Artificial intelligence in gastrointestinal endoscopy: a comprehensive review

Hassam Ali[a], Muhammad Ali Muzammil[b], Dushyant Singh Dahiya[c], Farishta Ali[d], Shafay Yasin[e], Waqar Hanif[e], Manesh Kumar Gangwani[f], Muhammad Aziz[g], Muhammad Khalaf[a], Debargha Basuli[h], Mohammad Al-Haddad[i]

ECU Health Medical Center/Brody School of Medicine, Greenville, North Carolina, USA; Dow University of Health Sciences, Sindh, PK; The University of Kansas School of Medicine, Kansas City, Kansas, USA; Khyber Girls Medical College, Peshawar, PK; Quaid-e-Azam Medical College, Punjab, PK; University of Toledo Medical Center, Toledo, OH, USA; East Carolina University/Brody School of Medicine, Greenville, North Carolina, USA; Indiana University School of Medicine, Indianapolis, IN, USA

Abstract

Integrating artificial intelligence (AI) into gastrointestinal (GI) endoscopy heralds a significant leap forward in managing GI disorders. AI-enabled applications, such as computer-aided detection and computer-aided diagnosis, have significantly advanced GI endoscopy, improving early detection, diagnosis and personalized treatment planning. AI algorithms have shown promise in the analysis of endoscopic data, critical in conditions with traditionally low diagnostic sensitivity, such as indeterminate biliary strictures and pancreatic cancer. Convolutional neural networks can markedly improve the diagnostic process when integrated with cholangioscopy or endoscopic ultrasound, especially in the detection of malignant biliary strictures and cholangiocarcinoma. AI's capacity to analyze complex image data and offer real-time feedback can streamline endoscopic procedures, reduce the need for invasive biopsies, and decrease associated adverse events. However, the clinical implementation of AI faces challenges, including data quality issues and the risk of overfitting, underscoring the need for further research and validation. As the technology matures, AI is poised to become an indispensable tool in the gastroenterologist's arsenal, necessitating the integration of robust, validated AI applications into routine clinical practice. Despite remarkable advances, challenges such as operator-dependent accuracy and the need for intricate examinations persist. This review delves into the transformative role of AI in enhancing endoscopic diagnostic accuracy, particularly highlighting its utility in the early detection and personalized treatment of GI diseases.

Keywords Medical imaging, gastroenterology, artificial intelligence, machine learning, deep learning

Ann Gastroenterol 2024; 37 (2): 133-141

Correspondence to: Hassam Ali, MD, Department of Gastroenterology and Hepatology, ECU Health Medical Center/Brody School of Medicine, Greenville, North Carolina, 27834, USA, e-mail: alih20@ecu.edu

Introduction

Gastrointestinal (GI) endoscopy, an integral tool in gastroenterology, enables real-time imaging and treatment of various GI disorders. Its integration with artificial intelligence (AI) holds immense potential for augmenting endoscopy's capabilities, enhancing patient care and outcomes. Early detection and personalized treatment are key in managing GI diseases, with GI cancers representing a significant portion of the global cancer incidence and mortality [1]. Enhanced endoscopic performance and quality screening are vital in reducing these statistics [2]. While GI endoscopy has advanced substantially, challenges remain, such as operator-dependent accuracy [3] and the need for detailed examinations. Personalized

treatments, considering individual patient differences, are essential [4]. AI's ability to analyze large volumes of endoscopic data aids in detecting pathologies, including complex GI and liver cancer histopathology [5]. Early AI applications in GI endoscopy include tumor detection, staging, lesion pathology prediction, and identifying *Helicobacter pylori* (*H. pylori*) infections [6]. AI also supports personalized treatment by assimilating patient-specific data, including genetic profiles, to develop tailored treatment plans [7]. One other challenge is the accurate diagnosis of indeterminate biliary strictures (IDBS) and pancreatic cancer. Conventional methods, such as endoscopic retrograde cholangiopancreatography (ERCP) with brushing and biopsy, often yield suboptimal results, necessitating a multi-disciplinary approach that includes advanced endoscopic techniques, such as cholangioscopy and endoscopic ultrasound (EUS). This review explores AI-driven techniques in transforming gastroenterology, emphasizing early detection and personalized treatment.

AI in GI endoscopy

The past decade has witnessed significant advances in the application of AI in GI endoscopy. Early AI systems primarily used computer-aided diagnosis (CADx) and computer-aided detection (CADe) algorithms. Their role was to identify, characterize and differentiate suspicious lesions, such as polyps, tumors, ulcers, and areas of dysplasia [8]. These AI systems deployed predetermined algorithms to pinpoint deviations from "normal" tissue, presenting them to the interpreter. In GI endoscopy, the processing speed of CAD algorithms and their visual output, i.e., graphics, are paramount—initially, computational power was CAD's main limitation. But the rise of faster, more affordable, portable and reliable digital platforms has improved CAD's efficiency [9]. Simultaneously, there have been notable advances in machine learning and deep learning (DL) methods, paving the way for more refined AI models. These AI/machine learning models are instrumental in the prevention, diagnosis, management and prognosis of various GI lesions, encompassing both premalignant and malignant types.

aDepartment of Gastroenterology and Hepatology, ECU Health Medical Center/Brody School of Medicine, Greenville, North Carolina, USA (Hassam Ali, Muhammad Khalaf); bDepartment of Internal Medicine, Dow University of Health Sciences, Sindh, PK (Muhammad Ali Muzammil); cDivision of Gastroenterology, Hepatology & Motility, The University of Kansas School of Medicine, Kansas City, Kansas, USA (Dushyant Singh Dahiya); dDepartment of Internal Medicine, Khyber Girls Medical College, Peshawar, PK (Farishta Ali); eDepartment of Internal Medicine, Quaid-e-Azam Medical College, Punjab, PK (Shafay Yasin, Waqar Hanif); fDepartment of Medicine, University of Toledo Medical Center, Toledo, OH, USA (Manesh Kumar Gangwani); gDepartment of Gastroenterology and Hepatology, The University of Toledo Medical Center, Toledo, OH, USA (Muhammad Aziz); hDepartment of Internal Medicine, East Carolina University/Brody School of Medicine, Greenville, North Carolina, USA (Debargha Basuli); iDivision of Gastroenterology and Hepatology, Indiana University School of Medicine, Indianapolis, IN, USA (Mohammad Al-Haddad)

