식도 및 위 질환의 인공지능 적용

방창석

한림대학교 의과대학 내과학교실

Artificial Intelligence in the Analysis of Upper Gastrointestinal Disorders

Chang Seok Bang

Department of Internal Medicine, Hallym University College of Medicine, Chuncheon, Korea

In the past, conventional machine learning was applied to analyze tabulated medical data while deep learning was applied to study afflictions such as gastrointestinal disorders. Neural networks were used to detect, classify, and delineate various images of lesions because the local feature selection and optimization of the deep learning model enabled accurate image analysis. With the accumulation of medical records, the evolution of computational power and graphics processing units, and the widespread use of open-source libraries in large-scale machine learning processes, medical artificial intelligence (AI) is overcoming its limitations. While early studies prioritized the automatic diagnosis of cancer or pre-cancerous lesions, the current expanded scope of AI includes benign lesions, quality control, and machine learning analysis of big data. However, the limited commercialization of medical AI and the need to justify its application in each field of research are restricting factors. Modeling assumes that observations follow certain statistical rules, and external validation checks whether assumption is correct or generalizable. Therefore, unused data are essential in the training or internal testing process to validate the performance of the established AI models. This article summarizes the studies on the application of AI models in upper gastrointestinal disorders. The current limitations and the perspectives on future development have also been discussed. (**Korean J Helicobacter Up Gastrointest Res 2021;21:300-310**)

Key Words: Artificial intelligence; Convolutional neural network; Deep learning; Endoscopy; Gastroenterology

서론

누구나 인공지능을 사용할 수 있는 시대이다. 또한 누구나 인공지능을 필요에 의해 만들어 사용할 수 있는 시대가 다가오고 있다. 본 종설은 빠르게 발전하고 있는 인공지능의 의학분야 적용에 대한 내용을 정리하고자 기존에 저자가 제시한 상부위장관 소화기학 분야의 인공지능 적용에 대한 종설에¹⁻³ 최근에 발표된 위장관 내시경 인공지능에 대한 내용을 추가하고자 한다. 또한 인공지능 모델을 개발하거나 개발과정에 참여할 때 의료인으로서 염두에 두어야 할 내용을 기술하고자 한다.

본론

1. 상부위장관 소화기학 분야의 인공지능 적용사례

미국 소화기내시경학회(American Society of Gastrointestinal Endoscopy)에서는 2020년 10월 위장관 내시경 인공지능 모델의 개발을 위한 position statement를 발표하였고, 3가지 요소의 향상이 이루어져야 임상에 사용 가능한 인공지능 모델이라고 발표하였다.⁴ 이는 임상적인 지표의 향상(augment clinical performance), 치료계획의 향상(establish better treatment plans), 또는 환자 치료지표의 향상(improve patient outcomes)이다.⁴ 이외에도 반복적이고 많은 시간을 소모하는 단순업무처리에서 인공지능을 사용하여 의료인 고유의 의학술기를 수행할 수 있는 시간을 확보하여 업무 효율성을 높일 수 있는 방식으로 인공지능이 적용되어야 한다고 기술하였다.⁴ 이미 영상의학 분야에서는 영상판독에 판독자의 음성을 인식(speech recognition)하여 판독 레포트를 자동으로 작성하는 인공지능 모델이 적용되어 왔으며 위장관 내시경 판독기록 작성에 적용이 가능하다. 또한 내시경 검사는 검사자가 갖고 있는 지식이나 술기 능력에 따라 육안진단이나 조직 검사의 정확도(accuracy)가 달라질 수 있

Received: July 12, 2021 Revised: August 17, 2021 Accepted: August 30, 2021

Corresponding author: Chang Seok Bang

Department of Internal Medicine, Hallym University College of Medicine, 1 Hallymdaehak-gil, Chuncheon 24252, Korea

Tel: +82-33-240-5821, Fax: +82-33-241-8064, E-mail: csbang@hallym.ac.kr

*Funding for this research was provided by the Bio & Medical Technology Development Program of the National Research Foundation (NRF) & by the Korean government, Ministry of Science and ICT (MSIT) (grant number NRF2020R1F1A1071494).

Copyright © 2021 Korean College of *Helicobacter* and Upper Gastrointestinal Research

© The Korean Journal of *Helicobacter* and Upper Gastrointestinal Research is an Open-Access Journal. All articles are distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

는 특성이 있으며 학습(training)에 따라 이를 향상시킬 수 있지만 오랜 시간이 필요한 한계도 있어 인공지능 모델의 적용이 이를 보완할 수 있는지 연구가 되었다.

대부분의 초기 위장관분야 인공지능 모델 개발 연구들이 컴퓨터 비전(computer vision)을 주제로 하고 있으며 이는 병변을 자동으로 발견하거나 경계를 구별해주고 병변의 사진을 인공지능 모델에 입력하는 형태로 제시하면 정답을 예측하는 판별모델(discriminative model) 방식이다(Fig. 1).⁴ 대표적으로 position statement에서 제시한 예로는 육안으로 정확히 발견하기 어려운 암 전구 병변(pre-cancerous lesion)이나 암 자체를 발견하는 모델이며 바렛식도(Barrett's esophagus)에서 신생물의 발생을 내시경 사진에서 발견하거나 위암 또는 그 전구 병변을 내시경 사진이나 영상에서 발견(detection)하고 진단 또는 분류(classification, categorization)하는 방식이다. 이와 같은 인공지능 모델이 필요한 당위성은 실제 내시경 검사 도중 실시간으로 병변으로 의심되는 부위를 검사자에게 알려주면 이 부위를 집중하여 조직 검사 또는 절제술을 시행하는 것이 환자의 예후에 영향을 미칠 것으로 예측되기 때문이다.⁴ Position statement에서 추천하는 인공지능 학습용 데이터는 상부위장관 내시경 사진인데 백색광 내시경(white-light imaging) 이외에도 narrow-band imaging, blue-light imaging, i-SCAN, linked color imaging과 같은 영상 강조 내시경(image-enhanced endoscopy) 사진과 색소 내시경(chromoendoscopy), 확대 내시경(magnification endoscopy) 사진 등이며, 각 해당 병변의 조직학적 진단이 이루어진 사진을 다루도록 권고하고 있다. 병변의 자동 발견을 예시로 들었지만, 이는 병변의 발견이 인공지능 모델에 의해 이루어진 후 각 병변을 자동 진단 및 분류하는 것을 염두해두고 있기 때문에 학습 데이터는 범주형(category) 자료로 예를 들고 있다(예시: 1. gastritis; 2. gastric atrophy; 3. intestinal metaplasia; 4. gastric dysplasia; 5. adenocarcinoma).⁴

