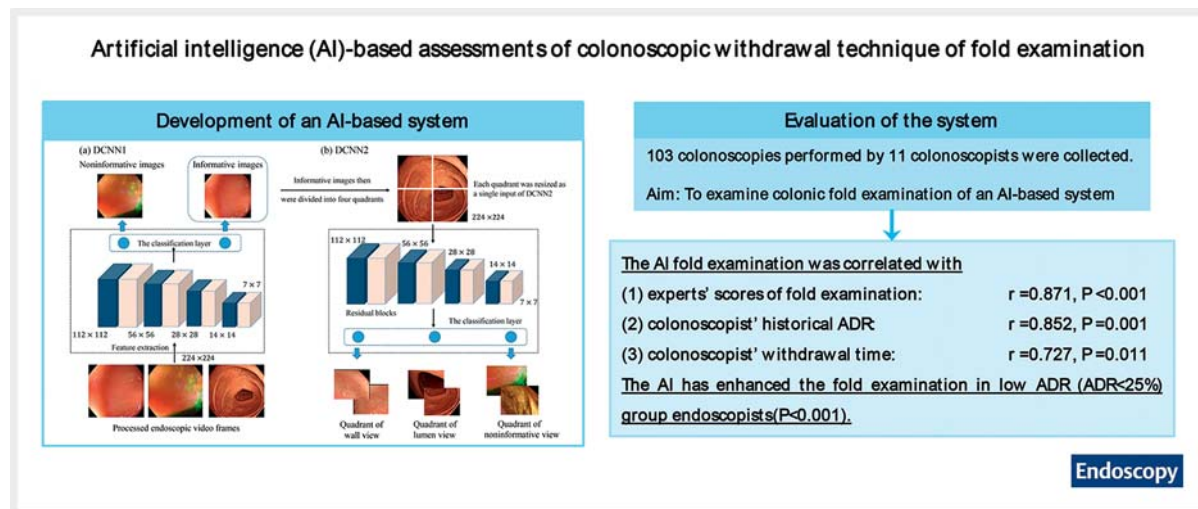**Original article**

Thieme

Artificial intelligence-based assessments of colonoscopic withdrawal technique: a new method for measuring and enhancing the quality of fold examination ▶

OPEN ACCESS

INFOGRAPHIC

Authors

Wei Liu[1,*], Yu Wu[2,*], Xianglei Yuan[1], Jingyu Zhang[3], Yao Zhou[2], Wanhong Zhang[4] ⓞ, Peipei Zhu[5], Zhang Tao[6] ⓞ, Long He[1], Bing Hu[1] ⓞ, Zhang Yi[2]

Institutions

1 Department of Gastroenterology, West China Hospital, Sichuan University, Chengdu, Sichuan, China
2 Machine Intelligence Laboratory, College of Computer Science, Sichuan University, Chengdu, Sichuan, China
3 State Key Laboratory of Ultrasound in Medicine and Engineering, College of Biomedical Engineering, Chongqing Medical University, Chongqing, Sichuan, China
4 Department of Gastroenterology, Cangxi People's Hospital, Guangyuan, Sichuan, China
5 Department of Gastroenterology, Dazhou Integrated Traditional Chinese and Western Medicine Hosptial, Dazhou, Sichuan, China
6 Department of Gastroenterology, Nanchong Central Hospital, Nanchong, Sichuan, China

🌐 Figs. 1 s–3 s, Tables 1 s-4 s
Supplementary material is available under
https://doi.org/10.1055/a-1799-8297

Corresponding author

Bing Hu, MD, Department of Gastroenterology, West China Hospital, Sichuan University, Chengdu 610041, Sichuan, PR China
hubingnj@163.com

---

* Co-first authors

ABSTRACT

Background This study aimed to develop an artificial intelligence (AI)-based system for measuring fold examination quality (FEQ) of colonoscopic withdrawal technique. We also examined the relationship between the system's evaluation of FEQ and FEQ scores from experts, and adenoma detection rate (ADR) and withdrawal time of colonoscopists, and evaluated the system's ability to improve FEQ during colonoscopy.

Methods First, we developed an AI-based system for measuring FEQ. Next, 103 consecutive colonoscopies performed by 11 colonoscopists were collected for evaluation. Three experts graded FEQ of each colonoscopy, after which the recorded colonoscopies were evaluated by the system. We further assessed the system by correlating its evaluation of FEQ against expert scoring, historical ADR, and withdrawal time of each colonoscopist. We also conducted a prospective observational study to evaluate the system's performance in enhancing fold examination.

Results The system's evaluations of FEQ of each endoscopist were significantly correlated with experts' scores (r = 0.871, $P < 0.001$), historical ADR (r = 0.852, $P = 0.001$), and withdrawal time (r = 0.727, $P = 0.01$). For colonoscopies performed by colonoscopists with previously low ADRs (< 25%), AI assistance significantly improved the FEQ, evaluated by both the AI system (0.29 [interquartile range (IQR) 0.27–0.30] vs. 0.23 [0.17–0.26]) and experts (14.00 [14.00–15.00] vs. 11.67 [10.00–13.33]) (both $P < 0.001$).

