scientific data

Check for updates

OPEN

DATA DESCRIPTOR

A comprehensive dataset of magnetic resonance enterography images with intestinal segment annotations

Zhangnan Zhong[1,2,4], Li Huang[3,4], Shi-Ting Feng[3], Haiwei Lin[1], Xinyue Wang[3], Baolan Lu[3], Kangyang Cao[1], Xuehua Li[3] ✉ & Bingsheng Huang[1] ✉

Inflammatory bowel disease (IBD) is a recurrent bowel disease that usually requires magnetic resonance enterography (MRE) for diagnosis and monitoring. However, recognition of bowel segments from MRE images by a radiologist is challenging and time-consuming. Deep learning-based medical image segmentation has shown the potential to reduce manual effort and provide automated tools to assist in disease management; however, it requires a large-scale fine-annotated dataset for training. To address this gap, we collected MRE data, including half-Fourier acquisition single-shot turbo spin-echo(HASTE) sequences with coronal orientation, from 114 patients with IBD, who received 1600–2000 mL of 2.5% mannitol. The bowel images per patient were contoured and annotated into ten segments (stomach, duodenum, small intestine, appendix, cecum, ascending colon, transverse colon, descending colon, sigmoid colon, and rectum), with fine pixel-level annotations labeled by experienced radiologists. Furthermore, we validated the efficiency of several state-of-the-art segmentation methods using this dataset. This study established a high-quality, publicly available whole-bowel segment MR dataset with benchmark results and laid the groundwork for AI research on IBD.

Background & Summary

Inflammatory bowel disease (IBD) is one of the most challenging diseases of the 21st century, affecting >10 million people worldwide. Inflammatory bowel disease (IBD), encompassing Crohn's disease (CD) and Ulcerative Colitis (UC), is a common complex digestive system disease with relapsing and remitting conditions that can be challenging to diagnose and manage. For instance, CD can affect the entire gastrointestinal tract, with a predilection for the terminal ileum and ascending colon, and is characterized by segmental and transmural granulomatous inflammation. CD diagnosis typically relies on a comprehensive evaluation of clinical symptoms, laboratory tests, endoscopic examinations, radiological imaging, and pathological tissue examination[1].

Cross-sectional imaging has long complemented the endoscopic assessment of IBD. The patient underwent multiple follow-up examinations to monitor the condition and treatment effectiveness. However, endoscopic assessments are often burdensome for the patient[2]. Cross-sectional enterography techniques serve as a complementary tool to ileocolonoscopy, enabling the visualization of intramural or proximal small bowel inflammation in approximately 50% of patients with CD whose endoscopic examinations appear normal[3,4]. Thus, cross-sectional enterography plays a vital role in diagnosis and monitoring the disease course. Computed tomography enterography (CTE) and magnetic resonance enterography (MRE) have emerged as the most effective for imaging the small bowel to diagnose small intestinal CD in terms of detecting the extent of lesions and nature of luminal strictures, which helps assess disease distribution, staging, and detecting extraintestinal complications. Cross-sectional enterography can also be used to monitor treatment response, lesion healing, and disease progression[5]. In particular, MRE, with its excellent soft-tissue resolution and multiparametric imaging capabilities, has shown better performance[2].

[1]Medical AI Lab, School of Biomedical Engineering, Shenzhen University Medical School, Shenzhen University, Shenzhen, 518060, China. [2]Neusoft Institute Guangdong, Foshan, 528225, China. [3]Department of Radiology, The First Affiliated Hospital, Sun Yat-Sen University, Guangzhou, 510080, China. [4]These authors contributed equally: Zhangnan Zhong, Li Huang. ✉e-mail: lxueh@mail.sysu.edu.cn; huangb@szu.edu.cn

