J Gastric Cancer. 2024 Jul;24(3):327-340
https://doi.org/10.5230/jgc.2024.24.e28
pISSN 2093-582X·eISSN 2093-5641

jgc Journal of Gastric Cancer

Original Article

Real-World Application of Artificial Intelligence for Detecting Pathologic Gastric Atypia and Neoplastic Lesions

Young Hoon Chang [1], **Cheol Min Shin** [1], **Hae Dong Lee** [1], **Jinbae Park** [2], **Jiwoon Jeon** [2], **Soo-Jeong Cho** [3], **Seung Joo Kang** [4], **Jae-Yong Chung** [5], **Yu Kyung Jun** [1], **Yonghoon Choi** [1], **Hyuk Yoon** [1], **Young Soo Park** [1], **Nayoung Kim** [1], **Dong Ho Lee** [1]

[1]Department of Internal Medicine, Seoul National University Bundang Hospital, Seongnam, Korea
[2]Ainex Co., LTD., Seoul, Korea
[3]Department of Internal Medicine and Liver Research Institute, Seoul National University College of Medicine, Seoul, Korea
[4]Department of Internal Medicine and Healthcare Research Institute, Healthcare System Gangnam Center, Seoul National University Hospital, Seoul, Korea
[5]Department of Clinical Pharmacology and Therapeutics, Seoul National University Bundang Hospital, Seongnam, Korea

Received: Apr 23, 2024
Revised: Jun 11, 2024
Accepted: Jun 17, 2024
Published online: Jun 28, 2024

Correspondence to
Cheol Min Shin
Department of Internal Medicine, Seoul National University Bundang Hospital, 82 Gumi-ro 173beon-gil, Bundang-gu, Seongnam 13620, Korea.
Email: scm6md@gmail.com

ORCID iDs
Young Hoon Chang
https://orcid.org/0000-0001-9221-0993
Cheol Min Shin
https://orcid.org/0000-0003-2265-9845
Hae Dong Lee
https://orcid.org/0000-0002-4395-6577
Jinbae Park
https://orcid.org/0009-0003-6635-0345
Jiwoon Jeon
https://orcid.org/0000-0001-7424-3740
Soo-Jeong Cho
https://orcid.org/0000-0001-7144-0589
Seung Joo Kang
https://orcid.org/0000-0002-7401-8356

ABSTRACT

Purpose: Results of initial endoscopic biopsy of gastric lesions often differ from those of the final pathological diagnosis. We evaluated whether an artificial intelligence-based gastric lesion detection and diagnostic system, ENdoscopy as AI-powered Device Computer Aided Diagnosis for Gastroscopy (ENAD CAD-G), could reduce this discrepancy.
Materials and Methods: We retrospectively collected 24,948 endoscopic images of early gastric cancers (EGCs), dysplasia, and benign lesions from 9,892 patients who underwent esophagogastroduodenoscopy between 2011 and 2021. The diagnostic performance of ENAD CAD-G was evaluated using the following real-world datasets: patients referred from community clinics with initial biopsy results of atypia (n=154), participants who underwent endoscopic resection for neoplasms (Internal video set, n=140), and participants who underwent endoscopy for screening or suspicion of gastric neoplasm referred from community clinics (External video set, n=296).
Results: ENAD CAD-G classified the referred gastric lesions of atypia into EGC (accuracy, 82.47%; 95% confidence interval [CI], 76.46%–88.47%), dysplasia (88.31%; 83.24%–93.39%), and benign lesions (83.12%; 77.20%–89.03%). In the Internal video set, ENAD CAD-G identified dysplasia and EGC with diagnostic accuracies of 88.57% (95% CI, 83.30%–93.84%) and 91.43% (86.79%–96.07%), respectively, compared with an accuracy of 60.71% (52.62%–68.80%) for the initial biopsy results (P<0.001). In the External video set, ENAD CAD-G classified EGC, dysplasia, and benign lesions with diagnostic accuracies of 87.50% (83.73%–91.27%), 90.54% (87.21%–93.87%), and 88.85% (85.27%–92.44%), respectively.
Conclusions: ENAD CAD-G is superior to initial biopsy for the detection and diagnosis of gastric lesions that require endoscopic resection. ENAD CAD-G can assist community endoscopists in identifying gastric lesions that require endoscopic resection.

Keywords: Gastric cancer; Endoscopic mucosal resection; Endoscopy; Artificial intelligence

jgc Journal of Gastric Cancer

Jae-Yong Chung
https://orcid.org/0000-0003-4188-2786
Yu Kyung Jun
https://orcid.org/0000-0002-4046-8578
Yonghoon Choi
https://orcid.org/0000-0002-1331-969X
Hyuk Yoon
https://orcid.org/0000-0002-2657-0349
Young Soo Park
https://orcid.org/0000-0003-1893-7726
Nayoung Kim
https://orcid.org/0000-0002-9397-0406
Dong Ho Lee
https://orcid.org/0000-0002-6376-410X

Funding
This research was supported by Ainex Co. Ltd.

Conflict of Interest
Jinbae Park and Jiwoon Jeon are employees of Ainex Co., LTD. The other authors have no conflict of interest or financial arrangement that could potentially influence the presented research.

Author Contributions
Conceptualization: S.C.M., C.S.J., K.S.J.; Data curation: S.C.M., C.S.J., K.S.J., J.Y.K., C.Y., Y.H., P.Y.S., K.N., L.D.H.; Formal analysis: C.Y.H., L.H.D., P.J., J.J.; Investigation: C.Y.H., L.H.D., P.J., J.J.; Supervision: C.Y.H., S.C.M., L.H.D., P.J., J.J., C.S.J., K.S.J., C.J.Y., J.Y.K., C.Y., Y.H., P.Y.S., K.N., L.D.H.; Writing - original draft: C.Y.H.; Writing - review & editing: C.Y.H., S.C.M., L.H.D., P.J., J.J., C.S.J., K.S.J., C.J.Y., J.Y.K., C.Y., Y.H., P.Y.S., K.N., L.D.H.

INTRODUCTION

Gastric cancer is a leading cause of death worldwide [1,2]. Nevertheless, with the development of endoscopic treatments, the overall survival rate of early gastric cancer (EGC) has reached over 90% [3]. For this, early diagnosis is essential [4,5]; however, the detection of subtle mucosal changes of malignancy in the stomach depends largely on the skills of experienced endoscopists [6]. Although several enhanced endoscopic techniques, such as magnifying endoscopy with narrow-band imaging, have been developed over the last few decades to aid EGC diagnosis, not every endoscopy laboratory has these facilities. Furthermore, less experienced endoscopists encounter difficulties in distinguishing malignant gastric lesions from other gastric abnormalities [7-10].

Over the last decade, a vast array of computer-aided detection (CADe) and diagnosis (CADx) systems for gastric lesions have been developed owing to the surge in artificial intelligence (AI) [11-13]. Most studies have been conducted in Asia because of the high prevalence of gastric cancer and the numerous upper gastrointestinal endoscopies performed in the East. Wu et al. [14], Nam et al. [15], and Gong et al. [16] demonstrated that an AI-based convolutional neural network (CNN) can not only reduce blind spots but also detect and diagnose gastric abnormalities with a sensitivity and specificity of >90%. Furthermore, Wu et al. [17] mentioned that their AI system surpassed endoscopists in predicting the EGC differentiation status and invasion depth.

