



A Pipeline for Predicting the Treatment Response of Neoadjuvant Chemoradiotherapy for Locally Advanced Rectal Cancer Using Single MRI Modality: Combining Deep Segmentation Network and Radiomics Analysis Based on “Suspicious Region”

OPEN ACCESS

Edited by:

Niccolo Petrucciari,
Sapienza University of Rome, Italy

Reviewed by:

Marc Gollub,
Memorial Sloan Kettering Cancer
Center, United States
Xueliang Wu,
First Affiliated Hospital of Hebei North
University, China
Luciano Izzo,
Sapienza University of Rome, Italy

*Correspondence:

Xinjuan Fan
fanxjuan@mail.sysu.edu.cn

†These authors have contributed
equally to this work and
share first authorship

Specialty section:

This article was submitted to
Gastrointestinal Cancers,
a section of the journal
Frontiers in Oncology

Received: 19 May 2021

Accepted: 06 July 2021

Published: 04 August 2021

Citation:

Pang X, Wang F, Zhang Q,
Li Y, Huang R, Yin X and Fan X
(2021) A Pipeline for Predicting the
Treatment Response of Neoadjuvant
Chemoradiotherapy for Locally
Advanced Rectal Cancer Using
Single MRI Modality: Combining
Deep Segmentation Network
and Radiomics Analysis Based
on “Suspicious Region”.
Front. Oncol. 11:711747.
doi: 10.3389/fonc.2021.711747

Xiaolin Pang^{1,2†}, Fang Wang^{2†}, Qianru Zhang^{3†}, Yan Li^{3†}, Ruiyan Huang^{2,4}, Xinke Yin²
and Xinjuan Fan^{2,4*}

¹ Department of Radiation Oncology, The Sixth Affiliated Hospital of Sun Yat-sen University, Guangzhou, China, ² Guangdong Institute of Gastroenterology, Guangdong Provincial Key Laboratory of Colorectal and Pelvic Floor Diseases, Supported by National Key Clinical Discipline, Guangzhou, China, ³ School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou, China, ⁴ Department of Pathology, The Sixth Affiliated Hospital of Sun Yat-sen University, Guangzhou, China

Patients with locally advanced rectal cancer (LARC) who achieve a pathologic complete response (pCR) after neoadjuvant chemoradiotherapy (nCRT) typically have a good prognosis. An early and accurate prediction of the treatment response, i.e., whether a patient achieves pCR, could significantly help doctors make tailored plans for LARC patients. This study proposes a pipeline of pCR prediction using a combination of deep learning and radiomics analysis. Taking into consideration missing pre-nCRT magnetic resonance imaging (MRI), as well as aiming to improve the efficiency for clinical application, the pipeline only included a post-nCRT T2-weighted (T2-w) MRI. Unlike other studies that attempted to carefully find the region of interest (ROI) using a pre-nCRT MRI as a reference, we placed the ROI on a “suspicious region”, which is a continuous area that has a high possibility to contain a tumor or fibrosis as assessed by radiologists. A deep segmentation network, termed the two-stage rectum-aware U-Net (tsraU-Net), is designed to segment the ROI to substitute for a time-consuming manual delineation. This is followed by a radiomics analysis model based on the ROI to extract the hidden information and predict the pCR status. The data from a total of 275 patients were collected from two hospitals and partitioned into four datasets: Seg-T (N = 88) for training the tsraUNet, Rad-T (N = 107) for building the radiomics model, In-V (N = 46) for internal validation, and Ex-V (N = 34) for external validation. The proposed method achieved an area under the curve (AUC) of 0.829 (95% confidence interval [CI]: 0.821, 0.837) on In-V and 0.815 (95% CI, 0.801, 0.830) on Ex-V. The performance of the method was

considerable and stable in two validation sets, indicating that the well-designed pipeline has the potential to be used in real clinical procedures.

Keywords: LARC, nCRT, MRI, radiomics analysis, deep learning

INTRODUCTION

Colorectal cancer is currently still the third most common cancer and the second most fatal cancer in the world (1). Nearly 30% sufferers are rectal cancer patients (2), great numbers of which are in the locally advanced stage at initial diagnosis (3).

To date, for patients with locally advanced rectal cancer (LARC), neoadjuvant chemoradiotherapy (nCRT) followed by total mesorectal excision (TME) has been the standard clinical treatment (4–6). The purpose of nCRT is to improve the feasibility of surgical procedures for LARC and reduce the incidence of complications, as it not only improves the local tumor control rate but also exhibits less toxicity to the human body (7). Clinically, the pathological response of LARC patients after nCRT treatment has demonstrated obvious heterogeneity (8). For a large percentage of patients (approximately 70–80%), the tumor will have been found to be shrunken or down-staged, and some patients may even have complete regression. It has been reported that approximately 20% of patients, defined as pathologic complete response (pCR) patients, contain no residual surviving tumor cells after nCRT and surgery (9, 10). These patients have a favorable long-term prognosis with superb local control and disease-free survival (11). Therefore, for pCR patients, the option of organ-saving treatment could be developed to replace surgery. However, currently, the only way to accurately diagnose pCR is to utilize a pathological diagnosis after TME surgery, which presents an insoluble dilemma (12–14). As a result, a prediction method before surgery would greatly assist doctors in evaluating the treatment effects of nCRT and construct a tailored plan for each patient.

In recent years, magnetic resonance imaging (MRI) has been widely used as a non-genetic and non-invasive diagnostic method to assess the tumor condition due to its superior soft-tissue contrast and high spatial resolution. The T2-weighted (T2-w) MRI is recognized as the most important modality for rectal cancer assessments (15). Some previous works have attempted to evaluate the response of nCRT on T2-w MRI using assessments of post-treatment T staging (ymrT), tumor regression grading (mrTRG), volume reduction post-treatment, and other characteristics (16). However, the results were dependent on the experience of doctors, thus introducing subjectivity. The development of more objective ways to extract information from the MRI and guide clinical diagnosis is required.

Radiomics is a mathematical technique that utilizes high-throughput extraction of shape, intensity, and texture features from images, and transforms this visual information into high dimensional features for quantitative analysis (17). Radiomics analysis can help obtain additional image information with reliability and objectivity that may be invisible in human assessments (18). Applying a radiomics analysis to predict the

pCR on an MRI has drawn increasing attention. In many studies (19–21), researchers have used a pre-nCRT MRI to analyze the relationship of the radiomics features and the pCR status. The use of a pre-nCRT MRI could provide a clear tumor region for analysis. However, considering that the nCRT could affect the tumor and change its properties, the use of a pre-nCRT MRI is indirect and may not reflect the true condition of the patient after nCRT treatment.

Analysis of a post-nCRT MRI might be a more direct method. However, the problem still remains that the region of interest (ROI) delineation on a post-nCRT is much more difficult due to tumor recession and the appearance of the fibrosis region. In previous studies (22, 23), researchers applied a pre-nCRT MRI to provide a reference of the primary tumor region or the treated region for a post-nCRT MRI. However, a pre-nCRT MRI is not always available in real clinical practice, as some patients may be diagnosed with LARC using proctoscopy, and some patients may be transferred from other hospitals without access to the two previously scanned MRIs. Currently, only a few of studies have utilized a post-nCRT MRI to predict the pCR status. The work of Horvat et al. (15) obtained considerable results by applying a radiomics analysis on a post-nCRT T2 and diffusion-weighted imaging (DWI) MRI; however, the ROI was still obtained due to a careful discussion by at least two experienced physicians. An accurate delineation might be difficult to obtain for less experienced physicians without the reference of a pre-nCRT MRI. Additionally, it is time-consuming and resource-wasting if each delineation requires at least two physicians in clinical practice. Inspired by this, the aim of this study is to explore a pipeline that only uses the information from a single post-nCRT T2 MRI combined with a new method to provide a fast and reliable ROI. Deep learning uses multiple layers as a portion of a broader family of machine learning methods and has been successfully applied to various medical tasks (24–28).

In this study, we introduce a deep learning model for ROI delineation. A novel two stage model, termed the two-stage rectum-aware U-Net (tsraU-Net), is proposed to replace human evaluation. The ROI should be feasible for a deep learning model to find and contain sufficient information relating to the pCR status; hence, it is defined on a continuous region having abnormal intensity signals on a T2-w MRI assessed by radiologists that has a high possibility to contain a tumor or fibrosis. It is considered a rougher region, as the further identification of a tumor, fibrosis, or other tissues like edema is not defined. The following analysis extracts a great number of radiomics features, including texture, first-order statistics, and shape, on the ROI and its wavelet decompositions to represent certain properties. Machine learning and statistical techniques are later applied to select the most representative features and construct a final model to predict the pCR status.

