FRONTIER

2921  Fecal microbiota transplantation for irritable bowel syndrome: An intervention for the 21st century
   El-Salhy M, Patcharatrakul T, Gonlachanvit S

REVIEW

2944  Interplay between nuclear factor erythroid 2-related factor 2 and inflammatory mediators in COVID-19-related liver injury

2963  Mucosal lesions of the upper gastrointestinal tract in patients with ulcerative colitis: A review
   Sun Y, Zhang Z, Zheng CQ, Sang LX

2979  Application of artificial intelligence-driven endoscopic screening and diagnosis of gastric cancer

2994  Hepatocellular carcinoma in viral and autoimmune liver diseases: Role of CD4+ CD25+ Foxp3+ regulatory T cells in the immune microenvironment

MINIREVIEWS

3010  Role of bile acids in liver diseases mediated by the gut microbiome
   Shao JW, Ge TT, Chen SZ, Wang G, Yang Q, Huang CH, Xu LC, Chen Z

3022  Liver injury in COVID-19: Detection, pathogenesis, and treatment
   Cai Y, Ye LP, Song YQ, Mao XL, Wang L, Jiang YZ, Que WT, Li SW

3037  Role of imaging in evaluating the response after neoadjuvant treatment for pancreatic ductal adenocarcinoma
   Zhang Y, Huang ZX, Song B

3050  Current approach to treatment of minimal hepatic encephalopathy in patients with liver cirrhosis
   Moran S, López-Sánchez M, Milke-Garcia MDP, Rodríguez-Leal G

3064  COVID-19 and pediatric fatty liver disease: Is there interplay?
**ORIGINAL ARTICLE**

**Basic Study**

3073 Enhancer of zeste homolog 2 contributes to apoptosis by inactivating janus kinase 2/signal transducer and activator of transcription signaling in inflammatory bowel disease


**Retrospective Study**

3085 Quinone oxidoreductase 1 is overexpressed in gastric cancer and associated with outcome of adjuvant chemotherapy and survival

Jiang ZN, Ahmed SMU, Wang QC, Shi HF, Tang XW

3097 Idiopathic mesenteric phlebosclerosis associated with long-term oral intake of geniposide

Wen Y, Chen YW, Meng AH, Zhao M, Fang SH, Ma YQ

3109 Early serum albumin changes in patients with ulcerative colitis treated with tacrolimus will predict clinical outcome

Ishida N, Miyazu T, Tamura S, Tani S, Yamade M, Iwaizumi M, Hamaya Y, Osawa S, Furuta T, Sugimoto K

**Observational Study**

3121 Preservation of superior rectal artery in laparoscopically assisted subtotal colectomy with ileorectal anastomosis for slow transit constipation

Wu CW, Pu TW, Kang JC, Hsiao CW, Chen CY, Hu JM, Lin KH, Lin TC

**Prospective Study**

3130 High fecal calprotectin levels are associated with SARS-CoV-2 intestinal shedding in COVID-19 patients: A proof-of-concept study


**CORRECTION**

3138 Helicobacter pylori promotes invasion and metastasis of gastric cancer by enhancing heparanase expression

Liu LP, Sheng XP, Shuai TK, Zhao YX, Li B, Li YM
ABOUT COVER
Editorial Board Member of World Journal of Gastroenterology, Francisco Rodriguez-Frias, MS, PhD, Senior Investigator, Associate Professor at Sciences and Medicine Schools, Autonoma University of Barcelona; Head of Special Units (Biochemistry Department) and Head of Liver Pathology Laboratory (Biochemistry and Microbiology Departments) of Vall d’Hebron University Hospital; Head of Biochemistry Research Group or Vall d’Hebron Institute of Research, Passeig Vall d’Hebron, 119-129, Barcelona 08035, Spain. frarodri@vhebron.net

AIMS AND SCOPE
The primary aim of World Journal of Gastroenterology (WJG, World J Gastroenterol) is to provide scholars and readers from various fields of gastroenterology and hepatology with a platform to publish high-quality basic and clinical research articles and communicate their research findings online. WJG mainly publishes articles reporting research results and findings obtained in the field of gastroenterology and hepatology and covering a wide range of topics including gastroenterology, hepatology, gastrointestinal endoscopy, gastrointestinal surgery, gastrointestinal oncology, and pediatric gastroenterology.

