

RESEARCH

Open Access



Computed tomography enterography radiomics and machine learning for identification of Crohn's disease

Qiao Shi^{1*}, Yajing Hao¹, Huixian Liu¹, Xiaoling Liu¹, Weiqiang Yan², Jun Mao³ and Bihong T. Chen⁴ 

Abstract

Background Crohn's disease is a severe chronic and relapsing inflammatory bowel disease. Although contrast-enhanced computed tomography enterography is commonly used to evaluate crohn's disease, its imaging findings are often nonspecific and can overlap with other bowel diseases. Recent studies have explored the application of radiomics-based machine learning algorithms to aid in the diagnosis of medical images. This study aims to develop a non-invasive method for detecting bowel lesions associated with Crohn's disease using CT enterography radiomics and machine learning algorithms.

Methods A total of 139 patients with pathologically confirmed Crohn's disease were retrospectively enrolled in this study. Radiomics features were extracted from both arterial- and venous-phase CT enterography images, representing both bowel lesions with Crohn's disease and segments of normal bowel. A machine learning classification system was constructed by combining six selected radiomics features with eight classification algorithms. The models were trained using leave-one-out cross-validation and evaluated for accuracy.

Results The classification model demonstrated robust performance and high accuracy, with an area under the curve of 0.938 and 0.961 for the arterial- and venous-phase images, respectively. The model achieved an accuracy of 0.938 for arterial-phase images and 0.961 for venous-phase images.

Conclusions This study successfully identified a radiomics machine learning method that effectively differentiates Crohn's disease bowel lesions from normal bowel segments. Further studies with larger sample sizes and external cohorts are needed to validate these findings.

Keywords Crohn's disease, Machine learning, Radiomics, Computed tomography enterography (CTE)

Introduction

Crohn's disease (CD) is a severe and chronic inflammatory bowel disease characterized by relapsing symptoms, including abdominal pain, bloody stool, fever, and signs of bowel obstruction [1–3]. The prevalence of CD is higher among young adults aged 20 to 29 years compared to other age groups [3, 4], with rates ranging from 0.54 to 3.44 per 100,000 in developed countries [5]. The prognosis of CD is uncertain due to the limited efficacy of conventional treatments such as surgery, steroid therapy, and biological agents [6–8]. Early diagnosis plays a critical role in optimizing treatment

*Correspondence:

Qiao Shi
docshi@126.com

¹ Department of Radiology, Shenzhen Baoan Women's and Children's Hospital, #56, Yulv St., Baoan District, Shenzhen, Guangdong 518102, People's Republic of China

² Department of Medical Imaging, Peking University Shenzhen Hospital, Shenzhen, Guangdong 518036, People's Republic of China

³ Zhuhai People's Hospital (Affiliated With Jinan University), Shenzhen, Guangdong 519000, People's Republic of China

⁴ Department of Diagnostic Radiology, City of Hope National Medical Center, Duarte, CA 91010, USA



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

and preventing complications. However, current diagnostic methods, such as invasive ileocolonoscopy, have limitations and associated risks. This study introduces a novel approach by utilizing contrast-enhanced computed tomography (CT) enterography (CTE), along with radiomics-based machine learning algorithms and predictive modeling, to evaluate CD bowel lesions.

CTE is frequently employed in clinical practice to evaluate CD due to its superior spatial resolution [1, 3, 9–13]. CTE images can help to assess transmural bowel lesions, extraluminal disease, and lesions in the proximal and middle ileum, which are either inaccessible or undetectable by ileocolonoscopy [14]. Typical CTE findings associated with CD include bowel luminal narrowing, mural thickening, pre-stenotic dilation [13, 15], mucosal hyperenhancement, and mesenteric fat stranding [16]. However, these imaging findings can overlap with other bowel diseases such as tumors, bacterial enteritis and ulcerative colitis. Therefore, further research is necessary to enhance the accuracy of clinically acquired CTE in identifying CD bowel lesions.

Radiomics involves computerized extraction of data from medical images obtained during clinical practice to evaluate unique imaging features that are imperceptible to the human eye, enabling differentiation between normal and abnormal tissue [17, 18]. Radiomics-based machine learning algorithms have been utilized to distinguish patients with CD from those without CD using genomic information [19]. These computational methods have also been employed to assess remission [20], predict genetic factors associated with risk, and characterize the extra-intestinal manifestations of CD [21]. In the field of oncology, machine learning techniques have been applied for tumor classification, grading, and survival analysis [22–24]. Nonetheless, there is a dearth of information regarding the utilization of radiomics features from CTE images and machine learning algorithms to distinguish CD lesions from normal tissues.

Thus, this research endeavor seeks to develop a diagnostic tool by integrating radiomics-based machine learning algorithms and predictive modeling. By extracting data from clinically acquired CTE images, we aim to identify distinctive imaging features that can accurately differentiate CD lesions from normal bowel tissues. Such a tool would prove invaluable for physicians lacking extensive expertise in interpreting CD-related imaging findings. Consequently, this study holds significant promise in facilitating timely and accurate lesion recognition for patients and ultimately improving treatment outcomes. Moreover, it has the potential to broaden the application of radiomics and machine learning in

gastroenterology, a field where limited information exists on their utility for distinguishing CD lesions.

Material and methods

Patient cohort and CTE acquisition

The experimental protocol was approved by ethics committee/Institutional Review Board of both Shenzhen Baoan Women's and Children's Hospital and Zhuhai People's Hospital. The need for informed consent was waived by the ethics committee/Institutional Review Board of Shenzhen Baoan Women's and Children's Hospital and Zhuhai People's Hospital, because of the retrospective nature of the study. The study was conducted in accordance with the ethical principles described in the 1964 Declaration of Helsinki and its amendments.

The inclusion criteria were:

- 1) Cases with enhanced CT images;
- 2) Cases with histology confirmed diagnosis of CD;
- 3) No prior history of any treatment.

A total of 139 patients spanning from 2010 to 2020, met the preliminary selection criteria. The exclusion criteria included the absence of arterial or venous phase CTE images and insufficient image quality. However, no cases were excluded based on these criteria.

