

# World Journal of *Gastroenterology*

*World J Gastroenterol* 2019 April 14; 25(14): 1640-1782



**REVIEW**

- 1640** Precision surgical approach with lymph-node dissection in early gastric cancer  
*Kinami S, Nakamura N, Tomita Y, Miyata T, Fujita H, Ueda N, Kosaka T*

**MINIREVIEWS**

- 1653** Update on hepatocellular carcinoma: Pathologists' review  
*El Jabbour T, Lagana SM, Lee H*
- 1666** Application of artificial intelligence in gastroenterology  
*Yang YJ, Bang CS*

**ORIGINAL ARTICLE****Basic Study**

- 1684** Clinical significance of programmed death 1/programmed death ligand 1 pathway in gastric neuroendocrine carcinomas  
*Yang MW, Fu XL, Jiang YS, Chen XJ, Tao LY, Yang JY, Huo YM, Liu W, Zhang JF, Liu PF, Liu Q, Hua R, Zhang ZG, Sun YW, Liu DJ*
- 1697** Functional role of long non-coding RNA CASC19/miR-140-5p/CEMIP axis in colorectal cancer progression *in vitro*  
*Wang XD, Lu J, Lin YS, Gao C, Qi F*
- 1715** Seven-senescence-associated gene signature predicts overall survival for Asian patients with hepatocellular carcinoma  
*Xiang XH, Yang L, Zhang X, Ma XH, Miao RC, Gu JX, Fu YN, Yao Q, Zhang JY, Liu C, Lin T, Qu K*

**Case Control Study**

- 1729** Chronic functional constipation is strongly linked to vitamin D deficiency  
*Panarese A, Pesce F, Porcelli P, Riezzo G, Iacovazzi PA, Leone CM, De Carne M, Rinaldi CM, Shahini E*

**Retrospective Study**

- 1741** Clinical characteristics of non-alcoholic fatty liver disease in Chinese adult hypopituitary patients  
*Yuan XX, Zhu HJ, Pan H, Chen S, Liu ZY, Li Y, Wang LJ, Lu L, Yang HB, Gong FY*

**Observational Study**

- 1753** Measurement of prostaglandin metabolites is useful in diagnosis of small bowel ulcerations  
*Matsuno Y, Umeno J, Esaki M, Hirakawa Y, Fuyuno Y, Okamoto Y, Hirano A, Yasukawa S, Hirai F, Matsui T, Hosomi S, Watanabe K, Hosoe N, Ogata H, Hisamatsu T, Yanai S, Kochi S, Kurahara K, Yao T, Torisu T, Kitazono T, Matsumoto T*

**Prospective Study**

- 1764** Endoscopic response to tumor necrosis factor inhibitors predicts long term benefits in Crohn's disease  
*Alfaro I, Masamunt MC, Planell N, López-García A, Castro J, Gallego M, Barastegui R, Giner A, Vara A, Salas A, Ricart E, Panés J, Ordás I*

**CASE REPORT**

- 1775** Application of a 3D-printed "fistula stent" in plugging enteroatmospheric fistula with open abdomen: A case report  
*Xu ZY, Ren HJ, Huang JJ, Li ZA, Ren JA*

## ABOUT COVER

Editorial board member of *World Journal of Gastroenterology*, Jun-Te Hsu, MD, Professor, Surgeon, Surgical Oncologist, Department of Surgery, Chang Gung Memorial Hospital, Taoyuan County and Chang Gung Medicine College, Tao Yuan 333, Taiwan

## AIMS AND SCOPE

*World Journal of Gastroenterology* (*World J Gastroenterol*, *WJG*, print ISSN 1007-9327, online ISSN 2219-2840, DOI: 10.3748) is a peer-reviewed open access journal. The *WJG* Editorial Board consists of 642 experts in gastroenterology and hepatology from 59 countries.

The primary task of *WJG* is to rapidly publish high-quality original articles, reviews, and commentaries in the fields of gastroenterology, hepatology, gastrointestinal endoscopy, gastrointestinal surgery, hepatobiliary surgery, gastrointestinal oncology, gastrointestinal radiation oncology, etc. The *WJG* is dedicated to become an influential and prestigious journal in gastroenterology and hepatology, to promote the development of above disciplines, and to improve the diagnostic and therapeutic skill and expertise of clinicians.

## 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, Scopus and Directory of Open Access Journals. The 2018 edition of Journal Citation Report® cites the 2017 impact factor for *WJG* as 3.300 (5-year impact factor: 3.387), ranking *WJG* as 35<sup>th</sup> among 80 journals in gastroenterology and hepatology (quartile in category Q2).

RESPONSIBLE EDITORS  
FOR THIS ISSUE

Responsible Electronic Editor: Han Song

Proofing Editorial Office Director: Ze-Mao Gong

## NAME OF JOURNAL

*World Journal of Gastroenterology*

## ISSN

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

## LAUNCH DATE

October 1, 1995

## FREQUENCY

Weekly

## EDITORS-IN-CHIEF

Subrata Ghosh, Andrzej S Tarnawski

## EDITORIAL BOARD MEMBERS

<http://www.wjgnet.com/1007-9327/editorialboard.htm>

## EDITORIAL OFFICE

Ze-Mao Gong, Director

## PUBLICATION DATE

April 14, 2019

## COPYRIGHT

© 2019 Baishideng Publishing Group Inc

## INSTRUCTIONS TO AUTHORS

<https://www.wjgnet.com/bpg/gerinfo/204>

## GUIDELINES FOR ETHICS DOCUMENTS

<https://www.wjgnet.com/bpg/GerInfo/287>

## GUIDELINES FOR NON-NATIVE SPEAKERS OF ENGLISH

<https://www.wjgnet.com/bpg/gerinfo/240>

## PUBLICATION MISCONDUCT

<https://www.wjgnet.com/bpg/gerinfo/208>

## ARTICLE PROCESSING CHARGE

<https://www.wjgnet.com/bpg/gerinfo/242>

## STEPS FOR SUBMITTING MANUSCRIPTS

<https://www.wjgnet.com/bpg/GerInfo/239>

## ONLINE SUBMISSION

<https://www.f6publishing.com>



## Application of artificial intelligence in gastroenterology

Young Joo Yang, Chang Seok Bang

**ORCID number:** Young Joo Yang (0000-0001-6325-1104); Chang Seok Bang (0000-0003-4908-5431).

**Author contributions:** Yang YJ collected the data and drafted the manuscript. Bang CS conceptualized, collected the data, drafted the manuscript, performed critical revision and approved the final manuscript.

**Conflict-of-interest statement:** The authors declare no conflicts of interest.

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

**Manuscript source:** Invited manuscript

**Received:** February 9, 2019

**Peer-review started:** February 12, 2019

**First decision:** February 26, 2019

**Revised:** March 4, 2019

**Accepted:** March 16, 2019

**Article in press:** March 16, 2019

**Published online:** April 14, 2019

**P-Reviewer:** Chiu KW, Triantafyllou K

**S-Editor:** Ma RY

**L-Editor:** A

**Young Joo Yang, Chang Seok Bang,** Department of Internal Medicine, Hallym University College of Medicine, Chuncheon, Gangwon-do 24253, South Korea

**Corresponding author:** Chang Seok Bang, MD, PhD, Assistant Professor, Doctor, Department of Internal Medicine, Hallym University College of Medicine, Sakju-ro 77, Chuncheon, Gangwon-do 24253, South Korea. [csbang@hallym.ac.kr](mailto:csbang@hallym.ac.kr)

**Telephone:** +82-33-2405821

**Fax:** +82-33-2418064

### Abstract

Artificial intelligence (AI) using deep-learning (DL) has emerged as a breakthrough computer technology. By the era of big data, the accumulation of an enormous number of digital images and medical records drove the need for the utilization of AI to efficiently deal with these data, which have become fundamental resources for a machine to learn by itself. Among several DL models, the convolutional neural network showed outstanding performance in image analysis. In the field of gastroenterology, physicians handle large amounts of clinical data and various kinds of image devices such as endoscopy and ultrasound. AI has been applied in gastroenterology in terms of diagnosis, prognosis, and image analysis. However, potential inherent selection bias cannot be excluded in the form of retrospective study. Because overfitting and spectrum bias (class imbalance) have the possibility of overestimating the accuracy, external validation using unused datasets for model development, collected in a way that minimizes the spectrum bias, is mandatory. For robust verification, prospective studies with adequate inclusion/exclusion criteria, which represent the target populations, are needed. DL has its own lack of interpretability. Because interpretability is important in that it can provide safety measures, help to detect bias, and create social acceptance, further investigations should be performed.

**Key words:** Artificial intelligence; Convolutional neural network; Deep-learning; Computer-assisted; Gastroenterology; Endoscopy

©The Author(s) 2019. Published by Baishideng Publishing Group Inc. All rights reserved.

**Core tip:** Artificial intelligence (AI) using deep-learning (DL) has emerged as a breakthrough computer technology. The convolutional neural network exhibited outstanding performance in image analysis. AI has been applied in the field of gastroenterology in terms of diagnosis, prognosis, and image analysis. However, potential inherent pitfalls of selection bias, overfitting, and spectrum bias (class imbalance) have the possibility of overestimating the accuracy and generalizing the

E-Editor: Song H



result. Therefore, external validation using unused datasets for model development, collected in a way that minimizes the spectrum bias, is mandatory. DL has its own lack of interpretability, and further investigations should be performed on this issue.

**Citation:** Yang YJ, Bang CS. Application of artificial intelligence in gastroenterology. *World J Gastroenterol* 2019; 25(14): 1666-1683

**URL:** <https://www.wjgnet.com/1007-9327/full/v25/i14/1666.htm>

**DOI:** <https://dx.doi.org/10.3748/wjg.v25.i14.1666>

## INTRODUCTION

Recently, artificial intelligence (AI) using deep-learning (DL) has emerged as a breakthrough computer technology, and numerous research studies, using AI applications to identify or differentiate images in various medical fields including radiology, neurology, orthopedics, pathology, ophthalmology, and gastroenterology, have been published<sup>[1]</sup>. However, AI, the display of intelligent behavior indistinguishable from that of a human being, was already mentioned in the 1950s<sup>[2]</sup>. Although AI has waxed and waned over the past six decades with seemingly little improvement, it was constantly applied to the medical field using various models of machine learning (ML) including Bayesian inferences, decision trees, linear discriminants, support vector machines (SVM), logistic regression, and artificial neural networks (ANNs).

By the era of big data, the accumulation of enormous digital images and medical records drove a need for the utilization of AI to efficiently deal with these data, which also become fundamental resources for the machine to learn by itself. Furthermore, the evolution of computing power with graphic processing units can overcome the limitations of traditional ML, particularly overtraining for input data (overfitting). This led to a revival of AI, especially when using DL technology, a new form of ML. Among several DL methods, the convolutional neural network (CNN), which consists of multilayers of ANN with step-by-step minimal processing, showed outstanding performance in image analysis and has received attention in AI (Figure 1 and Table 1).

