

Automatic Segmentation of Clinical Target Volume and Organs-at-Risk for Breast Conservative Radiotherapy Using a Convolutional Neural Network

Zhikai Liu¹
 Fangjie Liu^{2,*}
 Wanqi Chen^{1,*}
 Yinjie Tao¹
 Xia Liu¹
 Fuquan Zhang¹
 Jing Shen¹
 Hui Guan¹
 Hongnan Zhen¹
 Shaobin Wang³
 Qi Chen³
 Yu Chen³
 Xiaorong Hou¹

¹Department of Radiation Oncology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences, Beijing, 100730, People's Republic of China; ²Department of Radiation Oncology, Sun Yat-sen University Cancer Center, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Guangzhou, 510060, People's Republic of China; ³MedMind Technology Co., Ltd., Beijing, 100055, People's Republic of China

*These authors contributed equally to this work

Objective: Delineation of clinical target volume (CTV) and organs at risk (OARs) is important for radiotherapy but is time-consuming. We trained and evaluated a U-ResNet model to provide fast and consistent auto-segmentation.

Methods: We collected 160 patients' CT scans with breast cancer who underwent breast-conserving surgery (BCS) and were treated with radiotherapy. CTV and OARs were delineated manually and were used for model training. The dice similarity coefficient (DSC) and 95th percentile Hausdorff distance (95HD) were used to assess the performance of our model. CTV and OARs were randomly selected as ground truth (GT) masks, and artificial intelligence (AI) masks were generated by the proposed model. Two clinicians randomly compared CTV score differences of the contour. The consistency between two clinicians was tested. Time cost for auto-delineation was evaluated.

Results: The mean DSC values of the proposed method were 0.94, 0.95, 0.94, 0.96, 0.96 and 0.93 for breast CTV, contralateral breast, heart, right lung, left lung and spinal cord, respectively. The mean 95HD values were 4.31mm, 3.59mm, 4.86mm, 3.18mm, 2.79mm and 4.37mm for the above structures, respectively. The average CTV scores for AI and GT were 2.89 versus 2.92 when evaluated by oncologist A ($P=0.612$), and 2.75 versus 2.83 by oncologist B ($P=0.213$), with no statistically significant differences. The consistency between two clinicians was poor ($\kappa=0.282$). The time for auto-segmentation of CTV and OARs was 10.03 s.

Conclusion: Our proposed model (U-ResNet) can improve the efficiency and accuracy of delineation compared with U-Net, performing equally well with the segmentation generated by oncologists.

Keywords: clinical target volume, organ at risk, auto-segmentation, breast cancer radiotherapy, clinical evaluation

Key Points

1. A U-ResNet model can auto-delineate for breast conservative radiotherapy.
2. CTV and OARs generated by our model can meet the clinical requirements.
3. AI assistance can effectively improve consistency in contouring radiotherapy workflow.

Introduction

Breast cancer (BC) is one of the most common cancers for women throughout the world.¹ Breast radiotherapy after breast-conserving surgery (BCS) is an essential treatment for early breast cancer patients.^{2,3} Radiotherapy of tumors requires

Correspondence: Xiaorong Hou
 Tel +86 138 1196 3013
 Email hxr_pumch@163.com

Received: 8 August 2021
 Accepted: 4 October 2021
 Published: 2 November 2021



accurate, individualized contouring of clinical target volume (CTV) and organs at risk (OARs) to deliver high radiation doses to the target and to spare healthy tissues.⁴ Therefore, computer-assisted automatic segmentation techniques are highly desired and useful for relieving radiation oncologists from labor-intensive work as well as reducing considerable inter- and intra-observer variability in delineation of the regions of interest (ROIs).^{5,6}

Current automatic approaches can be generally categorized into two groups: atlas-based auto-segmentation (ABAS) and convolutional neural network (CNN) based segmentation. Acceptable results have been reported using ABAS for OARs in head and neck cancer and prostate cancer.⁷⁻⁹ However, CTV is not a region with clear boundaries but includes tissues of potential tumor or subclinical diseases that are barely detectable in CT images.¹⁰ Moreover, the inconsistencies in body shape, organ size, and density of mammary glandular tissue remain large from person to person.^{11,12} Therefore, various kinds of CNN models¹³⁻¹⁶ have been presented for different cancers,^{8,16-20} showing better performance than ABAS.

A deep dilated residual network (DD-ResNet) was previously proposed by Men et al¹⁶ to perform automatic breast CTV contouring. A 0.91 DSC was reported for both the right and left breast CTV, but no clinical evaluation was performed. Moreover, this method was focused on CTV contouring; the OARs were not considered.

Here, we constructed a new CNN model based on the 2D U-Net model to solve the large inconsistencies between source and target image, even with a scarce amount of labelled training data. The proposed model was trained and then compared against U-Net. The accuracy and effectiveness were evaluated by both performance metrics and qualified radiation oncologists.

