

RCpCR-Net: A DCGAN&3DCNN-based Framework for Rectal Cancer Pathological Complete Response Prediction

Lanlan Li¹, Mingshi Wang¹, Ziyue Wang¹, Tao Chen¹,
Chongyang Wang¹, Bin Xu¹, Juan Li^{2*}, Decao Niu^{3*}

¹Fujian Key Lab for Intelligent Processing and Wireless Transmission of Media Information, College of Physics and Information Engineering, Fuzhou University, Fuzhou, 350108, Fujian, China.

²Department of Endoscopic Surgery, The Sixth Affiliated Hospital, Sun Yat-sen University, Guangzhou, 510655, Guangdong, China.

³Department of Urological Surgery, Guangdong Second Provincial General Hospital, Guangzhou, 510317, Guangdong, China.

*Corresponding author(s). E-mail(s): lijuan67@mail.sysu.edu.cn;
ndcdoot@sr.gxmu.edu.cn;

Abstract

Purpose: Deep learning can predict the pathological complete response (pCR) in rectal cancer patients. This prediction helps doctors make informed decisions before radiotherapy and surgical resection. However, most current relative studies use private datasets with small data volumes. Moreover, most studies still use Two-dimensional (2D) image data, ignoring the continuity between the overall data slices.

Methods: To address these issues, this study proposes a novel prediction framework, RCpCR-Net, based on Three-dimensional Magnetic Resonance Imaging (3D MRI) for rectal cancer pCR. To enhance prediction accuracy, we employ an improved DCGAN (RC-GAN) for data augmentation. We utilized upsampling and convolution in generator to diminish the impact of "artifacts" on generated images. We utilized an AlexNet-based discriminator to further enhanced generated images' quality. Besides, we use the Convolutional Block Attention Module (CBAM) for feature extraction and capturing spatial and channel information in 3DCNN to predict pCR.

Results: The experimental results demonstrate significant improvements in accuracy, specificity, and sensitivity through RCpCR-Net. In detail, the accuracy

is improved to 0.778, specificity to 0.796, and sensitivity to 0.754. Compared to the baseline network, these values have increased by 8.8%, 9.9%, and 9.1% respectively.

Conclusion: These findings indicate that RCpCR-Net overcomes the challenges of limited datasets and 2D images. By incorporating data augmentation and attention mechanisms, RCpCR-Net offers a potential tool for doctors to avoid unnecessary radiotherapy and surgical procedures.

Keywords: RCpCR-Net, Rectal Cancer, pCR, 3D MRI, DCGAN,3DCNN

1 Introduction

Currently, with the rapid development of society, people's dietary structure and lifestyle have been impacted. As a result, the incidence of colorectal cancer is showing a year-on-year rapid increase. According to the latest report published in the Journal of Cancer Science and Advances by the National Cancer Center, colorectal cancer has become the "new king of cancer" globally. It has ranked among the top five in both incidence and mortality rates[1]. However, early-stage colon cancer can be completely cured through surgery, and even small metastases can be managed satisfactorily through surgery, radiation therapy, and chemotherapy. Therefore, prevention and early detection of colorectal cancer are of utmost importance[2].

The detection of colorectal cancer can be divided into invasive and non-invasive methods. The former includes pathological sections and colonoscopy, while the latter includes computed tomography (CT) and magnetic resonance imaging (MRI)[3]. MRI uses the principle of nuclear magnetic resonance and computers to assist in constructing human body information. Its screening accuracy is relatively high. It can not only detect the condition of the surrounding tissues of the colorectum, but also provide a reliable basis for the staging of colorectal cancer based on the degree of tumor infiltration of the intestinal wall. It is currently a highly recommended colorectal cancer screening method.

In 70% of diagnosed cases of colorectal cancer, patients are in the locally advanced stage. For patients at this stage, it is recommended to use neoadjuvant chemoradiotherapy combined with total mesangial excision to reduce the local recurrence rate[4, 5]. Patients who receive this treatment may achieve a pathological Complete Response (pCR). For pCR patients, surgical resection may not lead to an overall increase in survival and disease-free survival, but it may result in extensive complications such as sexual, urinary, and bowel dysfunction, or even increase the risk of anastomotic leakage in pCR patients[6, 7]. Therefore, in clinical practice, a "watch-and-wait" approach is commonly used to avoid these risks and achieve similar patient survival rates as radical resection[8, 9]. However, pCR can only be confirmed by evaluating the postoperative resected specimens. Therefore, prior to treatment, early prediction of tumor response using non-invasive methods can help improve patient compliance and confidence in treatment. And it may also avoid unnecessary surgical resection after treatment, while providing valuable information for the next steps of treatment.

In recent years, utilizing deep learning techniques combined with medical imaging for cancer diagnosis has become a popular trend. However, medical images often suffer from limited dataset size and imbalanced positive and negative samples, which is an inherent challenge. Traditional data augmentation methods mainly involve random horizontal and vertical flipping, random rotation, etc. These methods indeed augment the data volume to some extent, but ultimately, they are simple transformations applied to existing images[10, 11]. Currently, most research on colorectal cancer is based on the 2D analysis of patient histopathology images. However, in reality, tumor images captured through CT or MRI are reconstructed into 3D data from multiple slices. Utilizing the 3D image data and networks directly for research and analysis can capture tumor information to a greater extent and facilitate more comprehensive results display. In recent years, researchers have started using 3D multiscale convolutional neural networks for 3D automatic segmentation and pathological response prediction of colorectal cancer tumors[12]. Overall, there is currently limited research on the use of 3D networks in neoadjuvant treatment of rectal cancer. Additionally, the quantity of available 3D medical image data for rectal cancer or other diseases is not substantial. Hence, there is a need to address how to increase the dataset size under existing conditions. Therefore, it is urgent to generate new images based on the original image features rather than simply applying rotations, which can bring new features for the network to learn. Moreover, the generated images should have a more similar feature distribution to the original images. By constructing more accurate and high-quality 3D data, it can meet the requirements of combining deep learning techniques with medical image research for colorectal cancer.