MATERIALS AND METHODS

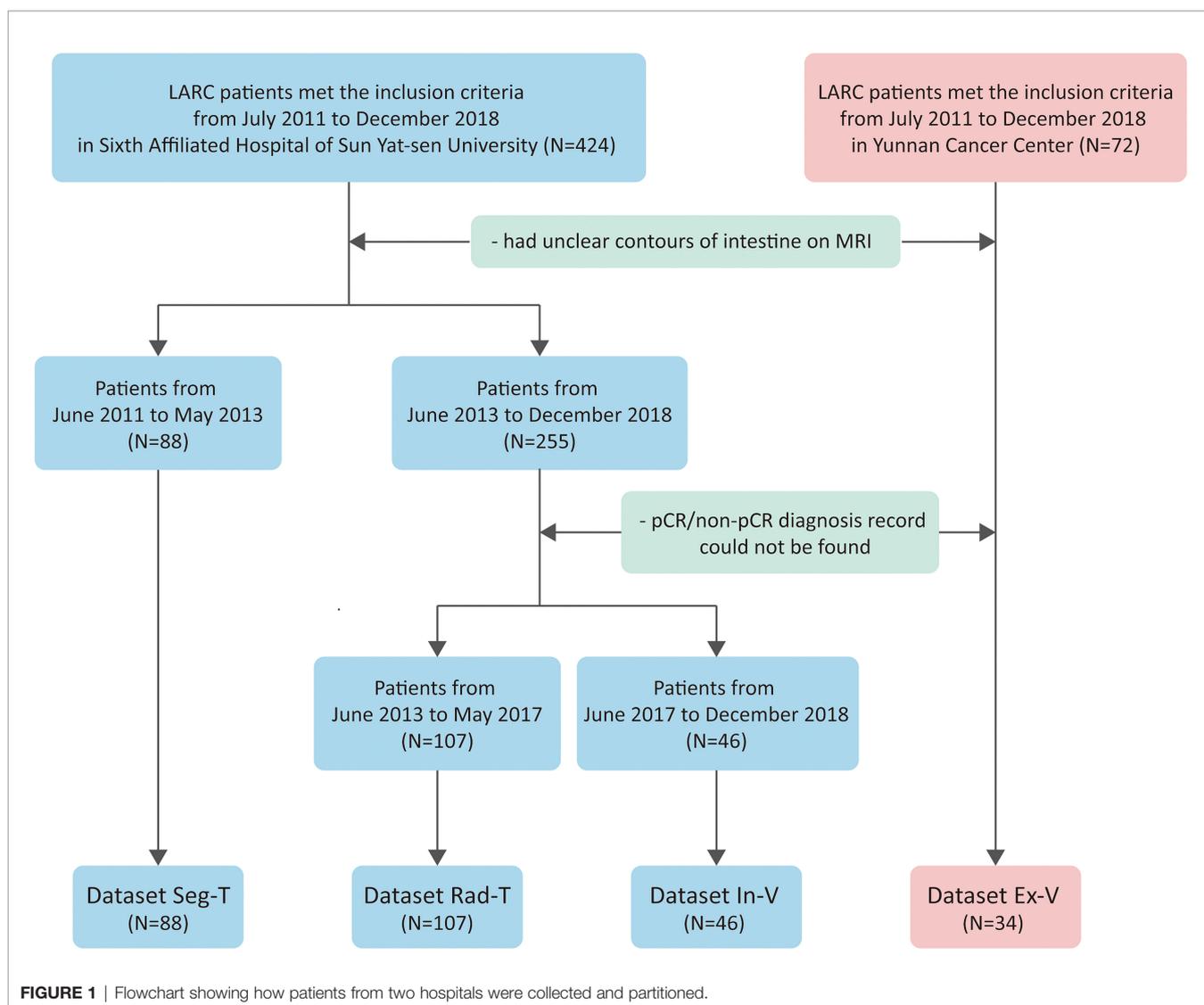
Patients

We included a total of 496 patients from multiple institutions who received nCRT treatment diagnosed with LARC. Patients were retrospectively enrolled from July 2011 to December 2018 from two hospitals (Guangdong Institute of Gastroenterology, Sixth Affiliated Hospital of Sun Yat-sen University and Yunnan Cancer Center). The inclusion criteria were as follows: (1) adenocarcinoma confirmed by pathologists (excluding mucinous adenocarcinoma); (2) tumor located within 15 cm from the edge of the anus; (3) received nCRT treatment and TME; (4) in the clinical stage of T3–4 or N-positive; (5) completed the restaging MRI; and (6) the restaging MRI was performed no more than 1 week before the TME. This clinical trial was approved by the Clinical Ethics Review Committee of Sixth Affiliated Hospital of Sun Yat-sen University (2020ZSLYEC-010).

Patients from the Sixth Affiliated Hospital of Sun Yat-sen University were separated into three groups. Dataset Seg-T

consisted of patients from July 2011 to June 2013 to train deep networks for ROI segmentation; dataset Rad-T consisted of patients from June 2013 to May 2017 to build the radiomics model for predicting the pCR status; and dataset In-V consisted of patients from June 2017 to December 2018 for internal validation. In addition, patients from Yunnan Cancer Center (dataset Ex-V) were used as an external validation set. Patients were further selected according to the following exclusion criteria: (1) poor MRI quality caused by severe inflammatory effusion, intestinal adhesions, or bowel movements, and (2) (for dataset Rad-T, Val, E-Val) the absence of a postoperative pathological diagnosis. **Figure 1** shows the flowchart of the patient selection process.

All of the patients' treatments were discussed by the multidisciplinary team (MDT). Patients were delivered the intensity-modulated radiotherapy treatment (IMRT), and a dose of 45 Gy for 25 fractions was delivered to the clinical target volume. Then, a boost dose of 5.4 Gy was delivered to the



gross tumor. The concurrent chemotherapy treatment was based on oral or intravenous 5-fluorouracil. Following the completion of neoadjuvant treatment, all of the patients received TME surgery. The majority of patients received adjuvant chemotherapy based on FOLFOX or CAPOX based on the decision of the members of the MDT.

Pathology Assessment of Response

In this study, the pCR diagnosis was confirmed by two pathologists with more than 12 years of experience. Following the recommendation of the NCCN Guidelines for rectal cancer (29), patients with no surviving tumor cells in the surgical pathological specimens were judged as pCR; otherwise, they were judged as non-pCR.

MRI Data Acquisition

All of the MRI images were scanned under a 1.5-Tesla MRI unit (30). Bowel preparation was not routinely used for most cases prior to the examination. However, some specific patients with relatively small tumors in the sagittal view were filled with some rectal gel, making it easier to identify tumors on the oblique axis. The MRIs in the datasets Seg-T, Rad-T, and In-V (Sixth Affiliated Hospital of Sun Yat-sen University) were acquired using GE OPTIMA MR360 with a 100 ms echo time, a 4000 ms repetition time, a 100 field of view, a 512×512 matrix, 0.4–0.5 pixel spacing, and 5 mm slice thickness. The MRIs in the dataset Ex-V (Yunnan Cancer Hospital) were acquired using Philips Ingenia with a 100 ms echo time, a 4000 ms repetition time, a 100 field of view, a 432×432 matrix, 0.4–0.5 pixel spacing, and 5 mm slice thickness.

Data Pre-Processing

As suggested by some researchers (31), we applied complex methods to the pre-process MRI to both improve the image quality and unify the geometric and intensity patterns with the aim to assure the success of our analysis. The steps included (1) all of the MRIs were resampled into $0.4 \text{ mm} \times 0.4 \text{ mm}$ pixel spacing using bilinear interpolation; (2) the size of each image matrix was unified into 544×544 by cutting or padding the background; (3) the intensity of each patient was adjusted using BiasCorrection to remove any inhomogeneity; and (4) the intensity histogram of each patient was matched to one selected patient (as template) who was from Seg-T. All of the procedures were implemented using the open-source python package “SimpleITK” (32).

Suspicious Region Definition

The “suspicious region” in our study was defined as a continuous region containing 129 abnormal intensity signals compared to a normal rectal wall, which are highly suspected to be cancer or fibrosis according to clinical experience. Following the guidance (33), the abnormal signals may have presented as slightly high, low, or mixed intensities. By such definition, the exact cancer and fibrosis region was not further distinguished. Instead, we relied on the radiomics analysis to elicit the hidden properties of the cancer or fibrosis and predict the pCR. In particular, the “suspicious region” was delineated by radiologists on only a

post-nCRT T2-w MRI. As the region was visible to human vision, we assumed it could be captured by deep learning as well. Therefore, the use of “suspicious region” was both sufficient and proper in the pipeline that combined a radiomics analysis and deep learning. **Figure 2** provides some examples of the suspicious region.

Deep Learning-Based Segmentation

The tsraU-Net Model

To provide a reliable ROI using deep learning for the radiomics analysis, the most important consideration was rectum localization. If the deep segmentation network misrecognized the rectum with other organs such as the colon, uterus, bladder, or prostate due to a morphology change or location shift of the rectum, it could still find a “suspicious region,” but not related to the pCR at all, thereby making the radiomics analysis totally meaningless. To address such a problem, we proposed a two-stage model, named the two-stage rectum-aware U-Net (tsraU-Net), which would first find the rectum region and then segment the ROI using the awareness of the rectum location. The overall framework of the tsraU-Net model is shown in **Figure 3**.

U-Net, a deep segmentation network expressly designed for biomedical segmentation tasks (24), was applied as a base model in both of the two stages of the tsraU-Net. Further improvements were made in each stage according to the task. The detailed descriptions were organized as follows. First, we briefly introduced the base model, U-Net. Next, we provided comprehensive explanations of the improvements in the two stages of tsraU-Net. Finally, we described other adjustments of the base model.

The Original U-Net Model

The original U-Net is a fully convolution network (FCN) containing an encoder to extract features and a decoder to reassemble features. Typically, both of them have five convolution blocks, s.t. each block consists of two 3×3 convolutions and a rectified linear unit (ReLU) for activation. After going through the convolution block, the number of features (more specifically, channels) would double in the encoder part and halve in the decoder part symmetrically. Between each convolution block in the encoder, the max pooling operation is applied to reduce the image resolution. Oppositely, an up-sample operation is inserted into the neighboring convolution blocks in the decoder to increase the image resolution. To fully utilize high resolution information, high resolution features in the encoder are concatenated to the corresponding convolution blocks in the decoder. This is named a “skip connection.” In addition, a 3×3 convolution is applied after the last convolution block in the decoder to combine the rest of the features and obtain the final segmentation result.

