ISSN 2689-7164 (online)

Artificial Intelligence in *Gastrointestinal Endoscopy*

Artif Intell Gastrointest Endosc 2021 April 28; 2(2): 12-49

Published by Baishideng Publishing Group Inc

AIGE

Artificial Intelligence in Gastrointestinal Endoscopy

Contents

Bimonthly Volume 2 Number 2 April 28, 2021

MINIREVIEWS

12 Application of deep learning in image recognition and diagnosis of gastric cancer

Li Y, Zhou D, Liu TT, Shen XZ

25 Application of artificial intelligence to endoscopy on common gastrointestinal benign diseases

Yang H, Hu B

36 Colonoscopy and artificial intelligence: Bridging the gap or a gap needing to be bridged?

Li JW, Ang TL

Artificial Intelligence in Gastrointestinal Endoscopy

Bimonthly Volume 2 Number 2 April 28, 2021

Contents

ABOUT COVER

AIMS AND SCOPE

The primary aim of *Artificial Intelligence in Gastrointestinal Endoscopy* (*AIGE, Artif Intell Gastrointest Endosc*) is to provide scholars and readers from various fields of artificial intelligence in gastrointestinal endoscopy with a platform to publish high-quality basic and clinical research articles and communicate their research findings online.

AIGE mainly publishes articles reporting research results obtained in the field of artificial intelligence in gastrointestinal endoscopy and covering a wide range of topics, including artificial intelligence in capsule endoscopy, colonoscopy, double-balloon enteroscopy, duodenoscopy, endoscopic retrograde cholangiopancreatography, endosonography, esophagoscopy, gastrointestinal endoscopy, gastroscopy, laparoscopy, natural orifice endoscopic surgery, proctoscopy, and sigmoidoscopy.

INDEXING/ABSTRACTING

There is currently no indexing.

RESPONSIBLE EDITORS FOR THIS ISSUE

NAME OF JOURNAL
Artificial Intelligence in Gastrointestinal Endoscopy

ISSN
ISSN 2689-7164 (online)

LAUNCH DATE
July 28, 2020

FREQUENCY
Bimonthly

EDITORS-IN-CHIEF
Krish Ragunath, Fatih Altintoprak, Sahin Coban

EDITORIAL BOARD MEMBERS
https://www.wjgnet.com/2689-7164/editorialboard.htm

PUBLICATION DATE
April 28, 2021

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 ETHICS
https://www.wjgnet.com/bpg/GerInfo/288

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

Artificial Intelligence in Gastrointestinal Endoscopy

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

Artif Intell Gastrointest Endosc 2021 April 28; 2(2): 12-24

DOI: 10.37126/aige.v2.i2.12

ISSN 2689-7164 (online)

MINIREVIEWS

Application of deep learning in image recognition and diagnosis of gastric cancer

Yu Li, Da Zhou, Tao-Tao Liu, Xi-Zhong Shen

ORCID number: Yu Li 0000-0002-4685-3418; Da Zhou 0000-0001-8838-1351; Tao-Tao Liu 0000-0002-4623-9012; Xi-Zhong Shen 0000-0003-3748-0709.

Author contributions: TT Liu and D Zhou contributed equally to conceptual development and supervision; Y Li and D Zhou contributed to the data collection and manuscript; XZ Shen supervised the paper; All authors have read and approved the final manuscript.

Supported by National Natural Science Foundation of China, No. 81800510; Shanghai Sailing Program, No. 18YF1415900.

Conflict-of-interest statement: The authors report no conflicts of interest in this work.

Yu Li, Da Zhou, Tao-Tao Liu, Xi-Zhong Shen, Department of Gastroenterology and Hepatology, Zhongshan Hospital Affiliated to Fudan University, Shanghai 200032, China

Corresponding author: Da Zhou, PhD, Doctor, Research Fellow, Department of Gastroenterology and Hepatology, Zhongshan Hospital Affiliated to Fudan University, 180 Feng Lin road, Shanghai 200032, China. mubing2007@foxmail.com

Abstract

In recent years, artificial intelligence has been extensively applied in the diagnosis of gastric cancer based on medical imaging. In particular, using deep learning as one of the mainstream approaches in image processing has made remarkable progress. In this paper, we also provide a comprehensive literature survey using four electronic databases, PubMed, EMBASE, Web of Science, and Cochrane. The literature search is performed until November 2020. This article provides a summary of the existing algorithm of image recognition, reviews the available datasets used in gastric cancer diagnosis and the current trends in applications of deep learning theory in image recognition of gastric cancer. covers the theory of deep learning on endoscopic image recognition. We further evaluate the advantages and disadvantages of the current algorithms and summarize the characteristics of the existing image datasets, then combined with the latest progress in deep learning theory, and propose suggestions on the applications of optimization algorithms. Based on the existing research and application, the label, quantity, size, resolutions, and other aspects of the image dataset are also discussed. The future developments of this field are analyzed from two perspectives including algorithm optimization and data support, aiming to improve the diagnosis accuracy and reduce the risk of misdiagnosis.

Key Words: Endoscope; Artificial intelligence; Algorithm optimization; Data support

Core Tip: Gastric cancer is a life-threatening disease with a high mortality rate. With the development of deep learning in the image processing of gastrointestinal endoscope, the efficiency and accuracy of gastric cancer diagnosis through imaging technology have been greatly improved. At present, there is no comprehensive

Manuscript source: Invited manuscript

Specialty type: Gastroenterology and Hepatology

Country/Territory of origin: China

Peer-review report's scientific quality classification

Grade A (Excellent): A
Grade B (Very good): B
Grade C (Good): 0
Grade D (Fair): D
Grade E (Poor): 0

Received: February 15, 2021
Peer-review started: February 15, 2021
First decision: March 16, 2021
Revised: March 30, 2021
Accepted: April 20, 2021
Article in press: April 20, 2021
Published online: April 28, 2021

P-Reviewer: Nayyar A, Taira K, Tanabe S
S-Editor: Wang JL
L-Editor: A
P-Editor: Wang LL

summary on the graphic recognition method for gastric cancer based on deep learning. In this review, some gastric cancer image databases and mainstream gastric cancer recognition models were summarized to make a prospect for the application of deep learning in this field.

Citation: Li Y, Zhou D, Liu TT, Shen XZ. Application of deep learning in image recognition and diagnosis of gastric cancer. *Artif Intell Gastrointest Endosc* 2021; 2(2): 12-24
URL: https://www.wjgnet.com/2689-7164/full/v2/i2/12.htm
DOI: https://dx.doi.org/10.37126/aige.v2.i2.12

INTRODUCTION

Gastric cancer is a life-threatening disease with a high mortality rate[1]. Globally, more than 900000 individuals develop gastric cancer each year out of which more than 700000 Lose their lives. Gastric cancer is second only to lung cancer in terms of mortality[2]. Unlike the developing countries, the number of diagnosed cases and the mortality rate of this cancer are declining in the developed countries such as those in the EU and North America[3,4].

Around 50% of the world's gastric cancer cases are diagnosed in Southeast Asia[5]. In China, gastric cancer is also second to lung cancer in terms of the number of annual cases, for instance, 424000 new patients are annually diagnosed with gastric cancer, accounting for more than 40% of the global total, out of which 392000 Lose their lives ranking the fifth and the sixth worldwide in annual morbidity and mortality, respectively[6].

The diagnosis of gastric cancer mainly relies on clinical manifestation, pathological images and medical imaging[7]. Compared with other methods such as pathological diagnosis, medical imaging provides a simple non-invasive and reliable method for the diagnosis of gastric cancer which is more accessible and efficient, easier to operate and has almost no side effects for the patients[8].

Doctors make a judgment based on medical imaging which mainly depend on their experience from similar cases, hence, occasional misdiagnosis is inevitable[9,10]. With the rapid development of computer technology and artificial intelligence, deep learning techniques are extremely effective in various branches of image processing and have been used in medical imaging to improve cancer diagnosis[11-13]. Danaee *et al*[14] established a deep learning model for colorectal cancer image recognition, the results showed that the deep learning method can achieve more effective information and is far more efficient than the way of manual extraction. Burke *et al*[15] found that deep learning could classify and predict mutations of NSCLC based on histopathological images, and the recognition efficiency of deep learning was much higher than that of manual recognition. Muhammad Owais *et al*[16] proposed a deep learning model to classify a variety of gastrointestinal diseases by recognizing endoscopic videos. This model can simultaneously extract spatiotemporal features to achieve better classification performance. Experimental results of the proposed models showed superior performance to the latest technology and indicated its potential in clinical application[16].

Endoscopic images are mostly used in gastric cancer diagnosis[17]. Endoscopic images contain a lot of useful structural information which can be used for deep learning algorithm, the algorithm can carry out purposeful image recognition[18]. Most of the image recognition based on gastric cancer diagnosis methods adopt supervised deep learning algorithms, mainly because the monitored network in supervised learning makes full use of the labeled sample data in the training and can obtain more accurate segmentation results[19].

In fact, the purpose of medical image recognition is to identify the tumor and we call this process image segmentation[20-22]. Accurate segmentation of tumor images is an important step in diagnosis, surgical planning and postoperative evaluation[23,24]. Endoscopic images segmentation can provide more comprehensive information for the diagnosis and treatment of gastric cancer, alleviate the doctor's heavy work for reading film and improve the accuracy of diagnosis[25]. However, due to the variety and complexity of gastric tumor types, segmentation has become an important and difficult problem in computer-aided diagnosis. Compared with the traditional

segmentation methods, the deep learning segmentation method of gastric tumor image has achieved obvious improved performance and rapid development[26,27].

As mentioned above, the deep learning method based on supervised learning can fully mine the effective information of existing data. However, when the amount of existing data cannot meet the requirements of model training, it is necessary to find ways to increase the data scale[28]. The deep learning based on unsupervised learning can generate samples, which are similar to the existing samples in dimension and structure, but not identical. At present, relevant research results have been obtained[29]. Researchers use semi-supervised and unsupervised image recognition algorithms to generate samples like training samples, to improve the accuracy of gastric cancer tumor recognition and enhance the robustness of the model[30].

In this paper, deep learning-based diagnosis of gastric cancer based on endoscope images is summarized and analyzed. The adopted segmentation networks in the previous works can be divided into three categories: the supervised network, semi-supervised network, and unsupervised network. The basic idea of the recognition method, the basic structure of the network, the experimental results, as well as their advantages and disadvantages are summarized. The performance of typical methods above-mentioned in recognition is compared. Finally, we hope to provide insights and concluding remarks on the development of deep-learning-based diagnosis of gastric cancer.

RELEVANT DATA SETS AND ALGORITHM EVALUATION INDEXES

Relevant datasets

To promote the progress of image recognition and make an objective comparison of available image recognition methods for gastric cancer diagnosis, we investigate the commonly used datasets including the GR-AIDS provided by Medical Image Computing and Computer Assisted Intervention Society as well as those internal datasets.

The GR-AIDS dataset established by Sun Yat-Sen University Cancer Center consists of 1036496 endoscopic images from 84424 individuals. This dataset is used according to the 8:1:1 pattern, the data is randomly selected for training and internal validation datasets for GR-AIDS development as well as for evaluating GR-AIDS performance[31].

Using clinical data collected from Gil Hospital, Jang Hyung Lee *et al*[32] also established a data set containing 200 normal cases, 367 cancer cases, and 220 ulcer cases. The data was divided into training sets of 180, 200, 337 images and test sets of 20, 30, 20 images. To improve the local contrast of the image and enhance the edge definition in each area of the image, histogram equalization was adopted to further enhance the image, the images' size was adjusted to 224 × 224 pixels [32].

Hirasawa *et al*[32] collected 13,584 endoscopic images of gastric cancer to build an image database. To evaluate the diagnostic accuracy, an independent test set of 2296 gastric images was collected from 69 patients with continuous gastric cancer lesions constructed as convolutional neural network (CNN). The image has an in-plane resolution of 512 × 512[33].

Cho *et al*[34] collected 5017 images from 1269 patients, of which 812 images from 212 patients were used as the test data set. An additional 200 images from 200 patients were collected and used for prospective validation. The resolution of the images is 512 × 512. The information for all major databases is shown in Table 1[34].

Introduction of evaluation indexes

To evaluate the effectiveness of each model in diagnosing gastric cancer, the following evaluation indicators are commonly used in the related literature (Table 2): DICE Similarity Coefficient (DICE, 1945), Jaccard Coefficient (Jaccard, 1912), Volumetric Over-lap Error (VOE), and Relative Volume Difference (RVD).

Here we define the following variables: P and N are used for judgment of the model results, T and F evaluation model of the judgment is correct, FP is on behalf of the false-positive cases, FN represents false-negative cases, TP is on behalf of the real example, TN represents true negative cases[38]. A represents the theory of segmentation, results for comparison with the resulting image. B represents the segmentation results[39]. The relationship among them is shown in Figure 1.

DICE coefficient: DICE coefficient also known as the overlap index, is one of the most commonly used indexes for verification of image segmentation. The DICE coefficient

Table 1 Commonly used databases in image recognition of gastric cancer

Database	Time collected	Number of samples	Resolution	Training set	Test set
GR-AIDS[31]	2019	1036496	512 × 512	829197	103650
Jang Hyung Lee[32]	2019	787	224 × 224	717	70
Toshiaki Hirasawa[33]	2018	13584	512 × 512	13584	2496
Bum-Joo Cho[34]	2019	5017	512 × 512	4205	812
Hiroya Ueyama[35]	2020	7874	512 × 512	5574	2300
Lan Li[36]	2020	2088	512 × 512	1747	341
Mads Sylvest Bergholt[37]	2011	1063	512 × 512	850	213

Table 2 Specific concepts of the main evaluation indicators

Index	Description	Usage	Unit
DICE	Repeat rate between the segmentation results and markers	Commonly	%
RMSD	The root mean square of the symmetrical position surface distance between the segmentation results and the markers	Commonly	mm
VOE	The degree of overlap between the segmentation results and the actual segmentation results represents the error rate	Commonly	%
RVD	The difference in volume between the segmentation results and the markers	Rarely	%

DICE: DICE Similarity Coefficient; RMSD: Root-Mean-Square Deviation; VOE: Volumetric Over-lap Error; RVD: Relative Volume Difference.

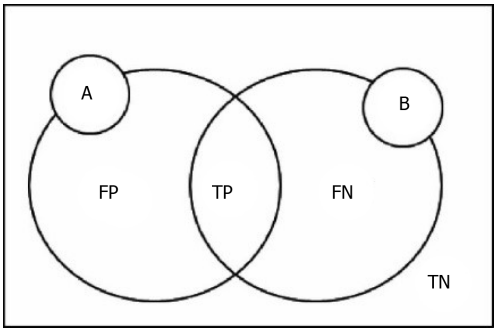

Figure 1 Schematic diagram of each evaluation index relationship. TP: True positive; FP: False-positive; TN: True negative; FN: False negative.

represents the repetition rate between the segmentation results and the markers. The value range of DICE is 0-1, where 0 indicates that the experimental segmentation result significantly deviates from the labeled result, and 1 indicates that the experimental segmentation result completely coincides with the labeled result[40]. DICE coefficient is defined as the following:

$$DICE = (2|A \cap B|)/(|A| + |B|) = (2TP)/(2TP + FP + FN)$$

Jaccard coefficient: Jaccard coefficient represents the similarity and difference between the segmentation result and the standard. The larger the coefficient, the higher the sample similarity. Besides, the Jaccard coefficient and DICE coefficient are correlated[41]. Jaccard coefficient is defined as the following:

$$JAC = (|A \cap B|)/(|A| \cup |B|) = TP/(TP + FP + FN) = DICE/(2 - DICE)$$

VOE: VOE stands for error rate, derived from Jaccard. VOE is represented as %, where 0% indicates complete segmentation. If there is no overlap between the segmentation result and the markers, the VOE is 100%[42]. VOE is defined as the following:

$$VOE = 1 - (|A \cap B|)/(|A| \cup |B|) = 1 - TP/(TP + FP + FN)$$

RVD: RVD represents the noise difference between the segmentation result and the markers. RVD is presented as %, where 0% denotes the same volume between the segmentation result and the markers[42]. The formula is:

$RVD = (|B| - |A|)/|A| = FP/(TP + FN)$

The specific concepts of all indicators are shown in Table 2.

CLASSIFICATION OF THE ALGORITHM

Supervised learning-based diagnosis of gastric cancer

Deep neural networks are often trained based on deep learning algorithms using large labeled datasets (*i.e.*, images in this case)[43]. The network is therefore able to learn how features are related to the target[44]. Since the data is already labeled, this learning method is referred to as supervised learning. Most of the existing studies on diagnosing gastric cancer are based on supervised learning in image recognition tasks[45-47]. This is because the network makes full use of the labeled dataset in the training, hence can obtain more accurate segmentation results.

Recent research works showed that CNN achieves outstanding performance in various image recognition tasks[48,49]. Toshiaki Hirasawa built a CNN-based diagnostic system based on a single-shot Multi-Box detector, Adejub, with a total sensitivity of 92.2% and trained their CNN using 13584 endoscopic images of gastric cancer. The trained CNN correctly called 71 out of 77 cases of gastric cancer, *i.e.*, a total sensitivity of 92.2%, also detected 161 non-cancerous samples as gastric cancer, *i.e.*, a positive predictive value of 30.6%. The CNN also correctly detected 70 of 71 cases of gastric cancer (98.6%) with a diameter of 6 mm or larger, as well as all invasive cancers[33]. Ueyama *et al*[35] also constructed an AI-assisted CNN based computer-aided diagnosis system with narrow band imaging-magnifying endoscopy images.

The above studies show that the CNN-based approach is far more accurate than human in recognition of cancer. This makes us believe that the method based on deep CNN can effectively solve the identification problem of gastric cancer.

However, the issue with the CNN is that only partial features could be extracted[50]. Due to the imbalanced information of gastric cancer image data, extracting the local features does not reflect all the information and might harm the efficiency of the image recognition. To address the problem, Shelhamer *et al*[51] proposed full convolutional neural network (FCN) for image segmentation. This network attempts to recover the category of each pixel from the abstract feature, in other words, instead of image-level classification, the network uses pixel-level classification[51]. This addresses the semantic level image segmentation problem and is the core component of many advanced semantic segmentation models[52,53].

The segmentation method of gastric cancer images based on the FCN network is mainly based on the idea of code-decoding design[54]. In practice, the image is classified at the pixel level and the network is pre-trained with supervision. In this method, the input image can have any arbitrary size and the output of the same size can be generated through effective reasoning and learning[55]. Typical FCN network-based image segmentation architecture for gastric cancer is shown in Figure 2.

The FCN is improved based on the CNN by transforming the last three full connections into three convolutional layers. The success of FCN network is largely attributed to the excellent ability of CNN network to extract hierarchical representation. In the concrete implementation process, the network realizes the segmentation of gastric tumor by down-sampling and up-sampling through convolution-deconvolution operation. The down-sampling path consists of convolution layer and maximum or average pooling layer, which can extract high-level semantic information, but its spatial resolution is often low. The up-sample path consists of convolution and a deconvolution layer (also known as transpose convolution) and uses the output of the down-sample path to predicting the fraction of each class at the pixel level[56,57]. However, the output image of deconvolution operations might be very rough and lost a lot of detail. The skip structure of the FCN network presented in the classified forecast comes from the deep layer (thick) semantic information and information from the appearance of the shallow layer (fine), thus, achieving a more accurate and robust segmentation result. As a deep neural network, FCN has shown good performance in many challenging medical image segmentation tasks, including liver tumor segmentation[58,59].

One of the most important features of the FCN is the use of skip structure. It is used to fuse the feature information of both the high and low layers. Through the cross-layer connection structure, the texture information of the shallow layer and the semantic information of the deep layer of the network are then combined to achieve the precise segmentation task[60,61]. Jang Hyung Lee improved the original FCN framework by applying the pre-trained Inception, Res-Net, and VGG-Net models on

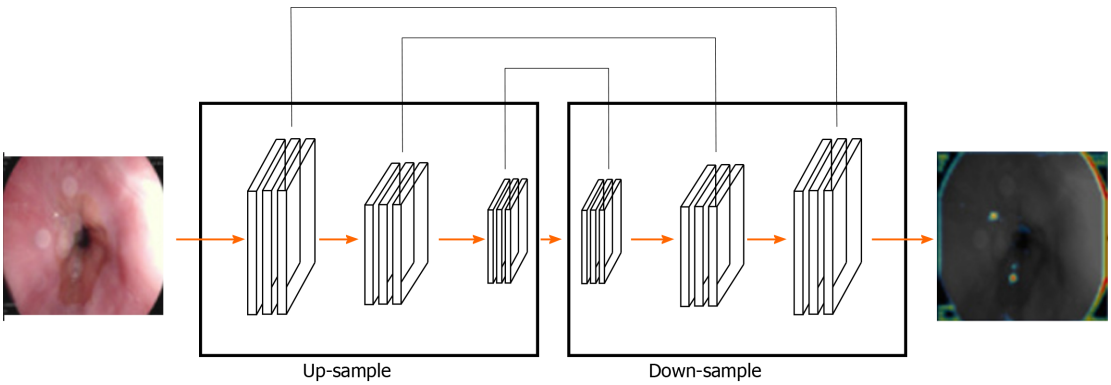

Figure 2 The basic architecture of image segmentation for gastric cancer based on full convolutional network.

