World Journal of Gastroenterology

Submit a Manuscript: https://www.f6publishing.com

World J Gastroenterol 2025 July 21; 31(27): 106819

DOI: 10.3748/wjg.v31.i27.106819

ISSN 1007-9327 (print) ISSN 2219-2840 (online)

ORIGINAL ARTICLE

Retrospective Study

Deep learning-based localization and lesion detection in capsule endoscopy for patients with suspected small-bowel bleeding

Yeong Seok Kwon, Tae Yong Park, So Eui Kim, Yehyun Park, Jae Gon Lee, Sang Pyo Lee, Kyeong Ok Kim, Hyun Joo Jang, Young Joo Yang, Bum-Joo Cho

Specialty type: Gastroenterology and hepatology

Provenance and peer review: Unsolicited article; Externally peer reviewed.

Peer-review model: Single blind

Peer-review report's classification

Scientific Quality: Grade B, Grade B

Novelty: Grade A, Grade B

Creativity or Innovation: Grade A, Grade B

Scientific Significance: Grade A, Grade B

P-Reviewer: Jamaluddin J; Mukundan A

Received: March 9, 2025

Revised: May 2, 2025

Accepted: July 1, 2025

Published online: July 21, 2025

Processing time: 135 Days and 20.4 Hours

Yeong Seok Kwon, Young Joo Yang, Department of Internal Medicine, Hallym University College of Medicine, Chuncheon-si 24253, South Korea

Tae Yong Park, So Eui Kim, Bum-Joo Cho, Medical Artificial Intelligence Center, Hallym University Medical Center, Anyang 14068, South Korea

Yehyun Park, Department of Internal Medicine, Ewha Womans University Seoul Hospital, Seoul 07804, South Korea

Jae Gon Lee, Department of Internal Medicine, Hanyang University Guri Hospital, Hanyang University College of Medicine, Guri 11923, South Korea

Sang Pyo Lee, Department of Internal Medicine, Hanyang University College of Medicine, Seoul 04763, South Korea

Kyeong Ok Kim, Division of Gastroenterology and Hepatology, Department of Internal Medicine, Yeungnam University College of Medicine, Daegu 42415, South Korea

Hyun Joo Jang, Department of Internal Medicine, Hallym University Dongtan Sacred Heart Hospital, Hallym University College of Medicine, Hwaseong 18450, South Korea

Bum-Joo Cho, Department of Ophthalmology, Hallym University Sacred Heart Hospital, Hallym University College of Medicine, Anyang 14068, South Korea

Co-first authors: Yeong Seok Kwon and Tae Yong Park.

Co-corresponding authors: Young Joo Yang and Bum-Joo Cho.

Corresponding author: Young Joo Yang, Assistant Professor, Department of Internal Medicine, Hallym University College of Medicine, 77 Sakju-ro, Chuncheon-si 25253, South Korea. yjyang@hallym.ac.kr

Abstract

BACKGROUND

Small-bowel capsule endoscopy (SBCE) is widely used to evaluate obscure gastrointestinal bleeding; however, its interpretation is time-consuming and reader-dependent. Although artificial intelligence (AI) has emerged to address these limitations, few models simultaneously perform small-bowel (SB) loca-

lization and abnormality detection.

AIM
To develop an AI model that automatically distinguishes the SB from the stomach and colon and diagnoses SB abnormalities.

METHODS
We developed an AI model using 87005 CE images (11925, 33781, and 41299 from the stomach, SB, and colon, respectively) for SB localization and 28405 SBCE images (1337 erosions/ulcers, 126 angiodysplasia, 494 bleeding, and 26448 normal) for abnormality detection. The diagnostic performances of AI-assisted reading and conventional reading were compared using 32 SBCE videos in patients with suspicious SB bleeding.

RESULTS
Regarding organ localization, the AI model achieved an area under the receiver operating characteristic curve (AUC) and accuracy exceeding 0.99 and 97%, respectively. For SB abnormality detection, the performance was as follows: Erosion/ulcer: 99.4% accuracy (AUC, 0.98); angiodysplasia: 99.8% accuracy (AUC, 0.99); bleeding: 99.9% accuracy (AUC, 0.99); normal: 99.3% accuracy (AUC, 0.98). In external validation, AI-assisted reading (8.7 minutes) was significantly faster than conventional reading (53.9 minutes; $P < 0.001$). The SB localization accuracies (88.6% *vs* 72.7%, $P = 0.07$) and SB abnormality detection rates (77.3% *vs* 77.3%, $P = 1.00$) of the conventional reading and AI-assisted reading were comparable.

CONCLUSION
Our AI model decreased SBCE reading time and achieved performance comparable to that of experienced endoscopists, suggesting that AI integration into SBCE reading enables efficient and reliable SB abnormality detection.

Key Words: Artificial intelligence; Obscure gastrointestinal bleeding; Capsule endoscopy; Detection; Localization; Small bowel

©The Author(s) 2025. Published by Baishideng Publishing Group Inc. All rights reserved.

Core Tip: Small-bowel capsule endoscopy (SBCE) is essential for diagnosing obscure gastrointestinal bleeding (OGIB), but its analysis is time-intensive and dependent on the reader's expertise. Despite advancements in artificial intelligence (AI), few models combine accurate small-bowel (SB) localization and abnormality detection. We developed an AI model that automatically distinguishes the SB from the stomach and colon and diagnoses SB abnormalities such as erosions/ulcers, angiodysplasia, and bleeding in patients with suspected OGIB. Our AI model significantly decreased SBCE reading time compared to that of conventional reading with comparable performance for SB abnormality detection, demonstrating the efficiency and reliability of AI integration into SBCE reading.

Citation: Kwon YS, Park TY, Kim SE, Park Y, Lee JG, Lee SP, Kim KO, Jang HJ, Yang YJ, Cho BJ. Deep learning-based localization and lesion detection in capsule endoscopy for patients with suspected small-bowel bleeding. *World J Gastroenterol* 2025; 31(27): 106819
URL: https://www.wjgnet.com/1007-9327/full/v31/i27/106819.htm
DOI: https://dx.doi.org/10.3748/wjg.v31.i27.106819

INTRODUCTION

Small-bowel capsule endoscopy (SBCE) is widely used to investigate SB abnormalities, including erosions/ulcers, angiodysplasia, and tumors, through direct visualization of the small-bowel (SB) mucosa[1,2]. As SBCE is noninvasive, it is a first-line modality for identifying the bleeding focus associated with obscure gastrointestinal bleeding (OGIB), which is defined as gastrointestinal bleeding of unknown origin with negative upper and lower endoscopy results, with a diagnostic yield of 38%-83%[3,4]. Additionally, SBCE enables localization of the bleeding focus to guide optimal enteroscopic approaches for therapeutic intervention. However, each SBCE video contains approximately 50000 frames; therefore, a thorough review requires 30 to 120 min[5,6]. Clinically relevant abnormalities can appear in only a few frames. Therefore, completing the review without missing abnormalities is a challenge for physicians. Additionally, SB abnormality detection depends strongly on the physician's experience[7]. Experts recommend a second review by an experienced reviewer, particularly for negative SBCE cases, to reduce interobserver variability[8]. The tedious and time-consuming nature of SBCE reading, as well as the interpretation variability, hinder diagnostic performance, strain human resources, and may yield adverse patient outcomes.