Machine learning allows AI algorithms to function like the human brain, refining its performance with every data interaction without being explicitly programmed to do so. DL, emerging as one of the most promising machine learning subsets, empowers AI to reason, discern intricate patterns, establish connections, and make decisions using raw data. Its structure resembles the human neural network, with an input layer, multiple hidden layers, and an output layer. Thus, DL offers significantly enhanced capabilities compared to its AI predecessors [10]. In supervised learning, techniques such as support vector machines, regression analysis, and random forests prevail.

Conversely, unsupervised learning adopts principal component and cluster analysis methods to detect recurring patterns within datasets. Based on these learning paradigms, algorithms emulate the human brain's function, leading to artificial neural networks under the machine learning umbrella. DL, a machine learning subset, harnesses multiple artificial neural networks to analyze datasets and predict outcomes directly. Convolutional neural networks (CNN) are another prominent neural network in medical imaging. Their interconnected layers reflect the workings of the human visual cortex [11].

In gastroenterology, a spectrum of imaging modalities is utilized to evaluate the digestive system and pinpoint GI tumors. These include histologic sections of GI specimens, radiography, and endoscopy. While machine learning is a go-to choice for the automated analysis of GI images, gastroenterology is open to exploring other AI branches, such as natural language processing [12,13]. There may often be a synergy between radiomic techniques and machine learning/DL algorithms. The term 'radiomics', introduced in 2012, denotes the computational, objective analysis of features within radiologic images, shedding light on insights typically elusive or challenging to quantify. A fundamental component of radiomic evaluations revolves around the analysis of image texture attributes [14]. Radiomics and AI share a symbiotic relationship, especially since the frequent deployment of machine learning strategies in radiomics for pattern identification in extensive datasets (Fig. 1) [15].

Esophageal disorders

AI has emerged as a transformative tool for investigating Barrett's disease (BE) and esophageal cancer. A recent study by Dumoulin *et al* has demonstrated the potential of AI systems in enhancing the screening and surveillance of BE and early esophageal adenocarcinoma, providing a nuanced approach to high-quality endoscopy and addressing the challenges of esophageal cancer in western societies [16]. Concurrently, advances in imaging technology, integrated with AI, are revolutionizing the detection and characterization of esophageal neoplastic lesions and diagnosing early esophageal squamous-cell carcinoma and precancerous lesions [17]. Innovations in surveillance methods, as explored by Iyer *et al*, including AI-powered approaches, are making strides in identifying and predicting dysplasia/adenocarcinoma in BE, with
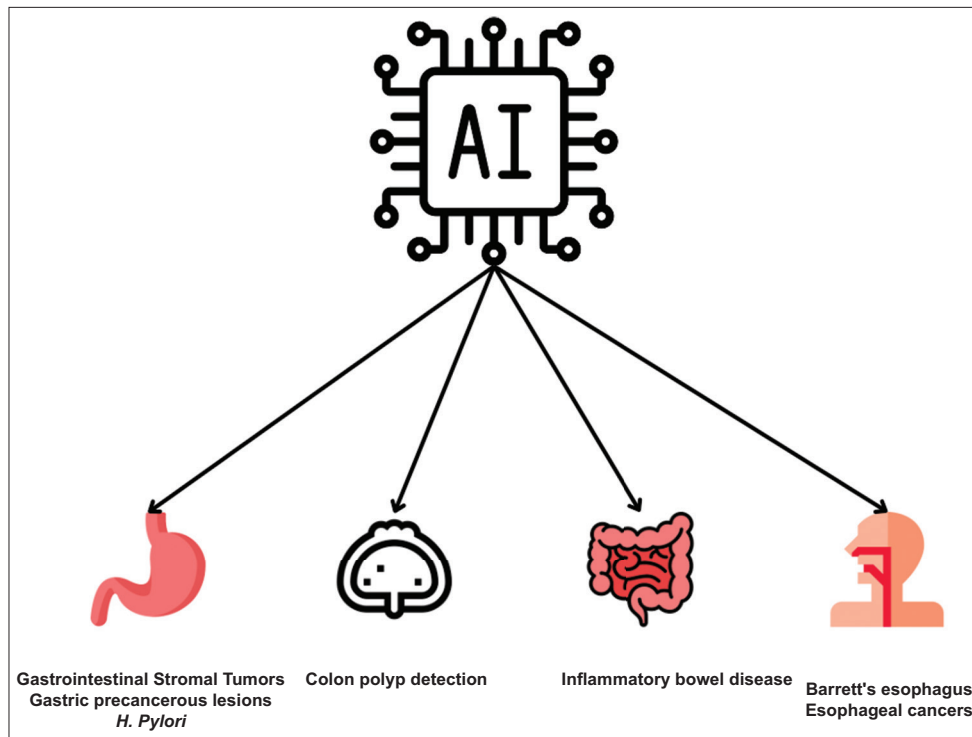
Figure 1 Overview of AI applications in gastrointestinal disorders
AI, artificial intelligence; H. pylori, Helicobacter pylori

emergent innovations spanning efficient sampling methods, advanced imaging tools and molecular marker-powered approaches [18]. As discussed by Vulpoi *et al*, DL models show promise in diagnosing and managing upper digestive tract diseases, including BE. They may also help diagnose and manage gastroesophageal reflux disease [19]. Furthermore, the establishment of AI models like YOLOv5l, as reported by Wang *et al*, is assisting in diagnosing esophageal squamous-cell carcinoma and precancerous lesions, thereby enhancing diagnostic accuracy and reducing missed cases, which can assist junior endoscopists in improving diagnosis [17].