However, most researchers have evaluated the performance of reported AI systems using selected images that depict a clear diagnosis [11,14-16]. Therefore, the diagnostic difficulties and clinical importance of the test lesions have not been comprehensively considered. Recently, we developed an AI software termed ENdoscopy as AI-powered Device Computer Aided Diagnosis for Gastroscopy (ENAD CAD-G) to locate, detect, and diagnose gastric lesions. Considering the above, we aimed to evaluate the performance of our AI software, especially in difficult clinical situations, such as patients with pathologic atypia referred from community clinics and patients referred for endoscopic resection because of gastric neoplasms.

MATERIALS AND METHODS

Study design and participants

We conducted a multicenter diagnostic study across 2 tertiary referral university hospitals and one health promotion center in South Korea, viz., Seoul National University Hospital (SNUH), Seoul National University Bundang Hospital (SNUBH), and Healthcare System Gangnam Center (HSGC). **Fig. 1** and **Supplementary Fig. 1** illustrate the workflow of the study. We retrospectively collected upper gastrointestinal endoscopic images from 3 hospitals to develop, validate, and assess the system. The images were manually annotated by endoscopists from each participating hospital. After randomly assigning the labeled images to 3 different datasets, we developed a compilation model called ENAD CAD-G, which consists of 3 CNN-based models, namely CNN1, CNN2, and CNN3, for classifying gastric lesions (**Fig. 1**). Additionally, we prospectively collected endoscopic videos from SNUBH, SNUH, and HSGC. Prospectively collected endoscopic videos from patients undergoing esophagogastroduodenoscopy (EGD) during health checkups at 3 hospitals were used to incorporate normal gastric images into our AI system. This study was approved by the ethics committee of each participating hospital (IRB No: B-2201-735-405 at SNUBH and H-2109-048-

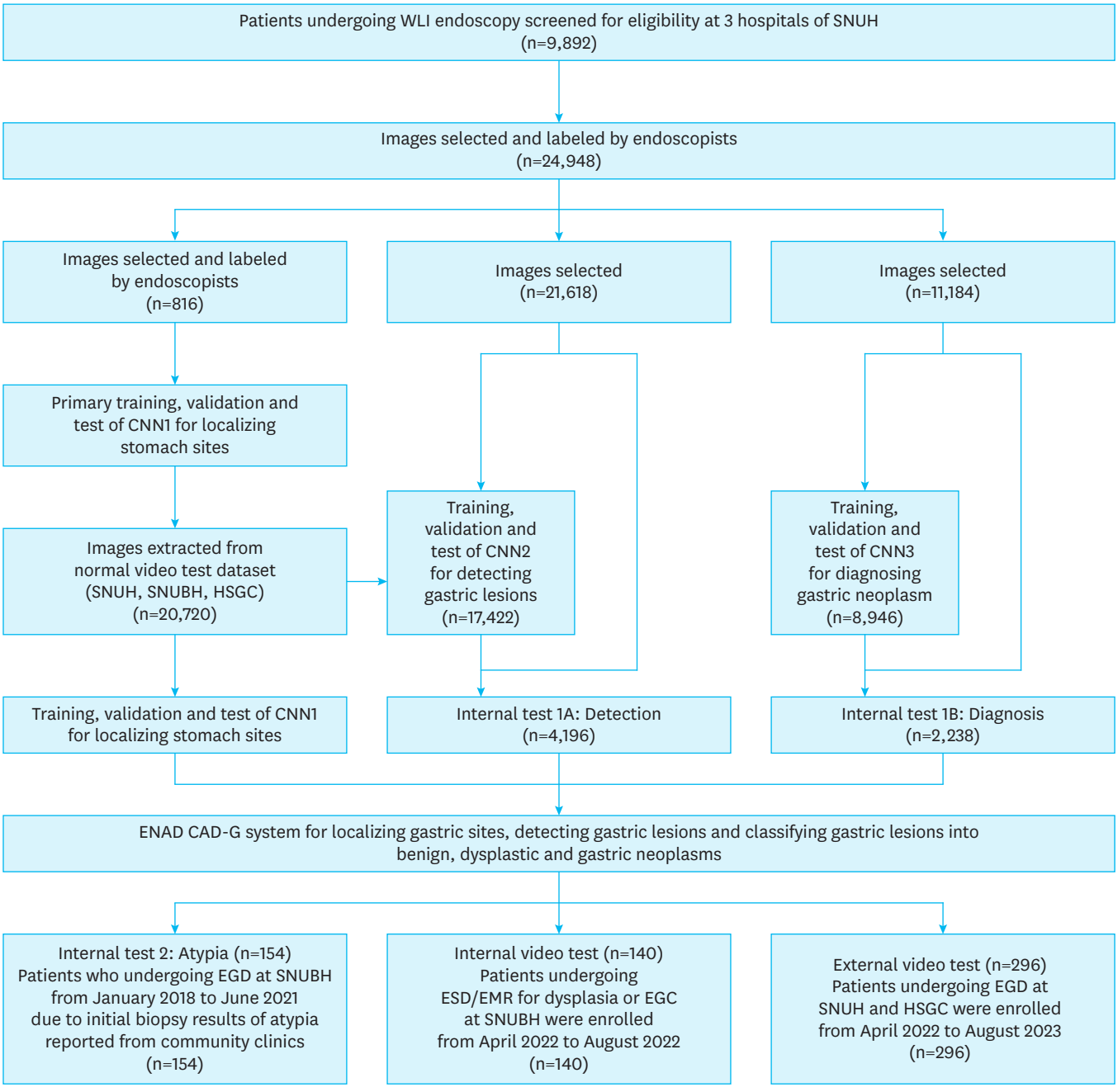

Fig. 1. Workflow chart for the development and evaluation of ENAD CAD-G.
ENAD CAD-G = ENdoscopy as AI-powered Device Computer-Aided Diagnosis for Gastroscopy; WLI = white-light imaging; SNUH = Seoul National University Hospital; CNN = convolutional neural network; EGD = esophagogastroduodenoscopy; SNUBH = Seoul National University Bundang Hospital; ESD = endoscopic submucosal dissection; EMR = endoscopic mucosal resection; EGC = early gastric cancer; HSGC = Healthcare System Gangnam Center.

1253 at SNUH/HSGC). The requirement for informed consent was waived for all participants included in the retrospective analyses. Written informed consent was obtained from the prospective patients at each hospital before video collection.