INDEXING/ABSTRACTING
The WJG is now indexed in Current Contents®/Clinical Medicine, Science Citation Index Expanded (also known as SciSearch®), Journal Citation Reports®, Index Medicus, MEDLINE, PubMed, PubMed Central, and Scopus. The 2020 edition of Journal Citation Report® cites the 2019 impact factor (IF) for WJG as 3.665; IF without journal self cites: 3.534; 5-year IF: 4.048; Ranking: 35 among 88 journals in gastroenterology and hepatology; and Quartile category: Q2. The WJG’s CiteScore for 2019 is 7.1 and Scopus CiteScore rank 2019: Gastroenterology is 17/137.

RESPONSIBLE EDITORS FOR THIS ISSUE
Production Editor: Ji-Hong Liu; Production Department Director: Yun-Xiaojian Wu; Editorial Office Director: Ze-Mao Gong.
Application of artificial intelligence-driven endoscopic screening and diagnosis of gastric cancer

Yu-Jer Hsiao, Yuan-Chih Wen, Wei-Yi Lai, Yi-Ying Lin, Yi-Ping Yang, Yueh Chien, Aliaksandr A Yarmishyn, De-Kuang Hwang, Tai-Chi Lin, Yun-Chia Chang, Ting-Yi Lin, Kao-Jung Chang, Shih-Hwa Chiou, Ying-Chun Jheng

ORCID number: Yu-Jer Hsiao 0000-0002-0783-605X; Yuan-Chih Wen 0000-0001-5364-9266; Wei-Yi Lai 0000-0003-2689-747X; Yi-Ying Lin 0000-0002-0954-4894; Yi-Ping Yang 0000-0002-7005-0631; Yueh Chien 0000-0002-1528-9252; Aliaksandr A Yarmishyn 0000-0002-0097-9541; De-Kuang Hwang 0000-0001-6346-8485; Tai-Chi Lin 0000-0003-1163-0457; Yun-Chia Chang 0000-0001-7964-9993; Ting-Yi Lin 0000-0001-7906-1612; Kao-Jung Chang 0000-0002-2238-7357; Shih-Hwa Chiou 0000-0002-8651-5755; Ying-Chun Jheng 0000-0002-2883-4432.

Author contributions: Chiou SH and Jheng YC designed the conception; Hsiao YJ, Wen YC, Yarmishyn AA and Jheng YC wrote the paper; Lai WY, Lin YY, Yang YP, Chien Y, Hwang DK, Lin TC, Chang YC, Lin TY and Chang KJ collected the data.

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

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)
Abstract

The landscape of gastrointestinal endoscopy continues to evolve as new technologies and techniques become available. The advent of image-enhanced and magnifying endoscopies has highlighted the step toward perfecting endoscopic screening and diagnosis of gastric lesions. Simultaneously, with the development of convolutional neural network, artificial intelligence (AI) has made unprecedented breakthroughs in medical imaging, including the ongoing trials of computer-aided detection of colorectal polyps and gastrointestinal bleeding. In the past demi-decade, applications of AI systems in gastric cancer have also emerged. With AI’s efficient computational power and learning capacities, endoscopists can improve their diagnostic accuracies and avoid the missing or mischaracterization of gastric neoplastic changes. So far, several AI systems that incorporated both traditional and novel endoscopy technologies have been developed for various purposes, with most systems achieving an accuracy of more than 80%. However, their feasibility, effectiveness, and safety in clinical practice remain to be seen as there have been no clinical trials yet. Nonetheless, AI-assisted endoscopies shed light on more accurate and sensitive ways for early detection, treatment guidance and prognosis prediction of gastric lesions. This review summarizes the current status of various AI applications in gastric cancer and pinpoints directions for future research and clinical practice implementation from a clinical perspective.

Key Words: Artificial intelligence; Diagnostic; Therapeutic; Endoscopy; Gastric cancer; Gastritis