In the field of gastroenterology, physicians handle large amounts of clinical data and various kinds of image devices such as esophagogastroduodenoscopy (EGD), colonoscopy, capsule endoscopy (CE), and ultrasound equipment. AI has been applied in the field of gastroenterology when making a diagnosis, predicting a prognosis, and analyzing images. Previous studies reported remarkable results of AI in gastroenterology. The rapid progression of AI demands that gastroenterologists learn the utility, strengths, and pitfalls of AI. In addition, physicians should prepare for the changes and effects of AI on real clinical practice in the near future. Hence, in this review, we aim to (1) briefly introduce an ML technology; (2) summarize an AI application in the field of gastroenterology, which is divided into two categories (statistical analysis for recognition of diagnosis or prediction of prognosis, and analyze images for patient applications excluding animal studies); and (3) discuss the challenges for the application and future directions of AI.

## ML TECHNOLOGY

Generally, AI is considered as a machine intelligence that has cognitive functions similar to those of humans including “learning” and “problem solving<sup>[3]</sup>”. Currently, ML is the most common approach of AI. It automatically builds mathematical algorithms from given data (known as input training data) and predicts or makes decisions in uncertain conditions without human instructions (Figure 1A)<sup>[4]</sup>. In the medical field, ML methods such as Bayesian networks, linear discriminants, SVMs, and ANNs have been used<sup>[5]</sup>. A naïve Bayes classifier that represents the probabilistic relationship between input and output data is a typical classification model<sup>[6]</sup>. The SVM, which was invented by Vladimir N Vapnik and Alexey Ya Chervonenkis in 1963<sup>[7]</sup>, is a discriminative model that uses a dividing hyperplane. Before DL development, SVM showed the best performance for classification and regression, which were achieved by optimizing a hyperplane with the largest functional margin (distance from the hyperplane in a high- or infinite-dimensional space to the nearest training data point of any class)<sup>[8]</sup>.



**Table 1 Artificial intelligence terminology**

Artificial intelligence	Machine intelligence that has cognitive functions similar to those of humans such as “learning” and “problem solving.”
Machine learning	Mathematical algorithms which is automatically built from given data (known as input training data) and predicts or makes decisions in uncertain conditions without being explicitly programmed
Support vector machines	Discriminative classifier formally defined by an optimizing hyperplane with the largest functional margin
Artificial neural networks	Multilayered interconnected network which consists of an input, hidden connection (between the input and output layer), and output layer
Deep learning	Subset of machine learning technique that composed of multiple-layered neural network algorithms
Convolutional neural networks	Specific class of artificial neural networks that consists of (1) convolutional and pooling layers, which are the two main components to extract distinct features; and (2) fully connected layers to make an overall classification
Overfitting	Modelling error which occurs when a certain learning model tailors itself too much on the training dataset and predictions are not well generalized to new datasets
Spectrum bias	Systematic error occurs when the dataset used for model development does not adequately represent or reflect the range of patients who will be applied in clinical practice (target population)

An ANN is a multilayered interconnected network inspired by the neuronal connections of the human brain. Although the ANN was introduced by McCulloch and Walter in 1943<sup>[9]</sup>, it was studied in 1957 by Frank Rosenblatt using the concept of the perceptron<sup>[10]</sup>. The ANN as a hierarchical structure consists of an input, hidden connection (between the input and output layer), and output layer. The connection in the hidden layer has a strength (known as weight) that is used for the learning process of the network (Figure 1B). Through an appropriate training process (learning process), the network can adjust the value of the connection weight to optimize the best result (Figure 1C).

In the 1980s, an ANN with several hidden layers between the input and output layer was introduced. This was known as a DL (or a deep neural network). Although the ANN showed remarkable performance in managing nonlinear datasets regarding diagnosis and prognostic prediction in the medical field, the ANN revealed several weaknesses as well: a vanishing gradient, overfitting, insufficient computing capacity, and lack of training data. These weaknesses hampered the advancement of the ANN. Finally, the recent availability of big data provided sufficient input data for training, and the rapid progression of computing power allowed researchers to overcome prior limitations. Among several AI methods, DL received the attention of the public and has shown excellent performance in the computer vision area using CNNs.

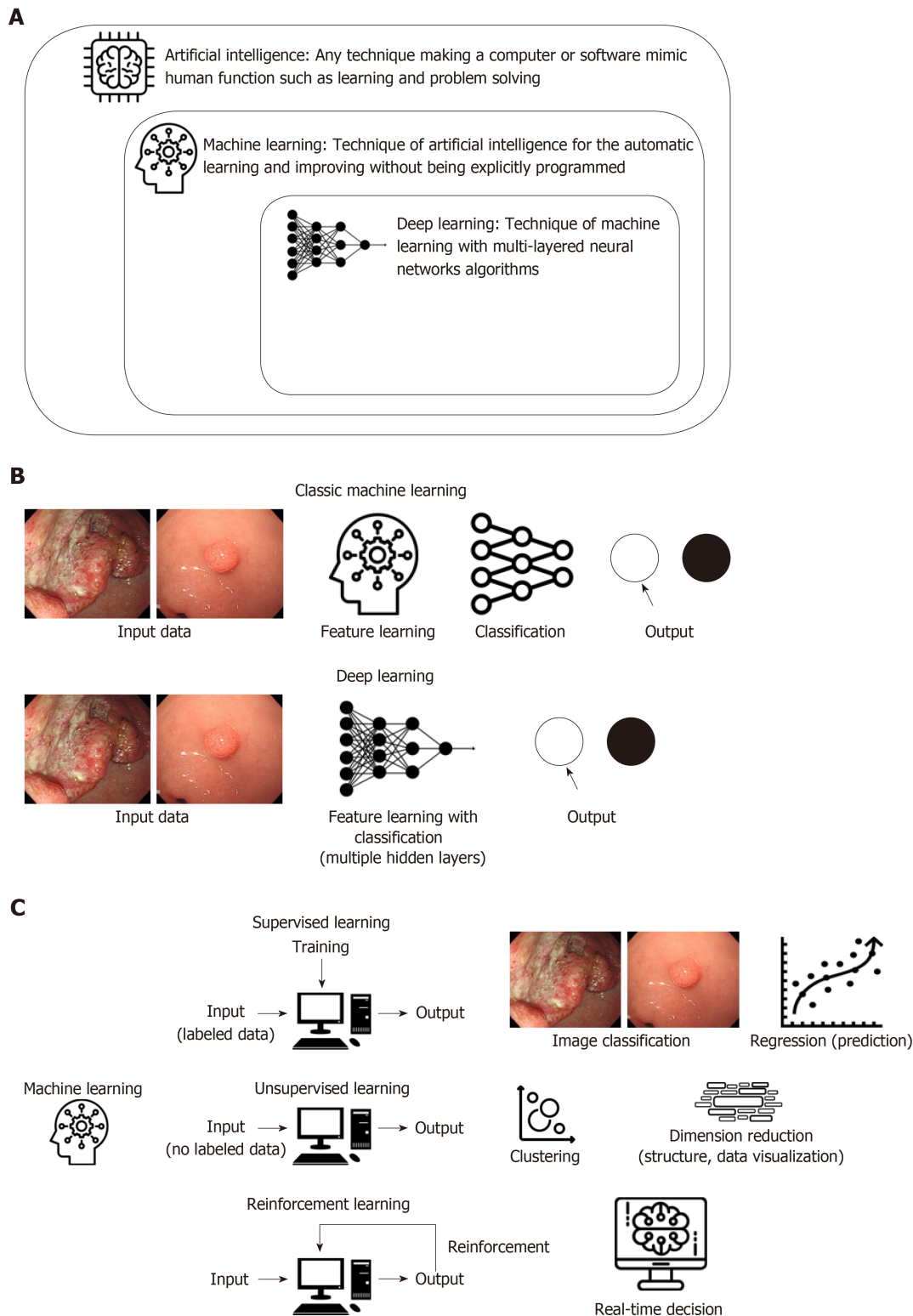
A CNN consists of (1) convolutional and pooling layers, which are the two main components to extract distinct features; and (2) fully connected layers to make an overall classification. The input images were filtered to extract specialized features using numerous specific filters, and to create multiple feature maps. This preprocessing operation for filtering is called convolution. A learning process for the convolution filter to make the best feature maps is essential for success in a CNN. These feature maps are compressed to smaller sizes by pooling the pixels to capture a larger field of the image, and these convolutional and pooling layers are iterated many times. Finally, fully connected layers combine all features and produce the final outcomes (Figure 1B).

The rapid growth of the CNN was demonstrated at the ImageNet Large Scale Visual Recognition Competition (ILSVRC) in 2012 by Geoffrey Hinton, and several CNNs such as Inception from google and ResNet from Microsoft have shown excellent performance. A graphical summary of AI, ML, and DL development is shown in Figure 1.

## APPLICATION OF AI IN GASTROENTEROLOGY

### *Recognition of diagnosis and prediction of prognosis*

Although AI in the field of gastroenterology recently focused on image analysis, several ML models have shown promising results in the recognition of diagnosis and prediction of prognosis. The ANN is appropriate for dealing with complex datasets to overcome the drawbacks of traditional linear statistics. In addition, the ANN can



**Figure 1 Schematic graphical summary for artificial intelligence, machine learning and deep learning development.** A: Definition of artificial intelligence, machine learning (ML) and deep learning (DL). B: Comparison of process between classic ML and DL. C: Modes of learning and examples of ML.

stand for the sophisticated interactions between demographic, environmental, and clinical characteristics.

In terms of diagnosis, Pace *et al*<sup>[11]</sup> demonstrated an ANN model in 2005 that made a diagnosis of gastroesophageal reflux disease using only 45 clinical variables in 159 cases with an accuracy of 100%. Lahner *et al*<sup>[12]</sup> performed a similar pilot study to recognize atrophic gastritis solely by using clinical and biochemical variables from 350 outpatients by using ANNs and linear discriminant analysis. This study showed great accuracy.

Regarding the prediction of prognosis, in 1998, Pofahl *et al*<sup>[13]</sup> compared an ANN



model to the Ranson criteria and the Acute Physiologic and Chronic Health Evaluation (APACHE II) scoring system to predict the length of stay for patients with acute pancreatitis. The authors used a backpropagation neural network that was trained using 156 patients. Although the highest specificity (94%) was observed in the Ranson criteria, the ANN model showed the highest sensitivity (75%) when predicting a length of stay more than 7 d. Similar accuracy was observed for the Ranson criteria and APACHE II scoring system<sup>[13]</sup>. In 2003, Das *et al*<sup>[14]</sup> used an ANN to predict the outcomes of acute lower gastrointestinal bleeding using 190 patients. The authors compared the performance of ANNs to a previously validated scoring system (BLEED), which revealed a significantly better predictive accuracy of mortality (87% *vs* 21%), recurrent bleeding (89% *vs* 41%), and the need for therapeutic intervention (96% *vs* 46%) in the ANN model.

Sato *et al*<sup>[15]</sup> presented an ANN model in 2005 to predict 1-year and 5-year survival using 418 esophageal cancer patients. This ANN model showed improved accuracy compared to the conventional linear discriminant analysis model.

Recently, the number of input training data items for ANNs was increased from hundreds to thousands of patients. Rotondano *et al*<sup>[16]</sup> compared the Rockall score to a supervised ANN model to predict the mortality of nonvariceal upper gastrointestinal bleeding using 2380 patients. This approach showed superior sensitivity (83.8% *vs* 71.4%), specificity (97.5% *vs* 52.0%), accuracy (96.8% *vs* 52.9%), and area under receiver operating characteristic (AUROC) of the predictive performance (0.95 *vs* 0.67) in the ANN model to those in the complete Rockall score.