CTE scans were conducted using either a 64-multidetector CT (MDCT) scanner (Discovery CT750HD, GE Healthcare, Milwaukee, WI, USA) or a 128-slice dual-source CT scanner (SOMATOM Definition Flash, Siemens, Germany). Intravenous administration of 1.1 mL/kg lopamidol (Lopamidol 370; Bracco Sine, Shanghai, P. T. China) at a rate of 3–6 mL/second (injection time=18 s) through an antecubital vein was employed for contrast. The MDCT scanning protocol encompassed pre-enhancement scan, late arterial-phase (hereafter refer as arterial-phase) and venous-phase scans. Details on CTE scanning parameters are available in Supplemental Table 1. Imaging data were reconstructed using adaptive statistical iterative reconstruction (ASIR) at ADW 4.5 workstations (GE Healthcare, Milwaukee, WI, USA), adhering to standardized patient preparation and imaging acquisition protocols in line with consensus statements for small bowel imaging [25].

Prior to undergoing CTE, all patients were required to fasted for 6 h and ingested a 1500 mL 2.5% mannitol solution (Kelun Pharmaceutical, Sichuan, China) as a negative enteric contrast. This was ingested in three doses of 500 mL every 15 min, starting 45 min before the CTE scan, to ensure adequate bowel luminal distension. Additionally, a rectal enema of 200–500 mL of water was administered on the scanning bed, with the volume adjusted according to patient tolerance.

To minimize bowel peristalsis, patients without contraindications received an intramuscular injections of 10 mg raceanisodamine hydrochloride (Minsheng pharmaceutical, Hangzhou, P. R. China) 30 min before the scan. The CTE images were subsequently retrieved from the Picture Archiving and Communication System (PACS) in Shenzhen Baoan Women's and Children's Hospital as well as Zhuhai People's Hospital.

CTE image pre-processing and segmentation

For each patient's CTE images, lesion delineation was performed on a selected slice featuring the Crohn's disease (CD) lesion, along with two additional slices showcasing normal colon tissues, across both arterial- and venous-phase scans. ITK-SNAP, an open-source software for 3D image analysis, facilitated the manual segmentation of the bowel lesions associated with CD and normal bowel tissues [26]. Regions of interest (ROIs) encompassing primary disease components were meticulously outlined on each CTE image by two radiologists, QS and GH, with over a decade of collective experience in abdominal CT imaging. These radiologists were unaware of the patients' clinical details. The study comprised 139 patients, with each patient's CTE scans including three normal and three lesion areas in both arterial phase and venous phase.

Inter-observer (between radiologist 1 and radiologist 2) and intra-observer (radiologist 1 performing the segmentation twice) reproducibility were assessed for both lesion and normal bowel segments by calculating inter- and intra-class correlation coefficients (ICC). This evaluation utilized all available CTE images. The segmentation was independently conducted by two trained researchers, who were blinded to each other's work. Additionally, one researcher repeated the segmentation within a week to measure consistency. Generally, an ICC greater than 0.80 was considered indicative of good agreement in segmentation accuracy.

Inter-observer (between radiologist 1 and radiologist 2) and intra-observer (radiologist 1 performing the segmentation twice) reproducibility were assessed for both lesions and normal bowel segments by calculating the inter- and intra-class correlation coefficients (ICC). This evaluation utilized all CTE images. The segmentation was independently conducted by two trained researchers who blinded to each other's work. Additionally, one researcher repeated the segmentation within a week to measure consistency. Generally, an ICC greater than 0.80 was considered indicative of good agreement in segmentation accuracy.

Radiomics feature extraction

Radiomics features were extracted from each ROI using the Pyradiomics package (<http://www.radiomics.io/pyradiomics.html>) for Python 3.7. Prior to feature extraction, various filters—Logarithm, Square, Laplacian of Gaussian (LoG), Exponential, Gradient, SquareRoot, Local Binary Pattern 2D (LBP2D), and Wavelet—were applied to enhance the images. For the LoG filter, sigma values were set to 1, 2, or 3 to vary the filter's sharpness. The histogram bin width was standardized was set to 20 for the discretization of image gray levels. A minimum ROI size of 5 voxels was required to ensure meaningful feature extraction. Additionally, the Distance parameter, defining the spacing from the center voxel to its neighbor for angles generation, was adjusted to 1, 2, or 3, optimizing the analysis of spatial relationships.

Feature selection and classification

Feature selection was conducted using various dimension reduction methods, including Ridge Regression (linear least squares with L2 regularization), Least Absolute Shrinkage and Selection Operator (LASSO), Linear Support Vector Classifier (LinearSVC), Logistic regression, Stochastic Gradient Descent Classifier (SGDClassifier), and Random Forest Classifier (RFC). For classification, we employed Adaptive Boosting Classifier (AdaBoost), Gradient Boosting Classifier (GradBoost), Extreme Gradient Boosting Classifier (XGBCClassifier), Ridge Classifier, SGDClassifier, RFC, Logistic Regression, LinearSVC, and Support Vector Classifier (SVC). This analysis was executed using Python 3.7 with scikit-learn version 0.21.3 and xgboost version 1.2.0. Both selection and classification algorithms are sourced from the scikit-learn and xgboost packages, applying the default parameters for consistency. The Pipeline class from scikit-learn facilitated the integration of selectors and classifiers, alongside data scaling and variance reduction as preprocessing steps. A grid search employing a leave-one-out strategy was utilized to systematically evaluate combinations of algorithms. In the leave-one-out approach, each instance is singularly used as the test set while the remainder form the training set.

Analysis of selected radiomics features

We applied hierarchical clustering to visualize radiomics features identified by best-performed models, encompassing 288 features for arterial-phase images and 286 for venous-phase images. Initially, each feature underwent z-score normalization, with the resulting score limited to a range between -1 and 1. Hierarchical clustering was then executed based on Euclidian distance to organize

both samples and features. To compare each feature between the two groups (Crohn's disease versus control normal bowel), we conducted the Wilcoxon signed-rank test within the R statistical programming environment (version 3.6.1). *P*-values were adjusted using Holm–Bonferroni method to account for multiple comparisons.

Statistical analysis

The performance of classification models was evaluated using the area under the receiver operating characteristic curve (AUC) and accuracy metrics. A *P*-value of less than 0.05 was deemed statistically significant. Statistical analyses were conducted in the R statistical programming environment (version 3.6.1), utilizing the 'verification' package (version 1.42) for this purpose. Furthermore, all figures presented in this study were generated using R (version 3.6.1).