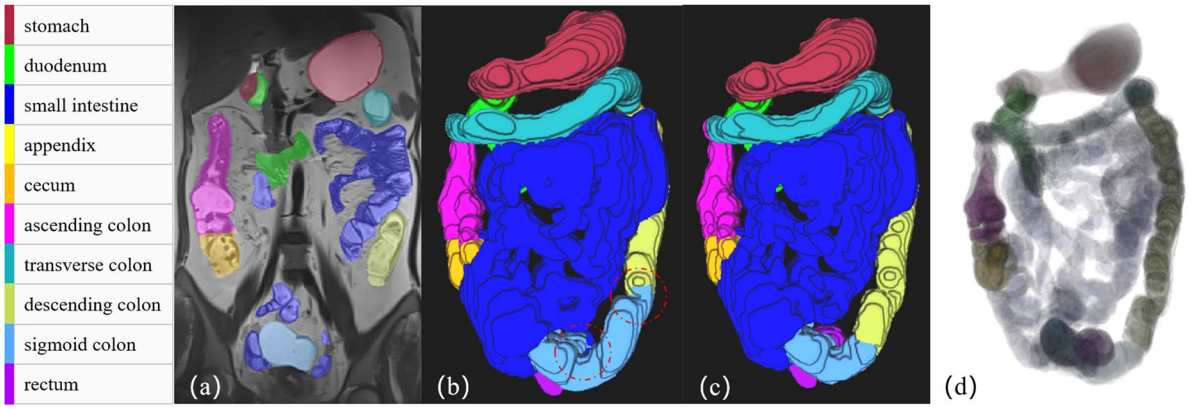

Fig. 1 An example of 10 gastrointestinal tract segments in an MRE. (**a**) T2WI MRE, (**b**) rough intestinal segmentation (primary prediction), (**c**) fine intestinal segmentation (physician-reviewed), (**d**) 3D visualization. T2WI, T2 weighted image; MRE, magnetic resonance enterography; 3D, three-dimensional.

T2-weighted imaging (T2WI) is a crucial sequence for MRE in IBD, as it not only provides information about the anatomical structure of the bowels but also detects imaging signs such as bowel wall thickening, intramural edema, and bowel lumen strictures when the bowels are in good filling status. Furthermore, coronal T2WI allows for scanning of the gastrointestinal tract across the entire abdomen, allowing a more intuitive view of the intestinal segmentation and path.

However, the complexity of IBD, variable radiological manifestations of intestinal disease, differences in scanning techniques, and uneven levels of radiologists' understanding of IBD imaging features combine to create challenges. The recognition of bowel segments from MRE images by radiologists can be challenging and time-consuming because of unclear boundaries, shape, size, and appearance variations, as well as uneven filling within the bowel. Consequently, accurate and standardized bowel segmentation is essential for the medical image analysis of IBD.

Deep-learning-based medical image segmentation has shown the potential to reduce manual effort and provide automated tools to assist in disease management. In recent years, a proliferation of deep learning-based methods has been proposed for the accurate and expedited segmentation of organs from abdominal CT volumes. However, the evaluation of these methods typically focuses on a limited number of organs. Although existing research have yielded commendable results in segmenting certain gastrointestinal tract segments, such as the liver, spleen, and kidneys, research specifically addressing intestinal segmentation remain scarce[6,7]. This issue often arises because the current studies tend to treat the entire colon as a single unit. However, this approach does not reflect real clinical scenarios because each segment of the colon has distinct functionalities (Fig. 1). By segmenting the colon into individual segments and simultaneously determining the position of the segment during the segmentation process, clinicians can conduct differential analyses of different segments, significantly enhancing the clinical utility and value of colon segmentation. The research gap in segmenting different intestinal segments stems primarily from the absence of a publicly available, large-scale, accurately annotated, and clinically relevant dataset for whole intestinal segmentation. Therefore, to advance intestinal segmentation research, it is crucial to develop high-quality task-specific datasets and establish benchmarks for this segmentation task.

Furthermore, the complex structure and difficulty in delineating the entire intestine pose significant challenges. Although the academic community has access to CT data and segmentation methods for intestinal sections, there are no publicly available MRI datasets for full intestinal segmentation. This gap has resulted in the lack of research in this field.

In this study, we curated a real-world clinical MRI dataset and annotated the intestinal regions for segmentation. All scans in our dataset were meticulously hand-labelled, encompassing ten segments of the gastrointestinal tract. Collecting real-world clinical data is challenging and time consuming, primarily because of privacy and ethical considerations.

Additionally, we explored both fully supervised segmentation methods and annotation-efficient strategies to assess the benchmark performance on our bowel dataset. Specifically, we evaluated several cutting-edge medical segmentation models, including nnU-Net[8], ResUNet[9], UCTransNet[10], and CoTr[11].

Such a dataset would possess significant research value and could be utilized to evaluate and enhance existing whole-intestine segmentation methods, thereby establishing a benchmark for organ segmentation problems. Furthermore, this dataset could serve as an effective testing platform for the development of advanced whole-intestine segmentation algorithms, thereby making a substantial contribution to research in this field.