Takayama *et al*<sup>[17]</sup> established an ANN model for the prediction of prognosis in patients with ulcerative colitis after cytoapheresis therapy and achieved a sensitivity and specificity for the need of an operation of 96% and 97%, respectively. Hardalaç *et al*<sup>[18]</sup> established an ANN model to predict mucosal healing by azathioprine therapy in patients with inflammatory bowel disease (IBD) and achieved 79.1% correct classifications. Peng *et al*<sup>[19]</sup> used an ANN model to predict the frequency of the onset, relapse, and severity of IBD. The researchers achieved an average accuracy to predict the frequency of onset and severity of IBD but a high accuracy in predicting the frequency of relapse of IBD (mean square error = 0.009, mean absolute percentage error = 17.1%).

SVMs have been used to analyze data and recognize patterns in classification analyses. Recently, Ichimasa *et al*<sup>[20]</sup> analyzed 45 clinicopathological factors in 690 endoscopically resected T1 colorectal cancer patients to predict lymph node metastasis using a SVM. This approach showed superior performance (sensitivity 100%, specificity 66%, accuracy 69%) compared to those of American (sensitivity 100%, specificity 44%, accuracy 49%), European (sensitivity 100%, specificity 0%, accuracy 9%), and Japanese (sensitivity 100%, specificity 0%, accuracy 9%) guidelines. A prediction model using a SVM model reduced the amount of unnecessary additional surgery (77%) when misdiagnosing lymph node metastasis than those of a prediction model using American (85%), European (91%), and Japanese guidelines (91%). Yang *et al*<sup>[21]</sup> constructed an SVM-based model using clinicopathological features and 23 immunologic markers from 483 patients who underwent curative surgery for esophageal squamous cell carcinoma. This study revealed reasonable performance in identifying high-risk patients with postoperative distant metastasis [sensitivity 56.6%, specificity 97.7%, positive predictive value (PPV) 95.6%, negative predictive value (NPV) 72.3%, and overall accuracy 78.7%] (Table 2).

### Analysis of images

Although endoscopic screening programs have reduced the mortality from gastrointestinal malignancies, they are still the leading cause of death worldwide and remain a global economic burden. To enhance the detection rate of gastrointestinal neoplasms and optimize the treatment strategies, a high-quality endoscopic examination for the recognition of gastrointestinal neoplasms and classifications between benign and malignant lesions are essential for the gastroenterologist. Thus, gastroenterologists are interested in the applications of AI, especially when using CNNs and SVMs for image analysis. Furthermore, AI has been increasingly adopted in terms of non-neoplastic gastrointestinal diseases including infection, inflammation, or hemorrhage.

**Upper gastrointestinal field:** Takiyama *et al*<sup>[22]</sup> constructed a CNN model that could recognize the anatomical location of EGD images with AUROCs of 1.00 for the larynx and esophagus, and 0.99 for the stomach and duodenum. This CNN model could also recognize specific anatomical locations within the stomach, with AUROCs of 0.99 for the upper, middle, and lower stomach.

To assist in the discrimination of early neoplastic lesions in Barrett's esophagus, van der Sommen *et al*<sup>[23]</sup> developed an automated algorithm to include specific

**Table 2** Summary of clinical studies using artificial intelligence for recognition of diagnosis and prediction of prognosis

Ref.	Published year	Aim of study	Design of study	Number of subjects	Type of AI	Input variables (number/type)	Outcomes
Pace <i>et al</i> <sup>[11]</sup>	2005	Diagnosis of gastroesophageal reflux disease	Retrospective	159 patients (10 times cross validation)	"backpropagation" ANN	101/clinical variables	Accuracy: 100%
Lahner <i>et al</i> <sup>[12]</sup>	2005	Recognition of atrophic corpus gastritis	Retrospective	350 patients (subdivided several times into training and test set equally)	ANN	37 to 3 /clinical and biochemical variables (experiment 1 to 5)	Accuracy: 96.6%, 98.8%, 98.4%, 91.3% and 97.7% (experiment 1-5, respectively)
Pofahl <i>et al</i> <sup>[13]</sup>	1998	Prediction of length of stay for patients with acute pancreatitis	Retrospective	195 patients (training set: 156, test set: 39)	"backpropagation" ANN	71/clinical variables	Sensitivity: 75 % (for prediction of a length of stay more than 7 d)
Das <i>et al</i> <sup>[14]</sup>	2003	Prediction of outcomes in acute lower gastrointestinal bleeding	Prospective	190 patients (training set: 120, internal validation set: 70, external validation set: 142)	ANN	26/clinical variables	Accuracy (external validation set): 97% for death, 93% for, recurrent bleeding, 94% for need for intervention
Sato <i>et al</i> <sup>[15]</sup>	2005	Prediction of 1-year and 5-year survival of esophageal cancer	Retrospective	418 patients (training-: validation-: test set = 53%: 27%: 20%)	ANN	199/ clinicopathologic, biologic, and genetic variables	AUROC for 1 year- and 5 year survival prediction: 0.883 and 0.884, respectively
Rotondano <i>et al</i> <sup>[16]</sup>	2011	Prediction of mortality in nonvariceal upper gastrointestinal bleeding	Prospective, multicenter	2380 patients (5 × 2 cross-validation)	ANN	68/clinical variables	Accuracy: 96.8%, AUROC: 0.95, sensitivity: 83.8%, specificity: 97.5%,
Takayama <i>et al</i> <sup>[17]</sup>	2015	Prediction of prognosis in ulcerative colitis after cytoapheresis therapy	Retrospective	90 patients (training set: 54, test set: 36)	ANN	13/clinical variables	Sensitivity: 96.0%, specificity: 97.0%
Hardalaç <i>et al</i> <sup>[18]</sup>	2015	Prediction of mucosal healing by azathioprine therapy in IBD	Retrospective	129 patients (training set: 103, validation set: 13, test set: 13)	"feed-forward back-propagation" and "cascade-forward" ANN	6/clinical variables	Total correct classification rate: 79.1%
Peng <i>et al</i> <sup>[19]</sup>	2015	Prediction of frequency of onset, relapse, and severity of IBD	Retrospective	569 UC and 332 CD patients (training set: data from 2003-2010, validation set: data in 2011)	ANN	5/meteorological data	Accuracy in predicting the frequency of relapse of IBD (mean square error = 0.009, mean absolute percentage error = 17.1%)
Ichimasa <i>et al</i> <sup>[20]</sup>	2018	Prediction of lymph node metastasis, thus minimizing the need for additional surgery in T1 colorectal cancer	Retrospective	690 patients (training set: 590, validation set: 100)	SVM	45/ Clinicopathologic al variables	Accuracy: 69%, sensitivity: 100%, specificity: 66%
Yang <i>et al</i> <sup>[21]</sup>	2013	Prediction of postoperative distant metastasis in esophageal squamous cell carcinoma	Retrospective	483 patients (training set: 319, validation set: 164)	SVM	30/7 clinicopathologica l variables and 23 immunomarkers	Accuracy: 78.7% sensitivity: 56.6%, specificity: 97.7%, PPV: 95.6%, NPV: 72.3%

AI: Artificial intelligence; ANN: Artificial neural network; AUROC: Area under receiver operating characteristic; IBD: Inflammatory bowel disease; UC: Ulcerative colitis; CD: Crohn's disease; SVM: Support vector machine; PPV: Positive predictive value; NPV: Negative predictive value.

textures, color filters, and ML from 100 endoscopic images. This algorithm reasonably detected early neoplastic lesions in a per-image analysis with a sensitivity and specificity of 83%. In 2017, the same group investigated a model to improve the detection rate of early neoplastic lesions in Barrett's esophagus by using 60 *ex vivo* volumetric laser endomicroscopy images. This novel computer model showed optimal performance compared with a clinical volumetric laser endomicroscopy prediction score with a sensitivity of 90% and specificity of 93%<sup>[24]</sup>.

Several studies evaluated the ML model using specialized endoscopy to differentiate neoplastic/dysplastic and non-neoplastic lesions. Kodashima *et al*<sup>[25]</sup> showed that computer-based analysis can easily identify malignant tissue at the cellular level using endocytoscopic images, which enables the microscopic visualization of the mucosal surface. In 2015, Shin *et al*<sup>[26]</sup> reported on an image analysis model to detect esophageal squamous dysplasia using high-resolution microendoscopy (HRME). The sensitivity and specificity of this model were 87% and 97%, respectively. During the following year, Quang *et al*<sup>[27]</sup> from the same study group evolved this model, which was incorporated in tablet-interfaced HRME with full automation for real-time analysis. As a result, the model reduced the costs compared to previous laptop-interfaced HRME and showed good diagnostic yields of esophageal squamous cell carcinoma with a sensitivity and specificity of 95% and 91%, respectively. However, there was a limitation for the application of this model owing to the unavailability of specialized endoscopy.

Finally, Horie *et al*<sup>[28]</sup> demonstrated the utility of AI using CNNs to make a diagnosis of esophageal cancer. This was trained with 8428 conventional endoscopic images including white-light images (WLIs) and narrow-band images (NBIs). This CNN model detected esophageal cancer with a sensitivity of 95% and could identify all small cancers of < 10 mm. This model also distinguished superficial esophageal cancer from advanced cancer with an accuracy of 98%.

*Helicobacter pylori* (*H. pylori*) infection is the most important risk factor of peptic ulcers and gastric cancer. Several researchers challenged AI to aid the endoscopic diagnosis of *H. pylori* infections. In 2004, Huang *et al*<sup>[29]</sup> investigated the predictability of *H. pylori* infection by refined feature selection with a neural network using related gastric histologic features in endoscopic images. This model was trained and analyzed with 84 image parameters from 30 patients. The sensitivity and specificity for the detection of *H. pylori* infection were 85.4% and 90.9%, respectively. In addition, the accuracy of this model for identifying gastric atrophy, intestinal metaplasia, and predicting the severity of *H. pylori*-related gastric inflammation was higher than 80%.

Recently, two Japanese researchers reported on the application of a CNN to make a diagnosis of *H. pylori* infection<sup>[30,31]</sup>. Itoh *et al*<sup>[31]</sup> developed a CNN model to recognize *H. pylori* infections by using 596 endoscopic images after the data augmentation of a prior set of 149 images. This CNN model showed promising results with a sensitivity and specificity of 86.7% and 86.7%, respectively. Shichijo *et al*<sup>[30]</sup> compared the performance of a CNN to that of 23 endoscopists for the diagnosis of *H. pylori* infection by using endoscopic images. The CNN model showed superior sensitivity (88.9% *vs* 79.0%), specificity (87.4% *vs* 83.2%), accuracy (87.7% *vs* 82.4%), and diagnostic time (194 s *vs* 230 s).

In 2018, a prospective pilot study was conducted for automated diagnosis of *H. pylori* infections using image-enhanced endoscopy such as blue laser imaging-bright and linked color imaging. The performance of the developed AI model was significantly higher with blue laser imaging-bright and linked color imaging training (AUROCs of 0.96 and 0.95) than WLI imaging training (0.66)<sup>[32]</sup>.