The First Stage

In the first stage, we designed a four-channel 2D U-Net aiming to guide the segmentation with a plentiful amount of information. Knowing that rectal regions would maintain certain continuity between neighboring MRI slices, the input of our model contains not only the currently input MRI slice but also its previous slice and the next slice to help detect the contours of the rectum. If the rectal wall is unclear compared with the neighboring region on

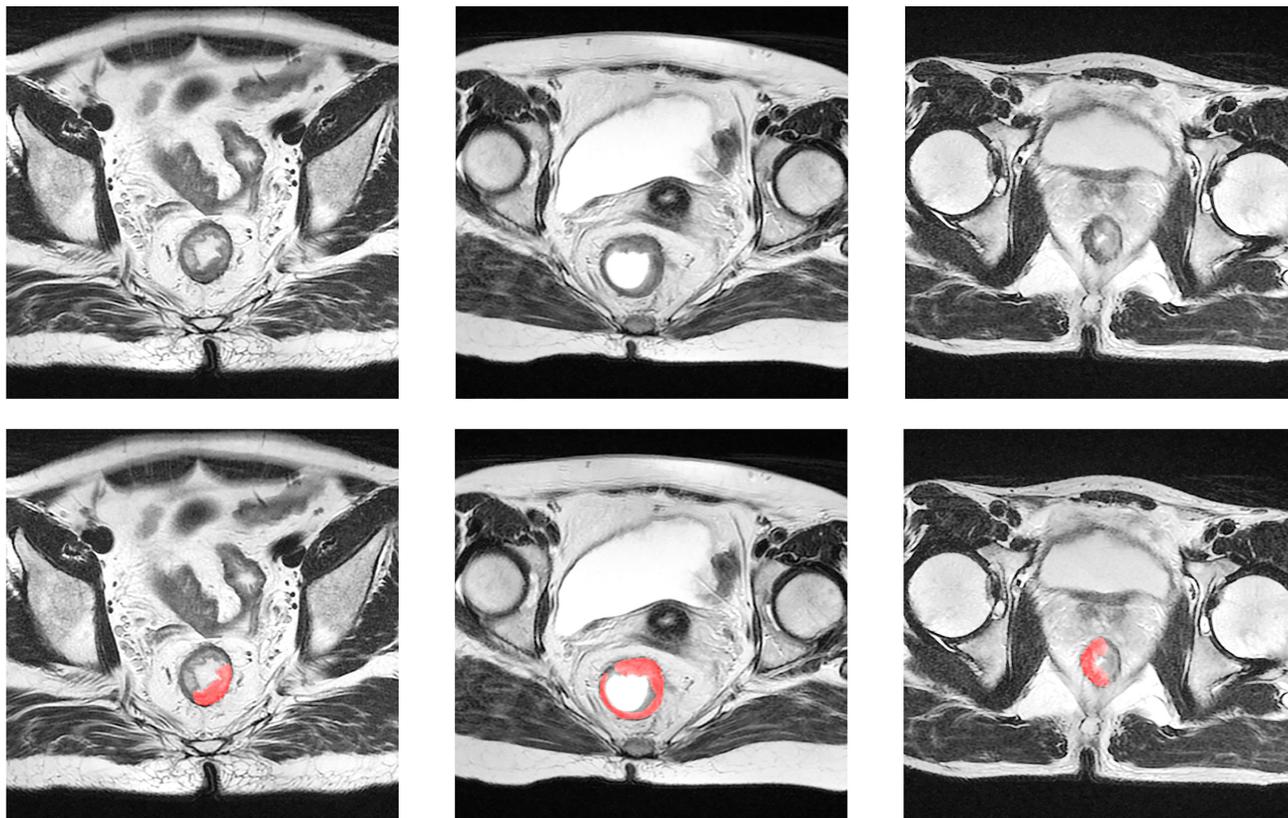


FIGURE 2 | Examples of suspicious region. The first row is the original T2-w MRI, and the second row displays the delineation of the suspicious region in red color.

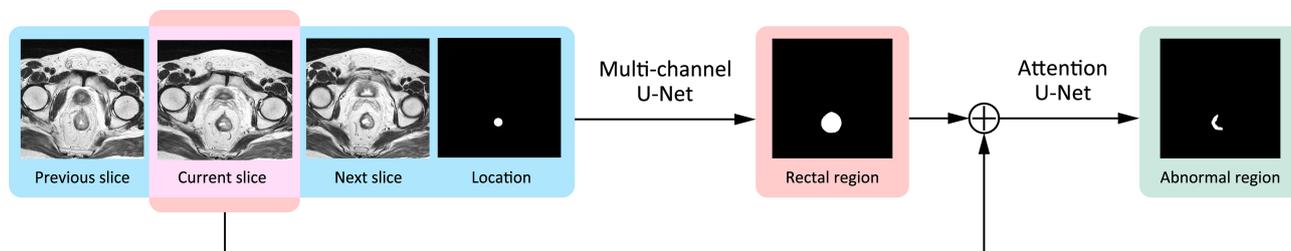


FIGURE 3 | The framework of tsraU-Net model.

the current slice, the other two slices may provide extra information. Furthermore, for each patient, we roughly marked two to four points inside the rectum for localization. Initially, the first and last MRI slices were given localization points, with bilinear interpolation applied between them to give localization information to the middle layers. If the shape of the rectum was not regular, another one to two points will be given in the middle layers. The previous slice, the current slice, the next slice, and the position information were combined into four channels as the input of our first stage model. The output of the first stage was the region of the rectum.

The Second Stage

The second stage would use the currently inputted slice and the predicted rectal region in the first stage to find the “suspicious region.” In this stage, we focused on strengthening the model awareness of the abnormal intensity signals and, hence, applied an “attention” mechanism. “Attention” was first introduced in the natural language process (NLP) tasks to encourage models to pay more attention to efficacious information and suppress irrelevant information. There are two types of attention, i.e., soft attention and hard attention. In this task, we used the soft attention mechanism for the model (34). This method would

update the propagated features from “skip connection” by point-wise multiplication with a weight matrix given by an attention gate. An attention gate is a block containing several 1×1 convolutions and activation functions. It uses both the propagated features and the features from the corresponding former decoder block as input. **Figure 4** provides a detailed explanation of an attention gate. Specifically, we inserted four attention gates into the original U-Net model s.t. every “skip connection” was followed by an attention gate.

Adjustments on Base Model

In addition to the above methods, we also made some small adjustments to the original U-Net model: (1) the addition of image padding during convolution so that the image size would not change; (2) the addition of instance normalization after each convolution block to accelerate the convergence; and (3) the replacement of the ReLU activation function with the Leaky Rectified Linear Unit (Leaky ReLU) to prevent the vanishing gradient problem.

Loss Function

The loss function in both two stages is the Dice Loss, which is widely used in medical image segmentation. It is defined from the dice coefficient, which essentially measures the overlap of two sets. The dice coefficient has a range of 0–1, where 1 means complete overlap. It is defined as Equation 1:

$$\text{Dice} = \frac{2|A \cap B|}{|A| + |B|} \quad (1)$$

where $A \cap B$ is the intersection of sets A and B, $||$ represents the number of elements in the set.

As for Dice Loss, it is simply defined by the following Equation 2.

$$\text{Dice Loss} = 1 - \text{Dice} \quad (2)$$

Experiment Setup

To provide the gold standard of segmentation, two radiologists (one had 6 years of experience and one had 9 years of experience) reviewed the post-nCRT T2-w MRI of the Dataset Seg-T and InV (for validating the performance), and they jointly provide delineations of the rectum and the “suspicious region.” When

delineating the “suspicious region,” the normal rectal wall should be avoided, and for some patients with rectal gel filling, the gel should also be avoided. Both radiologists were completely blinded to the histopathology information, as well as the pre-nCRT MRI of patients. Following the instructions in the guidance (33), they used the rectal wall as a reference to find abnormal signals. This work was performed *via* ITK-snap version 3.4.0 software (<http://itk-snap.org>).

While training, the initial hyper-parameters were established identically in two stages. The optimizer we used was the Adaptive Moment Estimation (Adam) (35) with an initial learning rate $\alpha = 2 \times 10^{-4}$, $\beta = (0.9, 0.999)$ and would decay 30% every 20 epochs. The maximum training epoch was set to 120, the early stopping method was applied to prevent overfitting, and the training would stop if performance on the minority set did not improve over 30 epochs. The model was implemented with PyTorch 1.8.1 (36) on a Nvidia Titan X GPU with 12 G of memory.

Radiomics Analysis

After the segmentation model was well-trained, it was directly applied to the Rad-T, In-V, and Ex-V to obtain the ROI. Then, the following procedures of the radiomics analysis were applied to build the pCR prediction model.

Feature Extraction

In order to extract useful information related to the pCR status, a large feature space was generated, which included features not only from the original image but also from its wavelet decomposition images. A total of 93 types of features were calculated on each original image and its four Harr wavelet subbands, i.e., HH, HL, LH, and LL. The 93 features include 18 intensity features and 75 texture features. In addition, nine shape features were extracted on the original image. Together, 474 features were generated from each MRI slice.

Concerning that each patient had a different number of MRI slices, we used the arithmetic mean, three quartiles (Q1, Q2, and Q3 points), and the standard deviation of the features extracted from all of the slices as representations. Therefore, each patient had 2370 features in total.

All of the features were extracted using the Python package “PyRadiomics” (37). As announced in its document, most

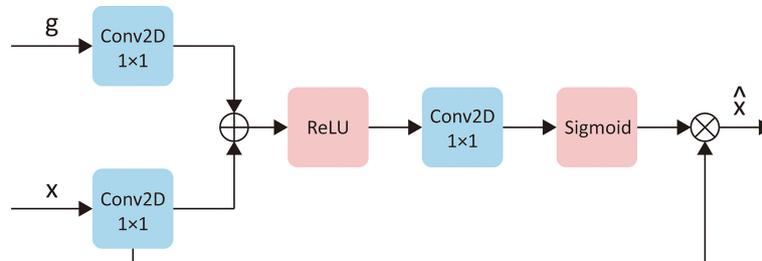


FIGURE 4 | The detailed explanation of attention gates. Let x represents the “skip connect” features from encoder, g represents the features from corresponding former decoder layer, then \hat{x} is the updated of x . “Conv2D 1×1 ” represents a 2D convention with kernel size of 1×1 . “ReLU” and “Sigmoid” are two different activation functions.

features meet the Introduction Intellidyne Business Systems (IBSI) standard, which would increase the reliability to our experiments. More details are shown on the “PyRadiomics documentation” website (<http://pyradiomics.readthedocs.io>).

Feature Selection

Two feature selection steps were applied in our study to increase the robustness and avoid overfitting. First, we evaluated the discriminative power of each single feature by calculating the Harrell’s concordance index (C-index) (38) between the features and the pCR status. The univariate Cox analysis is a commonly used procedure in survival analysis, and we adopted this method because it does not require the features to follow a normal distribution compared with a *t*-test. Let $A = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ denote the features of patients, β present the pCR status of patients, for each feature $\alpha_f \in A$, the C-index can be computed with the Equation 3.

$$C - \text{index}(\alpha_f, \beta) = \frac{\sum_{i,j} U(\alpha_{fi} < \alpha_{fj}) \cdot U(\beta_j < \beta_i)}{\sum_{i,j} U(\beta_j < \beta_i)}, \quad (3)$$

where α_{fi} is the *i*-th value of α_f , β_i is the *i*-th value of β , and $U(a < b) = 1$ if $a < b$ else 0.

