Original Article
Yonsei Med J 2026 Apr;67(4):358-367
https://doi.org/10.3349/ymj.2025.0133
Yonsei Medical Journal
YMJ
pISSN: 0513-5796 · eISSN: 1976-2437

Deep Learning for Automated Tumor Segmentation of Rectal Cancer on T2-Weighted Magnetic Resonance Images

Miri Seo[1], YongDae Lee[2], and Myung-Won You[3]

[1]Department of Medicine, Kyung Hee University College of Medicine, Seoul;
[2]Department of Radiology, Kyung Hee University Hospital, Kyung Hee University College of Medicine, Seoul;
[3]Department of Radiology, Severance Hospital, Yonsei University College of Medicine, Seoul, Korea.

Purpose: To develop a deep learning (DL)-based automated segmentation model for rectal cancer on T2-weighted (T2W) magnetic resonance (MR) images.

Materials and Methods: A total of 458 patients who underwent baseline rectal MR imaging were retrospectively enrolled. An experienced radiologist labeled the tumor on each slice of the selected T2W axial images covering the entire tumor mass, and tumor volume was measured during labeling. Another radiologist labeled the rectum on the same images. Attention U-Net was trained on the T2W images from the training dataset to classify each voxel as tumor or non-tumor. Segmentation with or without rectum guidance was compared using evaluation metrics, including the Dice similarity coefficient (DSC), in the test dataset.

Results: The tumor segmentation DSC without rectum guidance was 73.35% (71.97–74.71); that with the predicted rectum was 71.79% (70.53–73.05); and that with the overlaid rectum was 75.52% (74.32–76.62), which was the highest among the three segmentation models. Tumor volume showed a positive correlation with tumor segmentation accuracy with the predicted rectum (r=0.333, p=0.001; overlaid rectum, r=0.319, p=0.002) and without rectum guidance (r=0.219, p=0.036).

Conclusion: A DL-based automated segmentation model can predict RC with over 70% accuracy, which was further improved with overlaid rectum, and larger tumor volume.

Key Words: Deep learning, artificial intelligence, tumor segmentation, rectal cancer, magnetic resonance imaging

INTRODUCTION

Magnetic resonance (MR) imaging plays a critical role in the evaluation of rectal cancer (RC), not only for tumor diagnosis and staging but also for preoperative planning, treatment response assessment, post-treatment restaging, and prognosis prediction. T2-weighted (T2W) MR imaging is an essential component of the MR protocol due to its superior soft tissue contrast, which provides detailed anatomic depiction of RC and its surrounding areas.[1] MR evaluation of RC involves assessment of T staging and resectability to determine a suitable treatment plan, including whether to undergo neoadjuvant chemoradiation therapy (CRT) or upfront surgery. Accurate identification of the cancerous lesion is important both for preoperative planning and for determining the target region for radiation therapy in these patients.

In recent years, MR image-based artificial intelligence (AI) models have been developed and increasingly applied in many fields related to locally advanced RC. Machine learning-based AI analyzes MR image data to make effective decisions and predictions at each step of RC management, such as tumor staging,[2,3] segmentation,[4,5] post-CRT restaging,[6] treatment response assessment,[7] and prognosis prediction.[8] Tumor segmentation

Received: May 23, 2025 **Revised:** October 16, 2025
Accepted: November 18, 2025 **Published online:** February 12, 2026
Corresponding author: Myung-Won You, MD, Department of Radiology, Severance Hospital, Yonsei University College of Medicine, 50-1 Yonsei-ro, Seodaemun-gu, Seoul 03722, Korea.
E-mail: twrad2282@yuhs.ac

•The authors have no potential conflicts of interest to disclose.

is the preliminary step for these AI applications in RC management. The ability to accurately detect and segment tumors on MR images is critical for diagnosis and treatment planning, as AI applications increasingly rely on such segmentation for training and implementation. Indeed, segmentation of RC on MR images is a time-consuming process, and the accuracy largely depends on the reader's knowledge and experience. In this regard, AI-driven automated segmentation is a promising solution for reducing the time and labor required for these tedious tasks and for providing a standardized method of identifying tumors. It enables quantitative analysis of tumor volume and signal intensity, thereby facilitating the quantitative response assessment such as changes in tumor volume and radiomics. Deep learning (DL), one of the fundamental technologies of AI, enables the construction of highly effective machine learning algorithms based on extracted features. The field of medical segmentation is dominated by DL-based approaches, among which U-Net has shown the strongest performance for medical image segmentation. DL-based segmentation algorithms are end-to-end structures in which, once the model architecture is complete, radiologists only need to focus on the input and output ends of the model during training and application. The attention U-Net model incorporates attention gates into the standard U-Net architecture to highlight salient features that are passed through skip connections.[9] The addition of an attention mechanism to U-Net allows the convolutional neural network to focus on the region of interest (ROI) while suppressing features in the background region, thereby improving the model's classification performance.[10] It is inspired by the human biological system, which tends to focus on unique parts when processing large amounts of information. The DL-based segmentation algorithms currently achieve excellent performance, and are widely applied in medical image processing.[4,11] In this study, we attempted to develop a DL-based automatic segmentation model using attention U-Net in patients with RC. In addition, we investigated factors influencing model performance, such as background rectum information and tumor volume.

MATERIALS AND METHODS

Patients

This retrospective study was approved by the institutional review board (Kyung Hee University Hospital, IRB No. 2025-03-002), and informed consent was waived. Between July 2007 and November 2018, 563 patients who underwent baseline rectal MR imaging for RC were identified. After excluding 105 patients (n=15 with tumor location in the sigmoid colon, n=7 with metallic stent insertion, n=60 with no primary rectal tumor or only a seeding mass, n=23 with metallic artifacts or bowel motion artifacts), a total of 458 patients who underwent rectal MR imaging were included (Fig. 1).