Conclusion The system's evaluation of FEQ was strongly correlated with FEQ scores from experts, historical ADR, and withdrawal time of each colonoscopist. The system has the potential to enhance FEQ.

Introduction

Early detection and removal of adenomatous polyps via colonoscopy are still considered the gold standard method for prevention of colorectal cancer (CRC). Yet, 25% of adenomas are missed during the examination, which is significantly associated with interval CRC [1, 2]. Some studies have suggested that higher quality colonoscopic withdrawal technique is associated with a lower miss rate for adenomas and that four complementary skills contribute to inspection quality in screening colonoscopy: 1) fold examination, 2) mucosal cleaning, 3) luminal distension, and 4) adequacy of time spent viewing [3]. As the leading factor, fold examination has been reported to be significantly related to the polyp that does not appear in the field of view because of colonoscopy blind spots [4]. Thus, fold examination is strongly recommended for evaluation of colonoscopic withdrawal technique during colonoscopy examination. However, the lack of a quality supervision system makes it challenging to provide quality control of fold examination.

Over recent years, deep convolutional neural networks (DCNNs) have been successfully used for real-time detection of polyps, as well as assessment of bowel preparation, withdrawal speed, and withdrawal time [5–8]. These studies have indicated that artificial intelligence (AI) could indirectly increase quality control during colonoscopy examinations. However, to date, no studies have reported on the development of a DCNN in the assessment of colonoscopic withdrawal technique in terms of fold examination quality (FEQ).

This study aimed to develop an AI-based system for assessment of FEQ of colonoscopic withdrawal technique and to determine the relationship between the system's evaluation of FEQ and whole-colon FEQ scores determined by experts. We also aimed to analyze the relationship between FEQ scores and historical adenoma detection rates (ADRs) and mean withdrawal times of individual colonoscopists, and to evaluate whether use of the AI-based system could improve FEQ in clinical practice.

Methods

Model design

We developed two DCNNs to assess FEQ during colonoscope withdrawal, called GINets. First, DCNN1 was developed to identify informative frames (clear images) or noninformative frames (bubble, sliding, and fuzzy images). Second, DCNN2 was developed to identify whether an informative frame had a lumen view and wall view (indicating a close up examination of the colon mucosa), thus achieving automatic evaluation of FEQ during colonoscopy. Considering that distinguishing the lumen view from the wall view on a single image may be challenging, DCNN2 was used to divide each informative image into four quadrants and identify lumen view, wall view, and noninformative quadrant view based on each quadrant. A lumen view was defined as a quadrant in which the distant colon lumen can be seen; a clear view of a quadrant without the distant colon lumen was defined as a wall view; a noninformative quadrant view was defined as a quadrant with no clear view. Next, each four quadrants of wall view was defined as a whole endoscopic image of wall view to compute the proportions of wall view of each video; this proportion was used for assessment of FEQ and was determined for the whole colon rather than per colon segment. Therefore, the system's evaluation of FEQ (the percentage of wall views) was defined as follows: the total number of quadrants of wall views / (4 × the total number of images in the video stream). The percentage of noninformative frames was defined as follows: the total number of noninformative frames / the total number of images in the video stream. The total number of images in the video stream included all the informative frames and noninformative frames. Finally, we used the percentage of wall views in the video stream to estimate FEQ (▶ Fig. 1, ▶ Fig. 2, Fig. 1 s). The development of the DCNN models is provided in the online-only Supplementary Material.
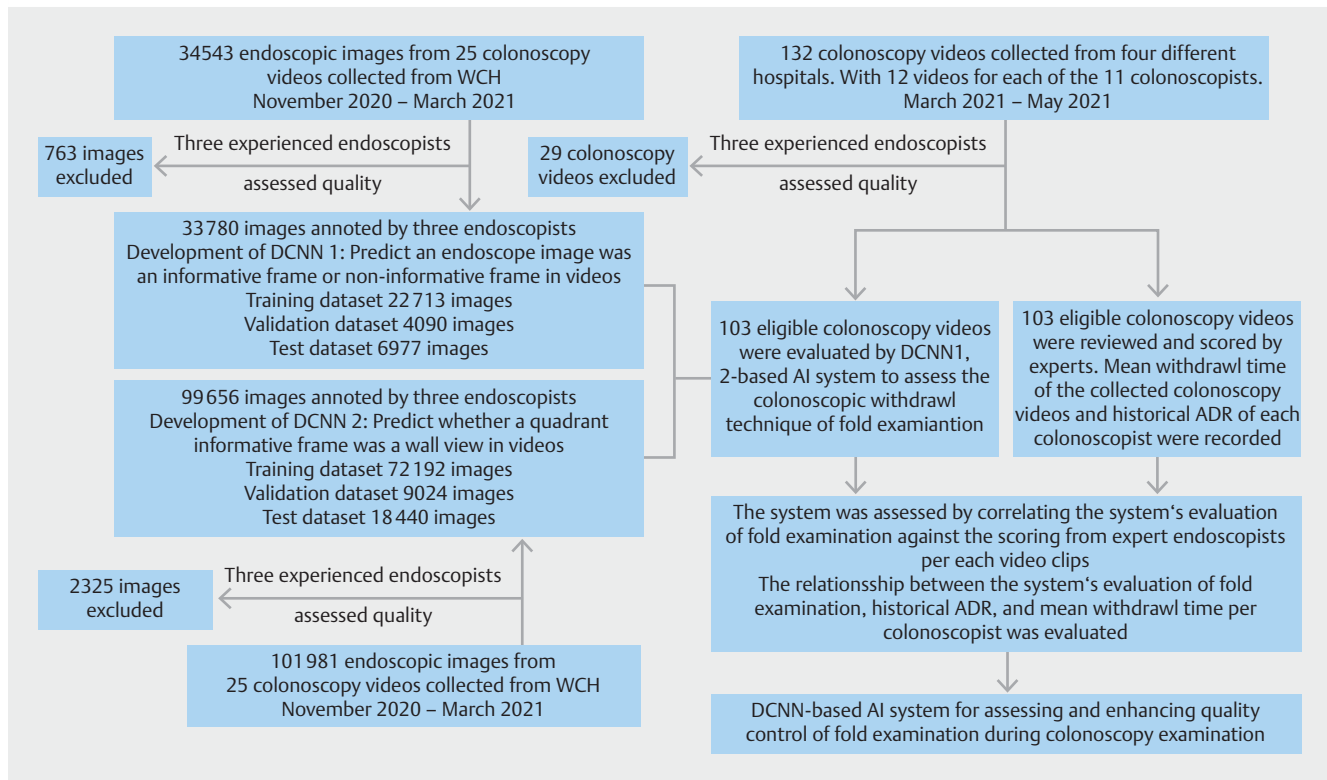
► **Fig. 1** Flowchart of the dataset for preprocessing, training, validating, and testing of the system. WCH, West China Hospital.