By definition, C-index equals to 1 means the best discriminative power and C-index equals to 0.5 represents a theoretically result of random prediction. We calculate the maximum of $C\text{-index}(\alpha_f, \beta)$ and $C\text{-index}(-\alpha_f, \beta)$ as the predictive score of feature α_f . After calculation, scores are sorted and features with lower score are excluded.

In the second step, the remaining features were put into the least absolute shrinkage and selection operator (LASSO) (39) for further selection. LASSO is a logistic regression model with L1 regularization as a penalty of the coefficients. It will encourage the regression use of sparse features. The objective function of LASSO is:

$$\min \frac{1}{2} \|Ax - \beta\|_2^2 + \lambda \|x\|_1, \quad (4)$$

where A is the matrix of radiomics features, x is the coefficient of each feature, β is the pCR status, and λ is the regularization penalty coefficient.

Due to the L1 regularization, the LASSO forces the sum of the absolute value of the regression coefficients to be less than a fixed value, minimizing the residual sum of the squares. Such an operation forces the certain coefficient to zero. After the LASSO regression, features with a coefficient of non-zero are retained. Here, λ was determined using a grid search and 5-fold cross validation on 100 iterations between 0.01 and 0.2.

The pCR Status Prediction

In our study, the support vector machine (SVM) (40) was applied to predict the pCR status. As suggested in study (41), the radial basis function (RBF) kernel was used. The RBF kernel is defined as Equation 5:

$$K(a, b) = e^{(-\gamma \|a-b\|^2)} \quad (5)$$

where a, b are two samples from dataset and γ is a hyper-parameter. γ and the regularization coefficient C of SVM were

determined also by grid search on 5-fold cross validation within set $\{1/16, 1/8, 1/4, 1/2, 1\}$. After choosing the best γ and C , the SVM model was trained for the pCR prediction.

Performance Evaluation

Various evaluation metrics of suspicious region segmentation and pCR prediction are listed below.

Three metrics—dice coefficient, sensitivity, and specificity—were applied to evaluate the performance of segmentation. The dice coefficient’s definition has been given in Equation 1. Sensitivity (SEN) and specificity (SPC) are defined as Equation 6 and Equation 7, where TP, FP, TN, and FN denoted true positive, false positive, true negative, and false negative, respectively.

$$SEN = \frac{TP}{TP + FN} \quad (6)$$

$$SPC = \frac{TN}{TN + FP} \quad (7)$$

As for the pCR status prediction, five metrics—the area under receiver operating characteristic (ROC) curve (AUC), accuracy, sensitivity, specificity, and the F-score—were applied for the evaluation. The F-score is a weighted harmonic mean that comprehensively considers sensitivity and specificity, which can be calculated as Equation 8.

$$F_{\beta} - \text{score} = (1 + \beta^2) \cdot \frac{SEN \times SPC}{\beta^2 \cdot SPC + SEN} \quad (8)$$

Here, we included $F_{0.5}$, F_1 , and $F_{1.5}$ in order to provide a multiple trade-off between specificity and sensitivity under different situations.

RESULTS

Clinical Characteristics

In our study, 241 of 424 patients from Guangdong Institute of Gastroenterology, Sixth Affiliated Hospital of Sun Yat-sen University and 34 of 72 patients from Yunnan Cancer Center met the inclusion criteria and did not meet the exclusion criteria. After selection, the number of patients in SegT, Rad-T, In-V, and Ex-V were 88, 107, 46, and 34, respectively. More clinical information of the Rad-T, In-V, and Ex-V groups is provided in **Table 1**. Statistical comparisons were performed for each clinical characteristic between the two response groups (pCR vs. non-pCR). There were no statistical differences between the pCR and non-pCR in sex, age, and the pre-CRT N stage in all three sets.

Segmentation Performance

When training the segmentation model, Seg-T was randomly separated into two sets with percentages of 70% and 30%. The s.t. of the majority set was used for updating the model, and the minority set was used for selecting the best network parameters.

TABLE 1 | The clinical characteristics of patients in dataset Rad-T, In-V, and Ex-V.

Characteristics	Dataset Rad-T		P-value	Dataset In-V		p-value	Dataset Ex-V		p-value
	pCR ^a (n = 36)	NonpCR (n = 71)		pCR (n = 8)	NonpCR (n = 38)		pCR (n = 6)	NonpCR (n = 28)	
Sex			0.72			0.24			0.18
Male	8 (22.2%)	18 (25.4%)		5 (62.5%)	31 (81.6%)		2 (33.3%)	19 (67.9%)	
Female	28 (77.8%)	53 (74.6%)		3 (37.5%)	7 (18.4%)		4 (66.7%)	9 (32.1%)	
Age (mean ± SD ^b , year)	53.4 ± 12.1	56.9 ± 9.9	0.11	57.9 ± 8.4	55.0 ± 11.1	0.49	60.0 ± 10.8	59.2 ± 8.2	0.83
Pre-CRT ^c T stage			0.51			<0.01			0.31
T0	0 (0%)	0 (0%)		0 (0%)	0 (0%)		0 (0%)	0 (0%)	
T1	0 (0%)	0 (0%)		0 (0%)	0 (0%)		0 (0%)	0 (0%)	
T2	1 (2.8%)	3 (4.2%)		0 (0%)	0 (0%)		1 (16.7%)	0 (0%)	
T3	31 (86.1%)	55 (77.5%)		8 (100%)	30 (78.9%)		2 (33.3%)	8 (28.6%)	
T4	4 (11.1%)	13 (18.3%)		0 (0%)	8 (21.1%)		3 (50.0%)	20 (71.4%)	
Pre-CRT N stage			0.59			0.2			0.33
N0	4 (11.1%)	13 (18.4%)		0 (0%)	13 (34.2%)		1 (16.7%)	1 (3.6%)	
N1	17 (47.2%)	29 (40.8%)		4 (50.0%)	9 (23.7%)		1 (16.7%)	1 (3.6%)	
N2	15 (41.7%)	29 (40.8%)		4 (50.0%)	16 (42.1%)		4 (66.6%)	26 (92.8%)	
Post-CRT T stage			< 0.01			<0.01			< 0.01
T0	36 (100%)	0 (0%)		8 (100%)	0 (0%)		6 (100%)	0 (0%)	
T1	0 (0%)	7 (9.9%)		0 (0%)	2 (5.3%)		0 (0%)	4 (14.4%)	
T2	0 (0%)	18 (25.4%)		0 (0%)	11 (28.9%)		0 (0%)	6 (21.4%)	
T3	0 (0%)	43 (60.6)		0 (0%)	22 (57.9%)		0 (0%)	9 (32.1%)	
T4	0 (0%)	3 (4.2%)		0 (0%)	3 (7.9%)		0 (0%)	9 (32.1%)	
Post-CRT N stage			0.01			<0.01			0.17
N0	36 (100%)	50 (70.4%)		8 (100%)	30 (78.9%)		6 (100%)	21 (75.0%)	
N1	0 (0%)	20 (28.2%)		0 (0%)	8 (21.1%)		0 (0%)	6 (21.4%)	
N2	0 (0%)	1 (1.4%)		0 (0%)	0 (0%)		0 (0%)	1 (3.6%)	
Post-CRT CRM ^d			0.48			/			/
Negative	36 (100%)	70 (98.6%)		8 (100%)	38 (100%)		6 (100%)	28 (100%)	
Positive	0 (0%)	1 (1.4%)		0 (0%)	0 (0%)		0 (0%)	0 (0%)	

^apCR, pathologic complete response. ^bSD, standard deviation. ^cCRT, chemoradiotherapy, ^dCRM, circumferential resection margin.

After training, the model was validated on In-V, and the results are listed below.

The First Stage

The numerical results of stage one, segmentation of the rectum, are presented in **Table 2**. The results of the original U-Net are also provided for comparison. It can be seen that the 4-channel U-Net had a significant improvement compared to the original U-Net on the rectum segmentation. Moreover, as the 4-channel U-Net achieved (0.942, 0.965) 95% CI of the dice coefficient, it could be inferred that this network could provide a stable 300 and accurate rectum segmentation.

The visual results of the 4-channel U-Net and the original U-Net are also presented. **Figure 5** shows eight typical cases of rectum segmentation. The 4-channel U-Net had successfully segmented the rectum regions in all of the cases, while the U-Net showed different defects. In case (A), the U-Net was able to find the rectum, but the morphology lacked accuracy. In case (B), because the U-Net had no position information, the prostate was mistakenly judged as the rectum. In case (C), the U-Net outputted a continuous region containing both the rectum and the uterus. In case (D), due to the unclear rectal wall, the U-Net produced a bad result with an undesirable shape. In case (E), the prediction of the U-Net was less regular compared with the 4-channel U-Net. In case (G), the U-Net seriously under-segmented the rectum region. In case (H), the U-Net found two separated regions with similar sizes. Among those defects,

(A) and (E) might be improved by post-processing methods, but for the rest, even postprocessing methods such as image dilation or the removal of the smaller region seems useless to obtain the correct rectum region. Thus, our design in the first stage has great importance for guaranteeing the success of the following analysis.

The Second Stage

Numerical results of stage two, segmentation of the “suspicious region,” are presented in **Table 3**. For a fair comparison, the U-Net model in this stage was also given the rectum localization information from the first stage, and we intended to evaluate the use of the attention mechanism. From the results, the attention slightly improved the result. The dice coefficient and specificity may not be considerably high, but the sensitivity achieved nearly 0.8, indicating that the network could find the major portion of the “suspicious region.”

Figure 6 shows a visual display of the tsraU-Net, including the last attention maps, the final segmentation results, and the overlapping region compared with the gold standard. In some cases, such as (B), (D), and (F), the segmentation results are oversized. However, the morphology between them is still similar, and the segmentation results do not neglect most of the gold standard. Consequently, we believe the model is capable of providing the ROI with enough information for radiomics analysis.