1) 식도암 및 식도 신생물의 내시경 인공지능 적용사례

상기 이론적인 배경을 바탕으로 현재까지 개발된 식도암이나

신생물에 대한 인공지능 모델은 신생물이나 암 자체를 발견하거나 비 신생물과 분류하여 진단하는 방식이 대표적이다(Table 1).⁵⁻²¹ 식도암이나 신생물의 경우 상부위장관 내시경 검사 도중 우연히 발견되거나 진단되는 경우가 대부분이고, 식도의 타액이나 점액, 정상 연동운동, 환자의 심장박동 및 생리적 협소부위를 고려하면 천천히 내시경을 회수하면서 면밀히 관찰하는 것이 중요하다.²² 색소 내시경이 상기 제한점을 보완할 수 있지만 이를 모든 환자에게 적용하는 것은 불가능하며, 영상 강조 내시경이 대안이 될 수 있지만 경험이 부족한 의사가 사용할 경우 민감도(sensitivity)가 떨어지는 한계가 있다.^{12,23} 미국 소화기내시경 학회는 식도에서 영상 강조 내시경이 고등급 선종(high-grade dysplasia) 또는 표재성 식도암(superficial esophageal cancer) 등의 병변에 대한 조직 검사를 대체할 수 있는 광학적 진단(optical diagnosis)을 하려면 진단 수치인 preservation and incorporation of valuable endoscopic innovations (PIVI) performance thresholds를 만족해야 한다고 제시하였는데, 환자당 분석으로 병변을 발견하는 민감도가 90%, 특이도(specificity)가 80%, 음성예측도(negative predictive value)가 98%를 넘어야 한다.²⁴ 저자가 수행한 체계적 문헌고찰 및 진단 메타분석(systematic review and diagnostic test accuracy meta-analysis)에서는 인공지능 모델이 식도암이나 신생물을 진단하는 민감도, 특이도, 음성예측도가 환자별 분석에서 각각 93%, 85%, 91%로 계산되었다(Table 1).²² 현재의 인공지능 모델은 백색광 내시경이나 영상 강조 내시경에 관계없이 또는 편평 세포암(squamous cell carcinoma)이나 선암(adenocarcinoma)에 관계없이 식도암을 구별하는 데 높은 정확도를 보여주었다.²² 전체적으로 민감도와 특이도의 경우 PIVI 역치를 만족하지만 음성예측도는 낮은 수치를 보였는데, 영상 강조 내시경 역시 아직 상기 수치를 만족하지 못하지만 인공지능 모델의 기능적인 발전을 고려하면 향후 발표될 추가적인 연구들의 결과가 기대된다. 또한 전체적인 모델의 성능이 내부검증 자료이기 때문에 일반화가 어렵고 상용화를 고려한다면 전향적인 외부검증이 필요하다(모델이 일반화되는지 알 수 있는 방법은 새로운 데

Fig. 1. Representative task of computer vision in upper gastrointestinal endoscopy. The classification task signifies the differentiation of input data into certain categories. Detection is a task where a green box is generated when the previously trained image is seen in the new image (localization). Semantic segmentation accurately identifies the boundary of the learned image without predicting a bounding box and can blur the background or distinguish the boundary. M, mucosa; SM, submucosa.

Table 1. Summary of Studies Using Artificial Intelligence in the Diagnosis of Esophageal Cancers or Neoplasms

Study	Nationality of datasets	Type of artificial intelligence	Type of endoscopic image	Type of controls	Total number of images	Number of cases in test dataset	Number of controls in test dataset
Image-based analysis							
de Groof et al. ⁵ (2020)	Europe	CNN	WLI	Non-dysplastic BE	144	33 Barrett's neoplasias	111 non-dysplastic BEs
Guo et al. ⁶ (2020)	Multi-national data	CNN	NBI	Non-cancer	6,671	1,480 precancerous and ESCCs (ESCN)	5,191 non-cancers
García-Peraza-Herrera et al. ⁷ (2020)	Asia	CNN	ME-NBI	Normal IPCL	67,740	39,662 abnormal IPCLs (ESCN)	28,078 normal IPCLs
Hashimoto et al. ⁸ (2020)	US	CNN	WLI	Non-dysplastic BE	448	225 Barrett's neoplasias	223 non-dysplastic BEs
de Groof et al. ⁹ (2020)	Multi-national data (Europe)	CNN	WLI	Non-dysplastic BE	297	129 Barrett's neoplasias	168 non-dysplastic BEs
Everson et al. ¹⁰ (2019)	Asia	CNN	ME-NBI	Normal IPCL	1,437	791 abnormal IPCLs (ESCN)	646 normal IPCLs
Cai et al. ¹¹ (2019)	Asia	CNN	WLI	Normal image	187	91 ESCN	96 normal images
Horie et al. ¹² (2019)	Asia	CNN	WLI with NBI	Non-cancer	97	47 esophageal cancers	50 non-cancers
Liu et al. ¹³ (2016)	Asia	SVM	WLI	Normal image	400	150 early esophageal cancers	250 normal images
van der Sommen et al. ¹⁴ (2016)	Europe	SVM	WLI	Non-dysplastic BE	100	60 Barrett's neoplasias	40 non-dysplastic BEs
Patient-based analysis							
de Groof et al. ⁵ (2020)	Europe	CNN	WLI	Non-dysplastic BE	20	10 Barrett's neoplasias	10 non-dysplastic BEs
Ohmori et al. ¹⁵ (2020)	Asia	CNN	Non-ME detection, ME diagnosis, NBI, BLI	Non-cancer or normal	102	52 superficial ESCC	50
Ebigbo et al. ¹⁶ (2020)	Europe	CNN	WLI	Non-dysplastic BE	62	36 early EACs	26 non-dysplastic BE
de Groof et al. ⁹ (2020)	Multi-national data (Europe)	CNN	WLI	Non-dysplastic BE	297	129 Barrett's neoplasias	168 non-dysplastic BEs
Zhao et al. ¹⁷ (2019)	Asia	CNN	ME-NBI	Non-cancerous IPCL	1,383	1,176 IPCLs (early ESCC)	207 non-cancerous IPCLs
Ebigbo et al. ¹⁸ (2019)	Europe	CNN	WLI with NBI	Non-dysplastic BE	74	33 early EACs	41 non-dysplastic BE
de Groof et al. ¹⁹ (2019)	Europe	SVM	WLI	Non-dysplastic BE	60	40 Barrett's neoplasias	20 non-dysplastic BE
Sehgal et al. ²⁰ (2018)	Europe	Decision tree algorithm	WLI	Non-dysplastic BE image	40	17 Barrett's neoplasias	23 non-dysplastic BE
van der Sommen et al. ²¹ (2014)	Europe	SVM	WLI	Non-dysplastic BE	64	32 early EAC	32 non-dysplastic BEs

CNN, convolutional neural network; WLI, white light imaging; BE, Barrett's esophagus; NBI, narrow-band imaging; BLI, blue-light imaging; ESCC, esophageal squamous cell carcinoma; ESCN, early squamous cell neoplasia; ME, magnification endoscopy; IPCL, intrapapillary capillary loop classification; SVM, support vector machine; EAC, early adenocarcinoma.

