scientific reports

Check for updates

OPEN

A deep learning model based on Mamba for automatic segmentation in cervical cancer brachytherapy

Lele Zang[1,5], Jing Liu[1,5], Huiqi Zhang[3,5], Shitao Zhu[4], Mingxuan Zhu[3], Yuqin Wang[3], Yaxin Kang[3], Jihong Chen[2✉] & Qin Xu[1✉]

This study developed and evaluated an automatic segmentation model based on the Mamba framework (AM-UNet) for rapid and precise delineation of high-risk clinical target volume (HRCTV) and organs at risk (OARs) in cervical cancer brachytherapy. Using 694 CT scans from 179 cervical cancer patients, the performance of five models (AM-UNet, UNet, DeepLab V3, UNETR and nnU-Net) was compared. The models were assessed using the Dice similarity coefficient (DSC), 95% Hausdorff distance (HD95), and dose-volume index (DVI). AM-UNet achieved mean DSCs of 0.862, 0.937, 0.823, and 0.725 for HRCTV, bladder, rectum, and sigmoid, respectively. Subjective evaluations showed 93.07% of AM-UNet predicted HRCTV were rated as clinically acceptable or needing minor adjustments, with no unacceptable cases. Dosimetric differences between AM-UNet-generated and manually delineated contours were within 1%, highlighting its potential for improving clinical workflows in brachytherapy.

Keywords Cervical cancer, Brachytherapy, Auto-segmentation, Deep learning, Computed

Abbreviations
CNN	Convolutional neural networks
DSC	Dice similarity coefficient
DVH	Dose-volume histograms
DVI	Dose-volume index
HD95	95% Hausdorff distance
HRCTV	High-risk clinical target volume
OAR	Organs at risk
ReLU	Rectified linear unit
SSM	State space model
CCA	Convolutional channel attention

Brachytherapy plays a crucial role in the comprehensive treatment of tumors, especially for cervical cancer. Accurately delineating target volumes and organs at risk (OARs) is essential for the efficacy and safety evaluating of brachytherapy[1–3]. According to reports and guidelines from the Groupe Européen de Curiethérapie (GEC) and the European Society for Radiotherapy and Oncology (ESTRO)[4–6], there is subjective inconsistency in the delineation of OARs in 3D image-guided brachytherapy, including inter-observer and intra-observer variations. Furthermore, the 3D image-guided brachytherapy process is complex and labor-intensive, with limited time for contouring. Therefore, rapid and accurate automatic delineation of target volumes and OARs is crucial in the brachytherapy for cervical cancer.

In recent years, deep learning models based on convolutional neural networks (CNNs) have significantly advanced in the automatic segmentation of various tumor radiotherapy applications[1,7–12]. Among these, the

[1]Department of Gynecology, Clinical Oncology School of Fujian Medical University, Fujian Cancer Hospital, Fuzhou 350011, Fujian, China. [2]Department of Radiation Oncology, Clinical Oncology School of Fujian Medical University, Fujian Cancer Hospital, Fuzhou, China. [3]Department of Radiation Oncology, Fujian Cancer Hospital, Clinical Oncology School of Fujian Medical University, Fuzhou 350011, Fujian, China. [4]College of Computer and Data Science, Fuzhou University, Fuzhou 350116, China. [5]Lele Zang, Jing Liu and Huiqi Zhang contributed equally to this work. ✉email: whuhandaxue@163.com; xuqin@fjmu.edu.cn

U-net architecture[13] has been widely adopted and has demonstrated exceptional performance in clinical target volume segmentation for cancers such as prostate cancer[12], cervical cancer[1,11,14], and breast cancer[8,9]. Despite their effectiveness, CNNs have inherent limitations related to their local focus, posing challenges in capturing global contextual information[15,16]. Transformer-based models adeptly capture long-range dependencies in input sequences using features like self-attention mechanisms and position encoding. However, they often require extensive computational resources, particularly when applied to high-resolution 3D medical image segmentatio[17,18].

To enhance global modeling and reduce computational overhead, Mamba, based on the state space model (SSM)[19], was designed to facilitate long-range dependency modeling while enhancing training speed and inference efficiency[20]. This approach has been extensively applied in medical image segmentation. U-Mamba[21] used this approach to medical image segmentation for the first time by constructing a novel SSM-CNN hybrid model. Xing et al.[22] introduced a 3D Mamba model by combining SSM with CNNs, achieving excellent performance in colorectal cancer segmentation.

To our knowledge, research on the automatic segmentation of HRCTV and OARs for cervical cancer brachytherapy remains sparse and largely relies on conventional methods. Therefore, we developed the AM-UNet model based on the Mamba and CNN architectures, which utilizes 3D CT as input to achieve improved automatic segmentation performance at lower computational costs. Our experimental results demonstrate that the proposed AM-UNet accurately and efficiently segments HRCTV and OARs with reduced computational costs.

Methods

The study was approved by the Ethics Committee of Fujian Cancer Hospital (No.SQ2022-191), and the requirement for informed consent was waived. All research involving human participants was conducted in accordance with the Declaration of Helsinki. The dataset included patients CT images used for model training and testing the model and the model's performance was evaluated both quantitatively and qualitatively. Additionally, the impact of automatic segmentation on dose distribution was assessed. The overall workflow of the study design is illustrated in Fig. 1.