The performance of classification models was evaluated using the area under the receiver operating characteristic curve (AUC) and accuracy metrics. A *P*-values of less than 0.05 was deemed statistically significant. Statistical analysis was conducted in the R statistical programming environment (version 3.6.1), utilizing the 'verification' package (version 1.42) for this purpose. Furthermore, all figures presented in this study were generated using R (version 3.6.1).

Results

Patient information

Table 1 summarizes the demographic and clinical characteristics of the study cohort, comprising 139 consecutive patients (median age = 37 years, age range = 15–77 years). The gender distribution included 36 females and 103 males. The majority, 62%, reported a history of smoking. Diarrhea was the most common symptom, affecting 86% of patients, followed by fever in 69%, and weight loss in 64%. Diagnosis was typically confirmed through enteroscopy histology after three months from symptom onset for most patients. According to the Montreal classification of inflammatory bowel disease [27], the ileal (L1) and ileocolonic (L3) regions were predominantly affected locations. Notably, 110 of the 139 patients, exhibited the "comb sign"—characterized by multiple tubular and tortuous opacities in contrast-enhanced CT images of the ileum [28]—suggesting mesenteric hypervascularity.

Radiomics feature extraction and classification algorithms

Both arterial- and venous-phase CTE images were collected from the 139 patients (Table 1). Figure 1 describes the workflow. A total of 1037 radiomics features were extracted (Fig. 1B) and were evaluated with 54 machine learning approaches using the combination of six feature

Table 1 Clinical characteristics of the patients with Crohn's disease

Characteristics (n = 139)	
Median age at diagnosis, years	37 (range: 15 to 77)
Gender	
Female	36 (26%)
Male	103 (74%)
Smoking history	
Yes	86 (62%)
No	53 (38%)
Symptom onset to diagnosis	
< 1 month	16 (12%)
1–3 months	35 (25%)
> 3 months	88 (63%)
Clinical presentations	
Fever	96 (69%)
Weight loss (> 5 kg)	89 (64%)
Diarrhea	120 (86%)
Histology diagnosis made by	
Enteroscopy	135 (97%)
Surgery	4 (3%)
Location of lesion, number of patients ^a	
L1	36
L2	7
L3	32
L4	7
L1, L2, and L3	7
L1 and L3	7
L3 and L4	4
Undetermined	39
Disease behavior, number of patients ^b	
B1	96
B2	35
B3	4
Undetermined	4
Thickness of bowel wall in mm, mean ± Std	9.56 ± 4.07
Luminal stenosis, n (%)	39 (28%)
Mural stratification, n (%)	53 (38%)
Comb sign, n (%) ^c	110 (79%)
Fibrofatty proliferation, n (%)	89 (64%)

^a Location of lesion (The Montreal classification): L1, terminal ileum (TI) without colonic involvement; L2, colonic involvement without TI involvement; L3, ileocolonic disease involving both the TI and colon; L4, disease proximal to the TI without TI or colonic involvement

^b Disease behavior: B1, non-stricturing and non-penetrating; B2 stricturing; B3 penetrating

^c Comb sign: multiple tubular, tortuous opacities on contrast enhanced computed tomography scan of the ileum

reductions and nine classifiers (Fig. 1C). We compared the AUC and accuracy values for each group and analyzed radiomic features (Fig. 1D).