Artificial intelligence (AI) for endoscopic imaging analyses in gastroenterology has advanced markedly[9,10]. AI models have been developed to reduce the burden of SBCE image reading, improve diagnostic accuracy, and allow consistent performance[11]. Recently, relevant AI-based research has shifted from single to multiple SB-abnormality detection[12-14]. Additionally, various AI systems have been developed for other purposes, such as assessing cleaning scores and distinguishing the SB from the stomach or colon[15-17]. Finally, the clinical benefits and feasibility of AI-assisted SBCE reading compared to conventional reading have been investigated[18-20].

However, most AI models have been trained and validated on the single task of SBCE reading. Comprehensive AI models performing SB localization and multiple abnormality detection similar to conventional reading are rare. Moreover, most AI systems have been validated using captured static images, which limits their direct clinical implementation. Therefore, this study presents a convolutional neural network (CNN) model capable of automatically distinguishing the SB from the stomach and colon and diagnosing multiple SB abnormalities (erosions/ulcers, angiodysplasia, and bleeding). We specifically focus on patients who underwent SBCE for suspected OGIB, as this is a common and clinically relevant indication for SBCE in real-world practice. Using complete SBCE videos from those patients, we evaluate the AI models performance for SB localization and abnormality detection and compare the diagnostic performance of AI-assisted reading with the results of conventional readings by experienced endoscopists.

MATERIALS AND METHODS

Training data collection

We retrospectively collected 101 full-length SBCE videos from 101 patients (MiroCam MC 1200 and MC 2000; Intromedic Co., Ltd., Seoul, Korea) at Chuncheon Sacred Heart Hospital (n = 81) and Dongtan Sacred Heart Hospital (n = 20), recorded between 2011 and 2022. An experienced gastroenterologist (Young Joo Yang) reviewed all SBCE videos and extracted 87005 images of the stomach (n = 11925), SB (n = 33781), and colon (n = 41299) (Supplementary Table 1). Among the SB images, 28279 were further classified into four categories: Erosions/ulcers (n = 1337), angiodysplasia (n = 126), bleeding (active bleeding or visible fresh blood, n = 494), and normal (including debris, bile, and bubbles, n = 26322) (Supplementary Table 2). To develop the CNN model, 87005 images were divided into training, tuning, and test datasets at a 7:1:2 ratio for localization, and 28279 images were divided into training, tuning, and test datasets at an 8:1:1 ratio for abnormality classification. No SB images overlapped in the datasets. This study was approved by the Institutional Review Board of Chuncheon Sacred Heart Hospital (IRB No. 2018-05) and performed in accordance with the Declaration of Helsinki.

Deep learning-based organ localization

To distinguish the stomach, SB, and colon, we used a CNN with DenseNet161 for organ localization[21]. DenseNet161, selected through an experimental analysis, is known for its dense connectivity, which enhances the gradient propagation and parameter efficiency. After pretraining on the ImageNet data set, the model was trained with data augmentation on data resized to 224 × 224 pixels. The learning rate begin at 1e-5, and was reduced by 0.1 if the validation loss did not decrease for 20 epochs. Training was conducted using an Ubuntu 20.04.5 server with an Intel Xeon Silver 4216 CPU, 256GB RAM, and an NVIDIA GeForce RTX 3090.

For post-processing, CNN predictions were listed sequentially along the temporal axis of the video and smoothing and noise reduction using a Savitzky–Golay filter[22]. Then, we defined windows, slid them over the graph to calculate the averages, and assigned class labels based on the highest average for each window. To identify SB and colon entry points, we searched for continuous segments exceeding a set time threshold (5 minutes); an entry point was defined as the first frame in which the SB or colon class was identified.

Dual-camera SBCE (MC 2000; Intromedic Co., Ltd., Seoul, Korea) generated left- and right-side videos. Discrepancies between the predicted entry points for these two videos were used for error correction. If the time differences in the identified regions exceeded 10 minutes, binary classification models (stomach *vs* SB and SB *vs* colon) were used for repeat error-region predictions. Finally, the predictions were integrated using the unweighted average ensemble technique.

Deep learning-based lesion classification

To overcome the frame similarities and imbalances between normal and abnormal SBCE data, ensemble methods were used to combine individual binary classification models, ensuring robust and consistent abnormality detection. CNNs were used to develop binary models for the three following classifications: Normal *vs* ulcer, *vs* bleeding, and *vs* angiodysplasia. Based on experimental analysis, DenseNet201, pretrained on the ImageNet data set, was further trained with data augmentation on images resized to 320 × 320 pixels, using a batch size of 32 for 100 epochs and the Adam optimizer with a learning rate of 1e-4[21]. For ulcers and angiodysplasia, the learning rate was reduced by 0.1 after 10 epochs without a decrease in validation loss using the cross-entropy loss. For bleeding, a cosine-shaped learning rate and 9:1 weighted cross-entropy loss were used. These binary classifiers were organized into a two-step architecture: The first step detects whether a frame is abnormal, and the second step determines the specific lesion type for abnormal frames.

To improve the SBCE video analysis, the prediction outcomes of binary classification models were first arranged along the time axis and then smoothed using a Savitzky–Golay filter for noise reduction[22]. The results were combined using a two-step ensemble method: Frames labeled as abnormal by any model were classified as abnormal; the highest-probability class was chosen as the final label. To further address noise and irregularities, the ensemble results were refined using a window mask of n frames. If multiple different lesions were detected within the window mask, they were considered as candidates for correction. Abnormalities within the current window mask were confirmed and compared

with the surrounding masks to update the information based on class counts and probability values. Note that consistent observation of the same abnormality across multiple frames increases certainty, ensuring consistent tracking and enhanced model accuracy and reliability. The window mask size and threshold were experimentally set to 5 and 0.5, respectively. The CNN model is outlined in Figure 1.

External validation

To evaluate the clinical usefulness of the CNN model, we conducted external validation using 32 full-length SBCE videos of patients with suspected SB bleeding at Dongtan Sacred Heart Hospital (n = 15), Yeungnam University Medical Center (n = 13), and Ewha University Medical Center (n = 4) between January 2021 and August 2022. None of these images were included in the training dataset. For external validation, we generated a summary version of SBCE video (SBCE by CNN) comprising three parts: A video segment recorded for 3 minutes before and after the first SB image, video segments showing CNN model-identified SB abnormalities, and a video segment recorded for 3 minutes before and after the first colon image. If the CNN model determined that the colon was not reached, it recorded the last 3 minutes of the SBCE.