Dataset and data preprocessing

In this study, we retrospectively analyzed 645 groups of CT scan from 179 patients who underwent brachytherapy for cervical cancer at Fujian Cancer Hospital between 2006 and 2020. Prior to each CT scan, each patient needs to take an oral contrast agent (iopamidol) for small bowel preparation. Additionally, the bladder is filled with 100 cc of saline, and CT images are obtained on a Brilliance CT Big Bore (Philips Medical Systems Inc., Cleveland, OH, USA) with a dimension of 512×512 pixels and a slice thickness of 2.5 mm.
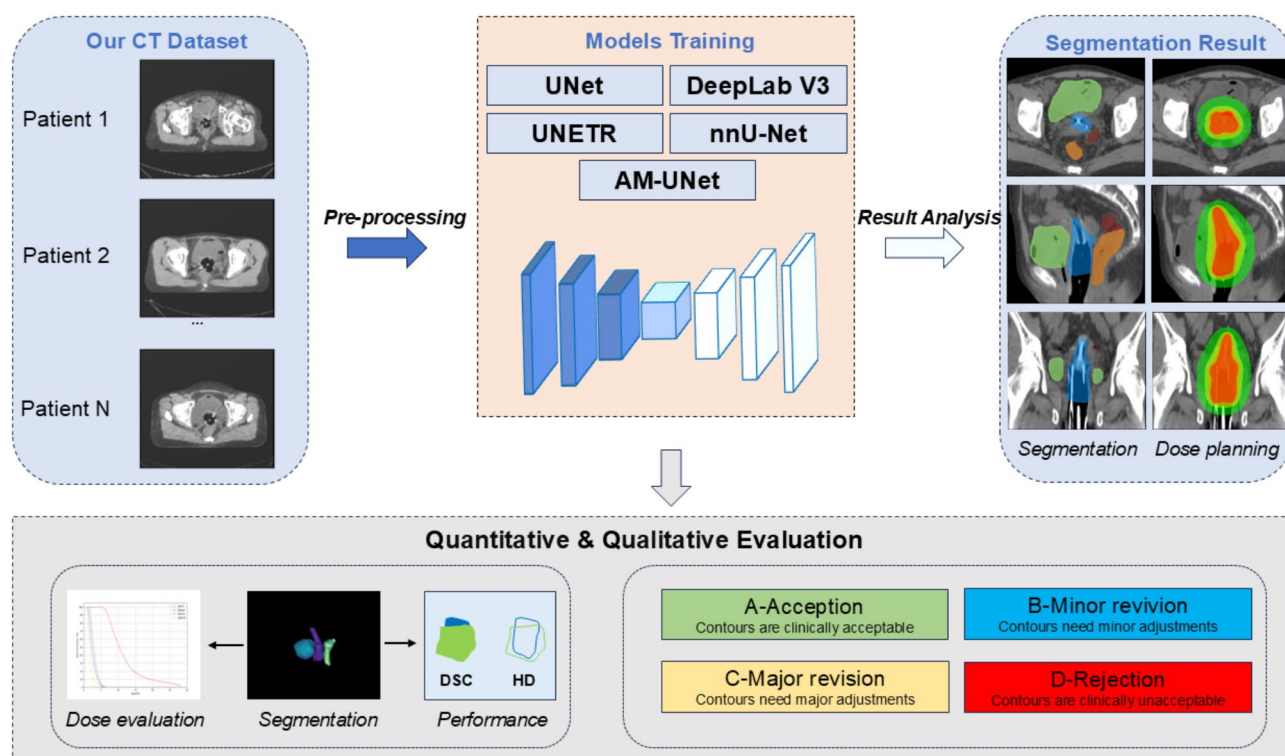

Fig. 1. Flowchart of automatic segmentation and dosimetric evaluation of the clinical target volume in this study.

According to the GEC-ESTRO definition of HRCTV, the prescription dose is administered to the surface of the uterus, cervix, parametrium and the upper vagina using intrauterine applicator and interstitial needles. The OARs in the brachytherapy plan include the bladder, rectum, and sigmoid. According to ICRU Report No. 89[23], the volumetric dose to 2–3 cm^3 of normal tissue and organ walls is linked to brachytherapy side effects. The dose exposure to normal tissue, assessed by outlining the organ's outer wall with a single line, is comparable to using a double line. The contours of the HRCTV and OARs were delineated by two experienced radiation oncologists (LLZ and JL with ten and six years of clinical experience, respectively) using 3D Slicer, under the guidance of a senior radiation oncologist (QX). To reduce the computational cost and time required for model training, we cropped all images based on the contour locations to a size of $64 \times 224 \times 224$. The radiotherapy plans were developed based on the manual contours and automatic contours by a medical physicist (JHC, with seven years of clinical experience) with the Oncentra (Nucletron, Elekta AB, Stockholm, Sweden, V.4.3). The prescribed dose were 7 Gy to the HRTCV.

3D model AM-UNet

The AM-UNet primarily consists of three components: (1) a 3D feature encoder composed of 3D Mamba blocks and downsampling convolutions, (2) a 3D decoder that performs upsampling through transposed convolutions, and (3) a channel attention mechanism composed of convolutional layers and residual connections. Figure 2 provides a detailed illustration of the overall architecture of the proposed AM-UNet. The encoder section primarily consists of a stem layer and four 3D Mamba layers. The stem layer is composed of two convolutional layers with a kernel size of 3, padding of 1, and a stride of (1) Each 3D Mamba layer comprises a Mamba layer, MLP, and a downsampling convolutional layer with a kernel size of 3, padding of 1, and a stride of (2) The specific structure of the Mamba layer is detailed in Supplementary Material 1. The convolutional channel attention (CCA) module consists of five blocks of convolutional layers and Rectified Linear Unit (ReLU) activation functions, leveraging convolutional layers and residual connections to further learn spatial relationships and inter-channel correlations within the feature maps. Finally, the output from each layer is processed through convolutional layers to generate a four-channel segmentation result.