Data preparation for AI model development

Data acquisition and preprocessing

We collected endoscopic images from 9,892 patients who underwent upper gastrointestinal

https://doi.org/10.5230/jgc.2024.24.e28

jgc Journal of Gastric Cancer

endoscopy at SNUH, SNUBH, and HSGC between January 2011 and June 2021. Standard video endoscopes (GIF-Q260, GIF-H260, or GIF-H290; Olympus Medical Systems, Tokyo, Japan) were used for all procedures. For the 3-class diagnosis, we included patients with dysplasia or EGC who underwent endoscopic resection or curative surgery. Our definition of dysplasia included low- and high-grade gastric adenomas. Furthermore, we included patients diagnosed with benign lesions such as erosions, ulcers, and polyps. Normal gastric images were used to calculate the specificity of the proposed detection model. Seventeen endoscopists (5 from SNUBH, 6 from SNUH, and 6 from HSGC) removed images of low-quality, non-white-light images, and those without definite pathology results (e.g., atypia, suspicion of malignancy). Subsequently, our endoscopists manually labeled all lesions with rectangular boxes and independently labeled each lesion as a neoplasm or non-neoplasm according to both pathologic results and endoscopic features. The 10 experts enrolled in the dataset preprocessing had at least 5 years of experience and had performed more than 15,000 examinations. In total, 24,948 annotated images were included in this study. Of these, 20, 720, 21, 618, and 11, 184 images were assigned for training and validating CNN1, CNN2, and CNN3, respectively.

Dataset for gastric site localization (CNN1)

To train and test the CNN1 model for localizing gastric sites, 816 gastric images were selected, and their gastric sites were labeled by endoscopists. After primary training, validation, and assessment of CNN1 for localizing stomach sites, an additional 20,720 images were extracted from normal videos obtained at HSGC, SNUH, and SNUBH to finalize CNN1 for localizing stomach sites (**Fig. 1**, **Supplementary Table 1**).

Dataset for gastric lesion detection (CNN2)

To train and test the CNN2 model for detecting gastric lesions, we divided the assigned 21,618 images into 17,422 and 4,196 images (**Supplementary Table 2**). We used 17,422 images for training and validation. In addition to the manually selected images, we generated images of EGCs, dysplasia, erosions, and ulcers using generative adversarial networks (GAN). The remaining 4,196 images were used for Internal Test 1A (**Supplementary Fig. 2**).

Dataset for gastric lesion classification (CNN3)

We divided the assigned 11,184 images into 8,946 and 2,238 images, respectively, to train and test the CNN3 model for diagnosing gastric lesions. We used 8,946 images for training and validation. The remaining 2,238 images were used for Internal Test 1B (**Supplementary Table 2**).

Development of AI model: Architecture of ENAD CAD-G and GAN incorporation

We preprocessed the labeled images using the contour detection function provided by OpenCV, an image processing library, before applying the annotated images to CNN training. We used the contour detection function in OpenCV to crop out necessary portions of the images and conceal personal information and other sensitive content. The annotated images contained noise, such as forceps or excessive bleeding. Therefore, data filtering was required before training to reduce the number of images in the estimated distribution. We excluded the following types of images: 1) those with a red channel value above a certain threshold, 2) objects other than the gastric mucosa in the image, and 3) those with severe light exposure.

After image pre-processing, 3 CNN models were trained: location classification (CNN1), gastric lesion detection (CNN2), and gastric lesion diagnosis (CNN3). In addition, we incorporated GAN images into CNN2. YOLOv5 was used for CNN2, EfficientNetB0 [18] was used for CNN1 and CNN3, and stylegan2 [19] was used for the GAN.

https://doi.org/10.5230/jgc.2024.24.e28

jgc Journal of Gastric Cancer

To train CNN1, EfficientNetB0 was implemented to classify the gastric location into the following 10 sites: gastroesophageal junction, antrum, duodenal bulb, duodenal second portion, angle retrograde view, mid-to-high-body retrograde view, fundus retrograde view, lower-body antegrade view, mid-to-high-body antegrade view, and esophagus.

To construct CNN2, we modified YOLOv5 to seamlessly detect gastric lesions in real time. The following data augmentation methods were applied to prevent overfitting and improve the performance: flip up-down, flip left-right, mix-up, copy-paste, and mosaic. There was a performance trade-off when using GAN-generated images; however, we used a model with high sensitivity, which was important for our study.

CNN3 classified lesions detected by CNN2 as EGC, dysplasia, or benign lesions; it extracts the features of endoscopic images using pretrained weights based on EfficientNetB0. The computing environment for training and assessing our deep learning models was as follows: Intel(R) Xeon(R) Silver 4210R CPU @ 2.40GHz (Intel Corporation, Santa Clara, CA, USA), NVIDIA RTX A5000 24GB X2 (NVIDIA, Santa Clara, CA, USA), and 128 GB RAM.

Datasets for AI model evaluation

Still image tests: Internal Test sets 1A and 1B

For Internal Test 1A, 4,196 gastric images with lesions from 1,183 patients (1,080 neoplasms and 803 non-neoplastic lesions) were selected from a retrospective collection of 21,618 gastric images assigned for the training, validation, and assessment of CNN2. For Internal Test 1B, 2,238 gastric images with lesions from 1,145 patients (758 neoplasms and 387 non-neoplastic lesions) were selected from a retrospective collection of 11,184 gastric images assigned for training, validation, and assessment of CNN3 (**Table 1**, **Supplementary Fig. 3**).

Internal Test 2: Atypia set

For Internal Test 2, we reviewed 154 patients referred from community clinics for further evaluation because of biopsy results of atypia or atypical glands. After referral, the patients underwent second-look upper endoscopic examinations at SNUBH between January 2018 and June 2021. The referred gastric lesions were re-biopsied. Pathologically confirmed abnormalities comprising 848 clear gastric images of 154 lesions were selected as follows:

Table 1. Baseline characteristics in test datasets

Characteristics	Internal Test 1A (Detection)	Internal Test 1B (Diagnosis)	Internal Test 2: Atypia	Internal video test	External video test
No. of patients	1,833	1,145	154	140	296
No. of images	4,196	2,238	351	N/A	N/A
Age (yr)	63.60±11.01	64.69±10.42	63.27±10.08	68.40±9.16	58.20±10.66
Sex					
Female	715 (39.0)	501 (43.8)	42 (27.3)	35 (25.0)	116 (39.2)
Male	1,118 (61.0)	644 (56.2)	112 (72.7)	105 (75.0)	180 (60.8)
No. of neoplasm	1,080	758	90	139	156
Dysplasia	452 (41.9)	281 (37.1)	35 (38.9)	45 (32.4)	50 (16.9)
Early gastric cancer	628 (58.1)	477 (62.9)	55 (61.1)	94 (67.6)	106 (35.8)
No. of non-neoplastic abnormalities	803	387	64	1	140
Ulcer	192 (24.0)	145 (37.5)	0 (0.0)	0 (0.0)	11 (3.7)
Polyps	104 (13.0)	23 (0.06)	2 (3.0)	0 (0.0)	11 (3.7)
Erosions	505 (63.0)	218 (56.3)	56 (87.5)	1 (100.0)	96 (32.4)
Others*	2 (0.0)	1 (0.0)	6 (9.5)	0 (0.0)	22 (7.4)

Values are mean ± standard deviation or number (%), unless specified otherwise.
N/A = not applicable.
*Others: scars, submucosal lesions, chronic gastritis.

https://doi.org/10.5230/jgc.2024.24.e28

jgc Journal of Gastric Cancer

90 neoplasms and 64 non-neoplastic images (**Supplementary Fig. 3**). ENAD CAD-G was assessed using the collected still images. A receiver operating characteristic (ROC) curve was constructed by plotting the sensitivity against the false-positive rate for each threshold.