Table 2. Summary of Clinical Studies Using Artificial Intelligence in the Diagnosis of Gastric Neoplasms

Study	Nationality of datasets	Type of artificial intelligence	Type of endoscopic image	Aim of study	Design of study	Number of cases	Outcomes
Cho et al. ²⁵ (2019)	Asia	CNN	WLI	Diagnosis of gastric neoplasms	Retrospective model establishment and prospective validation	Training and testing: 5,017 images, validation: 200 images	AUCs of classifying gastric cancer: 0.877; gastric neoplasm: 0.927
Cho et al. ²⁶ (2020)	Asia	CNN	WLI	Diagnosis of depth of invasion in gastric neoplasms	Retrospective model establishment and prospective validation	Training and testing: 2,899 white-light endoscopic images, validation: 206 images	External test accuracy 77.3%
Yoon et al. ²⁷ (2019)	Asia	CNN	WLI	Classification of endoscopic images as early gastric cancer (T1a or T1b) or non-cancer	Retrospective	11,539 endoscopic images (896 T1a-, 809 T1b-, and 9,834 non-early gastric cancer)	AUC of early gastric cancer detection: 0.981, depth prediction: 0.851
Zhu et al. ²⁸ (2019)	Asia	CNN	WLI	Diagnosis of depth of invasion in gastric cancer (mucosa/SM1/deeper than SM1)	Retrospective	Training: 790 images, testing: 203 images	Accuracy: 89.2%, AUC: 0.94
Kubota et al. ²⁹ (2012)	Asia	ANN	WLI	Diagnosis of depth of invasion in gastric cancer	Retrospective	902 images	Accuracy: 77.2%, 49.1%, 51.0%, and 55.3% for T1-4 staging, respectively
Hirasawa et al. ³⁰ (2018)	Asia	CNN	WLI, chromoendoscopy, narrow-band imaging	Detection of gastric cancers	Retrospective	Training: 13,584 images, testing: 2,296 images	Accurate detection rate with a diameter of 6 mm or more: 98.6%
Kanesaka et al. ³¹ (2018)	Asia	SVM	Magnifying narrow-band imaging	Diagnosis and delineation of early gastric cancer using magnifying narrow-band imaging images	Retrospective	Training: 126 images, testing: 81 images	Accuracy: 96.3%
Lee et al. ³² (2019)	Asia	CNN	WLI	Classification of normal, benign ulcer, and gastric cancer	Retrospective	200 normal, 367 cancer, and 220 ulcer cases	Accuracy: normal vs. ulcer/normal vs. cancer: above 90%; ulcer vs. cancer: 77.1%

Modified from tables in the study of Bang et al.¹

CNN, convolutional neural network; WLI, white light imaging; AUC, area under the curve; SM, submucosa; ANN, artificial neural network; SVM, support vector machine.

이터에 적용해 보는 것이며 외부검증에서 성능이 낮다면 이는 모델이 학습 데이터에 과대적합(overfitting)되었다는 것을 의미한다).

2) 위암 및 위 신생물의 내시경 인공지능 적용사례

두 번째로 위암이나 위 신생물의 경우 상부위장관 내시경 검사 도중 육안 소견으로 병변을 의심하고 조직 검사용 검자(biopsy forcep)를 이용한 병리학적 분석을 통해 일차 진단이 이루어진다. 하지만 육안 소견과 병리 검사가 일치하는 비율은 다양하며, 조직 검사를 통한 일차 진단과 수술 및 내시경 절제술을 통한 치료 후 최종 병리 소견의 불일치 비율 또한 병변의 상태와 모양에 따라 다양하게 보고되고 있다. 또한 내시경 검사 도중 병변을 발견하지 못하고 놓칠 경우 환자의 예후에 직접 영향을 미치게 된다. 진단 이외에도 위 신생물의 치료방침을 결정하기 위해서는 위벽 침윤 깊이를 예측해야 한다. 이는 주로 의사의 육안 소견과 내시경 초음파(endoscopic ultrasound) 등의 검사를 통해 판단하는데, 위 신생물이 점막내에 국한되었거나 점막하 침윤 깊이가 500 μm 이하인 경우에만 내시경 절제술의 대상이 된다. 하지만 육안 소견의 정확도는 개인에 따라 차이가 크고 내시경 절제술 이후에 점막하 침윤이 깊은 암으로 최종 판단되는 경우 추가적인 위 절제 수술이 필요하여 위 신생물의 위벽 침윤 깊이 판단에 마충족 의료요구가 있는 실정이다. 현재까지 개발된 위 신생물 및 위암에 대한 인공지능 모델의 경우 위 신생물 자체를 비 신생물과 (또는 위암을 분류) 분류하여 진단하는 모델과²⁵ 조기위암의 침윤 깊이를 진단하여 내시경 치료의 대상이 될 가능성이 있는지 확인하는 모델이 개발되었다.²⁶⁻²⁹ 또한 위암의 발견과^{30,31} 범위를 지정하는 연구³¹ 및 궤양형 병변 중에서 위암을 구별하는 연구에³² 인공지능 모델이 활용되었으며, 합성곱 신경망(convolutional neural network)을 이용한 딥러닝(deep learning) 학습법 또는 support vector machine 방식의 머신러닝(machine learning) 모델이 적용되었다(Table 2). 소화기 내시경분야에서 인공지능 모델이 가장 활발히 개발되어 있는 대장 용종(colon polyp)의 발견이나 진단 분야와 비교해서 위암이나 위 신생물의 경우 주변 점막의 염증의 정도나 암종의 조직학적 분화도 또는 임상 병기에 따라 다양한 외형을 보이기 때문에 인공지능 모델의 개발이 상대적으로 어렵다.