Participating colonoscopists

We recruited 11 colonoscopists from four different medical centers (West China Hospital contributed one colonoscopist; Cangxi People's Hospital contributed three colonoscopists; Dazhou Integrated Traditional Chinese and Western Medicine Hospital contributed three colonoscopists; Nanchong Central Hospital contributed four colonoscopists). Each colonoscopist had performed at least 500 annual screening colonoscopies in the 5 years preceding the study. A study investigator organized the recording of at least 12 consecutive screening colonoscopies (3 March 2021 to 11 May 2021) performed by each colonoscopist; colonoscopists were unaware that their colonoscopies were being video recorded. Colonoscopies including inflammatory bowel disease, polyposis syndrome, CRC, and colonoscopies with a Boston Bowel Preparation Score <6 were excluded from the selection. For each colonoscopy, withdrawal time was defined as the time for visual examination of the colon from the cecum to the rectum, excluding any time spent performing endoscopic treatment or biopsy. The mean withdrawal time for each colonoscopist was calculated as the average withdrawal time of all collected videos for each colonoscopist. ADR was defined as the proportion of screening colonoscopies with at least one adenoma detected. The historical ADR for each colonoscopist was calculated using 12-month historical data (1 May 2020 to 30 April 2021) of screening colonoscopies performed by each colonoscopist (► **Fig. 1**, ► **Table 1**).