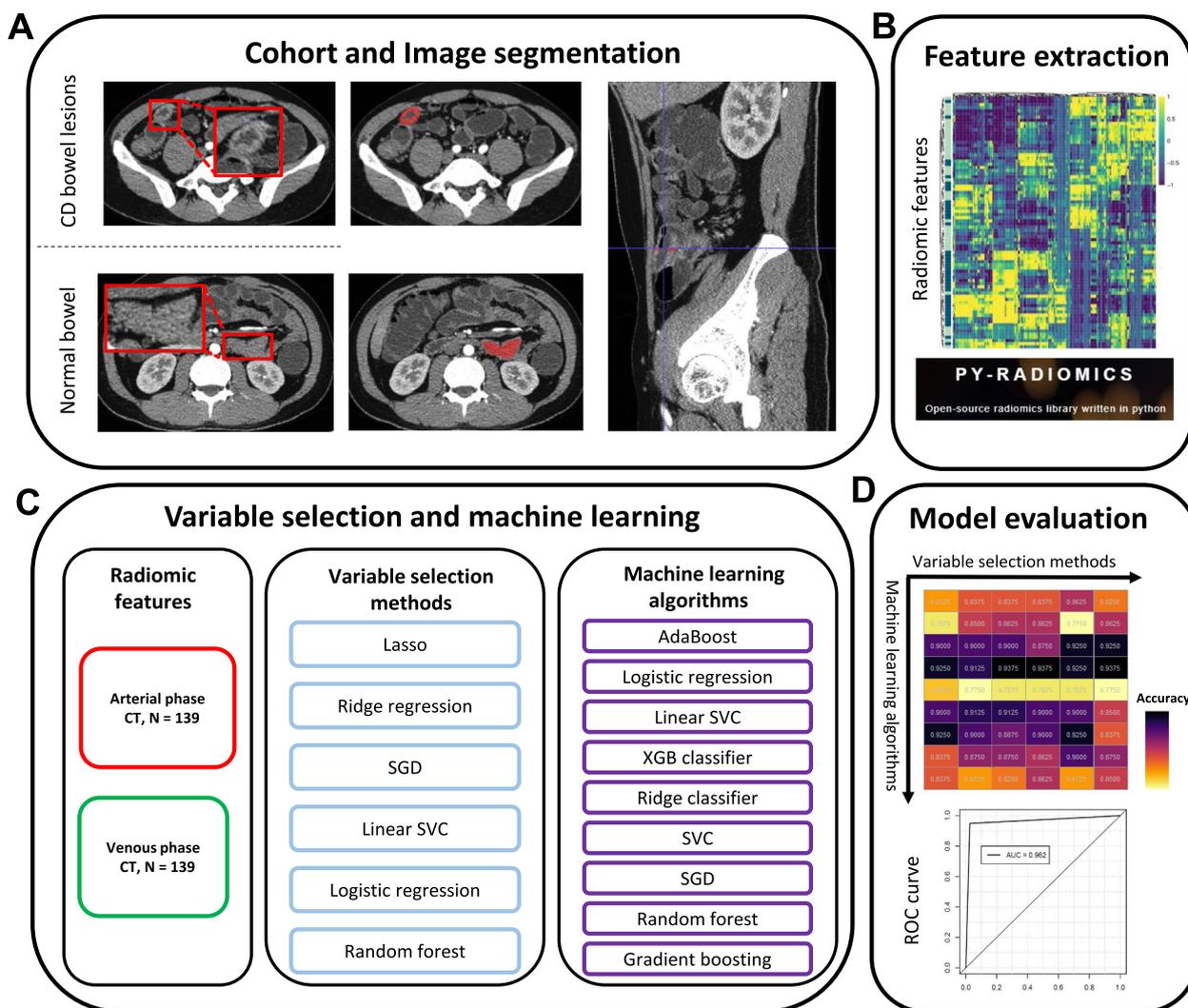


Fig. 1 Workflow of Bowel Lesion Segmentation and Predictive Modeling Using CTE Images. **a** Segmentation of both bowel lesions and normal bowel tissues using CTE images, highlighting the initial step in distinguishing pathological from healthy tissue. **b** Extraction of radiomics features from the segmented areas, focusing on quantifying lesion characteristics. **c** Selection of relevant features and application of machine learning algorithm to refine the diagnostic model. **d** Development of predictive models and evaluation of their performance, with metrics such as the AUC of ROC indicating diagnostic accuracy. Abbreviations: AUC, Area Under the Curve; CD, Crohn’s Disease; CTE, Computed Tomography Enterography; LASSO, Least Absolute Shrinkage and Selection Operator; ROC, Receiver Operating Characteristic; SGD, Stochastic Gradient Descent; SVC, Support Vector Classifier; XGB, Extreme Gradient Boosting

Figure 2 demonstrates the validation accuracy of various feature selection and classification method using leave-one-out cross-validation. For arterial-phase images, the highest accuracy (0.9375) was achieved by the combination of Logistic Regression+Logistic Regression and RFC+Logistic Regression. In contrast, for venous-phase images, the combination of Logistic Regression+Logistic Regression, LinearSVC+Logistic Regression, Logistic Regression+LinearSVC, and Ridge+Logistic Regression reached the peak performance with an accuracy of 0.9615. The results of Logistic Regression+Logistic

Regression for both arterial- and venous-phase images are further illustrated in Fig. 3.

A total of 107 features without filters was extracted including shape features ($n=14$), intensity features($n=18$), Gray Level Cooccurrence Matrix (GLCM) features ($n=24$), Gray Level Dependence Matrix (GLDM) features ($n=14$), Gray Level Run Length Matrix (GLRLM) features ($n=16$), Gray Level Size Zone Matrix (GLSZM) features($n=16$) and Neighboring Gray Tone Difference Matrix (NGTDM) features($n=5$). Without these filters, our model achieved a good

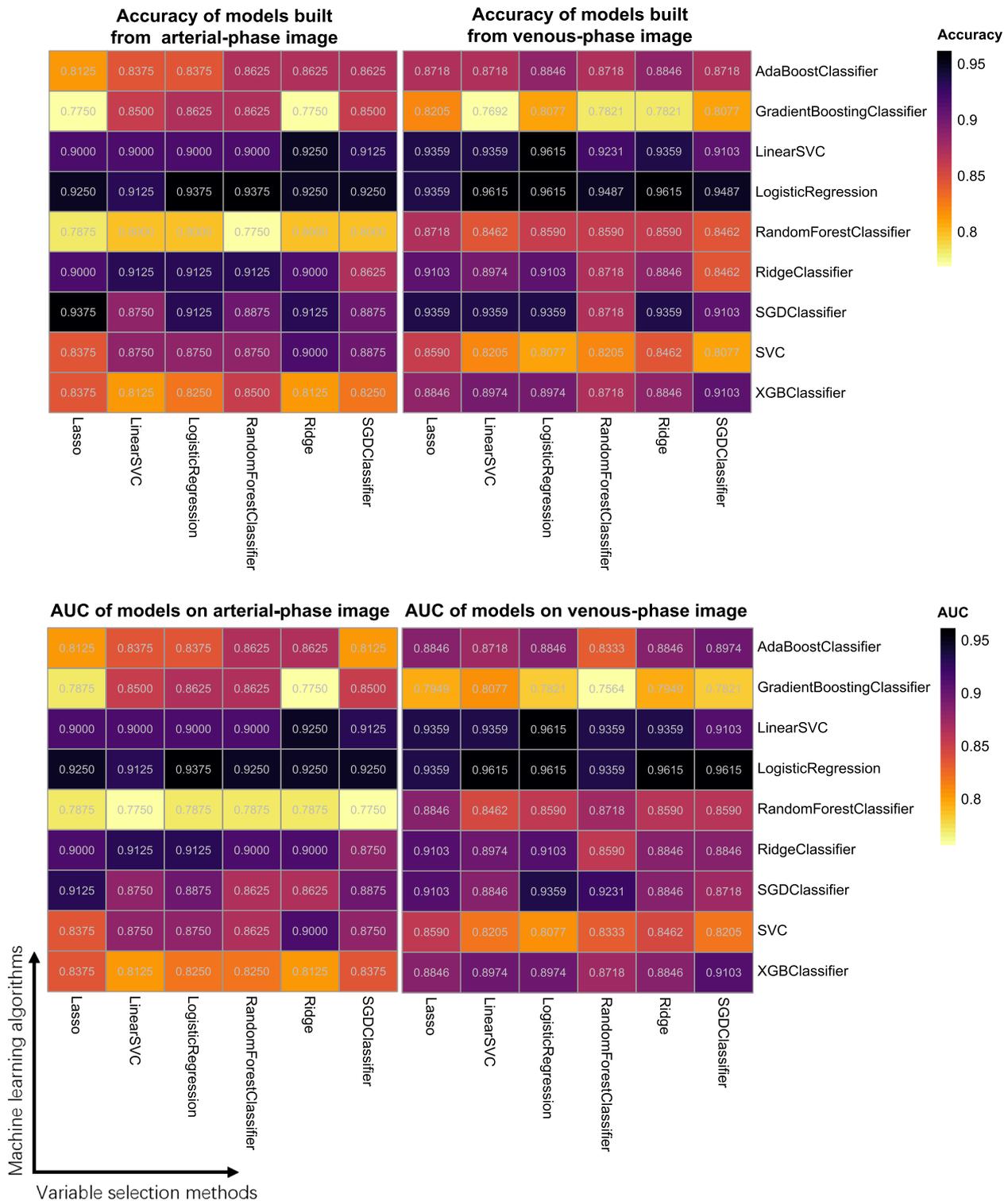


Fig. 2 Evaluation of Machine Learning Models Using Arterial- and Venous-Phase Images. This figure presents the accuracy assessment and heatmaps for the machine learning models, utilizing leave-one-out cross-validation. It compares the performance on arterial-phase images (left) and the venous-phase images (right), illustrating the predictive accuracy and pattern recognition capabilities of each model in different imaging phases