In this study, we combined the improved DCGAN and the improved 3DCNN to construct a framework, RCpCR-Net, for predicting pCR patients by 3D MRI images. We conducted experiments using 3D MRI images of rectal cancer patients to test the performance of RCpCR-Net. The dataset came from The Sixth Affiliated Hospital, Sun Yat-sen University, Guangdong Provincial Hospital, collected from 2010 to 2015. To address the issue of insufficient 3D MRI images in current colorectal cancer research, we proposed an improved DCGAN network, RC-GAN, to augment the data. To alleviate the issue of prominent checkerboard artifacts in the generator, we introduced an Upsample Convolution Block. Additionally, we replaced the deconvolutional layers with the overlapping pooling structure of AlexNet for the discriminator. This approach better preserved the original information, enhanced the performance of the discriminator, and improved the quality of generated images, thereby enhancing the overall performance of the DCGAN network. Through the aforementioned technological improvements, RC-GAN successfully constructed more accurate and high-quality 3D data resources in colorectal cancer research, satisfying the research needs in this field. Besides, we utilized different attention mechanism modules to optimize the 3DCNN network, so that RCpCR-Net can accurately predict pCR patients through 3D MRI images.

2 Materials and Methods

2.1 Data Collection and Preparation

This study used an 3D MRI image dataset of rectal cancer patients from The Sixth Affiliated Hospital, Sun Yat-sen University, Guangdong Provincial Hospital, from 2010 to 2015, and was registered in the US Clinical Trials Registry (ClinicalTrials.gov) with registry number NCT01211210. The population and related strategies of cancer treatment in the trial are detailed in references[13, 14]. All locally advanced rectal cancer patients received neoadjuvant therapy prior to surgical treatment. The resected tumor specimens were evaluated for pathological response using the Tumor Response Grading (TRG) system by two pathologists[15]. A total of 99 3D samples (with 1889 2D slices) were divided into two different groups: pCR group (TRG 0, no viable residual tumor cells) and npCR group (TRG 1-4, from rare residual cancer cells to extensive residual cancer cells). The pCR group contains 80 samples and the npCR group contains 19 samples. Experienced experts determined whether patients achieved pCR based on the pathological report post-surgery. The initial dataset is in Nearly Raw Raster Data (NRRD) format, 3D MRI images composed of 2D slices. This article has converted the NRRD images to PNG format for subsequent processing and extracted the 2D slices with tumor regions. Partial data is shown in Fig.1.

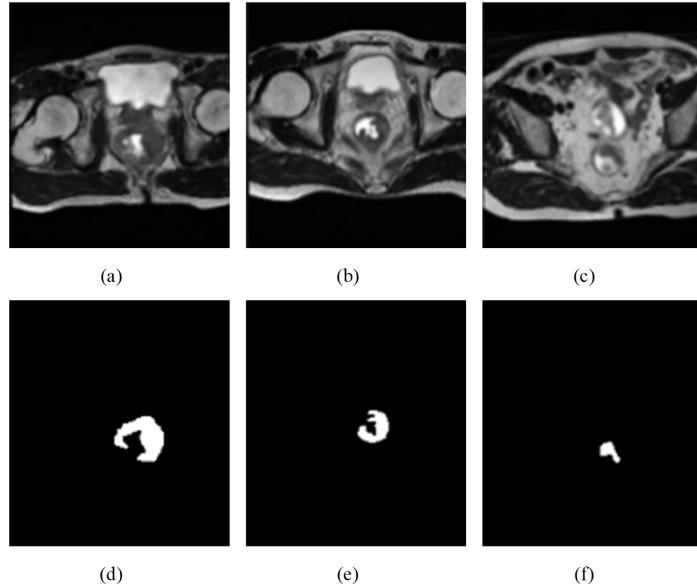


Fig. 1: 3D MRI images and their corresponding mask images from the dataset of the Sixth Affiliated Hospital of Sun Yat-sen University

2.2 Overall Structure of RCpCR-Net Framework

In this study, we addressed the long-standing issues of data scarcity and class imbalance in the field of medical imaging by proposing an improved data augmentation network, RC-GAN, based on DCGAN. RC-GAN created new images based on the original image features and applied them to predict 3D images of rectal cancer. Compared to traditional methods, this approach generated new data with novel features and similar feature distributions for the network to learn from. Furthermore, this study directly focused on predicting 3D MRI images of rectal cancer, which better preserved the continuity between slices. The research methodology involved using RC-GAN for data augmentation and training/predicting with an improved Three-dimensional Convolutional Neural Network (3DCNN) using the augmented data. The overall structure of this novel framework, RCpCR-Net is illustrated in Fig.2. The main process involved cropping 2D tumor slices from existing real rectal 3D MRI images and using RC-GAN to generate virtual tumor images, thereby increasing the size of the dataset. Due to hardware limitations, the generation of images was done in a 2D manner. Therefore, after generating the images, the corresponding 2D slices were processed to compose multiple 2D slices into a single 3D case. The augmented 3D case dataset, constructed from the real and generated 3D MRI images, was then inputted into the improved 3DCNN for pCR and npCR prediction.