Protocol of MR images

MR images were acquired using 3T MAGNETOM Vida (Siemens Healthineers Ltd., Erlangen, Germany) and 3T Achieva (Philips Healthcare, Amsterdam, Netherlands) with a 32-channel coil. The protocol for rectal MR consisted of axial, coronal, and sagittal planes of T2W images; T1-weighted (T1W) axial images with or without fat suppression; diffusion-weighted images (DWI, b=0, 50, 700) with apparent diffusion coefficient maps; and contrast-enhanced T1W axial images. Rectal gel was not administered during rectal MR acquisition, as it can affect measurement of the distance between the tumor and the mesorectal fascia. Either a T2W true axial scan obtained perpen-

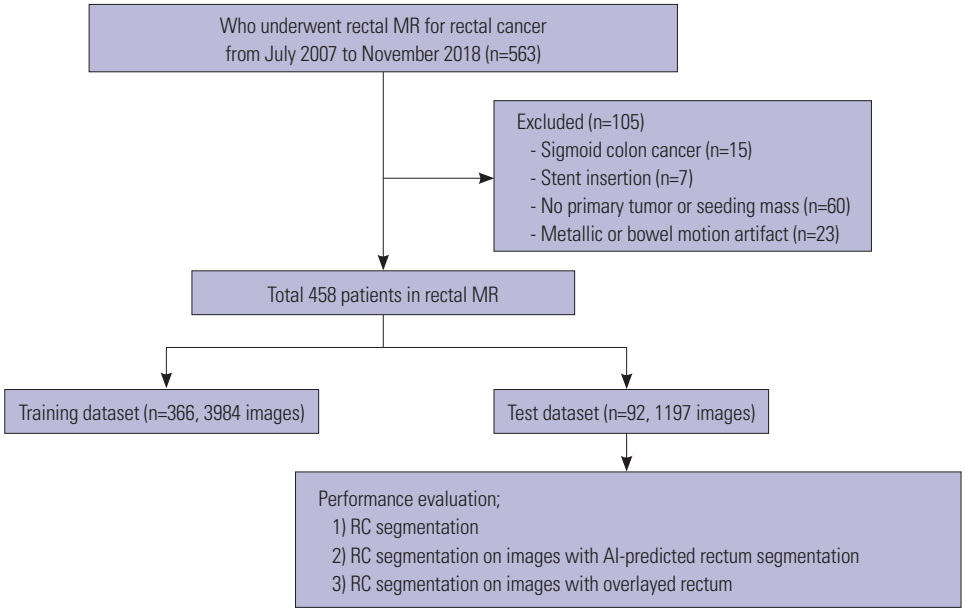

Fig. 1. Selection of patient data. MR, magnetic resonance; AI, artificial intelligence; RC, rectal cancer.

dicular to the long axis of the tumor or a T2W axial scan (when a true axial scan was unavailable) was selected for tumor and rectum labeling. The detailed scan parameters of T2W axial images are listed in Table 1.

Tumor and rectum labeling

An abdominal radiologist (M.W.Y., with 15 years of experience in rectal imaging) manually labeled the entire tumor mass on each slice of the T2W axial or true axial images using MEDIP Pro software (v2.0.0, MEDICALIP Co. Ltd., Seoul, Korea). The ROI was drawn along the margin of the tumor signal so as not to include non-rectal tissue or the normal rectum. The margin of tumor mass was determined by reviewing other MR sequences, such as DWI or the coronal and sagittal planes of T2W images, as appropriate. Another radiologist (Y.D.L., with 4 years of training) independently drew the ROI along the rectum, blinded to the tumor labeling. The rectum was labeled from the rectosigmoid junction (where the sigmoid take-off is visible) to the anorectal junction (superior margin of the puborectalis), and a few additional images were included or excluded in the longitudinal plane, depending on tumor location. When the tumor mass involved the entire rectal wall, the ROI was drawn to cover the presumed outer margin of the rectum, with reference to the upper and lower slices of T2W axial images.

DL-based automated tumor segmentation using attention U-Net

Raw MRI data of the included patients in Digital Imaging and Communications in Medicine (DICOM) format were retrieved from the Picture Archiving and Communication System of our institution. The ground truth labels for the rectum and tumor corresponding to the raw MRI data in Neuroimaging Informatics Technology Initiative (NIfTI) format (.nii) were retrieved from MEDIP Software (v2.0.0, MEDICALIP Co. Ltd., Seoul, Korea), which served as the annotation tool. Data preprocessing was carried out prior to training. As a first step, the MR images in DICOM format and the reference masks for the rectum and tumor (ground truth labels) in NIfTI (.nii) format were converted into Portable Network Graphics (PNG) images using Python-based algorithms for DL training. Following the conversion, the PNG MRI images and corresponding PNG tumor and rectum images were reviewed to ensure proper alignment. The verified PNG images were then rescaled to 256×256 pixels according to the input requirements of the training model. The entire patient dataset was split into training and test datasets in an 8:2 ratio, and 10% of the training dataset was used as the validation dataset.

Training was executed in Google Colaboratory (Colab) using NVIDIA Tesla T4 Tensor Core Graphics Processing Unit (GPU) with 16 GB of memory. Hyperparameters included a learning rate of 1e-3, a batch size of 16, 85 epochs, and Dice similarity coefficient (DSC) loss as the loss function. To enhance data variability and prevent overfitting, a data augmentation process was applied, including rotation range=0.2, width shift range=0.05, height shift range=0.05, shear range=0.05, zoom range=0.05, horizontal flip=true, and fill mode="nearest."

Attention U-Net is a DL model consisting of downsampling (encoding), skip connections using attention gates, and upsampling (decoding), and it can be effectively utilized in medical research, such as lesion or organ segmentation. Specifically, downsampling is a compression process that extracts features and reduces the spatial dimensions of an input image through convolution and pooling operations. Each convolution block in the compression path consists of consecutive 3×3 convolutions, followed by a ReLU activation unit and a max pooling layer. This structure is iteratively applied multiple times. Attention gates filter the features propagated through the skip connections, allowing only the most relevant information to pass forward. In the subsequent expansion path, each stage employs 2×2 deconvolution to upsample the feature maps. The upsampled feature maps are concatenated with their corresponding counterparts from the compression path via skip connections. Segmentation is performed through the upsampling process, which restores the processed data from the skip connection pathway to the original image size.[9,12]

The primary objective was tumor segmentation, but three different training strategies were employed. 1) Tumor segmen-

Table 1. Acquisition Parameters for the T2W Magnetic Resonance Images

Parameters	Sequences			
	T2W true short axial		T2W axial	
	Philips achieva (n=283)	Siemens vida (n=23)	Philips achieva (n=133)	Siemens vida (n=20)
Field strength	3T	3T	3T	3T
Field of view (mm)	160×160	160×160	220×220	220×220
Voxel size (mm^3)	0.25×0.25×3.0	0.2×0.2×3.0	0.34×0.34×3.0	0.2×0.2×3.0
Repetition time/echo time (ms)	3111–4800/90	3300–4470/102	3200–4800/80	4170–4280/92
Number of slices	35–40	31–40	35–40	35–39
Slice thickness (mm)	3	3	4–5	4
Slice gap (mm)	0-0.3	0.3	0–0.5	0.4
Number of signal averages	1	2	1	2

T2W, T2-weighted.

https://doi.org/10.3349/ymj.2025.0133

tation without guidance of the rectum: the model was trained using PNG MRI images and corresponding labeled tumor masks. 2) Tumor segmentation using the predicted outer margin of the rectum on MR images: rectum segmentation was trained as a first step (predicting the outer rectal margin), and the predicted rectum masks generated by a trained model were provided during tumor segmentation training. 3) Tumor segmentation using an overlaid rectum label on MR images: rectum masks labeled with the outer rectal margin were provided as anatomical guidance when training tumor segmentation, and segmentation was performed on this overlaid rectum. For training types 2) and 3), the influence of rectum masks on tumor segmentation performance was evaluated (Fig. 2).