Materials and Methods

Data Acquisition

CT scans of patients with early-stage BC who underwent BCS in Peking Union Medical College Hospital were collected from January 2019 to December 2019. This

study was approved by the Institutional Review Board of Peking Union Medical College Hospital. Informed consent/assent from the patient and/or parent/guardian, as appropriate, was obtained before enrollment. This study was conducted in accordance with the Declaration of Helsinki. The inclusion criteria are as follows: (1) Patients who were diagnosed with early-stage BC and underwent breast conservative surgery. (2) All the patients met the indication for radiotherapy and received whole-breast irradiation. Patients who underwent axilla or supra-clavicular lymph nodes radiotherapy were excluded.

In total, 12,640 CT slices were collected from 160 patients; 79 patients had left-sided BC and the remainder had right-sided BC. All the CT scans followed the digital imaging and communications in medicine (DICOM) protocol and were scanned using a Philips Brilliance Big Bore CT scanner. CT images were reconstructed using a matrix size of 512×512 and a thickness of 5 mm. The pixel spacing of the data was 1.1543 mm × 1.1543 mm.

Contouring of the CTV and OARs (contralateral breast, lungs, heart, and spinal cord) were delineated manually by trained radiation oncologists following the European Society for Radiotherapy and Oncology (ESTRO)²¹ and the Radiation Therapy Oncology Group (RTOG)²² protocols. The specific sketching standards for CTV are shown in Table 1. All the contours were reviewed and approved by two professional radiation oncologists with more than 10 years' experience in our center.

Network Architecture

Our model, called U-ResNet, is originated from the 2D U-Net model, which is composed of encoder and decoder paths. To conduct the segmentation task for BC radiotherapy, especially for the CTV segmentation, a deep network should be added to the U-Net to extract features at different abstraction levels. At the same time, the vanishing gradients of deep convolutional networks should be avoided. Therefore, ResNet is used as the encoder part. It encodes low-, middle- and high- level features and passes these features to the decoder part via four shortcut

Table 1 The Standard Delineation of CTV After BCS

Borders	Cranial	Caudal	Ventral	Dorsal	Lateral	Medial
Residual breast CTV	Upper border of palpable/visible breast tissue; maximally up to the inferior edge of the sterno- clavicular joint	Most caudal CT slice with visible breast	5 mm below the skin surface	Major pectoral muscle or costae and intercostal muscles where no muscle	Lateral breast fold; anterior to the lateral thoracic artery	The edge of sternum

connections. In the decoder part, the upscaling is achieved using nearest neighbour interpolation, followed by a convolutional layer and a residual block. In this way, multiple-level features in the encoder and decoder parts are concatenated. The overall architectures of DD-ResNet and our proposed method are shown in Figure 1. DD-ResNet has no shortcut connections between the encoder and the decoder. The output of the sum layer was interpolated to the original size with a factor of 8, which may result in information loss.

The breast is a continuous and smooth surface. A 2D architecture may result in a rough segmentation result in a 3D view. To obtain the 3D information of CT scans, the network is designed as a 2.5D architecture by assigning three adjacent slices into three channels as the input.

Implementation Details

The dataset, composed of 160 patients, was randomly assigned in 8:1:1 to three cohorts: 1) a training set of 128 patients was used to construct the segmentation model, 2) a validation set of 16 patients was used to optimize the parameters and 3) a testing set of 16 patients was used to obtain artificial intelligence-generated contouring for performance assessment. During the testing phase, all the CT slices of the 16 testing cases were tested individually.

We constructed our model using Python 3.6 and PyTorch 1.0. An Adam optimization algorithm²³ was

used for the optimization. The learning rate was 0.001. We trained and evaluated our model with a GTX 1080 GPU. The proposed model was trained over 50 circles to select the best model according to the lowest validation loss score. The convolutional layers are initialized using Xavier Uniform Initialization, and batch normalization layers were added after convolution layers to improve the training speed and to prevent overfitting.²⁴

Performance Measurement

Performance of the proposed method was evaluated using the dice similarity coefficient (DSC) and the 95th percentile Hausdorff distance (95HD) to quantify the results. The mean and standard deviation were also calculated.

The DSC was used to measure the spatial overlap between AI and GT contours, which is defined in Equation (1).

$$DSC(A, B) = \frac{2|A \cap B|}{|A| + |B|} \tag{1}$$

where A represents the volume of the human-generated contour; B is the volume of an AI contour; and $A \cap B$ is the intersect volume that A and B have in common. The DSC value was between 0 and 1 (0 = no overlap, 1 = complete overlap).

The 95HD is defined as follows:

