MDA-TransUNet: A Deep Learning-Based Automatic Segmentation Method for Cervical Cancer Brachytherapy

Dezheng Cao[1]; Xinye Ni[1]

[1]The Second People's Hospital of Changzhou, The Third Affiliated Hospital of Nanjing Medical University, Changzhou 213003, China

2026 JULY 19–22 VANCOUVER, BC
JOINT AAPM | COMP MEETING

ABSTRACT

Purpose:
Accurate delineation of the high-risk clinical target volume (HR-CTV) and organs at risk (OARs) is critical for cervical cancer brachytherapy. However, treatment planning is time-consuming, and prolonged waiting can lead to organ displacement and patient discomfort. Additionally, the steep dose gradients around HR-CTV amplify segmentation errors in HR-CTV and OARs. Therefore, achieving rapid and precise delineation of HR-CTV and OARs remains challenging. This study proposes a novel network model, MDA-TransUNet, for fast segmentation of HR-CTV and OARs in cervical cancer.
Materials:
We applied MDA-TransUnet, a CNN-Transformer hybrid model, to segment the bladder, colon, rectum, small bowel, and HR-CTV on cervical cancer CT images. 122 cervical cancer brachytherapy patients' CT images from three clinical centers were utilized for training and testing, with 80 cases allocated to training, 22 to testing, and 20 to external validation. Segmentation accuracy was quantified using the Dice Similarity Coefficient (DSC), Hausdorff Distance (HD95), and Average Surface Distance (ASD). Dosimetric differences were analyzed via paired t-tests.
Results:
Compared to other methods, MDA-TransUnet achieved superior segmentation performance on the test dataset. The DSCs for the bladder, colon, rectum, small bowel, and HR-CTV were 94.54%, 79.27%, 79.27%, 88.90%, and 82.35%, respectively. Paired t-tests on five dosimetric metrics (D_{5cc}, D_{2cc}, $D_{0.1cc}$, $D_{90\%}$, and D_{mean}) showed no significant differences. For OARs, the average difference in D_{2cc} was less than 12%. For HR-CTV, the average difference in Dmean was less than 8%, and $D_{90\%}$ was less than 11%.
Conclusion:
This work demonstrates the superiority of MDA-TransUnet in segmenting OARs and HR-CTV for cervical cancer brachytherapy, with robust performance across multi-center datasets.

CONTACT

Xinye Ni
The Second People's Hospital of Changzhou, The Third Affiliated Hospital of Nanjing Medical University, Changzhou, China
No. 68, Gehu Middle Road, Wujin District, Changzhou City, Jiangsu Province, China
nxy@njmu.edu.cn
+86 13776856838

INTRODUCTION

Accurate contouring is essential in cervical cancer brachytherapy because steep dose gradients make treatment planning highly sensitive to segmentation errors. However, manual delineation of HR-CTV and OARs is time-consuming and may prolong the interval between applicator insertion and treatment delivery.
Although deep learning has improved medical image segmentation, robust auto-segmentation for cervical cancer brachytherapy remains challenging due to variable pelvic anatomy, low CT soft-tissue contrast, and inter-center differences.
This study developed a multi-center CNN–Transformer framework for automatic segmentation of HR-CTV and OARs on brachytherapy CT images.

METHODS

This retrospective multi-center study included cervical cancer patients who underwent CT-based brachytherapy planning. Manual contours of the bladder, colon, rectum, small bowel, and HR-CTV were used as reference standards.
MDA-TransUNet was built on a CNN–Transformer segmentation framework. The CNN component extracted local anatomical features, while the Transformer component captured long-range spatial relationships.
Two modules were incorporated to improve robustness: MASAG enhanced multi-scale feature fusion between encoder and decoder layers, and DCAM improved adaptation to irregular organ shapes and anatomical deformation.

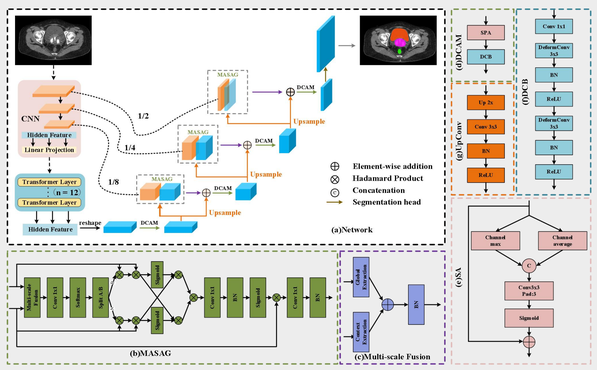

Figure 1. MDA-TransUnet network architecture diagram. (a) Overall framework of MDA-TransUnet, (b) Multi-Scale Adaptive Spatial Attention Gate (MASAG), (c) Multi-scale Feature Fusion (MSF), (d) Deformable Convolutional Attention Module (DCAM), (e) Spatial Attention (SA), (f) Deformable Convolutional Block (DCB), (g) UpConv.