ImageNet. The areas under the operating characteristic curves of each receiver are 0.95, 0.97, and 0.85, respectively, hence, Res-Net shows the highest level of performance. Under normal conditions, the classification between normal and ulcer or cancer, is more than 90 percent accurate[32].

The deep network structure leads to the problem of decreased training accuracy[62]. In Sun *et al*[63] the basic form of convolution is replaced with the deformable convolution and Atrous convolution in a specific layer to adapt to the non-rigid characteristics and large receiving fields. The Atrous space pyramid pooling module and the semantic-level embedded network based on encoder/decoder are used for multi-scale segmentation. Besides, they proposed a lightweight decoder to fuse the context information and further used dense up-sampled convolution for the boundary optimization at the end of the decoder. The model achieves 91.60% pixel-level accuracy and 82.65% average degree of the intersection[63].

Cho *et al*[34] established the Inception-ResNET-V2 model, which is an FCN model. In this model, they divided the images into five categories: advanced gastric cancer, early gastric cancer, high atypical hyperplasia, low atypical hyperplasia and non-neoplastic. For the above five categories, the Inception- ResNet-v2 model has a weighted average accuracy of 84.6%. The mean area under the curve of the model for differentiating gastric cancer and neoplasm was 0.877 and 0.927, respectively[34].

The above works show that FCN addresses the issue with the CNN hence can extract the local features. This is why the FCN is considered as the mainstream in gastric cancer image classification methods.

In addition to the application of FCN to address the shortcomings of CNN, researchers also tried other approaches such as fusion of multiple CNN methods to obtain an Ensemble of CNN algorithm to get more accurate classification results. Nguyen *et al*[64] trained three different CNN model architectures, including VGG-based, Inception-based Network and Dense-Net. In their study, the VGG-based network was used as a conventional deep CNN for classification problems, which consists of a linear stack of the convolutional layer. The network-based on Dense-net can be used as a very deep CNN with a short path, which is also helpful to train the network and extract more abstract and effective image features easily. The three models were trained separately, the AVERAGE combination rule is then used to combine the classification results of the three CNN-based Models. The final result was 70.369% of overall classification accuracy, 68.452% of sensitivity and 72.571% of specificity. The overall classification accuracy is higher than that generated by the listed model based on a single CNN[64].

Both the use of a fully convolutional network and the fusion of several CNN algorithms are significantly effective in improving the accuracy of gastric cancer image recognition. They are also effective in addressing the issues with the quality of images in the database. Table 3 shows the performance comparison of gastric cancer image recognition by using CNN, FCN, and Ensemble CNN.

Image recognition based on semi-supervised and unsupervised learning in gastric cancer

Most gastric cancer image recognition methods adopt supervised learning algorithms because the monitored network makes full use of the labeled sample data in the training and can obtain more accurate segmentation results. Nevertheless, there are very few accurately labeled image datasets, hence researchers have carried out studies

Table 3 Comparison of recognition performance of convolutional neural network, full convolutional neural network, and ensemble convolutional neural network models

Methods	DICE/%	VOE/%	RMSD/mm
Toshiaki Hirasawa (CNN)	0.5738	0.5977	6.491
Hiroya Ueyama (CNN)	0.6327	0.5373	7.257
Jang Hyung Lee (FCN)	0.8102	0.319	2.468
Bum-Joo Cho (FCN)	0.9350	0.1221	-
Dat Tien Nguyen (ECNN)	0.8947	0.113	-

CNN: Convolutional neural network; FCN: Full convolutional neural network; ECNN: Evolutionary convolutional neural network; DICE: DICE Similarity Coefficient; VOE: Volumetric Over-lap Error; RMSD: Root-Mean-Square Deviation.

based on semi-supervised and unsupervised image recognition algorithms for gastric cancer. In such studies, they trained a small number of samples through generative models to generate similar samples to improve the accuracy and robustness of gastric cancer tumor recognition[65].

Generative adversarial network (GAN) is a generative model proposed by Goodfellow *et al*[66] It uses an unsupervised training method that is trained by adversarial learning. The objective is to estimate the potential distribution of data samples and generate new data samples. GAN is composed of a generation model (Goodfellow *et al*[66], 2014) and a discrimination model (Denton *et al*, 2015). The generation model learns the distribution of a given noise (generally refers to uniform distribution or normal distribution) and synthesizes it, whereas the discrimination model distinguishes the real data from generated data. In theory, the former is trying to produce data that is closed to the real data. The latter is also constantly strengthening the "counterfeit detection" ability[67]. The success of GAN lies in its ability to capture high-level semantic information using adversarial learning techniques. Luc *et al*[68] first applied GAN to image segmentation. However, GAN has several drawbacks: (1) Crash problem: when the generation model crashes, all different inputs are mapped to the same data[69]; and (2) Instability: It causes the same input to produce different outputs. The main reason is due to gradient vanishing problem during the optimization process[66,70].

Although batch normalization is often used to solve the instability of GAN, it is often not enough to achieve optimal stability of GAN performance. Therefore, many GAN derived models have emerged to solve these gaps, *e.g.*, conditional GAN, deep convolutional GAN, information maxi-mizing GAN, Wassertein GAN, *etc*[71]. In the GAN-based image recognition for gastric cancer, the generator is used to perform the segmentation task. The discriminator is then used to train the refining generator. A typical gastric cancer image recognition architecture based on the generative adversarial network is illustrated in Figure 3.

Since its proposal generative adversarial network has been widely considered and rapidly developed in different application areas. In medical image processing, it is very challenging to construct a large enough dataset due to the difficulty of data acquisition and annotation[72]. To overcome this problem, traditional image enhancement technology such as geometric transformation is often used to generate new data. This technique cannot learn biological changes in medical data and can produce images that are not credible[73]. Although GAN is unable to know in advance hypothesis distribution due to the limitation of segmentation performance improvement. it can automatically infer real data sets, further expand the scale and diversity of data, and provide a new method for data expansion, thus improving the efficiency of model training[74,75].

Almalioglu *et al*[76] showed that the poor resolution of the capsule endoscope is a limiting factor in the accuracy of diagnosis. They designed an image synthesis technology based on GAN to enrich the training data. First, the standard data expansion method was used to enlarge the dataset. Then the dataset was used to train GAN and the proposed Endol2h method was used to synthesize gastric cancer images with higher resolution[76]. Wang proposed an unsupervised image classification method for tumors based on prototype migration generated against the network (Prototype Transfer Generative Adversarial Network). Using different data acquisition devices and parameter settings caused differences in the stye of tumor image and data distribution. These differences can be reduced by designing the target domain to

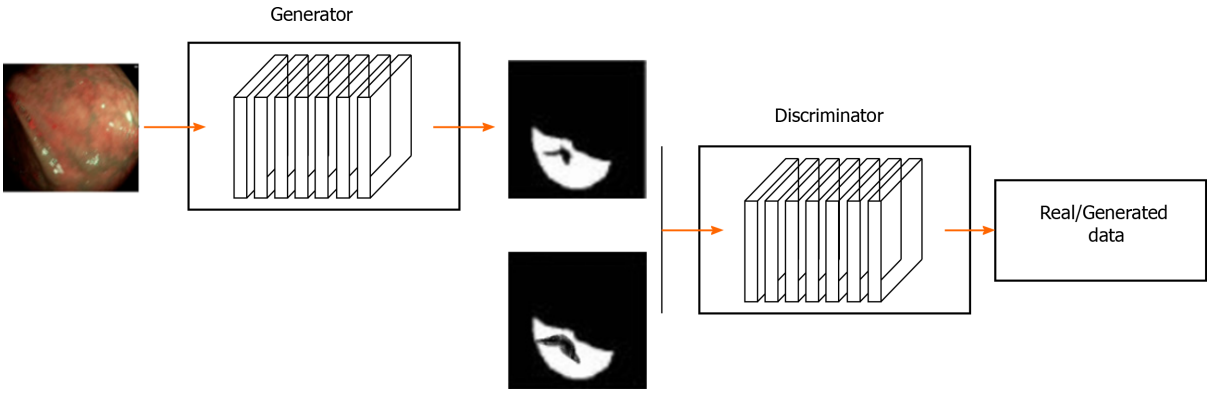

Figure 3 Basic architecture of gastric cancer image recognition methods based on generative adversarial network.

generate network, training process through the domain discriminant and performing generator reconstruction between source domain and target domain. The method achieved an average accuracy of 87.6% for unsupervised breast tumor image dichotomy under different magnifications and shows good scalability[77].

In conclusion, the GAN-based image segmentation method for gastric cancer can generate realistic gastric cancer images through the GAN network in the training stage, thus avoid the imbalance of the training samples. Moreover, due to the amplification of limited labeled sample data, the deep network is well-trained and achieves a high segmentation efficiency. However, there are still many problems in GAN, such as the instability of training and the breakdown of the training network. Therefore, researchers have optimized the original GAN network to reduce data noise or deal with class imbalance and other problems. In order to solve the problem that medical images are often polluted by different amounts and types of praise, T.Y Zhang *et al*. propose a novel Noise Adaptation Generative Adversarial Network (NAGAN), which contains a generator and two discriminators. The generator aims to map the data from source domain to target domain. Among the two discriminators, one discriminator enforces the generated images to have the same noise patterns as those from the target domain, and the second discriminator enforces the content to be preserved in the generated images. They apply the proposed NAGAN on both optical coherence tomography images and ultrasound images. Results show that the method is able to translate the noise style[74]. In the traditional GAN network training, the small number of samples of the minority classes in the training data makes the learning of optimal classification challenging, while the more frequently occurring samples of the majority class hamper the generalization of the classification boundary between infrequently occurring target objects and classes. Mina Rezaei *et al*. developed a novel generative multi-adversarial network, called Ensemble-GAN, for mitigating this class imbalance problem in the semantic segmentation of abdominal images. The Ensemble-GAN framework is composed of a single-generator and a multi-discriminator variant for handling the class imbalance problem to provide a better generalization than existing approaches[73]. In addition, there are other studies on the optimization of GAN network in medical image segmentation. Klages *et al*[78] proposed the patch-based generative adversarial neural network models, this model can significantly reduce errors in data generation. Nuo Tong *et al*[79] proposed the self-paced Dense-Net with boundary constraint for automated multi-organ segmentation on abdominal CT images. Specifically, a learning-based attention mechanism and dense connection block are seamlessly integrated into the proposed self-paced Dense-Net to improve the learning capability and efficiency of the backbone network. In a word, in the process of optimizing GAN network, whether it is optimizing generator or discriminant, the purpose of optimization is to generate new data which is as equal to the real data as possible. Therefore, more studies will be devoted to the optimization of GAN network to provide strong support for improving the image recognition of gastric cancer.

Table 4 shows comparison results of the three current mainstream methods for image recognition of gastric cancer.

Table 4 Comparison of convolutional neural network, full convolutional neural network, and generative adversarial network models

Model features	Contributions	Advantages	Disadvantages	Scope of application
CNN	The topology can be extracted from a two-dimensional image, and the backpropagation algorithm is used to optimize the network structure and solve the unknown parameters in the network	Shared convolution kernel, processing high-dimensional data without pressure; Feature extraction can be done automatically	When the network layer is too deep, the parameters near the input layer will be changed slowly by using BP propagation to modify parameters. A gradient descent algorithm is used to make the training results converge to the local minimum rather than the global minimum. The pooling layer will lose a lot of valuable information	Suitable for data scenarios with similar network structures
FCN	The end-to-end convolutional network is extended to semantic segmentation. The deconvolution layer is used for up-sampling; A skip connection is proposed to improve the roughness of the upper sampling	Can accept any size; Input image; Jump junction; The structure combines fine layers and coarse; Rough layers, generating precise segmentation	The receptive field is too small to obtain the global information;Small storage overhead	Applicable to large sample data
GAN	With adversarial learning criteria, there are two No's: The same network, not a single network	Can produce a clearer, more realistic sample; any generated network can be trained	Training is unstable and difficult to train; GAN is not suitable for processing data in discrete form	Suitable for data generation (*e.g.*, there are not many data sets with labels), image style transfer; Image denoising and restoration; Used to counter attacks

CNN: Convolutional neural network; FCN: Full convolutional neural network; GAN: Generative adversarial network.

CONCLUSION

At the present, the development direction of deep learning in image recognition of gastric cancer mainly focuses on the following aspects: (1) Training of deep learning algorithms relies on the availability of large datasets, because medical images are often difficult to obtain, medical professionals need to spend a lot of time on data collection and annotation which is time-consuming and costly. Besides, medical workers need not only to provide a large amount of data support but also to make use of all the effective information in the data as much as possible. Deep neural networks enable full mining of the information content of the data. Using deep networks seems to be the dominant future research direction in this field; (2) Multimodal gastric image segmentation combined with several different deep neural networks are used to extract the deeper information of the image and improve the accuracy of tumor segmentation and recognition. This is a promising major research direction in this field; and (3) Currently, most of the medical image segmentation techniques use supervised deep learning algorithms. However, for some of the rare diseases lacking a large number of data samples, supervised deep learning algorithms cannot reach their full efficiency. To overcome the issue with the lack of large datasets, some researchers utilize semi-supervised or unsupervised techniques such as GAN and combine the generated adversarial network with other higher performance networks. This might be another emerging research trend in this area.

REFERENCES

1 **Higgins AJ**, Lees P. Arachidonic acid metabolites in carrageenin-induced equine inflammatory exudate. *J Vet Pharmacol Ther* 1984; **7**: 65-72 [PMID: 6423835 DOI: 10.1111/j.1365-2885.1984.tb00881.x]

2 **Torre LA**, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. *CA Cancer J Clin* 2015; **65**: 87-108 [PMID: 25651787 DOI: 10.3322/caac.21262]

3 **Ferlay J**, Colombet M, Soerjomataram I, Dyba T, Randi G, Bettio M, Gavin A, Visser O, Bray F. Cancer incidence and mortality patterns in Europe: Estimates for 40 countries and 25 major cancers in 2018. *Eur J Cancer* 2018; **103**: 356-387 [PMID: 30100160 DOI: 10.1016/j.ejca.2018.07.005]

4 **Howlader N NA**, Krapcho M, Miller D, Brest A, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen HS, Feuer EJ, Cronin KA. SEER Cancer Statistics Review, 1975-2017. Bethesda: National Cancer Institute, 2019. Available from: https://seer.cancer.gov/csr/1975_2017/

5 **Rahman R**, Asombang AW, Ibdah JA. Characteristics of gastric cancer in Asia. *World J Gastroenterol* 2014; **20**: 4483-4490 [PMID: 24782601 DOI: 10.3748/wjg.v20.i16.4483]

6 **Chen W**, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, Jemal A, Yu XQ, He J. Cancer statistics in China, 2015. *CA Cancer J Clin* 2016; **66**: 115-132 [PMID: 26808342 DOI: 10.3322/caac.21338]

7 **Homeida AM**, Cooke RG. Pharmacological aspects of metaldehyde poisoning in mice. *J Vet Pharmacol Ther* 1982; **5**: 77-81 [PMID: 6178838 DOI: 10.1111/j.1365-2885.1982.tb00500.x]

8 **Luo X**, Mori K, Peters TM. Advanced Endoscopic Navigation: Surgical Big Data, Methodology, and Applications. *Annu Rev Biomed Eng* 2018; **20**: 221-251 [PMID: 29505729 DOI: 10.1146/annurev-bioeng-062117-120917]

9 **de Groof J**, van der Sommen F, van der Putten J, Struyvenberg MR, Zinger S, Curvers WL, Pech O, Meining A, Neuhaus H, Bisschops R, Schoon EJ, de With PH, Bergman JJ. The Argos project: The development of a computer-aided detection system to improve detection of Barrett's neoplasia on white light endoscopy. *United European Gastroenterol J* 2019; **7**: 538-547 [PMID: 31065371 DOI: 10.1177/2050640619837443]

10 **Qu JY**, Li Z, Su JR, Ma MJ, Xu CQ, Zhang AJ, Liu CX, Yuan HP, Chu YL, Lang CC, Huang LY, Lu L, Li YQ, Zuo XL. Development and Validation of an Automatic Image-Recognition Endoscopic Report Generation System: A Multicenter Study. *Clin Transl Gastroenterol* 2020; **12**: e00282 [PMID: 33395075 DOI: 10.14309/ctg.0000000000000282]

11 **Gan T**, Liu S, Yang J, Zeng B, Yang L. A pilot trial of Convolution Neural Network for automatic retention-monitoring of capsule endoscopes in the stomach and duodenal bulb. *Sci Rep* 2020; **10**: 4103 [PMID: 32139758 DOI: 10.1038/s41598-020-60969-5]

12 **Shin HC**, Roth HR, Gao M, Lu L, Xu Z, Nogues I, Yao J, Mollura D, Summers RM. Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. *IEEE Trans Med Imaging* 2016; **35**: 1285-1298 [PMID: 26886976 DOI: 10.1109/TMI.2016.2528162]

13 **Bernal J**, Kushibar K, Asfaw DS, Valverde S, Oliver A, Martí R, Lladó X. Deep convolutional neural networks for brain image analysis on magnetic resonance imaging: a review. *Artif Intell Med* 2019; **95**: 64-81 [PMID: 30195984 DOI: 10.1016/j.artmed.2018.08.008]

14 **Danaee P**, Ghaeini R, Hendrix DA. A Deep Learning Approach for Cancer Detection and Relevant Gene Identification. *Pac Symp Biocomput* 2017; **22**: 219-229 [PMID: 27896977 DOI: 10.1142/9789813207813_0022]

15 **Burke HB**, Goodman PH, Rosen DB, Henson DE, Weinstein JN, Harrell FE Jr, Marks JR, Winchester DP, Bostwick DG. Artificial neural networks improve the accuracy of cancer survival prediction. *Cancer* 1997; **79**: 857-862 [PMID: 9024725 DOI: 10.1002/(sici)1097-0142(19970215)79:4<857::aid-cncr24>3.0.co;2-y]

16 **Owais M**, Arsalan M, Choi J, Mahmood T, Park KR. Artificial Intelligence-Based Classification of Multiple Gastrointestinal Diseases Using Endoscopy Videos for Clinical Diagnosis. *J Clin Med* 2019; **8** [PMID: 31284687 DOI: 10.3390/jcm8070986]

17 **Lu S**, Cottone CM, Yoon R, Jefferson FA, Sung JM, Okhunov Z, Tapiero S, Patel RM, Landman J, Clayman RV. Endockscope: A Disruptive Endoscopic Technology. *J Endourol* 2019; **33**: 960-965 [PMID: 31195831 DOI: 10.1089/end.2019.0252]

18 **Wang KW**, Dong M. Potential applications of artificial intelligence in colorectal polyps and cancer: Recent advances and prospects. *World J Gastroenterol* 2020; **26**: 5090-5100 [PMID: 32982111 DOI: 10.3748/wjg.v26.i34.5090]

19 **Yasuda Y**, Tokunaga K, Koga T, Sakamoto C, Goldberg IG, Saitoh N, Nakao M. Computational analysis of morphological and molecular features in gastric cancer tissues. *Cancer Med* 2020; **9**: 2223-2234 [PMID: 32012497 DOI: 10.1002/cam4.2885]

20 **Yao Y**, Gou S, Tian R, Zhang X, He S. Automated Classification and Segmentation in Colorectal Images Based on Self-Paced Transfer Network. *Biomed Res Int* 2021; **2021**: 6683931 [PMID: 33542924 DOI: 10.1155/2021/6683931]

21 **Baig R**, Bibi M, Hamid A, Kausar S, Khalid S. Deep Learning Approaches Towards Skin Lesion Segmentation and Classification from Dermoscopic Images - A Review. *Curr Med Imaging* 2020; **16**: 513-533 [PMID: 32484086 DOI: 10.2174/1573405615666190129120449]

22 **Zhang K**, Liu X, Liu F, He L, Zhang L, Yang Y, Li W, Wang S, Liu L, Liu Z, Wu X, Lin H. An Interpretable and Expandable Deep Learning Diagnostic System for Multiple Ocular Diseases: Qualitative Study. *J Med Internet Res* 2018; **20**: e11144 [PMID: 30429111 DOI: 10.2196/11144]

23 **Weng S**, Xu X, Li J, Wong STC. Combining deep learning and coherent anti-Stokes Raman scattering imaging for automated differential diagnosis of lung cancer. *J Biomed Opt* 2017; **22**: 1-10 [PMID: 29086544 DOI: 10.1117/1.JBO.22.10.106017]

24 **Xu Y**, Jia Z, Wang LB, Ai Y, Zhang F, Lai M, Chang EI. Large scale tissue histopathology image classification, segmentation, and visualization *via* deep convolutional activation features. *BMC Bioinformatics* 2017; **18**: 281 [PMID: 28549410 DOI: 10.1186/s12859-017-1685-x]

25 **Yonekura A**, Kawanaka H, Prasath VBS, Aronow BJ, Takase H. Automatic disease stage classification of glioblastoma multiforme histopathological images using deep convolutional neural network. *Biomed Eng Lett* 2018; **8**: 321-327 [PMID: 30603216 DOI: 10.1007/s13534-018-0077-0]