Video test: Internal and external video sets

To evaluate the applicability of our AI system in real-time, we assessed prospectively collected videos at SNUBH from April 2022 to August 2022 of patients with gastric dysplasia or EGC undergoing endoscopic resection (ER) who provided informed consent (Internal video set, **Supplementary Fig. 3**). Patients who had previously undergone gastrectomy or esophagectomy and those with contraindications to biopsy (e.g., bleeding tendency and anticoagulant use) were excluded. Informed consent was obtained after providing sufficient explanation before and after the endoscopy. All patients underwent ER for previously diagnosed dysplasia or EGC based on the initial biopsies performed at community clinics or SNUBH. After ER, the resected specimens were sent for final pathological diagnosis. ENAD CAD-G was assessed using the collected videos containing the entire endoscopic procedure. The ROC curve was drawn by plotting the sensitivity against the false-positive rate for each threshold.

To evaluate the performance of ENAD CAD-G in an external setting, we assessed videos prospectively collected at HSGC and SNUH from April 2022 to August 2023 from patients who provided informed consent (External video set, **Supplementary Fig. 3**). All patients who underwent EGD for screening purposes or suspected malignancy were referred from community clinics.

AI methods used for video test

Object tracking is a key deep learning method for analyzing moving objects in real-time. We applied a state-of-the-art multi-object tracking method called observation-centric simple online and real-time tracking (OC-SORT) in a video test [20]. OC-SORT comprises of 2 stages. In our experiment, lesion detection was performed using a fast region-based CNN method. Subsequently, a Kalman filter was used to estimate the position of the lesion in the next frame. This tracking method measures the intersection of union (IOU) distance between actual and predicted bounding boxes. By assigning the predicted and current positions using the Hungarian algorithm, the method recognized the lesion as an identical object when the IOU threshold exceeded 0.3. We set the lesions to be recognized only upon tracking more than 6 frames. OC-SORT reduces the number of false-positive cases by continuously tracking and observing the primary lesion in videos. Finally, soft voting was performed to classify the tracked lesions as EGC, dysplastic, or benign (**Supplementary Fig. 4**).

Statistical analysis

The metrics of accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and area under the curve (AUC) with 95% confidence intervals (CIs) were used to evaluate the performance of ENAD CAD-G and the initial biopsy results. We compared the performances of the initial biopsy results and AI using the Student's t-test. A P-value <0.05 was considered statistically significant. All statistical analyses were performed using Python 3.10.11, with Scikit-learn (Python Software Foundation, Beaverton, OR, USA).

https://doi.org/10.5230/jgc.2024.24.e28

jgc Journal of Gastric Cancer

RESULTS

Baseline characteristics of test datasets

Table 1 shows the baseline characteristics of the patients and lesions in the test datasets. We assigned images of 1,833 and 1,145 patients to evaluate the detection (Internal Test 1A) and diagnosis (Internal Test 1B) models, respectively. The atypia set comprised 154 patients. In the Internal video set, 140 patients with final pathologic diagnoses of 139 neoplastic lesions and 1 non-neoplastic lesion were included based on the eligibility criteria. The External video set comprised 259 patients with 140 non-neoplastic and 156 neoplastic lesions, selected based on the eligibility criteria.

Performance of ENAD CAD-G on the classification of gastric locations, Internal Test 1A and 1B

With CNN1, ENAD CAD-G demonstrated an accuracy of 96.66 (95% CI, 96.21%–97.11%) for classifying the gastric location into 10 parts (**Supplementary Table 1**). To detect gastric lesions, CNN2 achieved a sensitivity, specificity, and PPV of 88.39% (86.92%–89.26%), 86.94% (83.74%–89.17%), and 87.59% (85.14%–89.33%), respectively, in Internal Test 1A at a confidence level of 0.3 (**Supplementary Table 3**). **Supplementary Figs. 5** and **6** depict the learning and ROC curves for CNN2. Upon incorporating 10% additional GAN images to CNN2, it achieved a sensitivity, specificity, and PPV of 88.55% (87.59%–89.44%), 86.73% (83.98%–89.12%), and 87.43% (85.34%–89.36%), respectively, in Internal Test 1A (**Supplementary Table 3**).

For classifying gastric lesions into 3 classes (EGC, dysplasia, and benign) in Internal Test 1B, CNN3 achieved an accuracy, sensitivity, and specificity of 92.81% (91.74%–93.88%), 91.92% (90.79%–93.05%), and 96.21% (95.42%–97.00%), respectively (**Supplementary Table 4**). **Supplementary Fig. 7** depicts AUC values of 0.9831, 0.9803, and 0.9882 for the diagnosis of EGC, dysplasia, and benign lesions, respectively.

Performance of ENAD CAD-G in the atypia set

Table 2 summarizes the performance of ENAD CAD-G in classifying gastric atypia lesions into EGC, dysplasia, and benign lesions. ENAD CAD-G classified the atypia lesions into EGCs with an accuracy, sensitivity, and specificity of 82.47% (76.46%–88.47%), 92.73% (85.86%–99.59%), and 68.92 (58.37%–79.46%), respectively; it classified the atypia lesions into dysplasia with an accuracy, sensitivity, and specificity of 88.31% (83.24%–93.39%), 60.00% (34.77%–76.23%), and 96.64% (93.40%–99.88%), respectively; finally, benign lesions could be identified with an accuracy, sensitivity, and specificity of 83.12% (77.20%–89.03%), 71.88% (60.86%–82.89%), and 91.11% (85.23%–94.66%), respectively. The AUC for diagnosing EGC, dysplasia, and benign lesions in patients with atypia were 0.7898, 0.7657, and 0.7722, respectively (**Fig. 2A**). In the age- and sex-stratified analyses, comparable performances were achieved among females and those aged under 55 years.

Performance of ENAD CAD-G in prospective videos: Internal and external video sets

Table 3 compares the performances of the initial biopsy results and the predictions of ENAD CAD-G for the Internal video set. For diagnosing EGCs, the initial biopsy results attained accuracy, sensitivity, and specificity of 60.71% (52.62%–68.80%), 41.49% (31.53%–51.45%), and 100.00% (100.00%–100.00%), respectively. In contrast, ENAD CAD-G attained an accuracy, sensitivity, and specificity of 88.57% (83.30%–93.84%), 91.49% (85.85%–97.13%),