In summary, our work provides the following key contributions:

1. We curated a unique, clinically focused dataset for comprehensive intestinal segmentation, comprising MRE data from 114 patients. This dataset offers a more detailed segmentation (10 intestinal segments) than previous studies.

Dataset		Total (n = 114)
Demographics	**Age in years**	33.7 (12, 74)
	Sex	
	Female	33 (28.95%)
	Male	81 (71.05%)
MRI characteristics	**MRI Magnetic Field**	
	3.0 T	114 (100%)
	Scanner	
	Siemens Prisma	60 (52.63%)
	Siemens Vida	54 (47.37%)
	T2WI image voxel size	
	1.4 × 1.4 × 4.0 mm	60 (52.63%)
	1.2 × 1.2 × 4.0 mm	54 (47.37%)
Imaging characteristics	**Location of the most severe bowel lesion**	
	Without abnormal findings	9 (7.89%)
	Stomach + duodenum	2 (1.75%)
	Proximal ileum + jejunum	8 (7.02%)
	Terminal ileum	43 (37.72%)
	Ileocecum	17 (14.91%)
	Colon	35 (30.70%)

Table 1. Information on demographics and MRI and imaging characteristics. Continuous variables were presented as means and ranges, and categorical variables were presented as numbers and percentages in each group.

2. We established a new benchmark for whole intestinal segmentation, which includes (1) evaluating the effectiveness of currently available fully supervised segmentation methods and (2) quantifying the difference in segmentation capability between deep learning models and radiologists.

Methods

Cohort. Clinical and magnetic resonance enterography (MRE) data were retrospectively obtained from 114 patients with IBD admitted to the Inflammatory Bowel Disease Center at the First Affiliated Hospital of Sun Yat-sen University between December 2019 and May 2022. This retrospective study was approved by the Institutional Ethics Review Board (approval number: 2022 [024]), which waived the requirement for informed consent.

The inclusion criteria were as follows: a) patients with a confirmed diagnosis of CD based on standard clinical, endoscopic, imaging, and histological criteria and b) patients over 12 years of age who had completed an MRE examination. The exclusion criteria were as follows: a) patients with incomplete clinical data or another concurrent intestinal disease, b) cases where MRI quality was insufficient for accurate observation, c) patients whose MR images did not include complete bowel segments, and d) patients who had undergone intestinal resection.

The population, images, and lesion profiles are shown in Table 1. The mean patient age was 33.67 (range, 12–74) years. There were more males (71.05%) than females. The location of the most severe bowel lesion was determined based on all MRE findings and was reported by experienced radiologists (L.H. and X.L., both with more than 10 years of experience).

MRIs. In this study, we retrospectively collected T2-weighted coronal magnetic resonance enterography (MRE) data from 114 patients with CD. Following the approach described in previous studies[12–14], the patients underwent bowel preparation. Subsequently, they received 1600–2000 mL of 2.5% mannitol solution one hour before the MRI to fill the bowels. Additionally, 10 mg raceanisodamine hydrochloride (Minsheng Pharmaceutical Group, Hangzhou, China) was administered intramuscularly to the buttocks 10 min before scanning to inhibit gastrointestinal peristalsis. MR was performed using a 3.0-T MRI scanner (Magnetom Vida or Prisma; Siemens Medical Solutions, Erlangen, Germany) with a high-performance gradient system (maximum gradient = 80 mT/m, maximum slew rate = 200 mT/m) and two 18-channel phased-array coils, ensuring the quality of the abdomen MR image and a high signal-to-noise ratio. The details of the MRI acquisition parameters are presented in Table 2.

Intestines annotation. The intestinal images per patient were segmented into ten segments (0 to background, 1 to the stomach, 2 to the duodenum, 3 to the small intestine, 4 to the appendix, 5 to the cecum, 6 to the ascending colon, 7 to the transverse colon, 8 to the descending colon, 9 to the sigmoid colon, and 10 to the rectum), with fine pixel-level annotations performed by two experienced radiologists (X.W. with 3 years of experience and B.L. with 6 years of experience). All labels were delineated in the T2-weighted MR images using ITK-SNAP[15] slice-by-slice in the coronal view using a pre-trained model for raw segmentation, which was subsequently refined by radiologists. The data annotation process is illustrated in Fig. 2. Subsequently, an abdominal imaging expert (L.H., with > 10 years of experience) carefully reviewed these annotations and resolved any