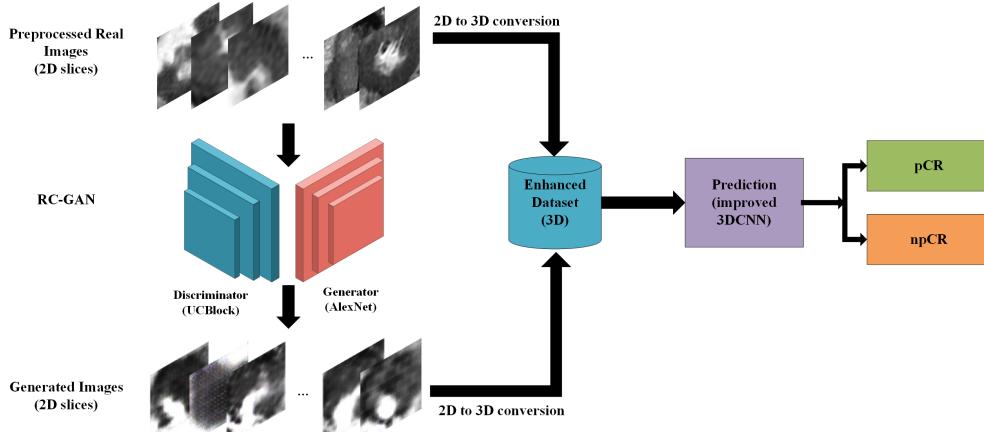


Fig. 2: Overall Structure of RCpCR-Net Framework

2.3 Structure of RC-GAN Data Augmentation Network

General DCGAN consists of two main modules: the generator and the discriminator. Unlike traditional GANs, DCGAN uses two convolutional neural networks as the generator and discriminator, respectively, which improves the sample quality and convergence speed. In DCGAN, the generator primarily processes one-dimensional noise data through an upsampling operation. Deconvolution is a crucial step in the

upsampling process. Deconvolution involves transposing and padding the kernel matrix before performing subsequent operations. Although deconvolution enlarges the size of the feature maps and changes the number of channels, its major drawback is that it often produces clear checkerboard artifacts in the feature maps. If multiple layers of transpose convolution are stacked, the checkerboard artifacts become more prominent, resulting in generated images with "ghosting" artifacts and reduced clarity. Therefore, instead of using multiple deconvolutions, RC-GAN employs upsampling and convolution blocks (UCBlock) in the generator[16], with bilinear interpolation used in the upsampling process. The improved generator with UCBlock first takes a 100-dimensional noise vector as input and passes it through fully connected and reshaping layers to obtain a $4 \times 4 \times 1024$ feature matrix. Next, it goes through four UC modules to obtain a $64 \times 64 \times 3$ feature matrix. Finally, a Tanh activation function is applied to generate the image. The discriminator, essentially a classifier, outputs a one-dimensional probability for the image. Hence, this study replaces the original discriminator structure with the structure of AlexNet, as the overlapping pooling used in AlexNet better preserves the original information, thereby improving the performance of the discriminator and the quality of generated images. The overall RC-GAN structure is illustrated in Fig.3.

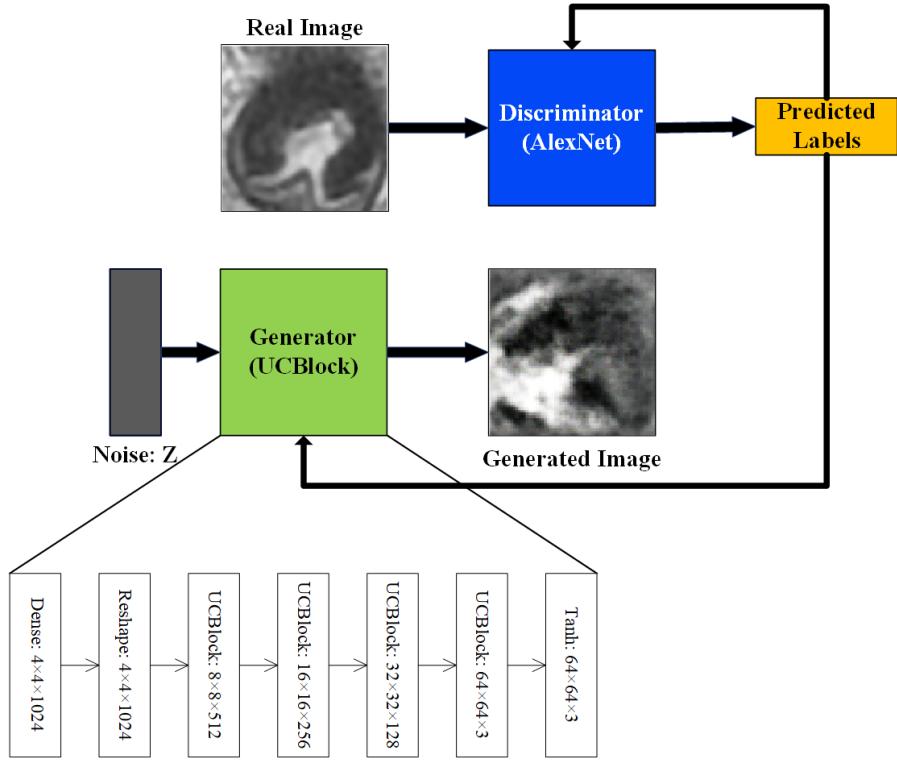


Fig. 3: Structure of RC-GAN Data Augmentation Network

2.4 Data Preprocessing for Improved 3DCNN Prediction Network

Due to the fact that the output of RC-GAN is a 2D image of rectal cancer tumors, it is necessary to convert the generated and original images from 2D to 3D. The process of converting 2D images to 3D images is shown in Fig.4, with the specific steps as follows:

(1) The generated and the original rectal cancer MRI data are classified by case. There are dozens of MRI images in a case, and only a few of them contain the tumor slice, so the slices containing the tumor need to be taken out. Considering that adjacent slices may also contain relevant data information, 2 slices before and after the tumor-containing slice are also taken respectively.