Expert FEQ scoring

The criteria for FEQ scoring developed by Duloy et al. [9] were used to further assess the FEQ: score 0 = very poor performance (not looking behind any folds, "straight pull-back" technique), 1 = poor, 2 = fair, 3 = good, 4 = very good, and 5 = excellent performances (looking behind all folds to allow for ideal mucosal visualization). Each withdrawal video was reviewed and independently evaluated by three experts (ADR of >40 %, total number of colonoscopies performed >10 000 cases) using a blind method. Five different colon segments were scored and calculated (from 0 to 25) based on the adequacy of fold examination (cecum, appendiceal orifice, or ileocecal valve; ascending colon; transverse colon; descending colon; and sigmoid or rectum; **Table 1 s**). To standardize the review process, five videos were randomly selected and simultaneously reviewed by all the three experts; the raters subsequently discussed score variation to arrive at a consensus of scoring criteria. Each video was then reviewed and independently scored by three experts, and the whole-colon FEQ score of each video was defined as the average score of the three experts. The mean whole-colon FEQ score from experts for each colonoscopist was calculated as the average whole-colon FEQ score of the collected videos of each colonoscopist.
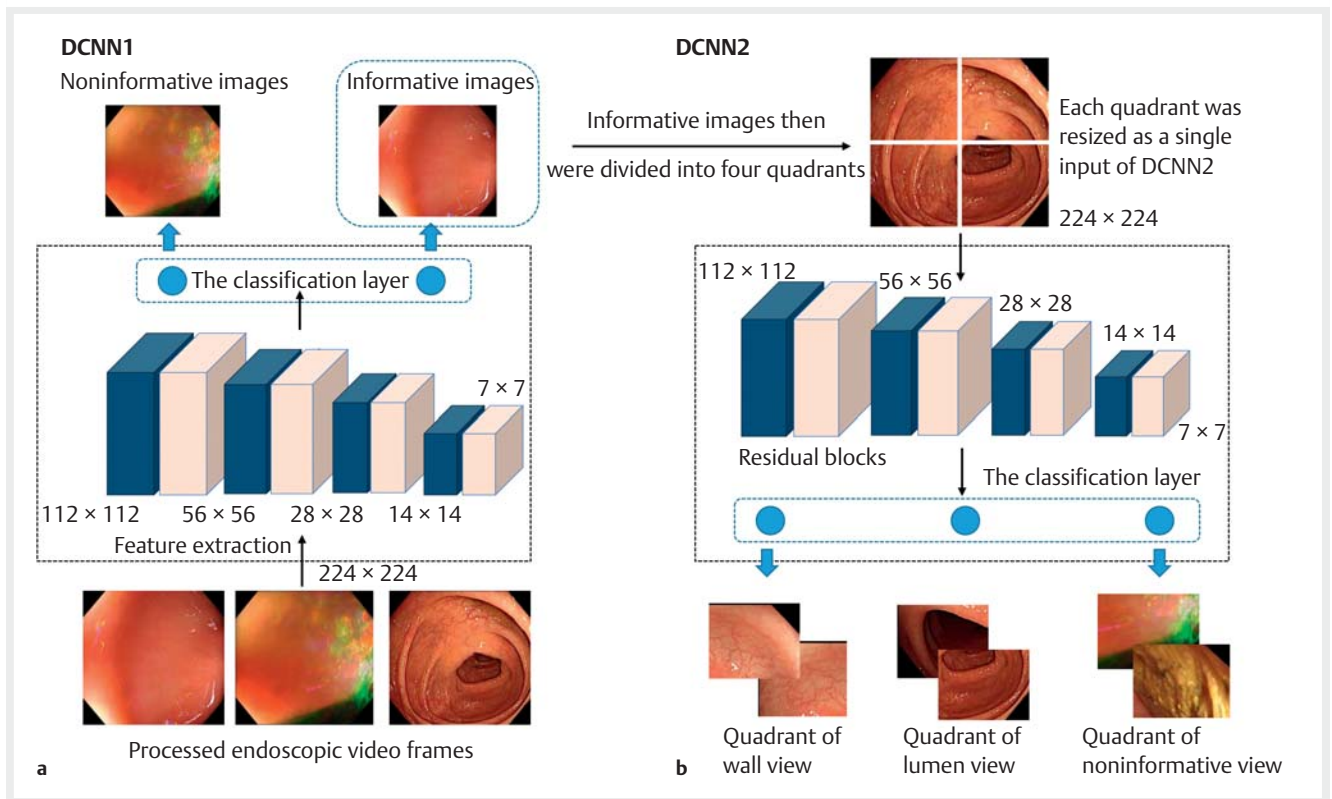
▶ **Fig. 2** The GINets architecture. The blue cubes represent the residual down-sampling convolutional layers and the pink cubes denote the bottleneck residual blocks. The number of neurons in the classification layer denote the number of categories in this task. There were two subnetworks in this architecture: the DCNN1 subnetwork (**a**) was used to identify informative images, and the DCNN2 (**b**) was used for recognition of lumen view, wall view, or noninformative view, based on the output of the previous subnetwork. The final results of DCNN2 can predict the quadrant of lumen view for each video clip.

▶ **Table 1** Colonoscopists' characteristics in different adenoma detection rate groups.

Characteristics	ADR <25% (n=5)	ADR ≥25% (n=6)	P value[1]
ADR[2], median (range), %	20.0 (18.0–24.0)	35.0 (28.0–48.0)	0.004
Colonoscopist age, median (range), years	43.0 (37.0–50.0)	38.5 (35.0–48.0)	0.25
Endoscopy experience, median (range), years	7.0 (6.0–13.0)	11.0 (8.0–15.0)	0.05

ADR, adenoma detection rate.
[1] P value, Mann–Whitney U test.
[2] The ADR was calculated based on 12-month historical data (1 May 2020 to 30 April 2021; the 12-month historical data of each colonoscopist not shown) of screening colonoscopies performed by each colonoscopist.

AI system FEQ evaluation and relationship with expert FEQ scores and ADR

All the collected colonoscopy videos were assessed by the system to obtain the AI system's evaluation of FEQ. We further assessed the system by correlating the system's evaluation of FEQ against the whole-colon expert FEQ score for each video. The AI system's mean evaluation of FEQ for each colonoscopist was calculated as the average system's evaluation of FEQ of the collected videos of each colonoscopist. Then, we evaluated the relationship between the AI system's mean evaluation of FEQ and mean whole-colon expert FEQ score, historical ADR, and mean withdrawal time of each colonoscopist.

Comparison of expert FEQ scores and ADR of endoscopists with low and high AI system FEQ evaluation

In order to compare the expert FEQ scores and ADRs of endoscopists with low and high AI system FEQ evaluation, based on the average value of AI FEQ evaluation of each colonoscopist, two groups were established: lower FEQ group and higher FEQ group. Data analysis was performed by comparing the mean score from experts and historical ADR of each colonoscopist in the two groups.