Gastric precancerous lesions and *H. pylori* infection

Gastric precancerous lesions and *H. pylori* infection are recognized risk factors for gastric cancer (GC). In addressing these concerns, the rise of AI, backed by its computational prowess and learning capacities, has been marked as a potent asset.

AI algorithms have notably boosted diagnostic precision for GI conditions such as early-stage GC. A study by Tang *et al* is has highlighted AI's diagnostic supremacy over endoscopists, particularly in detecting early GC [20]. This emphasizes how the judicious deployment of AI in GI endoscopy can equip clinicians with the means to identify lesions at earlier stages, elevate diagnostic accuracy, and personalize treatments—all converging towards optimized therapeutic results for the patient.

Diving into the predictive abilities of AI in the realm of endoscopy, Huang *et al* probed its capability to anticipate

H. pylori gastritis [21]. Their findings charted a sensitivity and specificity of 85.4% and 90.9%, respectively, for *H. pylori* detection. Furthermore, this study shed light on AI's commendable accuracy in predicting gastric atrophy, intestinal metaplasia, and the intensity of *H. pylori*-linked gastritis. Most observations hover above 80% accuracy, although no specific study provided an exact figure. The evolution of this methodology could trim down the resources expended on needless biopsies for *H. pylori* diagnosis. Supporting this perspective, a meta-analysis by Bang *et al*, consolidating 8 studies and data from 1719 patients, reported that AI attained a sensitivity of 0.87 (95% confidence interval [CI] 0.72-0.94) and a specificity of 0.86 (95%CI 0.77-0.92) in detecting *H. pylori* infections [22]. Additionally, AI exhibited an 82% accuracy in differentiating non-infected gastric mucosa images from their post-eradication counterparts [22].

Exploring this domain further, a distinct study encompassing school children in Ethiopia employed machine learning to discern risk factors and predict *H. pylori* infection on the basis of those risks [23]. The outcomes spotlighted the efficiency of machine learning tactics, such as the XGBoost classifier, in forecasting *H. pylori* infection status—a performance that surpassed traditional statistical models.

GC

Previous studies have reported that the miss rates for GC detection by endoscopists ranged from 4.6-25.8% [24].

Misses in GC detection during endoscopy are influenced by tumor characteristics and the endoscopists' level of experience. A definitive diagnosis of GC currently relies on visual examination of whole slide imaging (WSI) pathological images, which pose challenges for pathologists, necessitating long-term concentration and tedious efforts. While advanced image-enhanced endoscopy techniques show potential in improving GC detection [25], their widespread implementation is hindered by the need for additional training and expertise. AI has the potential to assist in automatic, precise, and rapid endoscopic detection, as well as histopathological examination, addressing these challenges. In GC detection, AI-based image classification has achieved a sensitivity of 92.2% in classifying endoscopic images [26].

In the field of GC diagnosis, traditional machine-learning methods have typically relied on the extraction of handcrafted features from images. For instance, Miyaki *et al* [27] employed densely sampled scale-invariant feature transform descriptors as local features, and applied a hierarchical k-means clustering algorithm to obtain a bag-of-features representation. They combined this with a support vector machine-based classifier to achieve a cancer detection sensitivity of 84.8% and specificity of 87.0%, using a cutoff value of 0.59 [28]. However, these approaches faced limitations, including reliance on human-selected cutoff values and small data sets. DL has emerged as a powerful tool to address these challenges, enabling end-to-end training that automates feature extraction and classification. Hirasawa *et al* [26] explored the use of a CNN model, specifically a single-shot multi-box detector, for detecting GC lesions. The CNN model achieved a sensitivity of 92.2% for automatic GC diagnosis, outperforming traditional machine-learning methods. Luo *et al* [29] developed a GI AI diagnostic system (GRAIDS) based on DL techniques, using a large dataset of endoscopic images. GRAIDS demonstrated a performance similar to expert endoscopists and showed potential in improving the detection of gastric tumors, particularly for trainee endoscopists and low-volume hospitals.

DL approaches have been utilized in GC detection on WSI. Most successful approaches extract and use small image patches instead of the whole image as input. Patch selection is a key research area, and existing approaches can be categorized based on whether they employ patch- or slide-level annotations. Li *et al* [30] proposed GastricNet, a GC detection model on WSI. The model combines shallow and deep layers to extract multi-scale features and achieve slice-level prediction by averaging the scores of the top 10 patches with the highest cancer probability. Wang *et al* [31] developed a method called recalibrated multi-instance DL (RMDL) for detecting GC in WSIs. They used a 2-stage framework, where a ResNet-based network was used to generate discriminative features and abnormal probabilities for each patch instance, and then a local-global feature fusion method and attention-based approach were used to aggregate the instance features. The RMDL method outperformed other multi-instance learning algorithms in classifying gastric slices into cancer, dysplasia and normal states. Machine learning has successfully diagnosed and guided treatment decisions for patients with GI stromal tumors, as well using integrated radiomics analysis [32-37].

Colorectal polyps and cancer

AI-assisted computer vision applications, particularly CADe and CADx, have emerged as critical tools for prevention of colorectal cancer (CRC) through colonoscopy screening and monitoring [30]. These technological advances are not without their limitations, as adenoma miss rates (AMR) range between 6% and 28% [38], mainly because of challenges in lesion detection influenced by operator expertise. The pressing need to overcome such obstacles has led to a surge in the validation and development of these tools, notably for AI-assisted polyp detection studies. Here, outcome measures such as adenoma detection rate, adenomas per colonoscopy, and AMR serve as pivotal assessment parameters.

Although most studies are still in their preliminary stages, they exhibit promising clinical relevance. One recent study showcased a DL-based AI system's significant improvement in real-time polyp detection rates [39]. Supporting this, a systematic review and a meta-analysis found that CADe usage during colonoscopy boosted the detection rates of adenomas and polyps compared to a control cohort [40]. In a randomized controlled trial, CADe was found to reduce AMR and miss rates for sessile-serrated lesions, and increase the count of adenomas per colonoscopy during the initial sweep [41].