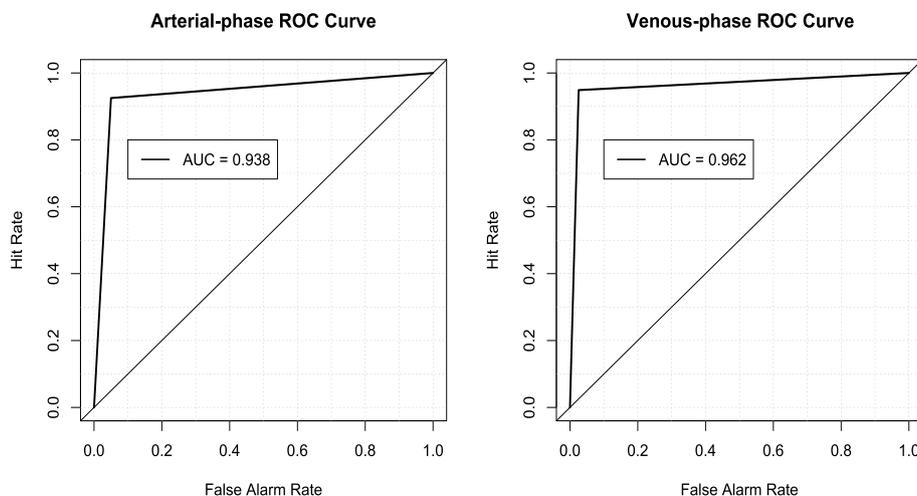


Fig. 3 Comparative Performance of Machine Learning Methods Using ROC Curves and AUC Values. This figure depicts the evaluation of machine learning methods through ROC curves and AUC metrics for arterial-phase images (left) and venous-phase images (right). The comparison highlights the diagnostic ability of each method across different imaging phases

performance, with accuracy of 0.91 and 0.86 for the arterial- and venous-phase images, respectively. After implementing the filters such as Logarithm, Wavelet, Square, LoG, Exponential, Gradient, Square Root, and LBP2D, the model performance was improved, and the accuracy was increased to 0.94 and 0.96. The results suggested that modifying the CTE images with filters could enhance the extraction of radiomics information [29, 30].

Reproducibility of CTE image segmentation

The inter-observer ICCs for all image ROIs for reader 1 and reader 2 were 0.933 ± 0.025 and 0.927 ± 0.032 for arterial- and venous-phase scans, respectively (Supplemental Tables 2). The intra-observer ICCs for all normal tissues and CD lesions for reader 1's first and second evaluations were 0.987 ± 0.004 and 0.986 ± 0.005 for the arterial- and venous-phase scans, respectively (Supplemental Tables 3). We used the following additional consistency tests for CD bowel lesions and normal bowel: DICE (Dice Coefficient), VOLSMITY (Volumetric Similarity Coefficient) and JACRD (Jaccard Coefficient), and the results indicated good reproducibility (Supplemental Tables 2–3).

Radiomics feature clustering

Clustering analysis of the radiomics features from the CTE images for both arterial- and venous-phase images showed that lesions and normal bowel have distinct difference in radiomics features (Fig. 4). The radiomics features, including shape-based, gray level-based, and intensity-based features, reflected the intrinsic differential imaging features between CD bowel lesions and

normal bowel on contrast-enhanced CTE images. Most features selected were significantly different between groups as determined by the Wilcoxon test. The top 10 significant radiomic features are shown in Table 2. Among these features, the first feature represented the shape feature, indicating that lesion size plays a crucial role in CD. GLDM features (2 to 5 in the ranking), representing the second-order statistics of images, showed significant values for model classification. These features emphasized the spatial relationships of lesion pixels and their neighborhoods.

Discussion

In this study, we combined radiomics and machine learning methods to distinguish CD bowel lesions from normal bowel. Our machine learning models reached an AUC of over 0.93 and an accuracy of over 0.93 for both the arterial-phase and venous-phase images. Thus, our pilot data showed that CTE radiomics may be a potentially useful method to improve the non-invasive diagnosis and monitoring of CD.

Typical CT imaging features for CD include thickening and unusual enhancement of the bowel wall, bowel stricture, and extraluminal changes, which correlates with active inflammation and response to therapy [9, 31–35]. However, it can be challenging to identify these features on bowel segments with CD lesions, especially in the less distended bowel or collapsed bowel, which may mimic soft tissue masses. Our results indicate that CTE radiomics and machine learning algorithms could be helpful to distinguish the diseased bowel with CD from the collapsed normal bowel. This should not be surprising as