Evaluation FEQ improvement when using the AI system

The 11 colonoscopists were divided into two groups: low ADR group (ADR<25%) and high ADR group (ADR≥25%) (▶ **Table 1**). Each colonoscopist then performed six consecutive screening colonoscopies (three of the six colonoscopies were randomly performed with the assistance of the AI system, and the remaining colonoscopies were performed without AI assistance) between 27 May 2021 and 11 June 2021. During the experiment, the AI system was integrated into the endoscopy model to process endoscopy frames of the video stream synchronously. All 11 colonoscopists were aware of the function of the AI system to measure the FEQ during colonoscopic withdrawal procedures, and whether or not the AI system was opened because the AI system's evaluation of FEQ was presented on the endoscopy monitor when it was activated. In the AI-assisted group, the system was opened manually when the endoscope reached the cecum. In addition to the original videos, the AI system's evaluation of FEQ was presented on the monitor, providing real-time feedback for each colonoscopist during withdrawal procedures (**Fig. 2 s**). No information of FEQ evaluation was presented on the monitor in the control group. When the experiment was finished, each video was reviewed and scored by three experts. Meanwhile, each video was evaluated by the AI system. We then evaluated whether the system's performance could improve the system's evaluation of FEQ and scores of FEQ from experts in low ADR group and high ADR group endoscopists (**Fig. 3 s**).

Outcomes

The main outcomes were the system's evaluation of FEQ and the expert scores of FEQ for each colonoscopy withdrawal video, and the mean system's evaluation of FEQ and mean whole-colon expert FEQ score for each colonoscopist. Secondary outcomes included the historical ADR and mean withdrawal time of each colonoscopist.

Ethics

The study protocol was approved on 10 November 2020 by the Ethics Committee of the West China Hospital, Sichuan University.

Statistical analysis

The primary analysis of the relationship between the system's evaluation of FEQ and experts' scores of each video was performed using Pearson's correlation analysis. A linear mixed model for repeated measures with a random term to evaluate the institution variable was used. The whole-colon expert FEQ score was defined as the dependent variable representing the quality of fold examination of each video, the AI system's evaluation of FEQ was the fixed effect, and the institutions were the random effect. The mean system's evaluation of FEQ and mean whole-colon expert FEQ score for each colonoscopist were averaged based on the collected videos of each colonoscopist. Pearson's correlation analysis was also used to analyze the relationship between the mean system's evaluation of FEQ and mean FEQ scores of experts, historical ADR, and mean withdrawal time of each colonoscopist.

The sample size was calculated based on analyses of the relationship between the system's evaluation of fold examination and the scores from expert endoscopists for each colonoscopist by using PASS 15.0 (NCSS Statistical Software, Kaysville, Utah, USA). Assuming $\alpha=0.05$ and power=0.90 is calculated by Pearson's correlation tests (r=0.8), we estimated a necessary sample size of 11 endoscopists, with at least 12 recordings (dropout rate 20%) to detect a 4-point difference in whole-colon FEQ score for each endoscopist (a standard deviation in FEQ score of 3). If the data followed a normal distribution, the variables were expressed as mean (SD). When the data did not follow a normal distribution, the variables were expressed as the median within a range or interquartile range (IQR), and Mann–Whitney U test was used when the data did not follow a normal distribution. The point estimates for rates were presented with 95%CIs. $P<0.05$ was set as the statistical significance. Statistical analyses were conducted using SPSS, version 22.0 (IBM Corp., Armonk, New York, USA).

Results

Colonoscopists' characteristics

The 11 participating colonoscopists from four different hospital endoscopy units were divided into an ADR<25% group and an ADR≥25% group, with median ADRs of 20.0% (range 18.0%–24.0%) and 35.00% (range 28.0%–48.0%; $P=0.004$), respectively. The age and endoscopy experience of the colonoscopists were not statistically different between the two groups (▶ **Table 1**).

AI system FEQ evaluation and relationship with expert FEQ scores and ADR

A total of 103 videos were graded. Based on each colonoscopy withdrawal video, the mean whole-colon expert FEQ score was 14.98 (SD 2.96), the AI system's mean evaluation of FEQ was 0.36 (SD 0.09), and the mean withdrawal time was 6.79 (SD 3.85) minutes. Based on each colonoscopist, the mean whole-colon expert FEQ score was 14.63 (SD 2.90), the AI system's mean evaluation of FEQ was 0.35 (SD 0.07), and the mean withdrawal time was 6.40 (SD 2.69) minutes. The AI system's evaluation of FEQ was significantly correlated with whole-colon expert FEQ score based on each colonoscopy withdrawal video (r=0.706, $P<0.001$) (▶ **Fig. 3**). Potential additional variation was present owing to the 103 colonoscopy videos being performed by 11 endoscopists from four different institutions, which may introduce statistical dependencies. During the process of evaluating the institution variable, we defined the whole-colon expert FEQ score as the dependent variable representing the quality of fold examination, the AI system's evaluation of FEQ as the fixed effect, and the four institutions as the random effect. The results of the mixed effect model showed that the AI system's evaluation of FEQ was significantly correlated with whole-colon expert FEQ scores even after controlling for the random effect of different institutions ($P=0.007$) (**Table 2 s**). The mean FEQ of colonoscopy withdrawal videos as-
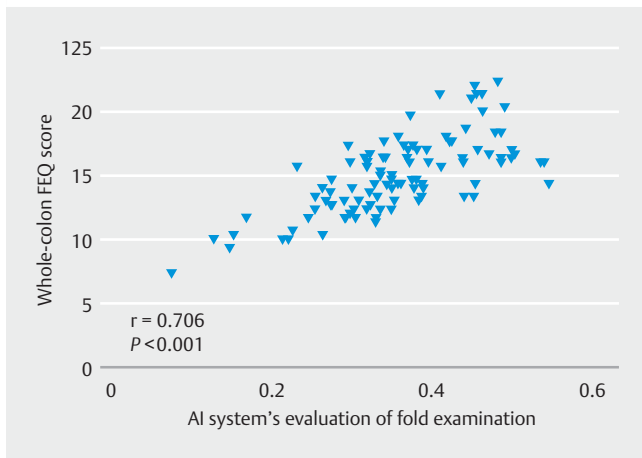
▶ **Fig. 3** Correlations between the artificial intelligence (AI) system's evaluation of fold examination quality (FEQ) and whole-colon FEQ from experts of each video. The AI system's evaluation was significantly associated with whole-colon FEQ (r = 0.706; *P* < 0.001, Pearson's correlation analysis) for each video clip (n = 103).