INTRODUCTION

Diagnostic and therapeutic endoscopies play a major role in the management of gastric cancer (GC). Endoscopy is the mainstay for the diagnosis and treatment of early adenocarcinoma and lesions and the palliation of advanced cancer[1-5]. GC, being the fifth most common cancer and the third leading cause of cancer-related deaths worldwide, affects more than one million people and causes approximately 780000 deaths annually[6-10]. Continued development in endoscopy aims to strengthen its quality indicators. These developments include using higher resolution and magnification endoscopies, chromoendoscopy and optical techniques based on the modulation of the light source, such as narrow-band imaging (NBI), fluorescence endoscopy and elastic scattering spectroscopy[11-13]. New tissue sampling methods to identify the stages of a patient’s risk for cancer are also being developed to decrease the burden on patients and clinicians during endoscopy.
Statistically, the relative 5-year survival rate of GC is less than 40%[7,9,14,15], often attributed to the late onset of symptoms and delayed diagnosis[10]. Although early diagnosis is difficult as most patients are asymptomatic in the early stage, the diagnosis point largely determines the patient’s prognosis[16,17]. In other words, endoscopic detection of GC at an earlier stage is the only and most effective way to reduce its recurrence and to prolong patient survival. This early diagnosis of GC provides the opportunity for minimally invasive therapy methods such as endoscopic mucosal resection or submucosal dissection[18-20]. The 5-year survival rate was reported to be more than 90% among patients with GC detected at an early stage[21-23]. Yet, the false negative rate of GC detected by esophagogastroduodenoscopy, the current standard diagnostic procedure, was reported to be between 4.6% and 25.8% [24-29]. In terms of the common diagnostic methods, esophagogastroduodenoscopy is the preferred diagnostic modality for patients with suspected GC; the combination of lymph node dissection, endoscopic ultrasonography and computed tomographic scanning is involved in staging the tumor[30,31]. From the differential diagnosis between GC and gastritis, prediction of the horizontal extent of GC to characterizing the depth of invasion of GC, the early abnormal symptoms of GC and its advanced aggressive malignancy as well as the heavy workload of image analysis present ample inevitable challenges for endoscopists[32-34]. With large variations in the diagnostic ability of endoscopists, long-term training and experience may not guarantee their consistency and accuracy of diagnosis[35-37].

In recent years, artificial intelligence (AI) has caught considerable attention in various medical fields, including skin cancer classification[38-41], diagnosis in radiation oncology[42-45] and analysis of brain magnetic resonance imaging[46-49]. Although its applications have shown impressive accuracy and sensitivity identifying and characterizing imaging abnormalities, its improved sensitivity also meant the detection of subtle and indeterminately significant changes[50,51]. For example, in the analysis of brain magnetic resonance imaging, despite the promise of early diagnosis with machine learning, the relationship between subtle parenchymal brain alterations detected by AI and its neurological outcomes is unknown in the absence of a well-defined abnormality[52]. In other words, the use of AI in diagnostic imaging in various medical fields is continuously undergoing extensive evaluation.

In the field of gastroenterology, AI applications in capsule endoscopy[53-56] and in the detection, localization and segmentation of colonic polyps have been reported as well[57-59]. In particular, in the late 2010s, there was an explosion of interest in GC. The use of AI has proven to provide better diagnostic capabilities, although further validation and extensions are necessary to augment their quality and interpretability. An AI system’s quality is often described with statistical measures of sensitivity, specificity, positive predictive value and accuracy.

Among the different AI models, the convolutional neural network (CNN) is a method most commonly used in medical imaging[60,61] as it allows the detection, segmentation and classification of image patterns[62] (Figure 1). CNN uses the mathematical operation of convolution to classify the images after recognizing patterns from the raw image pixel. Because the 7-layer Le-Net-5 program was first pioneered by LeCun et al in 1998, CNN architectures have been rapidly developing. Today, other widely-used CNNs include AlexNet (2012) with about 15.3% error rate, 22-layer GoogLeNet (2014) with a 6.67% error rate but only 4 million parameters, 19-layer visual geometry group (VGG) Net (2014) with 7.3% error rate and 138 million parameters, and Microsoft’s ResNet (2015) with an error rate of 3.6% that can be trained with as many as 152 layers[63-65]. While scholars have lauded AI for the potential and performance it has displayed, some have cast doubts on its generalizability and role in the holistic assessment of gastric abnormalities.

In the beginning of an AI-assisted diagnostic imaging revolution, we have to anticipate and meticulously assess the potential perils, in the context of its capabilities, to ensure effective and safe incorporation into clinical practice[66]. In this paper, we thereby review the current status of AI applications in screening and diagnosing GC. We explore with emphasis on two broad categories: namely, the identification of pathogenic infection and the qualitative diagnosis of GC. Finally, we considered some directions for further research and the future of its introduction into clinical practice.