Evaluation metrics

The performance of automated tumor segmentation was evaluated using quantitative parameters such as binary accuracy, intersection over union (IoU), and DSC; "accuracy" indicates the ratio of correctly classified pixels to the total number of pixels, "IoU" indicates the ratio of the intersection of the true (G) and predicted (P) values to the union of the true and predicted values, and "DSC" indicates a measure of the similarity between the two sets, and is used to measure the similarity between network segmentation results (P) and the gold standard (G).[10,13]

$$DSC=\frac{2|G\cap P|}{|G|+|P|},\ IoU=\frac{|G\cap P|}{|G|+|P|-|G\cap P|}$$

where |G| and |P| denote the number of pixels in the ground truth and the predicted result, and |G ∩ P| represents the number of common pixels in both pixel sets.

Precision, recall, and Hausdorff distance were also calculated to evaluate the segmentation model's performance; "precision" represents the proportion of correctly predicted disease pixels in the automated segmentation results. "Recall" represents the proportion of disease pixels in the ground truth that are correctly predicted. The "95th percentile Hausdorff distance" is a distance metric that measures the maximum of the minimum distances between the predicted segmentation and the ground truth at the 95th percentile, and was computed as a complementary parameter.[14]

Statistical analysis

Categorical data are presented as percentages, while numerical data are presented as means with standard deviations or medians with interquartile ranges, as appropriate. Indeterminate or missing data were excluded from the analysis. Repeated-measures analysis of variance was used to compare model performance among the three segmentation methods, and pairwise comparisons were performed as a post hoc test with Bonferroni correction. Pearson correlation analysis was performed to evaluate the association between tumor volume and tumor segmentation accuracy. Additionally, we investigated the correlation between tumor volume and tumor segmentation accuracy in a subgroup stratified by tumor volume with a cutoff of 20 cm^3. The cutoff value for the subgroup was

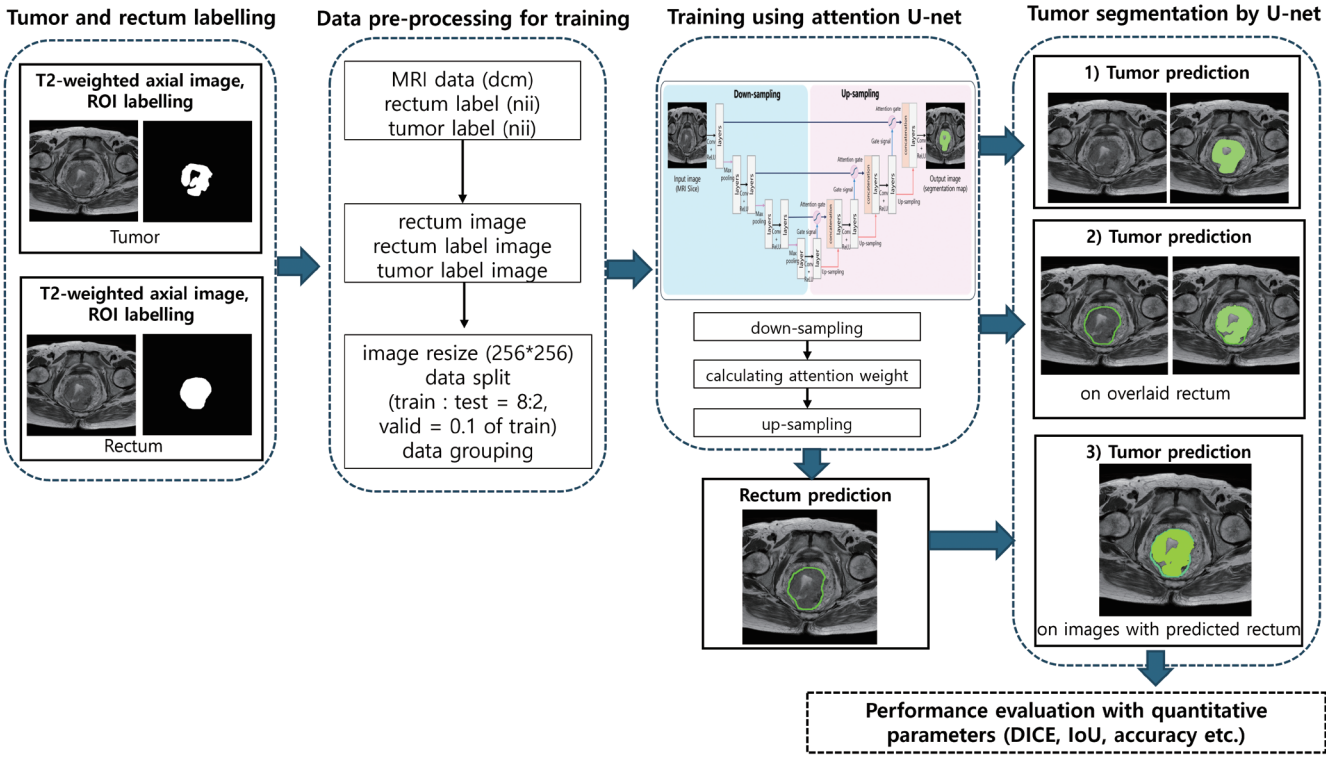

Fig. 2. Development of a deep learning-based automated segmentation model using attention U-Net.

determined based on the mean and median values of the tumor volume distribution in order to create groups of equal size. Statistical analyses were conducted using Python 3.9.0 and MedCalc (ver. 20.111 MedCalc software Ltd., Ostend, Belgium), with statistical significance set at $p<0.05$.

RESULTS

Patient dataset

Data for the training and test datasets are listed in Table 2. Training data comprised 366 patients with 3984 images, a mean age of 64.56±12.08 years, and a mean tumor volume of 23.09 cm^3 (19.22–26.96). Test data comprised 92 patients with 1197 images, a mean age of 62.52±12.15 years, and a mean tumor volume of 28.24 cm^3 (20.88–35.59).