Parameter	T2-weighted HASTE (Siemens Prisma)	T2-weighted HASTE (Siemens Vida)
Orientation	Coronal	Coronal
Acquisition matrix	320 × 224	384 × 384
Field of view (mm × mm)	450 × 422	450 × 422
Section thickness (mm)	4	4
Echo time (msec)	83	84
Repetition time (msec)	800	800
No. of sections	27	29
Respiratory control	Breath hold	Breath hold
Acquisition time (sec)	24	23

Table 2. MR imaging protocols. Note. HASTE, half-Fourier acquisition single-shot turbo spin-echo; MR, magnetic resonance.

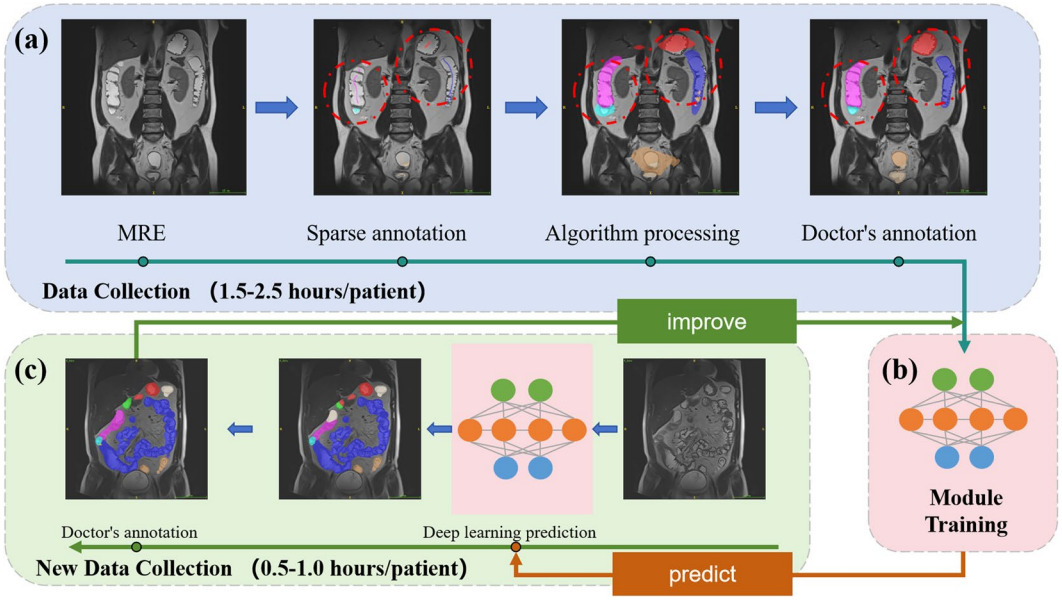

Fig. 2 Data collection procedure. (**a**) Initial Annotation: MRE data undergoes sparse annotations by doctors, algorithm processing for coarse segmentation, and final refinement by doctors (1.5–2.5 h/patient). (**b**) Model Training: The refined annotations are used to train a deep learning segmentation model. (**c**) Iterative Refinement: The model predicts segments on new data, which are corrected by doctors to further improve the model (0.5–1.0 h/patient).

disagreements through discussion, resulting in consensus annotations that ensured annotation quality. Finally, these consensus labels were released and used for subsequent model building.

The original annotated method is labor-intensive, taking 1.5–2.5 h per patient. To streamline this process, we first trained a deep-learning model using an initial batch of fully annotated data. This model was subsequently used to predict annotations for new MR data, resulting in more precise labels with minimal expert revisions and thus reducing the annotation time to 0.5–1.5 h.

Data Records

The dataset was hosted by Zenodo (https://zenodo.org/records/13839321)[16]. We made all the datasets described earlier available. They comprised 114 cases, each annotated with ten distinct labels corresponding to the abdominal digestive tract.

The data and corresponding label files are systematically named as "xx_data.nii.gz" and "xx_label.nii.gz." In the label files, key anatomical regions—the stomach, duodenum, small intestine, appendix, cecum, ascending colon, transverse colon, descending colon, sigmoid colon, and rectum— were numerically labeled from 1 to 10 in sequential order.