Recent advances in CADe, particularly demonstrated by a systematic review and meta-analysis by Hassan *et al*, encompassed 21 randomized trials involving 18,232 patients. This study showed that the adenoma detection rate was notably higher in the CADe group (44.0%) compared to the standard colonoscopy group (35.9%), indicating a 55% relative reduction in the AMR [42]. Another systematic review and a meta-analysis found that CADe usage during colonoscopy boosted the detection rates of adenomas and polyps compared to a control cohort [43]. In a randomized controlled trial, CADe was found to reduce AMR and miss rates of sessile-serrated lesions and increase the count of adenomas per colonoscopy [44].

AI-driven CNN models could refine bowel preparation quality and enhance polyp detection [45]. Moreover, real-time polyp identification systems rooted in DL algorithms can potentially improve polyp detection and localization during colonoscopy sessions [46,47]. Notably, incorporating AI into routine colonoscopy can halve the AMR and augment tumor detection, even for those under 10 mm, as highlighted in a recent global study [48].

On the other hand, in tumor disease diagnosis and radiation treatment planning, manual tumor boundary delineation is time-consuming, requires expertise, and suffers from interobserver variability. However, fully convolutional networks, such as SegNet, U-Net, and DenseVNet, have proven to be efficient frameworks for semantic image segmentation. U-Net, a specialized CNN, and DenseVNet, a fully convolutional network, enable precise and quick biomedical

image segmentation [49-51]. T2-weighted, dynamic contrast enhanced and multiparametric magnetic resonance imaging, and computed tomography-based multi-organ segmentation methods, often paired with multi-atlas label fusion techniques, yield more reproducible treatment area measurements, leading to more consistent radiation doses [52]. Some notable trials of AI models in colonic polyps and CRC detection are described in Table 1 [45-62].

Supplementing the above illustrations, AI's potential to significantly refine endoscopy spans various avenues:

Real-time support: AI stands ready to offer instantaneous aid to endoscopists, parsing live endoscopic imagery with a velocity that outstrips human cognition. It can highlight dubious territories, guiding pivotal decision making [63].

Elevated clinician efficacy: By automating tasks like image captures, data extractions, and report generation, AI amplifies clinician efficiency during endoscopies [64].

Superior image clarity: AI enhances endoscopic image fidelity by bolstering contrast, attenuating ambient noise, and discerning nuanced alterations [65].

Inflammatory bowel disease (IBD)

AI has found its use in diagnosing and managing IBD. A recent systematic review highlighted the promising technical results of AI-assisted endoscopy in IBD, emphasizing its growth as a research field and its additional benefits in experimental clinical scenarios [66]. The integration of AI with capsule endoscopy, as illustrated by Mascarenhas *et al*, is enhancing the future of IBD management, particularly in the investigation of obscure hemorrhagic lesions [67]. The current advancement of AI in diagnosing IBD from imaging data highlights its importance in the coming decades [68]. The feasibility of accurately predicting adverse outcomes using complex and novel AI models on large longitudinal data sets of patients with IBD has been demonstrated by Zand *et al*, suggesting potential applications for risk stratification and implementation of preemptive measures [69]. Furthermore, an overview by Takenaka *et al* provides insights into how AI can improve clinical practice in endoscopy for IBD, with some components already beginning to shape our understanding [70]. The proposed 2-tier IBDN classification, as evaluated by Ang *et al*, helps detect the "CRC/high-grade dysplasia" group for consideration of proctocolectomy. However, it does not differentiate low-grade dysplasia or sporadic adenomas from normal mucosa [71]. A recent study has taken a significant step toward enhancing dysplasia detection in patients with IBD. The researchers developed the first known model for CADe of colorectal lesions in patients with IBD. This model was retrained using high-definition white-light endoscopy (HDWLE) images and dye-based chromoendoscopy images of IBD-associated colorectal lesions. The retrained IBD-CADe model demonstrated impressive performance metrics, including a sensitivity of 95.1% and an accuracy of 96.8% for HDWLE. The study concluded that this model represents the first step toward developing AI-based endoscopic tools to enhance the detection of polypoid and non-polypoid dysplasia, potentially reducing CRC rates in patients with IBD. The ability to detect lesions as small as 5 mm with 93% sensitivity further emphasizes the potential of this AI model in early detection and intervention [72].

Pancreato-biliary diseases

Diagnosing IDBS and pancreatic cancer poses significant clinical challenges. The differentiation between malignant and benign biliary strictures remains difficult, despite advances in endoscopic techniques. Conventional ERCP techniques, involving cholangiogram appearance and brush cytology, often fail to differentiate reliably between malignant and benign biliary strictures because of their limited sensitivity [73,74]. Probe-based confocal laser endomicroscopy (pCLE) represents a novel imaging technique offering "optical biopsies" and providing *in vivo* cellular-level architectural information.

Table 1 Studies demonstrating use of AI in colon cancer and polyp detection

Author, year [ref.]	Key findings
Zhang *et al*, 2023 [53]	Showed potential in distinguishing malignant from benign lesions
Gong *et al*, 2023 [54]	Addressed challenges in differentiating foreground and background in polyp images
Nogueira-Rodríguez *et al*, 2023 [55]	Highlighted the importance of including negative samples to reduce false positives
Sadagopan *et al*, 2023 [56]	Focused on identifying polyps of various sizes and shapes
Gilabert *et al*, 2022 [57]	Developed a multi-platform web application with CNN for polyp detection
González-Bueno Puyal *et al*, 2022 [58]	Enhanced polyp detection using deep learning approaches
Lu *et al*, 2022 [45]	Focused on bowel preparation quality and polyp detection rate
Adjei *et al*, 2022 [59]	Explored the impact of synthetic data on training AI models for polyp detection
Tanwar *et al*, 2022 [60]	Developed a system for detecting and classifying colorectal polyps
Nogueira-Rodríguez *et al*, 2022 [61]	Assessed CNN performance for polyp localization
Jheng *et al*, 2022 [62]	Developed an AI algorithm for detecting colon polyps and identifying normal colon landmarks