TABLE 2 | Comparison between U-Net and 4-channel U-Net in the dice coefficient, sensitivity and specificity between the gold standard and the results from stage one, that is, 4-channel U-Net, in tsraU-Net as well as the baseline, original U-Net.

Model	Dice	Sensitivity	Specificity
U-Net	0.861 (95% CI ^a : 0.850, 0.873)	0.876 (95% CI: 0.864, 0.888)	0.867 (95% CI: 0.852, 0.882)
4-channel U-Net	0.954 (95% CI: 0.942, 0.965)	0.967 (95% CI: 0.955, 0.980)	0.96 (95% CI: 0.945, 0.976)

^aCI, confidence intervals.

Treatment Response Prediction

A total of 2370 features representing certain properties of the “suspicious region” predicted by tsraUNet were extracted from each patient. These features were progressively selected in the initial univariate analysis, and only approximately the top 2.5% features, which was 63 features, remained according to their predictive scores. The histogram of the predicted scores with the number of features is illustrated in **Figure 7A**. A distinct gap was found between the remaining features and the excluded features with a corresponding threshold of 0.622.

The remaining features were then put into LASSO. After a grid search and cross-validation, the best λ was 0.293. The grid search of λ in the LASSO regression to minimize the residual mean square error (MSE) is visually provided in **Figures 7B, C**, which provides the coefficients of the features during the grid search. Ten features were finally chosen, and they are presented in **Table 4**. The detailed descriptions of these features can be found in the “PyRadiomics documentation.”

The remaining 10 features was used to build a SVM classifier. The hyperparameters were decided after grid search and cross-validation: $C = 1$ and $\gamma = 0.125$. After training, the SVM achieved 0.924 of the AUC on Rad-T, 0.829 on In-V, and 0.815 on Ex-V. More numerical results are displayed in **Table 5**. In addition,

Figure 8 gives a visual display of the AUC and SVM scores [provided by the Python package “scikit-learn” (42)].

We further applied *t*-test on the three datasets to test the distribution of SVM score between real pCR and non-pCR patients. The *p*-values on Rad-T, In-V, and Ex-V are 1.26×10^{-10} , 1.41×10^{-3} , and 3.26×10^{-3} , respectively, indicating that the SVM score between real pCR and non-pCR is from different distribution under significance level $\alpha = 0.05$.

From the above results, we conclude that the radiomics model could extract information related to the pCR and predict its status with certain reliability. Furthermore, the results indicated that the “suspicious region” is capable to be the ROI in this research, which means a single post-nCRT T2-w MRI has the ability to predict the pCR status without the help of a pre-nCRT MRI or other post-nCRT modalities. The overall pipeline is provided in **Figure 9**.

DISCUSSION

In this study, we proposed a method for predicting the pCR status of LARC patients after nCRT that only requires the post-nCRT T2-w MRI and a few manual operations. We provided a

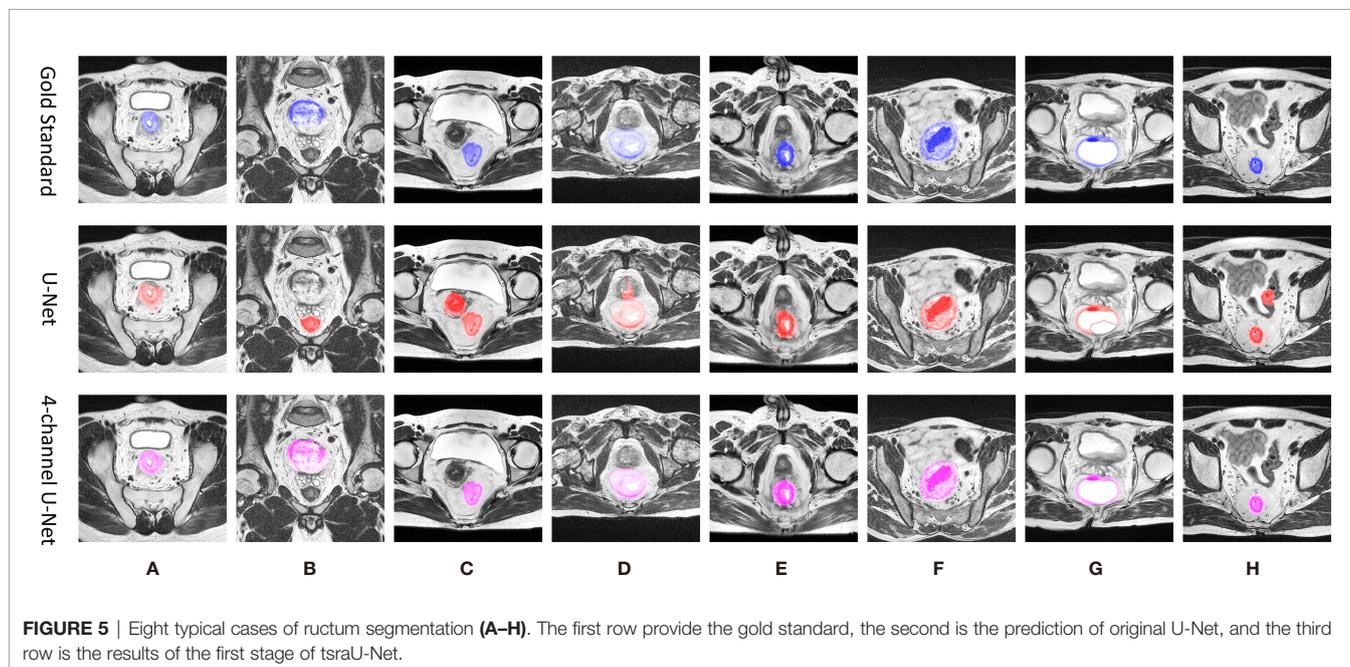
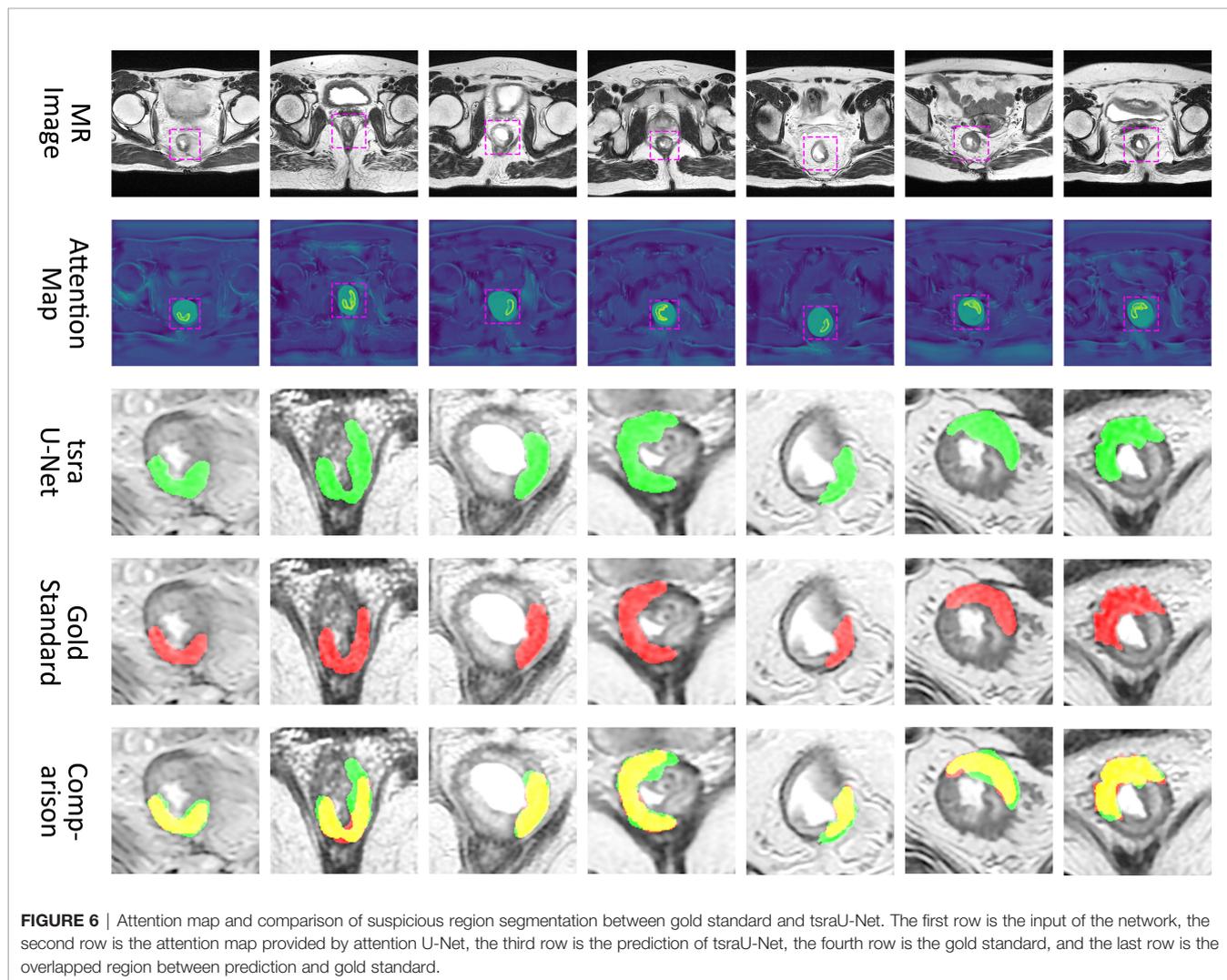


FIGURE 5 | Eight typical cases of rectum segmentation (A–H). The first row provide the gold standard, the second is the prediction of original U-Net, and the third row is the results of the first stage of tsraU-Net.

TABLE 3 | Comparison between U-Net and Attention U-Net in dice, sensitivity and specificity.