▶ **Table 2** Correlations between the artificial intelligence system's evaluation and mean whole-colon expert fold examination quality score, historical adenoma detection rates, and mean withdrawal time per colonoscopist.

Characteristics	Pearson's correlation	95 %CI	P value*
AI system evaluation			
▪ Whole-colon expert FEQ score	0.871	0.673–1.000	<0.001
▪ Historical ADR	0.852	0.642–1.000	0.001
▪ Withdrawal time	0.727	0.463–0.990	0.01

AI, artificial intelligence; FEQ, fold examination quality; ADR, adenoma detection rate.
* *P* value, Pearson's correlation analysis.

sessed by the AI system was significantly correlated with mean whole-colon expert FEQ score (r = 0.871, *P* < 0.001), historical ADR (r = 0.852, *P* = 0.001), and mean withdrawal time (r = 0.727, *P* = 0.01) of each colonoscopist (▶ **Table 2**, **Table 3 s**).

Comparison the expert FEQ scores and ADR of endoscopists with low and high AI system FEQ evaluation

The difference between the mean expert FEQ scores (12.71 [IQR 10.93–13.69] vs. 16.22 [IQR 14.81–17.94]) of lower (n = 5) and higher (n = 6) FEQ groups, determined by AI system evaluation, was statistically significant (*P* = 0.01). The difference between colonoscopists' median historical ADRs in these two groups (20.0 % [range 18.0 %–28.0 %] vs. 35.0 % [range 24.0 %–48.0 %]) was also statistically significant (*P* = 0.02) (**Fig. 4 s**, ▶ **Video 1**).
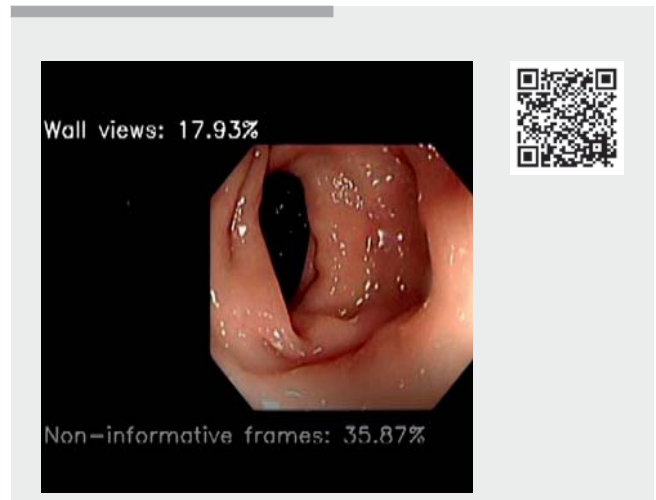
▶ **Video 1** Clip 1 Colon segment with poor fold examination. Clip 2 Colon segment with good fold examination. Online content viewable at: https://doi.org/10.1055/a-1799-8297

Evaluation of FEQ improvement when using the AI system

The 11 participating colonoscopists were divided into low ADR (ADR < 25 %, n = 5) and high ADR (ADR ≥ 25 %, n = 6) groups. Each colonoscopist performed six consecutive screening colonoscopies, resulting in 66 colonoscopies for analysis; 30 colonoscopies (15 in the control group, 15 in the AI-assisted group) were included in the lower ADR group, and 36 colonoscopies (18 in the control group, 18 in the AI-assisted group) were included in the higher ADR group. For colonoscopies performed in the low ADR group, AI assistance significantly improved the median FEQ evaluated by both the AI system (0.29 [IQR 0.27–0.30] vs. 0.23 [IQR 0.17–0.26]; *P* < 0.001) and experts (14.00 [IQR 14.00–15.00] vs. 11.67 [IQR 10.00–13.33]; *P* < 0.001). However, for colonoscopies performed in the high ADR group, AI assistance did not significantly improve the median FEQ evaluated by either the AI system (0.41 [IQR 0.39–0.43] vs. 0.40 [IQR 0.39–0.42]; *P* = 0.44) or experts (16.00 [IQR 15.00–18.50] vs. 16.67 [IQR 14.25–17.67]; *P* = 0.67) (▶ **Table 3**).