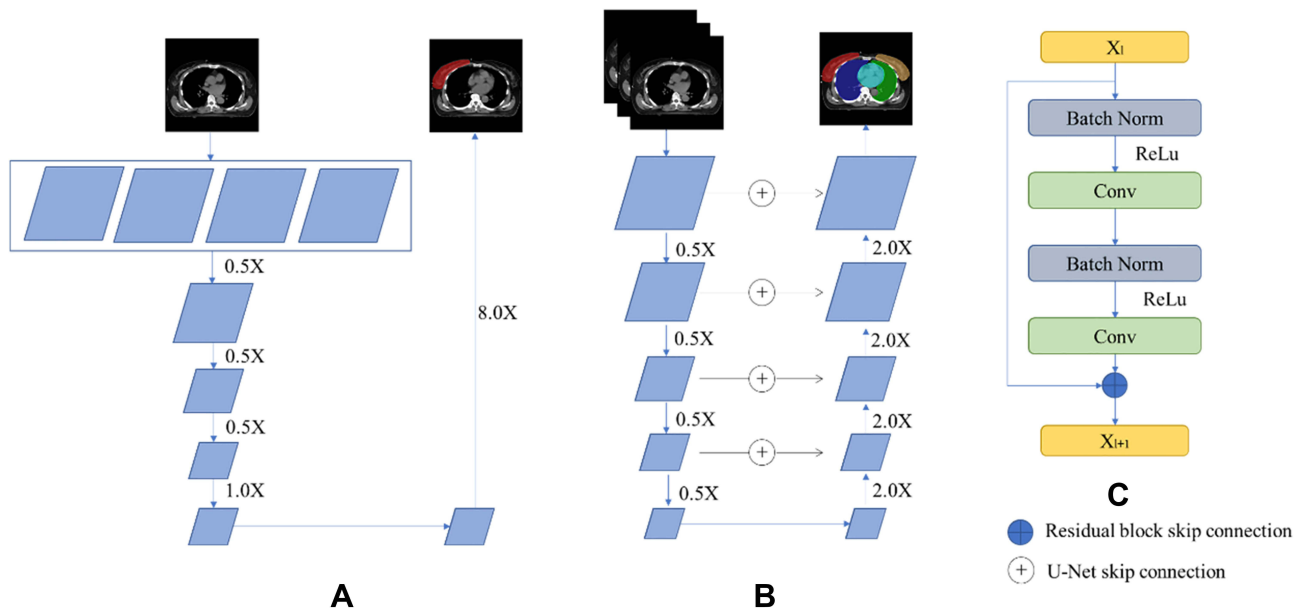


Figure 1 Architecture of (A) deep dilated convolutional neural network (DD-CNN), (B) our proposed network, and (C) the residual block used in decoder part of our network.

$$95HD(A, B) = \max(h(A, B), h(B, A), 95th) \quad (2)$$

$$h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\| \quad (3)$$

$$h(B, A) = \max_{b \in B} \min_{a \in A} \|b - a\| \quad (4)$$

where A represents the human-generated contour; B is the AI contour, $\|\cdot\|$ means the Euclidean norm of the points of A and B . The 95HD means the 95 percentile maximum mismatch between A and B . When the 95HD value decreases, the similarity between A and B increases.

Oncologist Evaluation

OARs Evaluation

Considering that evaluation metrics cannot provide a comprehensive insight into whether the contours need to be modified in clinical practice, another 20 cases in clinical practice from our center were randomly collected. Each case was delineated with GT and AI contours for OARs and then distributed to two radiation oncologists with more than 10 years of clinical experience for further evaluation. Each slice was carefully evaluated, and the results were graded in four levels: 3 points (no need to be edited), 2 points (the number of layers need to be edited ≤ 4), 1 point (the number of layers need to be edited ≥ 4) and 0 point (not acceptable).

CTV Evaluation

CTV segmentations generated by AI and GT were also evaluated blindly slice by slice. The test data contained 10 patients and 650 slices in total (AI: 327 slices vs GT: 323 slices). The representative results were also graded on four levels: 3 points (acceptable for subsequent treatment), 2 points (Minor Revision), 1 point (Major Revision) and 0 point (Not Acceptable for treatment). When the score ≥ 2 , it was defined as suitable for clinical application.

Furthermore, to verify the consistency of the judgment of two oncologists, we collected the CTV score of each slice evaluated by two oncologists, constituting a total of 650 slices of data sets. The data were classified into the same group if the slice was evaluated by two oncologists with the same CTV score. We calculated the weighted kappa coefficient to analyze for consistency.

Time Cost

Processing time was measured for the AI tool and pre- and post-AI assistance in the delineation of CTV and OARs for BC radiotherapy.

Statistical Analysis

The Wilcoxon matched-pairs signed-rank test was used to compare DSC and 95HD between our proposed model and U-Net and the differences between the two oncologists during the evaluation of CTV and OARs segmentation. McNemar test and kappa test were used to assess the consistency of the two oncologists. Statistical significance was set at two-tailed $P < 0.05$.

Results

Performance of U-ResNet and Comparison with U-Net

The median age for all the 160 patients in dataset was 49 [42, 58]. The average CTV volume was $494.41 \pm 198.51 \text{ cm}^3$.

For CTV segmentation, the average DSC values of U-ResNet and U-Net were 0.94 vs. 0.93 ($P=0.001$), and the average 95HD value was 4.31 mm vs 4.88 mm separately ($P=0.030$). Both differences were statistically significant, implying better accuracy of CTV contouring by U-ResNet.

Among all OARs, significant differences between U-ResNet and U-Net were achieved for the spinal cord (DSC: 0.93 vs 0.92 ($P=0.015$), 95HD: 4.37 mm vs 5.07 mm ($P=0.003$)) and the contralateral breast (DSC: 0.95 vs 0.93 ($P < 0.001$), 95HD: 3.59 mm vs 4.15 mm ($P=0.010$)). The right lung contouring also displayed a statistically significant difference in 95HD (3.18 mm vs 2.98 mm ($P=0.041$)).