이전 저자의 연구에서 상부위장관 내시경 사진을 자동으로 진행성 위암(advanced gastric cancer), 조기위암(early gastric cancer), 고등급 및 저등급 선종(high-grade or low-grade dysplasia), 양성 병변(non-neoplasm)의 5단계로 분류하는 인공지능 모델을 생성하였다.²⁵ 이는 사진에서 자동으로 병변의 조직학적 진단을 제시함으로써 시술자에게 조직학적 진단의 필요성과 치료적 계획을 세우도록 이차 의견을 제시하는 역할을 수행

할 수 있다. 인공지능 모델이 5가지 병변을 구별하는 평균 내부시험(internal-test) 정확도는 84.6%였고 외부시험(external-test) 정확도는 76.4%였다.²⁵ 위 신생물의 침윤 깊이의 경우 인공지능 모델이 점막내 국한된 병변과 점막하 이상을 침범한 병변을 구별하는 외부시험 정확도가 77.3%였다.²⁶ 후향적인 임상 simulation에서 인공지능 모델을 사용하여 위 신생물에 대한 수술적 치료와 내시경 절제술을 결정한다고 가정할 경우 일부 임상적인 지표의 향상이 있음을 보고하였지만 실제 임상에 적용하기에는 정확도가 다소 낮은 제한점이 있었다. 이를 보완하기 위해 인공지능 모델의 생성시 하이퍼 변수(hyperparameter)를 자동으로 최적화해주고 복잡한 코딩이 없이도 단순한 drag & drop 방식이나 사용자 인터페이스(user interface)의 버튼을 누르는 방식으로 인공지능 모델을 생성하는 no-code platform을 이용한 연구를 수행하였다.³³ 이전에 사용하였던 같은 학습용 사진을 이용하여 위 신생물의 침윤 깊이를 판단하는 인공지능 모델을 생성하였고 외부시험 정확도가 89.3%로 기존 모델의 77.3%에 비해 향상되었다.³³ 모델의 생성 시간 역시 수개월에서 수주 또는 수일내로 단축되었다. 현재는 향상된 성능의 일반화가 가능한지 확인하기 위해 전국 다기관 연구로 성능에 대한 외부검증이 이루어지고 있다.

인공지능 모델이 정확한 답을 제시한다고 가정하더라도 실제 임상 현장에서 내시경 검사를 시행하는 의사는 어떤 반응을 나타낼 지 알 수 없다. 이를 확인하기 위해 내시경 의사의 숙련도에 따라 소화기 내시경 전문의, 전임의, 수련의로 분류하여 각각 200여 장의 위 신생물 사진에 대한 침윤 깊이를 예측하는 연구를 하였다.³³ 총 3회의 시험을 하였는데, 첫 번째는 인공지능 모델의 도움 없이 스스로 정답을 예측하게 하였고, 두 번째는 정확도가 50% 밖에 되지 않는 부정확한 인공지능 모델의 답을 생성하여 시험자에게 알려주고 답안의 변화가 있는지 확인하였다. 세 번째는 본 연구에서 생성된 외부시험 정확도가 89.3%인 인공지능 모델의 답을 알려주고 시험자의 답안에 변화가 있는지 확인하였다. 위 신생물의 침윤 깊이를 예측하는 정확도는 소화기 내시경 전문의, 전임의, 수련의 순으로 높았으며, 총 3회의 시험 중 내시경 전문의의 정확도는 인공지능 모델의 정답에 관계없이 통계학적 차이가 없는 일정한 정확도를 유지하였다. 하지만 전임의와 수련의는 부정확한 인공지능 모델의 정답을 확인하고 침윤 깊이 예측 정확도가 감소하였고, 정확도가 높은 인공지능 모델의 정답을 확인한 후 예측 정확도가 상승하는 현상을 보였다. 예측 정확도의 상승 정도는 전임의에서 수련의에 비해 더욱 높아서 소화기 내시경 사진 판별 인공지능 모델의 경우 숙련도 중 또는 하의 내시경 의사가 가장 큰 이득을 얻을 수 있을 것으로 추측하였다.³³

Table 3. Summary of Studies Using Artificial Intelligence in the Diagnosis of *Helicobacter pylori* (*H. pylori*) Infections in Endoscopic Images

Study	Nationality of datasets	Type of artificial intelligence	Type of endoscopic image	Diagnostic method of <i>H. pylori</i> infection	Number of cases in test dataset	Number of controls in test dataset	Unit of analysis
Yasuda et al. ⁴⁸ (2020)	Japan	SVM	Linked color imaging	More than 2 different tests in each case (histology, serum antibody, stool antigen, urea breath test)	42 <i>H. pylori</i> patients	63 controls (46 post-eradication patients and 17 uninfected patients)	Patient-based
					210 <i>H. pylori</i> positive images	315 control images (230 post-eradication and 85 uninfected images)	Image-based
					210 <i>H. pylori</i> positive images	85 uninfected images (<i>H. pylori</i> naïve)	Image-based (infected vs. uninfected)
					210 <i>H. pylori</i> positive images	230 after eradication images	Image-based (infected vs. after-eradication)
Zheng et al. ⁴⁹ (2019)	China	CNN	WLI	Histology with immunohistochemistry (if negative, urea breath test was done)	85 uninfected images	230 after eradication images	Image-based (uninfected vs. after-eradication)
					2,575 <i>H. pylori</i> positive images	1,180 control images (whether post-eradication or uninfected images are unknown)	Image-based
					70 <i>H. pylori</i> positive patients	777 controls (284 post-eradication and 493 uninfected images)	Patient-based
					59 <i>H. pylori</i> positive images	477 uninfected images (<i>H. pylori</i> naïve)	Image-based (infected vs. uninfected)
Shichijo et al. ⁵⁰ (2019)	Japan	CNN	WLI	Serum or urine antibody, stool antigen, urea breath test	55 <i>H. pylori</i> positive images	182 after eradication images	Image-based (infected vs. after-eradication)
					481 uninfected images	249 after eradication images	Image-based (uninfected vs. after-eradication)
					30 <i>H. pylori</i> patients	30 controls (uninfected patients) (<i>H. pylori</i> naïve)	Patient-based
							Patient-based
Nakashima et al. ⁵¹ (2018)	Japan	CNN	WLI	Serum antibody (<i>H. pylori</i> IgG ≥ 10 U/mL was considered positive)			Patient-based
			Linked color imaging				Patient-based
			Blue laser imaging -bright				Patient-based
Itoh et al. ⁵² (2018)	Japan	CNN	WLI	Serum antibody (<i>H. pylori</i> IgG ≥ 10 U/mL was considered positive)	15 <i>H. pylori</i> positive images	15 control images (uninfected patients) (<i>H. pylori</i> naïve)	Image-based