DL-based automated segmentation of RC

Table 3 shows the performance of the AI-based automated tumor segmentation models. The accuracy, IoU, and DSC of tumor segmentation were 98.93% (98.87–98.99), 58.30% (56.56–59.85), and 73.35% (71.97–74.71). Precision, recall, F1-score, and 95th percentile Hausdorff distance of tumor segmentation were 73.31% (71.80–74.81), 74.21% (72.11–76.32), 73.34% (71.98–74.69), and 4.46 mm (4.21–4.72). Sequential tumor segmentation following rectum segmentation showed similar performance to tumor segmentation alone; accuracy 98.82% (98.77–98.88), IoU 56.24% (54.76–57.78), DSC 71.79% (70.53–73.05), precision 69.50% (67.82–71.19), recall 75.41% (73.85–76.97), F1-score 71.98% (70.80–73.17), and 95th percentile Hausdorff distance 4.55 mm (4.23–4.88). Tumor segmentation on the images of overlaid rectum showed the best performance; accuracy 99.06% (99.00–99.11), IoU 60.89% (59.38–62.33), DSC 75.52% (74.32–76.62), precision 79.90% (78.42–81.25), recall 72.06% (70.62–73.51), F1-score 75.53% (84.44–76.62), 95th percentile Hausdorff distance 3.63 mm (3.40–3.86). We compared the mean DSC among the three segmentation models and found significant differences among them ($p=0.001$). In pairwise comparisons, the mean DSC of tumor segmentation on overlaid rectum was significantly higher than that of tumor segmentation on predicted rectum ($p<0.0001$) and higher than that of tumor segmentation without rectum guidance, with borderline statistical significance ($p=0.05$). Example cases of the three segmentation models are presented in Figs. 3-6.

Table 2. Dataset of Included Patients

	Training dataset	Test dataset
Number of patients	366	92
Number of slices	3984	1197
Sex (male/female)	242/125	58/34
Age (yr, mean±SD)	64.56±12.08	62.52±12.15
Mean tumor volume (cm^3, 95% CI)	23.09 (19.22–26.96)	28.24 (20.88–35.59)

SD, standard deviation; CI, confidence interval.

Effect of tumor volume on the performance of tumor segmentation

Tumor volume showed a positive correlation with tumor segmentation accuracy measured by DSC in all three models as follows: tumor segmentation DSC without rectum guidance [r=0.219, 95% confidence interval (CI) 0.015–0.406, $p=0.036$], tumor segmentation on images with predicted rectum (r=0.333, 95% CI 0.138–0.504, $p=0.001$), and tumor segmentation on images with overlaid rectum (r=0.319, 95% CI 0.122–0.492, $p=0.002$) (Fig. 7). In the subgroup analysis, the group with tumor volume ≤20 cm^3 showed a significant positive correlation with tumor segmentation accuracy in all three models (r=0.398–0.432, $p<0.01$), whereas the group with tumor volume >20 cm^3 did not show a significant correlation (Table 4).

DISCUSSION

We developed an AI-based automated segmentation model for RC and investigated the effect of providing anatomical guidance of the rectum on tumor segmentation performance in this

Table 3. Automated Segmentation Accuracy for Rectal Cancer

	Tumor segmentation accuracy without guidance of rectum (95% CI)	Tumor segmentation accuracy after rectum segmentation		Tumor segmentation accuracy with guidance of rectum (95% CI)
		Rectum (95% CI)	Tumor (95% CI)	
Accuracy (%)	98.93 (98.87–98.99)	98.85 (98.76–98.93)	98.82 (98.77–98.88)	99.06 (99.00–99.11)
Intersection over union (%)	58.30 (56.56–59.85)	78.17 (77.03–79.30)	56.24 (54.76–57.78)	60.89 (59.38–62.33)
Dice similarity coefficient (%)	73.35 (71.97–74.71)	85.82 (84.82–86.81)	71.79 (70.53–73.05)	75.52 (74.32–76.62)
Precision (%)	73.31 (71.80–74.81)	86.73 (85.59–87.87)	69.50 (67.82–71.19)	79.90 (78.42–81.25)
Recall (%)	74.21 (72.11–76.32)	85.76 (84.41–87.11)	75.41 (73.85–76.97)	72.06 (70.62–73.51)
F1-score (%)	73.34 (71.98–74.69)	86.05 (85.19–86.92)	71.98 (70.80–73.17)	75.53 (84.44–76.62)
95th percentile Hausdorff distance (mm)	4.46 (4.21–4.72)	4.44 (4.16–4.73)	4.55 (4.23–4.88)	3.63 (3.40–3.86)

CI, confidence interval.