The results of the comparison are summarized in Table 2 and Figure 2.

Figure 3 shows the visualization segmentation samples in GT, U-Net and U-ResNet, respectively. The auto-segmented contours with U-ResNet were in good concordance with the GT contours.

Oncologist Evaluation

Tables 3 and 4 show the oncologist evaluation results of OAR and CTV contours. Scores ≥ 2 were defined as suitable for clinical application. When using our grading criteria for contour evaluation, the majority of AI- and GT-generated OAR contours were deemed acceptable by the experts. Only one contour (5%) of the heart was assessed to require major revision by oncologist A.

Regarding CTV contours, 99.4% of those generated by AI were clinically acceptable by oncologist A, compared with 98.1% of GT segmentations. For oncologist B, the

Table 2 DSC and 95HD for CTV and All OARs

ROI	CTV	Contra-Lateral Breast	Spinal Cord	Heart	Right Lung	Left Lung
DSC±STD						
U-ResNet	0.94±0.019	0.94±0.016	0.93±0.024	0.94±0.023	0.96±0.017	0.96±0.025
U-Net	0.93±0.022	0.93±0.020	0.92±0.029	0.93±0.026	0.96±0.018	0.96±0.024
P value	0.001	< 0.001	0.015	0.066	0.257	0.942
95HD±STD (mm)						
U-ResNet	4.31±1.76	3.59±1.56	4.37±2.13	4.86±1.48	3.18±1.64	2.79±1.62
U-Net	4.88±1.64	4.15±1.28	5.07±2.26	5.17±2.32	2.89±1.26	2.76±1.52
P value	0.030	0.010	0.003	0.559	0.041	0.781

results were 99.4% for both methods. The average CTV scores for AI and GT were 2.89 vs 2.92 when evaluated by oncologist A ($P=0.612$) and 2.75 vs 2.83 by oncologist B ($P=0.213$), with no statistical differences.

Wilcoxon matched-pairs test was performed for the evaluation of the two oncologists for AI and GT contours separately. The results indicated that the average score of oncologist A was higher than that of oncologist B in AI contours, with a significantly statistical difference ($P=0.009$ for AI contours and $P=0.314$ for GT contours).

The comparison of the average CTV scores evaluated by two oncologists is shown in [Figure 4](#).

The evaluation results of the two oncologists were further analyzed for consistency. We independently collected the CTV scores of all the 650 slices generated by both AI and GT segmentations, and used the data to calculate weighted kappa coefficient. The results are shown in [Table 5](#). Of all the 650 slices, 532 (81.8%) were evaluated with the same CTV score. The consistency between two oncologists was not good (kappa=0.282).

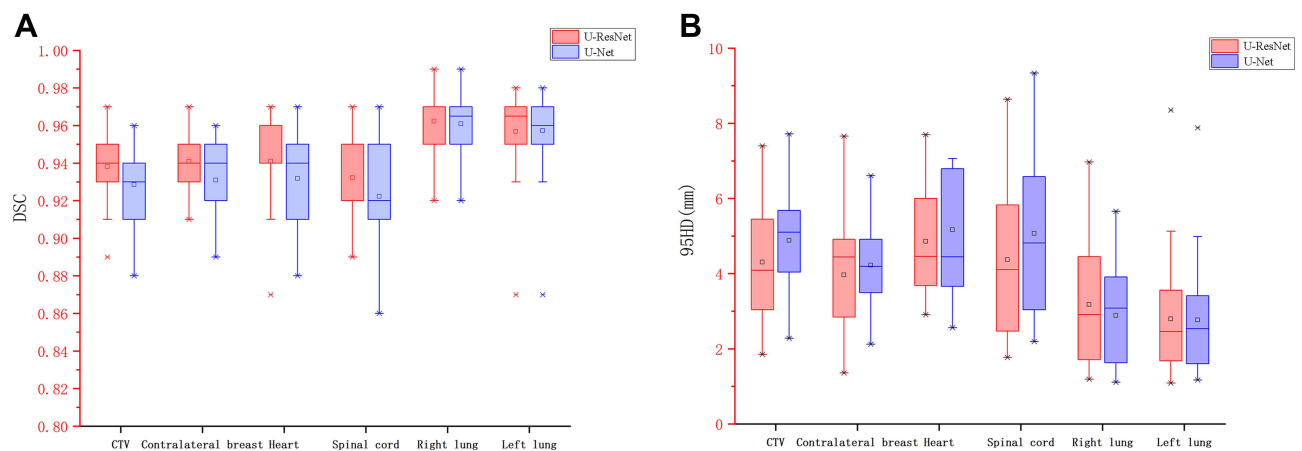
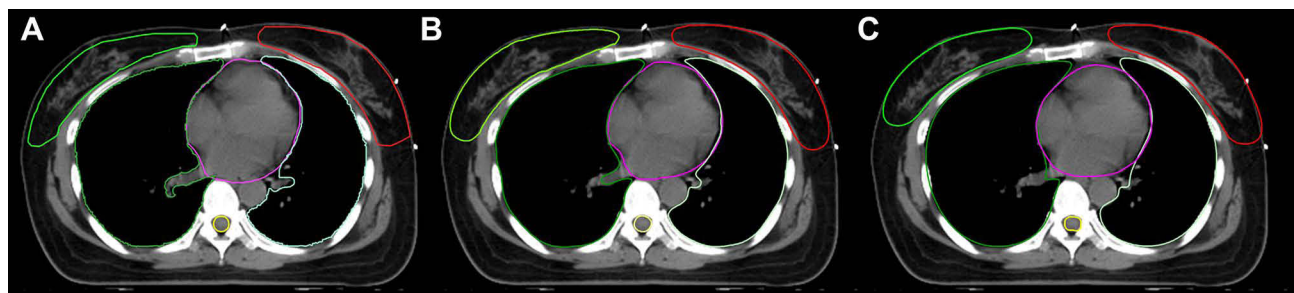
**Figure 2** Boxplots obtained for DSC and 95HD analyses of U-ResNet and U-Net. (A) DSC analyses, (B) 95HD analyses.**Figure 3** CTV and OAR contours generated by (A) GT, (B) U-ResNet, and (C) U-Net after breast conservative surgery.