Table 3. Continued

Study	Nationality of datasets	Type of artificial intelligence	Type of endoscopic image	Diagnostic method of <i>H. pylori</i> infection	Number of cases in test dataset	Number of controls in test dataset	Unit of analysis
Shichijo et al. ⁵³ (2017)	Japan	CNN	WLI	Serum or urine antibody, stool antigen, urea breath test	72 <i>H. pylori</i> patients	325 controls (uninfected patients) (<i>H. pylori</i> naïve)	Patient-based
Huang et al. ⁵⁴ (2008)	Taiwan	Sequential forward floating selection with SVM	WLI	Histology (3 pairs of samples from the topographic sites, including antrum, body, and cardia were obtained in a uniform way)	130 <i>H. pylori</i> patients	106 controls (whether post-eradication or uninfected patients are unknown)	Patient-based
Huang et al. ⁵⁵ (2004)	Taiwan	Refined feature selection with neural network	WLI	Histology (3 pairs of samples from the topographic sites, including antrum, body, and cardia were obtained in a uniform way)	41 <i>H. pylori</i> patients	33 controls (whether post-eradication or uninfected patients are unknown)	Patient-based

Modified from tables in the study of Bang et al.³⁹

SVM, support vector machine; CNN, convolutional neural network; WLI, white light imaging.

3) 기타 최신 내시경 인공지능 적용사례

최근에 해당 분야에 출판된 연구로는 내시경 절제 대상이 되는 조기위암 병변의 indigo carmine 색소 내시경과 백색광 내시경 사진으로 암의 경계(절제면 경계; resection margin)를 인공지능 모델에 학습시키고 확대 내시경과 narrow-band imaging 사진으로 분류능력을 시험하는 연구이다.³⁴ 내부시험 정확도는 색소 내시경 사진으로 학습시킨 인공지능 모델의 경우 85.7%, 백색광 내시경 사진으로 학습시킨 인공지능 모델의 경우 88.9%였다. 내부시험에서 전문가에 의해 수행된 병변과 절제면 marking의 최소거리는 3.32 ± 2.32 mm였고, 인공지능 모델에 의해 수행된 병변과 절제면 marking의 최소거리는 3.40 ± 1.49 mm로 통계적인 차이는 없다고 보고하였다. 또한 확대 내시경과 narrow-band imaging을 이용해서 조기위암의 분화도(differentiation)를 예측하는 인공지능 모델에 대한 연구가 발표되었다.³⁵ 내부시험 정확도가 83.3%였고 5명의 전문가와 인공지능 모델을 비교시험하였을 때 인공지능 모델이 통계학적으로 더 조기위암의 분화도를 잘 구별하는 것으로 발표되었다(정확도; 86.2%, vs. 69.7%). 진행성 위암의 경우 복막전이(peritoneal metastasis)가 때때로 진단이 되지만 수술적인 치료 전에는 이를 완벽히 예측하기가 어렵다. 수술 전 전산화단층촬영(CT) 사진을 학습하여 복막전이(occult peritoneal metastasis)를 예측하는 인공지능 모델이 발표되었고 외부시험에서 민감도 87.5%, 특이도 98.2%의 성적을 발표하였다.³⁶

의학 인공지능 모델의 생성 후에 의사와 인공지능 모델의 진단 성적을 비교하는 연구(man-machine contest)가 다수 출판되었다.²⁵ 연구 결과 대부분 인공지능 모델이 의사들에 비해 더 높은 성능을 보이거나 통계학적으로 차이가 없는 정도의 성능을 보여주었다.²² 각 연구 주제에 따라 비슷한 연구 결과들이 축적되면서 인공지능 모델이 잘 분류할 수 있는 분야에 대한 성능을 사람과 비교하는 것은 의미가 퇴색되고 있긴 하지만, 소화기 내시경 의사들의 진단 성적을 일률적으로 수치화하는 것도 어려운 것이 사실이다. 최근 의사와 인공지능 모델의 진단 성적을 비교하는 연구들을 체계적 문헌고찰을 통해 정리하고 분석하여 의사들의 상부위장관 신생물에 대한 분류 정확도, 민감도, 특이도 등을 제시한 연구가 출판되었다.³⁷ 상부위장관의 특정 분야에서 인공지능 모델의 생성을 염두에 두고 있다면 해당 연구의 결과를 참조하여 인공지능 모델의 성능 목표치를 제시하는 것도 고려해 볼만 하다.

2. 향후 발전 방향과 한계점

최근에는 암이나 암 전구 병변, 신생물 이외에도 양성 병변(benign lesions)의 진단에 인공지능이 적용되는 예가 발표되고 있다. 식도와 위의 정맥류(esophageal or gastric varices)의 경우 출혈(bleeding) 소견이 있는 경우 환자의 예후와 직접적인 연관이 있는데, 이를 자동으로 발견하고 크기, 형태, 색조, 출혈 가

능성을 분석하는 인공지능 모델이 발표되었다.³⁸ 이 모델은 정맥류를 발견하는 정확도가 의사보다 높았고, 출혈의 위험 요소를 판단하는 정확도는 의사와 비슷하다고 보고하였다.³⁸ 상부위장관 내시경 사진에서 헬리코박터 파일로리(*Helicobacter pylori*, *H. pylori*) 감염을 진단하는 인공지능 모델이 개발되었다. 8개의 개별 연구에 대한 진단 메타분석에서는 인공지능 모델이 *H. pylori* 감염을 내시경 사진에서 진단하는 민감도, 특이도가 환자별 분석에서 각각 87%, 86%로 계산되었다(Table 3).³⁹ 일반적인 상부위장관 내시경 이외에도 캡슐 내시경(wireless capsule endoscopy)을 이용하여 위 내의 궤양(ulcer), 용종, 출혈과 식도염(esophagitis) 등을 진단하는 인공지능 모델이 발표되었고 내부시험 정확도는 96.5%였다.⁴⁰ Magnetically controlled capsule endoscopy과 faster region-based convolutional neural network를 이용하여 위 내의 미란(erosion), 궤양, 용종, 점막하 종양(submucosal tumor), 황색종(xanthoma), 정상점막(normal mucosa) 등을 진단하는 인공지능 모델도 발표되었고 내부시험 정확도는 77.1%였다.⁴¹

전체적으로 위장관 내시경 영상기술의 발전으로 병변의 표면 구조 모양이나 혈관구조를 확대해서 확인하는 기술이 가능하기 때문에 인공지능 모델이 학습할 만한 특성(feature)을 더 많이 가진 영상 강조 내시경 사진이나 영상을 이용한 인공지능 모델 개발이 지속적으로 시도되리라 생각한다.