(2) The number of slices included in each case is uniformly standardized. Since the number of slices containing the tumor varies from case to case, and the input of the network needs to maintain consistency, the number of slices in each case, i.e., the depth dimension, needs to be standardized.

(3) Import the multiple slices of each case, create a four-dimensional matrix containing the attributes of image width, height, depth (number of slices), and channel number, output each image data as a four-dimensional matrix.

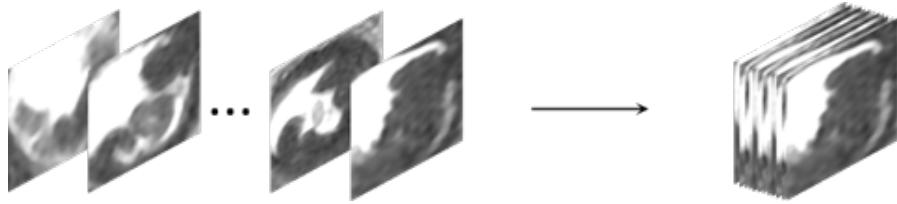


Fig. 4: Converting 2D MRI Images to 3D MRI Images

2.5 Structure of Improved 3DCNN Prediction Network

This article adopts Spatial Attention (SA), Channel Attention (CA), Squeeze-and-Excitation Channel Attention Module (SE), and Convolutional Block Attention Module (CBAM) as attention modules for the network, respectively. Since the prediction network has a 3D convolutional structure, these attention mechanism modules have been modified based on 3D. The 3DCNN designed in this paper mainly consists of four convolutional blocks, each of which includes a 3D convolutional layer, a batch normalization layer, an activation layer and a max pooling layer. Except for the first convolutional block, the other three convolutional blocks also include four optional attention mechanism modules.

In between each convolutional layer and max pooling layer, batch normalization and activation layer are usually used, followed by a max pooling operation. Dropout layers are added in the first two Fully Connected (FC) layers to discard some useless features and prevent overfitting. In addition, to provide more attention to the relevant features in the input data, different attention modules are inserted to form different networks. The structure of improved 3DCNN prediction network is shown in Fig.5.

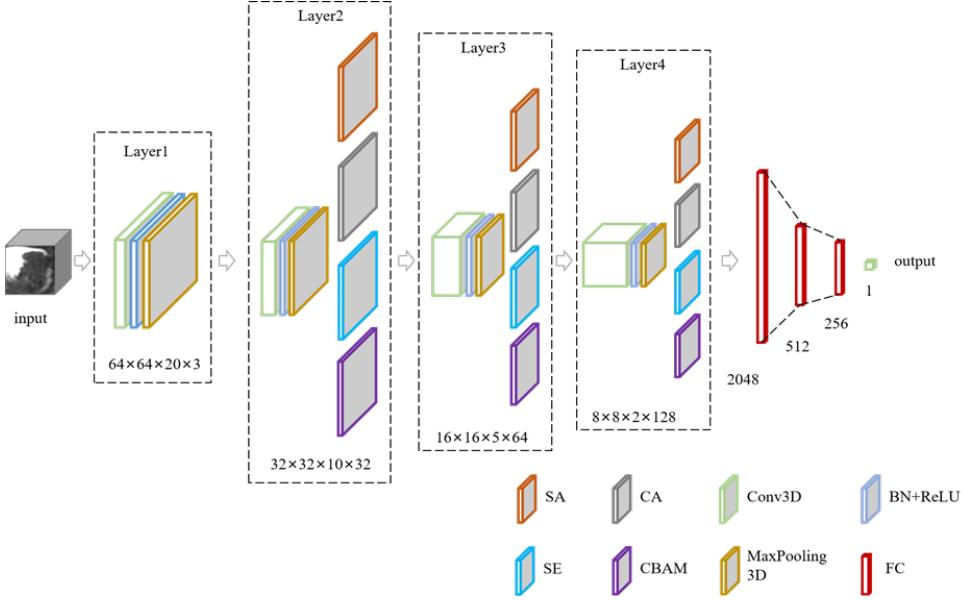


Fig. 5: Structure of Improved 3DCNN Prediction Network

3 Results

3.1 Comparison Experiments

This paper presents three sets of comparative experiments. In the first set, our improved 3DCNN (3DCNN+CBAM) is compared with Control Group 1[17], Control Group 2[18], and Control Group 3[19], both before and after data augmentation by RC-GAN. The second set compares the performance of the model with different attention modules. The third set is a comparison experiment of the improved 3DCNN network on complete 3D MRI images and region-of-interest (ROI) 3D MRI images.

3.1.1 Performance Comparison between Improved 3DCNN Prediction Networks and Other Networks

To verify the influence of each model using the original data and the generated data by RC-GAN, two different experiments were designed to evaluate the model's performance on the original real dataset and the augmented dataset, respectively.