AI, artificial intelligence; CNN, convoluted neural networks

Table 2 Recent trials of use of AI in the diagnosis of malignant biliary strictures and cholangiocarcinoma

Author (year) [ref.]	Image type	Pathology	Classifier	Performance metrics (accuracy, sensitivity, specificity, etc.)
Mascarenhas *et al* (2021) [75]	Cholangioscopy	Malignant biliary strictures	CNN	Accuracy: 95%, Sensitivity: 94.7%, Specificity: 92.1%, Precision: 0.6, F-1 Score: 0.7
Ribeiro *et al* (2021) [76]	Cholangioscopy	Intraductal papillary projections	CNN	Accuracy: 99%, Sensitivity: 99.7%, Specificity: 97.1%
Marya *et al* (2022) [74]	Cholangioscopy	Malignant biliary strictures	CNN (ResNet50V2)	Accuracy: 91%, Sensitivity: 93.3%, Specificity: 88.2%
Yao *et al* (2021) [77]	EUS	Bile duct lesion/ CCA	CNN (ResNet)	Accuracy: 90%, Sensitivity: 89.5%, Specificity: 82.4%

AI, artificial intelligence; EUS, endoscopic ultrasound; CCA, cholangiocarcinoma; CNN, convoluted neural networks

Table 3 Overview of key limitations of AI in digestive endoscopy

Limitation	Description
Insufficient temporal context consideration	AI algorithms may not effectively consider the temporal context of medical conditions, which could affect the accuracy of predictions
Single-source data collection	If all training and test datasets are gathered from a single source, there is a risk of limiting diversity and introducing bias in the AI's ability to identify and categorize abnormalities
Generalizability	To assess the true effectiveness of any AI system, ongoing, stringent external validation is required, using various imaging sources and techniques
Limited usage of CNN models	There is a need to expand the usage of CNN algorithms to include other currently available models, such as GoogLeNet and a few ResNet, to improve the AI's effectiveness
Need for high-quality datasets	A significant challenge in developing AI systems is the need for high-quality datasets, which persists in current clinical practice
Inclusion of low-quality images in datasets	Many datasets may include low-quality images, which can increase the likelihood of an AI system making a mistaken diagnosis or prediction
Inherent algorithm bias	AI can exhibit inherent biases based on the datasets it's trained on, which could affect the accuracy of diagnosis across diverse population sets
Risk of overfitting	Overfitting can limit the AI's performance on new, unseen data, affecting its effectiveness in practical applications
Lack of interpretability	AI's "black box" nature can make it difficult to interpret its decisions, potentially challenging trust and improvement based on feedback
Vulnerability to image manipulation	AI algorithms can be susceptible to perturbations in input data, with alterations such as cropping, rotating, or adding noise potentially affecting the output
Static post-training models	Post-training, AI models typically do not improve or learn from new data unless re-trained, which can prevent the system from adapting to new forms or presentations of medical conditions
Limited contextual understanding	AI may lack the ability to understand the context crucial in medical diagnostics
Dependence on image quality and preparation	The performance of AI can heavily depend on the quality and preparation of images, which can vary due to differences in equipment, preparation protocols, and staining methods
Economic and logistic constraints	Implementing AI in clinical practice can require substantial investment and personnel training, posing a challenge in resource-limited settings
Ethical and privacy concerns	Significant ethical and privacy issues are associated with using patient data for AI training, particularly in ensuring the anonymization and security of patient data

AI, artificial intelligence; CNN, convoluted neural networks

However, its specificity can be compromised by previous endoscopic interventions. Cholangioscopy, especially with the advent of single-operator systems, has significantly advanced the diagnostic capabilities for biliary strictures. Endoscopic-ultrasound-guided fine-needle aspiration (EUS-FNA) has been shown to be more effective than conventional ERCP techniques

in diagnosing malignant biliary strictures, with higher sensitivity and accuracy [73,74]. However, it is crucial to note that negative results from these techniques do not completely exclude malignancy, given their low negative predictive value. AI can play a transformative role in this context. AI algorithms can assist in the analysis of complex imaging data obtained from these advanced endoscopic techniques. By integrating AI with pCLE, cholangioscopy and EUS-FNA data, it may be possible to enhance the diagnostic accuracy for IDBS and pancreatic cancer. This integration could lead to more reliable differentiation between malignant and benign strictures, better informing treatment decisions, and improving patient outcomes [73,74].

Cholangiocarcinoma is notoriously difficult to diagnose early, since traditional techniques like ERCP have limited sensitivity [73]. With its ability to process and analyze complex image data at an unprecedented scale, AI presents a viable solution to this challenge. Recent studies have shown that CNNs, when used in conjunction with cholangioscopy or EUS, can significantly improve the accuracy of diagnosing malignant biliary strictures and cholangiocarcinoma. For instance, CNN with EUS imaging has demonstrated superior clinical performance, potentially streamlining the diagnostic process and providing real-time, reliable feedback during endoscopic procedures. However, the application of AI in this context is not without its challenges [73]. Issues such as data quality, inconsistency, and the risk of overfitting need to be carefully managed to ensure the reliability of AI algorithms in clinical settings. Moreover, while the initial results are promising, further comparative studies and external validations are essential to firmly establish the role of AI in the routine diagnostic process. The potential of AI to reduce procedure length, enhance diagnostic accuracy and possibly negate the need for invasive testing, such as biopsies, is significant, particularly for reducing procedure-associated adverse events [73]. Some recent trials using AI in pancreato-biliary disorders are described in Table 2 [74-77].