Outcome measures

The outcome measures were the SBCE reading time and diagnostic performance for SB localization and SB abnormality detection, which were assessed per patient on the external validation data set using both AI-assisted reading and conventional reading. All gastroenterologists (Jang HJ, Lee SP, Lee JG, Kim KO, Park Y, and Yang YJ) involved in this study, including those who participated in the AI-assisted reading, conventional reading, and ground truth establishment, had more than 10 years of clinical experience in gastroenterology, with subspecialty expertise in lower gastrointestinal diseases, and had interpreted more than 200 SBCE cases prior to participating in the study. The AI-assisted reading time was defined as the total time required for a gastroenterologist to review the CNN-provided SBCE summary (SBCE by CNN) only and make a final diagnosis. The CNN-generated SBCE summary video displayed the anatomical location of the SBCE using the original SBCE frames, and showed small-bowel abnormalities with both the original frames and the corresponding Gradient-Weighted Class Activation Mapping (Grad-CAM)-based heatmaps, with the associated probability scores displayed below each image (Supplementary material, Video). This layout allowed readers to intuitively assess the AI predictions by providing both the anatomical location and confidence level for each detected lesion. To facilitate efficient and flexible review, the video playback speed was adjustable at the reader's discretion. Although the AI system provided diagnostic suggestions, all final diagnoses were independently made by the reviewing gastroenterologists, who could accept or override the AI findings based on their clinical judgment. The conventional reading time was defined as the total time required for the gastroenterologist to review the SBCE videos without AI assistance at a playback speed of less than 15 frames per second. Each SBCE case was independently evaluated by different gastroenterologists for the conventional and AI-assisted readings, with readers blinded to each other's findings.

To establish the ground truth for the time of the first SB and colon image, as well as for SB abnormality detection, an experienced gastroenterologist (Yang YJ) independently reviewed the full-length SBCE videos, incorporating results from both AI-assisted and conventional readings. In cases where findings were ambiguous (*e.g.*, for unclear lesion location, difficulty in lesion classification, or poor image quality), the final ground truth was determined through discussion and consensus with another experienced gastroenterologist (Jang HJ or Lee SP) who was not involved in the AI-assisted or conventional readings of that particular case. The SB localization was considered correct if the predicted location was within 3 minutes of the ground truth. The SB abnormality detection score was defined as the proportion of correctly detected abnormality types compared to the ground truth.

Statistical analysis

We estimated the area under the receiver operating characteristic curve (AUC) and calculated the accuracy, sensitivity, specificity, and positive and negative predictive values using standard formulas. Continuous and categorical variables were expressed as mean ± SD and percentages with a 95%CI, respectively. Categorical variables and AUC values were compared using the Fisher exact test and DeLong test, respectively. $P < 0.05$ (two-tailed) was considered statistically significant. All analyses were performed using SPSS (version 24.0; SPSS Inc., Chicago, IL, United States) and MedCalc version 19.1 (MedCalc Software, Ostend, Belgium).

RESULTS

CNN model performance for SB localization and abnormality detection on internal test dataset

The internal test dataset comprised 21524 images from 17 patients. The AI model showed good performance for anatomical localization, with AUCs exceeding 0.99 for the stomach, SB, and colon, and organ differentiation accuracy exceeding 97% (Table 1). Specifically, the sensitivity and specificity were 98.7%, and 99.6%, respectively, for stomach localization; 96.1% and 97.9%, respectively, for the SB; and 97.1% and 97.8%, respectively, for the colon. For detection of SB abnormalities using ensemble methods, the AUCs for erosions/ulcers, angiodysplasia, bleeding, and normal detection all exceeded 0.98, with detection accuracy rates of 99.4%, 99.8%, 99.9%, and 99.3%, respectively (Figure 2 and Table 1). The AI model also achieved high sensitivity (90.1% for erosions/ulcers, 100% for angiodysplasia, and 100% for bleeding) and specificity (99.7% for erosions/ulcers, 99.8% for angiodysplasia, and 99.9% for bleeding).

Table 1 Diagnostic performance of artificial intelligence model for localization and lesion detection methods on the test dataset

	Diagnostic performance (%) (95%CI)					
	Accuracy	Sensitivity	Specificity	PPV	NPV	AUC
Localization						
Stomach	99.4 (99.4-99.4)	98.7 (98.6-98.7)	99.6 (99.6-99.6)	98.6 (98.6-98.6)	99.6 (99.6-99.7)	100 (100.0-100.0)
Small-bowel	97.4 (97.4-97.4)	96.1 (96.1-96.1)	97.9 (97.9-97.9)	94.7 (94.7-94.7)	98.5 (98.5-98.5)	99.5 (99.5-99.5)
Colon	97.5 (97.4-97.5)	97.1 (97.1-97.1)	97.8 (97.8-97.8)	97.9 (97.9-97.9)	97.0 (97.0-97.0)	99.6 (99.6-99.6)
Detection of small-bowel abnormalities						
Normal	99.3 (99.3-99.3)	99.5 (99.5-99.5)	95.3 (95.2-95.5)	99.8 (99.8-99.8)	86.4 (86.5-86.9)	99.1 (99.0-99.1)
Abnormal						
Erosions/ulcers	99.4 (99.4-99.4)	90.1 (89.8-90.2)	99.7 (99.7-99.7)	89.1 (88.8-89.3)	99.7 (99.7-99.7)	98.9 (98.8-98.9)
Angiodysplasia	99.8 (99.8-99.8)	100 (100.0-100.0)	99.8 (99.8-99.8)	58.8 (64.0-65.3)	100 (100.0-100.0)	99.7 (99.7-99.7)
Bleeding	99.9 (99.9-99.9)	100 (100.0-100.0)	99.9 (99.9-99.9)	66.7 (71.9-73.6)	100 (100.0-100.0)	99.9 (99.9-99.9)

PPV: Positive predictive value; NPV: Negative predictive value; AUC: Area under receiver-operating characteristic curve.