Model training

All data were divided into training, validation, and test sets by patient in a ratio of 6:2:2. Specifically, 426 groups of CT scan from 107 patients were used for training, 137 groups of CT scan from 36 patients for validation, and 130 groups of CT scan from 36 patients for independent testing. During training, all models were optimized using the AdamW optimizer, with an initial learning rate ranging from 1e−3 to 1e−2 and a weight decay of 1e−4. The best hyperparameters for other baseline models were determined based on the settings from the original literature and further fine-tuned through our experiments. The loss function combined cross-entropy and Dice coefficient, with dynamic weight adjustment based on validation performance. Data augmentation was applied with a 30% probability, including random flipping, scaling, and other operations to enhance model generalization. The batch size was set to 2, and training was conducted over 100 epochs. The specific training
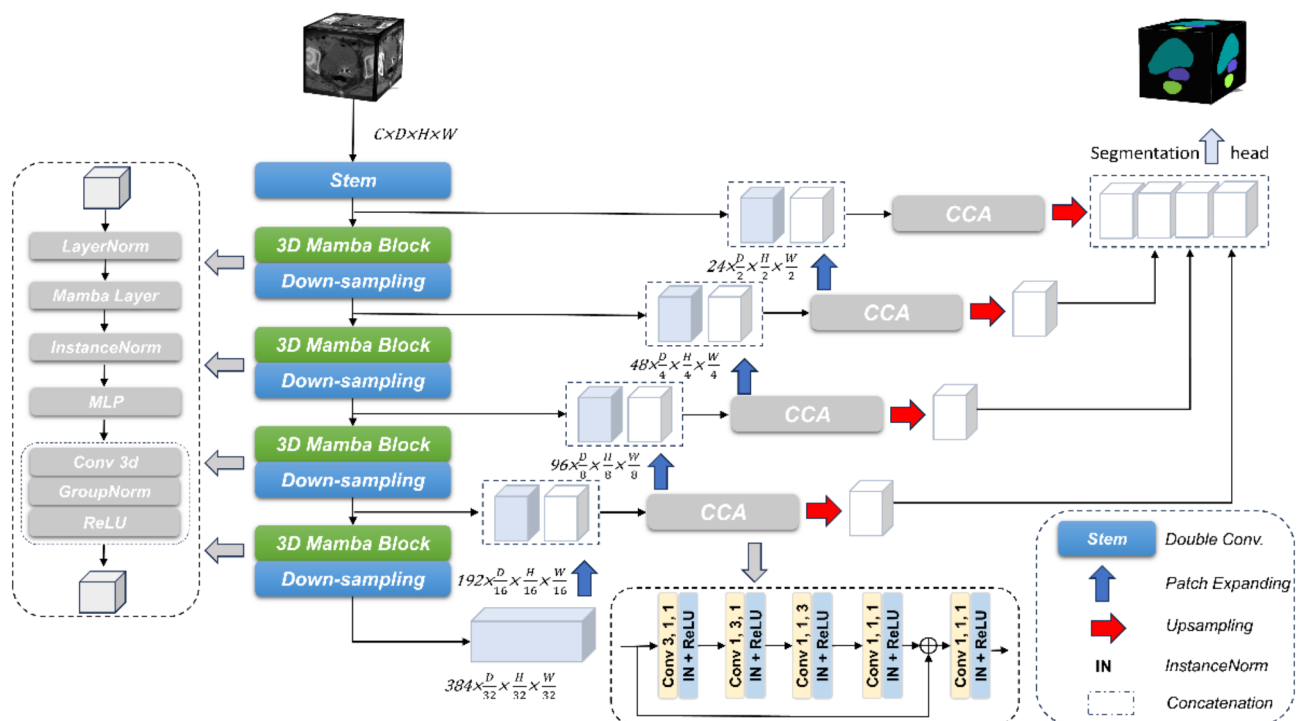

Fig. 2. The overview of the proposed AM-UNet.

hyperparameters for all models are detailed in Supplementary Material 2. To ensure reproducibility, random seeds were set for all training processes. All experiments were performed on an NVIDIA RTX A5000 24GB GPU. The code is available online: https://github.com/khuanging/AM-UNet.

Dosimetric assessment

To evaluate the impact of geometric errors in automatic segmentation on dose distribution, two independent brachytherapy plans were generated for each CT scan in the test set: (1) manually segmented contours and (2) automatically segmented contours using the AM-UNet model. Dose-volume indices (DVI) and dose-volume histograms (DVH) were used for dose quantification analysis. For HRCTV, we compared the differences in D_{mean} and $D_{90\%}$ between the two plans, where D_{mean} and $D_{90\%}$ represent the mean dose and minimum dose to 90% of the HRCTV, respectively. For OAR, D_{mean}, $D_{0.1cc}$, D_{2cc}, and D_{5cc} were used to compare dose differences between the different brachytherapy plans, where D_{xcc} represents the minimum dose received by x cm³ of the OAR. The specific brachytherapy plan structures and corresponding dosimetric indices are detailed in Supplementary Material 3.

Performance evaluation

For analyzing the performance of the automatic segmentation model, we used both quantitative metrics and subjective evaluations to assess its clinical usability. The quantitative analysis involved using the Dice similarity coefficient (DSC) and 95% Hausdorff distance (HD95) to measure the differences between manual delineations and model-predicted contours. Details of the performance indicator calculations are in Supplementary Material 4.

To further analyze the clinical applicability of the automatic segmentation model, model predictions were visually evaluated by two experienced radiation oncologists (Observer #1: LLC and Observer #2: JL) in collaboration. In this part of the analysis, the oncologists subjectively scored the prediction results of model predictions based on the following four-grade clinical criteria[11]: A—Acceptance (Contours are clinically acceptable), B—Minor Revision (Contours need minor adjustments), C—Major Revision (Contours need major adjustments), and D—Rejection (Contours are clinically unacceptable).

Statistical analysis

The nonparametric Wilcoxon 2-sided test (Wilcox) was used to compare the performance differences between different models. A p-value strictly less than 0.05 was considered statistically significant. All statistical analyses were performed using Python software (version 3.8.18, Anaconda Inc.).

Results

Automatic segmentation performance

Figure 3 shows the automatic segmentation results of HRCTV and OARs for each model in the same patient undergoing cervical cancer brachytherapy. Table 1 summarizes the specific metric results of each model on the test set. Among all models, AM-UNet demonstrated superior performance in both quantitative metrics for HRCTV and OAR segmentation. For HRCTV, AM-UNet achieved the best DSC (0.862), and HD95 (3.638 mm) significantly outperformed other models ($p = 0.108$) except nnU-Net. AM-UNet demonstrated superior performance in both DSC (0.937) and HD95 (2.578 mm) for bladder segmentation, surpassing other models ($p < 0.05$ and $p = 0.096$), except for nnU-Net. For rectum, AM-UNet also had the best metrics (DSC, 0.823; HD95, 5.312 mm), significantly outperforming other models ($p < 0.05$). Although all models performed moderately in sigmoid segmentation, AM-UNet still showed the best performance with a DSC of 0.725 and an HD95 of 18.739 mm, which was statistically significant ($p < 0.05$). The box plots of the results of different models on the test set are shown in Supplementary Material 5.