Limitations

The field of medical diagnostics has a great deal of potential for AI, but there are restrictions on how it can be used. One of the major limitations is that we cannot completely rely on AI; clinical judgment is still essential, as demonstrated by inability of algorithms to take contextual information into account. Computational errors may also be made worse by the issue of automation bias, which occurs when clinicians accept the AI's conclusions even when they are incorrect. Another obstacle is the need for more clarity in the data collection process. All the training and test datasets were derived from a single source, restricting diversity and potentially introducing bias in the AI's diagnosis capabilities. Additionally, the AI's ability to generalize its learned knowledge to new, unseen data is challenging, necessitating continuous and stringent external validation using various imaging and endoscopy sources. Furthermore, there is a need to expand the use of CNN algorithms to include currently available models, such as

GoogLeNet and a few ResNet, to improve the effectiveness of the AI. Additional limitations are described in Table 3. These limitations underscore the need for cautious optimism and continuous development in this promising field.

Concluding remarks

Endoscopy is central to gastroenterology, and AI's integration with various forms of endoscopy, including white-light imaging, linked color imaging and blue laser imaging, holds immense promise. AI can be trained to recognize distinct clinical markers in gastric mucosa, enhancing early detection and diagnosis. The use of AI in predicting treatment outcomes and personalizing interventions, especially in gastroenterological conditions, is an active research area. Challenges such as data standardization, ownership and protection must be addressed, and federated learning may offer a solution for privacy-preserving data mining. Despite the higher labor and storage costs, creating a globally shared database is essential for advancing AI in gastroenterology. Future studies should focus on clinically relevant applications, large real-world datasets, and improving the "explainability" of AI results to clinicians and patients. Collaboration between researchers, doctors, and AI professionals will be vital to fully realize the potential of AI and multimodal imaging in gastroenterology, leading to transformative advances in surveillance, diagnosis and therapeutic interventions.

References

1. Arnold M, Abnet CC, Neale RE, et al. Global burden of 5 major types of gastrointestinal cancer. *Gastroenterology* 2020;**159**: 335-349.
2. Li C, Li L, Shi J. Gastrointestinal endoscopy in early diagnosis and treatment of gastrointestinal tumors. *Pak J Med Sci* 2020;**36**:203-207.
3. de Lange T, Halvorsen P, Riegler M. Methodology to develop machine learning algorithms to improve performance in gastrointestinal endoscopy. *World J Gastroenterol* 2018;**24**:5057-5062.
4. Schork NJ. Artificial intelligence and personalized medicine. *Cancer Treat Res* 2019;**178**:265-283.
5. Calderaro J, Kather JN. Artificial intelligence-based pathology for gastrointestinal and hepatobiliary cancers. *Gut* 2021;**70**:1183-1193.
6. Okagawa Y, Abe S, Yamada M, Oda I, Saito Y. Artificial intelligence in endoscopy. *Dig Dis Sci* 2022;**67**:1553-1572.
7. Martin L, Peine A, Gronholz M, Marx G, Bickenbach J. [Artificial intelligence: challenges and applications in intensive care medicine]. *Anasthesiol Intensivmed Notfallmed Schmerzther* 2022;**57**:199-209.
8. Sundaram S, Choden T, Mattar MC, Desai S, Desai M. Artificial intelligence in inflammatory bowel disease endoscopy: current landscape and the road ahead. *Ther Adv Gastrointest Endosc* 2021;**14**:26317745211017809.
9. Leggett CL, Wang KK. Computer-aided diagnosis in GI endoscopy: looking into the future. *Gastrointest Endosc* 2016;**84**:842-844.
10. El Hajjar A, Rey JF. Artificial intelligence in gastrointestinal endoscopy: general overview. *Chin Med J (Engl)* 2020;**133**:326-334.
11. Dey D, Slomka PJ, Leeson P, et al. Artificial intelligence in cardiovascular imaging: JACC state-of-the-art review. *J Am Coll*

Cardiol 2019;**73**:1317-1335.

12. Berbís MA, Aneiros-Fernández J, Mendoza Olivares FJ, Nava E, Luna A. Role of artificial intelligence in multidisciplinary imaging diagnosis of gastrointestinal diseases. *World J Gastroenterol* 2021;**27**:4395-4412.

13. Ali H. The potential of GPT-4 as a personalized virtual assistant for bariatric surgery patients. *Obes Surg* 2023;**33**:1605.

14. Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer* 2012;**48**:441-446.

15. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology* 2016;**278**:563-577.

16. Dumoulin FL, Rodriguez-Monaco FD, Ebigbo A, Steinbrück I. Artificial intelligence in the management of barrett's esophagus and early esophageal adenocarcinoma. *Cancers (Basel)* 2022;**14**:1918.

17. Wang SX, Ke Y, Liu YM, et al. [Establishment and clinical validation of an artificial intelligence YOLOv51 model for the detection of precancerous lesions and superficial esophageal cancer in endoscopic procedure]. *Zhonghua Zhong Liu Za Zhi* 2022;**44**:395-401.

18. Iyer PG, Chak A. Surveillance in Barrett's esophagus: challenges, progress, and possibilities. *Gastroenterology* 2023;**164**:707-718.

19. Vulpoi RA, Luca M, Ciobanu A, Olteanu A, Barboi OB, Drug VL. Artificial intelligence in digestive endoscopy-where are we and where are we going? *Diagnostics (Basel)* 2022;**12**:927.

20. Tang D, Wang L, Ling T, et al. Development and validation of a real-time artificial intelligence-assisted system for detecting early gastric cancer: a multicentre retrospective diagnostic study. *EBioMedicine* 2020;**62**:103146.

21. Huang CR, Sheu BS, Chung PC, Yang HB. Computerized diagnosis of *Helicobacter pylori* infection and associated gastric inflammation from endoscopic images by refined feature selection using a neural network. *Endoscopy* 2004;**36**:601-608.

22. 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.

23. Tran V, Saad T, Tesfaye M, et al. *Helicobacter pylori* (*H. pylori*) risk factor analysis and prevalence prediction: a machine learning-based approach. *BMC Infect Dis* 2022;**22**:655.

24. Cao R, Tang L, Fang M, et al. Artificial intelligence in gastric cancer: applications and challenges. *Gastroenterol Rep (Oxf)* 2022;**10**:goac064.