https://doi.org/10.5230/jgc.2024.24.e28

jgc Journal of Gastric Cancer

Table 2. Diagnostic performance of ENAD CAD-G in Internal Test 2: Atypia

Final Pathology	Index	ENAD CAD-G				
		All atypia patients (n=154)	Male (n=112)	Female (n=42)	Age ≥55 yr (n=122)	Age <55 yr (n=32)
EGC (n=55)	Accuracy (%)	82.47 (76.46–88.47)	83.04 (76.08–89.99)	80.95 (69.08–92.83)	82.79 (76.09–89.49)	81.25 (67.73–94.77)
	Sensitivity (%)	92.73 (85.86–99.59)	91.11 (82.80–99.43)	100.00 (100.00–100.00)	91.11 (82.80–99.43)	100.00 (100.00–100.00)
	Specificity (%)	68.92 (58.37–79.46)	77.61 (67.63–87.59)	75.00 (60.00–90.00)	77.92 (68.66–87.19)	72.73 (54.12–91.34)
	PPV (%)	68.92 (58.37–79.46)	73.21 (64.21–82.21)	55.56 (38.85–72.26)	70.69 (61.81–79.57)	62.50 (43.89–81.11)
	NPV (%)	95.00 (90.22–99.78)	92.86 (86.11–99.90)	100.00 (100.00–100.00)	93.75 (87.82–99.68)	100.00 (100.00–100.00)
	AUC	0.7898 (0.7254–0.8542)	0.7634 (0.6847–0.8421)	0.6634 (0.5205–0.8063)	0.7882 (0.7157–86.07)	0.5822 (0.4113–0.7531)
Dysplasia (n=35)	Accuracy (%)	88.31 (83.24–93.39)	91.96 (86.93–97.00)	78.57 (66.16–90.98)	86.07 (79.92–92.21)	96.88 (90.85–100.00)
	Sensitivity (%)	60.00 (34.77–76.23)	65.22 (45.75–84.68)	50.00 (21.71–78.29)	60.61 (43.93–77.28)	50.00 (0.00–100.00)
	Specificity (%)	96.64 (93.40–99.88)	98.88 (96.69–100.00)	90.00 (79.26–100.00)	95.51 (91.20–99.81)	100.00 (100.00–100.00)
	PPV (%)	84.00 (69.63–98.37)	93.75 (89.08–98.42)	66.67 (50.58–94.98)	83.33 (76.20–90.46)	100.00 (100.00–100.00)
	NPV (%)	89.15 (83.78–94.51)	91.67 (86.14–97.20)	81.82 (68.66–94.98)	86.73 (80.02–93.45)	96.77 (90.55–100.00)
	AUC	0.7657 (0.6988–0.8326)	0.7621 (0.6832–0.8410)	0.7043 (0.5663–0.8423)	0.7772 (0.7034–0.8510)	0.3816 (0.2133–0.5499)
Benign (n=64)	Accuracy (%)	83.12 (77.20–89.03)	89.29 (83.56–95.01)	66.67 (52.41–80.92)	84.43 (77.99–90.86)	78.13 (63.80–92.45)
	Sensitivity (%)	71.88 (60.86–82.89)	81.82 (70.42–93.21)	50.00 (28.09–71.97)	72.73 (57.57–85.89)	70.00 (49.92–90.08)
	Specificity (%)	91.11 (85.23–94.66)	94.12 (88.53–99.71)	81.82 (65.70–97.94)	91.03 (84.68–97.37)	91.67 (76.03–100.00)
	PPV (%)	85.19 (75.71–94.66)	90.00 (84.12–95.88)	71.43 (54.70–88.16)	82.05 (74.64–89.46)	93.33 (83.56–100.00)
	NPV (%)	82.00 (74.47–89.53)	88.89 (81.63–96.15)	64.29 (46.54–82.03)	85.54 (77.98–93.11)	64.71 (41.99–87.42)
	AUC	0.7722 (0.7060–0.8384)	0.7876 (0.7119–0.8633)	0.5471 (0.3966–69.76)	0.7130 (0.6327–0.7933)	0.7206 (56.51–87.61)

Values in parentheses represent 95% confidence intervals.
EGC = early gastric cancer; PPV = positive predictive value; NPV = negative predictive value; AUC = area under the curve.

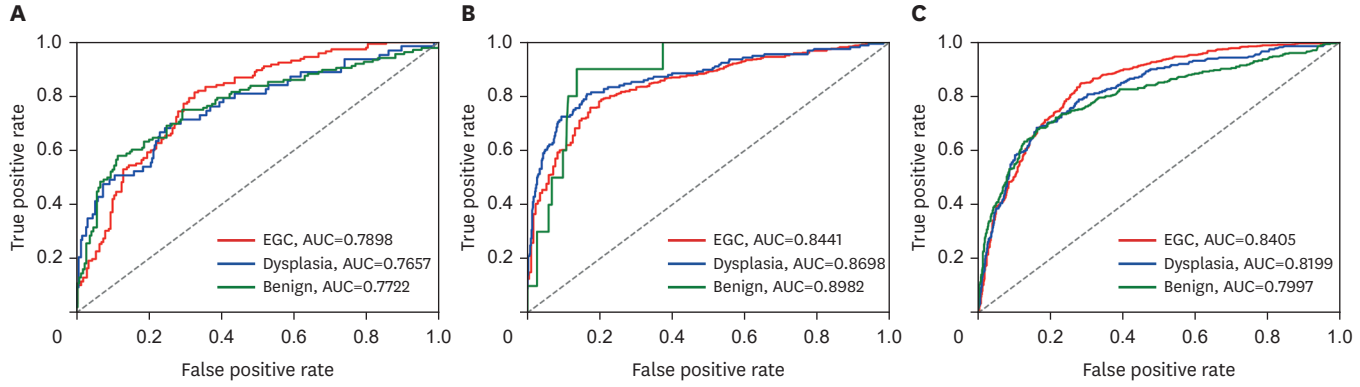

Fig. 2. Receiver operating characteristic curves demonstrating the performance of ENAD CAD-G in 3 real-world datasets. (A) Atypia dataset, (B) Internal video dataset, and (C) External video dataset.
ENAD CAD-G = ENdoscopy as AI-powered Device Computer Aided Diagnosis for Gastroscopy; EGC = early gastric cancer; AUC = area under the curve.

and 82.61% (71.66%–93.56%), respectively. ENAD CAD-G reached high AUC values of 0.8698 and 0.8982 for diagnosing dysplasia and benign lesions, respectively (**Fig. 2B**). The overall performances were comparable in the subgroup analyses stratified by age (≥55 vs. <55 years) and sex. Initial biopsy results indicated 101 cases of dysplasia and 39 cases of EGCs. Of the 55 cases in which the final diagnosis changed from dysplasia to EGC after endoscopic resection, ENAD CAD-G accurately predicted EGC in 48 cases (**Supplementary Table 5**).

In the External video set, ENAD-CAD-G could differentiate EGC, dysplasia, and benign lesions with diagnostic accuracies of 87.50% (83.73%–91.27%), 90.54% (87.21%–93.87%), and 88.85% (85.27%–92.44%), respectively (**Supplementary Table 6**). The AUC values of the External video set are shown in **Fig. 2C**. Videos are provided as supplementary material to depict representative cases diagnosed using ENAD CAD-G as neoplasms or non-neoplasms (**Supplementary Videos 1-3**).

https://doi.org/10.5230/jgc.2024.24.e28

Table 3. Comparison between the diagnostic performance of initial biopsy and ENAD CAD-G results in the Internal video dataset