Technical Validation

Experiment setup and evaluation metrics. In this study, all methods were implemented using the PyTorch framework on GPUs, including NVIDIA GTX1080TI, NVIDIA TITAN RTX, and GeForce GTX 1080 Ti. We selected nnUNet as the baseline for a fair comparison. nnUNet is a self-configuring segmentation framework that requires no manual intervention for data processing, training planning (network architecture, parameter settings, and so on), or postprocessing. It encompasses both 3D and 2D methods. Although nnUNet initially

method	nnUNet (2D)	nnUNetV2 (2D)	ResUNet (2D)	UCTransNet (2D)	nnUNet (3D)	nnUNetV2 (3D)	CoTr (3D)
Stomach	0.934 ± 0.054	0.939 ± 0.046	0.801 ± 0.101	0.908 ± 0.069	0.956 ± 0.048	**0.962 ± 0.041**	0.938 ± 0.067
Duodenum	0.798 ± 0.080	0.805 ± 0.068	0.603 ± 0.105	0.739 ± 0.093	0.858 ± 0.069	**0.865 ± 0.062**	0.830 ± 0.068
Small Intestine	0.889 ± 0.033	0.892 ± 0.035	0.743 ± 0.064	0.853 ± 0.042	0.926 ± 0.031	**0.931 ± 0.026**	0.906 ± 0.035
Appendix	0.125 ± 0.233	0.177 ± 0.278	0.209 ± 0.388	0.126 ± 0.239	0.151 ± 0.253	**0.479 ± 0.469**	0.091 ± 0.188
Cecum	0.425 ± 0.263	0.452 ± 0.254	0.127 ± 0.173	0.370 ± 0.222	**0.602 ± 0.269**	0.600 ± 0.280	0.410 ± 0.277
Ascending Colon	0.766 ± 0.157	0.759 ± 0.164	0.609 ± 0.181	0.699 ± 0.175	0.823 ± 0.152	**0.834 ± 0.152**	0.772 ± 0.186
Transverse Colon	0.757 ± 0.137	0.769 ± 0.128	0.523 ± 0.161	0.687 ± 0.151	0.798 ± 0.166	**0.822 ± 0.150**	0.765 ± 0.180
Descending Colon	0.725 ± 0.138	0.719 ± 0.153	0.559 ± 0.180	0.660 ± 0.174	0.756 ± 0.173	**0.780 ± 0.146**	0.731 ± 0.164
Sigmoid Colon	0.654 ± 0.165	0.673 ± 0.157	0.481 ± 0.161	0.586 ± 0.172	0.731 ± 0.169	**0.754 ± 0.165**	0.706 ± 0.175
Rectum	0.817 ± 0.155	0.820 ± 0.139	0.697 ± 0.187	0.774 ± 0.162	0.866 ± 0.131	**0.877 ± 0.144**	0.849 ± 0.162
Mean	0.689 ± 0.230	0.700 ± 0.215	0.535 ± 0.170	0.640 ± 0.150	0.747 ± 0.221	**0.791 ± 0.164**	0.700 ± 0.246

Table 3. Performance comparison (DSC) of 10 gastrointestinal tract segments segmentation using seven recent segmentation methods.