업무의 효율성이나 질을 향상시키는 연구들은 내시경 검사 도중 맹점(blind spot)을 모니터링해주는 인공지능 모델 개발을 예로 들 수 있다. Wu 등⁴²은 상부위장관 내시경 검사 도중 실시간으로 맹점을 모니터링해주는 인공지능 모델을 개발하여 무작위 대조군 연구를 통해 검사에 사용할 경우 맹점을 줄일 수 있음을 발표하였다. 또한 같은 연구그룹에서 의식화 진정 내시경에 인공지능 모델을 사용했을 경우 비 진정 상태의 검사보다 맹점을 더욱 줄일 수 있음을 발표하였다.⁴³ 최근에는 다기관 무작위 대조 연구를 통해 실제 진료에 인공지능 모델을 사용하였을 때 맹점을 줄이고 신생물도 높은 정확도로 진단할 수 있음을 확인하였다(병변당 정확도 84.69%, 민감도: 100%, 특이도: 84.29%).⁴⁴ 국내에서도 유럽 소화기내시경학회(European Society of Gastrointestinal Endoscopy)에서 제시한 8장의 표준 상부위장관 내시경 촬영 부위를 gold standard로 정하고, 이 8장의 사진을 자동으로 분류하여 이를 모두 촬영하였으면 내시경 검사가 충실히 시행된 것으로 판단하는 인공지능 모델을 개발하였다. 8장의 표준 사진을 분류하는 내부시험 정확도는 97.58%였고 검사가 충실히 시행되었는지 확인하는 내부시험 정확도는 89.2%였다.⁴⁵

사진이나 영상 이외에 빅데이터를 기반으로 분석하여 위암의 발생 가능성이나 치료 지표를 예측하는 연구들이 발표되었다.

나이와 성별, 혈액검사, 암표지자 등을 조합하여 위암의 발생을 예측하는 연구가 발표되었다.⁴⁶ 이런 예측모델을 생성하는 것은 기존에도 통계학적인 모델링을 통해 시행되어 왔으나 본 연구에서는 머신러닝 분석기법 중 gradient boosting decision tree를 이용하여 모델을 생성하였고 내부시험 정확도가 83%로 보고되었다. 미분화형 위암(early gastric cancer with undifferentiated histology)의 경우 내시경 절제 전에 치료적 절제(curative resection)를 예측하기 어려운 경우가 많은데 국내의 다기관 데이터를 이용하여 미분화형 위암의 치료적 절제를 예측하는 모델이 발표되었다.⁴⁷ 환자의 나이, 성별, 병변의 크기, 위치, 궤양 동반 유무 등을 바탕으로 예측모델이 생성되었고 extreme gradient boosting 분류모델의 경우 내부시험 정확도가 93.4%, 외부시험 정확도가 89.8%로 보고되었다.

결국에는 생성적 적대 신경망(generative adversarial network), diffusion model과 같은 생성모델(generative model)이 인공지능 모델이 수행할 수 있고 사람이 일반적으로 생각하지 못한 창의적인 업무를 수행하게 될 가능성이 높다. 향후 딥러닝에서 특징적으로 성능 향상을 보이는 앙상블기법(ensemble, 단일 모델 대신 여러 모델의 출력값을 평균하여 결과를 도출)이나 지식 증류(knowledge distillation, 작은 모델(student model)로 큰 모델(teacher model)과 비슷한 성능을 낼 수 있도록 학습), 자가 증류(self-distillation, 별도의 teacher model 없이 스스로 distillation을 수행) 등에 대한 추가 연구를 통해 수리적인 방식으로 딥러닝의 작동방식을 좀 더 깊이 이해할 수 있리라 기대된다.

3. 인공지능 모델을 개발 시 고려할 점

실제 임상에 존재하는 영상이나 환자 데이터는 질병이나 예후를 시사하는 소견을 포함하고 있지만 노이즈가 가득한 정제되지 않은 데이터이다. 또한 질병상태 또는 이상 병변보다는 정상 소견이 절대적으로 많은 의학데이터의 특성을 고려할 때 인공지능이 주로 학습하기를 기대하는 소견은 전체 데이터 중에서 소량인 경우가 많다. 인공지능 모델 개발을 원하는 데이터가 가진 특성에 따라 어떤 방식의 인공지능 모델을 적용해야 하는지 다를 수 있으며 실제 사용을 위해서는 모델 배포(model deployment)까지도 고려해야 한다. 또한 현재는 대부분의 연구자들이 특정 인공지능 모델의 성능을 높이려는 작업에 집중하고 있지만, 인공지능 모델의 생성에 사용된 데이터에 대한 분석으로 초점을 바꿀 필요가 있다. 즉 데이터 중심의 인공지능 개발인데, 어떤 재료로 학습하는지에 따라 모델의 성과와 그 사용방식이 달라진다. 의학 데이터의 특성상 실제 임상현장에서 사용할 수 있으려면 분류의 대상이 되는 특징이 충분히 존재하는

학습 데이터를 구축해야 한다. 학습 데이터를 어떤 방식으로 레이블링(labeling)하는지에 따라 실제 사용 시에 유용성이 달라질 수 있다. 기술적인 방식으로 모델의 성능을 향상시키는 것도 중요하지만 데이터가 갖고 있는 특성을 이해하고 데이터의 분포나 질에 따른 성능의 변화에 초점을 맞추어 현실에 사용하기에 가장 적합한 데이터를 가공하는 것이 중요하다.

결 론

딥러닝을 이용하여 내시경 영상이나 사진을 바탕으로 암이나 그 전구 병변인 신생물의 자동발견, 진단 및 병변의 경계를 확인하는 컴퓨터 비전 분석법은 높은 정확도를 보이고 있다. 하지만 전향적인 외부 검증 연구가 부족하고 임상적인 이득이 있는지 확인하는 연구가 부족하다. 개발된 대부분의 딥러닝 모델들은 연구 목적으로 생성되어 실용성이 떨어지는 단점이 있지만 최근에는 실제 내시경 의사의 검사를 방해하지 않으면서 원활하게 정보를 제공하는 딥러닝 모델의 개발이 이루어지고 있다. 내시경 영상 분석은 검사 도중 실시간으로 분석이 필요하기 때문에 모델의 경량화 및 전용 모듈이 필요한데 이에 대한 연구가 증가하리라 생각된다. 영상 강조 내시경 및 확대 내시경 기법 등 위장관 내시경 영상의 질 또한 향상되고 있기 때문에 이에 따라 개발되는 딥러닝 모델의 수준 또한 향상될 가능성이 높다. 향후에는 임상 의사결정지원 시스템(clinical decision support system)의 일부로서 이런 영상 분석 딥러닝 모델이 확장되어 사용되리라 예측된다.