Dosimetric evaluation

Table 2 compares the dose metrics of HRCTV and OARs between AM-UNet automatic segmentation and manual contouring in brachytherapy planning. For HRCTV, the relative differences in D_{mean} and $D_{90\%}$ between manual and automatic segmentation were less than 1%, with no significant difference ($p > 0.05$). Among all OARs, the radiation doses for Bladder_$D_{0.1cc}$, Bladder_D_{2cc}, and Bladder_D_{5cc} showed significant differences ($p < 0.05$), while the other metrics showed no significant differences. Except for Bladder_$D_{0.1cc}$, which had a relatively substantial difference (− 8.36%), the other differences were all less than 5%. Figure 4 illustrates the dose distribution and DVH for cases with the best and worst DSC. In Fig. 4A, B, differences in dose distributions for the best DCS patients were minimal between manual and automated segmentation, while the worst results showed some gaps between dose distributions; in the DVH curves of Fig. 4C, D, the dosimetric curves for the best DSC nearly overlapped, while the worst DSC showed slight differences in the sigmoid and bladder dose curve.

Qualitative evaluation

The oncologist evaluated and confirmed all predictions of the test cases on a slice-by-slice basis. Figure 5 shows the experts' qualitative assessment of the model segmentation results. Among all cases in the test set, 0.77% ($n = 1$) were deemed acceptable, 63.85% ($n = 83$) required minor revisions and 35.38% ($n = 46$) required major revisions; no patients were rejected. When considering specific site evaluations, HRCTV achieved the best results (Acceptance, 15.38%; Minor revision, 77.69%; Major revision, 6.92%), followed by rectum (10.77%, 84.62%, 4.62%). In the automatic segmentation of sigmoid, one patient CT was deemed to require re-outlining. The specific results of the oncologists' qualitative scores are presented in Supplementary Material 6. Additionally, the visualization results for patients of each grade are provided in Supplementary Material 7.
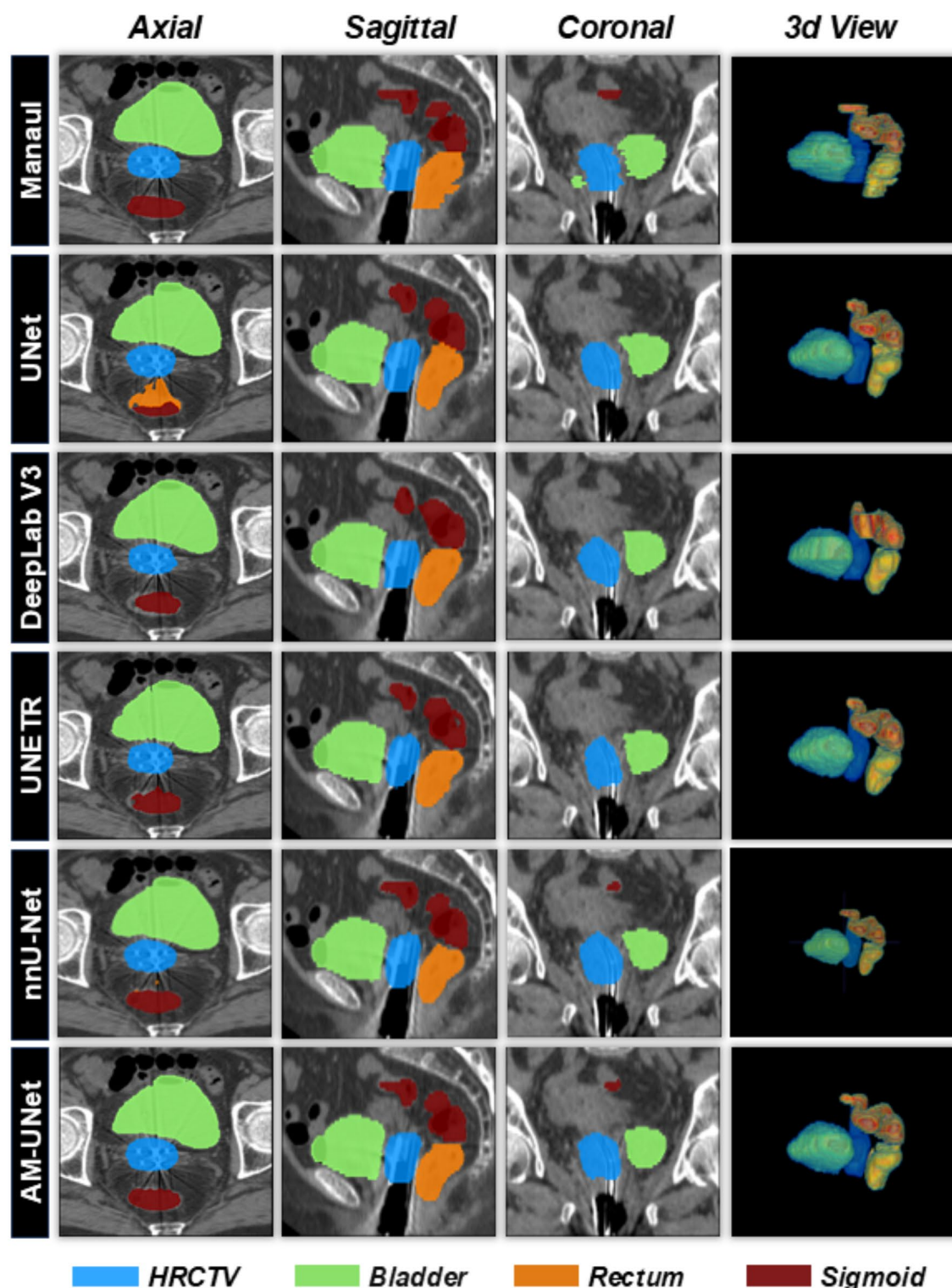
Fig. 3. Axial, sagittal, coronal, and 3d views showing HRCTV, bladder, rectum and sigmoid predicted by the model and delineated by manual. *HRCTV* high-risk clinical target volume.