26 **Horiuchi Y**, Aoyama K, Tokai Y, Hirasawa T, Yoshimizu S, Ishiyama A, Yoshio T, Tsuchida T, Fujisaki J, Tada T. Convolutional Neural Network for Differentiating Gastric Cancer from Gastritis Using Magnified Endoscopy with Narrow Band Imaging. *Dig Dis Sci* 2020; **65**: 1355-1363 [PMID: 31584138 DOI: 10.1007/s10620-019-05862-6]

27 **Guimarães P**, Keller A, Fehlmann T, Lammert F, Casper M. Deep-learning based detection of gastric precancerous conditions. *Gut* 2020; **69**: 4-6 [PMID: 31375599 DOI: 10.1136/gutjnl-2019-319347]

28 **Wu T**, Tegmark M. Toward an artificial intelligence physicist for unsupervised learning. *Phys Rev E* 2019; **100**: 033311 [PMID: 31639888 DOI: 10.1103/PhysRevE.100.033311]

29 **Zhang R**, Zhang Y, Li X. Unsupervised Feature Selection *via* Adaptive Graph Learning and Constraint. *IEEE Trans Neural Netw Learn Syst* 2020; **PP** [PMID: 33361001 DOI: 10.1109/TNNLS.2020.3042330]

30 **Wang Z**, Li M, Xu Z, Jiang Y, Gu H, Yu Y, Zhu H, Zhang H, Lu P, Xin J, Xu H, Liu C. Improvements to the gastric cancer tumor-node-metastasis staging system based on computer-aided unsupervised clustering. *BMC Cancer* 2018; **18**: 706 [PMID: 29970022 DOI: 10.1186/s12885-018-4623-z]

31 **Luo H**, Xu G, Li C, He L, Luo L, Wang Z, Jing B, Deng Y, Jin Y, Li Y, Li B, Tan W, He C, Seeruttun SR, Wu Q, Huang J, Huang DW, Chen B, Lin SB, Chen QM, Yuan CM, Chen HX, Pu HY, Zhou F, He Y, Xu RH. Real-time artificial intelligence for detection of upper gastrointestinal cancer by endoscopy: a multicentre, case-control, diagnostic study. *Lancet Oncol* 2019; **20**: 1645-1654 [PMID: 31591062 DOI: 10.1016/S1470-2045(19)30637-0]

32 **Lee JH**, Kim YJ, Kim YW, Park S, Choi YI, Park DK, Kim KG, Chung JW. Spotting malignancies from gastric endoscopic images using deep learning. *Surg Endosc* 2019; **33**: 3790-3797 [PMID: 30719560 DOI: 10.1007/s00464-019-06677-2]

33 **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]

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

35 **Ueyama H**, Kato Y, Akazawa Y, Yatagai N, Komori H, Takeda T, Matsumoto K, Ueda K, Hojo M, Yao T, Nagahara A, Tada T. Application of artificial intelligence using a convolutional neural network for diagnosis of early gastric cancer based on magnifying endoscopy with narrow-band imaging. *J Gastroenterol Hepatol* 2021; **36**: 482-489 [PMID: 32681536 DOI: 10.1111/jgh.15190]

36 **Li L**, Chen Y, Shen Z, Zhang X, Sang J, Ding Y, Yang X, Li J, Chen M, Jin C, Chen C, Yu C. Convolutional neural network for the diagnosis of early gastric cancer based on magnifying narrow band imaging. *Gastric Cancer* 2020; **23**: 126-132 [PMID: 31332619 DOI: 10.1007/s10120-019-00992-2]

37 **Bergholt MS**, Zheng W, Lin K, Ho KY, Teh M, Yeoh KG, Yan So JB, Huang Z. In vivo diagnosis of gastric cancer using Raman endoscopy and ant colony optimization techniques. *Int J Cancer* 2011; **128**: 2673-2680 [PMID: 20726002 DOI: 10.1002/ijc.25618]

38 **Liang Q**, Nan Y, Coppola G, Zou K, Sun W, Zhang D, Wang Y, Yu G. Weakly Supervised Biomedical Image Segmentation by Reiterative Learning. *IEEE J Biomed Health Inform* 2019; **23**: 1205-1214 [PMID: 29994489 DOI: 10.1109/JBHI.2018.2850040]

39 **El-Khatib H**, Popescu D, Ichim L. Deep Learning-Based Methods for Automatic Diagnosis of Skin Lesions. *Sensors (Basel)* 2020; **20** [PMID: 32245258 DOI: 10.3390/s20061753]

40 **Pereira S**, Pinto A, Alves V, Silva CA. Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images. *IEEE Trans Med Imaging* 2016; **35**: 1240-1251 [PMID: 26960222 DOI: 10.1109/TMI.2016.2538465]

41 **Polanski WH**, Zolal A, Sitoci-Ficici KH, Hiepe P, Schackert G, Sobottka SB. Comparison of Automatic Segmentation Algorithms for the Subthalamic Nucleus. *Stereotact Funct Neurosurg* 2020; **98**: 256-262 [PMID: 32369819 DOI: 10.1159/000507028]

42 **Liu T**, Liu J, Ma Y, He J, Han J, Ding X, Chen CT. Spatial feature fusion convolutional network for liver and liver tumor segmentation from CT images. *Med Phys* 2021; **48**: 264-272 [PMID: 33159809 DOI: 10.1002/mp.14585]

43 **Cuocolo R**, Caruso M, Perillo T, Ugga L, Petretta M. Machine Learning in oncology: A clinical appraisal. *Cancer Lett* 2020; **481**: 55-62 [PMID: 32251707 DOI: 10.1016/j.canlet.2020.03.032]

44 **LeCun Y**, Bengio Y, Hinton G. Deep learning. *Nature* 2015; **521**: 436-444 [PMID: 26017442 DOI: 10.1038/nature14539]

45 **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]

46 **Tan JW**, Wang L, Chen Y, Xi W, Ji J, Xu X, Zou LK, Feng JX, Zhang J, Zhang H. Predicting Chemotherapeutic Response for Far-advanced Gastric Cancer by Radiomics with Deep Learning Semi-automatic Segmentation. *J Cancer* 2020; **11**: 7224-7236 [PMID: 33193886 DOI: 10.7150/jca.46704]

47 **Cho BJ**, Bang CS, Lee JJ, Seo CW, Kim JH. Prediction of Submucosal Invasion for Gastric Neoplasms in Endoscopic Images Using Deep-Learning. *J Clin Med* 2020; **9** [PMID: 32549190 DOI: 10.3390/jcm9061858]

48 **Krizhevsky A**, Sutskever I, Hinton G. ImageNet Classification with Deep Convolutional Neural Networks. In: Pereira F, Burges CJC, Bottou L, Weinberger KQ, editors. Advances in Neural Information Processing Systems 25 (NIPS 2012). Red Hook: Curran Associates, 2012

49 **Karpathy A**, Toderici G, Shetty S, Leung T, Sukthankar R, Fei-Fei L. Large-Scale Video Classification with Convolutional Neural Networks. In: 2014 IEEE Conference on Computer Vision

and Pattern Recognition; 2014 June 23-28; Columbus, USA. IEEE, 2014 [DOI: 10.1109/CVPR.2014.223]

50 **Xiang Y**, Lin Z, Meng J. Automatic QRS complex detection using two-level convolutional neural network. *Biomed Eng Online* 2018; **17**: 13 [PMID: 29378580 DOI: 10.1186/s12938-018-0441-4]

51 **Shelhamer E**, Long J, Darrell T. Fully Convolutional Networks for Semantic Segmentation. *IEEE Trans Pattern Anal Mach Intell* 2016; **39**: 1 [DOI: 10.1109/tpami.2016.2572683]

52 **Kim HK**, Yoo KY, Park JH, Jung HY. Asymmetric Encoder-Decoder Structured FCN Based LiDAR to Color Image Generation. *Sensors (Basel)* 2019; **19** [PMID: 31694330 DOI: 10.3390/s19214818]

53 **Zhu H**, Adeli E, Shi F, Shen D; Alzheimer's Disease Neuroimaging Initiative. FCN Based Label Correction for Multi-Atlas Guided Organ Segmentation. *Neuroinformatics* 2020; **18**: 319-331 [PMID: 31898145 DOI: 10.1007/s12021-019-09448-5]

54 **Guo X**, Nie R, Cao J, Zhou D, Qian W. Fully Convolutional Network-Based Multifocus Image Fusion. *Neural Comput* 2018; **30**: 1775-1800 [PMID: 29894654 DOI: 10.1162/neco_a_01098]

55 **Zhou X**, Takayama R, Wang S, Hara T, Fujita H. Deep learning of the sectional appearances of 3D CT images for anatomical structure segmentation based on an FCN voting method. *Med Phys* 2017; **44**: 5221-5233 [PMID: 28730602 DOI: 10.1002/mp.12480]

56 **Wang R**, Cao S, Ma K, Zheng Y, Meng D. Pairwise learning for medical image segmentation. *Med Image Anal* 2021; **67**: 101876 [PMID: 33197863 DOI: 10.1016/j.media.2020.101876]

57 **Xue J**, He K, Nie D, Adeli E, Shi Z, Lee SW, Zheng Y, Liu X, Li D, Shen D. Cascaded MultiTask 3-D Fully Convolutional Networks for Pancreas Segmentation. *IEEE Trans Cybern* 2021; **51**: 2153-2165 [PMID: 31869812 DOI: 10.1109/TCYB.2019.2955178]

58 **Wu W**, Wu S, Zhou Z, Zhang R, Zhang Y. 3D Liver Tumor Segmentation in CT Images Using Improved Fuzzy *C*-Means and Graph Cuts. *Biomed Res Int* 2017; **2017**: 5207685 [PMID: 29090220 DOI: 10.1155/2017/5207685]

59 **Baazaoui A**, Barhoumi W, Zagrouba E. Semi-Automated Segmentation of Single and Multiple Tumors in Liver CT Images Using Entropy-Based Fuzzy Region Growing. *IRBM* 2017; **38**: 98-108 [DOI: 10.1016/j.irbm.2017.02.003]

60 **Öztürk Ş**, Özkaya U. Skin Lesion Segmentation with Improved Convolutional Neural Network. *J Digit Imaging* 2020; **33**: 958-970 [PMID: 32378058 DOI: 10.1007/s10278-020-00343-z]

61 **Oda M**, Tanaka K, Takabatake H, Mori M, Natori H, Mori K. Realistic endoscopic image generation method using virtual-to-real image-domain translation. *Healthc Technol Lett* 2019; **6**: 214-219 [PMID: 32038860 DOI: 10.1049/htl.2019.0071]

62 **de Groof AJ**, Struyvenberg MR, van der Putten J, van der Sommen F, Fockens KN, Curvers WL, Zinger S, Pouw RE, Coron E, Baldaque-Silva F, Pech O, Weusten B, Meining A, Neuhaus H, Bisschops R, Dent J, Schoon EJ, de With PH, Bergman JJ. Deep-Learning System Detects Neoplasia in Patients With Barrett's Esophagus With Higher Accuracy Than Endoscopists in a Multistep Training and Validation Study With Benchmarking. *Gastroenterology* 2020; **158**: 915-929. e4 [PMID: 31759929 DOI: 10.1053/j.gastro.2019.11.030]

63 **Sun M**, Zhang G, Dang H, Qi X, Zhou X, Chang Q. Accurate Gastric Cancer Segmentation in Digital Pathology Images Using Deformable Convolution and Multi-Scale Embedding Networks. *IEEE Access* 2019; **7**: 75530-75541 [DOI: 10.1109/ACCESS.2019.2918800]

64 **Nguyen DT**, Lee MB, Pham TD, Batchuluun G, Arsalan M, Park KR. Enhanced Image-Based Endoscopic Pathological Site Classification Using an Ensemble of Deep Learning Models. *Sensors (Basel)* 2020; **20** [PMID: 33105736 DOI: 10.3390/s20215982]

65 **Teramoto A**, Tsukamoto T, Yamada A, Kiriyama Y, Imaizumi K, Saito K, Fujita H. Deep learning approach to classification of lung cytological images: Two-step training using actual and synthesized images by progressive growing of generative adversarial networks. *PLoS One* 2020; **15**: e0229951 [PMID: 32134949 DOI: 10.1371/journal.pone.0229951]

66 **Goodfellow IJ**, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y. Generative adversarial nets. In: Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2. Montreal: MIT Press, 2014: 2672-2680

67 **Enokiya Y**, Iwamoto Y, Chen YW, Han XH. Automatic Liver Segmentation Using U-Net with Wasserstein GANs. *Int J Image Graph* 2019; **7**: 94-101 [DOI: 10.18178]

68 **Luc P**, Couprie C, Chintala S, Verbeek J. Semantic segmentation using adversarial networks. 2016 Preprint. Available from: arXiv:1611.08408

69 **Borji A**. Pros and cons of GAN evaluation measures. *CVIU* 2019; **179**: 41-65 [DOI: 10.1016/j.cviu.2018.10.009]

70 **Metz L**, Poole B, Pfau D, Sohl-Dickstein J. Unrolled generative adversarial networks. 2016 Preprint. Available from: arXiv:1611.02163

71 **Poorneshwaran JM**, Santhosh Kumar S, Ram K, Joseph J, Sivaprakasam M. Polyp Segmentation using Generative Adversarial Network. *Annu Int Conf IEEE Eng Med Biol Soc* 2019; **2019**: 7201-7204 [PMID: 31947496 DOI: 10.1109/EMBC.2019.8857958]

72 **Tang C**, Zhang W, Wang L, Cai A, Liang N, Li L, Yan B. Generative adversarial network-based sinogram super-resolution for computed tomography imaging. *Phys Med Biol* 2020; **65**: 235006 [PMID: 33053522 DOI: 10.1088/1361-6560/abc12f]

73 **Rezaei M**, Näppi JJ, Lippert C, Meinel C, Yoshida H. Generative multi-adversarial network for striking the right balance in abdominal image segmentation. *Int J Comput Assist Radiol Surg* 2020; **15**: 1847-1858 [PMID: 32897490 DOI: 10.1007/s11548-020-02254-4]

74 **Zhang T**, Cheng J, Fu H, Gu Z, Xiao Y, Zhou K, Gao S, Zheng R, Liu J. Noise Adaptation Generative Adversarial Network for Medical Image Analysis. *IEEE Trans Med Imaging* 2020; **39**: 1149-1159 [PMID: 31567075 DOI: 10.1109/TMI.2019.2944488]

75 **Han L**, Huang Y, Dou H, Wang S, Ahamad S, Luo H, Liu Q, Fan J, Zhang J. Semi-supervised segmentation of lesion from breast ultrasound images with attentional generative adversarial network. *Comput Methods Programs Biomed* 2020; **189**: 105275 [PMID: 31978805 DOI: 10.1016/j.cmpb.2019.105275]

76 **Almalioglu Y**, Bengisu Ozyoruk K, Gokce A, Incetan K, Irem Gokceler G, Ali Simsek M, Ararat K, Chen RJ, Durr NJ, Mahmood F, Turan M. EndoL2H: Deep Super-Resolution for Capsule Endoscopy. *IEEE Trans Med Imaging* 2020; **39**: 4297-4309 [PMID: 32795966 DOI: 10.1109/TMI.2020.3016744]

77 **Wang D**. Research on Key Technologies of Medical Image Classification Based on Unsupervised and Semi-supervised Framework. M.D. Thesis, Jilin University. 2020 Available from: http://cdmd.cnki.com.cn/Article/CDMD-10183-1020754186.htm

78 **Klages P**, Benslimane I, Riyahi S, Jiang J, Hunt M, Deasy JO, Veeraraghavan H, Tyagi N. Patch-based generative adversarial neural network models for head and neck MR-only planning. *Med Phys* 2020; **47**: 626-642 [PMID: 31733164 DOI: 10.1002/mp.13927]

79 **Tong N**, Gou S, Niu T, Yang S, Sheng K. Self-paced DenseNet with boundary constraint for automated multi-organ segmentation on abdominal CT images. *Phys Med Biol* 2020; **65**: 135011 [PMID: 32657281 DOI: 10.1088/1361-6560/ab9b57]

Artificial Intelligence in Gastrointestinal Endoscopy

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

Artif Intell Gastrointest Endosc 2021 April 28; 2(2): 25-35

DOI: 10.37126/aige.v2.i2.25

ISSN 2689-7164 (online)

MINIREVIEWS

Application of artificial intelligence to endoscopy on common gastrointestinal benign diseases

Hang Yang, Bing Hu

ORCID number: Hang Yang 0000-0002-7235-8162; Bing Hu 0000-0002-9898-8656.

Author contributions: All authors participated in the work; Yang H contributed to the design and draft of the manuscript; Hu B contributed to reviewing the manuscript; Yang H and Bing H contributed to revising the manuscript.

Supported by the 1·3·5 Project for Disciplines of Excellence Clinical Research Incubation Project, West China Hospital, Sichuan University, China, No. 20HXFH016.

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

Hang Yang, Bing Hu, Department of Gastroenterology, West China Hospital, Sichuan University, Chengdu 610041, Sichuan Province, China

Corresponding author: Bing Hu, MBBS, MD, Professor, Department of Gastroenterology, West China Hospital, Sichuan University, No. 37 Guo Xue Xiang, Wu Hou District, Chengdu 610041, Sichuan Province, China. hubingnj@163.com

Abstract

Artificial intelligence (AI) has been widely involved in every aspect of healthcare in the preclinical stage. In the digestive system, AI has been trained to assist auxiliary examinations including histopathology, endoscopy, ultrasonography, computerized tomography, and magnetic resonance imaging in detection, diagnosis, classification, differentiation, prognosis, and quality control. In the field of endoscopy, the application of AI, such as automatic detection, diagnosis, classification, and invasion depth, in early gastrointestinal (GI) cancers has received wide attention. There is a paucity of studies of AI application on common GI benign diseases based on endoscopy. In the review, we provide an overview of AI applications to endoscopy on common GI benign diseases including in the esophagus, stomach, intestine, and colon. It indicates that AI will gradually become an indispensable part of normal endoscopic detection and diagnosis of common GI benign diseases as clinical data, algorithms, and other related work are constantly repeated and improved.

Key Words: Artificial intelligence; Endoscopy; Common gastrointestinal benign diseases

Core Tip: In endoscopy, the application of artificial intelligence in early gastrointestinal cancer has been widely concerned. We provide a general conclusion of artificial intelligence endoscopy applications in common gastrointestinal benign diseases, such as Barrett's esophagus, atrophic gastritis, and colonic polyp. Studies indicate high accuracies and efficiencies. Further related work is needed to boost the real application of artificial intelligence in common gastrointestinal benign diseases in the future.

Citation: Yang H, Hu B. Application of artificial intelligence to endoscopy on common

Manuscript source: Invited manuscript

Specialty type: Gastroenterology and hepatology

Country/Territory of origin: China

Peer-review report's scientific quality classification

Grade A (Excellent): 0
Grade B (Very good): 0
Grade C (Good): C
Grade D (Fair): 0
Grade E (Poor): 0

Received: March 5, 2021
Peer-review started: March 5, 2021
First decision: March 14, 2021
Revised: March 17, 2021
Accepted: April 20, 2021
Article in press: April 20, 2021
Published online: April 28, 2021

P-Reviewer: Azimi P
S-Editor: Wang JL
L-Editor: Filipodia
P-Editor: Wang LL

gastrointestinal benign diseases. *Artif Intell Gastrointest Endosc* 2021; 2(2): 25-35
URL: https://www.wjgnet.com/2689-7164/full/v2/i2/25.htm
DOI: https://dx.doi.org/10.37126/aige.v2.i2.25

INTRODUCTION

Artificial intelligence (AI) is essentially a process of learning human thinking and transferring human experience based on mathematics and statistics. Iteration of algorithm, rising data, and improving computing power are cores of AI. Machine learning (ML) is a subset of AI[1], and deep learning is a subset of ML to realize ML[2], where multiple algorithms are structured together in complex layers. Artificial neural networks are one of the most common algorithms of AI[3]. Convolutional neural networks (CNNs) are a kind of supervised deep learning algorithm[4]. Its modified format is defined as deep convolutional neural networks[5]. Recognizing images based on artificial neural networks/CNNs promotes AI penetrating in medicine. Computer-aided diagnosis (CAD) systems are designed to interpret medical images using advances of AI from ML to deep learning[6].

In the field of gastroenterology, diseases of the liver, pancreases, and full digestive tract have been involved. Examples include a deep learning model based on computed tomography images to stage liver fibrosis, a deep learning model constructed to differentiate between precancerous lesions and pancreatic cancers, and a deep learning model used in endoscopy to detect early gastrointestinal (GI) cancers. A study covered five kinds of gastric diseases and showed the diagnostic specificity of the CNNs was higher than that of the endoscopists for early gastric cancer and high-grade intraepithelial neoplasia images (91.2% *vs* 86.7%). The diagnostic accuracy of the CNNs was close to those of the endoscopists for lesion-free, early gastric cancer and high-grade intraepithelial neoplasia, peptic ulcer (PU), advanced gastric cancer (GC), and gastric submucosal tumor images. The CNNs had an image recognition time of 42 s for all the test set images[7]. In this review, the application and research of AI on common GI benign lesions based on endoscopy were concluded.