Model	Dice	Sensitivity	Specificity
U-Net	0.656 (95% CI ^a : 0.630, 0.683)	0.781 (95% CI: 0.750, 0.812)	0.624 (95% CI: 0.590, 0.659)
4-channel	0.66	0.785	0.632
U-Net	(95% CI: 0.628, 0.691)	(95% CI: 0.752, 0.817)	(95% CI: 0.594, 0.668)

^aCI, confidence intervals.



novel definition of the suspicious region and used it as the ROI to build the radiomics analysis. Furthermore, we designed a deep learning model for suspicious region segmentation that could greatly reduce the workload of radiologists. Our experimental results, 0.829/0.815 AUC on the internal/external validation set, prove the feasibility and stability of our method in pCR prediction, indicating that our method has great potential for providing assistance to doctors in clinical diagnosis.

Some information was obtained from the 10 selected features. First of all, no shape feature remained after selection. The reason

might come from the fact that all of the ROIs were provided by the deep networks, which had a homogeneous shape regardless of the pCR status. Furthermore, half of the selected features were wavelet features, which might imply that some valuable information was hidden in the frequency domain. Many recent studies have also highlighted the importance of wavelet features in radiomics analysis (43–46). Finally, 8 of the 10 features were quantiles, and 1 was the standard deviation. This suggested that simply averaging the features from all of the slices per patient was insufficient compared with using multiple statistics.

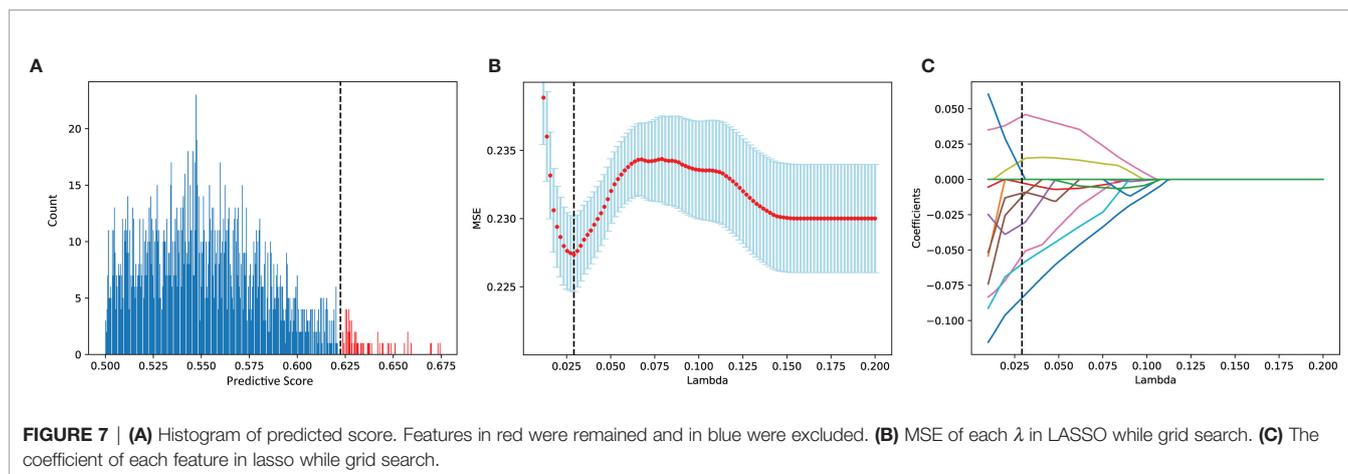


TABLE 4 | The features finally remained, and their coefficients.

Features Group	Abbreviation	Attribute	Coefficient
original firstorder	Maximum	Q2	-0.03166
original firstorder	Maximum	Q3	-0.009714
original glcm	MCC	SD	-0.083918
original glcm	LargeDependenceHighGrayLevelEmphasis	Q2	-0.002476
original glcm	LowGrayLevelEmphasis	Q1	0.044713
wavelet-LH ngtdm	Busyness	Q1	-0.012269
wavelet-LH ngtdm	Busyness	Q2	-0.054026
wavelet-LH ngtdm	Strength	Q2	0.013605
wavelet-HL firstorder	Median	AVG	-0.059893
wavelet-LL glszm	LargeAreaLowGrayLevelEmphasis	Q3	0.004648

TABLE 5 | The pCR^a status predicted performance on datasets Rad-T, In-V and Ex-V, in terms of AUC, accuracy, sensitivity, specificity, F_{0.5}-score, F₁-score, F_{1.5}-score.

Dataset	AUC ^b	Accuracy	Sensitivity	Specificity	F0.5-score	F ₁ -score	F _{1.5} -score
Rad-T	0.924 (95% CI: 0.923, 0.926)	0.860 (95% CI: 0.856, 0.863)	0.861 (95% CI: 0.855, 0.867)	0.859 (95% CI: 0.855, 0.863)	0.860 (95% CI: 0.856, 0.863)	0.860 (95% CI: 0.820, 0.880)	0.860 (95% CI: 0.825, 0.901)
In-V	0.829 (95% CI: 0.821, 0.837)	0.804 (95% CI: 0.794, 0.815)	0.750 (95% CI: 0.720, 0.780)	0.816 (95% CI: 0.805, 0.827)	0.802 (95% CI: 0.789, 0.811)	0.782 (95% CI: 0.689, 0.793)	0.769 (95% CI: 0.722, 0.794)
Ex-V	0.815 (95% CI: 0.801, 0.830)	0.853 (95% CI: 0.841, 0.865)	0.500 (95% CI: 0.453, 0.548)	0.929 (95% CI: 0.919, 0.938)	0.793 (95% CI: 0.746, 0.813)	0.650 (95% CI: 0.634, 0.678)	0.583 (95% CI: 0.555, 0.615)

^apCR, pathologic complete response. ^bAUC, area under the curve; ^cCI, confidence intervals.

To further address the thoughtful design of our study, we wanted to highlight the importance of independent training of the segmentation and radiomics model. As a model always tends to overfit more or less while training, if we use the same set to train the segmentation as well as the radiomics, the predicted ROI on the training set and validation set might have different distributions, and consequently affect the stability of the radiomics analysis. Due to the above consideration, we used a particular dataset, Seg-T, to train the segmentation model and choose the network parameters. We then applied the best model on Rad-T, In-V,

and Ex-V to assure the independence between segmentation and radiomics.

Unlike other studies (47, 48) that have utilized a two-sample *t*-test as the first step to select the features, we used univariate Cox analysis. In fact, prior to the feature selection, we examined whether each feature was normally distributed by calculating the skewness and kurtosis (49). A total of 34 features did not pass this simple normality test under a significance level $\alpha = 0.05$. Therefore, *t*-test could not be applied to all features in our study. Consequently, we utilized the concordance index as a replacement because it did not limit the distribution of data.

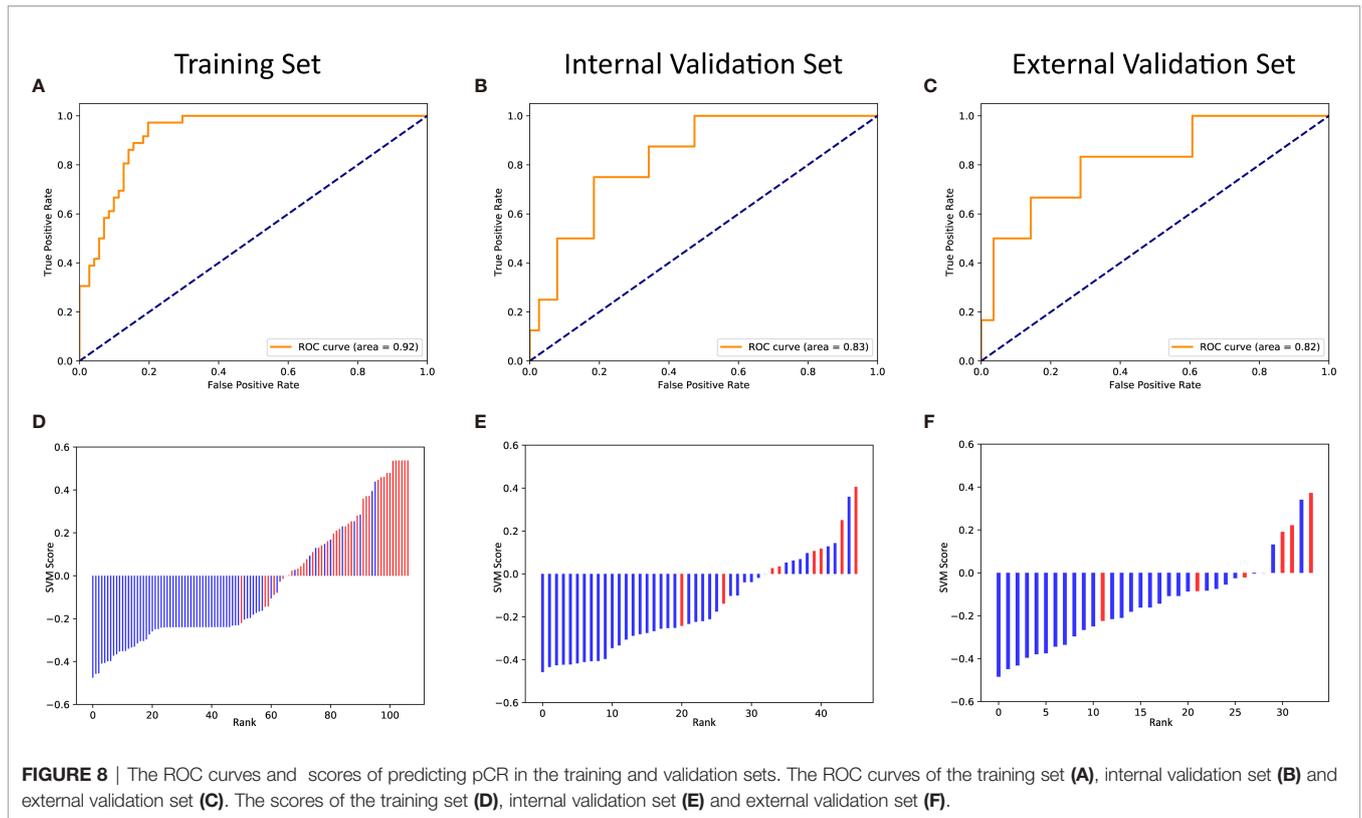


FIGURE 8 | The ROC curves and scores of predicting pCR in the training and validation sets. The ROC curves of the training set (A), internal validation set (B) and external validation set (C). The scores of the training set (D), internal validation set (E) and external validation set (F).