CONFLICT OF INTEREST

No potential conflict of interest relevant to this article was reported.

ORCID

Chang Seok Bang <https://orcid.org/0000-0003-4908-5431>

REFERENCES

1. Bang CS. Deep learning in upper gastrointestinal disorders: status and future perspectives. *Korean J Gastroenterol* 2020;75:120-131.
2. Yang YJ, Bang CS. Application of artificial intelligence in gastroenterology. *World J Gastroenterol* 2019;25:1666-1683.
3. Cho BJ, Bang CS. Artificial intelligence for the determination of a management strategy for diminutive colorectal polyps: hype, hope, or help. *Am J Gastroenterol* 2020;115:70-72.
4. Berzin TM, Parasa S, Wallace MB, Gross SA, Repici A, Sharma P. Position statement on priorities for artificial intelligence in GI endoscopy: a report by the ASGE Task Force. *Gastrointest Endosc* 2020;92:951-959.
5. de Groof AJ, Struyvenberg MR, Fockens KN, et al. Deep learning algorithm detection of Barrett's neoplasia with high accuracy during live endoscopic procedures: a pilot study (with video). *Gastrointest Endosc* 2020;91:1242-1250.
6. Guo L, Xiao X, Wu C, et al. Real-time automated diagnosis of precancerous lesions and early esophageal squamous cell carcinoma using a deep learning model (with videos). *Gastrointest Endosc* 2020;91:41-51.
7. García-Peraza-Herrera LC, Everson M, Lovat L, et al. Intrapapillary capillary loop classification in magnification endoscopy: open dataset and baseline methodology. *Int J Comput Assist Radiol Surg* 2020;15:651-659.
8. Hashimoto R, Requa J, Dao T, et al. Artificial intelligence using convolutional neural networks for real-time detection of early esophageal neoplasia in Barrett's esophagus (with video). *Gastrointest Endosc* 2020;91:1264-1271.e1.
9. de Groof AJ, Struyvenberg MR, van der Putten J, et al. Deep-learning system detects neoplasia in patients with Barrett's esophagus with higher accuracy than endoscopists in a multi-step training and validation study With Benchmarking. *Gastroenterology* 2020;158:915-929.e4.
10. Everson M, Herrera L, Li W, et al. Artificial intelligence for the real-time classification of intrapapillary capillary loop patterns in the endoscopic diagnosis of early oesophageal squamous cell carcinoma: a proof-of-concept study. *United European Gastroenterol J* 2019;7:297-306.
11. Cai SL, Li B, Tan WM, et al. Using a deep learning system in endoscopy for screening of early esophageal squamous cell carcinoma (with video). *Gastrointest Endosc* 2019;90:745-753.e2.
12. Horie Y, Yoshio T, Aoyama K, et al. Diagnostic outcomes of esophageal cancer by artificial intelligence using convolutional neural networks. *Gastrointest Endosc* 2019;89:25-32.
13. Liu DY, Gan T, Rao NN, et al. Identification of lesion images from gastrointestinal endoscope based on feature extraction of combinational methods with and without learning process. *Med Image Anal* 2016;32:281-294.
14. van der Sommen F, Zinger S, Curvers WL, et al. Computer-aided detection of early neoplastic lesions in Barrett's esophagus. *Endoscopy* 2016;48:617-624.
15. Ohmori M, Ishihara R, Aoyama K, et al. Endoscopic detection and differentiation of esophageal lesions using a deep neural network. *Gastrointest Endosc* 2020;91:301-309.e1.
16. Ebigbo A, Mendel R, Probst A, et al. Real-time use of artificial intelligence in the evaluation of cancer in Barrett's oesophagus. *Gut* 2020;69:615-616.
17. Zhao YY, Xue DX, Wang YL, et al. Computer-assisted diagnosis of early esophageal squamous cell carcinoma using narrow-band imaging magnifying endoscopy. *Endoscopy* 2019;51:333-341.
18. Ebigbo A, Mendel R, Probst A, et al. Computer-aided diagnosis