### IDENTIFICATION OF HELICOBACTER PYLORI INFECTION

AI applications in identifying pathogenic infections have been widely explored[67] (Table 1). Gastric epithelium *Helicobacter pylori* (*H. pylori*) infection is associated with
Table 1 Summary of artificial intelligence applications in predicting *Helicobacter pylori* infection

<table>
<thead>
<tr>
<th>Ref.</th>
<th>Endoscopic modality</th>
<th>Training dataset</th>
<th>Validation dataset</th>
<th>Accuracy</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Huang et al[78], 2004</td>
<td>WLI</td>
<td>30 patients</td>
<td>74 patients</td>
<td>85.1 (avg)(^1)</td>
<td>78.8 (avg)</td>
<td>90.2 (avg)</td>
<td>-</td>
</tr>
<tr>
<td>Shichijo et al[79], 2017</td>
<td>WLI</td>
<td>32208 images, 1768 patients</td>
<td>11481 images, 397 patients</td>
<td>87.7</td>
<td>88.9</td>
<td>87.4</td>
<td>-</td>
</tr>
<tr>
<td>Itoh et al[81], 2018</td>
<td>WLI</td>
<td>149 images, 139 patients</td>
<td>30 images, 30 patients</td>
<td>-</td>
<td>86.7</td>
<td>86.7</td>
<td>-</td>
</tr>
<tr>
<td>Nakashima et al[84], 2018</td>
<td>WLI, BLI and LCI</td>
<td>162 patients</td>
<td>60 patients</td>
<td>-</td>
<td>96.7</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Shichijo et al[80], 2019</td>
<td>WLI</td>
<td>98564 images, 4494 patients</td>
<td>23699 images, 847 patients</td>
<td>Infected: 66.0; post-eradication: 86.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Zheng et al[82], 2019</td>
<td>WLI</td>
<td>11729 images, 1307 patients</td>
<td>3755 images, 452 patients</td>
<td>84.5</td>
<td>81.4</td>
<td>90.1</td>
<td>-</td>
</tr>
<tr>
<td>Zhu et al[100], 2019</td>
<td>WLI</td>
<td>790 images</td>
<td>203 images</td>
<td>89.2</td>
<td>76.5</td>
<td>95.6</td>
<td>89.7</td>
</tr>
<tr>
<td>Nakashima et al[85], 2020</td>
<td>WLI, BLI and LCI</td>
<td>12887 images, 395 patients</td>
<td>120 patients</td>
<td>80.0 (avg)(^2)</td>
<td>61.3 (avg)</td>
<td>89.4 (avg)</td>
<td>74.7 (avg)</td>
</tr>
</tbody>
</table>

1 Histological characteristics were assessed for the various antrum, body and cardia locations.

2 White light imaging and linked color imaging-based images were both analyzed, with the linked color imaging obtaining significantly higher accuracy, sensitivity, specificity and positive predictive value. BLI: Blue laser imaging; LCI: Linked color imaging; PPV: Positive predictive value; WLI: White light imaging.

Figure 1 The convolutional neural network model. A convolutional neural network consists of an input layer, a few hidden layers and an output layer. It is commonly applied in medical imaging through the detection, segmentation and classification of image patterns.

functional dyspepsia, peptic ulcers, mucosal atrophy, intestinal metaplasia, atrophic gastritis and GC\[68,69\]. Because most gastric malignancies correlate with *H. pylori* infection, identifying *H. pylori* infection at its early stage is essential in preventing *H. pylori*-aggravated comorbidities\[70-74\]. Although physicians usually use the C13 urea breath test to diagnose *H. pylori*, most subclinical *H. pylori* infection cases still rely on the time-consuming and invasive biopsy examination to avoid the risk of a false negative diagnosis. Moreover, the Kyoto Classification, as the current gold standard of *H. pylori* severity classification, requires examiners to measure lesions by their bare eyes. Such a method is a subjective judgment that usually comes with interoperator bias\[75-77\]. Compelled by such ambiguity, researchers have turned to devising a next-generation semi-automatic standard examination protocol, that is AI.
As early as 2004, before CNN took the lead in machine-assisted image diagnosis, Huang et al.[78] deployed the refined feature selection neural network to process endoscopic images and return the results of *H. pylori* infection probability and severity. By training AI with 30 patients' endoscopic images including crops of antrum, body and cardia locations of the stomach, they established an algorithm that achieved an average of 78.8% sensitivity, 90.2% specificity and 85.1% accuracy in an independent cohort of 74 patient images. The overall prediction accuracy was better than the one demonstrated by young physicians and fellow doctors, who scored 68.4% and 78.4%, respectively. It was the first model demonstrating the potential of computer-aided diagnosis of *H. pylori* infection by endoscope images.