The utility of AI in the diagnosis of gastrointestinal neoplasms was classified into two main categories: detection and characterization. In 2012, Kubota *et al*<sup>[33]</sup> first evaluated a computer-aided pattern recognition system to identify the depth of the wall invasion of gastric cancer using endoscopic images. They used 902 endoscopic images and created a backpropagation model after a 10-time cross validation. As a result, the diagnostic accuracy was 77.2%, 49.1%, 51.0%, and 55.3% for T1-4 staging, respectively. In particular, the accuracy of T1a (mucosal invasion) and T1b staging (submucosal invasion) was 68.9% and 63.6%, respectively. Hirasawa *et al*<sup>[34]</sup> reported on the good performance of a CNN-based diagnostic system to detect gastric cancers in endoscopic images. The authors trained the CNN model using 13584 endoscopic images and tested it with 2296 images. The overall sensitivity was 92.2%. In addition, the detection rate with a diameter of 6 mm or more was 98.6%, and all invasive cancers were identified<sup>[34]</sup>. All missed lesions were superficially depressed and differentiated-type intramucosal cancers that were difficult to distinguish from gastritis even for experienced endoscopists. However, 69.4% of the lesions that the CNN diagnosed as gastric cancer were benign, and the most common reasons for misdiagnosis were gastritis with redness, atrophy, and intestinal metaplasia<sup>[34]</sup>.

Zhu *et al*<sup>[35]</sup> further applied a CNN system to discriminate the invasion depth of

gastric cancer (M/SM1 vs deeper than SM1) using conventional endoscopic images. They trained a CNN model with 790 images and tested it with another 203 images. The CNN model showed high accuracy (89.2%) and specificity (95.6%) when determining the invasion depth of gastric cancer. This result was significantly superior to that of experienced endoscopists. Kanesaka *et al*<sup>[36]</sup> studied a computer-aided diagnosis system using a SVM to facilitate the use of magnifying NBI to distinguish early gastric cancer. The study reported on remarkable potential in terms of diagnostic performance (accuracy 96.3%, PPV 98.3%, sensitivity 96.7%, and specificity 95%) and the performance of area concordance (accuracy 73.8%, PPV 75.3%, sensitivity 65.5%, and specificity 80.8%).

In terms of hepatology, the ultrasound has been challenged for the application of AI. Gatos *et al*<sup>[37]</sup> established a SVM diagnostic model of chronic liver disease using ultrasound shear wave elastography (70 patients with chronic liver disease and 56 healthy controls). The performance was promising, with an accuracy of 87.3%, sensitivity of 93.5%, and specificity of 81.2%, although the prospective validation was not conducted. Kuppli *et al*<sup>[38]</sup> established a fatty liver detection and characterization model using a single-layer feed-forward neural network, and validated this model with a higher accuracy than the previous SVM-based model. These researchers used ultrasound images of 63 patients, and the gold standard for labeling for each patient was the pathologic results of a liver biopsy.

The determination of liver cirrhosis was also challenged with ML technology. Liu *et al*<sup>[39]</sup> developed a CNN model with ultrasound liver capsule images (44 images from controls and 47 images from patients with cirrhosis), and classified these images using a SVM. The AUROC for the classification was 0.951, although the prospective validation was not conducted.

**Lower gastrointestinal field:** Among various gastrointestinal fields, the development of an AI model using colonoscopy has been the most promising area because polyp detection during colonoscopies is frequent. This provides sufficient sources for AI training, and a missed colorectal polyp is directly associated with interval colorectal cancer development.

In terms of polyp detection, Fernandez-Esparrach *et al*<sup>[40]</sup> established an automated computer-vision method using an energy map to detect colonic polyps in 2016. They used 24 videos containing 31 polyps and showed acceptable performance with a sensitivity of 70.4% and a specificity of 72.4% for polyp detection (Table 3). Recently, this performance was improved with a DL application for polyp detection<sup>[41,42]</sup>. Misawa *et al*<sup>[41]</sup> designed a CNN model using 546 short videos from 73 full-length videos, which were divided into two groups of training data (105 polyp-positive videos and 306 polyp-negative videos) and test data (50 polyp-positive videos and 85 polyp-negative videos). The researchers showed the possibility of the automated detection of colonic polyps in real time, and the sensitivity and specificity were 90.0% and 63.3%, respectively. Urban *et al*<sup>[42]</sup> also used a CNN system to identify colonic polyps. They used 8641 hand-labeled images and 20 colonoscopy videos in various combinations as training and test data. The CNN model detected polyps in real time with an AUROC of 0.991 and an accuracy of 96.4%. Moreover, it assisted in the identification of an additional nine polyps compared with expert endoscopists in the application of test colonoscopy videos.

Although there were many promising performances of the automated polyp detection models, a prospective validation was not conducted<sup>[43-45]</sup>. However, Klare *et al*<sup>[46]</sup> performed a prototype software validation under real-time conditions (55 routine colonoscopies), and the results were comparable between those of endoscopists and the established software. The endoscopists' polyp detection rates and adenoma detection rates were 56.4% and 30.9%, respectively, and these rates were 50.9% and 29.1% for the software, respectively). Wang *et al*<sup>[47]</sup> established a DL algorithm by using data from 1290 patients, and validated this model with 27113 newly collected colonoscopy images from 1138 patients. This model showed remarkable performance with a sensitivity of 94.38%, specificity of 95.2%, and AUROC of 0.984 for at least one polyp detection<sup>[47]</sup>.

For AI applications of polyp characterization, magnifying endoscopic images, which is useful when discriminating pit or vascular patterns, was first adopted to enhance the performance of AI. Tischendorf *et al*<sup>[48]</sup> developed an automated classification model of colorectal polyps by magnifying NBI images to evaluate vascular patterns in 2010. They reported that the overall accurate classification rates were 91.9% for a consensus decision between the human observers and 90.9% for a safe decision (classifying polyps as neoplastic in cases when there was an interobserver discrepancy)<sup>[48]</sup>. In 2011, Gross *et al*<sup>[49]</sup> compared the performances of a computer-based model for the differentiation of small colonic polyps of < 10 mm using NBI images. The expert endoscopists and computer-based model showed

**Table 3 Summary of clinical studies using artificial intelligence in the upper gastrointestinal field**

Ref.	Published year	Aim of study	Design of study	Number of subjects	Type of AI	Endoscopic or ultrasound modality	Outcomes
Takiyama <i>et al</i> <sup>[22]</sup>	2018	Recognition of anatomical locations of EGD images	Retrospective	Training set: 27335 images from 1750 patients. Validation set: 17081 images from 435 patients	CNN	White-light endoscopy	AUROC: 1.00 for the larynx and esophagus, and 0.99 for the stomach and duodenum recognition
van der Sommen <i>et al</i> <sup>[23]</sup>	2016	Discrimination of early neoplastic lesions in Barrett's esophagus	Retrospective	100 endoscopic images from 44 patients (leave-one-out cross-validation on a per-patient basis)	SVM	White-light endoscopy	Sensitivity: 83%, specificity: 83% (per-image analysis)
Swager <i>et al</i> <sup>[24]</sup>	2017	Identification of early Barrett's esophagus neoplasia on ex vivo volumetric laser endomicroscopy images.	Retrospective	60 volumetric laser endomicroscopy images	Combination of several methods (SVM, discriminant analysis, AdaBoost, random forest, etc)	Ex vivo volumetric laser endomicroscopy	Sensitivity: 90%, specificity: 93%
Kodashima <i>et al</i> <sup>[25]</sup>	2007	Discrimination between normal and malignant tissue at the cellular level in the esophagus	Prospective <i>ex vivo</i> pilot	10 patients	ImageJ program	Endocytoscopy	Difference in the mean ratio of total nuclei to the entire selected field, $6.4 \pm 1.9\%$ in normal tissues and $25.3 \pm 3.8\%$ in malignant samples
Shin <i>et al</i> <sup>[26]</sup>	2015	Diagnosis of esophageal squamous dysplasia	Prospective, multicenter	375 sites from 177 patients (training set: 104 sites, test set: 104 sites, validation set: 167 sites)	Linear discriminant analysis	HRME	Sensitivity: 87%, specificity: 97%
Quang <i>et al</i> <sup>[27]</sup>	2016	Diagnosis of esophageal squamous cell neoplasia	Retrospective, multicenter	Same data from reference number 26	Linear discriminant analysis	Tablet-interfaced HRME	Sensitivity: 95%, specificity: 91%
Horie <i>et al</i> <sup>[28]</sup>	2019	Diagnosis of esophageal cancer	Retrospective	Training set: 8428 images from 384 patients. Test set: 1118 images from 97 patients	CNN	White-light endoscopy with NBI	Sensitivity 98%
Huang <i>et al</i> <sup>[29]</sup>	2004	Diagnosis of <i>H. pylori</i> infection	Prospective	Training set: 30 patients. Test set: 74 patients	Refined feature selection with neural network	White-light endoscopy	Sensitivity: 85.4%, specificity: 90.9%
Shichijo <i>et al</i> <sup>[30]</sup>	2017	Diagnosis of <i>H. pylori</i> Infection	Retrospective	Training set: CNN1: 32208 images; CNN2: images classified according to 8 different locations in the stomach. Test set: 11481 images from 397 patients	CNN	White-light endoscopy	Accuracy: 87.7%, sensitivity: 88.9%, specificity: 87.4%, diagnostic time: 194 s.
Itoh <i>et al</i> <sup>[31]</sup>	2018	Diagnosis of <i>H. pylori</i> infection	Prospective	Training set: 149 images (596 images through data augmentation. Test set: 30 images)	CNN	White-light endoscopy	AUROC: 0.956, sensitivity: 86.7%, specificity: 86.7%



Nakashima <i>et al</i> <sup>[32]</sup>	2018	Diagnosis of <i>H. pylori</i> infection	Prospective pilot	222 patients (training set: 162, test set: 60)	CNN	White-light endoscopy and image-enhanced endoscopy, such as blue laser imaging-bright and linked color imaging	AUROC: 0.96 (blue laser imaging-bright), 0.95 (linked color imaging)
Kubota <i>et al</i> <sup>[33]</sup>	2012	Diagnosis of depth of invasion in gastric cancer	Retrospective	902 images (10 times cross validation)	"backpropagation" ANN	White-light endoscopy	Accuracy: 77.2%, 49.1%, 51.0%, and 55.3% for T1-4 staging, respectively
Hirasawa <i>et al</i> <sup>[34]</sup>	2018	Detection of gastric cancers	Retrospective	Training set: 13584 images. Test set: 2296 images.	CNN	White-light endoscopy, chromoendoscopy, NBI	Sensitivity: 92.2%, detection rate with a diameter of 6 mm or more: 98.6%
Zhu <i>et al</i> <sup>[35]</sup>	2018	Diagnosis of depth of invasion in gastric cancer (mucosa/SM1/deeper than SM1)	Retrospective	Training set: 790 images. Test set: 203 images	CNN	White-light endoscopy	Accuracy: 89.2%, AUROC: 0.94, sensitivity: 74.5%, specificity: 95.6%
Kanesaka <i>et al</i> <sup>[36]</sup>	2018	Diagnosis of early gastric cancer using magnifying NBI images	Retrospective	Training set: 126 images. Test set: 81 images	SVM	Magnifying NBI	Accuracy: 96.3%, sensitivity: 96.7%, specificity: 95%, PPV: 98.3%,
Gatos <i>et al</i> <sup>[37]</sup>	2017	Diagnosis of chronic liver disease	Retrospective	126 patients (56 healthy controls, 70 with chronic liver disease)	SVM	Ultrasound shear wave elastography imaging with a stiffness value-clustering	AUROC: 0.87, highest accuracy: 87.3%, sensitivity: 93.5%, specificity: 81.2%
Kuppili <i>et al</i> <sup>[38]</sup>	2017	Detection and characterization of fatty liver	Prospective	63 patients who underwent liver biopsy (10 times cross validation)	Extreme Learning Machine to train single-layer feed-forward neural network	Ultrasound liver images	Accuracy: 96.75%, AUROC: 0.97 (validation performance)
Liu <i>et al</i> <sup>[39]</sup>	2017	Diagnosis of liver cirrhosis	Retrospective	44 images from controls and 47 images from patients with cirrhosis	SVM	Ultrasound liver capsule images	AUROC: 0.951

AI: Artificial intelligence; EGD: Esophagogastroduodenoscopy; CNN: Convolutional neural network; AUROC: Area under receiver operating characteristic; SVM: Support vector machine; HRME: High-resolution microendoscopy; NBI: Narrow band image; *H. pylori*: *Helicobacter pylori*; ANN: Artificial neural network; PPV: Positive predictive value.

comparable diagnostic performance in sensitivity (93.4% *vs* 95.0%), specificity (91.8% *vs* 90.3%), and accuracy (92.7% *vs* 93.1%)<sup>[49]</sup>.