- using deep learning in the evaluation of early oesophageal adenocarcinoma. *Gut* 2019;68:1143-1145.
19. de Groof J, van der Sommen F, van der Putten J, et al. The Argos project: the development of a computer-aided detection system to improve detection of Barrett's neoplasia on white light endoscopy. *United European Gastroenterol J* 2019;7:538-547.
 20. Sehgal V, Rosenfeld A, Graham DG, et al. Machine learning creates a simple endoscopic classification system that improves dysplasia detection in Barrett's oesophagus amongst non-expert endoscopists. *Gastroenterol Res Pract* 2018;2018:1872437.
 21. van der Sommen F, Zinger S, Schoon EJ, Witha PHN. Supportive automatic annotation of early esophageal cancer using local gabor and color features. *Neurocomputing* 2014;144:92-106.
 22. Bang CS, Lee JJ, Baik GH. Computer-aided diagnosis of esophageal cancer and neoplasms in endoscopic images: a systematic review and meta-analysis of diagnostic test accuracy. *Gastrointest Endosc* 2021;93:1006-1015.e13.
 23. Ishihara R, Takeuchi Y, Chatani R, et al. Prospective evaluation of narrow-band imaging endoscopy for screening of esophageal squamous mucosal high-grade neoplasia in experienced and less experienced endoscopists. *Dis Esophagus* 2010;23:480-486.
 24. Sharma P, Savides TJ, Canto MI, et al. The American society for gastrointestinal endoscopy PIVI (preservation and incorporation of valuable endoscopic innovations) on imaging in Barrett's Esophagus. *Gastrointest Endosc* 2012;76:252-254.
 25. Cho BJ, Bang CS, Park SW, et al. Automated classification of gastric neoplasms in endoscopic images using a convolutional neural network. *Endoscopy* 2019;51:1121-1129.
 26. Cho BJ, Bang CS, Lee JJ, Seo CW, Kim JH. Prediction of submucosal invasion for gastric neoplasms in endoscopic images using deep-learning. *J Clin Med* 2020;9:1858.
 27. Yoon HJ, Kim S, Kim JH, et al. A lesion-based convolutional neural network improves endoscopic detection and depth prediction of early gastric cancer. *J Clin Med* 2019;8:1310.
 28. Zhu Y, Wang QC, Xu MD, et al. Application of convolutional neural network in the diagnosis of the invasion depth of gastric cancer based on conventional endoscopy. *Gastrointest Endosc* 2019;89:806-815.e1.
 29. Kubota K, Kuroda J, Yoshida M, Ohta K, Kitajima M. Medical image analysis: computer-aided diagnosis of gastric cancer invasion on endoscopic images. *Surg Endosc* 2012;26:1485-1489.
 30. Hirasawa T, Aoyama K, Tanimoto T, et al. Application of artificial intelligence using a convolutional neural network for detecting gastric cancer in endoscopic images. *Gastric Cancer* 2018;21:653-660.
 31. Kanesaka T, Lee TC, Uedo N, et al. Computer-aided diagnosis for identifying and delineating early gastric cancers in magnifying narrow-band imaging. *Gastrointest Endosc* 2018;87:1339-1344.
 32. Lee JH, Kim YJ, Kim YW, et al. Spotting malignancies from gastric endoscopic images using deep learning. *Surg Endosc* 2019;33:3790-3797.
 33. Bang CS, Lim H, Jeong HM, Hwang SH. Use of endoscopic images in the prediction of submucosal invasion of gastric neoplasms: automated deep learning model development and usability study. *J Med Internet Res* 2021;23:e25167.
 34. An P, Yang D, Wang J, et al. A deep learning method for delineating early gastric cancer resection margin under chromoendoscopy and white light endoscopy. *Gastric Cancer* 2020;23:884-892.
 35. Ling T, Wu L, Fu Y, et al. A deep learning-based system for identifying differentiation status and delineating the margins of early gastric cancer in magnifying narrow-band imaging endoscopy. *Endoscopy* 2021;53:469-477.
 36. Jiang Y, Liang X, Wang W, et al. Noninvasive prediction of occult peritoneal metastasis in gastric cancer using deep learning. *JAMA Netw Open* 2021;4:e2032269.
 37. Frazzoni L, Arribas J, Antonelli G, et al. Endoscopists' diagnostic accuracy in detecting upper gastrointestinal neoplasia in the framework of artificial intelligence studies. *Endoscopy* 2015. doi: 10.1055/a-1500-3730. [Epub ahead of print]
 38. Chen M, Wang J, Xiao Y, et al. Automated and real-time validation of gastroesophageal varices under esophagogastroduodenoscopy using a deep convolutional neural network: a multicenter retrospective study (with video). *Gastrointest Endosc* 2021;93:422-432.e3.
 39. Bang CS, Lee JJ, Baik GH. Artificial intelligence for the prediction of *Helicobacter pylori* infection in endoscopic images: systematic review and meta-analysis of diagnostic test accuracy. *J Med Internet Res* 2020;22:e21983.
 40. Majid A, Khan MA, Yasmin M, Rehman A, Yousafzai A, Tariq U. Classification of stomach infections: a paradigm of convolutional neural network along with classical features fusion and selection. *Microsc Res Tech* 2020;83:562-576.
 41. Xia J, Xia T, Pan J, et al. Use of artificial intelligence for detection of gastric lesions by magnetically controlled capsule endoscopy. *Gastrointest Endosc* 2021;93:133-139.e4.
 42. Wu L, Zhang J, Zhou W, et al. Randomised controlled trial of WISENSE, a real-time quality improving system for monitoring blind spots during esophagogastroduodenoscopy. *Gut* 2019;68:2161-2169.
 43. Chen D, Wu L, Li Y, et al. Comparing blind spots of unsedated ultrafine, sedated, and unsedated conventional gastroscopy with and without artificial intelligence: a prospective, single-blind, 3-parallel-group, randomized, single-center trial. *Gastrointest Endosc* 2020;91:332-339.e3.
 44. Wu L, He X, Liu M, 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.
 45. Choi SJ, Khan MA, Choi HS, et al. Development of artificial intelligence system for quality control of photo documentation in esophagogastroduodenoscopy. *Surg Endosc* 2021. doi: 10.1007/s00464-020-08236-6. [Epub ahead of print]
 46. Zhu SL, Dong J, Zhang C, Huang YB, Pan W. Application of machine learning in the diagnosis of gastric cancer based on non-invasive characteristics. *PLoS One* 2020;15:e0244869.
 47. Bang CS, Ahn JY, Kim JH, Kim YI, Choi IJ, Shin WG. Establishing machine learning models to predict curative resection in early

- gastric cancer with undifferentiated histology: development and usability study. J Med Internet Res 2021;23:e25053.
48. Yasuda T, Hiroyasu T, Hiwa S, et al. Potential of automatic diagnosis system with linked color imaging for diagnosis of *Helicobacter pylori* infection. Dig Endosc 2020;32:373-381.
49. Zheng W, Zhang X, Kim JJ, et al. High accuracy of convolutional neural network for evaluation of *Helicobacter pylori* infection based on endoscopic images: preliminary experience. Clin Transl Gastroenterol 2019;10:e00109.
50. Shichijo S, Endo Y, Aoyama K, et al. Application of convolutional neural networks for evaluating *Helicobacter pylori* infection status on the basis of endoscopic images. Scand J Gastroenterol 2019;54:158-163.
51. Nakashima H, Kawahira H, Kawachi H, Sakaki N. Artificial intelligence diagnosis of *Helicobacter pylori* infection using blue laser imaging-bright and linked color imaging: a single-center prospective study. Ann Gastroenterol 2018;31:462-468.
52. Itoh T, Kawahira H, Nakashima H, Yata N. Deep learning analyzes *Helicobacter pylori* infection by upper gastrointestinal endoscopy images. Endosc Int Open 2018;6:E139-E144.
53. Shichijo S, Nomura S, Aoyama K, et al. Application of convolutional neural networks in the diagnosis of *Helicobacter pylori* infection based on endoscopic images. EBioMedicine 2017;25:106-111.
54. Huang CR, Chung PC, Sheu BS, Kuo HJ, Popper M. *Helicobacter pylori*-related gastric histology classification using support-vector-machine-based feature selection. IEEE Trans Inf Technol Biomed 2008;12:523-531.
55. Huang CR, Sheu BS, Chung PC, Yang HB. Computerized diagnosis of *Helicobacter pylori* infection and associated gastric inflammation from endoscopic images by refined feature selection using a neural network. Endoscopy 2004;36:601-608.