Table 3 Evaluation for CTV and OARs by Oncologist A

Scores	CTV		Contra-Lateral Breast		Spinal Cord		Heart		Right Lung		Left Lung	
	AI	GT	AI	GT	AI	GT	AI	GT	AI	GT	AI	GT
1	0	0	0	0	0	0	0	0	0	0	0	0
2	0.6%	1.9%	0	0	0	0	0	5%	0	0	0	0
3	7%	7.4%	30%	35%	5%	20%	20%	20%	20%	20%	10%	0
4	92.4%	90.7%	70%	65%	95%	80%	80%	75%	80%	80%	90%	100%
P value	0.612		0.414		0.083		0.527		1.00		0.157	

Table 4 Evaluation for CTV and OARs by Oncologist B

Scores	CTV		Contra-Lateral Breast		Spinal Cord		Heart		Right Lung		Left Lung	
	AI	GT	AI	GT	AI	GT	AI	GT	AI	GT	AI	GT
1	0	0	0	0	0	0	0	0	0	0	0	0
2	0.6%	0.6%	0	0	0	0	0	0	0	0	0	0
3	23.5%	14.9%	10%	15%	0	0	0	0	0	0	0	0
4	75.9%	85.5%	90%	85%	100%	100%	100%	100%	100%	100%	100%	100%
P value	0.612		0.655		1.00		1.00		1.00		1.00	

Time Cost

Time for auto-segmentation of CTV and OARs with U-ResNet was 10.03 s, compared with 20 minutes and 30 minutes by experienced oncologists. With AI assistance, the delineation time can be reduced to 10 minutes and 5 minutes for CTV contouring and OAR contouring, respectively.

Discussion

Accurate and consistent delineation of CTV and OARs is a basic requirement for contemporary radiotherapy planning, while it is also the most burdensome step in the radiotherapy workflow.^{25,26} Manual delineation is a time-consuming process and has considerable inter- and intra-observer variability in anatomical contouring.^{27,28} In recent years, computer-assisted automatic segmentation techniques have made great breakthroughs in increasing reliability and accuracy as well as in relieving radiation oncologists from time-intensive contouring.

Among these automatic methods, CNNs are the most advanced method available for medical images and have shown better performance than atlas-based methods.²⁹ As a completely end-to-end network, only a limited amount of data is required. Among all the CNN-based models, U-Net has been a remarkable and the most popular deep network.

However, it may lose some abstract information with relatively lower level convolutional layers. To alleviate the disparity between the encoder-decoder features, we have trained and evaluated a new model based on the 2D U-Net model.

Our model was based on U-Net architecture. Therefore, the performance of the proposed model was also compared with U-Net. During the training stage, U-Net had the same training configuration as that of the proposed model. This was the first study assessing the performance of both CTV and OARs segmentation for breast-conservative radiotherapy. The analysis revealed that U-ResNet algorithm outperformed the U-Net algorithms. Moreover, U-ResNet performed well, with good agreement to the segmentations contoured manually by oncologists.

In our previous work, we have demonstrated that CNN architecture could facilitate the delineation of CTV for radical mastectomy radiotherapy.³⁰ In this piece of work, we focused on both CTV and OARs for breast conservative radiotherapy. We also made some modifications in CNN architecture. Firstly, the whole encoder part was completely replaced with the ResNet34 instead of a residual block. Moreover, we also proposed a self-adaptive weighted cross entropy loss function to tackle our multi-class structures segmentation problem, so that