Takemura *et al*<sup>[50]</sup> retrospectively compared the identification of pit patterns of a computer-based model with shape descriptors such as area, perimeter, fit ellipse, or circularity in reference to endoscopic diagnosis by using magnified endoscopic images with crystal violet staining in 2010. The accuracies of the type I, II, III<sub>L</sub>, and IV pit patterns of colorectal lesions were 100%, 100%, 96.6%, and 96.7%, respectively. In 2012, the authors applied an upgraded version of a computer system *via* SVM to distinguish neoplastic and non-neoplastic lesions by using endoscopic NBI images, which showed a detection accuracy of 97.8%<sup>[51]</sup>. They further demonstrated the availability of a real-time image recognition system in 2016, and the accuracy between the pathologic results of diminutive polyps and diagnosis by a real-time image recognition model was 93.2%<sup>[52]</sup>.

Byrne *et al*<sup>[53]</sup> developed a CNN model for the real-time differentiation of diminutive colorectal polyps by using only NBI video frames in 2017. This model discriminated adenomas from hyperplastic polyps with an accuracy of 94%, and identified the adenoma with a sensitivity of 98% and a specificity of 83%<sup>[53]</sup>. Likewise, Chen *et al*<sup>[54]</sup> made a CNN model trained with 2157 images to identify neoplastic or hyperplastic polyps of < 5 mm with a PPV and NPV of 89.6% and 91.5%, respectively. In 2017, Komeda *et al*<sup>[55]</sup> reported on the preliminary data of a CNN model to distinguish adenomas from non-adenomatous polyps. The CNN model was trained with 1800 conventional endoscopic images with WLI, NBI, and chromoendoscopy, and the accuracy of a 10-fold cross-validation was 75.1%.

To enhance the differentiation of polyps, a Japanese study group reported several articles for AI application with endocytoscopy images, which enables the observation of nuclei on site, and showed comparable diagnostic results to those of pathologic examinations. In 2015, these researchers first developed a computer-aided diagnosis system using endocytoscopy for the discrimination of neoplastic changes in small polyps. This approach showed a comparable sensitivity (92.0%) and accuracy (89.2%) with those of expert endoscopists<sup>[56]</sup>. In 2016, this research team developed a second-generation model that could (1) evaluate both nuclei and ductal lumens, (2) use an SVM instead of multivariate analysis, (3) provide the confidence levels of the decisions, and (4) provide a more rapid process of discriminating neoplastic changes from 0.3 s to 0.2 s. The endocytoscopic microvascular patterns could be effectively evaluated by staining with dye<sup>[57]</sup>. These researchers also developed endocytoscopy with NBI without staining to evaluate microvascular findings. This approach showed an overall accuracy of 90%<sup>[57]</sup>. The same group performed a prospective validation of a real-time computer-aided diagnosis system using endocytoscopy with NBI or stained images to identify neoplastic diminutive polyps. The researchers reported a pathologic prediction rate of 98.1%, and the time required to assess one diminutive polyp was about 35 to 47 s<sup>[59]</sup>.

The application of a computer-aided ultrahigh (approximately 400 ×) magnification endocytoscopy system for the diagnosis of invasive colorectal cancers was investigated by Takeda *et al*<sup>[60]</sup>. This system was trained with 5543 endocytoscopic images from 238 lesions and reported a sensitivity of 89.4%, specificity of 98.9%, and accuracy of 94.1% using 200 test images<sup>[60]</sup>.

For the application of AI in IBD, Maeda *et al*<sup>[61]</sup> developed a diagnosis system using a SVM after refining previous computer-aided endocytoscopy systems<sup>[56-58]</sup>. They evaluated the diagnostic performance of this model for the prediction of persistent histologic inflammation in ulcerative colitis patients. This model showed good performance with a sensitivity of 74%, specificity of 97%, and an accuracy of 91%<sup>[61]</sup>.

Currently, the resolution of images is relatively low in capsule endoscopy compared to other digestive endoscopies. Moreover, the interpretation and diagnosis of capsule endoscopy images highly depends on the reviewer's ability and effort. It is also a time-consuming process. Therefore, several conditions were attempted for the automated diagnosis of capsule endoscopy images including angiectasia, celiac disease, or intestinal hookworms, or for small intestinal motility characterization<sup>[62-65]</sup>.

Leenhardt *et al*<sup>[62]</sup> developed a gastrointestinal angiectasia detection model using semantic segmentation images with a CNN. They used 600 control images and 600 typical angiectasia images to form 4166 small bowel capsule endoscopy videos, which were divided equally into training and test data sets. The CNN-based model revealed a high diagnostic performance with a sensitivity of 100%, specificity of 96%, PPV of 96%, and NPV of 100%<sup>[62]</sup> (Table 4). Zhou *et al*<sup>[63]</sup> established a CNN model for the classification of celiac disease from control with capsule endoscopy clips from six celiac disease patients and five controls. The researchers achieved 100% sensitivity and specificity for the test data set. Moreover, the evaluation confidence was related to the severity level of small bowel mucosal lesions, reflecting the potential for the quantitative measurement of the existence and degree of pathology throughout the small intestine<sup>[63]</sup>. Intestinal hookworms are difficult to find with direct visualization because they have small tubular structures with a whitish color and semitransparent features similar to background intestinal mucosa. Moreover, the presence of intestinal secretory materials makes them difficult to detect. He *et al*<sup>[64]</sup> established a CNN model for the detection of hookworms in capsule endoscopy images. The CNN-based model showed a reasonable performance with a sensitivity of 84.6%, specificity of 88.6% and only 15% hookworm images and 11% non-hookworm image were falsely detected.

The interpretation of wireless motility capsule endoscopy is a complex task. Seguí *et al*<sup>[65]</sup> established a CNN model for small-intestine motility characterization and achieved a mean classification accuracy of 96% for six intestinal motility events ("turbid", "bubbles", "clear blob", "wrinkles", "wall", and, "undefined"). This outperformed the other classifiers by a large margin (a 14% relative performance increase).

## CHALLENGES AND FUTURE DIRECTIONS FOR APPLICATION OF AI

Although many researchers have investigated the utility of AI and have shown promising results, most studies were designed in retrospective manner: as a case-control study from a single center, or by using endoscopic images that were chosen from specific endoscopic modalities unavailable from many institutions. Potential

Table 4 Summary of clinical studies using artificial intelligence in the lower gastrointestinal field

Ref.	Published year	Aim of study	Design of study	Number of subjects	Type of AI	Endoscopic modality	Outcomes
Fernandez-Esparrach <i>et al</i> <sup>[40]</sup>	2016	Detection of colonic polyps	Retrospective	24 videos containing 31 polyps	Window Median Depth of Valleys Accumulation maps	White-light colonoscopy	Sensitivity: 70.4%. Specificity: 72.4%
Misawa <i>et al</i> <sup>[41]</sup>	2018	Detection of colonic polyps	Retrospective	546 short videos (training set: 105 polyp-positive videos and 306 polyp-negative videos, test set: 50 polyp-positive videos and 85 polyp-negative videos) from 73 full length videos	CNN	White-light colonoscopy	Accuracy: 76.5%. Sensitivity: 90.0%. Specificity: 63.3%.
Urban <i>et al</i> <sup>[42]</sup>	2018	Detection of colonic polyps	Retrospective	8641 images with 20 colonoscopy videos	CNN	White-light colonoscopy with NBI	Accuracy: 96.4%. AUROC: 0.991
Klare <i>et al</i> <sup>[46]</sup>	2019	Detection of colonic polyps	Prospective	55 patients	Automated polyp detection software	White-light colonoscopy	Polyp detection rate: 50.9%. Adenoma detection rate: 29.1%
Wang <i>et al</i> <sup>[47]</sup>	2018	Detection of colonic polyps	Retrospective	Training set: 5545 images from 1290 patients. Validation set A: 27113 images from 1138 patients. Validation set B: 612 images. Validation set C: 138 video clips from 110 patients. Validation set D: 54 videos from 54 patients	CNN	White-light colonoscopy	Dataset A: AUROC: 0.98 for at least one polyp detection, per-image sensitivity: 94.4%, per-image specificity: 95.2%. Dataset B: per-image sensitivity: 88.2%. Dataset C: per-image sensitivity: 91.6%, per-polyp sensitivity: 100%. Dataset D: per-image specificity: 95.4%
Tischendorf <i>et al</i> <sup>[48]</sup>	2010	Classification of colorectal polyps on the basis of vascularization features.	Prospective pilot	209 polyps from 128 patients	SVM	Magnifying NBI images	Accurate classification rate: 91.9%
Gross <i>et al</i> <sup>[49]</sup>	2011	Differentiation of small colonic polyps of < 10 mm	Prospective	434 polyps from 214 patients	SVM	Magnifying NBI images	Accuracy: 93.1%. Sensitivity: 95.0%. Specificity: 90.3%.
Takemura <i>et al</i> <sup>[50]</sup>	2010	Classification of pit patterns	Retrospective	Training set: 72 images. Validation set: 134 images	HuPAS software version 1.3	Magnifying endoscopic images with crystal violet staining	Accuracies of the type I, II, III <sub>L</sub> , and IV pit patterns of colorectal lesions: 100%, 100%, 96.6%, and 96.7%, respectively
Takemura <i>et al</i> <sup>[51]</sup>	2012	Classification of histology of colorectal tumors	Retrospective	Training set: 1519 images. Validation set: 371 images	HuPAS software version 3.1 using SVM	Magnifying NBI images	Accuracy: 97.8%
Kominami <i>et al</i> <sup>[52]</sup>	2016	Classification of histology of colorectal polyps	Prospective	Training set: 2247 images from 1262 colorectal lesion. Validation: 118 colorectal lesions	SVM with logistic regression	Magnifying NBI images	Accuracy: 93.2%, Sensitivity: 93.0%, Specificity: 93.3%, PPV: 93%, NPV: 93.3%
Byrne <i>et al</i> <sup>[53]</sup>	2017	Differentiation of histology of diminutive colorectal polyps	Retrospective	Training set: 223 videos, Validation set: 40 videos. Test set: 125 videos	CNN	NBI video frames	Accuracy: 94%, Sensitivity: 98%, Specificity: 83%