LITERATURE SEARCH

This review aimed to make a qualitative only review of the application of AI on common GI benign diseases. We searched the PubMed database for articles that were published in the last 5 years using the term combinations of artificial intelligence and common GI benign lesions [Barrett's esophagus (BE), esophageal varices (EV), atrophic gastritis (AG), PU, gastric polyp, small bowel capsule endoscopy, colonic polyp/adenoma, and inflammatory bowel diseases (IBDs)]. Articles based on radiological images or other samples, review articles, research articles of early or advanced GI cancers or other cancers, and articles only related to either GI benign diseases or AI were excluded. Two authors independently extracted data. Any disagreement was resolved by discussion until consensus was reached or by consulting a third author. Endoscopic-related results were qualitatively concluded in Table 1. The flowchart was presented in Figure 1.

SEARCH RESULTS

Initially, a total of 555 articles were identified. After manually screening and reading, only research articles related to the application of AI to common GI benign lesions (BE, EV, AG, PU, gastric polyp, small bowel capsule endoscopy, colonic polyp/adenoma, and IBDs) based on different endoscopic images or tissue slides from endoscopic biopsies were included. Finally, 35 studies were tabulated in Table 1. Six studies demonstrated the application of AI on esophageal benign diseases (5 BE and 1 EV). Seven studies were about gastric benign diseases (3 AG, 3 PU, and 1 polyp). Seven studies were about intestinal diseases. Fifteen studies were about colonic benign diseases (11 polyp/adenoma and 4 IBDs).

Table 1 Application of artificial intelligence on common gastrointestinal benign diseases

Ref.	Aim and disease	Prospective/retrospective	AI method	Endoscopy image	Training dataset	Validation dataset	Result sensitivity	Result specificity	Result accuracy/AUC
Esophageal benign diseases									
de Groof *et al*[12]	Detecting Barrett'sneoplasia	Retrospective	CAD	WLI images	40 images	A leave one out cross validation	92%	95%	85%[1]
Jisu *et al*[39]	Distinguishing BE	Retrospective	CNNs	Endomicroscopic images	262 images	Image distortion methods			80.77%[1]
Ebigbo *et al*[40]	Distinguishing BE	Retrospective	CNNs (ResNet)	WLI images	129 images	62 images	83.7%	100.0%	89.9%[1]
Sehgal *et al*[41]	Detecting dysplasia in BE	Retrospective	ML (decision trees)	Video recordings(AAC)	40 patients with NDBE and DBE		97%	88%	92%[1]
de Groof *et al*[14]	Detecting Barrett'sneoplasia	Retrospective	CNN (CAD (ResNet-UNet))	WLI images	494364 images	1704 images (early stage neoplasia in BE and NDBE from 669 patients)	90%	88%	89%[1]
Dong *et al*[16]	Screening high risk EV	Retrospective	ML (Random forest)		238 patients	109 patients			Training set (0.84); Validation set (0.82)
Gastric benign diseases									
Zhang *et al*[42]	Diagnosing CAG	Retrospective	CNNs (DenseNet)	WLI images	5470 images	Five-fold cross validation	94.5%	94.0%	94.2%[1]
Guimarães *et al*[43]	DiagnosingCAG	Retrospective	CNNs (VGG16)	WLI images	200 images	70 images(ten-fold cross validation)			93%[1]/0.98
Horiuchi *et al*[44]	Differentiating CAG	Retrospective	CNNs (GoogLeNet)	ME-NBI images	1078 images	107 images	95.4%	71.0%	85.3%[1]/0.85
Zhang *et al*[7]	Diagnosing PU	Retrospective	CNNs (ResNet34)	WLI images	4200 images	228 images	78.9%	88.4%	86.4%[1]
Lee *et al*[45]	Differentiating PU	Retrospective	CNNs (ResNet-50/ Inception v3/VGG16 model)	WLI images	200 images	20 images			92.6%[1]/85.24%[1] /91.2%[1]
Namikawa *et al*[46]	Classifying gastriccancers and ulcers	Retrospective	CNNs (SSD)	WLI/NBI/chromoendoscopy images	373 images	720 images	93.3%	99.0%	93.3 %[1]
Zhang *et al*[26]	Detecting GP	Retrospective	CNNs (SSD-GPNet)	WLI images	404 images	50 images			93.92%[1]
Intestinal benign diseases									

Hwang *et al* [29]	Classifying hemorrhagic and ulcerations	Retrospective	CNNs (VGGNet)	Capsule endoscopy	7556 images	5760 images	Model 1 *vs* Model 2; 97.61% *vs* 95.07%	Model 1 *vs* Model 2; 96.04% *vs* 98.18%	Model 1 *vs* Model 2; 96.83%[1] *vs* 96.62%[1]
Aoki *et al*[47]	Detecting erosions and ulcerations	Retrospective	CNNs (SSD)	Capsule endoscopy	5360 images	10440 images	88.2%	90.9%	90.8%[1]/0.958
Aoki *et al*[48]	Detecting erosions and ulcerations	Retrospective	CNNs (SSD)	Capsule endoscopy		20 videos			
Ding *et al*[49]	Detecting small bowel diseases	Retrospective	CNNs (ResNet)	Capsule endoscopy	158235 images	5000 patients	99.88% per patient99.90% per lesion	100% per patient100 % per lesion	
Fan *et al*[50]	Detecting erosions and ulcerations	Retrospective	CNNs (AlexNet)	Capsule endoscopy	Ulcer 2000; Erosion 2720	Ulcer 500; Erosion 690	Ulcer: 96.80%; Erosion: 93.67%	Ulcer: 94.79%; Erosion: 95.98%	Ulcer: 95.16%[1]; Erosion: 95.34%[1]/0.98
Leenhardt *et al* [51]	Detecting small bowel angiectasia	Retrospective	CNNs	Capsule endoscopy	300 videos with angiectasia	300 videos with angiectasia	100%	96%	
Tsuboi *et al*[52]	Detecting small bowel angiectasia	Retrospective	CNNs (SSD)	Capsule endoscopy	141 patients	28 patients	98.8%	98.4%	0.998
Colonic benign diseases									
Lui *et al*[34]	Detecting missed colonic lesions	Retrospective and prospective	R-FCN (ResNet101)	Endoscopic videos (WLI)	52 videos	Real-time AI detected at least 1 missed adenoma in 14 patients (26.9%) and increased the total number of adenomas detected by 23.6%.			
Rodriguez-Diaz *et al*[53]	Histologically classifying CP	Retrospective	CAD	NBI	745 images +65000 images		96%	84%	
Komeda *et al* [54]	Diagnosing CP	Retrospective	CNNs-CAD	WLI/NBI/ chromoendoscopy images	1200 images	10-fold cross validation			75.1%[1]
Akbari *et al*[55]	Classifying CP	Retrospective	FCNs	WLI images	200 images	300 images			
Chen *et al*[56]	Classifying diminutive CP	Retrospective	DCNNs-CAD	NBI images	96 images + 188 images		96.3%	78.1%	
Gong *et al*[57]	Detecting CA	Prospective	DCNNs	WLI images	DCNNs system (n = 355) or unassisted (control) colonoscopy (n = 349)		58 (16%) of 35527 (8%) of 349		
Byrne *et al*[58]	Differentiating adenomatous and hyperplastic polyps	Retrospective	DCNNs	Videos and NBI images	223 polyp videos	40 videos	98%	83%	
Mori *et al*[59]	Identifying diminutive CP	Prospective	CAD	NBI/stained images	791 consecutive patients undergoing colonoscopy and 23 endoscopists				Pathologic prediction rate of 98.1%[1]
Misawa *et al*[60]	DetectingCP	Retrospective	CAD	WLI images	105 positive and 306 negative	50 positive and 85 negative videos	90.0%	63.3%	76.5%[1]

					videos				
Taunk *et al*[61]	Classifying polyp histology	Retrospective	CAD	pCLE images	125 images	189 images	95%	94%	94%[1]
Wang *et al*[62]	Detecting CA	Prospective	CAD	WLI images	484 patients in the CADe group and 478 in the sham group		165 (34%) of 484; 132 (28%) of 478		
Tong *et al*[63]	Differentiating UC, CD, and ITB	Retrospective	CNNs/RF	WLI images	6399 consecutive patients (5128 UC, 875 CD and 396 ITB)		RF (UC 97%, CD 65%, and ITB 68%); CNN (UC 99%, CD 87%, and ITB 52%)	RF (UC 97%, CD 53%, and ITB 76%); CNN (UC 97%, CD 83%, and ITB 81%)	RF (UC 0.97, CD 0.58, and ITB 0.72); CNN (UC 0.98, CD 0.85, and ITB 0.63)
Ozawa *et al* [36]	Diagnosing UC	Retrospective	CAD	WLI images	26304 images	3981 images			0.86 (Mayo 0); 0.98 (Mayo 0–1)
Stidham *et al*[37]	Grading the severity of ulcerative colitis	Retrospective	CNNs	WLI images	2465 patients	308 patients	83.0%	96.0%	0.966
Maeda *et al*[38]	Identifying histologic inflammation associated with UC	Retrospective	CAD	Endocytoscopic images	87 patients	100 patients	74%	97%	91%[1]

[1]Results accuracy. AAC: Acetic acid chromoendoscopy; AI: Artificial intelligence; AUC: Area under the curve; BE: Barrett's esophagus; CA: Colorectal adenomas; CAD: Computer-aided diagnosis; CAG: Chronic atrophic gastritis; CD: Crohn's disease; CNN: Convolutional neural network; CP: Colorectal polyp; DBE: Dysplastic Barrett's esophagus; DCNNs: Deep convolutional neural networks; EV: Esophageal Varices; FCNs: Fully convolutional networks; GP: Gastric polyp; ITB: Intestinal tuberculosis; ME-NBI: Magnifying narrow-band imaging; ML: Machine learning; NBI: Narrow-band imaging; NDBE: Non-dysplastic Barrett's esophagus; pCLE: Probe-based confocal laser endomicroscopy; PU: Peptic ulcer; RF: Random forest; R-FCNs: Region-based fully connected convolutional neural networks; SSD: Single shot detector; UC: Ulcerative colitis; WLI: White-light imaging.

AI AND ESOPHAGEAL BENIGN DISEASES: BARRETT'S ESOPHAGUS AND ESOPHAGEAL VARICES

BE is a precursor to esophageal adenocarcinoma. Intestinal metaplasia and gastric metaplasia are two pathological subclasses of BE. Intestinal metaplasia can progress to esophageal cancer. The ablation of dysplastic BE will reduce the risk of progression to cancer[8]. Endoscopic surveillance, including white-light imaging (WLI), narrow-band imaging, and chromoendoscopy, is performed to detect dysplasia in BE. Approximately 5% of the esophageal mucosa is found at risk by random biopsies sample[9].

Recently, AI has been applied in some studies of BE. For example, CAD based on deep learning and different algorithms trained by WLI and endomicroscopic images to detect, diagnose, and distinguish BE with achievable results (the accuracy from 80.77% to 92%, specificity from 88% to 100%, and sensitivity from 83.7% to 97%) (Table 1). On pathology, CAD with wide area transepithelial sampling could increase the detection of high-grade dysplasia/esophageal adenocarcinoma (absolute increase: 14.4%)[10]. Deep convolutional neural networks were used in the whole-slide tissue histopathology images-based diagnosis of dysplastic and non-dysplastic BE[11]. Moreover, distinguishing BE adenocarcinoma by AI methods has been studied based on different endoscopic images such as WLI and volumetric laser endomicroscopic

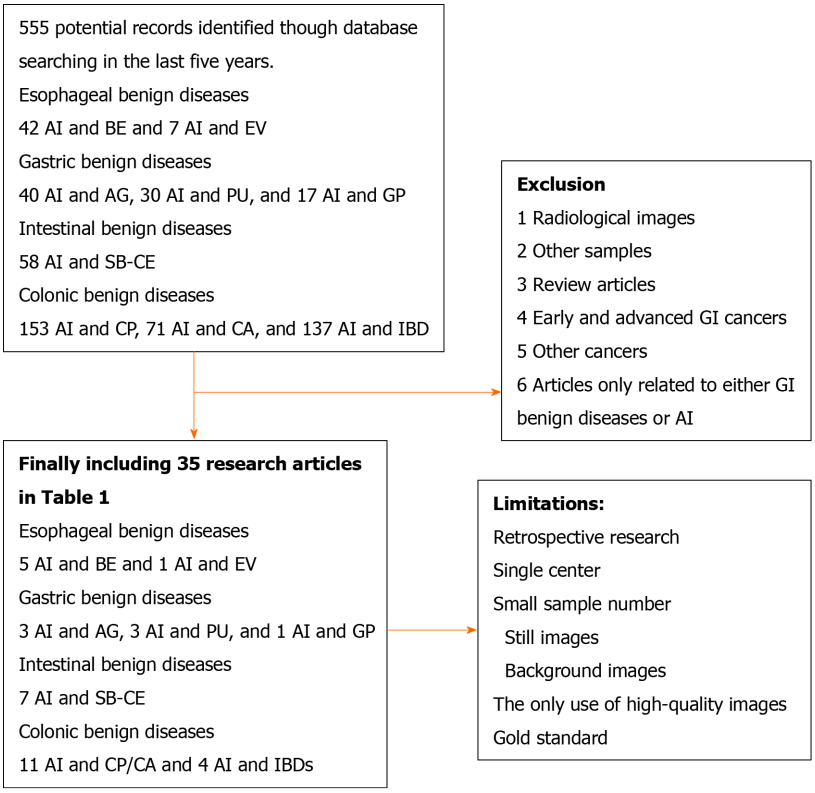

Figure 1 Flow chart of study selection and logic arrangement of review. AG: Atrophic gastritis; AI: Artificial intelligence; BE: Barrett's esophagus; CA: Colonic adenoma; CP: Colonic polyp; EV: Esophageal varices; GI: Gastrointestinal; GP: Gastric polyp; IBDs: Inflammatory bowel diseases; PU: Peptic ulcer; SB-CE: Small bowel capsule endoscopy.

images with accuracy from 88% to 92%, specificity from 88% to 93%, and sensitivity from 90% to 95%[12-14].

As another common esophageal benign disease, EV are associated with cirrhosis and portal hypertension, and variceal hemorrhage is a substantial cause of mortality[15]. However, related AI research is limited. A score system based on ML was built on the data of 238 patients with cirrhosis to reliably identify patients with varices that needed treatments and achieved an area under the curve (AUC) from 0.75 to 0.84 in different groups[16]. Another study of the index of spleen volume-to-platelet ratio based on deep learning-measured spleen volume on computed tomography to assess high-risk varices in B-viral compensated cirrhosis had a sensitivity of 69.4% and specificity of 78.5%[17]. There is little research of AI on esophagitis, although it is also a common esophageal disease associated with BE and esophageal cancer.

AI AND GASTRIC BENIGN LESIONS: ATROPHIC GASTRITIS, PEPTIC ULCER, AND POLYP

Gastritis, peptic ulcer, polyp and adenoma, and vascular lesion are common gastric benign diseases. The detection and diagnosis of these lesions account for a large part of daily endoscopic work. If AI can be applied in this field, then the rate of detection and accuracy will be improved. Moreover, the rapid identification of simple lesions can fill the lack of endoscopists and reduce the workload.

Early diagnosis of chronic AG, a precancerous lesion, is important to prevent the occurrence and development of GC. AI-assisted detection and diagnosis has been related to endoscopic images (Table 1), histological images[18,19], and X-ray images[20,21]. The accuracy was from 85.3% to 94.2%, the specificity was from 71% to 94%, and the sensitivity was from 94.5% to 95.4%. *Helicobacter pylori* infection, as a dominant cause of chronic AG and GC, has also been detected *via* AI methods based on endoscopic images, such as CNNs (GoogLeNet) and CNNs (ResNet-50 model), which achieved an accuracy up to 93.8% in a considerably short time of less than 200 s[22-24].

A CNN method was constructed to diagnose PU and differentiate GC from PU mainly based on WLI, narrow-band imaging, and chromoendoscopic images with an accuracy from 85.2% to 93.3%, specificity from 88.4% to 99%, and sensitivity from 78.9% to 93.3% (Table 1). In addition, a ML model was built on six parameters, such as age and the presence of PU, to predict recurrent ulcer bleeding within 1 year with an AUC of 0.775 and an accuracy of 84.3%[25].

There were only a few applications of AI on detecting gastric hyperplastic polyps and adenomas. A 93.92% accuracy was achieved when detecting polyps by CNNs (SSD-GPNet) based on WLI images[26]. A CNN method was trained to detect adenomas and showed an AUC of 0.99 based on histopathology whole-slide images[27]. Research and application of AI on gastric benign lesions are limited, although these diseases make up a considerable part of daily work. Some of them are usually prone to severe outcomes and risks despite the relative ease to diagnose. Indeed, the study of AI on this aspect will assist endoscopists to improve early detection rates and bring the opportunity of early treatment to benefit patients.

AI AND INTESTINAL DISEASES: CAPSULE ENDOSCOPY

The application of AI in small bowel diseases has been concentrated on capsule endoscopy. It includes image enhancement using ML algorithms to reduce artifact interference as well as three-dimensional luminal map reconstruction and localization[28]. AI-assisted capsule endoscopy in detecting ulcer, erosion, bleeding, polyps, parasite, diverticulum, and angiectasia with an accuracy more than 90.0%, specificity from 90.9% to 100%, and sensitivity from 88.2% to 100% in a short time (about 6 min) (Table 1). Furthermore, a gradient class activation map was used to visualize and detect lesions by CNNs-VGGNet to improve the classification and localization[29]. In addition, a CNN method based on conventional abdominal radiographs was trained to detect high-grade small bowel obstruction with an AUC of 0.84, a sensitivity of 83.8%, and a specificity of 68.1%[30]. In another study, it achieved an AUC of 0.971, a sensitivity of 91.4%, and a specificity of 91.9% using region-based CNNs[31]. The limited research indicated CNNs could recognize specific images among a large variety with high efficiency and accuracy. The application of AI will relieve the clinical workload as capsule endoscopy reading is a time-consuming process.

AI AND COLONIC BENIGN LESIONS: POLYP, ADENOMA, AND IBDS

A 1.0% increase of adenoma detection rate has been associated with a 3.0% decrease in the risk of interval colorectal cancer[32]. To improve colorectal polyp and adenoma detection, AI has been widely applied in the detection, real-time histological classification, segmentation, localization, and distinguishing of diminutive polyps and adenomas based on different methods trained by videos and images in retrospective or prospective and in multicenter or single center clinical trials (Table 1). Deep learning was also used to automatically classify colorectal polyps on histopathologic slides[33]. For the internal evaluation, the accuracy of the deep CNN method was 93.5%, which was comparable to the pathologists accuracy of 91.4%. On the external test, it achieved an accuracy of 87.0%, which was comparable to the pathologists accuracy of 86.6%. The application of AI in colorectal polyps has gained more concerns and practice, and it is deeper and closer to the clinical use to further increase the detection rate of polyps. For example, real-time AI detected at least one missed adenoma in 14 patients (26.9%) and increased the total number of adenomas detected by 23.6%[34].

AI methods have been trained in grading endoscopic disease severity of patients with ulcerative colitis and in predicting remission in patients with moderate to severe Crohn's disease[35]. For example, a CNN-CAD system based on GoogLeNet was robustly promising to identify normal mucosa (Mayo 0) and mucosal healing state with an accuracy of 0.86 of Mayo 0 and of 0.98 of Mayo 0-1[36]. Another similar system could differentiate remission (Mayo 0 or 1) from moderate or severe disease (Mayo 2 or 3) with an AUC of 0.966, a specificity of 96.0%, and a sensitivity of 83.0%[37]. A CAD was constructed to identify the presence of histologic inflammation associated with ulcerative colitis using endocytoscopy with an accuracy of 91%, a specificity of 97%, and a sensitivity 74%[38] (Table 1).

FUTURE PERSPECTIVES OF AI APPLICATION ON COMMON GI BENIGN LESIONS

We summarized the application and research of AI on common GI benign diseases. Limited studies are promising as most of the studies showed comparatively high accuracies and efficiencies. As studies of AI application on gastroenterology continue to increase, there are several areas of interest that will hold significant value in the future. First, the technical integration of AI systems will be important to optimize clinical workflow. New AI applications can easily "read in" data from a video input, allowing the systems to use the data for training and real time decision support. Second, AI systems will continue to expand the clinical applications. Some promising studies have demonstrated how AI can improve our performance on clinical tasks such as polyp identification, detection of small bowel bleeding, and endoscopic recognition of *Helicobacter pylori* infection, *etc*. More research, especially randomized controlled trials, on how to train and validate up-to-date algorithms will be continued on the present work to find more precise methods and identify new clinical tasks after practice. Third, further research will be needed to describe the most effective training methods for physician practices beginning to adopt AI technology because AI will be an indispensable helper of normal endoscopic detection and diagnosis of common GI benign lesions in the future.