25. Diao W, Huang X, Shen L, Zeng Z. Diagnostic ability of blue laser imaging combined with magnifying endoscopy for early esophageal cancer. *Dig Liver Dis* 2018;**50**:1035-1040.

26. Hirasawa T, Aoyama K, Tanimoto T, et al. Application of artificial intelligence using a convolutional neural network for detecting gastric cancer in endoscopic images. *Gastric Cancer* 2018;**21**:653-660.

27. Miyaki R, Yoshida S, Tanaka S, et al. Quantitative identification of mucosal gastric cancer under magnifying endoscopy with flexible spectral imaging color enhancement. *J Gastroenterol Hepatol* 2013;**28**:841-847.

28. Nagao S, Tsuji Y, Sakaguchi Y, et al. Highly accurate artificial intelligence systems to predict the invasion depth of gastric cancer: efficacy of conventional white-light imaging, nonmagnifying narrow-band imaging, and indigo-carmine dye contrast imaging. *Gastrointest Endosc* 2020;**92**:866-873.

29. Luo H, Xu G, Li C, et al. Real-time artificial intelligence for detection of upper gastrointestinal cancer by endoscopy: a multicentre, case-control, diagnostic study. *Lancet Oncol* 2019;**20**:1645-1654.

30. Byrne MF, Chapados N, Soudan F, et al. 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.

31. Wang S, Zhu Y, Yu L, et al. RMDL: Recalibrated multi-instance deep learning for whole slide gastric image classification. *Med Image Anal* 2019;**58**:101549.

32. Wang J, Xie Z, Zhu X, et al. Differentiation of gastric schwannomas from gastrointestinal stromal tumors by CT using machine learning. *Abdom Radiol (NY)* 2021;**46**:1773-1782.

33. Chen T, Liu S, Li Y, et al. Developed and validated a prognostic nomogram for recurrence-free survival after complete surgical resection of local primary gastrointestinal stromal tumors based on deep learning. *EBioMedicine* 2019;**39**:272-279.

34. Wen Q, Yang Z, Zhu J, et al. Pretreatment CT-based radiomics signature as a potential imaging biomarker for predicting the expression of PD-L1 and CD8+TILs in ESCC. *Onco Targets Ther* 2020;**13**:12003-12013.

35. Wu L, Wang C, Tan X, et al. Radiomics approach for preoperative identification of stages I-II and III-IV of esophageal cancer. *Chin J Cancer Res* 2018;**30**:396-405.

36. Wu L, Yang X, Cao W, et al. Multiple level CT radiomics features preoperatively predict lymph node metastasis in esophageal cancer: a multicentre retrospective study. *Front Oncol* 2019;**9**:1548.

37. Cui Y, Yang X, Shi Z, et al. Radiomics analysis of multiparametric MRI for prediction of pathological complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer. *Eur Radiol* 2019;**29**:1211-1220.

38. Renna F, Martins M, Neto A, et al. Artificial intelligence for upper gastrointestinal endoscopy: a roadmap from technology development to clinical practice. *Diagnostics (Basel)* 2022;**12**:1278.

39. Luo Y, Zhang Y, Liu M, et al. Artificial intelligence-assisted colonoscopy for detection of colon polyps: a prospective, randomized cohort study. *J Gastrointest Surg* 2021;**25**:2011-2018.

40. Hassan C, Spadaccini M, Iannone A, et al. Performance of artificial intelligence in colonoscopy for adenoma and polyp detection: a systematic review and meta-analysis. *Gastrointest Endosc* 2021;**93**:77-85.

41. Glissen Brown JR, Mansour NM, Wang P, et al. Deep learning computer-aided polyp detection reduces adenoma miss rate: a United States multi-center randomized tandem colonoscopy study (CADeT-CS Trial). *Clin Gastroenterol Hepatol* 2022;**20**:1499-1507.

42. Hassan C, Spadaccini M, Mori Y, et al. Real-time computer-aided detection of colorectal neoplasia during colonoscopy: a systematic review and meta-analysis. *Ann Intern Med* 2023;**176**:1209-1220.

43. Mohan BP, Facciorusso A, Khan SR, et al. Real-time computer aided colonoscopy versus standard colonoscopy for improving adenoma detection rate: A meta-analysis of randomized-controlled trials. *EClinicalMedicine* 2020;**29-30**:100622.

44. Kamba S, Tamai N, Saitoh I, et al. Reducing adenoma miss rate of colonoscopy assisted by artificial intelligence: a multicenter randomized controlled trial. *J Gastroenterol* 2021;**56**:746-757.

45. Lu YB, Lu SC, Huang YN, et al. A novel convolutional neural network model as an alternative approach to bowel preparation evaluation before colonoscopy in the COVID-19 era: a multicenter, single-blinded, randomized study. *Am J Gastroenterol* 2022;**117**:1437-1443.

46. Quan SY, Wei MT, Lee J, et al. Clinical evaluation of a real-time artificial intelligence-based polyp detection system: a US multi-center pilot study. *Sci Rep* 2022;**12**:6598.

47. Lux TJ, Banck M, Saßmannshausen Z, et al. Pilot study of a new freely available computer-aided polyp detection system in clinical practice. *Int J Colorectal Dis* 2022;**37**:1349-1354.

48. Wallace MB, Sharma P, Bhandari P, et al. Impact of artificial intelligence on miss rate of colorectal neoplasia. *Gastroenterology* 2022;**163**:295-304.

49. Trebeschi S, van Griethuysen JJM, Lambregts DMJ, et al. Deep learning for fully-automated localization and segmentation of rectal cancer on multiparametric MR. *Sci Rep* 2017;**7**:5301.

50. Irving B, Franklin JM, Papież BW, et al. Pieces-of-parts for

supervoxel segmentation with global context: Application to DCE-MRI tumour delineation. *Med Image Anal* 2016;**32**:69-83.