However, since the introduction of the 7-layer Le-Net-5 program by LeCun et al in 1998, the CNNs have gradually taken over in the field of medical image processing. To name a few examples, Shichijo et al.[79] used 32208 images of 735 *H. pylori*-positive and 1015 *H. pylori*-negative cases to develop an *H. pylori* identifying AI system based on the architecture of 22-layer GoogleLeNet. The sensitivity, specificity, accuracy and time consumption were 81.9%, 83.4%, 83.1% and 198 s for the first CNN and 88.9%, 87.4%, 87.7% and 194 s for the secondary CNN developed, respectively, compared with that of 79.0%, 83.2%, 82.4% and 230 min by the endoscopists. Later, still using GoogleLeNet, Shichijo et al.[80] developed another system that further classified the current infection, post-eradication and current noninfection statuses of *H. pylori*, obtaining an accuracy of 48%, 84%, and 80%, respectively. In this system, the CNN was trained with 98564 images from 4494 patients and tested with 25699 images from 847 independent cases. Itoh et al[81] also developed a CNN based on GoogleLeNet, trained with 149 endoscopic images obtained from 139 patients and tested with 30 images from 30 patients, which could detect and diagnose *H. pylori* with sensitivity and specificity of 86.7%. Additionally, the use of ResNet CNN architecture was reported by Zheng et al[82] in 2019, achieving a sensitivity, specificity and accuracy of 81.4%, 90.1% and 84.5%, respectively. In this study, the system was trained with 11729 images from 1507 patients and tested with 3755 images from 452 patients using a 50-layer ResNet-50 (Microsoft) CNN system and PyTorch (Facebook) deep learning framework.

Recently, AI has also been applied to linked color imaging (LCI) and blue laser imaging, two novel image-enhanced endoscopy technologies[83]. It helped diagnose[84] and classify[85] the *H. pylori* infection and has shown greater effectiveness. In 2018, Nakashima et al[84] developed a system on a training set of 162 patients and a test set of 60 patients that could diagnose *H. pylori* infection with an area under the curve of 0.96 and 0.95 and sensitivity of 96.7% and 96.7% for blue laser imaging-bright and LCI, respectively. Such performance is superior to systems that use conventional white light imaging (WLI) (with 0.66 area under the curve and sensitivities as mentioned earlier in other studies) as well as that of experienced endoscopists. Another 2020 study by the same team also showed that classifying the *H. pylori* infection status (uninfected, infected and post-eradication) by incorporating deep learning and image-enhanced endoscopies yields more accurate results. The system was trained with 6639 WLI and 6248 LCI images from 395 patients and tested with images from 120 patients[85].

However, there are some limitations of AI in identifying *H. pylori* that remain to be overcome amongst the developed systems and findings. First, the histological time frame, especially for the eradicated infection, was not considered in the AI systems[86]. Second, both the training data sets and test data sets were obtained from a single center for all existing systems. A continued and even more rigorous external validation, which uses more diverse sources of images and endoscopies, is necessary to evaluate each system’s generalizability[87,88]. Additionally, the application of CNN algorithms is also still confined to the existing models of CNN algorithms (mostly GoogleLeNet and a few ResNet). Further technical refinements may overcome current limitations faced by endoscopists. They also shed light on the possibility of a system that distinguishes between *H. pylori*-infected and *H. pylori*-eradicated patients, determines different parts of the stomach (cardia, body, angle and pylorus) and provides real-time evaluations of *H. pylori*. These will be considerations vital for its implementation in clinical practice in the future.

**DETECTION OF GC**

Besides *H. pylori* infection, computer-aided pattern recognition with endoscopic images has also been applied to diagnose wall invasion depth (Table 2). An accurate diagnosis of invasion depth and subsequent staging is the basis for determining the
Table 2 Summary of artificial intelligence applications in prediction of invasion depth and differentiation of cancerous areas from noncancerous areas

