/10.1016/j.semradonc.2009.11.003.

[3] G. Sharp, K. D. Fritscher, V. Pekar, M. Peroni, N. Shusharina, H. Veeraraghavan, J. Yang, Vision 20/20: Perspectives on automated image segmentation for radiotherapy, Medical Physics 41 (5) (2014) 050902. doi:https://doi.org/10.1118/1.4871620.

[4] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 3431–3440.

[5] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical image computing and computer-assisted intervention, Springer, 2015, pp. 234–241. doi:https://doi.org/10.1007/978-3-319-24574-4_28.

[6] F. Isensee, P. F. Jaeger, S. A. Kohl, J. Petersen, K. H. Maier-Hein, nnu-net: a self-configuring method for deep learning-based biomedical image segmentation, Nature Methods 18 (2) (2021) 203–211. doi:https://doi.org/10.1038/s41592-020-01008-z.

[7] S. A. Taghanaki, K. Abhishek, J. P. Cohen, J. Cohen-Adad, G. Hamarneh, Deep semantic segmentation of natural and medical images: A review, Artificial Intelligence Review 54 (1) (2021) 137–178. doi:https://doi.org/10.1007/s10462-020-09854-1.

[8] D. Nie, Y. Gao, L. Wang, D. Shen, Asdnet: Attention based semi-supervised deep networks for medical image segmentation, in: A. F. Frangi, J. A. Schnabel, C. Davatzikos, C. Alberola-López, G. Fichtinger (Eds.), Medical Image Computing and Computer Assisted Intervention – MICCAI 2018, Springer International Publishing, Cham, 2018, pp. 370–378. doi:https://doi.org/10.1007/978-3-030-00937-3_43.

[9] A. Sinha, J. Dolz, Multi-scale self-guided attention for medical image segmentation, IEEE Journal of Biomedical and Health Informatics 25 (1) (2021) 121–130. doi:https://doi.org/10.1109/JBHI.2020.2986926.

[10] R. Gu, G. Wang, T. Song, R. Huang, M. Aertsen, J. Deprest, S. Ourselin, T. Vercauteren, S. Zhang, Ca-net: Comprehensive attention convolutional neural networks for explainable medical image segmentation, IEEE Transactions on Medical Imaging 40 (2) (2021) 699–711. doi:https://doi.org/10.1109/TMI.2020.3035253.

[11] M. Saad, I. H. Lee, T.-S. Choi, Automated delineation of non-small cell lung cancer: A step toward quantitative reasoning in medical decision science, International Journal of Imaging Systems and Technology 29 (4) (2019) 561–576. doi:https://doi.org/10.1002/ima.22336.

[12] U. Kamal, A. M. Rafi, R. Hoque, J. Wu, M. K. Hasan, Lung cancer tumor region segmentation using recurrent 3d-denseunet, in: J. Petersen, R. San José Estépar, A. Schmidt-Richberg, S. Gerard, B. Lassen-Schmidt, C. Jacobs, R. Beichel, K. Mori (Eds.), Thoracic Image Analysis, Springer International Publishing, Cham, 2020, pp. 36–47.

[13] J. Jiang, Y. C. Hu, C. J. Liu, D. Halpenny, M. D. Hellmann, J. O. Deasy, G. Mageras, H. Veeraraghavan, Multiple resolution residually connected feature streams for automatic lung tumor segmentation from ct images, IEEE Transactions on Medical Imaging 38 (1) (2019) 134–144. doi:https://doi.org/10.1109/TMI.2018.2857800.

[14] F. Zhang, Q. Wang, H. Li, Automatic segmentation of the gross target volume in non-small cell lung cancer using a modified version of resnet, Technology in Cancer Research & Treatment 19 (2020) 1–9. doi:https://doi.org/10.1177/1533033820947484.

[15] S. Byun, J. Jung, H. Hong, H. Oh, B. seog Kim, Lung tumor segmentation using coupling-net with shape-focused prior on chest CT images of non-small cell lung cancer patients, in: H. K. Hahn, M. A. Mazurowski (Eds.), Medical Imaging 2020: Computer-Aided Diagnosis, Vol. 11314, International Society for Optics and Photonics, SPIE, 2020, pp. 598 – 603. doi:https://doi.org/10.1117/12.2551280.

[16] A. P. Harrison, Z. Xu, K. George, L. Lu, R. M. Summers, D. J. Mollura, Progressive and multi-path holistically nested neural networks for pathological lung segmentation from ct images, in: M. Descoteaux, L. Maier-Hein, A. Franz, P. Jannin, D. L. Collins, S. Duchesne (Eds.), Medical Image Computing and Computer Assisted Intervention - MICCAI 2017, Springer International Publishing, Cham, 2017, pp. 621–629. doi:https://doi.org/10.1007/978-3-319-66179-7_71.

[17] A. Comelli, C. Coronnello, N. Dahiya, V. Benfante, S. Palmucci, A. Basile, C. Vancheri, G. Russo, A. Yezzi, A. Stefano, Lung segmentation on high-resolution computerized tomography images using deep learning: A preliminary step for radiomics studies, Journal of Imaging 6 (11) (2020). doi:https://doi.org/10.3390/jimaging6110125.

[18] F. Shariaty, S. Hosseinlou, V. Y. Rud', Automatic lung segmentation method in computed tomography scans, Journal of Physics: Conference Series 1236 (2019) 012028. doi:https://doi.org/10.1088/1742-6596/1236/1/012028.

[19] J. Tan, L. Jing, Y. Huo, L. Li, O. Akin, Y. Tian, Lgan: Lung segmentation in ct scans using generative adversarial network, Computerized Medical Imaging and Graphics 87 (2021) 101817. doi:https://doi.org/10.1016/j.compmedimag.2020.101817.

[20] J. Hofmanninger, F. Prayer, J. Pan, S. Röhrich, H. Prosch, G. Langs, Automatic lung segmentation in routine imaging is primarily a data diversity problem, not a methodology problem, European Radiology Experimental 4 (1) (2020) 50. doi:https://doi.org/10.1186/s41747-020-00173-2.

[21] O. Oktay, J. Schlemper, L. L. Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y. Hammerla, B. Kainz, et al., Attention u-net: Learning where to look for the pancreas, arXiv preprint arXiv:1804.03999 (2018).