CONCLUSION

Although AI is a relatively new technology, it has the potential to ease the daily workload of radiologists, pathologists, and sonographers. In endoscopy, AI related to early GI cancers and precancerous lesions has garnered more research than common GI benign diseases, despite the latter occupying a large proportion of daily work and being easier to detect and diagnose than early cancers. If models and diagnosing routes based on AI targeted at common GI benign diseases are well developed, then it will bring great benefits to patients and endoscopists, especially in primary hospitals where medical resources are lacking and core work is mainly focused on early diagnosis and treatment of common GI benign diseases. Furthermore, AI methods and technology targeted at common benign diseases will be easier for endoscopists to adopt professional education. More research is needed to overcome the challenges of integrating AI into the detection of common GI benign diseases by endoscopy, but the future is promising.

REFERENCES

1 **Bi Q**, Goodman KE, Kaminsky J, Lessler J. What is Machine Learning? *Am J Epidemiol* 2019; **188**: 2222-2239 [PMID: 31509183 DOI: 10.1093/aje/kwz189]

2 **LeCun Y**, Bengio Y, Hinton G. Deep learning. *Nature* 2015; **521**: 436-444 [PMID: 26017442 DOI: 10.1038/nature14539]

3 **Zador AM**. A critique of pure learning and what artificial neural networks can learn from animal brains. *Nat Commun* 2019; **10**: 3770 [PMID: 31434893 DOI: 10.1038/s41467-019-11786-6]

4 **Mieloszyk RJ**, Bhargava P. Convolutional Neural Networks: The Possibilities are Almost Endless. *Curr Probl Diagn Radiol* 2018; **47**: 129-130 [PMID: 29477265 DOI: 10.1067/j.cpradiol.2018.01.008]

5 **O'Toole AJ**, Castillo CD, Parde CJ, Hill MQ, Chellappa R. Face Space Representations in Deep Convolutional Neural Networks. *Trends Cogn Sci* 2018; **22**: 794-809 [PMID: 30097304 DOI: 10.1016/j.tics.2018.06.006]

6 **Ahmad OF**, Soares AS, Mazomenos E, Brandao P, Vega R, Seward E, Stoyanov D, Chand M, Lovat LB. Artificial intelligence and computer-aided diagnosis in colonoscopy: current evidence and future directions. *Lancet Gastroenterol Hepatol* 2019; **4**: 71-80 [PMID: 30527583 DOI: 10.1016/S2468-1253(18)30282-6]

7 **Zhang L**, Zhang Y, Wang L, Wang J, Liu Y. Diagnosis of gastric lesions through a deep convolutional neural network. *Dig Endosc* 2020 [PMID: 32961597 DOI: 10.1111/den.13844]

8 **Soh YSA**, Lee YY, Gotoda T, Sharma P, Ho KY; Asian Barrett's Consortium. Challenges to diagnostic standardization of Barrett's esophagus in Asia. *Dig Endosc* 2019; **31**: 609-618 [PMID: 30892742 DOI: 10.1111/den.13402]

9 **ASGE Technology Committee**. , Thosani N, Abu Dayyeh BK, Sharma P, Aslanian HR, Enestvedt BK, Komanduri S, Manfredi M, Navaneethan U, Maple JT, Pannala R, Parsi MA, Smith ZL, Sullivan SA, Banerjee S. ASGE Technology Committee systematic review and meta-analysis assessing the ASGE Preservation and Incorporation of Valuable Endoscopic Innovations thresholds for adopting

real-time imaging-assisted endoscopic targeted biopsy during endoscopic surveillance of Barrett's esophagus. *Gastrointest Endosc* 2016; **83**: 684-98. e7 [PMID: 26874597 DOI: 10.1016/j.gie.2016.01.007]

10 **Vennalaganti PR**, Kaul V, Wang KK, Falk GW, Shaheen NJ, Infantolino A, Johnson DA, Eisen G, Gerson LB, Smith MS, Iyer PG, Lightdale CJ, Schnoll-Sussman F, Gupta N, Gross SA, Abrams J, Haber GB, Chuttani R, Pleskow DK, Kothari S, Goldblum JR, Zhang Y, Sharma P. Increased detection of Barrett's esophagus-associated neoplasia using wide-area trans-epithelial sampling: a multicenter, prospective, randomized trial. *Gastrointest Endosc* 2018; **87**: 348-355 [PMID: 28757316 DOI: 10.1016/j.gie.2017.07.039]

11 **Sali R**, Moradinasab N, Guleria S, Ehsan L, Fernandes P, Shah TU, Syed S, Brown DE. Deep Learning for Whole-Slide Tissue Histopathology Classification: A Comparative Study in the Identification of Dysplastic and Non-Dysplastic Barrett's Esophagus. *J Pers Med* 2020; **10** [PMID: 32977465 DOI: 10.3390/jpm10040141]

12 **de Groof J**, van der Sommen F, van der Putten J, Struyvenberg MR, Zinger S, Curvers WL, Pech O, Meining A, Neuhaus H, Bisschops R, Schoon EJ, de With PH, Bergman JJ. The Argos project: The development of a computer-aided detection system to improve detection of Barrett's neoplasia on white light endoscopy. *United European Gastroenterol J* 2019; **7**: 538-547 [PMID: 31065371 DOI: 10.1177/2050640619837443]

13 **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]

14 **de Groof AJ**, Struyvenberg MR, van der Putten J, van der Sommen F, Fockens KN, Curvers WL, Zinger S, Pouw RE, Coron E, Baldaque-Silva F, Pech O, Weusten B, Meining A, Neuhaus H, Bisschops R, Dent J, Schoon EJ, de With PH, Bergman JJ. Deep-Learning System Detects Neoplasia in Patients With Barrett's Esophagus With Higher Accuracy Than Endoscopists in a Multistep Training and Validation Study With Benchmarking. *Gastroenterology* 2020; **158**: 915-929. e4 [PMID: 31759929 DOI: 10.1053/j.gastro.2019.11.030]

15 **Kovacs TOG**, Jensen DM. Varices: Esophageal, Gastric, and Rectal. *Clin Liver Dis* 2019; **23**: 625-642 [PMID: 31563215 DOI: 10.1016/j.cld.2019.07.005]

16 **Dong TS**, Kalani A, Aby ES, Le L, Luu K, Hauer M, Kamath R, Lindor KD, Tabibian JH. Machine Learning-based Development and Validation of a Scoring System for Screening High-Risk Esophageal Varices. *Clin Gastroenterol Hepatol* 2019; **17**: 1894-1901. e1 [PMID: 30708109 DOI: 10.1016/j.cgh.2019.01.025]

17 **Lee CM**, Lee SS, Choi WM, Kim KM, Sung YS, Lee S, Lee SJ, Yoon JS, Suk HI. An index based on deep learning-measured spleen volume on CT for the assessment of high-risk varix in B-viral compensated cirrhosis. *Eur Radiol* 2021; 31: 3355-3365 [PMID: 33128186 DOI: 10.1007/s00330-020-07430-3]

18 **Steinbuss G**, Kriegsmann K, Kriegsmann M. Identification of Gastritis Subtypes by Convolutional Neuronal Networks on Histological Images of Antrum and Corpus Biopsies. *Int J Mol Sci* 2020; **21** [PMID: 32932860 DOI: 10.3390/ijms21186652]

19 **Martin DR**, Hanson JA, Gullapalli RR, Schultz FA, Sethi A, Clark DP. A Deep Learning Convolutional Neural Network Can Recognize Common Patterns of Injury in Gastric Pathology. *Arch Pathol Lab Med* 2020; **144**: 370-378 [PMID: 31246112 DOI: 10.5858/arpa.2019-0004-OA]

20 **Li Z**, Togo R, Ogawa T, Haseyama M. Chronic gastritis classification using gastric X-ray images with a semi-supervised learning method based on tri-training. *Med Biol Eng Comput* 2020; **58**: 1239-1250 [PMID: 32221796 DOI: 10.1007/s11517-020-02159-z]

21 **Kanai M**, Togo R, Ogawa T, Haseyama M. Chronic atrophic gastritis detection with a convolutional neural network considering stomach regions. *World J Gastroenterol* 2020; **26**: 3650-3659 [PMID: 32742133 DOI: 10.3748/wjg.v26.i25.3650]

22 **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]

23 **Zheng W**, Zhang X, Kim JJ, Zhu X, Ye G, Ye B, Wang J, Luo S, Li J, Yu T, Liu J, Hu W, Si J. High Accuracy of Convolutional Neural Network for Evaluation of Helicobacter pylori Infection Based on Endoscopic Images: Preliminary Experience. *Clin Transl Gastroenterol* 2019; **10**: e00109 [PMID: 31833862 DOI: 10.14309/ctg.0000000000000109]

24 **Bang CS**, Lee JJ, Baik GH. Artificial Intelligence for the Prediction of Helicobacter Pylori Infection in Endoscopic Images: Systematic Review and Meta-Analysis Of Diagnostic Test Accuracy. *J Med Internet Res* 2020; **22**: e21983 [PMID: 32936088 DOI: 10.2196/21983]

25 **Wong GL**, Ma AJ, Deng H, Ching JY, Wong VW, Tse YK, Yip TC, Lau LH, Liu HH, Leung CM, Tsang SW, Chan CW, Lau JY, Yuen PC, Chan FK. Machine learning model to predict recurrent ulcer bleeding in patients with history of idiopathic gastroduodenal ulcer bleeding. *Aliment Pharmacol Ther* 2019; **49**: 912-918 [PMID: 30761584 DOI: 10.1111/apt.15145]

26 **Zhang X**, Chen F, Yu T, An J, Huang Z, Liu J, Hu W, Wang L, Duan H, Si J. Real-time gastric polyp detection using convolutional neural networks. *PLoS One* 2019; **14**: e0214133 [PMID: 30908513 DOI: 10.1371/journal.pone.0214133]

27 **Iizuka O**, Kanavati F, Kato K, Rambeau M, Arihiro K, Tsuneki M. Deep Learning Models for

Histopathological Classification of Gastric and Colonic Epithelial Tumours. *Sci Rep* 2020; **10**: 1504 [PMID: 32001752 DOI: 10.1038/s41598-020-58467-9]

28 **Turan M**, Almalioglu Y, Araujo H, Konukoglu E, Sitti M. A non-rigid map fusion-based direct SLAM method for endoscopic capsule robots. *Int J Intell Robot Appl* 2017; **1**: 399-409 [PMID: 29250588 DOI: 10.1007/s41315-017-0036-4]

29 **Hwang Y**, Lee HH, Park C, Tama BA, Kim JS, Cheung DY, Chung WC, Cho YS, Lee KM, Choi MG, Lee S, Lee BI. Improved classification and localization approach to small bowel capsule endoscopy using convolutional neural network. *Dig Endosc* 2020 [PMID: 32640059 DOI: 10.1111/den.13787]

30 **Cheng PM**, Tejura TK, Tran KN, Whang G. Detection of high-grade small bowel obstruction on conventional radiography with convolutional neural networks. *Abdom Radiol (NY)* 2018; **43**: 1120-1127 [PMID: 28828625 DOI: 10.1007/s00261-017-1294-1]

31 **Cheng PM**, Tran KN, Whang G, Tejura TK. Refining Convolutional Neural Network Detection of Small-Bowel Obstruction in Conventional Radiography. *AJR Am J Roentgenol* 2019; **212**: 342-350 [PMID: 30476452 DOI: 10.2214/AJR.18.20362]

32 **Corley DA**, Levin TR, Doubeni CA. Adenoma detection rate and risk of colorectal cancer and death. *N Engl J Med* 2014; **370**: 2541 [PMID: 24963577 DOI: 10.1056/NEJMc1405329]

33 **Wei JW**, Suriawinata AA, Vaickus LJ, Ren B, Liu X, Lisovsky M, Tomita N, Abdollahi B, Kim AS, Snover DC, Baron JA, Barry EL, Hassanpour S. Evaluation of a Deep Neural Network for Automated Classification of Colorectal Polyps on Histopathologic Slides. *JAMA Netw Open* 2020; **3**: e203398 [PMID: 32324237 DOI: 10.1001/jamanetworkopen.2020.3398]

34 **Lui TKL**, Hui CKY, Tsui VWM, Cheung KS, Ko MKL, Foo DCC, Mak LY, Yeung CK, Lui TH, Wong SY, Leung WK. New insights on missed colonic lesions during colonoscopy through artificial intelligence-assisted real-time detection (with video). *Gastrointest Endosc* 2021; **93**: 193-200. e1 [PMID: 32376335 DOI: 10.1016/j.gie.2020.04.066]

35 **Waljee AK**, Wallace BI, Cohen-Mekelburg S, Liu Y, Liu B, Sauder K, Stidham RW, Zhu J, Higgins PDR. Development and Validation of Machine Learning Models in Prediction of Remission in Patients With Moderate to Severe Crohn Disease. *JAMA Netw Open* 2019; **2**: e193721 [PMID: 31074823 DOI: 10.1001/jamanetworkopen.2019.3721]

36 **Ozawa T**, Ishihara S, Fujishiro M, Saito H, Kumagai Y, Shichijo S, Aoyama K, Tada T. Novel computer-assisted diagnosis system for endoscopic disease activity in patients with ulcerative colitis. *Gastrointest Endosc* 2019; **89**: 416-421. e1 [PMID: 30367878 DOI: 10.1016/j.gie.2018.10.020]

37 **Stidham RW**, Liu W, Bishu S, Rice MD, Higgins PDR, Zhu J, Nallamothu BK, Waljee AK. Performance of a Deep Learning Model *vs* Human Reviewers in Grading Endoscopic Disease Severity of Patients With Ulcerative Colitis. *JAMA Netw Open* 2019; **2**: e193963 [PMID: 31099869 DOI: 10.1001/jamanetworkopen.2019.3963]

38 **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]

39 **Jisu Hong**, Bo-Yong Park, Hyunjin Park. Convolutional neural network classifier for distinguishing Barrett's esophagus and neoplasia endomicroscopy images. *Annu Int Conf IEEE Eng Med Biol Soc* 2017; **2017**: 2892-2895 [PMID: 29060502 DOI: 10.1109/EMBC.2017.8037461]

40 **Ebigbo A**, Mendel R, Probst A, Manzeneder J, Prinz F, de Souza LA Jr, Papa J, Palm C, Messmann H. Real-time use of artificial intelligence in the evaluation of cancer in Barrett's oesophagus. *Gut* 2020; **69**: 615-616 [PMID: 31541004 DOI: 10.1136/gutjnl-2019-319460]

41 **Sehgal V**, Rosenfeld A, Graham DG, Lipman G, Bisschops R, Ragunath K, Rodriguez-Justo M, Novelli M, Banks MR, Haidry RJ, Lovat LB. Machine Learning Creates a Simple Endoscopic Classification System that Improves Dysplasia Detection in Barrett's Oesophagus amongst Non-expert Endoscopists. *Gastroenterol Res Pract* 2018; **2018**: 1872437 [PMID: 30245711 DOI: 10.1155/2018/1872437]

42 **Zhang Y**, Li F, Yuan F, Zhang K, Huo L, Dong Z, Lang Y, Zhang Y, Wang M, Gao Z, Qin Z, Shen L. Diagnosing chronic atrophic gastritis by gastroscopy using artificial intelligence. *Dig Liver Dis* 2020; **52**: 566-572 [PMID: 32061504 DOI: 10.1016/j.dld.2019.12.146]

43 **Guimarães P**, Keller A, Fehlmann T, Lammert F, Casper M. Deep-learning based detection of gastric precancerous conditions. *Gut* 2020; **69**: 4-6 [PMID: 31375599 DOI: 10.1136/gutjnl-2019-319347]

44 **Horiuchi Y**, Aoyama K, Tokai Y, Hirasawa T, Yoshimizu S, Ishiyama A, Yoshio T, Tsuchida T, Fujisaki J, Tada T. Convolutional Neural Network for Differentiating Gastric Cancer from Gastritis Using Magnified Endoscopy with Narrow Band Imaging. *Dig Dis Sci* 2020; **65**: 1355-1363 [PMID: 31584138 DOI: 10.1007/s10620-019-05862-6]

45 **Lee JH**, Kim YJ, Kim YW, Park S, Choi YI, Park DK, Kim KG, Chung JW. Spotting malignancies from gastric endoscopic images using deep learning. *Surg Endosc* 2019; **33**: 3790-3797 [PMID: 30719560 DOI: 10.1007/s00464-019-06677-2]

46 **Namikawa K**, Hirasawa T, Nakano K, Ikenoyama Y, Ishioka M, Shiroma S, Tokai Y, Yoshimizu S, Horiuchi Y, Ishiyama A, Yoshio T, Tsuchida T, Fujisaki J, Tada T. Artificial intelligence-based diagnostic system classifying gastric cancers and ulcers: comparison between the original and newly developed systems. *Endoscopy* 2020; **52**: 1077-1083 [PMID: 32503056 DOI: 10.1055/a-1194-8771]

47 **Aoki T**, Yamada A, Aoyama K, Saito H, Tsuboi A, Nakada A, Niikura R, Fujishiro M, Oka S, Ishihara S, Matsuda T, Tanaka S, Koike K, Tada T. Automatic detection of erosions and ulcerations in

wireless capsule endoscopy images based on a deep convolutional neural network. *Gastrointest Endosc* 2019; **89**: 357-363. e2 [PMID: 30670179 DOI: 10.1016/j.gie.2018.10.027]

48 **Aoki T**, Yamada A, Aoyama K, Saito H, Fujisawa G, Odawara N, Kondo R, Tsuboi A, Ishibashi R, Nakada A, Niikura R, Fujishiro M, Oka S, Ishihara S, Matsuda T, Nakahori M, Tanaka S, Koike K, Tada T. Clinical usefulness of a deep learning-based system as the first screening on small-bowel capsule endoscopy reading. *Dig Endosc* 2020; **32**: 585-591 [PMID: 31441972 DOI: 10.1111/den.13517]

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

50 **Fan S**, Xu L, Fan Y, Wei K, Li L. Computer-aided detection of small intestinal ulcer and erosion in wireless capsule endoscopy images. *Phys Med Biol* 2018; **63**: 165001 [PMID: 30033931 DOI: 10.1088/1361-6560/aad51c]

51 **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]

52 **Tsuboi A**, Oka S, Aoyama K, Saito H, Aoki T, Yamada A, Matsuda T, Fujishiro M, Ishihara S, Nakahori M, Koike K, Tanaka S, Tada T. Artificial intelligence using a convolutional neural network for automatic detection of small-bowel angioectasia in capsule endoscopy images. *Dig Endosc* 2020; **32**: 382-390 [PMID: 31392767 DOI: 10.1111/den.13507]

53 **Rodriguez-Diaz E**, Baffy G, Lo WK, Mashimo H, Vidyarthi G, Mohapatra SS, Singh SK. Real-time artificial intelligence-based histologic classification of colorectal polyps with augmented visualization. *Gastrointest Endosc* 2021; **93**: 662-670 [PMID: 32949567 DOI: 10.1016/j.gie.2020.09.018]

54 **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]

55 **Akbari M**, Mohrekesh M, Nasr-Esfahani E, Soroushmehr SMR, Karimi N, Samavi S, Najarian K. Polyp Segmentation in Colonoscopy Images Using Fully Convolutional Network. *Annu Int Conf IEEE Eng Med Biol Soc* 2018; **2018**: 69-72 [PMID: 30440343 DOI: 10.1109/EMBC.2018.8512197]

56 **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]

57 **Gong D**, Wu L, Zhang J, Mu G, Shen L, Liu J, Wang Z, Zhou W, An P, Huang X, Jiang X, Li Y, Wan X, Hu S, Chen Y, Hu X, Xu Y, Zhu X, Li S, Yao L, He X, Chen D, Huang L, Wei X, Wang X, Yu H. Detection of colorectal adenomas with a real-time computer-aided system (ENDOANGEL): a randomised controlled study. *Lancet Gastroenterol Hepatol* 2020; **5**: 352-361 [PMID: 31981518 DOI: 10.1016/S2468-1253(19)30413-3]

58 **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]

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 **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]

61 **Taunk P**, Atkinson CD, Lichtenstein D, Rodriguez-Diaz E, Singh SK. Computer-assisted assessment of colonic polyp histopathology using probe-based confocal laser endomicroscopy. *Int J Colorectal Dis* 2019; **34**: 2043-2051 [PMID: 31696259 DOI: 10.1007/s00384-019-03406-y]

62 **Wang P**, Liu X, Berzin TM, Glissen Brown JR, Liu P, Zhou C, Lei L, Li L, Guo Z, Lei S, Xiong F, Wang H, Song Y, Pan Y, Zhou G. Effect of a deep-learning computer-aided detection system on adenoma detection during colonoscopy (CADe-DB trial): a double-blind randomised study. *Lancet Gastroenterol Hepatol* 2020; **5**: 343-351 [PMID: 31981517 DOI: 10.1016/S2468-1253(19)30411-X]

63 **Tong Y**, Lu K, Yang Y, Li J, Lin Y, Wu D, Yang A, Li Y, Yu S, Qian J. Can natural language processing help differentiate inflammatory intestinal diseases in China? *BMC Med Inform Decis Mak* 2020; **20**: 248 [PMID: 32993636 DOI: 10.1186/s12911-020-01277-w]

Artificial Intelligence in Gastrointestinal Endoscopy

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

Artif Intell Gastrointest Endosc 2021 April 28; 2(2): 36-49

DOI: 10.37126/aige.v2.i2.36

ISSN 2689-7164 (online)

MINIREVIEWS

Colonoscopy and artificial intelligence: Bridging the gap or a gap needing to be bridged?