(1) Performance Comparison before Data Augmentation by RC-GAN Fig.6 provide a performance comparison between the improved 3DCNN (3DCNN+CBAM) and other networks, prior to data augmentation. The results show that the improved 3DCNN (3DCNN+CBAM) exhibits higher accuracy and specificity than the control groups, with only slightly lower sensitivity than control group 1. Overall, the performance of the improved 3DCNN (3DCNN+CBAM) is superior to that of other networks evaluated in the experiment. This suggests that the dataset itself is not particularly large in size and only measures 64x64x20, making it unnecessary to employ a complex

network architecture to achieve optimal results. Additionally, the CNN convolutional blocks constructed in this study incorporate dropout layers, which effectively mitigate the problem of overfitting. Furthermore, the inclusion of attention modules further enhances the prediction performance.

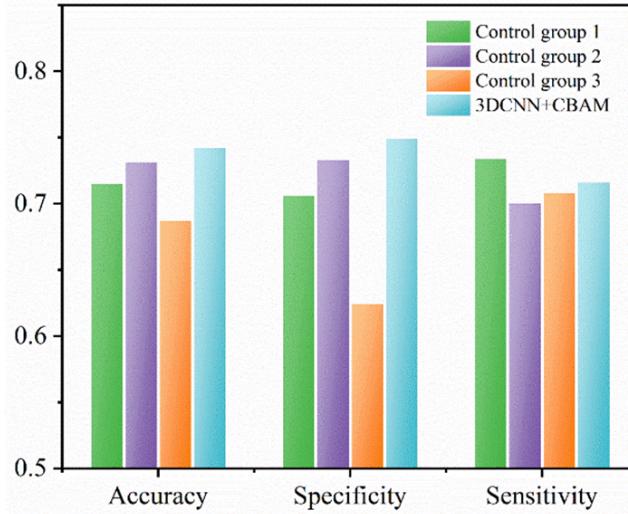


Fig. 6: Performance Comparison of Different Networks before Data Augmentation

(2) Performance Comparison After Data Augmentation by RC-GAN

To test the efficacy of the data augmentation approach by RC-GAN, a comparison experiment was conducted on models trained with the augmented dataset. Prior to the experiment, data augmentation was conducted utilizing RC-GAN. Fig.7 shows the comparison between generated rectal cancer tumors by RC-GAN and real rectal cancer tumors. It can be observed that as the training epochs increase, the generated images become progressively closer to real data, starting from noise. The RC-GAN architecture also addresses the issue of "checkerboard" in the deconvolution process. It can be seen that in the first half between the 1st epoch and 100th epoch, the generated data still exhibit some "checkerboard artifacts." However, in the subsequent epochs, the influence of "checkerboard" on the generated data diminishes. Subsequent to the aforementioned data augmentation, the same set of experiments was conducted with equivalent configurations and parameters, except the dataset. Fig.8 provides a comparative analysis of the performance of the improved 3DCNN and other networks following data augmentation.

From the analysis of Fig.8, it can be observed that after data augmentation, the overall performance of all models has improved. Notably, the proposed 3D convolutional structure utilizing the CBAM achieved an accuracy, specificity, and sensitivity of 0.778, 0.796, and 0.754, respectively, representing a significant increase of 4.9%, 6.3%, and 5.3% in comparison to performance prior to data augmentation. Moreover,

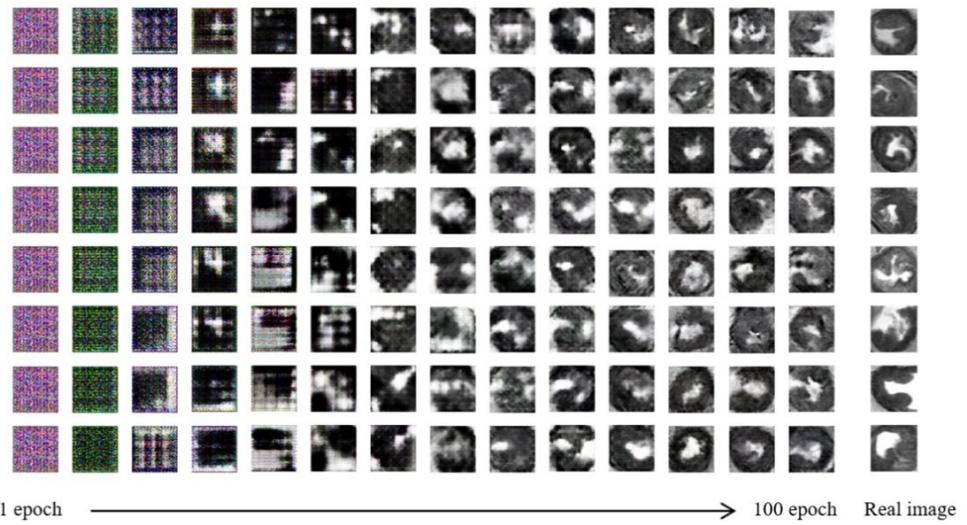


Fig. 7: Generated Rectal Cancer Tumor and Real Rectal Cancer Tumor

the control group demonstrated a corresponding improvement in performance, further illustrating the effectiveness of the data augmentation approach proposed in this paper for addressing the challenge of small data volume and improving the prediction performance of complete pathological response in rectal cancer.