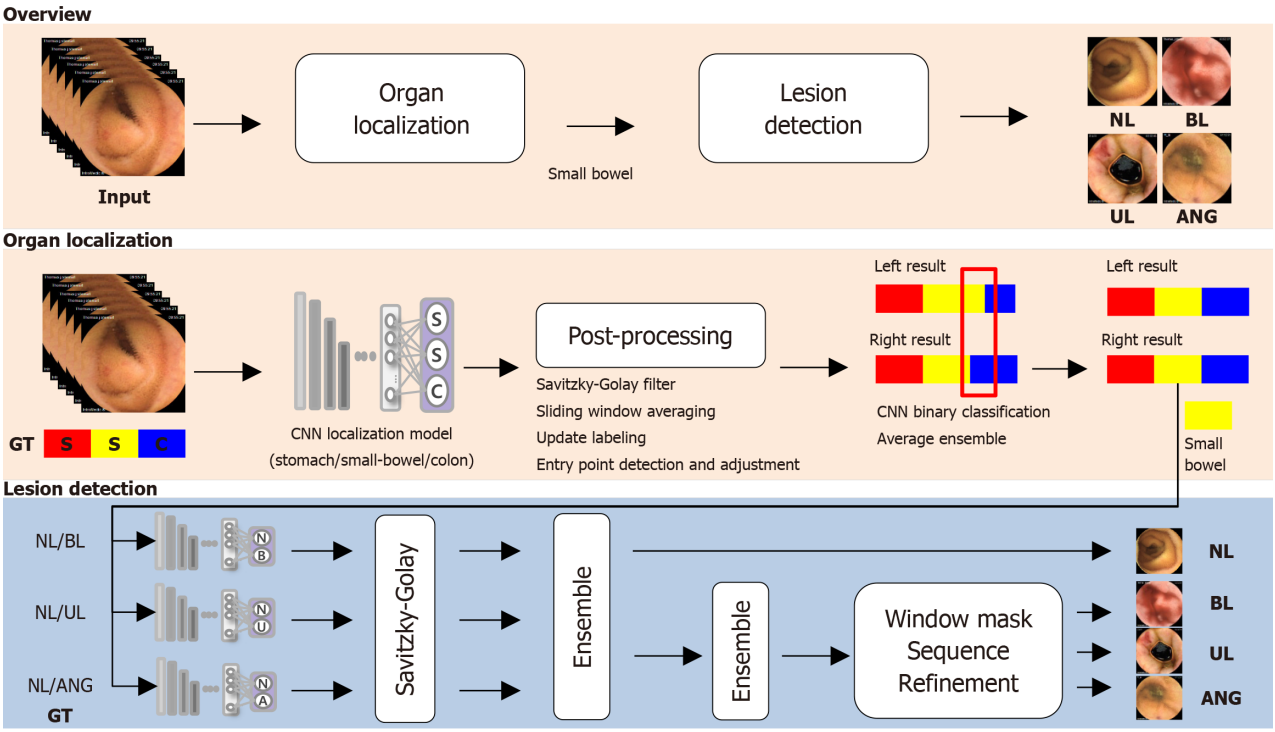


Figure 1 Overview of artificial intelligence model. The artificial intelligence system is a two-stage process that first localizes the small bowel in capsule endoscopy videos and then detects lesions such as bleeding, erosion/ulcers, and angiodysplasia. It utilizes convolutional neural network models along with post-processing and ensemble techniques to enhance accuracy and consistency across video frames. NL: Normal mucosa (including debris, bile, and bubbles); BL: Bleeding (active bleeding or visible fresh blood); GT: Ground truth; UL: Erosions/ulcers; CNN: Convolutional neural network; ANG: Angiodysplasia

Comparison of diagnostic performance between conventional reading and AI-assisted reading on external validation dataset

Thirty-two patients (mean age: 59.5 years; 68.8% male) who underwent SBCE were included in the external validation group. The indication for SBCE was suspected SB bleeding. Twenty-seven patients completed the SBCE. Capsule retention occurred in five. Among the 32 patients, seven had negative results, and the others had erosions/ulcers (n = 17), angiodysplasia (n = 9), or bleeding (n = 6). The mean AI-assisted reading time was 9.03 minutes, which was significantly shorter than the conventional reading time of 57.2 minutes (P < 0.001). The localization rates of the first SB (96.9% *vs* 87.5%, P = 0.18) and colon (81.2% *vs* 62.5%, P = 0.08) images were comparable for AI-assisted and conventional reading. The time differences between the ground truth and AI-assisted and conventional readings for the first SB image were

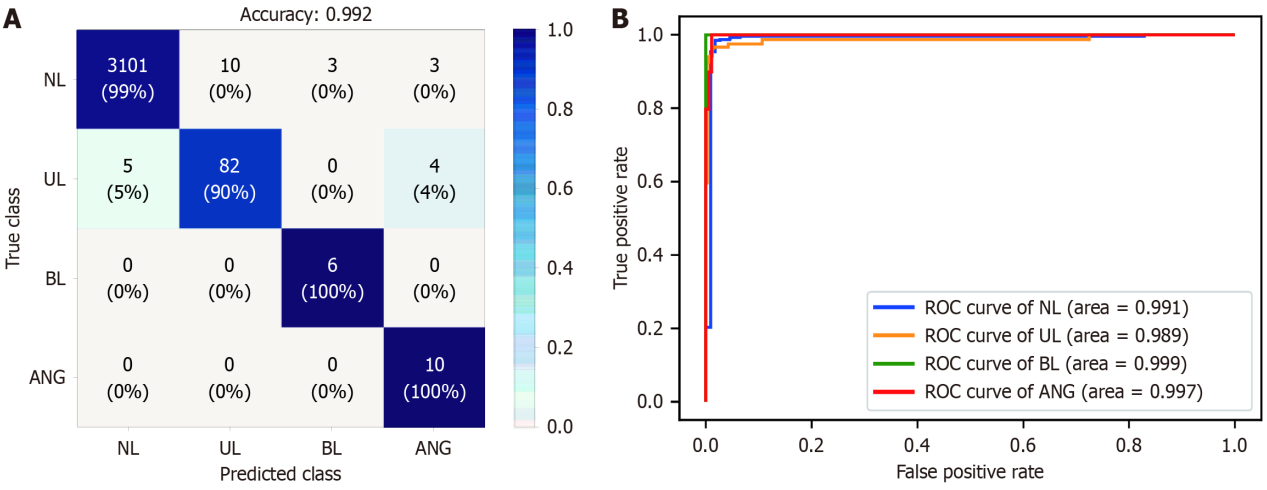


Figure 2 Performance of the artificial intelligence model for detecting small bowel abnormalities in the internal test dataset. The model demonstrated high performance in detecting small bowel abnormalities, with an overall accuracy of 0.992 and excellent area-under-the curve scores for all classes. A: Confusion matrix; B: Receiver-operating characteristic curve. NL: Normal mucosa (including debris, bile, and bubbles); BL: Bleeding (active bleeding or visible fresh blood); GT: Ground truth; UL: Erosions/ulcers, ANG: Angiodysplasia; ROC: Receiver-operating characteristic curve.

similar (100.7 ± 265.2 s *vs* 72.2 ± 303.3 s, P = 0.69). For the first colon image, the AI-assisted reading time difference exceeded that for conventional reading (622.8 ± 1132.0 s *vs* 93.8 ± 159.0 s, P = 0.03). The overall SB-abnormality detection scores for conventional and AI-assisted reading were similar (0.73 *vs* 0.75, P = 0.88) (Table 2). In the subgroup analysis, the detection of each SB abnormality type was comparable (Table 3).