Final pathology	Index	Initial biopsy	ENAD CAD-G					
			All video patients (n=140)	P-value (vs. initial biopsy)	Male (n=105)	Female (n=35)	Age ≥55 yr (n=125)	Age <55 yr (n=15)
EGC (n=94)	Accuracy (%)	60.71 (52.62–68.80)	88.57 (83.30–93.84)	<0.001	88.57 (82.49–94.66)	88.57 (78.03–99.11)	88.00 (82.30–93.70)	93.33 (80.71–100.00)
	Sensitivity (%)	41.49 (31.53–51.45)	91.49 (85.85–97.13)	<0.001	91.78 (85.48–98.08)	90.48 (77.92–100.00)	91.57 (85.59–97.54)	90.91 (73.92–100.00)
	Specificity (%)	100 (100.00–100.00)	82.61 (71.66–93.56)	<0.001	81.25 (67.73–94.77)	85.71 (67.38–100.00)	80.95 (69.08–92.83)	100.00 (100.00–100.00)
	PPV (%)	100 (100.00–100.00)	91.49 (85.85–97.13)	<0.001	91.78 (86.20–97.36)	90.48 (80.14–100.00)	90.48 (84.99–95.96)	100.00 (100.00–100.00)
	NPV (%)	45.54 (35.83–55.26)	82.61 (71.66–93.56)	<0.001	81.25 (67.73–94.77)	85.71 (87.38–100.00)	82.93 (71.41–94.44)	80.00 (44.94–100.00)
	AUC	N/A	0.8441 (0.7868–0.9014)	N/A	0.7319 (0.6472–0.8166)	0.5423 (0.3772–0.7074)	0.7645 (0.6901–0.8389)	0.5549 (0.3034–0.8064)
Dysplasia (n=45)	Accuracy (%)	60.71 (52.62–68.80)	91.43 (86.79–96.07)	<0.001	91.43 (86.07–96.78)	91.43 (82.15–100.00)	91.20 (86.23–96.17)	93.33 (80.71–100.00)
	Sensitivity (%)	100 (100.00–100.00)	75.56 (63.00–88.11)	<0.001	70.97 (54.99–86.95)	85.71 (67.38–100.00)	75.61 (62.46–88.75)	75.00 (32.56–100.00)
	Specificity (%)	42.11 (32.18–52.03)	98.95 (96.90–100.00)	<0.001	100.00 (100.00–100.00)	95.24 (86.13–100.00)	98.81 (96.49–100.00)	100.00 (100.00–100.00)
	PPV (%)	45.00 (35.25–54.75)	97.14 (91.62–100.00)	<0.001	100.00 (100.00–100.00)	92.31 (83.08–100.00)	96.88 (93.68–100.00)	100.00 (100.00–100.00)
	NPV (%)	100 (100.00–100.00)	89.52 (83.67–95.38)	0.02	89.16 (82.47–95.58)	90.91 (78.90–100.00)	89.25 (82.95–95.54)	91.67 (76.03–100.00)
	AUC	N/A	0.8698 (0.8166–0.9230)	N/A	0.8199 (0.8404–0.9560)	0.8497 (0.7313–0.9681)	0.8562 (0.7947–0.9177)	0.8042 (0.6034–1.0000)
Benign (n=1)	Accuracy (%)	97.89 (97.89–100.00)	92.86 (88.59–97.12)	<0.001	91.43 (86.07–96.78)	97.14 (91.62–100.00)	93.60 (89.31–97.89)	86.67 (69.46–100.00)
	Sensitivity (%)	0 (0.00–0.00)	100 (100.00–100.00)	<0.001	100.00 (100.00–100.00)	N/A	100.00 (100.00–100.00)	N/A
	Specificity (%)	100 (100.00–100.00)	92.81 (88.51–97.10)	<0.001	91.35 (85.94–96.75)	97.14 (97.14–100.00)	93.55 (89.22–97.87)	86.67 (69.46–100.00)
	PPV (%)	N/A	9.09 (0.00–26.08)	N/A	10.00 (4.00–16.00)	0.00 (0.00–0.00)	11.11 (5.42–16.81)	0.00 (0.00–0.00)
	NPV (%)	97.89 (97.89–100.00)	100 (100.00–100.00)	<0.001	100.00 (100.00–100.00)	100.00 (100.00–100.00)	100.00 (100.00–100.00)	100.00 (100.00–100.00)
	AUC	N/A	0.8982 (0.8504–0.9460)	N/A	0.8982 (0.8404–0.9560)	N/A	0.8982 (0.8452–0.9512)	N/A

Values in parentheses represent 95% confidence intervals.
ENAD CAD-G = ENdoscopy as AI-powered Device Computer Aided Diagnosis for Gastroscopy; EGC = early gastric cancer; PPV = positive predictive value; NPV = negative predictive value; AUC = area under the curve; N/A = not applicable.

https://doi.org/10.5230/jgc.2024.24.e28

Journal of Gastric Cancer

jgc Journal of Gastric Cancer

DISCUSSION

This novel multicenter diagnostic study applied deep learning-based CADe and CADx models to evaluate real clinical difficulties that require decisions on implementing gastric endoscopic resections. Even expert endoscopists encounter difficulties in detecting and diagnosing EGC. Moreover, it is even more difficult to differentiate EGC from dysplasia in the early stages. Therefore, the use of AI for the detection and diagnosis of EGC may be a key solution. Previous studies had limitations in the selection of training and testing datasets. Luo et al. [11] trained AI software using a single-center dataset, whereas Wu et al. [21] used training data from a single center and their test set comprised only 100 neoplasms.

The utility of AI depends not only on the accuracy of the system but also on its real-world clinical application. The rapid implementation of AI in gastroenterology was first achieved in colonoscopy; therefore, the evaluation of AI in the lower gastrointestinal tract using real-world datasets is ongoing. Contrary to high expectations, Wei et al. [22] demonstrated that endoscopists at community clinics did not benefit significantly from a CADe colonoscopy system. The need to select an appropriate test set population for the upper gastrointestinal tract to assist in real-world clinical difficulties arises from this unexpected result of human-AI interaction. Thus, we re-examined patients referred to a tertiary university hospital with initial biopsy results of gastric lesions inconclusively reported as atypia before final pathological diagnoses because >30% of gastric biopsy lesions are upgraded after endoscopic resection [23]. While all previous AI studies compared their deep learning performances with those of endoscopists, we compared the AI performance with actual biopsy results and achieved superior results across all metrics. This comparison is clinically significant because a biopsy is considered the gold standard for diagnosing gastric malignancies. For Internal Test 2, ENAD CAD-G classified the initial biopsy sites diagnosed as atypia into EGC, dysplasia, and benign, with overall accuracy, sensitivity, and specificity of 80%. The application of AI to atypical lesions has important clinical value because it facilitates accurate diagnosis at the community clinic level without further referral to tertiary hospitals. In addition to the atypia set, videos of patients who underwent endoscopic mucosal or submucosal dissection for gastric dysplasia or EGC were prospectively collected. The high AUC values of 0.8698 and 0.8982 obtained by ENAD CAD-G demonstrate the potential of our established system to help endoscopists differentiate patients who require a cancer staging workup, such as computed tomography scans, before endoscopic resection.