<table>
<thead>
<tr>
<th>Ref.</th>
<th>Application</th>
<th>Endoscopic modality</th>
<th>Training dataset</th>
<th>Validation dataset</th>
<th>Accuracy</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kubota et al</td>
<td>Prediction of invasion depth</td>
<td>WLI</td>
<td>344 patients, 902 images</td>
<td>-</td>
<td>77.2 (T1)</td>
<td>-</td>
<td>-</td>
<td>80.1</td>
<td>85.1</td>
</tr>
<tr>
<td>Miyaki et al</td>
<td>Differentiation of cancerous areas from noncancerous areas</td>
<td>WLI and magnified FICE</td>
<td>495 images</td>
<td>46 images</td>
<td>85.9</td>
<td>84.8</td>
<td>87.0</td>
<td>86.7</td>
<td>85.1</td>
</tr>
<tr>
<td>Hirasawa et al</td>
<td>Differentiation of cancerous areas from noncancerous areas</td>
<td>WLI</td>
<td>13584 images</td>
<td>2296 images, 69 patients</td>
<td>92.2</td>
<td>92.2</td>
<td>-</td>
<td>30.6</td>
<td>-</td>
</tr>
<tr>
<td>Kanesaka et al</td>
<td>Detection of EGC</td>
<td>Magnified NBI</td>
<td>126 images</td>
<td>81 images</td>
<td>96.3</td>
<td>96.7</td>
<td>95.0</td>
<td>98.3</td>
<td>-</td>
</tr>
<tr>
<td>Horiuchi et al</td>
<td>Differentiation of EGC from gastritis</td>
<td>Magnified NBI</td>
<td>2570 images</td>
<td>258 images</td>
<td>85.3</td>
<td>95.4</td>
<td>71.0</td>
<td>82.3</td>
<td>91.7</td>
</tr>
<tr>
<td>Yoon et al</td>
<td>Detection of EGC and prediction of EGC invasion depth</td>
<td>WLI</td>
<td>11686 images, 800 patients</td>
<td>-</td>
<td>79.2</td>
<td>77.8</td>
<td>79.3</td>
<td>77.7</td>
<td>-</td>
</tr>
<tr>
<td>Horiuchi et al</td>
<td>Detection of EGC</td>
<td>Magnified NBI</td>
<td>2570 images</td>
<td>174 videos, 82 patients</td>
<td>85.1</td>
<td>87.4</td>
<td>82.8</td>
<td>83.5</td>
<td>86.7</td>
</tr>
<tr>
<td>Li et al</td>
<td>Differentiation of EGC from noncancerous lesions</td>
<td>Magnified NBI</td>
<td>2088 images</td>
<td>342 images</td>
<td>90.9</td>
<td>91.2</td>
<td>90.6</td>
<td>90.6</td>
<td>91.2</td>
</tr>
<tr>
<td>Nagao et al</td>
<td>Prediction of invasion depth</td>
<td>WLI, nonmagnifying NBI and indigo-carmine dye contrast imaging (Indigo)</td>
<td>16457 images, 1084 patients</td>
<td>-</td>
<td>94.4</td>
<td>89.2</td>
<td>98.7</td>
<td>98.3</td>
<td>91.7</td>
</tr>
<tr>
<td>Namikawa et al</td>
<td>Differentiation of cancerous areas from noncancerous areas</td>
<td>WLI, nonmagnifying NBI and indigo-carmine dye contrast imaging (Indigo)</td>
<td>18410 images</td>
<td>1459 images</td>
<td>95.9</td>
<td>99.0</td>
<td>93.3</td>
<td>92.5</td>
<td>-</td>
</tr>
</tbody>
</table>

EGC: Early gastric cancer; FICE: Flexible spectral imaging color enhancement; NBI: Narrow-band imaging; NPV: Negative predictive value; PPV: Positive predictive value; WLI: White light imaging.

Appropriate treatment modality, especially for suspected early GCs (EGC)\[^{99-101}\]. Classified based on the 7\(^{th}\) TNM classification of tumors\[^{92,93}\], EGC is categorized as tumor invasion of the mucosa (T1a) or invasion of the submucosa (T1b) stages. While endoscopic ultrasonography is useful for T-staging of GC by delineating each gastric wall layer\[^{94,95}\], conventional endoscopy is still arguably superior to endoscopic ultrasonography for T-staging of EGC\[^{96,97}\]. However, there remains room for improvement, such as by utilizing AI, to increase its accuracy. In 2012, Kubota et al\[^{98}\] first explored the system with a relatively high sensitivity of 68.9% and 63.6% in T1a and T1b GCs, achieving high accuracy, especially in early tumors. The accuracy for T1 tumors was 77.2% compared to that of 49.1%, 51.0%, and 55.3% for T2, T3 and T4 tumors, respectively. Another system developed by Hirasawa et al\[^{99}\] achieved a high sensitivity of 92.2% of CNN, though at the expense of a low positive predictive value (30.6%). Zhu et al\[^{100}\] later demonstrated a CNN-computer assisted diagnosis system that achieved much higher accuracy (by 17.25%; 95% confidence interval: 11.63-22.59) and specificity (by 32.21%; 95% confidence interval: 26.78-37.44) compared to human endoscopists. These preliminary findings showed that AI is a potentially helpful diagnostic procedure in EGC detection and pointed towards developing an AI system that can differentiate between malignant and benign lesions.