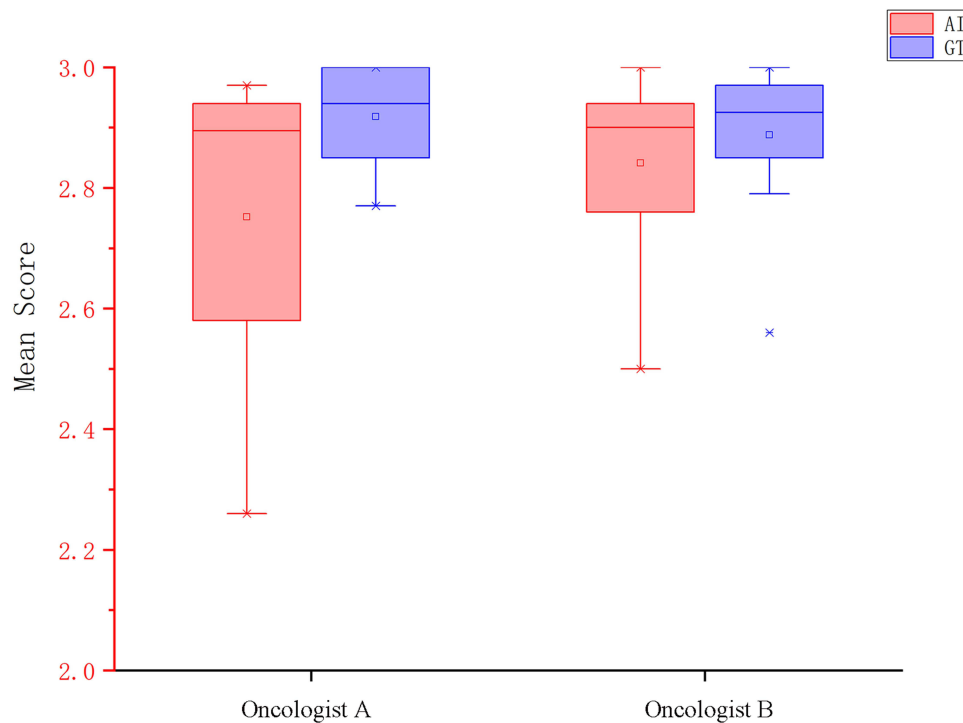


Figure 4 The average CTV scores evaluated by two oncologists.

we could train and predict all the CTV and OARs in one time, introducing greater efficiency. Notably, however, an automatic segmentation model of CTV may not take into account particular anatomical variations and clinical needs related to the multiple decision-making criticalities to be considered in the context of breast cancer treatment.³¹

The DSC and 95HD are used as quantitative evaluation metrics to assess the proposed method. DSC values are used more commonly than 95HD values in the literature. Numerous studies have reported DSC values of CTV ranges from 0.88 to 0.93.^{16,32–34} The average DSC of our proposed model was 0.94, which was higher than those in historical reports. The result indicated a strong concordance between our automatic model and human experts for CTV contouring. The corresponding average DSC of U-Net was 0.93. For our model, we also used the 95HD to

exclude the unreasonable distances caused by outliers, and the value was 4.31 mm for CTV. According to 95HD values, our model performed better than DDCNN (15.6 mm), DDNN (14.1 mm), and DD-ResNet (10.7 mm) in left breast CTV, as reported by Men et al.¹⁶ Moreover, we also evaluated the 95HD for OARs. Among all OARs, the significant results were achieved for spinal cord (95HD=4.37 mm, DSC=0.93). This is mainly because the spinal cord has good low-contrast visibility and its shape is regular. Both of the above metrics are statistically significant. Based on the results stated above, we can demonstrate that U-ResNet is superior to the U-Net architecture for this task. However, U-Net seems to be superior to U-ResNet in contouring right lung with a lower 95HD value. Considering the fact that the DSC for right lung with two architecture both reached 0.96, it remained to be discussed whether the significant difference resulted from the performance between two architectures or from the intrinsic nature of 95HD itself. In our following study, we will improve the delineation of right lung.

Except for performance metrics, we randomly distributed the AI and GT contours to experienced oncologists for further evaluation to verify their significance in clinical practice. The results showed that the majority of the AI contours can be accepted for subsequent treatment. There are no significant differences in scores grading between AI and GT contours

Table 5 The Consistency Test Between Two Oncologists

CTV Score		Doctor A				
		0	1	2	3	Total
Doctor B	0	0	0	0	0	0
	1	0	2	4	17	23
	2	0	5	26	75	106
	3	0	0	17	504	521
	Total	0	7	47	596	650

for CTV and all OARs, meaning that the U-ResNet model performed well, with good agreement for the manual contours. However, a Wilcoxon matched-pairs test indicated the significant difference between the evaluation of two oncologists. The inter-observer variability of the delineation standard is also one of the limitations of our study.

In addition, the consistency between the two radiation oncologists did not seem to be good. Among all the 650 slices of CTV marked with AI and GT contours, 532 slices (81.8%) were evaluated with the same CTV score, and the weighted kappa test ($\kappa = 0.282$) showed the poor consistency between the oncologists. The results proved the considerable inter- and intra-observer variability, resulting in different judgment criteria. This may illustrate the fact that the acceptance of AI contouring methods is still impacted by the opinion and expertise of the treating radiation oncologists. We also evaluated the time needed for segmentation. With an approximately the same time cost, U-ResNet enjoyed an advantage of auto-contouring CTV and OARs spontaneously, saving time for clinical work. With U-ResNet assistance, contouring could be much easier and more efficient.