Grad-CAM for false-positive or -negative SB localization and abnormality detection images

Representative Grad-CAM images of SB abnormalities correctly identified using the CNN model are shown in Figure 3. The AI-assisted and conventional readings concurred (Figure 3A), with CAM highlighting the red vascular architecture as angiodysplasia, red coloration as bleeding, and depressed lesions with yellow or whitish bases as erosions/ulcers. CNN model-detected SB abnormalities missed by conventional reading, such as those hidden by debris or shadows, are shown in Figure 3B. However, the CNN model occasionally missed or misclassified abnormalities (Figure 4). In some cases, debris was interpreted as ulcers or red spots as angiodysplasia (Figure 4A). In false-negative cases, foam hindered the detection of ulcers or angiodysplasia during the AI-assisted reading (Figure 4B).

DISCUSSION

We developed a CNN model capable of comprehensively reading SBCE images to localize SB and detect multiple SB abnormalities in patients undergoing SBCE for suspected OGIB. The AI-assisted reading time was significantly reduced compared to the conventional reading time. Comparable performance was maintained for SB segment identification, and SB abnormalities were detected in patients with OGIB. To the best of our knowledge, this is the first CNN model that mimics clinical SBCE readings to localize SB and detect abnormalities. Although numerous AI systems have detected various SB abnormalities with good diagnostic performance and reduced SBCE reading time, those systems performed only a single reading process to detect SB abnormalities and utilized static SBCE images to assess the performance[12-14,18,19,23,24]. Recently, an AI model performed multiple SBCE reading processes, including lesion detection, landmark identification, and bowel preparation assessment, thereby achieving significant time-saving and performance similar to that of conventional reading[20]. However, the AI-assisted reading was performed using static images. In contrast, we used summary version of SBCE videos to track suspected lesions, even when complicated by debris or fecal material, thereby improving diagnostic performance and demonstrating the practical utility of the model.

Several AI models to distinguish the SB from the stomach or colon have been developed[15,16]. Son *et al*[15] demonstrated a CNN model with 99.8% accuracy for binary SB classification (SB or not) using high-quality still SBCE images. Chung *et al*[16] reported an AI model for the automatic classification of gastrointestinal organs, including the esophagus, stomach, SB, and colon, that achieved an average accuracy of 0.87 for SB detection. Similarly, organ localization using our model yielded a > 97% accuracy rate on an internal dataset. For external validation, AI-assisted reading identified the first SB and colon images with performance comparable to that of conventional reading; however, for first colon-image identification, the AI-assisted reading performance was numerically lower. Challenges, such as debris, fecal material, or bloody material in the distal SB or colon may have been responsible. Therefore, the model should be trained under more varied conditions, including poor bowel preparation, to improve the colon identification performance.

Our model performed well for SB-abnormality detection on an internal dataset. Although the positive predictive values for angiodysplasia and bleeding were relatively low, this trade-off minimized the missed significant lesions because the

Table 2 Comparison of small-bowel capsule endoscopy reading performance of conventional reading and artificial intelligence-assisted reading, *n* (%)/mean ± SD

	Conventional reading (*n* = 32)	AI-assisted reading (*n* = 32)	*P* value
Reading time	57.19 ± 20.30	9.03 ± 4.25	< 0.001
Small-bowel localization			
First small-bowel localization	31 (96.9)	28 (87.5)	0.18
First colon localization	26 (81.2)	20 (62.5)	0.08
Detection of small-bowel abnormalities			
Detection score	0.73 ± 0.40	0.75 ± 0.42	0.88

AI: Artificial intelligence.

Table 3 Comparison of diagnostic performance for small-bowel abnormality detection between conventional and artificial intelligence-assisted reading

	Diagnostic performance (%) (95%CI)					*P* value
	Accuracy	Sensitivity	Specificity	PPV	NPV	
Erosions/ulcers						
Conventional reading	81 (80.5-81.3)	64 (63.4-64.8)	100 (100-100)	100 (100-100)	71 (70.4-71.6)	0.746
AI-assisted reading	84 (84.1-84.9)	71 (70.0-71.4)	100 (100-100)	100 (100-100)	75 (74.7-75.9)	
Angiodysplasia						
Conventional reading	87 (87.0-87.7)	56 (54.5-56.6)	100 (100-100)	100 (99.1-99.9)	85 (84.5-85.4)	0.719
AI-assisted reading	91 (90.3-90.9)	67 (65.1-67.0)	100 (100-100)	100 (99.5-100)	88 (88.0-88.8)	
Bleeding						
Conventional reading	94 (93.6-94.2)	67 (65.8-68.4)	100 (100-100)	98 (97.6-99.2)	93 (92.7-93.3)	0.071
AI-assisted reading	84 (83.8-84.6)	67 (65.6-68.2)	88 (87.9-88.6)	57 (55.8-58.2)	92 (91.7-92.4)	

AI: Artificial intelligence; PPV: Positive predictive value; NPV: Negative predictive value.

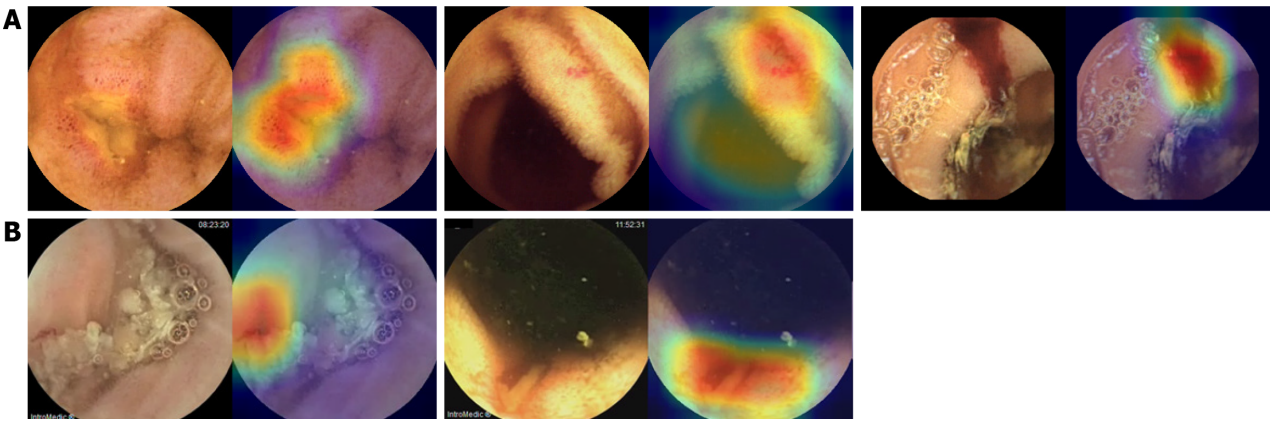


Figure 3 Representative images of small bowel abnormalities correctly detected using artificial intelligence-assisted reading in the validation data set. A: Correct detection of erosions/ulcers (left), angiodysplasia (middle), and bleeding (right) using both artificial intelligence (AI)-assisted reading and conventional readings; B: Correct detection of angiodysplasia (left) and erosions/ulcers (right) using AI-assisted reading; however, conventional reading missed these abnormalities.