Given that the difference in EGC depth in endoscopic images is subtler and more difficult to discern, Yoon et al\[^{101}\] identified that more sophisticated image classification methods but not merely conventional CNN models are required. The team developed a system that classifies endoscopic images into EGC (T1a or T1b) or non-EGC. This system used the combination of the CNN-based visual geometry group-16 network pretrained on ImageNet and a novel model of the weighted sum of gradient-weighted class activation mapping. This system focused on learning the visual features of EGC regions rather than those of other gastric textures, achieving both high accuracy of 91.0% and high area under the curve of 0.981. In another study in 2020, Nagao et al\[^{102}\] used the state-of-the-art ResNet50 CNN architecture to...
develop a system via training the images from different angles and distances. This system predicted the invasion depth of GC with an image-based accuracy as high as 94.5%.

However, when using these AI systems for invasion depth diagnosis, distinguishing superficially depressed and differentiated-type intramucosal cancers from gastritis remains a challenge. The diagnostic accuracy of determining invasion depth is largely affected by its histological characteristics. For instance, the system developed by Yoon et al.[101] achieved an accuracy of 77.1% for differentiated-type tumors in contrast to that of 65.5% for undifferentiated type. Horiuchi et al.[105] made a substantial effort and developed another system that could differentiate EGC from gastritis using magnifying endoscopy with NBI (M-NBI). The system achieved an accuracy of 85.3% and sensitivity and negative predictive value of 95.4% and 91.7%. Another approach was made by Namikawa et al.[104], who developed a system that was trained by gastritis images and tested to classify GC and gastric ulcers. A continued development of AI systems that consider the differentiated type histology will shed light on the future of AI-assisted differentiation of T1a from T1b GC and that of T1a and T1b cancers from the later stages of GC.

To bring AI-assisted systems one step closer to real-time clinical applications, video-based systems have also been explored. In 2020, Horiuchi et al.[105] used the video-based systems and achieved a comparable accuracy of 85.1% in distinguishing EGC and noncancerous lesions. Based on the CNN-CAD system, their system was trained with 2570 images (1492 cancerous and 1078 noncancerous images) and tested with 174 videos. This preliminary success in the video-based CNN-CAD system pointed out the potential of real-time AI-assisted diagnosis, which could be a promising technique for detecting EGC in clinical practice. Early detection of GC means an early treatment of endoscopic dissection in accordance with the works promoted by the Japanese Gastric Cancer Association since 2014[106].

**DISCUSSION**

Over the past decade, AI has displayed its potential diagnosing GC to amplify human endoscopist capacities. Although the diagnosis of GC requires a holistic set of assessments, AI is applicable and helpful in some parts. A system that detects GC with high sensitivity regardless of its accuracy in determining invasion depth could provide great clinical assistance for physicians to decide if biopsy and endoscopic submucosal dissection are necessary. In the near future, there should be some other diagnostic procedures that can be explored with AI. For example, macroscopic characteristics, namely the “nonextension signs” commonly used to distinguish between SM1 and SM2 invasion depths of GCs[107] have yet to be explored with AI.

Clinically, there are also some distinct markers that endoscopists use to evaluate gastric surface and color changes. Distinguishing the markers such as changes in light reflection and spontaneous bleeding are clinical skills[108,109] that AI could potentially learn and interpret. In clinical practice, antiperistaltic agents are suggested for polyethersulfone preparation, and indigo carmine chromoendoscopy could help diagnose elevated superficial lesions with an irregular surface pattern[110] with which their efficacy could be evaluated by real-time AI endoscopy in the near future.

Although several studies have attempted to apply AI in different types of endoscopies, ranging from WLI to LCI to blue laser imaging, these studies can also continue to extend AI to NBI and other nonconventional endoscopies. For instance, endoscopy with NBI has shown higher diagnostic accuracy compared to M-NBI [78.8% (76.4%-83.0%) vs 72.2% (69.3%-73.6%), \(P < 0.0001\)][111]. An AI system that is trained with WLI images and tested with NBI images instead will also have clinical significance[112,113]. Proposed in 2016 was the Magnifying Endoscopy Simple Diagnostic Algorithms for EGC that suggested a systematic approach to WLI magnifying endoscopy. It is recommended that if a suspicious lesion is detected, M-NBI should be performed to distinguish if the lesion is cancer or noncancer[114]. According to this algorithm, changing from WLI to M-NBI endoscopy is therefore critical for diagnosis, and the future development of AI systems can consider accounting for such changes.