Discussion

In the current study, we successfully developed an AI-based system to assess the colonoscopic withdrawal technique of fold examination. Our data demonstrated that the AI system had good accuracy in detecting the wall view at validation and test datasets. Strong correlations were found between the system's evaluation of FEQ and expert endoscopists' scoring of FEQ, historical ADR, and mean withdrawal time of each colonoscopist, suggesting that this system could be used to help colonoscopists improve their FEQ in clinical practice.

With the rapid progress in the development of AI in recent years, a remarkable performance of computer-aided diagnosis for detection of colorectal polyps has been achieved [5, 7, 8].

▶ **Table 3** Performance of the artificial intelligence system in enhancing colonoscopic withdrawal technique of fold examination during screening colonoscopy.

Characteristics	AI-assisted colonoscopy (n = 33)	Unassisted colonoscopy (n = 33)	P value*
Colonoscopies performed by lower-ADR colonoscopists (n = 30)	15	15	
AI system evaluation, median (IQR)	0.29 (0.27–0.30)	0.23 (0.17–0.26)	<0.001
Whole-colon expert FEQ score, median (IQR)	14.00 (14.00–15.00)	11.67 (10.00–13.33)	<0.001
Colonoscopies performed by higher-ADR colonoscopists (n = 36)	18	18	
AI system evaluation, median (IQR)	0.41 (0.39–0.43)	0.40 (0.39–0.42)	0.44
Whole-colon expert FEQ score, median (IQR)	16.00 (15.00–18.50)	16.67 (14.25–17.67)	0.67

AI, artificial intelligence; ADR, adenoma detection rate; IQR, interquartile range; FEQ, fold examination quality.
* P value, Mann–Whitney U test.

Zhou et al. and Su et al. have also proposed that a DCNN could evaluate bowel preparation, thus improving quality control performance [6, 7]. Furthermore, in combination with computer-aided diagnosis, Su et al. and Gong et al. proposed that their systems could be applied to further evaluate withdrawal speed and withdrawal stability, which can also help enhance colonoscopy quality [7, 8]. However, none of the existing techniques were proposed to measure the fold examination performance of colonoscopists during colonoscopy. Stanek et al. proposed that the four-quadrant analysis technique could be used for counting the withdrawal spiral motions of the endoscope by using a semi-automated technology based on default configuration parameters and default graph architecture [10]. However, our proposed method is based on deep neural networks for automatically learning feature representation by using the back-propagation algorithm to analyze mucosal inspection; the performance and stability of neural networks are better than traditional machine learning. Furthermore, we evaluated the performance of our model combined with clinical evaluation, which provided us with more convincing evidence.

Previous studies have suggested that the existence of blind spots is one of the most common risk factors for a missed diagnosis of colorectal polyps when performing colonoscopy [3]. Colonoscopic blind spots are also related to intraprocedural colonoscopy quality indicators [11]. The lack of mucosal fold examination, poor bowel preparation, inadequate luminal distension, and rapid colonoscopy withdrawal time have been considered among the most common factors leading to colonoscopic blind spots [12]. A variety of tools are available to assist endoscopists in exposing more mucosa during the inspection, including distal attachment exposure devices and new types of colonoscopes [13–15]. However, adequate inspection of all mucosal folds and flexures by rotating the endoscope is one of the basic and most common colonoscopic withdrawal techniques for reducing missed diagnoses [16]. However, measuring the FEQ in clinical practice may be challenging. In the current study, we established and validated an AI system for assessing the FEQ. The results showed that there was a strong correlation between the AI system's mean evaluation of FEQ and mean experts' scores of FEQ of each colonoscopist. Use of this AI system might reduce the blind areas during colonoscopy by increasing the colonoscopist's manipulation of the endoscope for more complete inspection of mucosal folds, and by enabling real-time quality analysis of colonoscopy examination and feedback during procedures.

Although ADR is highly variable among endoscopists, it is regarded as an important indicator of colonoscopy performance quality and is inversely associated with the incidence of interval CRC [17–19]. In addition, it is believed that low ADR is an indirect measure of an inadequate examination of the colon [13, 20]. Thus, we further analyzed the relationship between the AI system's mean evaluation of each video according to the colonoscopist's ADR, and found that the AI-based evaluation was significantly associated with colonoscopists' ADRs. Short withdrawal time has also been proposed as a reflection of poor examination technique [17, 21]. Previous studies have demonstrated that a withdrawal time of ≥6 minutes is associated with higher detection of neoplastic lesions during colonoscopy in patients with intact colons and reduces the risk of interval cancers [21, 22]. Additional analysis in the current study showed that the AI system's mean evaluation was significantly associated with the mean withdrawal time of each colonoscopist. These data suggest that our system may improve withdrawal time and ADR. Moreover, we analyzed the role of the system in improving FEQ among different groups of colonoscopists. Although it was demonstrated that the AI system did not enhance the fold examination in the group of endoscopists with high ADR, it is suggested that the system could significantly improve fold examination in endoscopists with low ADR.