The substantial revalidation performance achieved by ENAD CAD-G can be attributed to several factors. First, the key strength was the diversity of the training dataset. The atypia set and Internal video test included patients from SNUBH, a tertiary university referral hospital. However, the Internal video set includes not only patients suspected of having EGC or dysplasia but also those who underwent endoscopy for dyspepsia or gastric cancer screening. Furthermore, we conducted an additional external validation using prospective video data from SNUH and HSGC. Since patients at the HSGC primarily come for check-ups, the goal was to ensure that our AI program was able to detect gastric neoplasms with low false positives. Unlike HSGC, SNUH is a tertiary referral university hospital, and its video dataset mostly consists of EGC or dysplasia cases. Therefore, we believe that the performance of our AI model can be replicated in various clinical settings. We avoided a selection bias and constructed a highly accurate AI model by extensively collecting EGD data from diverse patient populations.

https://doi.org/10.5230/jgc.2024.24.e28

jgc Journal of Gastric Cancer

Second, our technological adaptation was unique because we introduced GAN images into the gastric detection dataset. Researchers have demonstrated the potential benefits of using GAN in several areas of medical imaging to overcome the shortage of high-quality training data [24]. In upper endoscopy, Kanayama et al. [25] improved the performance of their gastric cancer detection system using GAN images. However, studies on the application of GAN images for the detection and diagnosis of gastric malignancies are limited. The proposed CNN2 achieved a higher sensitivity with additional GAN images in the detection training dataset. However, increasing the dataset by 5% and 10% with GAN-generated images did not result in consistent improvements in detection performance. Furthermore, CNN3 demonstrated a minimal effect of additional GAN images in achieving higher metrics, and the model's diagnostic performance dramatically dropped to 66.22% with 100% replacement of real endoscopic images with GAN images. Therefore, an additional analysis is required to determine whether these GAN images improve the generalization capabilities of the model without overfitting. Further research is required to determine the optimal use of generative AI data without model collapse [26].

Third, we applied OC-SORT and soft voting methods to enable the real-time detection and diagnosis of gastric lesions. Few studies have evaluated their performances in video tests [15]. Furthermore, studies that validated their systems using videos either required further technical sophistication [27] or did not comprehensively explain the technical details of video testing [21]. Our AI system takes approximately 0.1 to 0.15 seconds to process a video input, perform AI predictions, and display the results. As the AI system detects, confirms, and displays gastric lesions faster than endoscopists, real-time applications are feasible.

Despite the rigorous data collection and validation of our AI model, our study had some limitations. First, our training data consisted of highly selected retrospective images as follows: white light and high-definition refined images without noise, such as bubbles and bleeding. However, to overcome this limitation and reduce the number of false positives, we retrained our AI system by including images of foreign bodies, forceps, bubbles, fluids, and debris. Second, our AI model did not incorporate electronically enhanced images such as narrow band images. This approach was used to ensure universal applicability of the model across diverse endoscopic machines. Third, as either endoscopic images (atypia set) or videos (internal and external validation sets) were used to validate ENAD CAD-G, inter-observer variability was inevitable, which might have impacted the diagnostic accuracy. To address this issue, more randomized-controlled tandem trials are necessary [28]. Fourth, although there were no cases of the proposed CADx missing the detection of EGC or dysplasia lesions in both the internal and External video sets, there were instances where the detection of cancer/dysplasia lesions was made, but they were misclassified as benign (presented in **Supplementary Fig. 8**). Currently, AI endoscopy systems cannot fully replace expert-level endoscopy [29]. Therefore, even if AI endoscopy suggests a high probability of benign lesions, a biopsy of AI-detected lesions should be performed when necessary. Fifth, we did not analyze the effects of *Helicobacter pylori* infection on the performance of our AI system. Sixth, our analysis included only South Koreans from selected healthcare institutions. Therefore, external validation across diverse healthcare systems among different ethnic groups is warranted in the near future. Seventh, although we achieved consistent and reliable performance across the training dataset with our current labeling methodologies, the integration of techniques, such as multilayer labeling and cross-validation annotations, can further improve the data quality and model robustness. Eighth, our AI system can perform real-time detection and diagnosis of gastric lesions; however, we did not perform a real-time

https://doi.org/10.5230/jgc.2024.24.e28

jgc Journal of Gastric Cancer

clinical study using ENAD CAD-G to aid endoscopists in daily clinical practice, where the overwhelming number of cases requiring screening endoscopy will be normal. Therefore, real-time clinical trials are necessary to confirm the usefulness of ENAD CAD-G. Finally, long-term follow-up of patients with non-neoplastic lesions in the test sets is required to assess the possibility of false negatives.

In conclusion, we developed a real-time AI system termed ENAD CAD-G that can locate, detect, and diagnose gastric lesions as EGC, dysplastic, or benign lesions. The proposed AI shows promise in identifying gastric lesions that require endoscopic resection or further work-up by evaluating difficult community-based clinical cases. Prospective studies are necessary to confirm the effectiveness of using ENAD CAD-G in routine clinical practice.

SUPPLEMENTARY MATERIALS

Supplementary Table 1
Distribution of images used to train and test ENAD CAD-G and performance of ENAD CAD-G for location classification (CNN1)

Supplementary Table 2
Distribution of images used to train and test ENAD CAD-G for gastric lesion detection (CNN2) and gastric lesion classification (CNN3)

Supplementary Table 3
Performance of ENAD CAD-G on detecting gastric lesions (CNN2)

Supplementary Table 4
Performance of ENAD CAD-G for gastric lesion classification (CNN3)

Supplementary Table 5
Lesion characteristics and performance of ENAD CAD-G in cases that changed final diagnoses after endoscopic resection in the Internal video dataset

Supplementary Table 6
Diagnostic performance of ENAD CAD-G in the External video test

Supplementary Fig. 1
Schematic diagram of the ENAD CAD-G system.

Supplementary Fig. 2
Representative images of GAN used for CNN2. (A) EGC, (B) Dysplasia, (C) Erosion, and (D) Ulcer.

Supplementary Fig. 3
Flow diagram for the test sets: the Internal Test 1A, the Internal Test 1B, the Internal Test 2: Atypia, the Internal video test, and the External video test.

Supplementary Fig. 4
Video test methodology.

https://doi.org/10.5230/jgc.2024.24.e28

jgc Journal of Gastric Cancer

Supplementary Fig. 5

Learning curves of gastric lesion detection and diagnosing models. (A) Precision-recall curves of detection model, (B) Accuracy curve of diagnosis model, and (C) Train loss curves of diagnosis model.

Supplementary Fig. 6

Receiver operator characteristic curve demonstrating the performance of ENAD CAD-G in detecting gastric lesions in Internal Test 1A.

Supplementary Fig. 7

Receiver operating characteristic curves demonstrating the performance of ENAD CAD-G in diagnosing gastric lesions in Internal Test 1B.

Supplementary Fig. 8

Endoscopic images of all the cases from the Internal (n=10) and External (n=4) validation video sets that were finally diagnosed as early gastric cancer or dysplasia but were misclassified as benign by the ENAD CAD-G system.

Supplementary Video 1

Use of ENAD CAD-G for detecting and diagnosing EGC lesions. ENAD CAD-G correctly identified the EGC lesions.

Supplementary Video 2

Use of ENAD CAD-G for detecting and diagnosing gastric adenoma lesions. ENAD CAD-G correctly identified gastric dysplasia.

Supplementary Video 3

Use of ENAD CAD-G for detecting and diagnosing benign gastric lesions. ENAD CAD-G correctly identified the benign gastric ulcers.

REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. CA Cancer J Clin 2018;68:7-30. PUBMED | CROSSREF
2. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018;68:394-424. PUBMED | CROSSREF
3. Imagawa A, Okada H, Kawahara Y, Takenaka R, Kato J, Kawamoto H, et al. Endoscopic submucosal dissection for early gastric cancer: results and degrees of technical difficulty as well as success. Endoscopy 2006;38:987-990. PUBMED | CROSSREF
4. Amin MB, Greene FL, Edge SB, Compton CC, Gershenwald JE, Brookland RK, et al. The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. CA Cancer J Clin 2017;67:93-99. PUBMED | CROSSREF
5. Sano T, Coit DG, Kim HH, Roviello F, Kassab P, Wittekind C, et al. Proposal of a new stage grouping of gastric cancer for TNM classification: International Gastric Cancer Association staging project. Gastric Cancer 2017;20:217-225. PUBMED | CROSSREF
6. Pimenta-Melo AR, Monteiro-Soares M, Libânio D, Dinis-Ribeiro M. Missing rate for gastric cancer during upper gastrointestinal endoscopy: a systematic review and meta-analysis. Eur J Gastroenterol Hepatol 2016;28:1041-1049. PUBMED | CROSSREF
7. Veitch AM, Uedo N, Yao K, East JE. Optimizing early upper gastrointestinal cancer detection at endoscopy. Nat Rev Gastroenterol Hepatol 2015;12:660-667. PUBMED | CROSSREF

8. Chiu PW, Uedo N, Singh R, Gotoda T, Ng EK, Yao K, et al. An Asian consensus on standards of diagnostic upper endoscopy for neoplasia. Gut 2019;68:186-197. PUBMED | CROSSREF
9. Hamashima C; Systematic Review Group and Guideline Development Group for Gastric Cancer Screening Guidelines. Update version of the Japanese Guidelines for Gastric Cancer Screening. Jpn J Clin Oncol 2018;48:673-683. PUBMED | CROSSREF
10. Nakanishi H, Doyama H, Ishikawa H, Uedo N, Gotoda T, Kato M, et al. Evaluation of an e-learning system for diagnosis of gastric lesions using magnifying narrow-band imaging: a multicenter randomized controlled study. Endoscopy 2017;49:957-967. PUBMED | CROSSREF
11. Luo Q, Yang H, Hu B. Application of artificial intelligence in the endoscopic diagnosis of early gastric cancer, atrophic gastritis, and Helicobacter pylori infection. J Dig Dis 2022;23:666-674. PUBMED | CROSSREF
12. Renna F, Martins M, Neto A, Cunha A, Libânio D, Dinis-Ribeiro M, et al. Artificial intelligence for upper gastrointestinal endoscopy: a roadmap from technology development to clinical practice. Diagnostics (Basel) 2022;12:1278. PUBMED | CROSSREF
13. Sharma P, Hassan C. Artificial intelligence and deep learning for upper gastrointestinal neoplasia. Gastroenterology 2022;162:1056-1066. PUBMED | CROSSREF
14. Wu L, He X, Liu M, Xie H, An P, Zhang J, et al. Evaluation of the effects of an artificial intelligence system on endoscopy quality and preliminary testing of its performance in detecting early gastric cancer: a randomized controlled trial. Endoscopy 2021;53:1199-1207. PUBMED | CROSSREF
15. Nam JY, Chung HJ, Choi KS, Lee H, Kim TJ, Soh H, et al. Deep learning model for diagnosing gastric mucosal lesions using endoscopic images: development, validation, and method comparison. Gastrointest Endosc 2022;95:258-268.e10. PUBMED | CROSSREF
16. Gong EJ, Bang CS, Lee JJ, Baik GH, Lim H, Jeong JH, et al. Deep learning-based clinical decision support system for gastric neoplasms in real-time endoscopy: development and validation study. Endoscopy 2023;55:701-708. PUBMED | CROSSREF
17. Wu L, Wang J, He X, Zhu Y, Jiang X, Chen Y, et al. Deep learning system compared with expert endoscopists in predicting early gastric cancer and its invasion depth and differentiation status (with videos). Gastrointest Endosc 2022;95:92-104.e3. PUBMED | CROSSREF
18. Tan M, Le Q. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. In: Kamalika C, Ruslan S, eds. Proceedings of the 36th International Conference on Machine Learning, PMLR, Volume 97; 2019 June 9-15; Long Beach, CA, USA. MLResearchPress, 2019:6105-6114.
19. Karras T, Laine S, Aittala M, Hellsten J, Lehtinen J, Aila T. Analyzing and improving the image quality of StyleGAN. arXiv. Forthcoming 2020.
20. Cao J, Pang J, Weng X, Khirodkar R, Kitani K. Observation-centric SORT: rethinking SORT for robust multi-object tracking. arXiv. Forthcoming 2022.
21. Wu L, Xu M, Jiang X, He X, Zhang H, Ai Y, et al. Real-time artificial intelligence for detecting focal lesions and diagnosing neoplasms of the stomach by white-light endoscopy (with videos). Gastrointest Endosc 2022;95:269-280.e6. PUBMED | CROSSREF
22. Wei MT, Shankar U, Parvin R, Abbas SH, Chaudhary S, Friedlander Y, et al. Evaluation of computer-aided detection during colonoscopy in the community (AI-SEE): a multicenter randomized clinical trial. Am J Gastroenterol 2023;118:1841-1847. PUBMED | CROSSREF
23. Won CS, Cho MY, Kim HS, Kim HJ, Suk KT, Kim MY, et al. Upgrade of lesions initially diagnosed as low-grade gastric dysplasia upon forceps biopsy following endoscopic resection. Gut Liver 2011;5:187-193. PUBMED | CROSSREF
24. You A, Kim JK, Ryu IH, Yoo TK. Application of generative adversarial networks (GAN) for ophthalmology image domains: a survey. Eye Vis (Lond) 2022;9:6. PUBMED | CROSSREF
25. Kanayama T, Kurose Y, Tanaka K, Aida K, Satoh S, Kitsuregawa M, et al. Gastric cancer detection from endoscopic images using synthesis by GAN. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2019: 22nd International Conference; 2019 October 13-17; Shenzhen, China. Shenzhen: Springer-Verlag, 2019:530-538.
26. Shumailov I, Shumaylov Z, Zhao Y, Gal Y, Papernot N, Anderson R. The curse of recursion: training on generated data makes models forget. arXiv. Forthcoming 2023.
27. Horiuchi Y, Hirasawa T, Ishizuka N, Tokai Y, Namikawa K, Yoshimizu S, et al. Performance of a computer-aided diagnosis system in diagnosing early gastric cancer using magnifying endoscopy videos with narrow-band imaging (with videos). Gastrointest Endosc 2020;92:856-865.e1. PUBMED | CROSSREF
28. Wu L, Shang R, Sharma P, Zhou W, Liu J, Yao L, et al. Effect of a deep learning-based system on the miss rate of gastric neoplasms during upper gastrointestinal endoscopy: a single-centre, tandem, randomised controlled trial. Lancet Gastroenterol Hepatol 2021;6:700-708. PUBMED | CROSSREF
29. Kim JH, Nam SJ, Park SC. Usefulness of artificial intelligence in gastric neoplasms. World J Gastroenterol 2021;27:3543-3555. PUBMED | CROSSREF