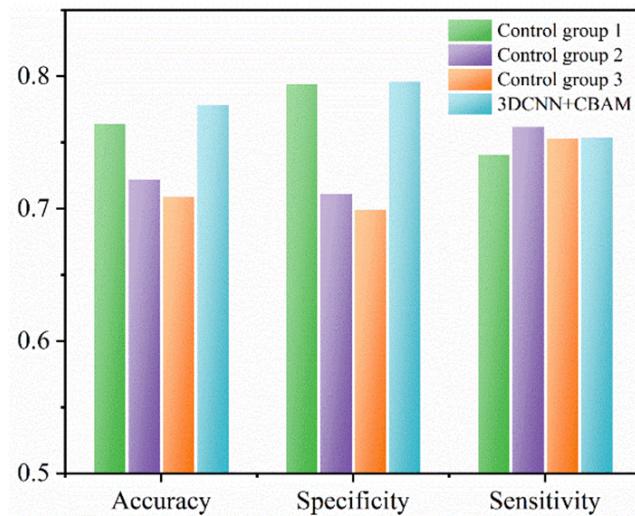


Fig. 8: Performance Comparison of Different Networks after Data Augmentation

3.1.2 Performance Comparison of Networks with Different Attention Modules

Following validation of the RC-GAN generated data, this paper proceeds to evaluate the performance of different attention mechanism modules. The performance comparison is shown in Fig.9. The results demonstrate that utilizing both SA and CA modules leads to an improvement in network performance, with CA proving particularly effective in enhancing prediction accuracy. Notably, the CBAM effectively combines both SA and CA to achieve relatively high accuracy, specificity, and sensitivity, which respectively recorded 0.778, 0.796, and 0.754.

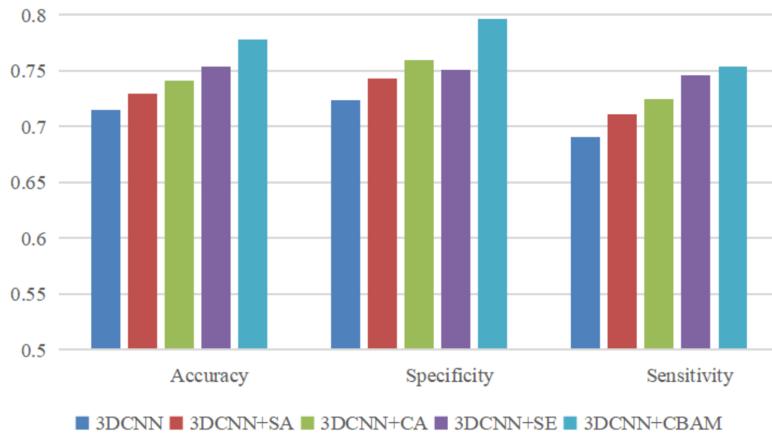


Fig. 9: Performance comparison of different attention mechanisms for 3DCNN

3.1.3 Comparison Experiment on Complete 3D MRI Images and ROI 3D MRI Images

The T2W MRI images of patients obtained by MRI scanners before treatment contain both complete organ information and some irrelevant data. To explore the influence of the complete MRI images compared to the 3D data composed of ROI (tumor) on the prediction performance, this article utilized the improved 3DCNN (3DCNN+CBAM), to make the comparison between the two. The comparison of different performance metrics between the two datasets is provided in Table1 and Fig.10.

As presented in Table 1, the prediction performance significantly improved after removing the background and other irrelevant factors from the complete MRI images. Specifically, the accuracy, specificity, and sensitivity increased by 9.3%, 13.3%, and 4.4%, respectively. Therefore, using the tumor region images ROI MRI images as the dataset was found to be more effective in predicting pCR.

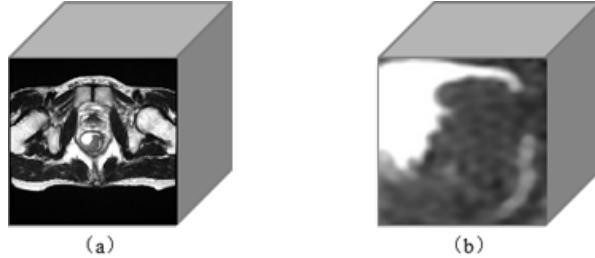


Fig. 10: (a) Complete 3D MRI Image (b) ROI 3D MRI Image

Table 1: Performance Comparison of Different Input Data

Input data	Accuracy	Specificity	Sensitivity
Full 3D MRI Images	0.679	0.661	0.686
ROI 3D MRI Images	0.742	0.749	0.716

3.2 Ablation Study

Given the scarcity of existing research on predicting pCR based on 3D MRI images in rectal cancer and the limited size of related datasets, we propose the RCpCR-Net framework to address these issues. To tackle the problem of small dataset sizes, we introduce RC-GAN, which generates synthetic data to augment the dataset. In the improved 3DCNN, we integrate the CBAM module to leverage spatial and channel information simultaneously, achieving greater improvements in prediction accuracy compared to other attention mechanisms. To validate the effectiveness of the RCpCR-Net framework, we designed and conducted ablation experiments. Table 2 demonstrates the performance comparison of the baseline network and baseline network with our proposed improvement.

Table 2: prediction performance of 3DCNN at different stages

Model name	Accuracy	Specificity	Sensitivity
BaseNet	0.715	0.724	0.691
BaseNet+CBAM	0.742	0.749	0.716
BaseNet+RC-GAN	0.734	0.741	0.729
BaseNet+RC-GAN+CBAM(RCpCR-Net)	0.778	0.796	0.754

Fig.11 provides a more intuitive depiction of the changes in accuracy, specificity, and sensitivity of the baseline network after adding various modules. These results, along with the data in Table 2, indicate that the proposed 3D prediction model in this study demonstrates significant enhancement in prediction performance after incorporating data augmentation by RC-GAN and CBAM. Specifically, the implementation of the CBAM module led to increases of 3.8%, 3.5%, and 3.6% in accuracy, specificity, and sensitivity, respectively. An improvement in these indicators was also observed

after data augmentation, with respective increases of 2.7%, 2.3%, and 5.5%. Nevertheless, the most significant improvements followed the use of both data augmentation by RC-GAN and the CBAM module, which resulted in increases of 8.8%, 9.9%, and 9.1% in accuracy, specificity, and sensitivity, respectively.