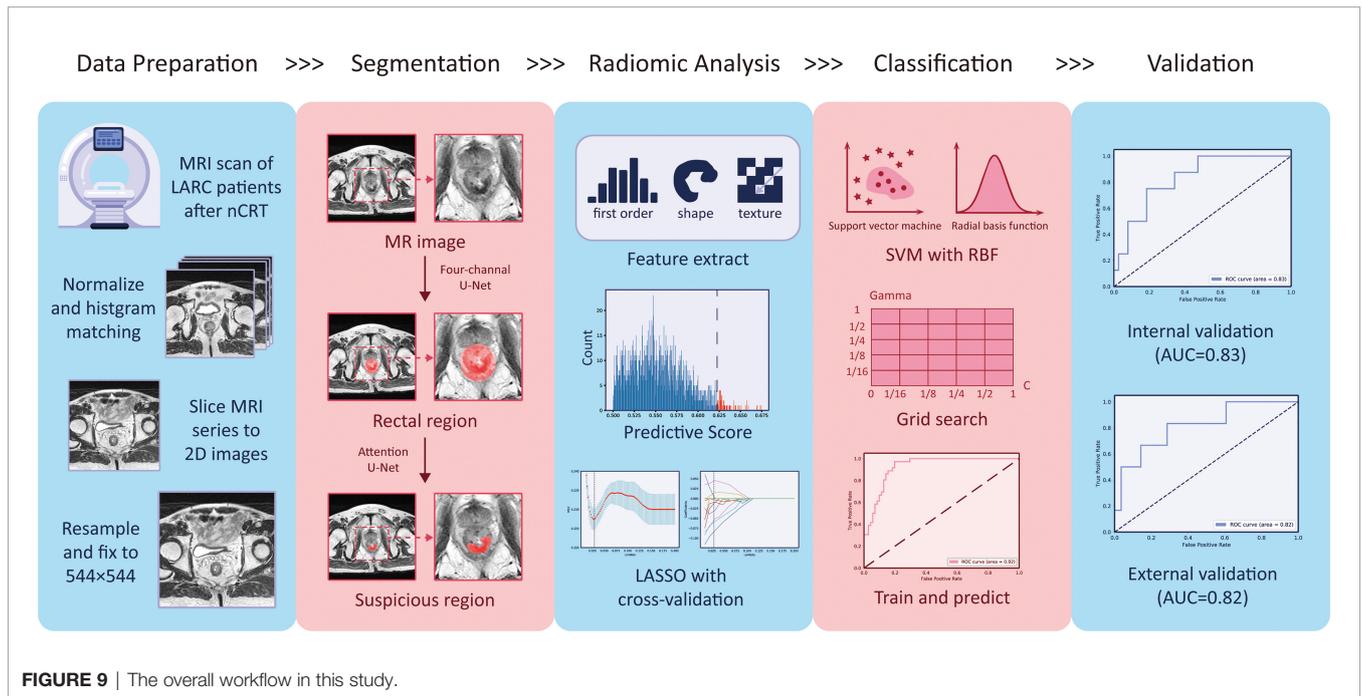


FIGURE 9 | The overall workflow in this study.

There were some limitations in our study. However, this was only a preliminary exploration. In the future, we will attempt further improvements. For instance, we could collect more unified standard data for analysis, and at the same time

carefully choose the year of patient recruitment to avoid data mixing. In addition, we could encourage radiologists to delineate more precisely so that our segmentation network could better learn the characteristics of the suspicious regions. Additionally,

our segmentation network could be adjusted and improved using techniques, such as combining different loss functions following previous works (50–53) or adding clinical characteristics for a joint analysis. In addition, if we could obtain MRI with smaller slice thickness ($\leq 1\text{mm}$), we could consider building a 3D model and studying the 3D radiomics features that may contain richer information of the suspicious region.

Finally, the motivation of this study was different than other related works. We wanted to explore the possibility of using a single post-nCRT T2 MRI for patients missing a pre-nCRT MRI or other modalities. In addition, we intended to improve the efficiency of the model and reduce the workload of doctors for clinical use. Our method has great potential to guide less-experienced doctors, as it does not require manual delineation of the ROI. Moreover, as our model is less restricted regarding data requirements and the prediction of the pCR was easily obtained, it can be combined with other studies for joint decision-making.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

REFERENCES

- Sung H, Ferlay J, Siegel R, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: Globocan Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: Cancer J Clin* (2021) 71:209–49. doi: 10.3322/caac.21660
- Siegel RL, Miller KD, Goding Sauer A, Fedewa SA, Butterly LF, Anderson JC, et al. Colorectal Cancer Statistics, 2020. *CA: Cancer J Clin* (2020) 70:145–64. doi: 10.3322/caac.21601
- Xu Q, Xu Y, Sun H, Chan Q, Shi K, Song A, et al. Quantitative Intravoxel Incoherent Motion Parameters Derived From Whole-Tumor Volume for Assessing Pathological Complete Response to Neoadjuvant Chemotherapy in Locally Advanced Rectal Cancer. *J Magnetic Resonance Imaging Jmri* (2018) 48:248–58. doi: 10.1002/jmri.25931
- Valentini V, Aristei C, Glimelius B, Minsky BD, Beets-Tan R, Borrás JM, et al. Multidisciplinary Rectal Cancer Management: 2nd European Rectal Cancer Consensus Conference (Eureca-Cc2). *Radiother Oncol* (2009) 92:148–63. doi: 10.1016/j.radonc.2009.06.027
- van de Velde CJ, Boelens PG, Borrás JM, Coebergh JW, Cervantes A, Blomqvist L, et al. EURECCA Colorectal: Multidisciplinary Management: European Consensus Conference Colon & Rectum. *Eur J Cancer* (2014) 50:1–e1. doi: 10.1016/j.ejca.2013.06.048
- van Gijn W, Marijnen CA, Nagtegaal ID, Kranenburg EMK, Putter H, Wiggers T, et al. Preoperative Radiotherapy Combined With Total Mesorectal Excision for Resectable Rectal Cancer: 12-Year Follow-Up of the Multicentre, Randomised Controlled Trial. *Lancet Oncol* (2011) 12:575–82. doi: 10.1016/S1470-2045(11)70097-3
- Sauer R, Becker H, Hohenberger W, Rödel C, Wittekind C, Fietkau R, et al. Preoperative Versus Postoperative Chemoradiotherapy for Rectal Cancer. *New Engl J Med* (2004) 351:1731–40. doi: 10.1056/NEJMoa040694
- Jia H, Shen X, Guan Y, Xu M, Tu J, Mo M, et al. Predicting the Pathological Response to Neoadjuvant Chemoradiation Using Untargeted Metabolomics in Locally Advanced Rectal Cancer. *Radiother Oncol* (2018) 128:548–56. doi: 10.1016/j.radonc.2018.06.022

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by This clinical trial was approved by the Clinical Ethics Review Committee of the Sixth Affiliated Hospital of Sun Yat-sen University (2020ZSLYEC-010).

AUTHOR CONTRIBUTIONS

XP, FW, QZ, and YL jointly wrote the manuscript. XP, RH, and XY collected all the input sources data. XP and FW annotated the images data. FW, QZ, and YL designed the model and implemented the main algorithm and other computational analysis. XF performed a review of the manuscript on the clinical aspects. All authors contributed to the article and approved the submitted version.

FUNDING

Guangdong Science and Technology Project (No. 2019B030316003) and The Sixth Affiliated Hospital of Sun Yat-Sen University Clinical Research 1010 Program [No. 1010PY (2020)-09] supported this study.

- Sanghera P, Wong D, McConkey C, Geh J, Hartley A. Chemoradiotherapy for Rectal Cancer: An Updated Analysis of Factors Affecting Pathological Response. *Clin Oncol* (2008) 20:176–83. doi: 10.1016/j.clon.2007.11.013
- Maas M, Nelemans PJ, Valentini V, Das P, Rödel C, Kuo LJ, et al. Long-Term Outcome in Patients With a Pathological Complete Response After Chemoradiation for Rectal Cancer: A Pooled Analysis of Individual Patient Data. *Lancet Oncol* (2010) 11:835–44. doi: 10.1016/S1470-2045(10)70172-8
- Lu W, Jing H, Ju-Mei Z, Shao-Lin N, Fang C, Xiao-Ping Y, et al. Intravoxel Incoherent Motion Diffusion-Weighted Imaging for Discriminating the Pathological Response to Neoadjuvant Chemoradiotherapy in Locally Advanced Rectal Cancer. *Sci Rep* (2017) 7:1–9. doi: 10.1038/s41598-017-09227-9
- Hupkens BJ, Martens MH, Stoot JH, Berbee M, Melenhorst J, Beets-Tan RG, et al. Quality of Life in Rectal Cancer Patients After Chemoradiation: Watch-and-Wait Policy Versus Standard Resection—a Matched-Controlled Study. *Dis Colon Rectum* (2017) 60:1032–40. doi: 10.1097/DCR.0000000000000862
- Renhan AG, Malcomson L, Emsley R, Gollins S, Maw A, Myint AS, et al. Watch-And-Wait Approach Versus Surgical Resection After Chemoradiotherapy for Patients With Rectal Cancer (the Oncore Project): A Propensity-Score Matched Cohort Analysis. *Lancet Oncol* (2016) 17:174–83. doi: 10.1016/S1470-2045(15)00467-2
- Marijnen CA. Organ Preservation in Rectal Cancer: Have All Questions Been Answered? *Lancet Oncol* (2015) 16:e13–22. doi: 10.1016/S1470-2045(14)70398-5
- Horvat N, Carlos Tavares Rocha C, Clemente Oliveira B, Petkovska I, Gollub MJ. MRI of Rectal Cancer: Tumor Staging, Imaging Techniques, and Management. *Radiographics* (2019) 39:367–87. doi: 10.1148/rg.2019180114
- Patel UB, Brown G, Rutten H, West N, Sebag-Montefiore D, Glynne-Jones R, et al. Comparison of Magnetic Resonance Imaging and Histopathological Response to Chemoradiotherapy in Locally Advanced Rectal Cancer. *Ann Surg Oncol* (2012) 19:2842–52. doi: 10.1245/s10434-012-2309-3
- Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, Van Stiphout RG, Granton P, et al. Radiomics: Extracting More Information From Medical Images Using Advanced Feature Analysis. *Eur J Cancer* (2012) 48:441–6. doi: 10.1016/j.ejca.2011.11.036