https://doi.org/10.3349/ymj.2025.0133

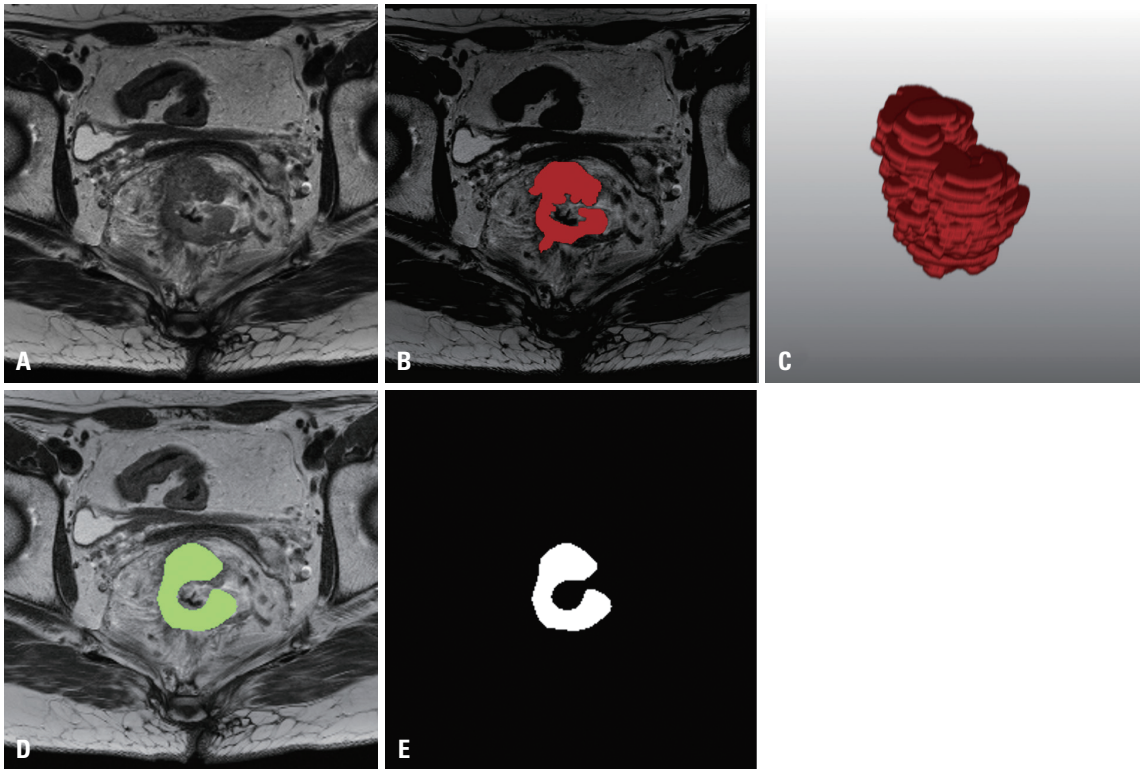

Fig. 3. A case of deep learning-based automated tumor segmentation without rectum guidance. (A) A T2-weighted axial image was selected, and tumor labeling was performed by placing a region of interest on the tumor mass in each slice, covering the entire tumor volume (B and C). (D and E) Automated tumor segmentation was performed, which predicted an area of tumor mass similar to the baseline tumor labeling.

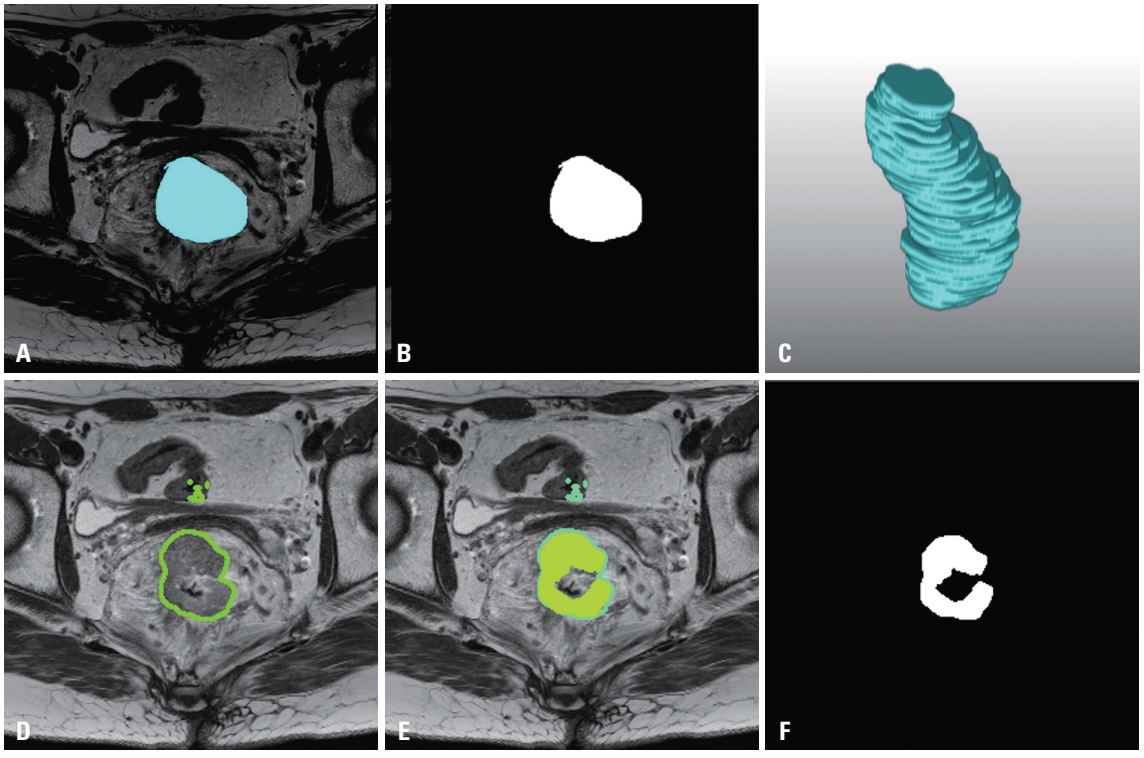

Fig. 4. A case of deep learning-based automated tumor segmentation on images with rectum segmentation. (A and B) Rectum labeling was performed on each slice of the T2-weighted axial image, (C) covering the rectum from the anorectal junction to the rectosigmoid junction. (D) Rectum segmentation was performed, which showed a slightly different shape compared with the baseline rectum labeling, but remained within the rectal outer wall. An additional erroneous prediction was noted in the adjacent pelvic ileum. (E and F) Subsequent tumor segmentation was performed using the segmented rectum.

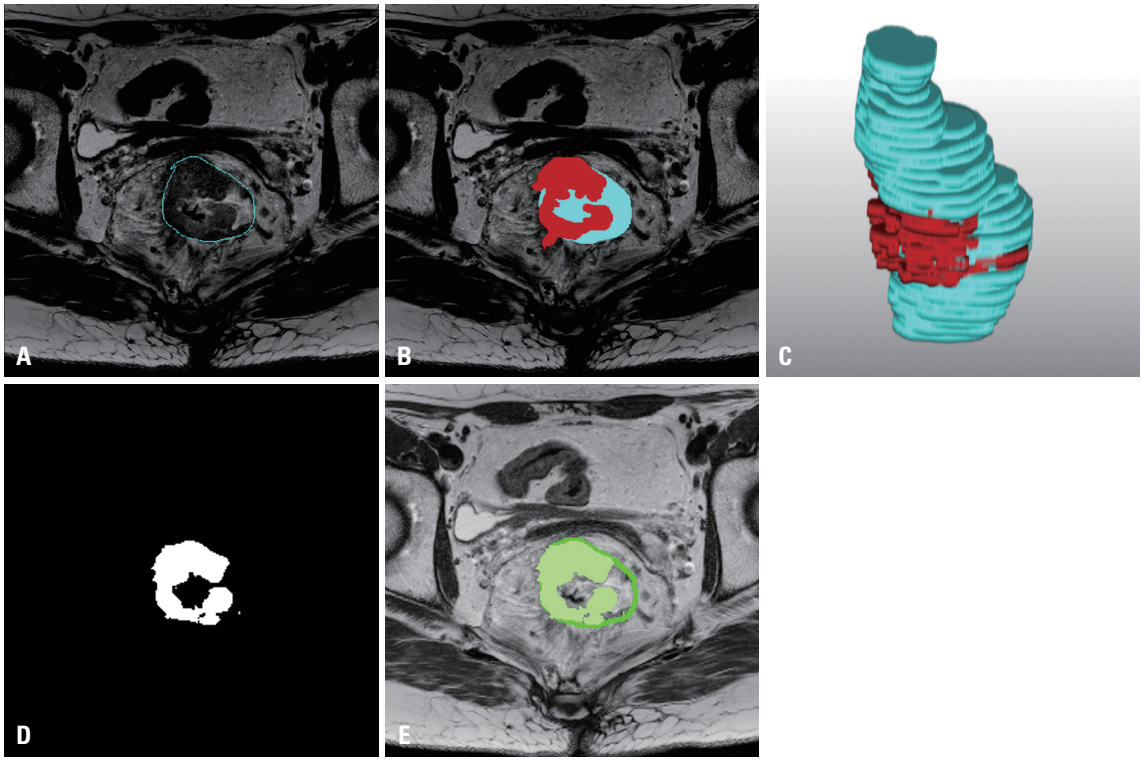

Fig. 5. A case of deep learning-based automated tumor segmentation with guidance of rectum location. (A) The outer wall margin of the rectum was overlaid on the baseline T2-weighted axial image. (B and C) Tumor labeling was performed with this guidance of rectum location, covering the entire tumor volume. (D and E) Tumor segmentation was performed on the overlaid rectum.