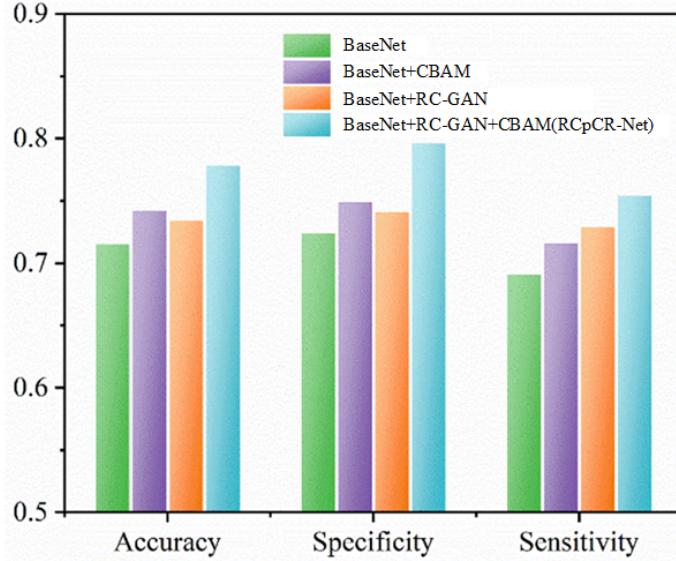


Fig. 11: Histogram of the performance of the proposed network at different stages

4 Discussion

New adjuvant chemotherapy combined with surgical resection is a conventional treatment strategy for locally advanced rectal cancer and high-risk tumors. After completion of chemoradiotherapy, surgical resection of the tumor is typically performed. However, if it is possible to predict prior to chemoradiotherapy whether a patient will achieve pCR, it could mean fewer surgeries and less chemotherapy for the patient. Therefore, predicting treatment outcomes, especially pCR, in rectal cancer patients is crucial for the development and optimization of new adjuvant chemoradiotherapy protocols and guiding subsequent treatments. Currently, most rectal cancer studies use proprietary datasets that require manual annotation by doctors and are often relatively small in size. Moreover, there is limited utilization of 3D data in studies related to pCR in rectal cancer, with most studies using 2D slices for analysis. To address the issues commonly seen in rectal cancer studies, such as the use of hospital-specific datasets with small sample sizes, this paper proposes RC-GAN for data augmentation to increase the effective amount of data and enhance prediction performance. Additionally, to overcome the loss of inter-slice continuity information in 2D image-based prediction studies, this paper utilizes 3D rectal cancer data and improved 3DCNN for

predicting pCR. RCpCR-Net is a framework combining the above methods. Although RCpCR-Net improves the accuracy, specificity, and sensitivity of the predictions with significant enhancements, there are still limitations that need to be addressed. Further research is required to overcome the challenges of small datasets and hardware constraints. One possible approach is to fuse data from different modalities, such as CT scans, pathological slides, or clinical information, to achieve better prediction performance.

5 Conclusion

This paper addresses the existing issues in research on predicting pCR in 3D rectal cancer MRI images, namely the scarcity of studies and the inadequate volume of datasets. To tackle these challenges, we propose our framework, RCpCR-Net. It involves using our proposed network, RC-GAN, for data augmentation. It also employs different attention mechanism modules to optimize the 3DCNN network. Specifically, RC-GAN replaces the original deconvolution blocks used in the DCGAN generator with upsampling and convolution blocks (UCBlock), which partially solves the "checkerboard" artifact problem commonly encountered in deconvolution processes. Additionally, we enhance the quality of generated images by adopting the AlexNet architecture for the discriminator of RC-GAN. In the improved 3DCNN, the CBAM module is added to extract features using both spatial and channel-wise information, demonstrating more significant effectiveness compared to other attention mechanism modules. In the final experimental section, we compare RCpCR-Net with other networks and validate the feasibility of the added modules through ablation experiments.

Declarations

Ethics approval and consent to participate

The study protocol was reviewed and approved by the Institutional Review Board of the Sixth Affiliated Hospital of Sun Yat-sen University (NO. 2021ZSLYEC-366). Informed consent to participate was obtained from all of the participants in the study. This study complied with the ethical requirements outlined in the Helsinki Declaration.

Consent for publication

Consent for publication is not applicable since our manuscript does not contain data from any individual person.

Availability of data and materials

The dataset of the patients used and analyzed in this study are available from Dr. Juan Li on reasonable request. The trained RCpCR-Net and relative codes developed in this study is not applicable.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the Guangzhou Science and Technology Project under Grant No. 2023A04J2240 and National Key Clinical Discipline.

Authors' contributions

Conception and design: LL, MW and BX. Financial support: JL. Administrative support: LL. Provision of study material or patients: JL and DN. Collection and assembly of data: LL, MW, ZW, TC, CW, BX, JL and DN. Data analysis and interpretation: LL, MW, ZW, TC, CW, BX, JL and DN. Manuscript writing: All authors. Final approval of manuscript: All authors. Accountable for all aspects of the work: All authors.

Acknowledgements

The authors would like to express their sincere appreciation to all the challenge organizers and dataset providers for making the public dataset available to the community.