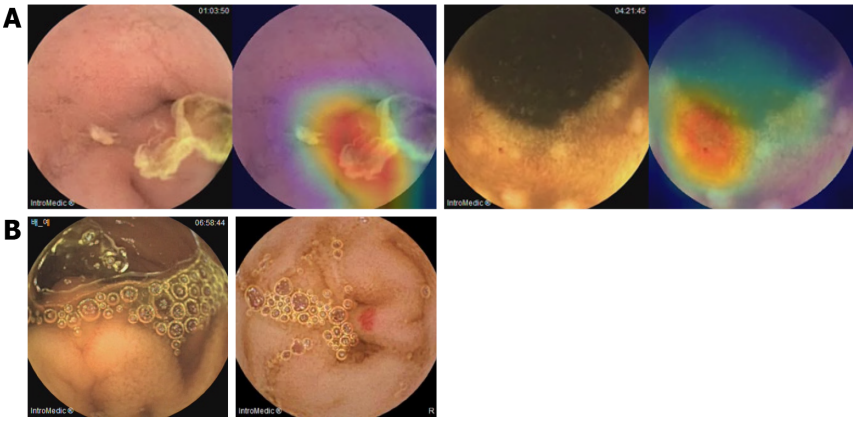


Figure 4 Representative images of incorrect recognition using artificial intelligence assisted reading on validation data set. A: False-positive cases of erosions/ulcers (left) and angiodysplasia (right); B: False-negative cases of erosions/ulcers (left) and angiodysplasia (right).

final diagnosis was established by a physician who reviewed the CNN-model results. For external validation, our CNN model seemed to exhibit lower sensitivity than previous models[12,20]. However, this was because we used real-world clinical data of patients with OGIB for external validation; therefore, several factors, such as the quality of bowel preparation and the size or number of abnormalities, could have affected the SB-abnormality detection. Previous studies have shown a wide range of abnormal lesion detection rates (38%–83%) for conventional reading of images of patients with OGIB[3,4]. This wide range encompasses the sensitivity of the conventional and AI-assisted readings observed in our study.

AI-assisted reading can reduce time and maintain a similar or superior diagnostic performance to conventional reading [18,20,24,25]. In our study, AI-assisted reading reduced interpretation time without compromising SBCE reading performance, aligning with previous studies. Although both the conventional and AI-assisted reading times in our study appeared longer than those of previous studies, this was likely attributable to the double-camera SBCE, which provided more images for review[18,20]. As experts suggest that 40-50 minutes minimum are required for one accurate reading, even for an experienced reader, the reading time in our study is acceptable[8].

The high performance observed in our study may raise concerns about potential overfitting or insufficient challenge in the test cases. To mitigate this, we evaluated model performance using a comprehensive set of metrics—including sensitivity, specificity, and precision—rather than relying solely on accuracy. The model was designed as a two-step architecture, in which abnormality detection and lesion classification were performed sequentially, and a post-processing method utilizing temporal context was applied to improve prediction consistency and compensate for the limitations of single-frame interpretation. With regard to data composition, we did not selectively construct or enrich the dataset with particularly difficult cases (*i.e.*, hard negative or ambiguous cases), but instead included all available SBCE data collected during the study period to reflect the actual clinical distribution. As a result, a range of lesion difficulty levels may have been included naturally; however, such cases could still be underrepresented. Further efforts to diversify training data and incorporate more challenging cases will be needed to improve the robustness and generalizability of the model.

Our study has several limitations. First, the training datasets were retrospectively collected from a limited number of patients, potentially causing class imbalance. To address this issue, we applied various augmentation techniques to the minority classes and adjusted the class distribution to approximate a 1:1 balance with the majority class. All augmented samples were reviewed by clinical experts to ensure clinical plausibility. In addition, class-aware loss functions such as weighted cross-entropy were employed to enhance the learning signal for underrepresented lesions. Furthermore, we adopted a two-step detection architecture consisting of coarse-level filtering and fine-grained lesion classification. The first step utilized a wide receptive field to detect abnormality candidates, and the second step refined the classification of potential lesions. This structure contributed to improved sensitivity for abnormal findings, reduced false positives, and mitigated challenges such as representation collapse and gradient conflict, which are common in multi-class classification with severe imbalance. Finally, post-processing based on Youden's Index was applied to determine the optimal threshold for abnormality classification, helping to balance sensitivity and specificity—particularly for minority classes. Second, our CNN model was developed using data from Asian patients, a specific SBCE platform (MiroCam; Intromedic Co., Ltd.), and cases limited to suspected OGIB, which may not fully represent the broader range of indications and pathologies encountered in routine SBCE practice; these aspects potentially limit the generalizability of our findings. However, our AI model was specifically developed for application to patients with suspected OGIB, which is one of the most common and clinically significant indications for SBCE in real-world practice. The external validation cases were collected from four independent institutions and accurately reflected the intended target population (patients with suspected SB bleeding for whom conventional endoscopies failed to identify a bleeding source). Nonetheless, we acknowledge that further validation in larger and more diverse patient cohorts is warranted to enhance the generalizability and extend the applicability of our AI model. Third, the external validation cohort included only 32 SBCE cases, which may have limited the statistical power to detect significant differences in outcomes. However, the primary outcome of reading time reduction was robustly demonstrated, with an extremely large effect size (Cohen's d = 3.28) and a post hoc statistical power of 1.0, indicating that the sample size was sufficient to support the primary findings. Further studies with larger and more

diverse cohorts are warranted to confirm the model's diagnostic performance, particularly for lesion detection and localization. Fourth, in our study, we focused on detecting three specific abnormalities in SBCE images: Erosions/ulcers, angiodysplasia, and bleeding. These abnormalities were selected based on their high prevalence as common causes of OGIB. Studies have shown that small-bowel lesions contribute to approximately 75% of OGIB cases, with vascular abnormalities, including angioectasia, constituting approximately 70% of these cases in Western populations[26]. Furthermore, ulcerations are a significant cause of OGIB, especially in Asian populations[27]. Although our model currently focuses on these three abnormalities, we recognize that other conditions, such as tumors, are also important causes of OGIB. Indeed, we had intended to include tumors in the present study, but the available data were too limited. Incorporating the detection of these additional abnormalities into future iterations of our AI model would enhance its clinical applicability, thereby providing a more comprehensive tool for SBCE interpretation. This expansion could facilitate a more reliable and efficient diagnostic approach, particularly in the context of patients with suspected OGIB. Finally, a core limitation of SBCE, including for the present AI model, is the lack of advanced image-enhancement technologies, such as narrow-band imaging (NBI). These technologies are available for conventional endoscopy and significantly aid the detection and characterization of subtle lesions, such as angiodysplasia or early neoplasia. However, SBCE is primarily employed in the evaluation of patients with suspected OGIB to detect bleeding lesions. Indeed, the non-invasive nature of SBCE renders it particularly suitable for initial diagnostic assessment of OGIB patients. When further intervention, such as endoscopic bleeding control or additional lesion differentiation is required, SB enteroscopy becomes necessary, with NBI and similar techniques being beneficial for detailed lesion characterization.