method	nnUNet (2D)	nnUNetV2 (2D)	ResUNet (2D)	UCTransNet (2D)	nnUNet (3D)	nnUNetV2 (3D)	CoTr (3D)
Stomach	8.377 ± 8.780	6.814 ± 6.143	15.001 ± 6.964	10.059 ± 10.677	4.397 ± 5.813	**3.918 ± 6.387**	6.191 ± 7.149
Duodenum	10.144 ± 6.443	9.548 ± 6.069	18.234 ± 10.759	12.279 ± 7.071	6.884 ± 4.991	**6.581 ± 5.617**	8.684 ± 7.117
Small Intestine	7.849 ± 3.870	7.735 ± 3.152	14.124 ± 16.772	10.414 ± 4.404	4.909 ± 2.083	**4.617 ± 1.806**	5.875 ± 2.394
Appendix	**19.001 ± 17.700**	30.046 ± 33.802	43.315 ± 33.692	43.234 ± 34.729	29.389 ± 35.432	23.154 ± 16.878	22.804 ± 17.842
Cecum	23.305 ± 20.841	23.797 ± 21.043	34.806 ± 20.504	29.540 ± 20.543	17.426 ± 15.334	**14.369 ± 9.783**	27.296 ± 45.524
Ascending Colon	19.623 ± 22.233	19.765 ± 22.556	23.478 ± 21.560	28.244 ± 31.417	14.268 ± 10.082	**13.191 ± 9.801**	35.804 ± 49.910
Transverse Colon	24.268 ± 21.675	20.616 ± 17.702	26.424 ± 12.427	29.639 ± 27.231	26.117 ± 30.410	**18.351 ± 23.310**	26.001 ± 30.062
Descending Colon	24.948 ± 17.258	26.111 ± 18.557	25.152 ± 10.336	**23.921 ± 13.514**	35.509 ± 35.106	28.105 ± 25.604	39.474 ± 45.864
Sigmoid Colon	27.999 ± 17.014	26.212 ± 15.834	33.462 ± 15.594	34.019 ± 23.425	25.093 ± 17.017	**23.376 ± 16.576**	27.110 ± 19.883
Rectum	10.729 ± 6.346	11.646 ± 9.814	16.260 ± 16.415	12.615 ± 10.029	8.586 ± 6.645	**8.248 ± 7.354**	8.916 ± 7.397
Mean	17.624 ± 7.267	18.229 ± 8.147	25.026 ± 16.502	23.396 ± 18.304	17.258 ± 10.629	**14.391 ± 12.312**	20.815 ± 11.879

Table 4. Performance comparison (HD95 [mm]) of 10 gastrointestinal tract segments segmentation using seven recent segmentation methods.

provided only a standard UNet implementation, we modified it to support additional network architectures. We used the default settings of nnUNet as our experimental settings, with a batch size of two for the 3D methods, 12 for the 2D methods, 1000 epochs, and a loss function combining cross-entropy and dice loss. All models were trained and tested based on these default settings, except that we did not use test-time augmentation owing to the extensive computational resources required—each model needed more than six GPU days to train, and each volume required more than five minutes to infer. We employed the Dice Similarity Coefficient (DSC), a widely used metric, to evaluate the segmentation quality by measuring the pixel overlap between the gold standard and prediction. In image segmentation, the Hausdorff distance is highly sensitive to the accuracy of the segmented boundaries, whereas the Dice coefficient focuses more on the consistency within the mask's interior. Therefore, we utilized the 95% Hausdorff distance (95Hd) to assess the quality of the boundaries in the image segmentation.

Evaluation of SOTA methods on the whole intestine dataset. Fully supervised learning is a fundamental and widely used approach for deep learning-based clinical applications, particularly in automatic multi-organ delineation systems. In this study, we explored several state-of-the-art (SOTA) methods for our dataset, including nnUNet[8], ResUNet[9], UCTransNet[10], and CoTr[11]. The quantitative segmentation results for DSC and HD95 are presented in Tables 3, 4, respectively. Results indicate that all SOTA methods can achieve promising results for large organs such as the stomach, duodenum, small intestine, ascending colon, transverse colon, descending colon, sigmoid colon, and rectum (DSC > 77%). However, the segmentation results for the appendix and cecum are notably poor, with almost all methods achieving a DSC of <70% and HD95 > 20 mm. These findings suggest that segmenting large organs is a well-addressed problem given sufficient high-quality annotated samples from MRI images that offer clear soft tissue delineation. Currently, the image quality is sufficient to clearly distinguish large organs.

The challenge remains in achieving satisfactory segmentation results for small organs such as the appendix and cecum, even with strong soft-tissue recognition capabilities on MRI datasets. Limited research has focused on addressing these issues, and many datasets lack annotations for these small organs. Moreover, in our study, we distinguished the duodenum from the rest of the small intestine without further distinguishing between the jejunum and ileum. This is because the jejunum and ileum often do not have an exact boundary on MR Images, making accurate segmentation difficult. Identification of the terminal ileum is clinically significant for inflammatory bowel, as IBD lesions usually occur in the terminal ileum. This is one of the limitations of this study and a direction for future research.

Usage Notes

In summary, we introduced a meticulously annotated whole-intestine MRI dataset and evaluated several SOTA methods using this dataset. Our clinical research highlights the need for further improvements in model performance, particularly in small organs. We also identify unresolved technical and clinical issues that suggest potential research directions. The segmentation model database can be further utilized for the classification of intestinal MR signals and establishment of disease prediction models for IBD. It is clinically important to detect lesions in the corresponding intestinal segments of IBD, including automatic measurement of intestinal wall thickness and intestinal lumen diameter, automatic morphology fitting of the intestinal lumen, and qualitative and quantitative analyses of the presence of stenosis or penetrating lesions in the intestinal wall. In the future, we aim to expand our dataset to encompass a more extensive and uneven range of cases.