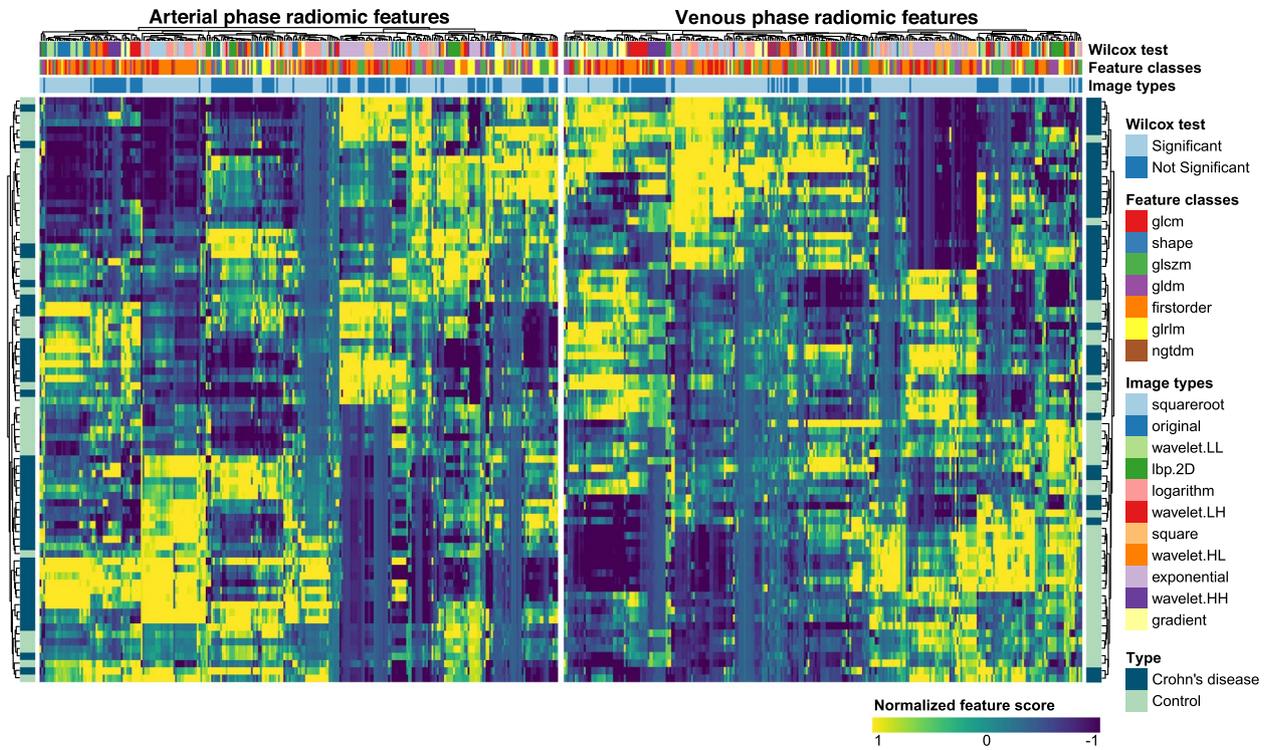


Fig. 4 Clustering Analysis and Heatmaps of Radiomics Features in Arterial- and Venous-Phase Images. This figure illustrates the clustering analysis and heatmaps, displaying radiomics features (columns) against each analyzed image (rows) for arterial-phase images (left) and venous-phase images (right). Significance is determined by an adjusted Wilcoxon *p* value, with a threshold of < 0.05 indicating significance and > 0.05 indicating non-significance

Table 2 Top 10 significant radiomic features selected in the predictive models

Feature source	Feature name	P-value	Wilcox statistics	Adjusted-P	Mean	SD
Arterial phase	original shape Maximum2DdiameterRow	0.00026491	1172	6.50e-04	27.76700798	11.07093432
Arterial phase	lbp-2D gldm DependenceEntropy	6.9297E-08	1332	9.19e-07	3.627823454	0.690042936
Arterial phase	lbp-2D gldm LargeDependenceEmphasis	2.422E-06	328	1.02e-05	1823.879162	196.5698259
Arterial phase	lbp-2D gldm LargeDependenceHighGrayLevelEmphasis	2.9176E-07	1309	1.95e-06	1355.76384	944.8137645
Arterial phase	lbp-2D gldm LargeDependenceLowGrayLevelEmphasis	8.8983E-05	402	2.50e-04	1145.433246	226.7343427
Arterial phase	squareroot glszm ZoneEntropy	6.9297E-08	1332	9.19e-07	5.44181741	0.540545674
Arterial phase	logarithm gldm DependenceEntropy	6.495E-08	1333	9.19e-07	6.304725307	0.607748159
Arterial phase	wavelet-HL firstorder 90Percentile	2.1525E-07	1314	1.67e-06	10.6712064	9.586738546
Arterial phase	wavelet-HH glszm GrayLevelNonUniformity	1.8803E-07	1342	1.50e-06	18.38896654	11.10963088
Arterial phase	wavelet-LL firstorder RobustMeanAbsoluteDeviation	5.9451E-07	1297	3.35e-06	18.38628947	6.411283722
Venous phase	original shape Maximum2DdiameterRow	0.00400135	1046	7.41e-03	28.94012286	10.50962803
Venous phase	lbp-2D gldm DependenceEntropy	2.9156E-07	1250	2.19e-06	3.599582042	0.688791176
Venous phase	lbp-2D gldm LargeDependenceEmphasis	5.7008E-06	322	2.42e-05	1835.016931	183.9222846
Venous phase	lbp-2D gldm LargeDependenceHighGrayLevelEmphasis	0.00013022	386	3.64e-04	1194.078438	241.9314389
Venous phase	lbp-2D gldm LargeDependenceLowGrayLevelEmphasis	9.1941E-06	1190	3.59e-05	1215.99752	633.5716877
Venous phase	squareroot gldm LargeDependenceHighGrayLevelEmphasis	3.306E-07	1248	2.39e-06	46252.93852	53311.24276
Venous phase	squareroot glszm ZoneEntropy	2.7373E-07	270	2.17e-06	4572.723736	6931.740139
Venous phase	logarithm gldm DependenceEntropy	8.2496E-09	1302	1.24e-07	6.088002214	0.773459549
Venous phase	wavelet-LL firstorder InterquartileRange	3.686E-06	1207	1.73e-05	42.82224848	11.5787774
Venous phase	wavelet-HL firstorder 90Percentile	2.3946E-06	288	1.18e-05	12.65935559	9.567709839

radiomics could extract many features reflective of the underlying CD pathology including intensity-related features, shape-based features, and texture features, which are not visible to human eyes through conventional radiological assessment.

Our strategy for model building and performance testing were generally in line with existing literature. Previous studies have explored the prognostic capabilities of radiomic features for CD and other bowel pathology such as cancer [17, 36, 37]. Feature selection and predictive modeling are critical for high throughput data-driven radiomics [22]. For instance, Ryan et al. assessed small bowel structure and morphology in CD using semi-automated radiomic measures with an accuracy of 0.876 [38]. Feng et al. built a radiomic model for characterizing intestinal fibrosis in CD and the model was robust across different medical centers [39]. Yi et al. used support vector machine algorithm combined with LASSO or RFC to regress radiomics and traditional radiological-clinico-pathological data for predicting locally advanced rectal tumor response to neoadjuvant chemoradiotherapy with high performance [40]. In our study, we tested different combinations of feature selection and classification algorithms to achieve optimal performance for detection of CD bowel lesion. In addition, our study was unique in that we extracted radiomic features from CTE, which provided more details about the bowel pathology than routine conventional CT scan of the abdomen and pelvis.