Several limitations of our study should be noted. First, the study is a single-center study, and the ground truth contours were approved by only two oncologists. Therefore, the proposed model cannot meet the contouring preferences of other centers and all clinicians. In the future, multi-center research should be conducted to obtain larger datasets to improve the generalization ability of the auto-segmentation model. Second, studies have suggested that 3D architecture showed better performance on segmentation compared with 2D.³⁵ We may consider extending our model to 3D U-Net in the future. Third, CT images with artifacts caused by pacemakers as well as contrast injected CT images were not included in our training set. We were unable to deal with these conditions so far. Moreover, the CT slice thickness used in our study was 5 mm for model training. The auto-contouring will be less accurate if it is used for images with some other slice thickness.

Conclusion

Accurate and consistent segmentation is important for improving radiotherapy outcomes. Our study implemented a U-ResNet model to auto-delineate the CTV and OARs for breast conservative radiotherapy on planning CT images. The results showed that AI assistance can effectively improve consistency in contouring and streamlining radiotherapy workflows. In the future, with the assistance of AI, we may be able to obtain an initial standard

segmentation to reduce the bias caused by observer preferences. However, for different patients and diseases, the contouring details should still be delineated specifically according to the principle of individualization.

Abbreviations

BC, breast cancer; BCS, breast conserving surgery; WBI, whole-breast irradiation; CTV, clinical target volume; OARs, organs at risk; ROIs, regions of interest; ABAS, atlas-based auto-segmentation; CNNs, convolutional neural networks; DIR, deformable image registration; GT, ground truth; AI, artificial Intelligence; DICOM, digital imaging and communications in medicine; ESTRO, European Society for Radiotherapy and Oncology; RTOG, Radiation Therapy Oncology Group.

Funding

The paper was funded by the National Foundation for Education Sciences Planning (grant number BLA200216).

Disclosure

Ms Qi Chen, Mr Shaobin Wang, and Mr Yu Chen are employees of MedMind Technology Co. The authors report no other conflicts of interest in this work.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin.* 2020;70:7–30.
2. Veronesi U, Cascinelli N, Mariani L, et al. Twenty-year follow-up of a randomized study comparing breast-conserving surgery with radical mastectomy for early breast cancer. *N Engl J Med.* 2002;347(16):1227–1232. doi:10.1056/NEJMoa020989
3. Darby S, McGale P, Correa C, et al. Effect of radiotherapy after breast-conserving surgery on 10-year recurrence and 15-year breast cancer death: meta-analysis of individual patient data for 10,801 women in 17 randomised trials. *Lancet.* 2011;378:1707–1716.
4. Andrianarison VA, Laouiti M, Fargier-Bochaton O, et al. Contouring workload in adjuvant breast cancer radiotherapy. *Cancer Radiother.* 2018;22:747–753. doi:10.1016/j.canrad.2018.01.008
5. Jensen NK, Mulder D, Lock M, et al. Dynamic contrast enhanced CT aiding gross tumor volume delineation of liver tumors: an interobserver variability study. *Radiother Oncol.* 2014;111:153–157. doi:10.1016/j.radonc.2014.01.026
6. Steenbergen P, Haustermans K, Lerut E, et al. Prostate tumor delineation using multiparametric magnetic resonance imaging: inter-observer variability and pathology validation. *Radiother Oncol.* 2015;115:186–190. doi:10.1016/j.radonc.2015.04.012
7. Wong WKH, Leung LHT, Kwong DLW. Evaluation and optimization of the parameters used in multiple-atlas-based segmentation of prostate cancers in radiation therapy. *Brit J Radiol.* 2016;89. doi:10.1259/bjr.20140732
8. Tong N, Gou SP, Yang SY, Ruan D, Sheng K. Fully automatic multi-organ segmentation for head and neck cancer radiotherapy using shape representation model constrained fully convolutional neural networks. *Med Phys.* 2018;45:4558–4567. doi:10.1002/mp.13147