Chen <i>et al</i> <sup>[54]</sup>	2018	Identification of neoplastic or hyperplastic polyps of < 5 mm	Retrospective	Training set: 2157 images. Test set: 284 images	CNN	Magnifying NBI images	Sensitivity: 96.3%, specificity: 78.1%, PPV: 89.6%, NPV: 91.5%
Komeda <i>et al</i> <sup>[55]</sup>	2017	Discrimination adenomas from non-adenomatous polyps	Retrospective	1200 images from the endoscopic videos (10 times cross validation)	CNN	White-light colonoscopy with NBI and chromoendoscopy	Accuracy in validation: 75.1%
Mori <i>et al</i> <sup>[56]</sup>	2015	Discrimination of neoplastic changes in small polyps	Retrospective	Test set: 176 polyps from 152 patients	Multivariate regression analysis	Endocytoscopy	Accuracy: 89.2%, Sensitivity: 92.0%
Mori <i>et al</i> <sup>[57]</sup>	2016	Development of 2 <sup>nd</sup> generation model, which was mentioned in reference number 56	Retrospective	Test set: 205 small colorectal polyps ( $\leq 10$ mm) from 123 patients	SVM	Endocytoscopy	Accuracy: 89% for both diminutive (< 5 mm) and small (< 10 mm) polyps
Misawa <i>et al</i> <sup>[58]</sup>	2016	Diagnosis of colorectal lesions using microvascular findings	Retrospective	Training set: 979 images, validation set: 100 images	SVM	Endocytoscopy with NBI	Accuracy: 90%
Mori <i>et al</i> <sup>[59]</sup>	2018	Diagnosis of neoplastic diminutive polyp	Prospective	466 diminutive polyps from 325 patients	SVM	Endocytoscopy with NBI and stained images	Prediction rate: 98.1%
Takeda <i>et al</i> <sup>[60]</sup>	2017	Diagnosis of invasive colorectal cancer	Retrospective	Training set: 5543 images from 238 lesions. Test set: 200 images	SVM	Endocytoscopy with NBI and stained images	Accuracy: 94.1% Sensitivity: 89.4%, Specificity: 98.9%, PPV: 98.8%, NPV: 90.1%
Maeda <i>et al</i> <sup>[61]</sup>	2018	Prediction of persistent histologic inflammation in ulcerative colitis patients	Retrospective	Training set: 12900 images. Test set: 9935 images	SVM	Endocytoscopy with NBI	Accuracy: 91%, Sensitivity: 74%, Specificity: 97%

AI: Artificial intelligence; CNN: Convolutional neural network; NBI: Narrow band image; AUROC: Area under receiver operating characteristic; SVM: Support vector machine; PPV: Positive predictive value; NPV: Negative predictive value.

inherent bias such as selection bias cannot be excluded in this situation. Therefore, it is crucial to meticulously validate the performance of AI before the application of AI in real clinical practice. To properly verify the accuracy of AI, physicians should understand the effects of overfitting and spectrum bias (class imbalance) on the performance of AI, and try to evaluate the performance by avoiding these biases.

Overfitting occurs when a learning model tailors itself too much on the training dataset and predictions are not well generalized to new datasets<sup>[66]</sup> (Table 5). Although several methods were used to reduce overfitting in the development of DL models, they did not guarantee the resolution of this problem. In addition, datasets that were collected by case-control design are particularly vulnerable to spectrum bias. Spectrum bias occurs when the dataset used for model development does not adequately represent the range of patients who will be applied in clinical practice (target population)<sup>[67]</sup>.

Because overfitting and spectrum bias may lead to overestimation of the accuracy and generalization, external validation using unused datasets for model development, collected in a way that minimizes the spectrum bias, is mandatory. For more robust clinical verification, well-designed multicenter prospective studies with adequate inclusion/exclusion criteria that represent the target population are needed. Furthermore, DL technology has its own “black box” nature (lack of interpretability or explainability), which means the decision mechanism of AI is not clearly demonstrated (Figure 2). Because interpretability is important in that it can provide safety measures, help to detect bias, and establish social acceptance, further investigation to solve this issue should be performed. However, there have been some methods to complement “black box” characteristics, such as the attention map and saliency region<sup>[68]</sup>.

It is obvious that the efficiency and accuracy of ML increases as the amount of data increases; however, it is challenging to develop an efficient ML model owing to the paucity of human labeled data given the issue of privacy with regard to private

**Table 5 Summary of clinical studies using artificial intelligence in the capsule endoscopy**

Ref.	Published year	Aim of study	Design of study	Number of subjects	Type of AI	Outcomes
Leenhardt <i>et al</i> <sup>[62]</sup>	2019	Detection of gastrointestinal angiectasia	Retrospective	600 control images and 600 typical angiectasia images (divided equally into training and test datasets)	CNN	Sensitivity: 100%, specificity: 96%, PPV: 96%, NPV: 100%.
Zhou <i>et al</i> <sup>[63]</sup>	2017	Classification of celiac disease	Retrospective	Training set: 6 celiac disease patients, 5 controls. Test set: additional 5 celiac disease patients, 5 controls	CNN	Sensitivity: 100%, specificity: 100% (for test dataset)
He <i>et al</i> <sup>[64]</sup>	2018	Detection of intestinal hookworms	Retrospective	440000 images	CNN	Sensitivity: 84.6%, specificity: 88.6%
Seguí <i>et al</i> <sup>[65]</sup>	2016	Characterization of small intestinal motility	Retrospective	120000 images (training set: 100000, test set: 20000)	CNN	Mean classification accuracy: 96%

AI: Artificial intelligence; CNN: Convolutional neural network; PPV: Positive predictive value; NPV: Negative predictive value.

medical records. To overcome this issue, data augmentation strategies (with synthetically modified data) have been proposed<sup>[69]</sup>. Spiking neural networks, which more closely mimic the real mechanisms of neurons, can potentially replace current ANN models with more powerful computing ability, although no effective supervised learning method currently exists<sup>[70]</sup>.

The precision of diagnosis or classification using AI does not always mean efficacy in real clinical practice. The actual benefit of the clinical outcome, the satisfaction of physicians, and the cost effectiveness beyond the academic performance must be proven by sophisticated investigation. Finally, the acquisition of reasonable regulations from responsible authorities and a reimbursement policy are essential for integrating AI technology in the current healthcare environment. Moreover, AI is not perfect. That's why "Augmented Intelligence" emerged emphasizing the fact that AI is designed to improve or enhance human intelligence rather than replace it. Although the aim of applying AI in medical practice is to improve the workflow with enhanced precision and to reduce the number of unintentional errors, established models with inaccuracy or exaggerated performance are likely to cause ethical issues owing to misdiagnosis or misclassification. Moreover, we do not know the impact of AI application on the doctor-patient relationship, which is an essential part of healthcare utilization and the practice of medicine. Therefore, ethical principles relevant to AI model development should be established in the current period when AI research begins to increase.

## CONCLUSION

Since AI was introduced in the 1950s, it has been persistently challenged in terms of statistical or image analyses in the field of gastroenterology. Recent evaluation of big data and computer science enabled the dramatic development of AI technology, particularly DL, which showed promising potential. Now, there is no doubt that the implementation of AI in the gastroenterology field will progress in various healthcare services. To utilize AI wisely, physicians should make great effort to understand its feasibility and ameliorate the drawbacks through further investigation.



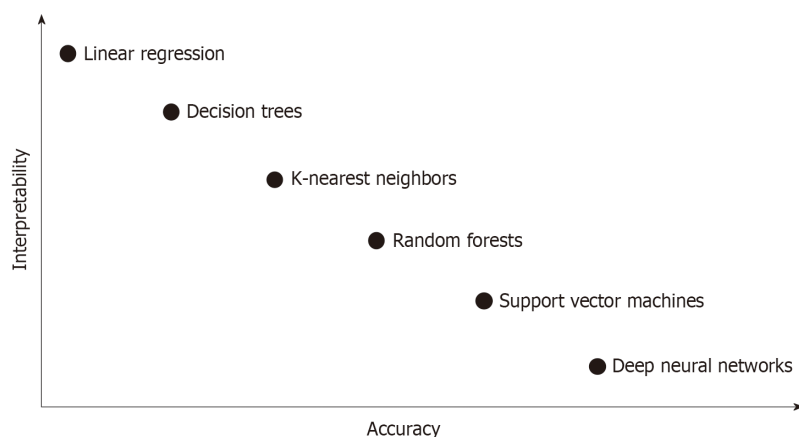


Figure 2 Interpretability-accuracy tradeoff in classification algorithms of machine learning.