In this study, we evaluated 6 feature selection methods and 9 classification methods and we found all selection methods with the classifier LogisticRegression achieved relatively high performance for both the arterial- and the venous-phase CTE images and all classifier with LogisticRegression as feature selection methods also achieved relatively high performance. However, the predictive performance of RFC was not optimal compared with other classifiers in our study although RFC has been frequently used as a robust algorithm in machine learning research [41, 42]. A similar finding has been reported by Sun et al. [43], which showed that RFC had lower performance than the XGBClassifier and gradient boosting decision tree (GBDT) for glioma grading differentiation. We speculate the less powerful performance of RFC in our study may be due to specific bowel pathology associated with CD.

In Table 2, we listed the top 10 radiomics features contributing to the model classification performance. These features showed high values in lesion recognition, however, the practical interpretations and correlations with clinical and biological characteristics still needed to be explored deeply. In this study, the selected features included the shape and texture features. For shape features, they provided the lesion compactness

(2D-diameter in Table 2), which prompt clinicians to carefully observe details such as the size and length of suspicious lesions. This, in turn, underscores the importance for radiologists to meticulously identify ROI. As for the texture features [44] they mainly describe the image pixel information and their spatial and gray-level relationships. Due to the focus of these features mainly on the images themselves, we still need to further explore their potential clinical and biological significance [45] in relation to patient course and CD-related pathological biomarkers in our subsequent research under the guidance of the criteria [46]. Nevertheless, our model's high validation performance underscores its significant clinical utility in assisting doctors to identify potential lesions in intestinal CT scans.

There were several limitations to this study. First, this was a retrospective study with a modest sample size, limiting the generalizability of the results. Our data will need to be validated in a large multicenter study [47]. Further prospective researches are warranted to broaden the clinical applicability and enhance the generalizability of the model. Secondly, our study aimed to distinguish severe CD bowel lesions, specifically ulcerated lesions from normal bowel. Future work to test the performance of these algorithms for less severe CD lesions that do not exhibit ulcerations. This is particularly important as CD lesions can vary in severity, especially in the early stage of the disease. Thirdly, it is worth noting that our model did not take into account for confounding clinical variables, such as disease grading, progression, treatment regimen and survival information. Future studies should explore how these clinical variables and other multi-omics features may affect the model performance. And we only implemented the radiomics features for CD lesion recognition. The interpretation between the features and clinical factors will be investigated. Lastly, deep learning algorithms might generate additional useful data for identification of CD, and these should be tested in the future studies [48, 49]. Lastly, the labeling of involved and uninvolved bowel segments was manually performed by two independent radiologists. This approach is not as accurate as diagnosis based on biopsy. In order to improve the accuracy of the segmentation, we evaluated the inter-observer reproducibility between the two radiologists and applied stringent criteria to filter the segmentation results. Future studies using histology-confirmed lesion labels could further increase the accuracy of the model.

Conclusions

In summary, we developed a classification model validated to distinguish between diseased and normal bowel in CD patients using radiomics features extracted from CTE images and grouping machine learning selection

and classification algorithms. The model achieved high validation performance with AUC for arterial-phase and venous-phase images. In the future, we will expand the robustness of the model by correlating the features to clinical factors, and multi-center and prospective studies will be conducted to facilitate our model application in clinical practice.

Abbreviations

AUC	Area under the curve
CD	Crohn's disease
CT	Computed tomography
CTE	Computed tomography enterography
ROI	Region of interest
ICC	Inter- and intra- class correlation coefficients
LoG	Laplacian of Gaussian
LBP2D	Local Binary Pattern 2D
LASSO	Least absolute shrinkage and selection operator
SVC	Support vector classifier
SGD	Stochastic gradient descent
RFC	Random forest classifier
GradBoost	Gradient boosting classifier
AdaBoost	Adaptive boosting classifier
XGBClassifier	Extreme gradient boosting classifier
GLCM	Gray level cooccurrence matrix
GLDM	Gray level dependence matrix
GLRLM	Gray level run length matrix
GLSZM	Gray level size zone matrix
NGTDM	Neighboring gray tone difference matrix

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12880-024-01480-5>.

Supplementary Material 1.

Acknowledgements

We wish to acknowledge our scientific writer Nancy Linford, PhD., for editing this manuscript.

Authors' contributions

JM and QS conceptualized this study. YH and HL contributed to data curation. XL and WY performed the segmentation of all CTE images. QS performed image analysis. QS and JM wrote the first draft of the manuscript. BC supervised this study and revised the manuscript. All authors read and approved the final version of the manuscript.

Funding

No funding was received for conducting this study.

Data availability

The datasets we used and all analysis codes during the current study are not publicly available due to the privacy issues but are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

The need for informed consent was waived by the ethics committee/Institutional Review Board of Shenzhen Baoan Women's and Children's Hospital, because of the retrospective nature of the study. The experimental protocol was approved by ethics committee/Institutional Review Board of Shenzhen Baoan Women's and Children's Hospital.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 25 August 2023 Accepted: 24 October 2024