9. Fortunati V, Verhaart RF, van der Lijn F, et al. Tissue segmentation of head and neck CT images for treatment planning: a multiatlas approach combined with intensity modeling. *Med Phys*. 2013;40:071905. doi:10.1118/1.4810971
10. Tao CJ, Yi JL, Chen NY, et al. Multi-subject atlas-based auto-segmentation reduces interobserver variation and improves dosimetric parameter consistency for organs at risk in nasopharyngeal carcinoma: a multi-institution clinical study. *Radiother Oncol*. 2015;115:407–411. doi:10.1016/j.radonc.2015.05.012
11. Men K, Dai JR, Li YX. Automatic segmentation of the clinical target volume and organs at risk in the planning CT for rectal cancer using deep dilated convolutional neural networks. *Med Phys*. 2017;44:6377–6389. doi:10.1002/mp.12602
12. Ibtihaz N, Rahman MS. MultiResUNet: rethinking the U-Net architecture for multimodal biomedical image segmentation. *Neural Netw*. 2020;121:74–87. doi:10.1016/j.neunet.2019.08.025
13. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. *Lect Notes Comput Sci*. 2015;9351:234–241.
14. Badrinarayanan V, Kendall A, Cipolla R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans Pattern Anal Mach Intell*. 2017;39:2481–2495. doi:10.1109/TPAMI.2016.2644615
15. Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation. *IEEE Trans Pattern Anal Mach Intell*. 2017;39:640–651. doi:10.1109/TPAMI.2016.2572683
16. Men K, Zhang T, Chen XY, et al. Fully automatic and robust segmentation of the clinical target volume for radiotherapy of breast cancer using big data and deep learning. *Phys Med*. 2018;50:13–19. doi:10.1016/j.ejmp.2018.05.006
17. Ibragimov B, Xing L. Segmentation of organs-at-risks in head and neck CT images using convolutional neural networks. *Med Phys*. 2017;44:547–557. doi:10.1002/mp.12045
18. Chan JW, Kearney V, Haaf S, et al. A convolutional neural network algorithm for automatic segmentation of head and neck organs at risk using deep lifelong learning. *Med Phys*. 2019;46:2204–2213. doi:10.1002/mp.13495
19. Men K, Chen X, Zhang Y, et al. Deep deconvolutional neural network for target segmentation of nasopharyngeal cancer in planning computed tomography images. *Front Oncol*. 2017;7:315. doi:10.3389/fonc.2017.00315
20. Liu Z, Liu X, Xiao B, et al. Segmentation of organs-at-risk in cervical cancer CT images with a convolutional neural network. *Phys Med*. 2020;69:184–191. doi:10.1016/j.ejmp.2019.12.008
21. Offersen BV, Boersma LJ, Kirkove C, et al. ESTRO consensus guideline on target volume delineation for elective radiation therapy of early stage breast cancer. *Radiother Oncol*. 2015;114:3–10. doi:10.1016/j.radonc.2014.11.030
22. R.T.O. Group. *Breast Cancer Atlas for Radiation Therapy Planning: Consensus Definition*; 2018.
23. Dubey SR, Chakraborty S, Roy SK, Mukherjee S, Singh SK, Chaudhuri BB. diffGrad: an optimization method for convolutional neural networks. *IEEE Trans Neural Netw Learn Syst*. 2019;31(11):4500–4511. doi:10.1109/TNNLS.2019.2955777
24. Kalayeh MM, Shah M. Training faster by separating modes of variation in batch-normalized models. *IEEE Trans Pattern Anal Mach Intell*. 2020;42:1483–1500. doi:10.1109/TPAMI.2019.2895781
25. Lennerts E, Coucke P. [The radiotherapy journey: from information to treatment]. *Rev Med Liege*. 2014;69(Suppl 1):3–8. Romanian.
26. Eldesoky AR, Yates ES, Nyeng TB, et al. Internal and external validation of an ESTRO delineation guideline - dependent automated segmentation tool for loco-regional radiation therapy of early breast cancer. *Radiother Oncol*. 2016;121:424–430. doi:10.1016/j.radonc.2016.09.005
27. Jameson MG, Holloway LC, Vial PJ, Vinod SK, Metcalfe PE. A review of methods of analysis in contouring studies for radiation oncology. *J Med Imaging Radiat Oncol*. 2010;54:401–410. doi:10.1111/j.1754-9485.2010.02192.x
28. Barillot I, Chauvet B, Hannoun Lévi JM, Lisbona A, Leroy T, Mahé MA. [The irradiation process]. *Cancer Radiother*. 2016;20(Suppl):S8–S19. French. doi:10.1016/j.canrad.2016.07.013
29. Lee LK, Liew SC, Thong WJ. A review of image segmentation methodologies in medical image. *Lect Notes Electr Eng*. 2015;315.
30. Liu Z, Liu F, Chen W, et al. Automatic segmentation of clinical target volumes for post-modified radical mastectomy radiotherapy using convolutional neural networks. *Front Oncol*. 2021;10:581347.
31. Gregucci F, Fozza A, Falivene S, et al; R. Italian Society of G. Clinical oncology breast, present clinical practice of breast cancer radiotherapy in Italy: a nationwide survey by the Italian Society of Radiotherapy and Clinical Oncology (AIRO) Breast Group. *Radiol Med*. 2020;125(7):674–682. doi:10.1007/s11547-020-01147-5
32. Batumalai V, Koh ES, Delaney GP, et al. Interobserver variability in clinical target volume delineation in tangential breast irradiation: a comparison between radiation oncologists and radiation therapists. *Clin Oncol*. 2011;23:108–113. doi:10.1016/j.clon.2010.10.004
33. Mast M, Coerkamp E, Heijnen M, et al. Target volume delineation in breast conserving radiotherapy: are co-registered CT and MR images of added value? *Radiat Oncol*. 2014;9(1):65. doi:10.1186/1748-717X-9-65
34. van der Leij F, Elkhuzen PHM, Janssen TM, et al. Target volume delineation in external beam partial breast irradiation: less inter-observer variation with preoperative- compared to postoperative delineation. *Radiother Oncol*. 2014;110:467–470. doi:10.1016/j.radonc.2013.10.033
35. Balagopal A, Kazemifar S, Nguyen D, et al. Fully automated organ segmentation in male pelvic CT images. *Phys Med Biol*. 2018;63:245015. doi:10.1088/1361-6560/aaf11c

Cancer Management and Research

Publish your work in this journal

Cancer Management and Research is an international, peer-reviewed open access journal focusing on cancer research and the optimal use of preventative and integrated treatment interventions to achieve improved outcomes, enhanced survival and quality of life for the cancer patient.

Submit your manuscript here: <https://www.dovepress.com/cancer-management-and-research-journal>

The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.