RESULTS

Metrics	TransUnet	DLKA-net	SelfRag-Unet	Unet	Cascade	EMCAD	Our Study
Bladder							
DSC	88.59±2.79	92.50±3.22	91.93±2.56	91.98±2.37	94.11±1.97	94.20±1.60	94.54±1.54
HD95	3.38±1.03	4.08±4.85	2.56±0.83	2.49±0.79	1.90±0.71	1.99±0.69	1.80±0.87
ASD	1.23±0.40	1.38±0.99	1.03±0.47	0.91±0.30	0.61±0.27	0.59±0.19	0.54±0.21
Colon							
DSC	64.57±7.73	70.25±11.11	67.70±10.01	68.46±10.41	77.77±8.00	78.46±8.13	79.27±8.05
HD95	19.08±15.18	14.49±16.83	15.92±13.25	17.57±16.83	11.86±16.37	11.28±14.80	11.66±14.50
ASD	4.99±3.35	4.00±4.59	4.38±3.75	4.86±4.51	2.99±4.61	2.45±2.01	2.74±4.09
Rectum							
DSC	70.22±11.24	73.57±7.00	74.76±7.57	75.92±7.41	78.80±6.32	78.41±5.96	79.27±6.79
HD95	6.83±3.03	6.86±3.72	6.39±3.07	5.70±2.67	5.79±3.62	6.02±4.12	6.28±4.24
ASD	1.93±0.64	2.41±1.37	1.91±1.10	1.60±0.74	1.45±0.86	1.56±0.75	1.35±0.77
Small intestine							
DSC	77.23±5.09	83.46±4.93	80.64±6.15	81.13±5.98	87.55±4.20	87.17±3.86	88.90±3.64
HD95	5.76±1.73	4.34±1.51	6.00±3.27	5.58±2.74	3.16±1.52	3.40±2.13	2.97±1.44
ASD	1.70±0.46	1.49±0.52	1.89±0.89	1.72±0.79	0.92±0.38	0.96±0.41	0.87±0.48
HR-CTV							
DSC	77.42±5.18	77.71±4.76	79.69±4.32	79.15±4.47	80.72±4.02	81.49±4.34	82.35±4.07
HD95	5.06±1.53	4.60±1.60	4.43±1.84	4.53±1.89	4.05±1.56	3.99±1.70	3.77±1.58
ASD	1.94±0.64	1.82±0.54	1.57±0.44	1.64±0.50	1.52±0.52	1.41±0.56	1.29±0.40

TABLE 1. Quantitative comparison between our proposed method and other methods, where "Our Study" denotes our approach.

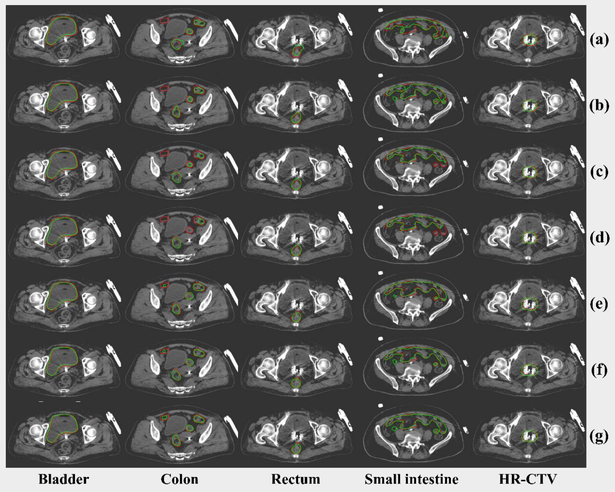

Figure 2. Visual comparison of our method with other methods, where red indicates the ground truth and green represents AI-based segmentation results. Subfigures correspond to: (a) TransUnet, (b) DLKA-net, (c) SelfRag-Unet, (d) Unet, (e) Trans-CASCADE, (f) EMCAD, and (g) Our method.

Components		Segmentation Targets				
MASAG	DCAM	Bladder	Colon	Rectum	Small intestine	HR-CTV
NO	NO	93.37	76.84	77.97	86.38	79.85
YES	NO	93.96	77.39	78.25	87.10	80.94
NO	YES	94.39	78.54	78.47	87.90	80.92
YES	YES	94.54	79.27	79.27	88.90	82.35

TABLE 2. Ablation study

Structure	Dosimetric parameters	Differences	P
Bladder	D_{5cc}	0.1951±0.1607	0.388
	D_{2cc}	0.2826±0.2277	0.435
	$D_{0.1cc}$	0.6186±0.5103	0.505
Colon	D_{2cc}	0.3173±0.4511	0.614
	$D_{0.1cc}$	0.4752±0.6740	0.507
Rectum	D_{5cc}	0.0975±0.0756	0.962
	D_{2cc}	0.1160±0.0695	0.158
	$D_{0.1cc}$	0.3238±0.2612	0.234
Small intestine	D_{2cc}	0.5215±0.6269	0.704
	$D_{0.1cc}$	1.1644±1.4556	0.571
HR-CTV	$D_{90\%}$	0.6367±0.5573	0.081
	D_{mean}	0.9659±0.8036	0.470

TABLE 3. Differences in dosimetric parameters between manual methods and our method within the original clinical brachytherapy plans, along with paired t-test results.

DISCUSSION

MDA-TransUNet showed reliable auto-segmentation performance for cervical cancer brachytherapy CT images, including both OARs and HR-CTV. The model performed well for anatomically clear structures such as the bladder and also improved segmentation of more challenging regions, including the colon, small bowel, and HR-CTV.
Dosimetric comparison showed no significant differences between manual and automatic contours, suggesting potential clinical applicability for treatment planning evaluation. By reducing contouring time, this method may help improve brachytherapy workflow efficiency and decrease patient waiting time.

© Poster Template by Genigraphics® 1.800.790.4001 www.genigraphics.com