James Weiquan Li, Tiing Leong Ang

ORCID number: James Weiquan Li 0000-0002-5241-4278; Tiing Leong Ang 0000-0001-9993-8549.

Author contributions: Li JW performed the literature search and wrote the manuscript; Ang TL performed the literature search and reviewed the manuscript.

Conflict-of-interest statement: The authors declare no conflict of interests for this article.

Manuscript source: Invited manuscript

Specialty type: Gastroenterology and hepatology

Country/Territory of origin: Singapore

James Weiquan Li, Tiing Leong Ang, Department of Gastroenterology and Hepatology, Changi General Hospital, Singapore 529889, Singapore

Corresponding author: Tiing Leong Ang, FASGE, FRCP, MBBS, Chief Physician, Professor, Department of Gastroenterology and Hepatology, Changi General Hospital, 2 Simei Street 3, Singapore 529889, Singapore. ang.tiing.leong@singhealth.com.sg

Abstract

Research in artificial intelligence (AI) in gastroenterology has increased over the last decade. Colonoscopy represents the most widely published field with regards to its use in gastroenterology. Most studies to date center on polyp detection and characterization, as well as real-time evaluation of adequacy of mucosal exposure for inspection. This review article discusses how advances in AI has bridged certain gaps in colonoscopy. In addition, the gaps formed with the development of AI that currently prevent its routine use in colonoscopy will be explored.

Key Words: Artificial intelligence; Endoscopy; Colonoscopy; Detection; Diagnosis

Core Tip: The use of artificial intelligence (AI) for colonoscopy has been studied most extensively for polyp detection and characterization. Despite advances made in this field, AI systems studied for these purposes represent only the machine learning domain of AI, and individual machine learning algorithms used in these studies are each focused on performing a very narrow task. While they may bridge existing gaps in polyp detection and real-time optical diagnosis of colorectal polyps, the introduction of AI into colonoscopy will also mean that there are new gaps that must be bridged for AI systems to be routinely used in clinical practice.

Citation: Li JW, Ang TL. Colonoscopy and artificial intelligence: Bridging the gap or a gap needing to be bridged? *Artif Intell Gastrointest Endosc* 2021; 2(2): 36-49

URL: https://www.wjgnet.com/2689-7164/full/v2/i2/36.htm

DOI: https://dx.doi.org/10.37126/aige.v2.i2.36

Peer-review report's scientific quality classification
Grade A (Excellent): 0
Grade B (Very good): 0
Grade C (Good): C
Grade D (Fair): 0
Grade E (Poor): 0

Received: March 19, 2021
Peer-review started: March 19, 2021
First decision: March 26, 2021
Revised: March 27, 2021
Accepted: April 20, 2021
Article in press: April 20, 2021
Published online: April 28, 2021

P-Reviewer: Viswanath YK
S-Editor: Wang JL
L-Editor: Filipodia
P-Editor: Wang LL

INTRODUCTION

The use of artificial intelligence (AI) in gastroenterology has gained momentum in the past decade. This is reflected in the increasing number of publications in the field of AI in endoscopy, most of which have been centered on colonoscopy. This is understandable as the unique role of colonoscopy in the prevention and management of colorectal cancer (CRC), together with the unmet needs in this field, has created the perfect milieu for the introduction of AI into world of endoscopy.

CRC represents one of the leading causes of cancer-related morbidity and mortality worldwide[1,2]. Colonoscopy decreases CRC-related mortality[3,4], with a 1% increase in adenoma detection rate (ADR) estimated to decrease interval CRC by 3%[5]. As such, a key barrier to overcome is the adenoma miss rate (AMR), which has been estimated in a meta-analysis to be a high as 22% overall, with a higher AMR when diminutive adenomas are considered[6]. Another unmet need in colonoscopy is the need for accuracy in the optical diagnosis of colonic polyps in relation to their actual histology. Up to 90% of lesions detected on colonoscopy consist of diminutive (≤ 5 mm) and small (6-9 mm) polyps, with the progression rates to advanced adenomas or CRC postulated to be low based on evidence from available studies[7]. It is therefore no surprise that most of the literature to date has focused on computer-assisted detection (CADe)[8,9] and computer-assisted diagnosis (CADx)[10-12] applications in colonoscopy.

This review article evaluates the areas in colonoscopy where AI may be a bridge for certain gaps in clinical practice. It will also explore in detail the current limitations and pitfalls in the application of AI in colonoscopy, highlighting how despite the proliferation of literature on this topic and what it promises to offer, AI may be a new gap in endoscopy which clinicians need to work to bridge.

LITERATURE SEARCH

We performed a comprehensive literature search in the PubMed, MEDLINE and EMBASE (up to March 17, 2021) electronic databases to identify relevant clinical trials that evaluated the roles of AI systems in colonoscopy. Electronic searches were also supplemented with manual searches of the references in the included studies and review articles.

AI TERMINOLOGY IN COLONOSCOPY

What does the term AI mean in colonoscopy?

The term "artificial intelligence" was first coined by John McCarthy in 1956 at the Dartmouth Summer Research Project. In essence, it is a branch in computer science where computer systems are designed to perform tasks which would ordinarily require human intelligence. This definition is extremely broad and often confuses clinicians to what exactly the capabilities, and by inference, the limitations of AI are in their respective fields[13]. There is therefore a need to define what AI means in colonoscopy as this is a prerequisite for meaningful discussion of its role in colonoscopy.

Published and ongoing studies incorporating AI in the context of colonoscopy involve the machine learning (ML) domain of AI. ML refers to the use of algorithms, which form predictive and descriptive models based on analysis of input data provided by investigators (the training set)[14]. These algorithms undergo multiple iterations of these models with the goal of performing a specific task, the aim of which is to come to a specified classification output (*e.g.*, polyp or no polyp) when the algorithms are tested on an unseen set of data (the test set). In practical terms and in the context of colonoscopy, this is achieved using either handcrafted models or deep learning (DL).

A useful mental model in understanding the scope of and roles which AI plays in colonoscopy is to regard the progress made in this field as "waves"[15]. It is crucial to understand that the methods, technologies, and results from earlier AI studies are not obsolete the moment a "better" or "faster" computer system is available based on results we as clinicians are familiar with such as the ADR and adenoma per colonoscopy (APC), or technical matrices that we may gravitate towards such as the processing speed of an algorithm. Rather, these "waves" are continuously interacting

and building on top of each other, and as a result, have a strong influence on the development of later technologies. The earlier "waves" remain relevant and may sometimes harbor solutions to certain issues faced with CADe and CADx support tools, which will be discussed later in this article. Having this mental model also helps us better understand the intrinsic biases present in all forms of ML regardless of advancements made in AI, which is essential for critical appraisal of literature surrounding AI in clinical practice.

AI terminology relevant to colonoscopy

Commonly used terms in AI which are relevant to this review article will be discussed here. This list is not meant to be exhaustive and is meant instead to highlight terms which will help the reader understand the later critiques and solutions offered in this paper.

AI can be categorized very broadly into weak (or narrow) AI and strong AI. The former refers to systems built to solve a specific problem or performing a single task extremely well, without an emphasis on elucidating how human reasoning works. This type of AI operates within significant constraints and a limited context. The latter term, also referred to as artificial general intelligence, aims to build systems which think like humans.

Features in ML refer to the set of numbers which quantitatively summarize and represent in a compact fashion the input data. For example, differences in morphology of polyps as defined in the Paris classification[16] and pit patterns[17] can be converted into different arrays of numbers which an ML algorithm can use to generate a prediction such as "polyp" or "no polyp" in a CADe application. Conventional learning by the ML algorithm may be supervised, where training takes place on labeled data sets, or unsupervised, where commonalities are used to identify groups within data. Supervised learning occurs on pre-established input and output pairs, enabling the ML algorithm to learn predictive mathematical models which can then map the input from unseen data into an outcome of interest (*e.g.*, neoplastic, or hyperplastic). In contrast, unsupervised learning predicts similarities between data points through looking at the underlying structure of the data provided, with no prior knowledge of its significance.

Handcrafted knowledge represents the first "wave" of AI. This consisted of knowledge-based methods where manual extraction and selection of characteristics of an object such as polyp shape and texture, are used to create mathematical models which can achieve a class or numerical output. This is labor-intensive and as a result, are usually implemented on small sets of data. These systems do not have the ability to learn and were of limited clinical use. DL is another form of ML where an artificial neural network (ANN) is used to perform the same task. ANNs are supervised ML models where interconnected artificial neurons form layered networks. Signals travel *via* weighted inputs from artificial neurons in the previous layer to the next layer, which then propagate the signal when a predefined threshold is reached, like how biological neurons work. Classification can be optimized, and the system enhanced by adjustment of the weights given to these inter-neuron connections.

Deep convolutional neural networks (DCNNs) have enabled more hidden layers to be added to the input and output layers of ANN, a development which has been facilitated by advancements made in other areas of computer science as this is computationally expansive. In addition, convolutional layers apply filters (a set of weights) in a systematic fashion to each overlapping part of the input data. In this manner, large numbers of filters can be applied to the training set of data in parallel under the constraints of the intended task, for example classification of an image as having a polyp or not in colonoscopy, allowing information to be extracted directly from images training data to form a feature map. DCNN usually require large amounts of labelled training data, which are derived wither from public databases or private collections in individual institutions.

Hyperparameters in ML refer to all parameters that have been arbitrarily set by the investigator and are used to configure the model for optimal performance at a specific task or on a specific dataset. As opposed to model parameters, which are learned automatically during training of the model, hyperparameters are manually set and affects the learning process and ultimately, the behavior of the model. This is useful in understanding the roles (and potential biases resulting from) the optimization and training process of AI models used in colonoscopy. The training set refers to the initial dataset used to determine optimal parameters after multiple rounds or iterations of adjustments. The validation set is mostly (but not always) a different dataset where these parameters are tested and adjusted. It is also used to optimize the hyperparameters in the model. Lastly, the test set refers to a new set of unseen data which is used

to test the model and its generalizability.

AI: BRIDGING THE GAP IN COLONOSCOPY

AI in the field of colonoscopy has been studied primarily for polyp detection, polyp characterization in terms of predicted histology, and for quality assurance in the performance of colonoscopy.

Polyp detection

The rate of missed polyps was mentioned earlier in the introduction. The AMR in influenced by different factors, among which the endoscopist is considered one of the major determinants[18-21]. These human biases may be due to distraction during colonoscopy, fatigue, or the inability to maintain a sustained level of alertness during withdrawal. These lead to errors in perception where the endoscopist may miss polyps which are visible on the monitor. The role of "second readers" in colonoscopy in increasing ADR[22,23] lends support to the hypothesis that CADe may help increase APC and ADR, and decrease AMR, during colonoscopy.

At the time of writing, there are six randomized controlled trials (RCTs)[24-29] to date that have evaluated the role of CADe in colonoscopy. Hassan *et al*[9] recently performed a systematic review and meta-analysis of five of these studies[24,25,27-29], which consisted of 4354 participants. The pooled ADR was significantly higher in the CADe group compared with the control group (36.6% *vs* 25.2%; relative risk [RR] 1.44; 95% confidence interval [CI]: 1.27-1.62; $P < 0.1$), with all of the included RCTs reporting a significant increase in ADR individually. APC, which is defined as the total number of adenomas found divided by the total number of colonoscopies and has good correlation with ADR[30,31], was also significantly higher in the CADe compared to the control group (0.58 *vs* 0.36; RR 1.70; 95%CI: 1.53-1.89; $P < 0.01$). The mean withdrawal time in the CADe and control groups was shown to be statistically different in this meta-analysis.

An interesting prospective study conducted by Wang *et al*[32] showed that the AMR was decreased with CADe. This study differed from the RCT mentioned above in that tandem colonoscopies were performed. Patients in this study were randomly assigned to colonoscopy with CADe or colonoscopy without CADe by an endoscopist, followed immediately by the other procedure. The study showed that the AMR and polyp miss rates were significantly lower in the CADe colonoscopy group compared to the routine colonoscopy group (13.89% *vs* 40.00%, $P < 0.0001$ and 12.98% *vs* 45.90%; $P < 0.0001$, respectively). These results were also consistent regardless of colonic segments, *i.e.* the AMR was significantly lower in the CADe group in the ascending, transverse, and descending colon.

Polyp characterization (optical prediction of polyp histology)

In contrast to CADe for polyp detection, CADx deals with the interpretation of polyp appearance during colonoscopy to determine the predicted histology. Polyp classification systems such as the Kudo pit pattern[17], Sano *et al*[33], NBI International Colorectal Endoscopic (NICE)[34], and Japan NBI Expert Team (JNET)[35] classifications were developed with the purpose of predicting polyp histology and severity of neoplasia to guide therapy. The use of these classification systems for optical prediction of colorectal polyp histology requires the proper equipment, structured training, and experience in clinical application. Studies have shown wide variation in the sensitivity and specificity of NICE and JNET classifications, with most studies reporting a moderate interobserver agreement at best[36-39].

With the clinical use of CADe, the detection of diminutive polyps is likely to increase exponentially, as demonstrated in the CADe RCT mentioned[24,25,27-29]. Most diminutive polyps tend to be hyperplastic in nature with low malignant potential. The "resect and discard" and "detect and leave" strategies for such polyps were previously studied to address these issues before the emergence of AI but have failed to gain traction due to the need for better quality training and quality assurance in the accurate optical diagnosis of colon polyps[40-42]. The threshold for optical biopsy technologies in high confidence predictions established by the American Society for Gastrointestinal Endoscopy (ASGE) Preservation and Incorporation of Valuable Endoscopic Innovations (PIVI)[43] are deemed appropriate targets for CADx support tools[44]. A systematic review and meta-analysis by ASGE[45] showed that these thresholds were met using NBI only among NBI experts, illustrating the difficulty and practical limitations of replying on the use of these forms of imaging by

endoscopists in general to achieve accurate optical diagnoses of colorectal polyps. Hence, this represents a significant clinical gap which AI has the potential to bridge in colonoscopy.

CADx is postulated to aid in this field of colorectal polyp management by using DL models to increase the accuracy of prediction of polyp histology during colonoscopy[46]. At the time of writing, there are currently no RCT evaluating CADx in colonoscopy. In a study by Jin *et al*[10], a DCNN was trained to differentiate between adenomatous and hyperplastic diminutive colorectal polyps with an overall accuracy of 86.7% using polyp histology as the gold standard. The system was tested on 22 endoscopists with varying expertise such as novice endoscopists, colonoscopy experts with differing levels of expertise in NBI, and NBI-trained experts. The use of CADx markedly improved the accuracy of novice endoscopists in differentiating adenomatous and hyperplastic polyps from 73.8% to 85.6% ($P < 0.05$), which was comparable to the baseline accuracy of NBI-trained experts (87.6%). However, in the colonoscopy expert and NBI-trained expert groups, this increase in accuracy was less impressive (83.8% to 89.0% and 87.6% to 90.0, respectively). The overall time to diagnosis per polyp was also decreased from 3.92 s to 3.37 s; $P = 0.42$).

A review of CADx predictions[47] for diminutive polyp histology which included 9 studies[48-56] showed a pooled sensitivity of 93.5% (95%CI: 90.7%-95.6%) and specificity of 90.8% (95%CI: 86.3%-95.9%), with a pooled area under the curve of 0.98. This pooled analysis of diminutive polyps had a negative predictive value (NPV) of 0.91 (95%CI: 0.89-0.94). This meets the 90% or greater threshold for NPV in adenomatous histology in rectosigmoid diminutive polyps recommended by the ASGE PIVI[43] and thus would in theory support a "diagnose and leave" strategy if these applications are validated in clinical use. However, most of these studies are retrospective in nature or, when conducted prospectively, involved the use of *ex vivo* video or still images.

Few prospective studies on CADx in real-time colonoscopy are currently available in the literature. In a single-center, open-label, prospective study of 791 consecutive patients undergoing colonoscopy in a university hospital, Mori *et al*[54] evaluated the performance of CADx in a clinical setting using endocytoscopy (CF-H290ECI; Olympus Corp, Tokyo, Japan). NBI was applied to visualize the microvascular pattern and methylene blue staining for cellular structure under these ultra-magnifying colonoscopes with 520X optical zoom capability. Of the 466 diminutive polyps found in this study, 250 polyps were in the rectosigmoid colon. The CADx system using endocytoscopy had an NPV for diminutive rectosigmoid adenomas ranging from 93.7% to 96.4% with methylene blue staining and 95.2% to 96.5% with NBI. This is well above the "diagnose and leave" threshold of 90% recommended by the ASGE PIVI[43] described. This prospective study also provides evidence for utilization of CADx for prediction of polyp histology in a clinical setting which may have an impact on decisions on polyp management real-time.

In an earlier study with a similar design by Horiuchi *et al*[56], CADx was evaluated with the use of autofluorescence imaging (AFI) to differentiate diminutive rectosigmoid polyps in real-time colonoscopies. The CADx system used software-based automatic color intensity analysis, which utilized AFI's ability to differentiate polyps based on the ratio of green to red tone intensities and was tested on 258 rectosigmoid polyps in 95 patients undergoing colonoscopy. The CAD-AFI system achieved an NPV for adenomatous polyps of 93.4% (95%CI: 89.0%-96.4%), which again exceeds the 90% "diagnose and leave" threshold[43]. In addition, the NPV using CAD-AFI was comparable to that of diagnoses made by endoscopists using AFI in the study (94.9%; 95%CI: 90.8%-97.5%).

Quality assurance in colonoscopy

Quality indices such as a high cecal intubation rate and adequate withdrawal time have been studied extensively[57,58]. However, these quality indices in colonoscopy performance and reporting are not always adhered to for a variety of factors such as training, lack of real-time feedback and failure of enforcement[59-61]. In an RCT of 704 patients by Gong *et al*[26], which used an AI system called ENDOANGEL, the withdrawal speed and time, as well as the adequacy of mucosal exposure, was monitored in real-time and in an automated fashion. The resulted in a significantly longer withdrawal time in the ENDOANGEL[62] *vs* the control group (mean 6.38 min *vs* 4.76 min, respectively; $P < 0.0001$). This translated into an increased ADR in the ENDOANGEL group and, more significantly, is the only RCT to date which demonstrates an AI system which can increase the rate of detection of adenomas 10 mm or larger in size (10/355 *vs* 1/349, respectively; odds ratio [OR] 9.50, 95%CI: 1.19-75.75; $P = 0.034$). Su *et al*[28] used both a CADe tool together with an automatic quality

control system (AQCS) to increase ADR and APC. The AQUS consisted of a timer on the monitor and audio prompts for the Endoscopist to slow down withdrawal speed when unstable and blurry frames were displayed or when the Boston Bowel Preparation Scale (BPPS) in a colonic segment was < 2. This study showed an improved withdrawal time (7.03 min *vs* 5.68 min; $P < 0.001$) and rate of adequate bowel preparation (87.34% *vs* 80.63%; $P = 0.023$) in the AQCS group in addition to the mentioned significant increase in ADR and APC.

AI: A GAP NEEDING TO BE BRIDGED IN COLONOSCOPY?

While AI has emerged in the world of endoscopy with much promise, there are several significant gaps which need to be bridged before it can be routinely applied in colonoscopy in a clinical setting.

Undefined and unspecified role in clinical environment

A major bridge which needs to be bridged before AI systems can be applied in routine environments is its generalizability. Three of the five CADe RCT[25,27,28] available involved senior endoscopists with extensive experience in colonoscopy. ADR is dependent on several factors, one of which includes experience. A more experienced endoscopist is not only skilled in recognition, but also in scope handling and consequent mucosal exposure during withdrawal. The role of a "second reader" in previous studies[22,23] in increasing small adenoma detection rates suggests that trainees and Nurses, who by inference have less "experience" than the senior endoscopist, have no issues recognizing a polyp visible on screen. In addition, as discussed in the ENDOANGEL study, one of the largest increments in ADR and the only increase in detection of adenomas larger then 10 mm was seen in the RCT by Gong *et al*[26], where real-time feedback on adequacy of mucosal exposure was studied. An obvious but less often mentioned fact is that any CADe algorithm is still completely dependent on the endoscopist to present optimal images with adequately exposed colonic mucosa in each real-time colonoscopy performed in a busy clinical setting. A polyp not visible on the screen will not be detected by a CADe tool, no matter how powerful the algorithm is[33]. This has implications on how generalizable available data is for clinical use, as more studies involving both "high detectors" and "low detectors" are required[25,63].