Published online: 06 November 2024

References

- Baumgart DC, Sandborn WJ. Crohn's disease. *Lancet*. 2012;380(9853):1590–605.
- Siegel CA, et al. Development of an index to define overall disease severity in IBD. *Gut*. 2018;67(2):244–54.
- Torres J, et al. Crohn's disease. *Lancet*. 2017;389(10080):1741–55.
- Molodecky NA, et al. Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review. *Gastroenterology*. 2012;142(1):46–54 e42 quiz e30.
- Ng SC, et al. Incidence and phenotype of inflammatory bowel disease based on results from the Asia-pacific Crohn's and colitis epidemiology study. *Gastroenterology*. 2013;145(1):158–165 e2.
- Freeman HJ. Natural history and long-term clinical course of Crohn's disease. *World J Gastroenterol*. 2014;20(1):31–6.
- Gajendran M, et al. A comprehensive review and update on Crohn's disease. *Dis Mon*. 2018;64(2):20–57.
- Buisson A, et al. Review article: the natural history of postoperative Crohn's disease recurrence. *Aliment Pharmacol Ther*. 2012;35(6):625–33.
- Puylaert CA, et al. Grading of Crohn's disease activity using CT, MRI, US and scintigraphy: a meta-analysis. *Eur Radiol*. 2015;25(11):3295–313.
- Higgins PD, et al. Computed tomographic enterography adds information to clinical management in small bowel Crohn's disease. *Inflamm Bowel Dis*. 2007;13(3):262–8.
- Minordi LM, et al. CT and MRI evaluations in crohn's complications: a guide for the radiologist. *Acad Radiol*. 2022;29(8):1206–27.
- Li X, et al. Development and validation of a novel computed-tomography enterography radiomic approach for characterization of intestinal fibrosis in Crohn's disease. *Gastroenterology*. 2021;160(7):2303–2316 e11.
- Bachour SP, et al. Test characteristics of cross-sectional imaging and concordance with endoscopy in postoperative Crohn's disease. *Clin Gastroenterol Hepatol*. 2022;20(10):2327–2336 e4.
- Veauthier B, Hornecker JR. Crohn's disease: diagnosis and management. *Am Fam Physician*. 2018;98(11):661–9.
- Bettenworth D, et al. Assessment of Crohn's disease-associated small bowel strictures and fibrosis on cross-sectional imaging: a systematic review. *Gut*. 2019;68(6):1115–26.
- Adler J, et al. Computed tomography enterography findings correlate with tissue inflammation, not fibrosis in resected small bowel Crohn's disease. *Inflamm Bowel Dis*. 2012;18(5):849–56.
- Lambin P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer*. 2012;48(4):441–6.
- Aerts HJ, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun*. 2014;5:4006.
- Romagnoni A, et al. Comparative performances of machine learning methods for classifying Crohn Disease patients using genome-wide genotyping data. *Sci Rep*. 2019;9(1):10351.
- Waljee AK, et al. Development and validation of machine learning models in prediction of remission in patients with moderate to severe Crohn disease. *JAMA Netw Open*. 2019;2(5):e193721.
- Bottigliengo D, et al. The role of genetic factors in characterizing extra-intestinal manifestations in crohn's disease patients: are bayesian machine learning methods improving outcome predictions? *J Clin Med*. 2019;8(6):865.
- Parmar C, et al. Machine learning methods for quantitative radiomic biomarkers. *Sci Rep*. 2015;5:13087.
- Bychkov D, et al. Deep learning based tissue analysis predicts outcome in colorectal cancer. *Sci Rep*. 2018;8(1):3395.
- Kather JN, et al. Predicting survival from colorectal cancer histology slides using deep learning: a retrospective multicenter study. *PLoS Med*. 2019;16(1):e1002730.

25. Taylor SA, et al. The first joint ESGAR/ ESPR consensus statement on the technical performance of cross-sectional small bowel and colonic imaging. *Eur Radiol.* 2017;27(6):2570–82.
26. Yushkevich PA, et al. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *Neuroimage.* 2006;31(3):1116–28.
27. Satsangi J, et al. The Montreal classification of inflammatory bowel disease: controversies, consensus, and implications. *Gut.* 2006;55(6):749–53.
28. Madureira AJ. The comb sign. *Radiology.* 2004;230(3):783–4.
29. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures. *They Are Data Radiology.* 2016;278(2):563–77.
30. Soufi M, Arimura H, Nagami N. Identification of optimal mother wavelets in survival prediction of lung cancer patients using wavelet decomposition-based radiomic features. *Med Phys.* 2018;45(11):5116–28.
31. Hara AK, et al. Using CT enterography to monitor Crohn's disease activity: a preliminary study. *AJR Am J Roentgenol.* 2008;190(6):1512–6.
32. Moy MP, Sauk J, Gee MS. The role of MR enterography in assessing Crohn's disease activity and treatment response. *Gastroenterol Res Pract.* 2016;2016:8168695.
33. Paquet N, et al. Crohn's disease activity: abdominal computed tomography histopathology correlation. *Eur J Radiol Open.* 2016;3:74–8.
34. Park EK, et al. Value of computerized tomography enterography in predicting Crohn's disease activity: correlation with Crohn's Disease activity index and c-reactive protein. *Iran J Radiol.* 2016;13(4):e34301.
35. Kim J, Kim SH, Kim TO. Evaluation of CT enterography findings for endoscopic complete remission after anti-TNF-alpha therapy in patients with Crohn's disease. *Acta Radiol.* 2019;60(10):1200–8.
36. Parmar C, et al. Robust radiomics feature quantification using semiautomatic volumetric segmentation. *PLoS ONE.* 2014;9(7):e102107.
37. Coroller TP, et al. CT-based radiomic signature predicts distant metastasis in lung adenocarcinoma. *Radiother Oncol.* 2015;114(3):345–50.
38. Stidham RW, et al. Assessing small bowel stricturing and morphology in Crohn's disease using semi-automated image analysis. *Inflamm Bowel Dis.* 2020;26(5):734–42.
39. Li X, Liang D, Meng J, et al. Development and Validation of a Novel Computed-Tomography Enterography Radiomic Approach for Characterization of Intestinal Fibrosis in Crohn's Disease. *Gastroenterology.* 2021;160(7):2303–2316.e11. <https://doi.org/10.1053/j.gastro.2021.02.027>. Epub 2021 Feb 17.
40. Yi X, et al. MRI-based radiomics predicts tumor response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer. *Front Oncol.* 2019;9:552.
41. Rodriguez-Galiano VF, et al. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS J Photogramm Remote Sens.* 2012;67:93–104.
42. Singh A, Lakshminathan R. Impact of different data types on classifier performance of random forest, naive bayes, and k-nearest neighbors algorithms. 2018.
43. Sun P, et al. Comparison of feature selection methods and machine learning classifiers for radiomics analysis in glioma grading. *IEEE Access.* 2019;7:102010–20.
44. Abbasian Ardakani A, et al. Interpretation of radiomics features-A pictorial review. *Comput Methods Programs Biomed.* 2022;215:106609.
45. Tomaszewski MR, Gillies RJ. The biological meaning of radiomic features. *Radiology.* 2021;299(2):E256.
46. Huang EP, et al. Criteria for the translation of radiomics into clinically useful tests. *Nat Rev Clin Oncol.* 2023;20(2):69–82.
47. Traverso A, et al. Repeatability and Reproducibility of Radiomic Features: A Systematic Review. *Int J Radiat Oncol Biol Phys.* 2018;102(4):1143–58.
48. Klang E, et al. Deep learning algorithms for automated detection of Crohn's disease ulcers by video capsule endoscopy. *Gastrointest Endosc.* 2020;91(3):606–613 e2.
49. Morilla I. A deep learning approach to evaluate intestinal fibrosis in magnetic resonance imaging models. *Neural Comput Appl.* 2020;32(18):14865–74.

Publisher's Note

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