Discussion

The delineation of HRCTV and OARs are crucial in brachytherapy for cervical cancer. In clinical practice, radiation oncologists are required to delineate this task volume through time-consuming and laborious manual contouring. Despite the availability of standard guidelines, the process remains highly dependent on the expertise of clinicians, and there are issues with intra- and inter-observer consistency[24]. Currently, there are few studies on the automatic segmentation of HRCTV and OARs in cervical cancer brachytherapy[1,11,25], most of

Metrics	Deep learning model				
	UNet	DeepLab V3	UNETR	AM-UNet	P-value
HRCTV					
DSC	0.860	0.849	0.827	0.862	0.10
HD95 (mm)	4.325	4.311	5.013	3.638	< 0.01
Bladder					
DSC	0.936	0.925	0.910	0.937	< 0.01
HD95 (mm)	2.583	2.881	3.532	2.578	0.08
Rectum					
DSC	0.814	0.774	0.786	0.823	< 0.01
HD95 (mm)	5.510	6.102	5.834	5.312	0.01
Sigmoid					
DSC	0.708	0.632	0.664	0.725	< 0.01
HD95 (mm)	21.829	21.497	24.461	18.739	0.02

Table 1. Summary of quantitative metrics for HRCTV and OAR automatic segmentation results on the test set by different models. *HRCTV* high-risk clinical target volume, *OAR* organs at risk, *DSC* dice similarity coefficient, *HD95* 95% Hausdorff DISTANCE.

OAR or target	Dosimetric parameters	Manual (Gy)	Auto (Gy)	Difference with manual	P-value
HRCTV	D_{mean}	12.377	12.260	− 0.95%	0.06
	$D_{90\%}$	6.976	6.994	0.26%	0.09
Bladder	D_{mean}	2.292	2.307	0.66%	0.86
	$D_{0.1\,cc}$	5.977	5.477	− 8.36%	< 0.01
	D_{2cc}	4.890	4.697	− 3.95%	< 0.01
	D_{5cc}	4.352	4.221	− 2.99%	0.01
Rectum	D_{mean}	2.752	2.776	0.85%	0.66
	$D_{0.1\,cc}$	5.743	5.456	− 5.00%	0.11
	D_{2cc}	4.277	4.211	− 1.53%	0.63
	D_{5cc}	3.519	3.507	− 0.37%	0.96
Sigmoid	D_{mean}	1.793	1.804	0.66%	0.50
	$D_{0.1\,cc}$	3.681	3.575	− 2.88%	0.34
	D_{2cc}	2.654	2.632	− 0.83%	0.84

Table 2. Mean dose for each clinical target volume and assessment segment of the OAR calculated from the manual contour (Manual), AM-UNet (Auto), and minimum value of the HRCTV. *OAR* organ at risk, *HRCTV* high-risk clinical target volume.

which use traditional convolutional neural networks. Therefore, we collected 645 groups of CT scan from 179 cervical cancer patients receiving brachytherapy at our institution. We proposed a novel deep learning network, AM-UNet, for the automatic segmentation of HRCTV and OARs.

The Mamba architecture, based on SSM, offers a novel design that combines the ability to capture long-range dependencies akin to Transformers, without being constrained to the local features of 3D images as seen in traditional CNNs[21,26,27]. Unlike the self-attention mechanism in Transformers, which involves quadratic computational complexity, Mamba significantly reduces computational demands, making it particularly advantageous for tasks like 3D image segmentation that are computationally intensive.

We found that the quantitative metrics for HRCTV, bladder, rectum, and sigmoid predicted by AM-UNet were superior to other models, except for nnU-Net. Due to nnU-Net's unique adaptive training and data processing methods, it outperforms AM-UNet in certain aspects. Among these, bladder segmentation performed exceptionally well across all deep-learning models, primarily due to the bladder's distinct shape and volume compared to HRCTV and other OARs. In contrast, all models performed poorly in segmenting the sigmoid. However, this outcome is consistent with findings from other studies[2,28,29], which have noted that the sigmoid is challenging to segment due to its complex anatomy and low image contrast[2,30]. AM-UNet, based on the Mamba architecture, effectively captures long-range three-dimensional features in CT images, which enhances segmentation performance for larger or irregularly shaped organs like the bladder, rectum, and sigmoid. Additionally, for HRCTV, which has blurred boundaries and a complex structure, the model employs convolutional attention and multi-scale feature fusion for more precise feature aggregation. As shown in Fig. 3, AM-UNet demonstrates better segmentation performance compared to other models.
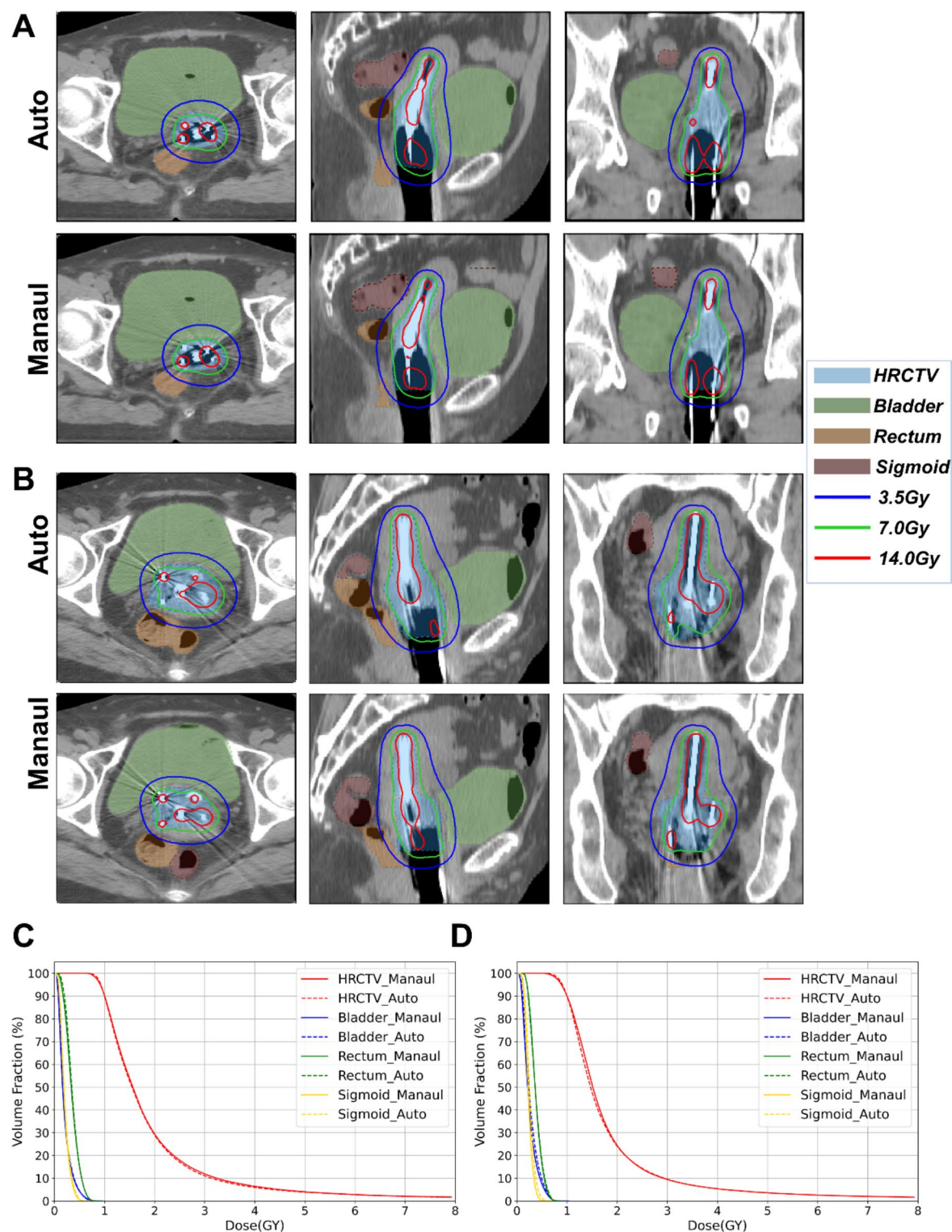
Fig. 4. Dose distributions and DVH curves for treatment evolution using manual contouring and automatic segmentation of contours. (**A,B**) Measurement distribution plots for the best and worst test patients for DSC, (**C,D**) DVH curves for DSC best and worst test patients. *DVH* dose-volume histogram, *DSC* dice similarity coefficient.