This study has a few limitations. First, previous studies reported that examination technique mainly contains four components: looking behind all folds, cleaning residual stool, providing adequate bowel distension, and withdrawing. However, in this study, we only assessed the fold examination performance of our system. Second, we conducted correlation analysis only for the system's FEQ evaluation against that of experts, and against colonoscopists' withdrawal time and ADR. However, we did not determine a relevant threshold of the system's evaluation of FEQ. Third, we only used colonoscopy videos and conducted a prospective observational study with small sample size to assess the performance of the system. Thus, more randomized controlled trials should be performed to further verify whether the system can improve FEQ and to determine its impact on ADR in clinical practice.

In conclusion, our AI-based system was successfully constructed to calculate FEQ during the withdrawal phase. By providing real-time feedback, our system can enhance the awareness of adequate FEQ, which is crucial to ensure quality control in endoscopy practice.

Acknowledgments

Funding

Competing interests

The authors declare that they have no conflict of interest.

References

[1] Leufkens AM, van Oijen MG, Vleggaar FP et al. Factors influencing the miss rate of polyps in a back-to-back colonoscopy study. Endoscopy 2012; 44: 470–475

[2] Kaminski MF, Regula J, Kraszewska E et al. Quality indicators for colonoscopy and the risk of interval cancer. N Engl J Med 2010; 362: 1795–1803

[3] Rex DK. Colonoscopic withdrawal technique is associated with adenoma miss rates. Gastrointest Endosc 2000; 51: 33–36

[4] Freedman D, Blau Y, Katzir L et al. Detecting deficient coverage in colonoscopies. IEEE Trans Med Imaging 2020; 39: 3451–3462

[5] Wang P, Xiao X, Glissen Brown JR et al. Development and validation of a deep-learning algorithm for the detection of polyps during colonoscopy. Nat Biomed Eng 2018; 2: 741–748

[6] Zhou J, Wu L, Wan X et al. A novel artificial intelligence system for the assessment of bowel preparation (with video). Gastrointest Endosc 2020; 91: 428–435

[7] Su J-R, Li Z, Shao X-J et al. Impact of a real-time automatic quality control system on colorectal polyp and adenoma detection: a prospective randomized controlled study (with videos). Gastrointest Endosc 2020; 91: 415–424

[8] Gong D, Wu L, Zhang J et al. Detection of colorectal adenomas with a real-time computer-aided system (ENDOANGEL): a randomised controlled study. Lancet Gastroenterol Hepatol 2020; 5: 352–361

[9] Duloy A, Yadlapati RH, Benson M et al. Video-based assessments of colonoscopy inspection quality correlate with quality metrics and highlight areas for improvement. Clin Gastroenterol Hepatol 2019; 17: 691–700

[10] Stanek SR, Tavanapong W, Wong J et al. SAPPHIRE: a toolkit for building efficient stream programs for medical video analysis. Comput Methods Programs Biomed 2013; 112: 407–421

[11] Rex DK, Schoenfeld PS, Cohen J et al. Quality indicators for colonoscopy. Am J Gastroenterol 2015; 110: 72–90

[12] Lee RH, Tang RS, Muthusamy VR et al. Quality of colonoscopy withdrawal technique and variability in adenoma detection rates (with videos). Gastrointest Endosc 2011; 74: 128–134

[13] May FP, Shaukat A. State of the science on quality indicators for colonoscopy and how to achieve them. Am J Gastroenterol 2020; 115: 1183–1190

[14] Rex DK. Polyp detection at colonoscopy: endoscopist and technical factors. Best Pract Res Clin Gastroenterol 2017; 31: 425–433

[15] Thirumurthi S, Ross WA, Raju GS. Can technology improve the quality of colonoscopy? Curr Gastroenterol Rep 2016; 18: 38

[16] Moons LMG, Gralnek IM, Siersema PD. Techniques and technologies to maximize mucosal exposure. Gastrointest Endosc Clin N Am 2015; 25: 199–210

[17] Kaminski MF, Robertson DJ, Senore C et al. Optimizing the quality of colorectal cancer screening worldwide. Gastroenterology 2020; 158: 404–417

[18] Gessl I, Waldmann E, Penz D et al. Evaluation of adenomas per colonoscopy and adenomas per positive participant as new quality parameters in screening colonoscopy. Gastrointest Endosc 2019; 89: 496–502

[19] Dawwas MF. Adenoma detection rate and risk of colorectal cancer and death. N Engl J Med 2014; 370: 2539–2540

[20] Chittleborough TJ, Luck A, Boussioutas A et al. The conundrum of quality in colonoscopy. ANZ J Surg 2018; 88: 263–264

[21] Barclay RL, Vicari JJ, Doughty AS et al. Colonoscopic withdrawal times and adenoma detection during screening colonoscopy. N Engl J Med 2006; 355: 2533–2541

[22] Shaukat A, Rector TS, Church TR et al. Longer withdrawal time is associated with a reduced incidence of interval cancer after screening colonoscopy. Gastroenterology 2015; 149: 952–957