51. Gibson E, Giganti F, Hu Y, et al. Automatic multi-organ segmentation on abdominal CT with dense V-networks. *IEEE Trans Med Imaging* 2018;**37**:1822-1834.

52. Jin D, Guo D, Ho TY, et al. DeepTarget: Gross tumor and clinical target volume segmentation in esophageal cancer radiotherapy. *Med Image Anal* 2021;**68**:101909.

53. Zhang S, Wu J, Shi E, et al. MM-GLCM-CNN: A multi-scale and multi-level based GLCM-CNN for polyp classification. *Comput Med Imaging Graph* 2023;**108**:102257.

54. Gong R, He S, Tian T, Chen J, Hao Y, Qiao C. FRCNN-AA-CIF: An automatic detection model of colon polyps based on attention awareness and context information fusion. *Comput Biol Med* 2023;**158**:106787.

55. Nogueira-Rodríguez A, Glez-Peña D, Reboiro-Jato M, López-Fernández H. Negative samples for improving object detection-a case study in ai-assisted colonoscopy for polyp detection. *Diagnostics (Basel)* 2023;**13**:966.

56. Sadagopan R, Ravi S, Adithya SV, Vivekanandhan S. PolyEffNetV1: A CNN based colorectal polyp detection in colonoscopy images. *Proc Inst Mech Eng H* 2023;**237**:406-418.

57. Gilabert P, Vitrià J, Laiz P, et al. Artificial intelligence to improve polyp detection and screening time in colon capsule endoscopy. *Front Med (Lausanne)* 2022;**9**:1000726.

58. González-Bueno Puyal J, Brandao P, Ahmad OF, et al. Polyp detection on video colonoscopy using a hybrid 2D/3D CNN. *Med Image Anal* 2022;**82**:102625.

59. Adjei PE, Lonseko ZM, Du W, Zhang H, Rao N. Examining the effect of synthetic data augmentation in polyp detection and segmentation. *Int J Comput Assist Radiol Surg* 2022;**17**:1289-1302.

60. Tanwar S, Vijayalakshmi S, Sabharwal M, Kaur M, AlZubi AA, Lee HN. Detection and classification of colorectal polyp using deep learning. *Biomed Res Int* 2022;**2022**:2805607.

61. Nogueira-Rodríguez A, Reboiro-Jato M, Glez-Peña D, López-Fernández H. Performance of convolutional neural networks for polyp localization on public colonoscopy image datasets. *Diagnostics (Basel)* 2022;**12**:898.

62. Jheng YC, Wang YP, Lin HE, et al. A novel machine learning-based algorithm to identify and classify lesions and anatomical landmarks in colonoscopy images. *Surg Endosc* 2022;**36**:640-650.

63. Taghiakbari M, Mori Y, von Renteln D. Artificial intelligence-assisted colonoscopy: a review of current state of practice and research. *World J Gastroenterol* 2021;**27**:8103-8122.

64. Alagappan M, Brown JRG, Mori Y, Berzin TM. Artificial intelligence in gastrointestinal endoscopy: The future is almost here. *World J Gastrointest Endosc* 2018;**10**:239-249.

65. Khalaf K, Terrin M, Jovani M, et al. A comprehensive guide to artificial intelligence in endoscopic ultrasound. *J Clin Med* 2023;**12**:3757.

66. Tontini GE, Rimondi A, Vernero M, et al. Artificial intelligence in gastrointestinal endoscopy for inflammatory bowel disease: a systematic review and new horizons. *Therap Adv Gastroenterol* 2021;**14**:17562848211017730.

67. Mascarenhas M, Afonso J, Andrade P, Cardoso H, Macedo G. Artificial intelligence and capsule endoscopy: unravelling the future. *Ann Gastroenterol* 2021;**34**:300-309.

68. Kawamoto A, Takenaka K, Okamoto R, Watanabe M, Ohtsuka K. Systematic review of artificial intelligence-based image diagnosis for inflammatory bowel disease. *Dig Endosc* 2022;**34**:1311-1319.

69. Zand A, Stokes Z, Sharma A, van Deen WK, Hommes D. Artificial intelligence for inflammatory bowel diseases (IBD); accurately predicting adverse outcomes using machine learning. *Dig Dis Sci* 2022;**67**:4874-4885.

70. Takenaka K, Kawamoto A, Okamoto R, Watanabe M, Ohtsuka K. Artificial intelligence for endoscopy in inflammatory bowel disease. *Intest Res* 2022;**20**:165-170.

71. Ang TL, Wang LM. Artificial intelligence for the diagnosis of dysplasia in inflammatory bowel diseases. *J Gastroenterol Hepatol* 2022;**37**:1469-1470.

72. Guerrero Vinsard D, Fetzer JR, Agrawal U, et al. Development of an artificial intelligence tool for detecting colorectal lesions in inflammatory bowel disease. *iGIE* 2023;**2**:91-101.

73. Njei B, McCarty TR, Mohan BP, Fozo L, Navaneethan U. Artificial intelligence in endoscopic imaging for detection of malignant biliary strictures and cholangiocarcinoma: a systematic review. *Ann Gastroenterol* 2023;**36**:223-230.

74. Marya NB, Powers PD, Petersen BT, et al. Identification of patients with malignant biliary strictures using a cholangioscopy-based deep learning artificial intelligence (with video). *Gastrointest Endosc* 2023;**97**:268-278.

75. Saraiva MM, Ribeiro T, Ferreira JPS, et al. Artificial intelligence for automatic diagnosis of biliary stricture malignancy status in single-operator cholangioscopy: a pilot study. *Gastrointest Endosc* 2022;**95**:339-348.

76. Ribeiro T, Saraiva MM, Afonso J, et al. Automatic identification of papillary projections in indeterminate biliary strictures using digital single-operator cholangioscopy. *Clin Transl Gastroenterol* 2021;**12**:e00418.

77. Yao L, Zhang J, Liu J, et al. A deep learning-based system for bile duct annotation and station recognition in linear endoscopic ultrasound. *EBioMedicine* 2021;**65**:103238.