Code availability

No custom code was used.

Received: 1 October 2024; Accepted: 5 March 2025;
Published online: 11 March 2025

References

1. Hong, S. M. & Baek, D. H. Diagnostic Procedures for Inflammatory Bowel Disease: Laboratory, Endoscopy, Pathology, Imaging, and Beyond. *Diagnostics (Basel)*. **14**(13) (2024).
2. Shaban, N. *et al*. Imaging in inflammatory bowel disease: current and future perspectives. *Frontline Gastroenterol* **13**(e1), e28–e34 (2022).
3. Rimola, J. *et al*. Magnetic resonance for assessment of disease activity and severity in ileocolonic Crohn's disease. *Gut* **58**(8), 1113–20 (2009).
4. Samuel, S. *et al*. Endoscopic skipping of the distal terminal ileum in Crohn's disease can lead to negative results from ileocolonoscopy. *Clin Gastroenterol Hepatol* **10**(11), 1253–9 (2012).
5. Bruining, D. H. *et al*. Consensus Recommendations for Evaluation, Interpretation, and Utilization of Computed Tomography and Magnetic Resonance Enterography in Patients With Small Bowel Crohn's Disease. *Radiology* **286**(3), 776–799 (2018).
6. Luo, X. *et al*. Deep learning-based accurate delineation of primary gross tumor volume of nasopharyngeal carcinoma on heterogeneous magnetic resonance imaging: A large-scale and multi-center study. *Radiother Oncol* **180**, 109480 (2023).
7. Wasserthal, J.A.-O. *et al*. *TotalSegmentator: Robust Segmentation of 104 Anatomic Structures in CT Images*. (2638–6100 (Electronic)).
8. Isensee, F. *et al*. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nature Methods* **18**(2), 203−+ (2021).
9. Diakogiannis, F. I. *et al*. ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data. *ISPRS Journal of Photogrammetry and Remote Sensing* **162**, 94–114 (2020).
10. Wang, H. *et al*. Uctransnet: rethinking the skip connections in u-net from a channel-wise perspective with transformer. in *Proceedings of the AAAI conference on artificial intelligence*. (2022).
11. Xie, Y. *et al*. Cotr: Efficiently bridging cnn and transformer for 3d medical image segmentation. in *Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part III 24*. Springer (2021).
12. Li, X. H. *et al*. Assessment of Activity of Crohn Disease by Diffusion-Weighted Magnetic Resonance Imaging. *Medicine (Baltimore)* **94**(43), e1819 (2015).
13. Li, X. H. *et al*. Diffusion-weighted MRI Enables to Accurately Grade Inflammatory Activity in Patients of Ileocolonic Crohn's Disease: Results from an Observational Study. *Inflamm Bowel Dis* **23**(2), 244–253 (2017).
14. Huh, J. *et al*. Diffusion-Weighted MR Enterography to Monitor Bowel Inflammation after Medical Therapy in Crohn's Disease: A Prospective Longitudinal Study. *Korean J Radiol* **18**(1), 162–172 (2017).
15. Yushkevich, P.A. *et al*. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. **31**(3): p. 1116–1128 (2006).
16. Zhong, H. *et al*. A comprehensive dataset of magnetic resonance enterography images with bowel segment annotations [Data set]. *Zenodo*. https://doi.org/10.5281/zenodo.13839321 (2024).

Acknowledgements

This work was supported by the National Key R&D Program of China (approval number: 2023YFC2507300), the National Natural Science Foundation of China (approval number: 62371303, 82270693), and the Key-Area Research and Development Program of Guangdong Province (approval number: 2023B1111040003).

Author contributions

Z.Z. and L.H. designed the study and drafted the manuscript. L.H., S.F., X.W., B.L. and X.L. collected and processed the datasets. Z.Z. and H.L. analyzed the data. X.L., B.H. and H.L. reviewed and edited the manuscript. B.H., L.H. and S.F. coordinated and supervised this study. All authors were involved in the critical revision of the manuscript and approved the final version.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41597-025-04760-z.

Correspondence and requests for materials should be addressed to X.L. or B.H.