CONCLUSION

Our comprehensive CNN model significantly reduced SBCE reading time and maintained performance comparable to conventional reading by experienced endoscopists when used for SB localization and multiple SB-abnormality detection in patients with OGIB. Therefore, AI-assisted SBCE reading can be implemented in clinical practice for improved efficiency and consistency.

FOOTNOTES

Author contributions: Yang YJ, Cho BJ contributed to the conceptualization and supervision of the manuscript; Yang YJ, Park TY was involved in the validation of this study; Kwon YS, Park Y, Lee JG, Lee SP, Kim KO, Jang HJ participated in the data curation; Kwon YS, Park TY, Kim SE contributed to the investigation; Kwon YS, Park TY took part in the visualization and contributed to the writing-original draft preparation; Yang YJ, Cho BJ, Park TY contributed to the methodology of the manuscript; Park TY, Kim SE was responsible to the software; Yang YJ, Cho BJ contributed to the writing-reviewing and editing.

Supported by The Bio and Medical Technology Development Program of the National Research Foundation, No. NRF-2022R1C1C1010643.

Institutional review board statement: The study was approved by the Institutional Review Board of Chuncheon Sacred Heart Hospital (IRB No. 2018-05).

Informed consent statement: All of the patients' personal identifiable information was removed before use, and individual consent for this retrospective analysis was waived.

Conflict-of-interest statement: The authors declare that they have no conflict of interest.

Data sharing statement: The data sets generated and/or analyzed during this study are available from the corresponding author on reasonable request.

Open Access: This article is an open-access article that was selected by an in-house editor and fully peer-reviewed by external reviewers. It is distributed in accordance with the Creative Commons Attribution NonCommercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: https://creativecommons.org/Licenses/by-nc/4.0/

Country of origin: South Korea

ORCID number: Jae Gon Lee 0000-0003-2930-8159; Kyeong Ok Kim 0000-0001-5799-7436; Young Joo Yang 0000-0001-6325-1104.

S-Editor: Liu H
L-Editor: A
P-Editor: Zhang L

REFERENCES

1 **Iddan G**, Meron G, Glukhovsky A, Swain P. Wireless capsule endoscopy. *Nature* 2000; **405**: 417 [PMID: 10839527 DOI: 10.1038/35013140] [FullText]

2 **Kopylov U**, Seidman EG. Diagnostic modalities for the evaluation of small bowel disorders. *Curr Opin Gastroenterol* 2015; **31**: 111-117 [PMID: 25635667 DOI: 10.1097/MOG.0000000000000159] [FullText]

3 **Pennazio M**, Rondonotti E, Despott EJ, Dray X, Keuchel M, Moreels T, Sanders DS, Spada C, Carretero C, Cortegoso Valdivia P, Elli L, Fuccio L, Gonzalez Suarez B, Koulaouzidis A, Kunovsky L, McNamara D, Neumann H, Perez-Cuadrado-Martinez E, Perez-Cuadrado-Robles E, Piccirelli S, Rosa B, Saurin JC, Sidhu R, Tacheci I, Vlachou E, Triantafyllou K. Small-bowel capsule endoscopy and device-assisted enteroscopy for diagnosis and treatment of small-bowel disorders: European Society of Gastrointestinal Endoscopy (ESGE) Guideline - Update 2022. *Endoscopy* 2023; **55**: 58-95 [PMID: 36423618 DOI: 10.1055/a-1973-3796] [FullText]

4 **Rondonotti E**, Villa F, Mulder CJ, Jacobs MA, de Franchis R. Small bowel capsule endoscopy in 2007: indications, risks and limitations. *World J Gastroenterol* 2007; **13**: 6140-6149 [PMID: 18069752 DOI: 10.3748/wjg.v13.i46.6140] [FullText]

5 **ASGE Technology Committee**, Wang A, Banerjee S, Barth BA, Bhat YM, Chauhan S, Gottlieb KT, Konda V, Maple JT, Murad F, Pfau PR, Pleskow DK, Siddiqui UD, Tokar JL, Rodriguez SA. Wireless capsule endoscopy. *Gastrointest Endosc* 2013; **78**: 805-815 [PMID: 24119509 DOI: 10.1016/j.gie.2013.06.026] [FullText]

6 **McAlindon ME**, Ching HL, Yung D, Sidhu R, Koulaouzidis A. Capsule endoscopy of the small bowel. *Ann Transl Med* 2016; **4**: 369 [PMID: 27826572 DOI: 10.21037/atm.2016.09.18] [FullText]

7 **Jang BI**, Lee SH, Moon JS, Cheung DY, Lee IS, Kim JO, Cheon JH, Park CH, Byeon JS, Park YS, Shim KN, Kim YS, Kim KJ, Lee KJ, Ryu JK, Chang DK, Chun HJ, Choi MG. Inter-observer agreement on the interpretation of capsule endoscopy findings based on capsule endoscopy structured terminology: a multicenter study by the Korean Gut Image Study Group. *Scand J Gastroenterol* 2010; **45**: 370-374 [PMID: 20148733 DOI: 10.3109/00365520903521574] [FullText]

8 **Rondonotti E**, Pennazio M, Toth E, Koulaouzidis A. How to read small bowel capsule endoscopy: a practical guide for everyday use. *Endosc Int Open* 2020; **8**: E1220-E1224 [PMID: 33015322 DOI: 10.1055/a-1210-4830] [FullText]

9 **Cho BJ**, Bang CS, Park SW, Yang YJ, Seo SI, Lim H, Shin WG, Hong JT, Yoo YT, Hong SH, Choi JH, Lee JJ, Baik GH. Automated classification of gastric neoplasms in endoscopic images using a convolutional neural network. *Endoscopy* 2019; **51**: 1121-1129 [PMID: 31443108 DOI: 10.1055/a-0981-6133] [FullText]

10 **Yang YJ**, Bang CS. Application of artificial intelligence in gastroenterology. *World J Gastroenterol* 2019; **25**: 1666-1683 [PMID: 31011253 DOI: 10.3748/wjg.v25.i14.1666] [FullText]