study. Our tumor segmentation model of RC without guidance of rectum showed a DSC of 73.35%, which is similar to that of previous studies.[4,15,16] Regarding rectum guidance for tumor segmentation, we compared two approaches: RC segmentation on images with an overlaid rectal outer margin showed better performance (DSC 75.52%), whereas segmentation on an AI-predicted rectal outer margin showed similar or suboptimal performance (DSC 71.79%) compared with RC segmentation without anatomical guidance. This may be due to imperfect localization of the rectum by the segmentation model (DSC 85.82%), although this was better than tumor segmentation (DSC 72%). Kim, et al.[17] also studied rectum and tumor segmentation using AI models and reported superior accuracy of rectum segmentation compared with tumor segmentation using U-Net (DSC 0.90 vs. 0.81), which might be due to easier recognition of the rectal outer margin than the tumor margin, which is sometimes ambiguous for labeling. However, the rectum segmentation model did not achieve the perfect localization or segmentation accuracy, as expected. Therefore, when this AI-predicted rectum segmentation was provided as anatomical guidance for tumor segmentation, model performance may have been degraded by the incomplete representation of rectum location, and it did not add any benefit to tumor segmentation.

On the other hand, the overlaid image with the rectal outer margin may improve tumor segmentation accuracy as it provides the complete rectal outer margin as anatomical guidance for tumor localization.

For tumor segmentation on images with overlaid rectum, the labeled rectal outer margin (ground truth) was superimposed on the T2W MR image and provided as an input image during tumor segmentation training. The labeled rectal outer margin could provide guidance on the location of the rectum as well as a focal point for AI learning. However, it could also limit the possible tumor area to within the rectal outer margin, and tumor portions extending beyond the rectal outer margin may be missed in cases of a large tumor mass.

Several studies reported RC segmentation models using T2W images,[5,17] DWI,[16] or multiparametric MR with combined T2W and DWI,[18] and they reported favorable DSC values of 0.81–0.91. We used only T2W images for AI model training as DWI may not help detect the precise tumor margin owing to their poor spatial resolution, and T2W images are the essential sequence for basic interpretation of RC. In most studies, U-Net has been used as the AI learning algorithm, and Kim, et al.[17] reported the best performance of tumor and rectum segmentation using U-Net compared with other convolutional neural networks such as FCN-8S or SegNet. We used attention U-Net, which adds an attention gate to the traditional U-Net, allowing the network to focus on the target to be learned and detected. While previous articles studied changes in model performance according to different convolutional neural network algo-