In the AI systems developed over the past decade, we summarize the following common limitations faced. First, there seems to be a common lack of high-quality datasets for machine learning development, a problem faced in clinical practice even without AI[115]. Simultaneously, some studies reported that low-quality images result in higher chances of misdiagnosis by the AI system[116], and most studies excluded...
large numbers of poor-quality images\cite{99-101}. The call for cross-validation with multicenter observational studies has also been discussed in several studies in hopes of picking out any potential overfitting and spectrum bias that is foreseeable in deep image classification models\cite{117-119}. Some authors have argued that the AI system they have developed is institution-specific and that the validation with the dataset from external sources is necessary\cite{103,105}. In this regard, multicenter studies have been widely used in other medical fields to evaluate deep learning systems\cite{120-122}, though there have been no such studies in the field of GC.

Another challenge that remains is seen in the imbalanced class distributions, a common classification problem in which the distribution of samples across the known classes is biased or skewed. For example, in the study reported by Hirasawa et al\cite{99} in 2018, there are few samples for the later stages (only 32.5% of samples were T2-T4 cancers) than for early cancer (67.5% of samples). Such imbalanced classifications pose a challenge as machine learning models are primarily designed on the assumption of an equal number of samples for each class\cite{123}. Without sufficient samples for certain classes of the training dataset, their existence might be misperceived as other classes as the AI model becomes more sensitive to classification errors. It may result in poor predictive performance, especially for the minority class and subsequently an overall increased misdiagnosis rate. For example, in the cases of the AI model for GC staging, a misdiagnosis of late-stage cancer for gastritis or nonmalignancy has dangerous implications\cite{124-126} if the AI system was used for its diagnosis alone. However, in most cases, advanced-stage GC might have already metastasized to other parts of the body\cite{127-129}, and its diagnosis based only on the AI system alone is unlikely. Nonetheless, in the development of AI systems for such medical applications, these technical problems of imbalanced classification should not be overlooked. It has been discussed by other reviews how modifications can be made to AI models to recognize targets, no matter how frequent or rare they are, to minimize the possibility of misdiagnosis\cite{130,131}.

Overall, the potential for AI applications in GC is extensive yet highly specific. In an upcoming era of AI-assisted diagnosis, by combining image information, medical history and laboratory data, endoscopists can look forward to the continued development of new systems for varying purposes (Figure 2). AI systems are specific and unlikely to be generalized\cite{132,133}, and it is fallacious to compare a single statistical performance measure across different AI systems. The efficacy of an AI system depends on the intended role it plays in clinical practice. For example, an AI system with a high positive predictive value is desirable in determining which multicancer to send for biopsy, while high sensitivity suffices for a system that helps differentiate cancerous from noncancerous clinical signs, especially for amateur endoscopists. In the foreseeable future, AI can be incorporated in the differential diagnosis of the malignancy and stages of gastric lesions, using various endoscopic technologies and techniques.

**CONCLUSION**

Overall, the application of AI in gastroenterology is in its infancy. At present, there exist several retrospective models applied in both images and videos and using both WLI and NBI endoscopies that have proven to have better performance for the same tasks carried out by experienced endoscopists. However, there have not been any attempts of clinical trials. In contrast to the ongoing trials for detecting colorectal polyps\cite{134-136}, AI applications in GC and its corresponding diagnostic methods are still preliminary. The limitations of existing efforts point towards the importance of continued research in the field that can go a long way in making quicker, more accurate and precise evaluations of GC risk. While we witnessed its rapid and steep growth in the past decade, future studies are needed to streamline the machine learning process and define its role in the computer-aided diagnosis of *H. pylori* infections and GC in real-life clinical scenarios.
In an upcoming era of artificial intelligence-assisted diagnosis, endoscopists can look forward to the continued development of new artificial intelligence systems for varying purposes. From determining multicancer via biopsy to real-time endoscopies, artificial intelligence has the potential of assisting physicians to improve their diagnostic accuracies.

REFERENCES


find colorectal polyps in colonoscopy: A systematic literature review.


Hsiao YJ et al. AI in gastric cancer


Harari YN. 21 Lessons for the 21st century. London: Jonathan Cape, 2018