In their study on automatic CT segmentation for cervical cancer brachytherapy patients, Zhang et al.[28] introduced the DSD-UNET model, achieving DSC of 0.83, 0.87, 0.82, and 0.65 for HRCTV, bladder, rectum, and sigmoid, respectively. These values are lower compared to our study, possibly due to limitations in CNN's ability to extract spatial contextual features, thereby struggling with the varied shapes and volumes of OARs in
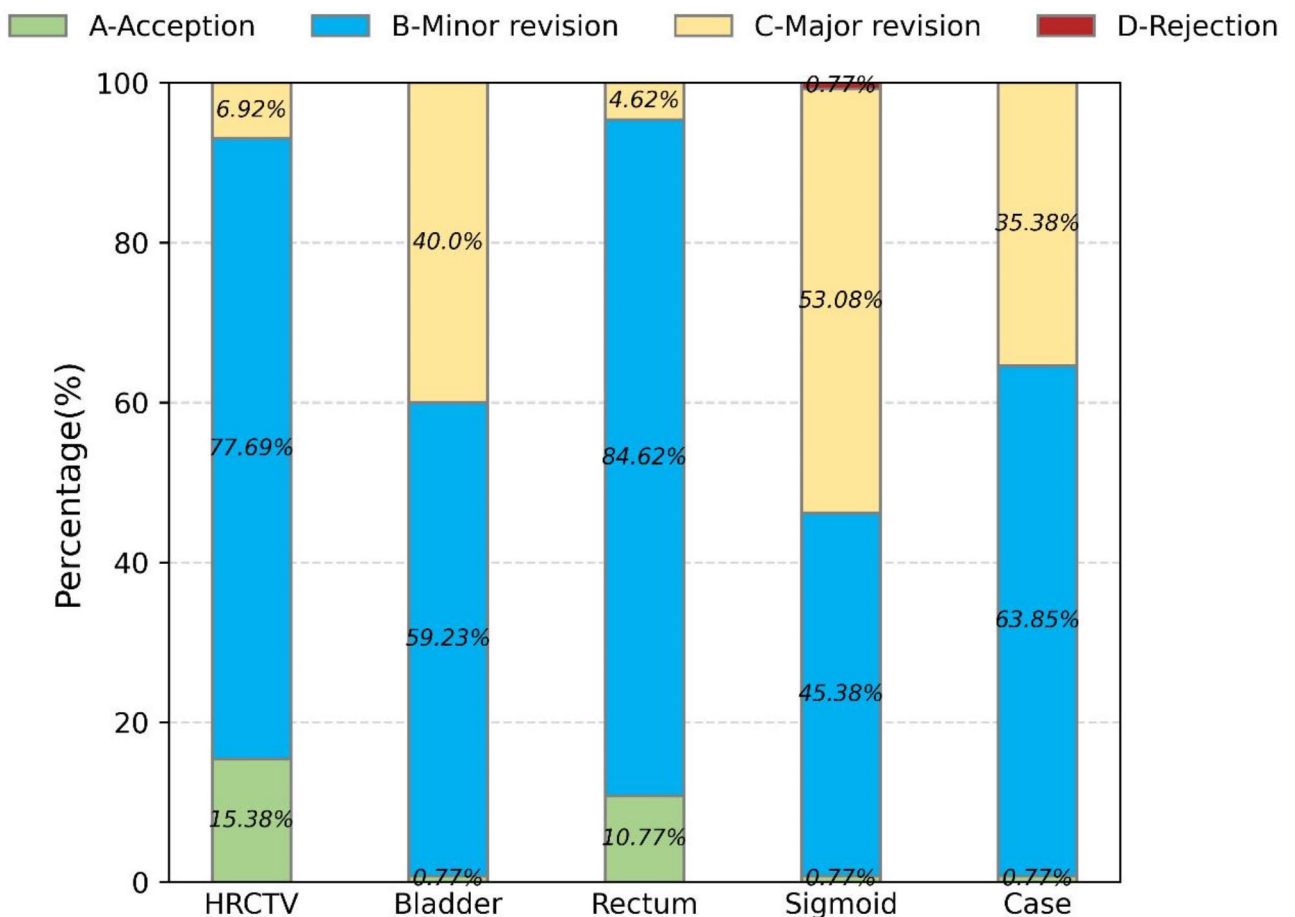
Fig. 5. Stacked histograms show the qualitative ratings of tumor experts on the segmentation results of AM-UNet.