https://doi.org/10.3349/ymj.2025.0133

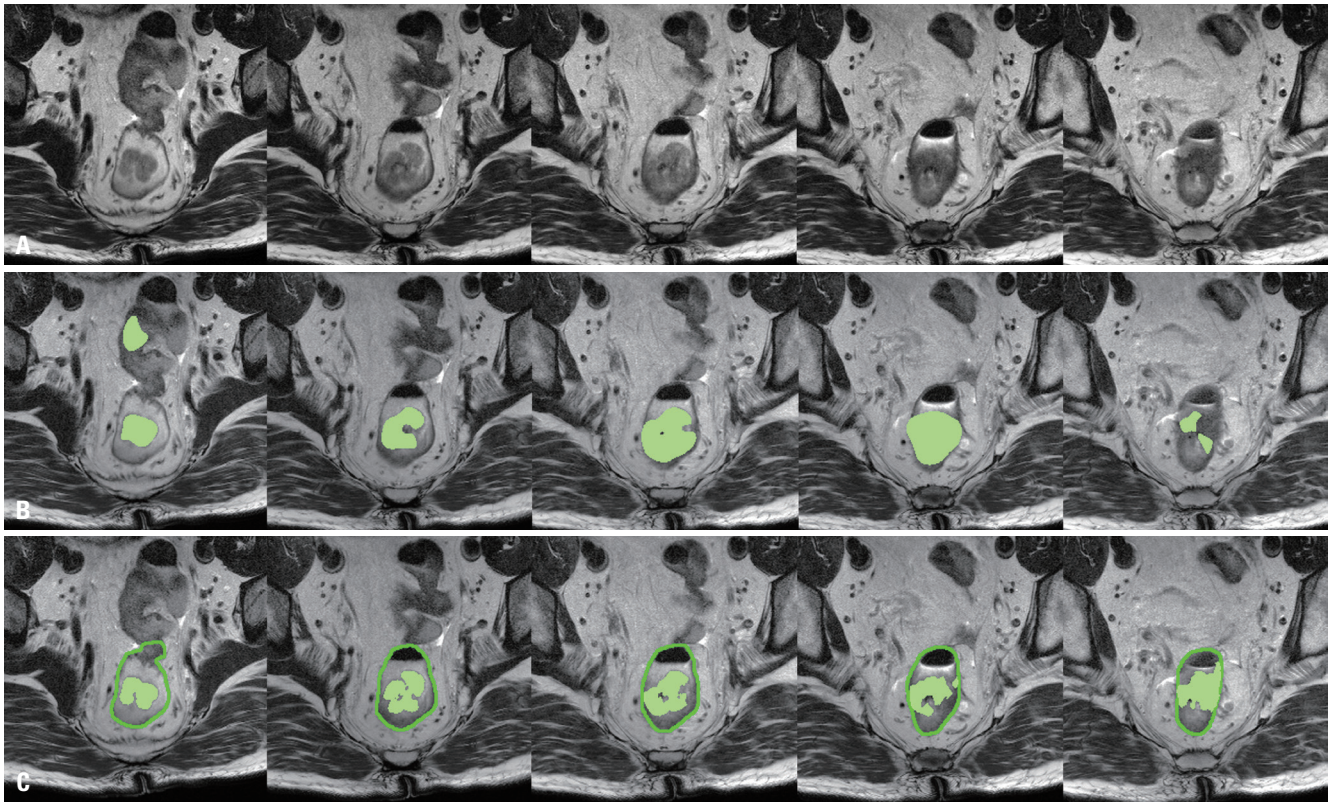

Fig. 6. An example showing improved performance of rectal cancer segmentation with rectum guidance (from left to right, images correspond to the upper to lower portions). (A) The baseline T2-weighted MR images show a polypoid cancerous mass in the rectum. (B) The tumor segmentation model marked the tumor mass in each image without rectum guidance. Errors include erroneous segmentation of the tumor outside the rectum (first image on the left), inaccurate segmentation of the tumor's shape (second and fourth images), and incomplete segmentation of the tumor lesion (last image on the right). (C) Segmenting the tumor on an overlaid rectum more accurately identifies the tumor mass than direct tumor segmentation does. Erroneous segmentation of the tumor outside the rectum is eliminated.

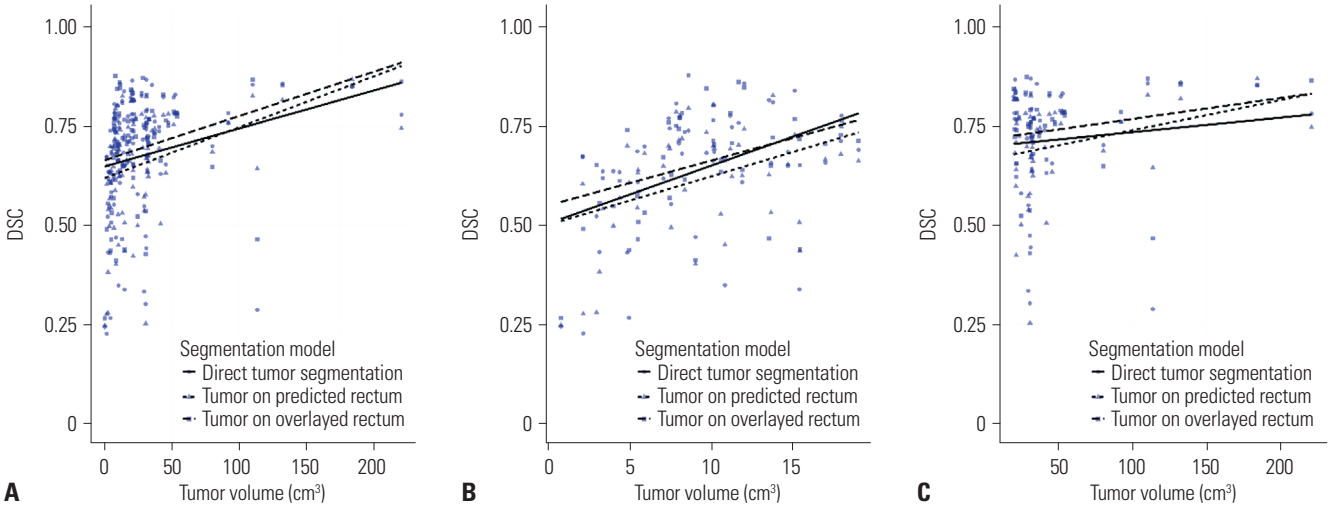

Fig. 7. Correlation between tumor volume and tumor segmentation accuracy measured by the DSC. (A) Tumor volume was positively correlated with the DSC of direct tumor segmentation (solid line, r=0.219, 95% CI 0.015–0.406, p=0.036), tumor segmentation on the predicted rectum (thin dotted line, r=0.333, 95% CI 0.138–0.504, p=0.001), and tumor segmentation on the overlaid rectum (bold dotted line, r=0.319, 95% CI 0.122–0.492, p=0.002). (B) In the subgroup analysis, the smaller tumor volume group (≤20 cm^3) showed a positive correlation with DSC in all three models: direct tumor segmentation (solid line, r=0.432, 95% CI 0.166–0.640, p=0.002), tumor on predicted rectum (thin dotted line, r=0.406, 95% CI 0.134–0.621, p=0.005), and tumor on overlaid rectum (bold dotted line, r=0.398, 95% CI 0.125–0.615, p=0.005). (C) The larger tumor volume group (>20 cm^3) showed no significant correlation with DSC in all three models. DSC, dice similarity coefficient; CI, confidence interval.

Table 4. Correlation between Tumor Volume and Tumor Segmentation Accuracy Measured by the Dice Similarity Coefficient

	Tumor volume (n=92) (95% CI)	*p*	Tumor volume ≤20 cm^3 (n=47) (95% CI)	*p*	Tumor volume >20 cm^3 (n=45) (95% CI)	*p*
Tumor segmentation without rectum guidance (r)	0.219 (0.015–0.406)	0.036*	0.432 (0.166–0.640)	0.002*	0.106 (-0.193–0.387)	0.488
Tumor segmentation on predicted rectum (r)	0.333 (0.138–0.504)	0.001*	0.406 (0.134–0.621)	0.005*	0.266 (-0.030–0.519)	0.077
Tumor segmentation on overlaid rectum (r)	0.319 (0.122–0.492)	0.002*	0.398 (0.125–0.615)	0.005*	0.222 (-0.076–0.484)	0.142

CI, confidence interval.
*$p<0.05$.

rithms such as U-Net,[5,16,17] attention-based fusion network,[18] or improved U-Net,[10] we studied changes in performance according to input factors such as rectal outer wall guidance and tumor volume.

The former—anatomical guidance of the rectal outer wall—could improve RC segmentation performance, but it depends on the performance of rectum segmentation. The latter—tumor volume—showed a positive correlation with tumor segmentation accuracy; that is, as tumor volume increased, segmentation accuracy also increased. However, in the subgroup analysis, only the smaller tumor group (≤20 cm^3) showed a positive correlation with tumor segmentation accuracy, whereas the larger tumor group did not. This may be due to an increased proportion of erroneous segmentation in smaller tumors. However, at a certain tumor volume, segmentation accuracy can remain stable regardless of tumor volume. Therefore, an automated tumor segmentation model would be particularly useful for larger tumors, as manual labeling requires more time and effort.