18. Li Z, Duan H, Zhao K, Ding Y. Stability of Mri Radiomics Features of Hippocampus: An Integrated Analysis of Test-Retest and Inter-Observer Variability. *IEEE Access* (2019) 7:97106–16. doi: 10.1109/ACCESS.2019.2923755
19. Cui Y, Yang X, Shi Z, Yang Z, Du X, Zhao Z, et al. Radiomics Analysis of Multiparametric Mri for Prediction of Pathological Complete Response to Neoadjuvant Chemoradiotherapy in Locally Advanced Rectal Cancer. *Eur Radiol* (2019) 29:1211–20. doi: 10.1007/s00330-018-5683-9
20. Giannini V, Mazzetti S, Bertotto I, Chiarenza C, Cauda S, Delmastro E, et al. Predicting Locally Advanced Rectal Cancer Response to Neoadjuvant Therapy With 18 F-Fdg Pet and Mri Radiomics Features. *Eur J Nucl Med Mol Imaging* (2019) 46:878–88. doi: 10.1007/s00259-018-4250-6
21. Shaish H, Aukerman A, Vanguri R, Spinelli A, Armenta P, Jambawalikar S, et al. Radiomics of Mri for Pretreatment Prediction of Pathologic Complete Response, Tumor Regression Grade, and Neoadjuvant Rectal Score in Patients With Locally Advanced Rectal Cancer Undergoing Neoadjuvant Chemoradiation: An International Multicenter Study. *Eur Radiol* (2020) 30:6263–73. doi: 10.1007/s00330-020-06968-6
22. Wan L, Peng W, Zou S, Ye F, Zhang H. Mri-Based Delta-Radiomics are Predictive of Pathological Complete Response After Neoadjuvant Chemoradiotherapy in Locally Advanced Rectal Cancer. *Acad Radiol* (2020). doi: 10.1016/j.acra.2020.10.026
23. Alvarez-Jimenez C, Antunes JT, Talasila N, Bera K, Brady JT, Gollamudi J, et al. Radiomic Texture and Shape Descriptors of the Rectal Environment on Post-Chemoradiation T2-Weighted Mri Are Associated With Pathologic Tumor Stage Regression in Rectal Cancers: A Retrospective, Multi-Institution Study. *Cancers (Basel)* (2020) 12:2027. doi: 10.3390/cancers12082027
24. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *Int Conf Med Image Comput Computer-assisted intervention (Springer)* (2015) 9351:234–41. doi: 10.1007/978-3-319-24574-4_28
25. Chen W, Liu B, Peng S, Sun J, Qiao X. S3d-Unet: Separable 3d U-Net for Brain Tumor Segmentation. *Int MICCAI Brainlesion Workshop (Springer)* (2018) 11384:358–68. doi: 10.1007/978-3-030-11726-9_32
26. Li X, Chen H, Qi X, Dou Q, Fu CW, Heng PA. H-Denseunet: Hybrid Densely Connected Unet for Liver and Tumor Segmentation From Ct Volumes. *IEEE Trans Med Imaging* (2018) 37:2663–74. doi: 10.1109/TMI.2018.2845918
27. Balagopal A, Kazemifar S, Nguyen D, Lin MH, Hannan R, Owangi A, et al. Fully Automated Organ Segmentation in Male Pelvic Ct Images. *Phys Med Biol* (2018) 63:245015. doi: 10.1088/1361-6560/aaf11c
28. Zhang L, Mohamed AA, Chai R, Guo Y, Zheng B, Wu S. Automated Deep Learning Method for Whole-Breast Segmentation in Diffusion-Weighted Breast Mri. *J Magnetic Resonance Imaging* (2020) 51:635–43. doi: 10.1002/jmri.26860
29. Benson AB, Venook AP, Al-Hawary MM, Cederquist L, Chen Y-J, Ciombor KK, et al. Nccn Clinical Practice Guidelines in Oncology. *J Natl Compr Canc Netw* (2018). *Rectal Cancer* (2018) 16:874–901. doi: 10.6004/jnccn.2018.0061
30. Beets-Tan RGH, Lambregts DMJ, Maas M, Bipat S, Barbaro B, Curvo-Semedo Luis, et al. Magnetic Resonance Imaging for Clinical Management of Rectal Cancer: Updated Recommendations From the 2016 European Society of Gastrointestinal and Abdominal Radiology (Esgar) Consensus Meeting. *Eur Radiol* (2018) 28:874–901. doi: 10.1007/s00330-017-5204-2
31. Martí-Bonmati L, Alberich-Bayarri A. *Imaging Biomarkers: Development and Clinical Integration*. Springer (2016) 376.
32. Lowekamp BC, Chen DT, Ibáñez L, Blezek D. The Design of Simpleitk. *Front Neuroinf* (2013) 7:45. doi: 10.3389/fninf.2013.00045
33. Barbaro B, Fiorucci C, Tebala C, Valentini V, Gambacorta MA, Vecchio FM, et al. Locally Advanced Rectal Cancer: Mr Imaging in Prediction of Response After Preoperative Chemotherapy and Radiation Therapy. *Radiology* (2009) 250:730–9. doi: 10.1148/radiol.2503080310
34. Oktay O, Schlemper J, Folgoc LL, Lee M, Heinrich M, Misawa K, et al. Attention U-Net: Learning Where to Look for the Pancreas. *ArXiv* (2018). Available at: <https://arxiv.org/abs/1804.03999>.
35. Kingma DP, Ba J. Adam: A Method for Stochastic Optimization. *ArXiv* (2014). Available at: <https://arxiv.org/abs/1412.6980>.
36. Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, et al. Pytorch: An Imperative Style, High-Performance Deep Learning Library. *ArXiv* (2019). Available at: <https://arxiv.org/abs/1912.01703>.
37. Van Griethuysen JJ, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Res* (2017) 77:e104–7. doi: 10.1158/0008-5472.CAN-17-0339
38. Harrell FE, Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the Yield of Medical Tests. *Jama* (1982) 247:2543–6. doi: 10.1001/jama.247.18.2543
39. Tibshirani R. Regression Shrinkage and Selection via the Lasso: A Retrospective. *J R Stat Soc: Ser B Stat Methodol* (2011) 73:273–82. doi: 10.1111/j.1467-9868.2011.00771.x
40. Suykens JA, Vandewalle J. Least Squares Support Vector Machine Classifiers. *Neural Process Lett* (1999) 9:293–300. doi: 10.1023/A:1018628609742
41. Smola AJ, Schölkopf B, Müller KR. The Connection Between Regularization Operators and Support Vector Kernels. *Neural Networks* (1998) 11:637–49. doi: 10.1016/S0893-6080(98)00032-X
42. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-Learn: Machine Learning in Python. *J Mach Learn Res* (2011) 12:2825–30.
43. Chen B, Zhang R, Gan Y, Yang L, Li W. Development and Clinical Application of Radiomics in Lung Cancer. *Radiat Oncol* (2017) 12:1–8. doi: 10.1186/s13014-017-0885-x
44. Su C, Jiang J, Zhang S, Shi J, Xu K, Shen N, et al. Radiomics Based on Multicontrast Mri can Precisely Differentiate Among Glioma Subtypes and Predict Tumour-Proliferative Behaviour. *Eur Radiol* (2019) 29:1986–96. doi: 10.1007/s00330-018-5704-8
45. Yang F, Ford JC, Dogan N, Padgett KR, Breto AL, Abramowitz MC, et al. Magnetic Resonance Imaging (Mri)-Based Radiomics for Prostate Cancer Radiotherapy. *Trans Androl Urol* (2018) 7:445. doi: 10.21037/tau.2018.06.05
46. Chetan MR, Gleeson FV. Radiomics in Predicting Treatment Response in Non-Small-Cell Lung Cancer: Current Status, Challenges and Future Perspectives. *Eur Radiol* (2020) 31:1049–58. doi: 10.1007/s00330-020-07141-9
47. Liu Z, Zhang XY, Shi YJ, Wang L, Zhu HT, Tang Z, et al. Radiomics Analysis for Evaluation of Pathological Complete Response to Neoadjuvant Chemoradiotherapy in Locally Advanced Rectal Cancer. *Clin Cancer Res* (2017) 23:7253–62. doi: 10.1158/1078-0432.CCR-17-1038
48. Wang H, Hu D, Yao H, Chen M, Li S, Chen H, et al. Radiomics Analysis of Multiparametric Mri for the Preoperative Evaluation of Pathological Grade in Bladder Cancer Tumors. *Eur Radiol* (2019) 29:6182–90. doi: 10.1007/s00330-019-06222-8
49. Bai J, Ng S. Tests for Skewness, Kurtosis, and Normality for Time Series Data. *J Bus Econ Stat* (2005) 23:49–60. doi: 10.1198/073500104000000271
50. Lin TY, Goyal P, Girshick R, He K, Dollár P. (2017). Focal Loss for Dense Object Detection. In: *Proceedings of the IEEE international conference on computer vision* (2017) 42:2980–8.
51. Rahman MA, Wang Y. Optimizing Intersection-Over-Union in Deep Neural Networks for Image Segmentation. *Int Symp Visual Comput (Springer)* (2016) 10072:234–44. doi: 10.1007/978-3-319-50835-1_22
52. Kervadec H, Bouchtiba J, Desrosiers C, Granger E, Dolz J, Ayed IB. Boundary Loss for Highly Unbalanced Segmentation. *ArXiv* (2019). Available at: <https://arxiv.org/abs/1812.07032>.
53. Berman M, Triki AR, Blaschko MB. The Lovász-Softmax Loss: A Tractable Surrogate for the Optimization of the Intersection-Over-Union Measure in Neural Networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. (2018) 4413–21.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Pang, Wang, Zhang, Li, Huang, Yin and Fan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.