CT images. On the other hand, Wang et al.[2] proposed an enhanced CNN model, achieving higher DSC values of 0.87, 0.94, 0.86, and 0.79 for HRCTV, bladder, rectum, and sigmoid, respectively, slightly outperforming our results. However, their study was limited by a small sample size of only 10 cases in the test group, making it challenging to fully assess the model's generalizability and stability. This issue is prevalent in deep learning research, particularly in medical imaging[30–32]. Deep-learning models require independent testing with broad datasets. Personalized studies based on multiple CT scans per patient better reflect clinical realities, confirming the effectiveness of our model in practical clinical applications.

Dose assessment is more informative than geometric assessment in automated segmentation studies for brachytherapy. Yoganathan et al.[33] found that geometric metrics for the sigmoid were poor but highly consistent with manually delineated dosimetry. This finding is similar to our study, where all dose metrics for the sigmoid showed no significant differences ($p > 0.05$). In Wang et al.'s study[2], substantial dose variations were observed in the bladder and rectum, similar to our findings where, despite favorable geometric parameters of the bladder, dosimetric parameters such as bladder_$D_{0.1cc}$, bladder_D_{2cc}, and bladder_D_{5cc} exhibited significant differences. These results emphasize the importance of considering both geometric and dosimetric parameters when assessing automatic segmentation performance, and the critical role of experienced oncology radiation experts in reviewing particularly high-dose region structures.

To further demonstrate the clinical applicability of AM-UNet, we conducted a subjective evaluation of the segmentation results. Upon comparing quantitative metrics with subjective evaluation results, we found that although the bladder had better results in the quantitative analysis, it had poorer results in the subjective assessment (only one patient was deemed clinically acceptable). This discrepancy may be because the bladder's distinctiveness prompts oncologists to adopt stricter evaluation criteria. Meanwhile, HRCTV performed best in subjective evaluation, with 15.38% of cases deemed clinically acceptable and 77.69% requiring only minor adjustments (Supplementary Material 6). This can be attributed to the complexity of HRCTV, as there is no detailed consensus or standard for its contouring. Different clinical centers or oncologists may have varying standards and preferences for HRCTV delineation.

One significant advantage of automated HRCTV and OARs segmentation is the reduction in manual contouring time for radiation oncologists, decreasing the complexity of the brachytherapy process. The average time to manually contour HRCTV and OARs for a cervical cancer patient is 90–120 min[34], while our AM-UNet model requires only 5 s. Even considering potential further manual adjustments to the contours, the

total processing time per patient does not exceed 10 min on average. The AM-UNet significantly simplifies the treatment workflow and reduces subjective inconsistencies among experts.

Despite the excellent results achieved in our work, all current data come from a single center. In the future, it will be necessary to include more centers for independent testing to evaluate the model's generalization performance. Although our model performed best in the segmentation of the sigmoid, the variability of this structure necessitates further consideration of data processing and model structure optimization to enhance its segmentation ability. On the other hand, considering our success in CT images of patients with cervical cancer patients, this approach may apply to the volumetric segmentation of other cancers or MRI images. We will explore this possibility in future work.

Conclusion

The delineation of HRCTV and OARs is a critical component of brachytherapy planning, directly impacting dose distribution to these organs and target volumes. Automatic segmentation technology enables rapid and repeatable contour delineation, effectively reducing the time and labor costs associated with the treatment process. The proposed AM-UNet model achieved good consistency between automatic and manual segmentation contours for HRCTV and OARs in cervical cancer brachytherapy, although some results still have clinical acceptability issues. Nonetheless, this study offers an efficient solution for the demanding clinical workload.

Data availability

The datasets generated and/or analysed during the current study are not publicly available due to hospital ethics committee regulations and potential involvement of patient privacy, but are available from the corresponding author on reasonable request.