This study has several limitations. First, the dataset was collected retrospectively, which introduced inherent selection bias and heterogeneous MR protocols across the dataset. We consistenyl ensured that the baseline input MR image was a T2W axial scan before labeling; if there was no appropriate T2W axial scan, T2W true short- or long-axis images, obtained perpendicular to the short or long axis of the tumor mass, were selected. We attempted to standardize the background MR image to an axial scan across the dataset, regardless of tumor shape or axis, as this could reduce confusion from surrounding pelvic organs during DL. Second, the included cases presented tumor masses with different shapes and axes, such as long-axis tumors located in the rectosigmoid junction or distal sigmoid, short-axis tumors located in the anorectum, or polypoid-shaped tumors. We performed a subgroup analysis for the majority of short-axis tumors, excluding other long-axis and polypoid tumors (not shown here); however, the segmentation accuracy was even lower than that of the overall dataset. Third, we labeled the rectum along the presumed outer margin across the continuous z-axis when the tumor encircled the entire rectal wall with minimal visible spared normal rectal tissue. Therefore, the tumor portion beyond the rectal wall could be ignored during the DL process. Fourth, segmentation accuracy measured by the DSC has an inherent bias toward larger tumors, as demonstrated in our correlation analysis results. We performed a subgroup analysis stratified by tumor volume to investigate the effect of tumor volume on model performance. Fifth, converting the original three-dimensional DICOM images into two-dimensional PNG images may have caused data loss. However, model performance may not have been substantially affected because our model was developed using axial images, and a three-dimensional reconstruction of the tumor was not performed. Sixth, external validation was not performed; however, results were reported from a test dataset that was not used during training.

In conclusion, our RC segmentation model demonstrated a DSC of 72%–73%, and both larger tumor volume and the use of an overlaid rectal outer margin improved tumor segmentation accuracy.

ACKNOWLEDGEMENTS

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (Ministry of Science and ICT, NRF-2022R1C1C1012477).

AUTHOR CONTRIBUTIONS

Conceptualization: Myung-Won You and Miri Seo. Data curation: YongDae Lee. Formal analysis: Miri Seo. Funding acquisition: Myung-Won You. Investigation: all authors. Methodology: Myung-Won You and Miri Seo. Project administration: Myung-Won You. Resources: YongDae Lee and Myung-Won You. Software: Miri Seo and YongDae Lee. Supervision: Myung-Won You. Validation: Myung-Won You and Miri Seo. Visualization: Miri Seo. Writing—original draft: Myung-Won You and Miri Seo. Writing—review & editing: Myung-Won You and Miri Seo. Approval of final manuscript: all authors.

ORCID iDs

Miri Seo https://orcid.org/0000-0003-2895-6787
YongDae Lee https://orcid.org/0009-0002-1276-9522
Myung-Won You https://orcid.org/0000-0001-6262-5784

REFERENCES

1. Horvat N, Carlos Tavares Rocha C, Clemente Oliveira B, Petkovska

https://doi.org/10.3349/ymj.2025.0133

I, Gollub MJ. MRI of rectal cancer: tumor staging, imaging techniques, and management. Radiographics 2019;39:367-87.
2. Wang D, Xu J, Zhang Z, Li S, Zhang X, Zhou Y, et al. Evaluation of rectal cancer circumferential resection margin using faster region-based convolutional neural network in high-resolution magnetic resonance images. Dis Colon Rectum 2020;63:143-51.
3. Zhao X, Xie P, Wang M, Li W, Pickhardt PJ, Xia W, et al. Deep learning-based fully automated detection and segmentation of lymph nodes on multiparametric-MRI for rectal cancer: a multicentre study. EBioMedicine 2020;56:102780.
4. Trebeschi S, van Griethuysen JJM, Lambregts DMJ, Lahaye MJ, Parmar C, Bakers FCH, et al. Deep learning for fully-automated localization and segmentation of rectal cancer on multiparametric MR. Sci Rep 2017;7:5301.
5. Wang J, Lu J, Qin G, Shen L, Sun Y, Ying H, et al. Technical note: a deep learning-based autosegmentation of rectal tumors in MR images. Med Phys 2018;45:2560-4.
6. Ferrari R, Mancini-Terracciano C, Voena C, Rengo M, Zerunian M, Ciardiello A, et al. MR-based artificial intelligence model to assess response to therapy in locally advanced rectal cancer. Eur J Radiol 2019;118:1-9.
7. Fu J, Zhong X, Li N, Van Dams R, Lewis J, Sung K, et al. Deep learning-based radiomic features for improving neoadjuvant chemoradiation response prediction in locally advanced rectal cancer. Phys Med Biol 2020;65:075001.
8. Cui Y, Liu H, Ren J, Du X, Xin L, Li D, et al. Development and validation of a MRI-based radiomics signature for prediction of KRAS mutation in rectal cancer. Eur Radiol 2020;30:1948-58.
9. Schlemper J, Oktay O, Schaap M, Heinrich M, Kainz B, Glocker B, et al. Attention gated networks: learning to leverage salient regions in medical images. Med Image Anal 2019;53:197-207.
10. Zhang K, Yang X, Cui Y, Zhao J, Li D. Imaging segmentation mechanism for rectal tumors using improved U-Net. BMC Med Imaging 2024;24:95.
11. Xia S, Li Q, Zhu HT, Zhang XY, Shi YJ, Yang D, et al. Fully semantic segmentation for rectal cancer based on post-nCRT MRl modality and deep learning framework. BMC Cancer 2024;24:315.
12. Yang M, Yang M, Yang L, Wang Z, Ye P, Chen C, et al. Deep learning for MRI lesion segmentation in rectal cancer. Front Med (Lausanne) 2024;11:1394262.
13. Rezatofighi H, Tsoi N, Gwak J, Sadeghian A, Reid I, Savarese S. Generalized intersection over union: a metric and a loss for bounding box regression. arXiv [Preprint]. 2019 [accessed on 2025 April 30]. Available at: https://doi.org/10.48550/arXiv.1902.09630.
14. Crum WR, Camara O, Hill DL. Generalized overlap measures for evaluation and validation in medical image analysis. IEEE Trans Med Imaging 2006;25:1451-61.
15. Lee J, Oh JE, Kim MJ, Hur BY, Sohn DK. Reducing the model variance of a rectal cancer segmentation network. IEEE Access 2019;7:182725-33.
16. Zhu HT, Zhang XY, Shi YJ, Li XT, Sun YS. Automatic segmentation of rectal tumor on diffusion-weighted images by deep learning with U-Net. J Appl Clin Med Phys 2021;22:324-31.
17. Kim J, Oh JE, Lee J, Kim MJ, Hur BY, Sohn DK, et al. Rectal cancer: toward fully automatic discrimination of T2 and T3 rectal cancers using deep convolutional neural network. Int J Imag Syst Technol 2019;29:247-59.
18. Dou M, Chen Z, Tang Y, Sheng L, Zhou J, Wang X, et al. Segmentation of rectal tumor from multi-parametric MRI images using an attention-based fusion network. Med Biol Eng Comput 2023;61:2379-89.