Most RCT in CADe to date were conducted in single centers. Moreover, except for the study by Wang *et al*[27] where a second monitor was used and visible only to an observer who reported the alerts, the rest of the RCT were non-blinded studies[24-26,28-29]. It is not known what the impact of the latter factor may be in actual clinical practice, as non-blinded endoscopists in these studies may put in more effort in exposing colonic mucosa for inspection when they are under observation. This Hawthorne Effect, together with the single-center experiences of most of these RCT, also limit their generalizability to routine clinical practice. While single monitors are encouraged[44] due to presumed gaze limitations of endoscopists and the need to reduce distractions, it is the opinion of the authors that a dual monitor setting in clinical trials plays a crucial role in achieving a double-blind and objective environment for assessment of the performance of the AI system and to bridge this gap. Furthermore, it resembles tandem colonoscopy in that the performance of the AI system can be compared directly against endoscopists of varying skill levels and experience. Useful information such as the AMR can be determined accurately without the patient having to go through an additional colonoscopy like in a traditional tandem study with this methodology.

Another limitation to the generalizability of the published results of AI systems for polyp detection and characterization is the differences in operational environments of different endoscopy suites and centers. These can vary greatly between institutions, even those located in the same country[64]. Unlike a new endoscopic method or classification system which can be taught or standardized in training or with major society guidelines, different AI algorithms have unique hardware and software requirements which must be fulfilled for technical integration into the operational environment. For instance, some may be fully integrated into the processing unit[65] while others may be web-based applications or require an additional laptop to be linked to the endoscopy stack to function. The latter may require cloud integration support, which in turn is likely to be vendor-specific and has implications in procurement and cybersecurity. This technical integration into the operational environment is key, as the development environment from which these AI systems are derived may be vastly

different[66]. Most clinical trials understandably focus on the clinical aspects like the ADR and APC and the outcomes will inevitably be based on these primary objectives. However, few studies have reported the technical specifications and limitations of the AI systems they are investigating. The rare studies that do report them, do so in varying details, most of which are insufficient for interpretation and contextualization into the operational environment. Moreover, most of the published trials have been conducted in academic or expert centers and in several instances, in the same institutions where the AI algorithm was developed, *i.e.* the development and operational environment are the same[3,47]. Individual institutions may have difficulty integrating these systems due to budgeting constraints, existence of legacy systems which are incompatible with the software and hardware requirements of the AI systems, logistical limitations such as space, and established workflows in endoscopy which does not cater to the introduction of an AI system.

The current scope of AI applications in colonoscopy in the literature is also largely skewed towards to polyp detection, characterization, and assessment of adequacy of mucosal exposure, which is ultimately linked to ADR. When translated to clinical practice, this effectively confines the indications for which AI should be used in colonoscopy to CRC screening or indications where one might expect to find colorectal polyps in the process of performing a colonoscopy. All systems developed in the field of AI in colonoscopy, from handcrafted models to the most complex DCNN, are fundamentally "weak AI." This is a term used to describe AI systems designed to solve a single problem or narrow task[15]. In a clinical setting, indications for colonoscopy are widely variable and the pre-test probability of finding of a polyp may be low. An endoscopist will be able to process the demographic data, clinical course, medical history, clinical condition, laboratory investigations and concerns of the patient and use this information during the colonoscopy. For example, an 85-year-old patient who is troubled by per rectal bleeding has a hugely different indication and clinical index of suspicion than a 50-year-old male with a family history of early CRC. In the former case, the endoscopist's focus may be on looking for angiodysplasia, diverticular disease or hemorrhoids as the etiology. A "strong AI" system would be able to think and adapt like a human and calibrate the weights in its layers to perform the task at hand, determine the appropriate classification output and achieve the correct alarm settings. However, current AI systems will continue looking for polyps and may present a distraction to the Endoscopist if used in this clinical example, prolonging the time taken for colonoscopy in an elderly patient, who may have multiple co-morbidities and for whom resection of small or diminutive adenomas may not have clinical relevance, much less answer the clinical question at hand. A trainee endoscopist or an experienced nurse, on the other hand, would be able to immediately recognize an unusual finding, such as multiple angiodysplasia or extensive diverticular disease, even if they were not formally trained to recognize these abnormalities.

It should be noted that AI has also been studied in colonoscopy outside the context of polyp detection, characterization, and quality assurance. Endocytoscopy has been used with AI to accurately detect persistent histologic inflammation in patients with ulcerative colitis (UC) which was reproducible based on static images[67]. A separate group used a deep neural network to predict endoscopic and histologic remission in UC patients based on evaluation of static images obtained from colonoscopy with high accuracy[68]. However, studies looking at indications other than polyp detection and characterization are few and far between.

Technical biases and lack of technical knowledge among clinicians

There is significant variability and a lack of standardization in reporting of the technical aspects of AI algorithms in clinical trials[69]. In addition, clinicians may not have the technical knowledge to critically appraise AI literature given that this has not been a formal part of training or an emphasis in clinical practice until relatively recently. A "minimum reporting standard" and practical knowledge of terms and potential biases on the part of investigators and clinicians, respectively, is required to bridge these gaps[70-72].

A practical knowledge of commonly used terms and how AI systems are derived is necessary for the clinician to appreciate the technical biases inherent to these algorithms. While the inclusion criteria of patients in clinical trials is clearly defined, the criteria for inclusion of the input data for the AI system during training and validation may not always be included in the methodology. This is crucial as most AI systems for CADe were tested in the same centers where they were developed[73]. This is often due to the ease with which large amounts of data are readily available for training and validation. Although the training, validation, and test datasets may be

different, they could be derived from the same database in a single, often expert, center, which is then split to form these datasets. The nature of the images used could be highly similar in terms of quality (*e.g.*, no confounding fecal material and bubbles and polyps always centered in the image) and labelling (*e.g.*, experts from different centers may mark out the most obvious abnormal area or delineate even the most minute detail which does not look like normal colonic mucosa for sessile serrated polyps depending on their level of skill and the training received, while experts from the same center are more likely to label lesions similarly). Prevalence and variability in presentations of disease may also differ depending on the populations studied, but the sample of images used in training and validating the AI algorithm may not necessarily reflect this natural variability of disease if data from a single center is used in the development of the AI system. This is a form of selection bias, as input data is not selected at random and hence is not fully representative of the study population in which the AI system is meant to function. This could impact the hyperparameters chosen during validation, and lead to overfitting, which occurs when the mathematical model derived is optimized to work on the training data and fits this data too tightly. This would limit its generalizability when new data is presented to the same AI algorithm.

Moreover, the proportion of "positive" to "normal" images used for training is not often mentioned in the published literature. For example, in a CADe application, polyps of various shapes, sizes and colors may be included in the training dataset to expose the AI algorithm to all possible eventualities when presented with an image with even the subtlest polyp. However, the "normal" images used may be disproportionately lower when compared to the natural prevalence of adenomas in the population. In addition, there may not be the same rigor in the selection of "normal" images for training. Variations in degrees of bowel preparation, bubbles, and artefacts due to the light source reflecting off normal colonic mucosa may thus not be reflected in images supplied to the AI algorithm for training. Positive and negative predictive values are determined by the prevalence of disease, and this may result in a higher proportion of false positives per true positive detected in clinical practice, depending on how the ratio of "positive" to "normal" images used in training compares with the true prevalence of the lesion of interest (*e.g.*, polyps) in the study population. This is a factor which needs to be adjusted for in the AI algorithm[74].

A certain form of publication bias may also exist as clinicians who wish to publish on the topic of AI will search for references almost exclusively from medical journals. For example, meta-analysis and systematic reviews on the use of AI in colonoscopy may take a very clinical slant, while publications in computer science and engineering journals which may add technical dept to the chosen topic on AI being discussed will not be included. Even if a search were performed for these articles, the inclusion criteria for the literature search will inevitably involve clinical-based endpoints like ADR and APC, and almost always exclude publications from computer science and engineering journals as a result. The barrier to entry in medical journals for these studies is high, as editors and reviewers, who themselves are clinicians, may not have enough technical knowledge to feel comfortable about accepting these articles for publication, and may also be compounded by fear of a lack of interest or understanding in the readership. On the other hand, AI and ML experts will not be familiar with the clinical aspects or relevance of their research and would not be able to pitch it at a level that would be acceptable to a Medical journal and its readership. This may result in a "reinforcement bias" of sorts, where only certain types of publications from a few expert centers and which revolve around common themes are published repeatedly and in different forms in Medical journals, whereas significant developments in AI and ML which may have the potential for changing clinical practice are missed out. The same technical terms specific to these publications will also be mentioned repeatedly, while novel approaches and new technical terms unfamiliar to clinicians may never see publication in a medical journal. The endoscopy readership may already have been "overfitted" towards polyp detection and characterization in the endoscopy literature[75], while neglecting the fact that, as mentioned, the use of AI in colonoscopy to date has utilized only an extremely limited aspect of AI and in a very narrow clinical context. Including computer science experts in the editorship and as reviewers for Medical journals may help to bridge the gap in these technical and publication biases.

Physician sentiment towards AI

Physician sentiment is a significant determinant on how quickly technologies and recommendations are deployed in a clinical setting. A recently conducted online survey among Gastroenterologists in the United States showed high overall interest in

CADe and perception that it would increase ADR (85.5% and 75.8%, respectively)[76]. However, the same survey also showed that majority of the respondents felt that CADe will prolong the time taken per colonoscopy, despite evidence to the contrary[9,24,25,27-29].

Concerns about operator dependence, or "deskilling", of the Endoscopist due to reliance on CADe and CADx for detection and characterization of polyps, respectively, are also mentioned in this survey[76] and other reviews[44,73]. Another major concern shown in the survey by Wadhwa *et al*[76] was the perceived increase in cost per procedure (75.2%). While concerns such as withdrawal time have been addressed independently in several RCT, others such as operator dependance and cost-effectiveness have not studied. Hence, physician sentiment may be another significant gap in AI which needs to be bridged in the field of colonoscopy.

Medicolegal challenges and future directions

AI algorithms which utilize DL are considered "black box" models, meaning that it is almost impossible to trace the decision-making process which led to the output determined by the algorithm when faced with a specific task (*e.g.*, polyp or no polyp in the image, hyperplastic or adenomatous). One of the major gaps in clinical use of AI systems in colonoscopy is medicolegal liability when a misdiagnosis or missed diagnosis occurs. While a clinician's account of events and the accompanying documentation can be helped up to scrutiny, the black box nature of DL algorithms means that the root cause and mitigating factors surrounding such a case may never be elucidated or even discovered. This has ethical implications in the event of harm to a patient[77], particularly if no clear protocol exists to define how an AI system should interface with its user and what its limits are, as the error may be due to deviation from safe use of the system or from an error of the AI system itself[78].

As AI systems, like other healthcare interventions, may have unpredictable errors, this inability to explain the errors or to detect them as they occur due to their black box nature may result in a perpetuation of systemic errors with unknown clinical implications if they are scaled up rapidly for routine clinical use in all colonoscopies. It is also unknown if the liability rests with the manufacturer, the regulatory body approving its use, or the clinician interfacing with the AI system. Having a reliable and accountable post-deployment surveillance plan is perhaps one of the strategies to minimize this risk.

Lastly, while AI systems have been shown to improve various quality indices associated with colonoscopy, one should remember that they are still limited most of all by our current expertise in this field. A useful example to illustrate this is the fact that there is currently no AI system capable of detecting dysplasia in UC. The availability of DCNN with high computing power and hardware to support the required processing speeds would have made this a rather simple task from an ML point of view. However, the optimal method of surveillance for dysplasia in UC and its optical features do not have the same clinical certainty as colorectal polyps in CRC screening, with resultant discrepancies in surveillance and biopsy practices[79,80]. Moreover, there is wide interobserver variability in the histological diagnosis of dysplasia in UC[81] and an inadequate understanding of its pathogenesis[82]. It is therefore understandable that there would be a paucity of expertly labelled data for "dysplasia" and "non-dysplasia" controls in UC patients for the training of an ML algorithm. Similarly, other potential AI applications in colonoscopy could include localization of diverticular bleeding and an automated scoring system for adequacy of bowel preparation which includes the BPPS[83] and the newly validated Colon Endoscopic Bubble Scale[84]. The clinical expertise and research in these fields must progress sufficiently for an accompanying increase in standardized and labelled data to be available for such future AI systems to be trained on and to materialize.

CONCLUSION

Despite the advances made in the field of AI, most notably for polyp detection and characterization in colonoscopy, there remain significant gaps which need to be bridged before its routine clinical use in colonoscopy.

REFERENCES

1 **Bray F**, Jemal A, Grey N, Ferlay J, Forman D. Global cancer transitions according to the Human Development Index (2008-2030): a population-based study. *Lancet Oncol* 2012; **13**: 790-801 [PMID: 22658655 DOI: 10.1016/S1470-2045(12)70211-5]

2 **Araghi M**, Soerjomataram I, Jenkins M, Brierley J, Morris E, Bray F, Arnold M. Global trends in colorectal cancer mortality: projections to the year 2035. *Int J Cancer* 2019; **144**: 2992-3000 [PMID: 30536395 DOI: 10.1002/ijc.32055]

3 **Kaminski MF**, Thomas-Gibson S, Bugajski M, Bretthauer M, Rees CJ, Dekker E, Hoff G, Jover R, Suchanek S, Ferlitsch M, Anderson J, Roesch T, Hultcranz R, Racz I, Kuipers EJ, Garborg K, East JE, Rupinski M, Seip B, Bennett C, Senore C, Minozzi S, Bisschops R, Domagk D, Valori R, Spada C, Hassan C, Dinis-Ribeiro M, Rutter MD. Performance measures for lower gastrointestinal endoscopy: a European Society of Gastrointestinal Endoscopy (ESGE) Quality Improvement Initiative. *Endoscopy* 2017; **49**: 378-397 [PMID: 28268235 DOI: 10.1055/s-0043-103411]

4 **Rex DK**, Schoenfeld PS, Cohen J, Pike IM, Adler DG, Fennerty MB, Lieb JG 2nd, Park WG, Rizk MK, Sawhney MS, Shaheen NJ, Wani S, Weinberg DS. Quality indicators for colonoscopy. *Gastrointest Endosc* 2015; **81**: 31-53 [PMID: 25480100 DOI: 10.1016/j.gie.2014.07.058]

5 **Corley DA**, Levin TR, Doubeni CA. Adenoma detection rate and risk of colorectal cancer and death. *N Engl J Med* 2014; **370**: 2541 [PMID: 24963577 DOI: 10.1056/NEJMc1405329]

6 **van Rijn JC**, Reitsma JB, Stoker J, Bossuyt PM, van Deventer SJ, Dekker E. Polyp miss rate determined by tandem colonoscopy: a systematic review. *Am J Gastroenterol* 2006; **101**: 343-350 [PMID: 16454841 DOI: 10.1111/j.1572-0241.2006.00390.x]

7 **Vleugels JLA**, Hazewinkel Y, Fockens P, Dekker E. Natural history of diminutive and small colorectal polyps: a systematic literature review. *Gastrointest Endosc* 2017; **85**: 1169-1176. e1 [PMID: 28024986 DOI: 10.1016/j.gie.2016.12.014]

8 **Barua I**, Vinsard DG, Jodal HC, Løberg M, Kalager M, Holme Ø, Misawa M, Bretthauer M, Mori Y. Artificial intelligence for polyp detection during colonoscopy: a systematic review and meta-analysis. *Endoscopy* 2021; **53**: 277-284 [PMID: 32557490 DOI: 10.1055/a-1201-7165]

9 **Hassan C**, Spadaccini M, Iannone A, Maselli R, Jovani M, Chandrasekar VT, Antonelli G, Yu H, Areia M, Dinis-Ribeiro M, Bhandari P, Sharma P, Rex DK, Rösch T, Wallace M, Repici A. Performance of artificial intelligence in colonoscopy for adenoma and polyp detection: a systematic review and meta-analysis. *Gastrointest Endosc* 2021; **93**: 77-85. e6 [PMID: 32598963 DOI: 10.1016/j.gie.2020.06.059]

10 **Jin EH**, Lee D, Bae JH, Kang HY, Kwak MS, Seo JY, Yang JI, Yang SY, Lim SH, Yim JY, Lim JH, Chung GE, Chung SJ, Choi JM, Han YM, Kang SJ, Lee J, Chan Kim H, Kim JS. Improved Accuracy in Optical Diagnosis of Colorectal Polyps Using Convolutional Neural Networks with Visual Explanations. *Gastroenterology* 2020; **158**: 2169-2179. e8 [PMID: 32119927 DOI: 10.1053/j.gastro.2020.02.036]

11 **Renner J**, Phlipsen H, Haller B, Navarro-Avila F, Saint-Hill-Febles Y, Mateus D, Ponchon T, Poszler A, Abdelhafez M, Schmid RM, von Delius S, Klare P. Optical classification of neoplastic colorectal polyps - a computer-assisted approach (the COACH study). *Scand J Gastroenterol* 2018; **53**: 1100-1106 [PMID: 30270677 DOI: 10.1080/00365521.2018.1501092]

12 **Kudo SE**, Mori Y, Misawa M, Takeda K, Kudo T, Itoh H, Oda M, Mori K. Artificial intelligence and colonoscopy: Current status and future perspectives. *Dig Endosc* 2019; **31**: 363-371 [PMID: 30624835 DOI: 10.1111/den.13340]

13 **Tang A**, Tam R, Cadrin-Chênevert A, Guest W, Chong J, Barfett J, Chepelev L, Cairns R, Mitchell JR, Cicero MD, Poudrette MG, Jaremko JL, Reinhold C, Gallix B, Gray B, Geis R; Canadian Association of Radiologists (CAR) Artificial Intelligence Working Group. Canadian Association of Radiologists White Paper on Artificial Intelligence in Radiology. *Can Assoc Radiol J* 2018; **69**: 120-135 [PMID: 29655580 DOI: 10.1016/j.carj.2018.02.002]

14 **Shalev-Shwartz S**, Ben-David S. Understanding Machine Learning: from theory to algorithms. New York: Cambridge University Press, 2014

15 **Wittenberg T**, Raithel M. Artificial Intelligence-Based Polyp Detection in Colonoscopy: Where Have We Been, Where Do We Stand, and Where Are We Headed? *Visc Med* 2020; **36**: 428-438 [PMID: 33447598 DOI: 10.1159/000512438]

16 Paris endoscopic classification of superficial neoplastic lesions: esophagus, stomach, and colon: November 30 to December 1, 2002. *Gastrointest Endosc* 2003; **58**: S3-43 [PMID: 14652541 DOI: 10.1016/s0016-5107(03)02159-x]

17 **Kudo S**, Tamura S, Nakajima T, Yamano H, Kusaka H, Watanabe H. Diagnosis of colorectal tumorous lesions by magnifying endoscopy. *Gastrointest Endosc* 1996; **44**: 8-14 [PMID: 8836710 DOI: 10.1016/s0016-5107(96)70222-5]

18 **Leufkens AM**, van Oijen MG, Vleggaar FP, Siersema PD. Factors influencing the miss rate of polyps in a back-to-back colonoscopy study. *Endoscopy* 2012; **44**: 470-475 [PMID: 22441756 DOI: 10.1055/s-0031-1291666]

19 **Rex DK**, Cutler CS, Lemmel GT, Rahmani EY, Clark DW, Helper DJ, Lehman GA, Mark DG. Colonoscopic miss rates of adenomas determined by back-to-back colonoscopies. *Gastroenterology* 1997; **112**: 24-28 [PMID: 8978338 DOI: 10.1016/s0016-5085(97)70214-2]

20 **Zhao S**, Wang S, Pan P, Xia T, Chang X, Yang X, Guo L, Meng Q, Yang F, Qian W, Xu Z, Wang Y, Wang Z, Gu L, Wang R, Jia F, Yao J, Li Z, Bai Y. Magnitude, Risk Factors, and Factors Associated

With Adenoma Miss Rate of Tandem Colonoscopy: A Systematic Review and Meta-analysis. *Gastroenterology* 2019; **156**: 1661-1674. e11 [PMID: 30738046 DOI: 10.1053/j.gastro.2019.01.260]

21 **Anderson R**, Burr NE, Valori R. Causes of Post-Colonoscopy Colorectal Cancers Based on World Endoscopy Organization System of Analysis. *Gastroenterology* 2020; **158**: 1287-1299. e2 [PMID: 31926170 DOI: 10.1053/j.gastro.2019.12.031]

22 **Buchner AM**, Shahid MW, Heckman MG, Diehl NN, McNeil RB, Cleveland P, Gill KR, Schore A, Ghabril M, Raimondo M, Gross SA, Wallace MB. Trainee participation is associated with increased small adenoma detection. *Gastrointest Endosc* 2011; **73**: 1223-1231 [PMID: 21481861 DOI: 10.1016/j.gie.2011.01.060]

23 **Lee CK**, Park DI, Lee SH, Hwangbo Y, Eun CS, Han DS, Cha JM, Lee BI, Shin JE. Participation by experienced endoscopy nurses increases the detection rate of colon polyps during a screening colonoscopy: a multicenter, prospective, randomized study. *Gastrointest Endosc* 2011; **74**: 1094-1102 [PMID: 21889137 DOI: 10.1016/j.gie.2011.06.033]

24 **Wang P**, Berzin TM, Glissen Brown JR, Bharadwaj S, Becq A, Xiao X, Liu P, Li L, Song Y, Zhang D, Li Y, Xu G, Tu M, Liu X. Real-time automatic detection system increases colonoscopic polyp and adenoma detection rates: a prospective randomised controlled study. *Gut* 2019; **68**: 1813-1819 [PMID: 30814121 DOI: 10.1136/gutjnl-2018-317500]