References

1. Mohammadi, R. et al. Deep learning-based auto-segmentation of organs at risk in high-dose rate brachytherapy of cervical cancer. *Radiother. Oncol.* **159**, 231–240 (2021).
2. Wang, J. et al. Evaluation of auto-segmentation for brachytherapy of postoperative cervical cancer using deep learning-based workflow. *Phys. Med. Biol.* **68** (5), 055012 (2023).
3. Rodríguez Outeiral, R., González, P. J., Schaake, E. E., van der Heide, U. A. & Simões, R. Deep learning for segmentation of the cervical cancer gross tumor volume on magnetic resonance imaging for brachytherapy. *Radiat. Oncol.* **18** (1), 1–7 (2023).
4. Saarnak, A. E., Boersma, M., van Bunningen, B. N. F. M., Wolterink, R. & Steggerda, M. J. Inter-observer variation in delineation of bladder and rectum contours for brachytherapy of cervical cancer. *Radiother. Oncol.* **56** (1), 37–42 (2000).
5. Hellebust, T. P., Dale, E., Skjønsberg, A. & Olsen, D. R. Inter fraction variations in rectum and bladder volumes and dose distributions during high dose rate brachytherapy treatment of the uterine cervix investigated by repetitive CT-examinations. *Radiother. Oncol.* **60** (3), 273–280 (2001).
6. Duane, F. K. et al. Impact of delineation uncertainties on dose to organs at risk in CT-guided intracavitary brachytherapy. *Brachytherapy* **13** (2), 210–218 (2014).
7. An Automated Treatment Planning Framework for Spinal Radiation Therapy and Vertebral-Level Second Check. *Int. J. Radiat. Oncol. Biol. Phys.* . **114**(3), 516–528 (2022).
8. Buelens, P. et al. Clinical evaluation of a deep learning model for segmentation of target volumes in breast cancer radiotherapy. *Radiother. Oncol.* **171**, 84–90 (2022).
9. Choi, M. S. et al. Clinical evaluation of atlas- and deep learning-based automatic segmentation of multiple organs and clinical target volumes for breast cancer. *Radiother. Oncol.* **153**, 139–145 (2020).
10. Deep Learning For Automatic Gross Tumor Volumes. Contouring in esophageal cancer based on Contrast-Enhanced computed tomography images: A Multi-Institutional study. *Int. J. Radiat. Oncol. Biol. Phys.* (2024).
11. Liu, Z. et al. Development and validation of a deep learning algorithm for auto-delineation of clinical target volume and organs at risk in cervical cancer radiotherapy. *Radiother. Oncol.* **153**, 172–179 (2020).
12. Kawula, M. et al. Dosimetric impact of deep learning-based CT auto-segmentation on radiation therapy treatment planning for prostate cancer. *Radiat. Oncol.* **17** (1), 1–12 (2022).
13. Ronneberger, O., Fischer, P. & Brox, T. U-Net: convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*, 234–241. https://doi.org/10.1007/978-3-319-24574-4_28 (Springer, 2015).
14. Automatic segmentation using deep learning to enable online dose optimization during adaptive radiation therapy of cervical cancer. *Int. J. Radiat. Oncol. Biol. Phys.* **109**(4), 1096–1110 (2021).
15. Xu, Z. et al. Single-Branch CNN with transformer semantic information for Real-Time segmentation. arXiv; [cited 2024 Feb 29]. http://arxiv.org/abs/2312.17071 (2024).
16. Xie, Y., Zhang, J., Shen, C., Xia, Y. & CoTr Efficiently bridging CNN and transformer for 3D medical image segmentation. In *Medical Image Computing and Computer Assisted Intervention—MICCAI 2021*, 171–180. https://doi.org/10.1007/978-3-030-87199-4_16 (Springer, 2021).
17. Tang, F. et al. arXiv.org. 2023 [cited 2024 Jan 12]. MobileUtr: Revisiting the relationship between light-weight CNN and Transformer for efficient medical image segmentation. https://arxiv.org/abs/2312.01740v1
18. Ruan, J. & Xiang, S. arXiv.org. 2024 [cited 2024 Mar 2]. VM-UNet: Vision Mamba UNet for Medical Image Segmentation. https://arxiv.org/abs/2402.02491v1
19. Basar, T. A new approach to linear filtering and prediction problems. In *Control Theory: Twenty-Five Seminal Papers (1932–1981)*, 167–179. https://ieeexplore.ieee.org/abstract/document/5311910 (IEEE, 2001).
20. Gu, A. & Dao, T. arXiv.org. 2023 [cited 2024 May 7]. Mamba: Linear-Time Sequence Modeling with Selective State Spaces. https://arxiv.org/abs/2312.00752v1
21. Ma, J., Li, F. & Wang, B. arXiv.org. 2024 [cited 2024 May 7]. U-Mamba: Enhancing Long-range Dependency for Biomedical Image Segmentation. https://arxiv.org/abs/2401.04722v1
22. Xing, Z., Ye, T., Yang, Y., Liu, G. & Zhu, L. arXiv.org. 2024 [cited 2024 Mar 26]. SegMamba: Long-range Sequential Modeling Mamba For 3D Medical Image Segmentation. https://arxiv.org/abs/2401.13560v3
23. Report 89. Journal of the ICRU [Internet]. Apr [cited 2024 Dec 24];13(1–2). https://pubmed.ncbi.nlm.nih.gov/27335498/ (2013).

24. Eminowicz, G. & McCormack, M. Variability of clinical target volume delineation for definitive radiotherapy in cervix cancer. *Radiother. Oncol.* **117** (3), 542–547 (2015).
25. Tian, M. et al. Delineation of clinical target volume and organs at risk in cervical cancer radiotherapy by deep learning networks. *Med. Phys.* **50** (10), 6354–6365 (2023).
26. S Z, L L, Q L, J L, Y S, Q X. Integrating a deep neural network and transformer architecture for the automatic segmentation and survival prediction in cervical cancer. *Quant. Imaging Med. Surg..* **14**(8) (2024).
27. Ruan, J. & Xiang, S. arXiv.org. 2024 [cited 2024 Mar 2]. VM-UNet: Vision Mamba UNet for Medical Image Segmentation. https://arxiv.org/abs/2402.02491v1
28. Zhang, D. et al. Automatic segmentation and applicator reconstruction for CT-based brachytherapy of cervical cancer using 3D convolutional neural networks. *J. Appl. Clin. Med. Phys.* **21** (10), 158–169 (2020).
29. Jiang, X., Wang, F., Chen, Y. & Yan, S. RefineNet-based automatic delineation of the clinical target volume and organs at risk for three-dimensional brachytherapy for cervical cancer. *Ann. Transl. Med.* **9** (23), 1721–1721 (2021).
30. Semi-automatic sigmoid. Colon segmentation in CT for radiation therapy treatment planning via an iterative 2.5-D deep learning approach. *Med. Image. Anal.* **68**, 101896 (2021).
31. Ramesh, K. K. et al. A fully automated post-surgical brain tumor segmentation model for radiation treatment planning and longitudinal tracking. *Cancers* **15** (15), 3956 (2023).
32. Said, D. et al. Semiautomated segmentation of hepatocellular carcinoma tumors with MRI using convolutional neural networks. *Eur. Radiol.* **33** (9), 6020–6032 (2023).
33. Yoganathan, S. et al. Automatic segmentation of magnetic resonance images for high-dose-rate cervical cancer brachytherapy using deep learning. *Med. Phys.* **49** (3), 1571–1584 (2022).
34. Chen Kaiqiang, C. et al. Systematic evaluation of atlas-based autosegmentation (ABAS) software for adaptive radiation therapy in cervical cancer. *Chin. J. Radiol. Med. Prot.* **35** (2), 111–113 (2015).

Acknowledgements

Author contributions

For this research paper with several authors, a short paragraph specifying their individual contributions was provided. X.Q. and C.J. developed the original idea and contributed to the research design for the study. Z.H. and Z.M. and W.Y. and K.Y. were responsible for data collecting. Z.L. and L.J. provided guidance and improving suggestion. Z.S., and Q.X. provided some guidance for the writing of the article. X.Q and Z.L provided ideas for the mapping in this paper. All authors have read and approved the final manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-94431-1.

Correspondence and requests for materials should be addressed to J.C. or Q.X.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.