## REFERENCES

- 1 Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med* 2019; **25**: 44-56 [PMID: 30617339 DOI: 10.1038/s41591-018-0300-7]
- 2 Turing AM. Computing machinery and intelligence. Oxford: Mind 1950; 433-460
- 3 Russell S, Norvig P. Artificial intelligence: a modern approach. 3<sup>rd</sup> ed. Harlow: Pearson Education 2009; 1-3
- 4 Kevin PM. Machine learning: a probabilistic perspective. 1<sup>st</sup> ed. Cambridge: The MIT Press 2012; 1-2
- 5 Bishop B. CM: Pattern recognition and machine learning. *J Electron Imaging* 2006; **16**: 140-155
- 6 Lee JG, Jun S, Cho YW, Lee H, Kim GB, Seo JB, Kim N. Deep Learning in Medical Imaging: General Overview. *Korean J Radiol* 2017; **18**: 570-584 [PMID: 28670152 DOI: 10.3348/kjr.2017.18.4.570]
- 7 Boser BE, Guyon IM, Vapnik VN. A training algorithm for optimal margin classifiers. Proceedings of the fifth annual workshop on Computational learning theory; 1992 July 1, Pittsburgh, US. New York: ACM 1992; 144-152
- 8 Hastie T, Tibshirani R, Friedman J. The elements of Statistical Learning. New York: Springer 2001;
- 9 McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity. *Bull Math Biophys* 1943; **5**: 115-133
- 10 ROSENBLATT F. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychol Rev* 1958; **65**: 386-408 [PMID: 13602029]
- 11 Pace F, Buscema M, Dominici P, Intraligi M, Baldi F, Cestari R, Passaretti S, Bianchi Porro G, Grossi E. Artificial neural networks are able to recognize gastro-oesophageal reflux disease patients solely on the basis of clinical data. *Eur J Gastroenterol Hepatol* 2005; **17**: 605-610 [PMID: 15879721]
- 12 Lahner E, Grossi E, Intraligi M, Buscema M, Corleto VD, Delle Fave G, Annibale B. Possible contribution of artificial neural networks and linear discriminant analysis in recognition of patients with suspected atrophic body gastritis. *World J Gastroenterol* 2005; **11**: 5867-5873 [PMID: 16270400]
- 13 Pofahl WE, Walczak SM, Rhone E, Izenberg SD. Use of an artificial neural network to predict length of stay in acute pancreatitis. *Am Surg* 1998; **64**: 868-872 [PMID: 9731816]
- 14 Das A, Ben-Menachem T, Cooper GS, Chak A, Sivak MV, Gonet JA, Wong RC. Prediction of outcome in acute lower-gastrointestinal haemorrhage based on an artificial neural network: internal and external validation of a predictive model. *Lancet* 2003; **362**: 1261-1266 [PMID: 14575969 DOI: 10.1016/S0140-6736(03)14568-0]
- 15 Sato F, Shimada Y, Selaru FM, Shibata D, Maeda M, Watanabe G, Mori Y, Stass SA, Imamura M, Meltzer SJ. Prediction of survival in patients with esophageal carcinoma using artificial neural networks. *Cancer* 2005; **103**: 1596-1605 [PMID: 15751017 DOI: 10.1002/ncr.20938]
- 16 Rotondano G, Cipolletta L, Grossi E, Koch M, Intraligi M, Buscema M, Marmo R; Italian Registry on Upper Gastrointestinal Bleeding (Progetto Nazionale Emorragie Digestive). Artificial neural networks accurately predict mortality in patients with nonvariceal upper GI bleeding. *Gastrointest Endosc* 2011; **73**: 218-226, 226.e1-226.e2 [PMID: 21295635 DOI: 10.1016/j.gie.2010.10.006]
- 17 Takayama T, Okamoto S, Hisamatsu T, Naganuma M, Matsuoka K, Mizuno S, Bessho R, Hibi T, Kanai T. Computer-Aided Prediction of Long-Term Prognosis of Patients with Ulcerative Colitis after Cytoapheresis Therapy. *PLoS One* 2015; **10**: e0131197 [PMID: 26111148 DOI: 10.1371/journal.pone.0131197]
- 18 Hardalaç F, Başaranoglu M, Yüksel M, Kutbay U, Kaplan M, Özderin Özın Y, Kılıç ZM, Demirbağ AE, Coşkun O, Aksoy A, Gangarapu V, Örmeci N, Kayaçetin E. The rate of mucosal healing by azathioprine therapy and prediction by artificial systems. *Turk J Gastroenterol* 2015; **26**: 315-321 [PMID: 26039001 DOI: 10.5152/tjg.2015.0199]
- 19 Peng JC, Ran ZH, Shen J. Seasonal variation in onset and relapse of IBD and a model to predict the frequency of onset, relapse, and severity of IBD based on artificial neural network. *Int J Colorectal Dis* 2015; **30**: 1267-1273 [PMID: 25976931 DOI: 10.1007/s00384-015-2250-6]
- 20 Ichimasa K, Kudo SE, Mori Y, Misawa M, Matsudaira S, Kouyama Y, Baba T, Hidaka E, Wakamura K, Hayashi T, Kudo T, Ishigaki T, Yagawa Y, Nakamura H, Takeda K, Haji A, Hamatani S, Mori K, Ishida F, Miyachi H. Artificial intelligence may help in predicting the need for additional surgery after endoscopic resection of T1 colorectal cancer. *Endoscopy* 2018; **50**: 230-240 [PMID: 29272905 DOI: 10.1055/s-0043-122385]
- 21 Yang HX, Feng W, Wei JC, Zeng TS, Li ZD, Zhang LJ, Lin P, Luo RZ, He JH, Fu JH. Support vector machine-based nomogram predicts postoperative distant metastasis for patients with oesophageal

- squamous cell carcinoma. *Br J Cancer* 2013; **109**: 1109-1116 [PMID: [23942069](#) DOI: [10.1038/bjc.2013.379](#)]
- 22 **Takiyama H**, Ozawa T, Ishihara S, Fujishiro M, Shichijo S, Nomura S, Miura M, Tada T. Automatic anatomical classification of esophagogastrroduodenoscopy images using deep convolutional neural networks. *Sci Rep* 2018; **8**: 7497 [PMID: [29760397](#) DOI: [10.1038/s41598-018-25842-6](#)]
  - 23 **van der Sommen F**, Zinger S, Curvers WL, Bisschops R, Pech O, Weusten BL, Bergman JJ, de With PH, Schoon EJ. Computer-aided detection of early neoplastic lesions in Barrett's esophagus. *Endoscopy* 2016; **48**: 617-624 [PMID: [27100718](#) DOI: [10.1055/s-0042-105284](#)]
  - 24 **Swager AF**, van der Sommen F, Klomp SR, Zinger S, Meijer SL, Schoon EJ, Bergman JJGHM, de With PH, Curvers WL. Computer-aided detection of early Barrett's neoplasia using volumetric laser endomicroscopy. *Gastrointest Endosc* 2017; **86**: 839-846 [PMID: [28322771](#) DOI: [10.1016/j.gie.2017.03.011](#)]
  - 25 **Kodashima S**, Fujishiro M, Takubo K, Kammori M, Nomura S, Kakushima N, Muraki Y, Goto O, Ono S, Kaminishi M, Omata M. Ex vivo pilot study using computed analysis of endo-cytoscopic images to differentiate normal and malignant squamous cell epithelia in the oesophagus. *Dig Liver Dis* 2007; **39**: 762-766 [PMID: [17611178](#) DOI: [10.1016/j.dld.2007.03.004](#)]
  - 26 **Shin D**, Protano MA, Polydorides AD, Dawsey SM, Pierce MC, Kim MK, Schwarz RA, Quang T, Parikh N, Bhutani MS, Zhang F, Wang G, Xue L, Wang X, Xu H, Anandasabapathy S, Richards-Kortum RR. Quantitative analysis of high-resolution microendoscopic images for diagnosis of esophageal squamous cell carcinoma. *Clin Gastroenterol Hepatol* 2015; **13**: 272-279.e2 [PMID: [25066838](#) DOI: [10.1016/j.cgh.2014.07.030](#)]
  - 27 **Quang T**, Schwarz RA, Dawsey SM, Tan MC, Patel K, Yu X, Wang G, Zhang F, Xu H, Anandasabapathy S, Richards-Kortum R. A tablet-interfaced high-resolution microendoscope with automated image interpretation for real-time evaluation of esophageal squamous cell neoplasia. *Gastrointest Endosc* 2016; **84**: 834-841 [PMID: [27036635](#) DOI: [10.1016/j.gie.2016.03.1472](#)]
  - 28 **Horie Y**, Yoshio T, Aoyama K, Yoshimizu S, Horiuchi Y, Ishiyama A, Hirasawa T, Tsuchida T, Ozawa T, Ishihara S, Kumagai Y, Fujishiro M, Maetani I, Fujisaki J, Tada T. Diagnostic outcomes of esophageal cancer by artificial intelligence using convolutional neural networks. *Gastrointest Endosc* 2019; **89**: 25-32 [PMID: [30120958](#) DOI: [10.1016/j.gie.2018.07.037](#)]
  - 29 **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 [PMID: [15243882](#) DOI: [10.1055/s-2004-814519](#)]
  - 30 **Shichijo S**, Nomura S, Aoyama K, Nishikawa Y, Miura M, Shinagawa T, Takiyama H, Tanimoto T, Ishihara S, Matsuo K, Tada T. Application of Convolutional Neural Networks in the Diagnosis of *Helicobacter pylori* Infection Based on Endoscopic Images. *EBioMedicine* 2017; **25**: 106-111 [PMID: [29056541](#) DOI: [10.1016/j.ebiom.2017.10.014](#)]
  - 31 **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 [PMID: [29399610](#) DOI: [10.1055/s-0043-120830](#)]
  - 32 **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 [PMID: [29991891](#) DOI: [10.20524/aog.2018.0269](#)]
  - 33 **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 [PMID: [22083334](#) DOI: [10.1007/s00464-011-2036-z](#)]
  - 34 **Hirasawa T**, Aoyama K, Tanimoto T, Ishihara S, Shichijo S, Ozawa T, Ohnishi T, Fujishiro M, Matsuo K, Fujisaki J, Tada T. Application of artificial intelligence using a convolutional neural network for detecting gastric cancer in endoscopic images. *Gastric Cancer* 2018; **21**: 653-660 [PMID: [29335825](#) DOI: [10.1007/s10120-018-0793-2](#)]
  - 35 **Zhu Y**, Wang QC, Xu MD, Zhang Z, Cheng J, Zhong YS, Zhang YQ, Chen WF, Yao LQ, Zhou PH, Li QL. 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 [PMID: [30452913](#) DOI: [10.1016/j.gie.2018.11.011](#)]
  - 36 **Kanesaka T**, Lee TC, Uedo N, Lin KP, Chen HZ, Lee JY, Wang HP, Chang HT. Computer-aided diagnosis for identifying and delineating early gastric cancers in magnifying narrow-band imaging. *Gastrointest Endosc* 2018; **87**: 1339-1344 [PMID: [29225083](#) DOI: [10.1016/j.gie.2017.11.029](#)]
  - 37 **Gatos I**, Tsantis S, Spiliopoulos S, Karnabatidis D, Theotokas I, Zoumpoulis P, Loupas T, Hazle JD, Kagadis GC. A Machine-Learning Algorithm Toward Color Analysis for Chronic Liver Disease Classification, Employing Ultrasound Shear Wave Elastography. *Ultrasound Med Biol* 2017; **43**: 1797-1810 [PMID: [28634041](#) DOI: [10.1016/j.ultrasmedbio.2017.05.002](#)]
  - 38 **Kuppili V**, Biswas M, Sreekumar A, Suri HS, Saba L, Edla DR, Marinho RT, Sanches JM, Suri JS. Extreme Learning Machine Framework for Risk Stratification of Fatty Liver Disease Using Ultrasound Tissue Characterization. *J Med Syst* 2017; **41**: 152 [PMID: [28836045](#) DOI: [10.1007/s10916-017-0797-1](#)]
  - 39 **Liu X**, Song JL, Wang SH, Zhao JW, Chen YQ. Learning to Diagnose Cirrhosis with Liver Capsule Guided Ultrasound Image Classification. *Sensors (Basel)* 2017; **17** [PMID: [28098774](#) DOI: [10.3390/s17010149](#)]
  - 40 **Fernández-Esparrach G**, Bernal J, López-Cerón M, Córdova H, Sánchez-Montes C, Rodríguez de Miguel C, Sánchez FJ. Exploring the clinical potential of an automatic colonic polyp detection method based on the creation of energy maps. *Endoscopy* 2016; **48**: 837-842 [PMID: [27285900](#) DOI: [10.1055/s-0042-108434](#)]
  - 41 **Misawa M**, Kudo SE, Mori Y, Cho T, Kataoka S, Yamauchi A, Ogawa Y, Maeda Y, Takeda K, Ichimasa K, Nakamura H, Yagawa Y, Toyoshima N, Ogata N, Kudo T, Hisayuki T, Hayashi T, Wakamura K, Baba T, Ishida F, Itoh H, Roth H, Oda M, Mori K. Artificial Intelligence-Assisted Polyp Detection for Colonoscopy: Initial Experience. *Gastroenterology* 2018; **154**: 2027-2029.e3 [PMID: [29653147](#) DOI: [10.1053/j.gastro.2018.04.003](#)]
  - 42 **Urban G**, Tripathi P, Alkayali T, Mittal M, Jalali F, Karnes W, Baldi P. Deep Learning Localizes and Identifies Polyps in Real Time With 96% Accuracy in Screening Colonoscopy. *Gastroenterology* 2018; **155**: 1069-1078.e8 [PMID: [29928897](#) DOI: [10.1053/j.gastro.2018.06.037](#)]
  - 43 **Wang H**, Liang Z, Li LC, Han H, Song B, Pickhardt PJ, Barish MA, Lascarides CE. An adaptive paradigm for computer-aided detection of colonic polyps. *Phys Med Biol* 2015; **60**: 7207-7228 [PMID: [26348125](#) DOI: [10.1088/0031-9155/60/18/7207](#)]