11 **Tziortziotis I**, Laskaratos FM, Coda S. Role of Artificial Intelligence in Video Capsule Endoscopy. *Diagnostics (Basel)* 2021; **11**: 1192 [PMID: 34209029 DOI: 10.3390/diagnostics11071192] [FullText]

12 **Ding Z**, Shi H, Zhang H, Meng L, Fan M, Han C, Zhang K, Ming F, Xie X, Liu H, Liu J, Lin R, Hou X. Gastroenterologist-Level Identification of Small-Bowel Diseases and Normal Variants by Capsule Endoscopy Using a Deep-Learning Model. *Gastroenterology* 2019; **157**: 1044-1054.e5 [PMID: 31251929 DOI: 10.1053/j.gastro.2019.06.025] [FullText]

13 **Aoki T**, Yamada A, Kato Y, Saito H, Tsuboi A, Nakada A, Niikura R, Fujishiro M, Oka S, Ishihara S, Matsuda T, Nakahori M, Tanaka S, Koike K, Tada T. Automatic detection of various abnormalities in capsule endoscopy videos by a deep learning-based system: a multicenter study. *Gastrointest Endosc* 2021; **93**: 165-173.e1 [PMID: 32417297 DOI: 10.1016/j.gie.2020.04.080] [FullText]

14 **Otani K**, Nakada A, Kurose Y, Niikura R, Yamada A, Aoki T, Nakanishi H, Doyama H, Hasatani K, Sumiyoshi T, Kitsuregawa M, Harada T, Koike K. Automatic detection of different types of small-bowel lesions on capsule endoscopy images using a newly developed deep convolutional neural network. *Endoscopy* 2020; **52**: 786-791 [PMID: 32557474 DOI: 10.1055/a-1167-8157] [FullText]

15 **Son G**, Eo T, An J, Oh DJ, Shin Y, Rha H, Kim YJ, Lim YJ, Hwang D. Small Bowel Detection for Wireless Capsule Endoscopy Using Convolutional Neural Networks with Temporal Filtering. *Diagnostics (Basel)* 2022; **12**: 1858 [PMID: 36010210 DOI: 10.3390/diagnostics12081858] [FullText]

16 **Chung J**, Oh DJ, Park J, Kim SH, Lim YJ. Automatic Classification of GI Organs in Wireless Capsule Endoscopy Using a No-Code Platform-Based Deep Learning Model. *Diagnostics (Basel)* 2023; **13**: 1389 [PMID: 37189489 DOI: 10.3390/diagnostics13081389] [FullText]

17 **Nam JH**, Hwang Y, Oh DJ, Park J, Kim KB, Jung MK, Lim YJ. Development of a deep learning-based software for calculating cleansing score in small bowel capsule endoscopy. *Sci Rep* 2021; **11**: 4417 [PMID: 33627678 DOI: 10.1038/s41598-021-81686-7] [FullText]

18 **Aoki T**, Yamada A, Oka S, Tsuboi M, Kurokawa K, Togo D, Tanino F, Teshima H, Saito H, Suzuki R, Arai J, Abe S, Kondo R, Yamashita A, Tsuboi A, Nakada A, Niikura R, Tsuji Y, Hayakawa Y, Matsuda T, Nakahori M, Tanaka S, Kato Y, Tada T, Fujishiro M. Comparison of clinical utility of deep learning-based systems for small-bowel capsule endoscopy reading. *J Gastroenterol Hepatol* 2024; **39**: 157-164 [PMID: 37830487 DOI: 10.1111/jgh.16369] [FullText]

19 **Choi KS**, Park D, Kim JS, Cheung DY, Lee BI, Cho YS, Kim JI, Lee S, Lee HH. Deep learning in negative small-bowel capsule endoscopy improves small-bowel lesion detection and diagnostic yield. *Dig Endosc* 2024; **36**: 437-445 [PMID: 37612137 DOI: 10.1111/den.14670] [FullText]

20 **O'Hara FJ**, Mc Namara D. Capsule endoscopy with artificial intelligence-assisted technology: Real-world usage of a validated AI model for capsule image review. *Endosc Int Open* 2023; **11**: E970-E975 [PMID: 37828977 DOI: 10.1055/a-2161-1816] [FullText]

21 **Huang G**, Liu Z, Van Der Maaten L, Weinberger KQ. Densely Connected Convolutional Networks. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 21-26 July 2017; United States, Honolulu, 2017: 2261-2269 [DOI: 10.1109/cvpr.2017.243] [FullText]

22 **Schafer R**. What Is a Savitzky-Golay Filter? [Lecture Notes]. *IEEE Signal Process Mag* 2011; **28**: 111-117 [DOI: 10.1109/msp.2011.941097] [FullText]

23 **Hwang Y**, Lee HH, Park C, Tama BA, Kim JS, Cheung DY, Chung WC, Cho YS, Lee KM, Choi MG, Lee S, Lee BI. Improved classification and localization approach to small bowel capsule endoscopy using convolutional neural network. *Dig Endosc* 2021; **33**: 598-607 [PMID: 32640059 DOI: 10.1111/den.13787] [FullText]

24 **Park J**, Hwang Y, Nam JH, Oh DJ, Kim KB, Song HJ, Kim SH, Kang SH, Jung MK, Jeong Lim Y. Artificial intelligence that determines the clinical significance of capsule endoscopy images can increase the efficiency of reading. *PLoS One* 2020; **15**: e0241474 [PMID: 33119718 DOI: 10.1371/journal.pone.0241474] [FullText]

25 **Li L**, Yang L, Zhang B, Yan G, Bao Y, Zhu R, Li S, Wang H, Chen M, Jin C, Chen Y, Yu C. Automated detection of small bowel lesions based on capsule endoscopy using deep learning algorithm. *Clin Res Hepatol Gastroenterol* 2024; **48**: 102334 [PMID: 38582328 DOI:

10.1016/j.clinre.2024.102334] [FullText]

26 **Liao Z**, Gao R, Xu C, Li ZS. Indications and detection, completion, and retention rates of small-bowel capsule endoscopy: a systematic review. *Gastrointest Endosc* 2010; **71**: 280-286 [PMID: 20152309 DOI: 10.1016/j.gie.2009.09.031] [FullText]

27 **Pasha SF**, Hara AK, Leighton JA. Diagnostic evaluation and management of obscure gastrointestinal bleeding: a changing paradigm. *Gastroenterol Hepatol (N Y)* 2009; **5**: 839-850 [PMID: 20567529] [FullText]

Published by **Baishideng Publishing Group Inc**
7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA
Telephone: +1-925-3991568
E-mail: office@baishideng.com
Help Desk: https://www.f6publishing.com/helpdesk
https://www.wjgnet.com

© 2025 Baishideng Publishing Group Inc. All rights reserved.