25 **Repici A**, Badalamenti M, Maselli R, Correale L, Radaelli F, Rondonotti E, Ferrara E, Spadaccini M, Alkandari A, Fugazza A, Anderloni A, Galtieri PA, Pellegatta G, Carrara S, Di Leo M, Craviotto V, Lamonaca L, Lorenzetti R, Andrealli A, Antonelli G, Wallace M, Sharma P, Rosch T, Hassan C. Efficacy of Real-Time Computer-Aided Detection of Colorectal Neoplasia in a Randomized Trial. *Gastroenterology* 2020; **159**: 512-520. e7 [PMID: 32371116 DOI: 10.1053/j.gastro.2020.04.062]

26 **Gong D**, Wu L, Zhang J, Mu G, Shen L, Liu J, Wang Z, Zhou W, An P, Huang X, Jiang X, Li Y, Wan X, Hu S, Chen Y, Hu X, Xu Y, Zhu X, Li S, Yao L, He X, Chen D, Huang L, Wei X, Wang X, Yu H. Detection of colorectal adenomas with a real-time computer-aided system (ENDOANGEL): a randomised controlled study. *Lancet Gastroenterol Hepatol* 2020; **5**: 352-361 [PMID: 31981518 DOI: 10.1016/S2468-1253(19)30413-3]

27 **Wang P**, Liu X, Berzin TM, Glissen Brown JR, Liu P, Zhou C, Lei L, Li L, Guo Z, Lei S, Xiong F, Wang H, Song Y, Pan Y, Zhou G. Effect of a deep-learning computer-aided detection system on adenoma detection during colonoscopy (CADe-DB trial): a double-blind randomised study. *Lancet Gastroenterol Hepatol* 2020; **5**: 343-351 [PMID: 31981517 DOI: 10.1016/S2468-1253(19)30411-X]

28 **Su JR**, Li Z, Shao XJ, Ji CR, Ji R, Zhou RC, Li GC, Liu GQ, He YS, Zuo XL, Li YQ. Impact of a real-time automatic quality control system on colorectal polyp and adenoma detection: a prospective randomized controlled study (with videos). *Gastrointest Endosc* 2020; **91**: 415-424. e4 [PMID: 31454493 DOI: 10.1016/j.gie.2019.08.026]

29 **Liu WN**, Zhang YY, Bian XQ, Wang LJ, Yang Q, Zhang XD, Huang J. Study on detection rate of polyps and adenomas in artificial-intelligence-aided colonoscopy. *Saudi J Gastroenterol* 2020; **26**: 13-19 [PMID: 31898644 DOI: 10.4103/sjg.SJG_377_19]

30 **Kahi CJ**, Vemulapalli KC, Johnson CS, Rex DK. Improving measurement of the adenoma detection rate and adenoma per colonoscopy quality metric: the Indiana University experience. *Gastrointest Endosc* 2014; **79**: 448-454 [PMID: 24246797 DOI: 10.1016/j.gie.2013.10.013]

31 **Hilsden RJ**, Bridges R, Dube C, McGregor SE, Naugler C, Rose SM, Rostom A, Heitman SJ. Defining Benchmarks for Adenoma Detection Rate and Adenomas Per Colonoscopy in Patients Undergoing Colonoscopy Due to a Positive Fecal Immunochemical Test. *Am J Gastroenterol* 2016; **111**: 1743-1749 [PMID: 27725649 DOI: 10.1038/ajg.2016.449]

32 **Wang P**, Liu P, Glissen Brown JR, Berzin TM, Zhou G, Lei S, Liu X, Li L, Xiao X. Lower Adenoma Miss Rate of Computer-Aided Detection-Assisted Colonoscopy vs Routine White-Light Colonoscopy in a Prospective Tandem Study. *Gastroenterology* 2020; **159**: 1252-1261. e5 [PMID: 32562721 DOI: 10.1053/j.gastro.2020.06.023]

33 **Sano Y**, Ikematsu H, Fu KI, Emura F, Katagiri A, Horimatsu T, Kaneko K, Soetikno R, Yoshida S. Meshed capillary vessels by use of narrow-band imaging for differential diagnosis of small colorectal polyps. *Gastrointest Endosc* 2009; **69**: 278-283 [PMID: 18951131 DOI: 10.1016/j.gie.2008.04.066]

34 **Rex DK**. Narrow-band imaging without optical magnification for histologic analysis of colorectal polyps. *Gastroenterology* 2009; **136**: 1174-1181 [PMID: 19187781 DOI: 10.1053/j.gastro.2008.12.009]

35 **Sano Y**, Tanaka S, Kudo SE, Saito S, Matsuda T, Wada Y, Fujii T, Ikematsu H, Uraoka T, Kobayashi N, Nakamura H, Hotta K, Horimatsu T, Sakamoto N, Fu KI, Tsuruta O, Kawano H, Kashida H, Takeuchi Y, Machida H, Kusaka T, Yoshida N, Hirata I, Terai T, Yamano HO, Kaneko K, Nakajima T, Sakamoto T, Yamaguchi Y, Tamai N, Nakano N, Hayashi N, Oka S, Iwatate M, Ishikawa H, Murakami Y, Yoshida S, Saito Y. Narrow-band imaging (NBI) magnifying endoscopic classification of colorectal tumors proposed by the Japan NBI Expert Team. *Dig Endosc* 2016; **28**: 526-533 [PMID: 26927367 DOI: 10.1111/den.12644]

36 **Hewett DG**, Kaltenbach T, Sano Y, Tanaka S, Saunders BP, Ponchon T, Soetikno R, Rex DK. Validation of a simple classification system for endoscopic diagnosis of small colorectal polyps using narrow-band imaging. *Gastroenterology* 2012; **143**: 599-607. e1 [PMID: 22609383 DOI: 10.1053/j.gastro.2012.05.006]

37 **Repici A**, Ciscato C, Correale L, Bisschops R, Bhandari P, Dekker E, Pech O, Radaelli F, Hassan C. Narrow-band Imaging International Colorectal Endoscopic Classification to predict polyp histology: REDEFINE study (with videos). *Gastrointest Endosc* 2016; **84**: 479-486. e3 [PMID: 26928372 DOI:

10.1016/j.gie.2016.02.020]
38 **Komeda Y**, Kashida H, Sakurai T, Asakuma Y, Tribonias G, Nagai T, Kono M, Minaga K, Takenaka M, Arizumi T, Hagiwara S, Matsui S, Watanabe T, Nishida N, Chikugo T, Chiba Y, Kudo M. Magnifying Narrow Band Imaging (NBI) for the Diagnosis of Localized Colorectal Lesions Using the Japan NBI Expert Team (JNET) Classification. *Oncology* 2017; **93** Suppl 1: 49-54 [PMID: 29258091 DOI: 10.1159/000481230]
39 **Kobayashi S**, Yamada M, Takamaru H, Sakamoto T, Matsuda T, Sekine S, Igarashi Y, Saito Y. Diagnostic yield of the Japan NBI Expert Team (JNET) classification for endoscopic diagnosis of superficial colorectal neoplasms in a large-scale clinical practice database. *United European Gastroenterol J* 2019; **7**: 914-923 [PMID: 31428416 DOI: 10.1177/2050640619845987]
40 **Kandel P**, Wallace MB. Should We Resect and Discard Low Risk Diminutive Colon Polyps. *Clin Endosc* 2019; **52**: 239-246 [PMID: 30661337 DOI: 10.5946/ce.2018.136]
41 **Neumann H**, Neumann Sen H, Vieth M, Bisschops R, Thieringer F, Rahman KF, Gamstätter T, Tontini GE, Galle PR. Leaving colorectal polyps in place can be achieved with high accuracy using blue light imaging (BLI). *United European Gastroenterol J* 2018; **6**: 1099-1105 [PMID: 30228899 DOI: 10.1177/2050640618769731]
42 **von Renteln D**, Kaltenbach T, Rastogi A, Anderson JC, Rösch T, Soetikno R, Pohl H. Simplifying Resect and Discard Strategies for Real-Time Assessment of Diminutive Colorectal Polyps. *Clin Gastroenterol Hepatol* 2018; **16**: 706-714 [PMID: 29174789 DOI: 10.1016/j.cgh.2017.11.036]
43 **Rex DK**, Kahi C, O'Brien M, Levin TR, Pohl H, Rastogi A, Burgart L, Imperiale T, Ladabaum U, Cohen J, Lieberman DA. The American Society for Gastrointestinal Endoscopy PIVI (Preservation and Incorporation of Valuable Endoscopic Innovations) on real-time endoscopic assessment of the histology of diminutive colorectal polyps. *Gastrointest Endosc* 2011; **73**: 419-422 [PMID: 21353837 DOI: 10.1016/j.gie.2011.01.023]
44 **Vinsard DG**, Mori Y, Misawa M, Kudo SE, Rastogi A, Bagci U, Rex DK, Wallace MB. Quality assurance of computer-aided detection and diagnosis in colonoscopy. *Gastrointest Endosc* 2019; **90**: 55-63 [PMID: 30926431 DOI: 10.1016/j.gie.2019.03.019]
45 **ASGE Technology Committee**. , Abu Dayyeh BK, Thosani N, Konda V, Wallace MB, Rex DK, Chauhan SS, Hwang JH, Komanduri S, Manfredi M, Maple JT, Murad FM, Siddiqui UD, Banerjee S. ASGE Technology Committee systematic review and meta-analysis assessing the ASGE PIVI thresholds for adopting real-time endoscopic assessment of the histology of diminutive colorectal polyps. *Gastrointest Endosc* 2015; **81**: 502.e1-502. e16 [PMID: 25597420 DOI: 10.1016/j.gie.2014.12.022]
46 **Song EM**, Park B, Ha CA, Hwang SW, Park SH, Yang DH, Ye BD, Myung SJ, Yang SK, Kim N, Byeon JS. Endoscopic diagnosis and treatment planning for colorectal polyps using a deep-learning model. *Sci Rep* 2020; **10**: 30 [PMID: 31913337 DOI: 10.1038/s41598-019-56697-0]
47 **Lui TKL**, Guo CG, Leung WK. Accuracy of artificial intelligence on histology prediction and detection of colorectal polyps: a systematic review and meta-analysis. *Gastrointest Endosc* 2020; **92**: 11-22. e6 [PMID: 32119938 DOI: 10.1016/j.gie.2020.02.033]
48 **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]
49 **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]
50 **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]
51 **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]
52 **Lui TKL**, To WP, Ko KK. Superiority of the artificial intelligence image classifier for histological prediction of diminutive colorectal polyps based on non-magnifying endoscopic images. *Hong Kong Med J* 2019; **25**: 31
53 **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]
54 **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]
55 **Kudo SE**, Misawa M, Mori Y, Hotta K, Ohtsuka K, Ikematsu H, Saito Y, Takeda K, Nakamura H, Ichimasa K, Ishigaki T, Toyoshima N, Kudo T, Hayashi T, Wakamura K, Baba T, Ishida F, Inoue H, Itoh H, Oda M, Mori K. Artificial Intelligence-assisted System Improves Endoscopic Identification of

Colorectal Neoplasms. *Clin Gastroenterol Hepatol* 2020; **18**: 1874-1881. e2 [PMID: 31525512 DOI: 10.1016/j.cgh.2019.09.009]

56 **Horiuchi H**, Tamai N, Kamba S, Inomata H, Ohya TR, Sumiyama K. Real-time computer-aided diagnosis of diminutive rectosigmoid polyps using an auto-fluorescence imaging system and novel color intensity analysis software. *Scand J Gastroenterol* 2019; **54**: 800-805 [PMID: 31195905 DOI: 10.1080/00365521.2019.1627407]

57 **Mathews SC**, Zhao N, Holub JL, Lieberman D. Improvement in colonoscopy quality metrics in clinical practice from 2000 to 2014. *Gastrointest Endosc* 2019; **90**: 651-655. e3 [PMID: 31207221 DOI: 10.1016/j.gie.2019.06.004]

58 **Clark BT**, Rustagi T, Laine L. What level of bowel prep quality requires early repeat colonoscopy: systematic review and meta-analysis of the impact of preparation quality on adenoma detection rate. *Am J Gastroenterol* 2014; **109**: 1714-23; quiz 1724 [PMID: 25135006 DOI: 10.1038/ajg.2014.232]

59 **de Jonge V**, Sint Nicolaas J, Cahen DL, Moolenaar W, Ouwendijk RJ, Tang TJ, van Tilburg AJ, Kuipers EJ, van Leerdam ME; SCoPE Consortium. Quality evaluation of colonoscopy reporting and colonoscopy performance in daily clinical practice. *Gastrointest Endosc* 2012; **75**: 98-106 [PMID: 21907986 DOI: 10.1016/j.gie.2011.06.032]

60 **Leyden JE**, Doherty GA, Hanley A, McNamara DA, Shields C, Leader M, Murray FE, Patchett SE, Harewood GC. Quality of colonoscopy performance among gastroenterology and surgical trainees: a need for common training standards for all trainees? *Endoscopy* 2011; **43**: 935-940 [PMID: 21997723 DOI: 10.1055/s-0030-1256633]

61 **Coe SG**, Panjala C, Heckman MG, Patel M, Qumseya BJ, Wang YR, Dalton B, Tran P, Palmer W, Diehl N, Wallace MB, Raimondo M. Quality in colonoscopy reporting: an assessment of compliance and performance improvement. *Dig Liver Dis* 2012; **44**: 660-664 [PMID: 22579446 DOI: 10.1016/j.dld.2012.03.022]

62 **Zhou J**, Wu L, Wan X, Shen L, Liu J, Zhang J, Jiang X, Wang Z, Yu S, Kang J, Li M, Hu S, Hu X, Gong D, Chen D, Yao L, Zhu Y, Yu H. A novel artificial intelligence system for the assessment of bowel preparation (with video). *Gastrointest Endosc* 2020; **91**: 428-435. e2 [PMID: 31783029 DOI: 10.1016/j.gie.2019.11.026]

63 **Hassan C**, Wallace MB, Sharma P, Maselli R, Craviotto V, Spadaccini M, Repici A. New artificial intelligence system: first validation study versus experienced endoscopists for colorectal polyp detection. *Gut* 2020; **69**: 799-800 [PMID: 31615835 DOI: 10.1136/gutjnl-2019-319914]

64 **Day LW**, Bhuket T, Inadomi JM, Yee HF. Diversity of endoscopy center operations and practice variation across California's safety-net hospital system: a statewide survey. *BMC Res Notes* 2013; **6**: 233 [PMID: 23767938 DOI: 10.1186/1756-0500-6-233]

65 **Weigt J**, Repici A, Antonelli G, Afifi A, Kliegis L, Correale L, Hassan C, Neumann H. Performance of a new integrated CADe/CADx system for detection and characterization of colorectal neoplasia. *Endoscopy* 2021 [PMID: 33494106 DOI: 10.1055/a-1372-0419]

66 **Kelly CJ**, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. *BMC Med* 2019; **17**: 195 [PMID: 31665002 DOI: 10.1186/s12916-019-1426-2]

67 **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]

68 **Takenaka K**, Ohtsuka K, Fujii T, Negi M, Suzuki K, Shimizu H, Oshima S, Akiyama S, Motobayashi M, Nagahori M, Saito E, Matsuoka K, Watanabe M. Development and Validation of a Deep Neural Network for Accurate Evaluation of Endoscopic Images From Patients With Ulcerative Colitis. *Gastroenterology* 2020; **158**: 2150-2157 [PMID: 32060000 DOI: 10.1053/j.gastro.2020.02.012]

69 **Nagendran M**, Chen Y, Lovejoy CA, Gordon AC, Komorowski M, Harvey H, Topol EJ, Ioannidis JPA, Collins GS, Maruthappu M. Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. *BMJ* 2020; **368**: m689 [PMID: 32213531 DOI: 10.1136/bmj.m689]

70 **Collins GS**, Moons KGM. Reporting of artificial intelligence prediction models. *Lancet* 2019; **393**: 1577-1579 [PMID: 31007185 DOI: 10.1016/S0140-6736(19)30037-6]

71 **CONSORT-AI and SPIRIT-AI Steering Group**. Reporting guidelines for clinical trials evaluating artificial intelligence interventions are needed. *Nat Med* 2019; **25**: 1467-1468 [PMID: 31551578 DOI: 10.1038/s41591-019-0603-3]

72 **Rivera SC**, Liu X, Chan AW, Denniston AK, Calvert MJ; SPIRIT-AI and CONSORT-AI Working Group. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI Extension. *BMJ* 2020; **370**: m3210 [PMID: 32907797 DOI: 10.1136/bmj.m3210]

73 **Mori Y**, Neumann H, Misawa M, Kudo SE, Bretthauer M. Artificial intelligence in colonoscopy - Now on the market. What's next? *J Gastroenterol Hepatol* 2021; **36**: 7-11 [PMID: 33179322 DOI: 10.1111/jgh.15339]

74 **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]

75 **Ang TL**, Carneiro G. Artificial intelligence in gastrointestinal endoscopy. *J Gastroenterol Hepatol* 2021; **36**: 5-6 [PMID: 33448513 DOI: 10.1111/jgh.15344]

76 **Wadhwa V**, Alagappan M, Gonzalez A, Gupta K, Brown JRG, Cohen J, Sawhney M, Pleskow D, Berzin TM. Physician sentiment toward artificial intelligence (AI) in colonoscopic practice: a survey of US gastroenterologists. *Endosc Int Open* 2020; **8**: E1379-E1384 [PMID: 33015341 DOI: 10.1055/a-1223-1926]

77 **Habli I**, Lawton T, Porter Z. Artificial intelligence in health care: accountability and safety. *Bull World Health Organ* 2020; **98**: 251-256 [PMID: 32284648 DOI: 10.2471/BLT.19.237487]

78 **Wiens J**, Saria S, Sendak M, Ghassemi M, Liu VX, Doshi-Velez F, Jung K, Heller K, Kale D, Saeed M, Ossorio PN, Thadaney-Israni S, Goldenberg A. Do no harm: a roadmap for responsible machine learning for health care. *Nat Med* 2019; **25**: 1337-1340 [PMID: 31427808 DOI: 10.1038/s41591-019-0548-6]

79 **Singh K**, Al Khoury A, Kurti Z, Gonczi L, Reinglas J, Verdon C, Kohen R, Bessissow T, Afif W, Wild G, Seidman E, Bitton A, Lakatos PL. High Adherence to Surveillance Guidelines in Inflammatory Bowel Disease Patients Results in Low Colorectal Cancer and Dysplasia Rates, While Rates of Dysplasia are Low Before the Suggested Onset of Surveillance. *J Crohns Colitis* 2019; **13**: 1343-1350 [PMID: 30918959 DOI: 10.1093/ecco-jcc/jjz066]

80 **Watanabe T**, Ajioka Y, Mitsuyama K, Watanabe K, Hanai H, Nakase H, Kunisaki R, Matsuda K, Iwakiri R, Hida N, Tanaka S, Takeuchi Y, Ohtsuka K, Murakami K, Kobayashi K, Iwao Y, Nagahori M, Iizuka B, Hata K, Igarashi M, Hirata I, Kudo SE, Matsumoto T, Ueno F, Watanabe G, Ikegami M, Ito Y, Oba K, Inoue E, Tomotsugu N, Takebayashi T, Sugihara K, Suzuki Y, Watanabe M, Hibi T. Comparison of Targeted vs Random Biopsies for Surveillance of Ulcerative Colitis-Associated Colorectal Cancer. *Gastroenterology* 2016; **151**: 1122-1130 [PMID: 27523980 DOI: 10.1053/j.gastro.2016.08.002]

81 **Allende D**, Elmessiry M, Hao W, DaSilva G, Wexner SD, Bejarano P, Berho M, Al-Qadasi M. Inter-observer and intra-observer variability in the diagnosis of dysplasia in patients with inflammatory bowel disease: correlation of pathological and endoscopic findings. *Colorectal Dis* 2014; **16**: 710-8; discussion 718 [PMID: 24836541 DOI: 10.1111/codi.12667]

82 **Beaugerie L**, Itzkowitz SH. Cancers complicating inflammatory bowel disease. *N Engl J Med* 2015; **372**: 1441-1452 [PMID: 25853748 DOI: 10.1056/NEJMra1403718]

83 **Lai EJ**, Calderwood AH, Doros G, Fix OK, Jacobson BC. The Boston bowel preparation scale: a valid and reliable instrument for colonoscopy-oriented research. *Gastrointest Endosc* 2009; **69**: 620-625 [PMID: 19136102 DOI: 10.1016/j.gie.2008.05.057]

84 **Taveira F**, Hassan C, Kaminski MF, Ponchon T, Benamouzig R, Bugajski M, de Castelbajac F, Cesaro P, Chergui H, Goran L, Minelli Grazioli L, Janičko M, Januszewicz W, Lamonaca L, Lenz J, Negreanu L, Repici A, Spada C, Spadaccini M, State M, Szlak J, Veseliny E, Dinis-Ribeiro M, Areia M. The Colon Endoscopic Bubble Scale (CEBuS): a two-phase evaluation study. *Endoscopy* 2020 [PMID: 33285583 DOI: 10.1055/a-1331-4325]

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