- 44 **K G**, c R. Automatic Colorectal Polyp Detection in Colonoscopy Video Frames. *Asian Pac J Cancer Prev* 2016; **17**: 4869-4873 [PMID: [28030914](#) DOI: [10.22034/APJCP.2016.17.11.4869](#)]
- 45 **Billah M**, Waheed S, Rahman MM. An Automatic Gastrointestinal Polyp Detection System in Video Endoscopy Using Fusion of Color Wavelet and Convolutional Neural Network Features. *Int J Biomed Imaging* 2017; **2017**: 9545920 [PMID: [28894460](#) DOI: [10.1155/2017/9545920](#)]
- 46 **Klare P**, Sander C, Prinzen M, Haller B, Nowack S, Abdelhazef M, Poszler A, Brown H, Wilhelm D, Schmid RM, von Delius S, Wittenberg T. Automated polyp detection in the colorectum: a prospective study (with videos). *Gastrointest Endosc* 2019; **89**: 576-582.e1 [PMID: [30342029](#) DOI: [10.1016/j.gie.2018.09.042](#)]
- 47 **Wang P**, Xiao X, Brown JR, Berzin TM, Tu M, Xiong F, Hu X, Liu P, Song Y, Zhang D, Yang X. Development and validation of a deep-learning algorithm for the detection of polyps during colonoscopy. *Nat Biomed Eng* 2018; **2**: 741 [DOI: [10.1038/s41551-018-0301-3](#)]
- 48 **Tischendorf JJ**, Gross S, Winograd R, Hecker H, Auer R, Behrens A, Trautwein C, Aach T, Stehle T. Computer-aided classification of colorectal polyps based on vascular patterns: a pilot study. *Endoscopy* 2010; **42**: 203-207 [PMID: [20101564](#) DOI: [10.1055/s-0029-1243861](#)]
- 49 **Gross S**, Trautwein C, Behrens A, Winograd R, Palm S, Lutz HH, Schirin-Sokhan R, Hecker H, Aach T, Tischendorf JJ. Computer-based classification of small colorectal polyps by using narrow-band imaging with optical magnification. *Gastrointest Endosc* 2011; **74**: 1354-1359 [PMID: [22000791](#) DOI: [10.1016/j.gie.2011.08.001](#)]
- 50 **Takemura Y**, Yoshida S, Tanaka S, Onji K, Oka S, Tamaki T, Kaneda K, Yoshihara M, Chayama K. Quantitative analysis and development of a computer-aided system for identification of regular pit patterns of colorectal lesions. *Gastrointest Endosc* 2010; **72**: 1047-1051 [PMID: [21034905](#) DOI: [10.1016/j.gie.2010.07.037](#)]
- 51 **Takemura Y**, Yoshida S, Tanaka S, Kawase R, Onji K, Oka S, Tamaki T, Raytchev B, Kaneda K, Yoshihara M, Chayama K. Computer-aided system for predicting the histology of colorectal tumors by using narrow-band imaging magnifying colonoscopy (with video). *Gastrointest Endosc* 2012; **75**: 179-185 [PMID: [22196816](#) DOI: [10.1016/j.gie.2011.08.051](#)]
- 52 **Kominami Y**, Yoshida S, Tanaka S, Sanomura Y, Hirakawa T, Raytchev B, Tamaki T, Koide T, Kaneda K, Chayama K. Computer-aided diagnosis of colorectal polyp histology by using a real-time image recognition system and narrow-band imaging magnifying colonoscopy. *Gastrointest Endosc* 2016; **83**: 643-649 [PMID: [26264431](#) DOI: [10.1016/j.gie.2015.08.004](#)]
- 53 **Byrne MF**, Chapados N, Soudan F, Oertel C, Linares Pérez M, Kelly R, Iqbal N, Chandelier F, Rex DK. Real-time differentiation of adenomatous and hyperplastic diminutive colorectal polyps during analysis of unaltered videos of standard colonoscopy using a deep learning model. *Gut* 2019; **68**: 94-100 [PMID: [29066576](#) DOI: [10.1136/gutjnl-2017-314547](#)]
- 54 **Chen PJ**, Lin MC, Lai MJ, Lin JC, Lu HH, Tseng VS. Accurate Classification of Diminutive Colorectal Polyps Using Computer-Aided Analysis. *Gastroenterology* 2018; **154**: 568-575 [PMID: [29042219](#) DOI: [10.1053/j.gastro.2017.10.010](#)]
- 55 **Komeda Y**, Handa H, Watanabe T, Nomura T, Kitahashi M, Sakurai T, Okamoto A, Minami T, Kono M, Arizumi T, Takenaka M, Hagiwara S, Matsui S, Nishida N, Kashida H, Kudo M. Computer-Aided Diagnosis Based on Convolutional Neural Network System for Colorectal Polyp Classification: Preliminary Experience. *Oncology* 2017; **93** Suppl 1: 30-34 [PMID: [29258081](#) DOI: [10.1159/000481227](#)]
- 56 **Mori Y**, Kudo SE, Wakamura K, Misawa M, Ogawa Y, Kutsukawa M, Kudo T, Hayashi T, Miyachi H, Ishida F, Inoue H. Novel computer-aided diagnostic system for colorectal lesions by using endocytoscopy (with videos). *Gastrointest Endosc* 2015; **81**: 621-629 [PMID: [25440671](#) DOI: [10.1016/j.gie.2014.09.008](#)]
- 57 **Mori Y**, Kudo SE, Chiu PW, Singh R, Misawa M, Wakamura K, Kudo T, Hayashi T, Katagiri A, Miyachi H, Ishida F, Maeda Y, Inoue H, Nimura Y, Oda M, Mori K. Impact of an automated system for endocytoscopic diagnosis of small colorectal lesions: an international web-based study. *Endoscopy* 2016; **48**: 1110-1118 [PMID: [27494455](#) DOI: [10.1055/s-0042-113609](#)]
- 58 **Misawa M**, Kudo SE, Mori Y, Nakamura H, Kataoka S, Maeda Y, Kudo T, Hayashi T, Wakamura K, Miyachi H, Katagiri A, Baba T, Ishida F, Inoue H, Nimura Y, Mori K. Characterization of Colorectal Lesions Using a Computer-Aided Diagnostic System for Narrow-Band Imaging Endocytoscopy. *Gastroenterology* 2016; **150**: 1531-1532.e3 [PMID: [27072671](#) DOI: [10.1053/j.gastro.2016.04.004](#)]
- 59 **Mori Y**, Kudo SE, Misawa M, Saito Y, Ikematsu H, Hotta K, Ohtsuka K, Urushibara F, Kataoka S, Ogawa Y, Maeda Y, Takeda K, Nakamura H, Ichimasa K, Kudo T, Hayashi T, Wakamura K, Ishida F, Inoue H, Itoh H, Oda M, Mori K. Real-Time Use of Artificial Intelligence in Identification of Diminutive Polyps During Colonoscopy: A Prospective Study. *Ann Intern Med* 2018; **169**: 357-366 [PMID: [30105375](#) DOI: [10.7326/M18-0249](#)]
- 60 **Takeda K**, Kudo SE, Mori Y, Misawa M, Kudo T, Wakamura K, Katagiri A, Baba T, Hidaka E, Ishida F, Inoue H, Oda M, Mori K. Accuracy of diagnosing invasive colorectal cancer using computer-aided endocytoscopy. *Endoscopy* 2017; **49**: 798-802 [PMID: [28472832](#) DOI: [10.1055/s-0043-105486](#)]
- 61 **Maeda Y**, Kudo SE, Mori Y, Misawa M, Ogata N, Sasanuma S, Wakamura K, Oda M, Mori K, Ohtsuka K. Fully automated diagnostic system with artificial intelligence using endocytoscopy to identify the presence of histologic inflammation associated with ulcerative colitis (with video). *Gastrointest Endosc* 2019; **89**: 408-415 [PMID: [30268542](#) DOI: [10.1016/j.gie.2018.09.024](#)]
- 62 **Leenhardt R**, Vasseur P, Li C, Saurin JC, Rahmi G, Cholet F, Becq A, Marteau P, Histace A, Dray X; CAD-CAP Database Working Group. A neural network algorithm for detection of GI angiectasia during small-bowel capsule endoscopy. *Gastrointest Endosc* 2019; **89**: 189-194 [PMID: [30017868](#) DOI: [10.1016/j.gie.2018.06.036](#)]
- 63 **Zhou T**, Han G, Li BN, Lin Z, Ciaccio EJ, Green PH, Qin J. Quantitative analysis of patients with celiac disease by video capsule endoscopy: A deep learning method. *Comput Biol Med* 2017; **85**: 1-6 [PMID: [28412572](#) DOI: [10.1016/j.compbiomed.2017.03.031](#)]
- 64 **He JY**, Wu X, Jiang YG, Peng Q, Jain R. Hookworm Detection in Wireless Capsule Endoscopy Images With Deep Learning. *IEEE Trans Image Process* 2018; **27**: 2379-2392 [PMID: [29470172](#) DOI: [10.1109/TIP.2018.2801119](#)]
- 65 **Seguí S**, Drozdal M, Pascual G, Radeva P, Malagelada C, Azpiroz F, Vitrià J. Generic feature learning for wireless capsule endoscopy analysis. *Comput Biol Med* 2016; **79**: 163-172 [PMID: [27810622](#) DOI: [10.1016/j.compbiomed.2016.10.011](#)]
- 66 **England JR**, Cheng PM. Artificial Intelligence for Medical Image Analysis: A Guide for Authors and Reviewers. *AJR Am J Roentgenol* 2019; **212**: 513-519 [PMID: [30557049](#) DOI: [10.2214/AJR.18.20490](#)]
- 67 **Park SH**, Han K. Methodologic Guide for Evaluating Clinical Performance and Effect of Artificial

- Intelligence Technology for Medical Diagnosis and Prediction. *Radiology* 2018; **286**: 800-809 [PMID: 29309734 DOI: 10.1148/radiol.2017171920]
- 68 **Park SH.** Artificial intelligence in medicine: beginner's guide. *J Korean Soc Radiol* 2018; **78**: 301-308 [DOI: 10.3348/jksr.2018.78.5.301]
- 69 **Bae HJ,** Kim CW, Kim N, Park B, Kim N, Seo JB, Lee SM. A Perlin Noise-Based Augmentation Strategy for Deep Learning with Small Data Samples of HRCT Images. *Sci Rep* 2018; **8**: 17687 [PMID: 30523268 DOI: 10.1038/s41598-018-36047-2]
- 70 **Tavanaei A,** Ghodrati M, Kheradpisheh SR, Masquelier T, Maida A. Deep learning in spiking neural networks. *Neural Netw* 2019; **111**: 47-63 [PMID: 30682710 DOI: 10.1016/j.neunet.2018.12.002]



Published By Baishideng Publishing Group Inc  
7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA  
Telephone: +1-925-2238242  
Fax: +1-925-2238243  
E-mail: [bpgoffice@wjgnet.com](mailto:bpgoffice@wjgnet.com)  
Help Desk: <http://www.f6publishing.com/helpdesk>  
<http://www.wjgnet.com>