References

- [1] Zheng, R., Zhang, S., Zeng, H., Wang, S., Sun, K., Chen, R., Li, L., Wei, W., He, J.: Cancer incidence and mortality in china, 2016. *Journal of the National Cancer Center* **2**(1), 1–9 (2022)
- [2] Xie, Y., Shi, L., He, X., Luo, Y.: Gastrointestinal cancers in china, the usa, and europe. *Gastroenterology report* **9**(2), 91–104 (2021)
- [3] Rabi, I.I., Zacharias, J.R., Millman, S., Kusch, P.: A new method of measuring nuclear magnetic moment. *Physical review* **53**(4), 318 (1938)
- [4] Maas, M., Nelemans, P.J., Valentini, V., Das, P., Rödel, C., Kuo, L.-J., Calvo, F.A., García-Aguilar, J., Glynne-Jones, R., Haustermans, K., *et al.*: Long-term outcome in patients with a pathological complete response after chemoradiation for rectal cancer: a pooled analysis of individual patient data. *The lancet oncology* **11**(9), 835–844 (2010)
- [5] Byrd, D.R., Brierley, J.D., Baker, T.P., Sullivan, D.C., Gress, D.M.: Current and future cancer staging after neoadjuvant treatment for solid tumors. *CA: a cancer journal for clinicians* **71**(2), 140–148 (2021)
- [6] Sluis, F.J., Couwenberg, A.M., Bock, G.H., Intven, M.P., Reerink, O., Leeuwen, B.L., Westreenen, H.L.: Population-based study of morbidity risk associated with

- pathological complete response after chemoradiotherapy for rectal cancer. *Journal of British Surgery* **107**(1), 131–139 (2020)
- [7] Menzies, A.M., Amaria, R.N., Rozeman, E.A., Huang, A.C., Tetzlaff, M.T., Wiel, B.A., Lo, S., Tarhini, A.A., Burton, E.M., Pennington, T.E., *et al.*: Pathological response and survival with neoadjuvant therapy in melanoma: a pooled analysis from the international neoadjuvant melanoma consortium (inmc). *Nature medicine* **27**(2), 301–309 (2021)
 - [8] Dossa, F., Chesney, T.R., Acuna, S.A., Baxter, N.N.: A watch-and-wait approach for locally advanced rectal cancer after a clinical complete response following neoadjuvant chemoradiation: a systematic review and meta-analysis. *The lancet Gastroenterology & hepatology* **2**(7), 501–513 (2017)
 - [9] Kather, J.N., Pearson, A.T., Halama, N., Jäger, D., Krause, J., Loosen, S.H., Marx, A., Boor, P., Tacke, F., Neumann, U.P., *et al.*: Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. *Nature medicine* **25**(7), 1054–1056 (2019)
 - [10] Hu, Y., Li, J., Zhuang, Z., Xu, B., Wang, D., Yu, H., Li, L.: Automatic treatment outcome prediction with deepinteg based on multimodal radiological images in rectal cancer. *Heliyon* **9**(2) (2023)
 - [11] Li, L., Xu, B., Zhuang, Z., Li, J., Hu, Y., Yang, H., Wang, X., Lin, J., Zhou, R., Chen, W., *et al.*: Accurate tumor segmentation and treatment outcome prediction with deeptop. *Radiotherapy and Oncology* **183**, 109550 (2023)
 - [12] Soomro, M.H., Coppotelli, M., Conforto, S., Schmid, M., Giunta, G., Del Secco, L., Neri, E., Caruso, D., Rengo, M., Laghi, A.: Automated segmentation of colorectal tumor in 3d mri using 3d multiscale densely connected convolutional neural network. *Journal of healthcare engineering* **2019**(1), 1075434 (2019)
 - [13] Deng, Y., Chi, P., Lan, P., Wang, L., Chen, W., Cui, L., Chen, D., Cao, J., Wei, H., Peng, X., *et al.*: Neoadjuvant modified folfox6 with or without radiation versus fluorouracil plus radiation for locally advanced rectal cancer: final results of the chinese fowarc trial. *Journal of Clinical Oncology* **37**(34), 3223–3233 (2019)
 - [14] Xie, Y., Lin, J., Wang, X., Wang, P., Zhuang, Z., Zou, Q., Cai, D., Huang, Z., Bai, L., Tang, G., *et al.*: The addition of preoperative radiation is insufficient for lateral pelvic control in a subgroup of patients with low locally advanced rectal cancer: a post hoc study of a randomized controlled trial. *Diseases of the Colon & Rectum* **64**(11), 1321–1330 (2021)
 - [15] Trakarnsanga, A., Gönen, M., Shia, J., Nash, G.M., Temple, L.K., Guillem, J.G., Paty, P.B., Goodman, K.A., Wu, A., Gollub, M., *et al.*: Comparison of tumor regression grade systems for locally advanced rectal cancer after multimodality treatment. *Journal of the National Cancer Institute* **106**(10), 248 (2014)

- [16] Odena, A., Dumoulin, V., Olah, C.: Deconvolution and checkerboard artifacts. *Distill* **1**(10), 3 (2016)
- [17] Tran, D., Bourdev, L., Fergus, R., Torresani, L., Paluri, M.: Learning spatiotemporal features with 3d convolutional networks. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 4489–4497 (2015)
- [18] Hara, K., Kataoka, H., Satoh, Y.: Learning spatio-temporal features with 3d residual networks for action recognition. In: Proceedings of the IEEE International Conference on Computer Vision Workshops, pp. 3154–3160 (2017)
- [19] Carreira, J., Zisserman, A.: Quo vadis, action recognition? a new model and the kinetics dataset. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 6299–6308 (2017)