Role of artificial intelligence in multidisciplinary imaging diagnosis of gastrointestinal diseases

M Alvaro Berbís, José Aneiros-Fernández, F Javier Mendoza Olivares, Enrique Nava, Antonio Luna

ORCID number: M Alvaro Berbís 0000-0002-0331-7762; José Aneiros-Fernández 0000-0001-7281-8045; F. Javier Mendoza Olivares 0000-0001-9450-4127; Enrique Nava 0000-0001-7817-6442; Antonio Luna 0000-0001-9358-3396.

Author contributions: All authors contributed to this paper with literature review and analysis and approval of the final version.

Conflict-of-interest statement: Authors declare no conflict of interest for this article.

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

Manuscript source: Invited manuscript

Specialty type: Gastroenterology and hepatology

Abstract

The use of artificial intelligence-based tools is regarded as a promising approach to increase clinical efficiency in diagnostic imaging, improve the interpretability of results, and support decision-making for the detection and prevention of diseases. Radiology, endoscopy and pathology images are suitable for deep-learning analysis, potentially changing the way care is delivered in gastroenterology. The aim of this review is to examine the key aspects of different neural network architectures used for the evaluation of gastrointestinal conditions, by discussing how different models behave in critical tasks, such as lesion detection or characterization (i.e. the distinction between benign and malignant lesions of the esophagus, the stomach and the colon). To this end, we provide an overview on recent achievements and future prospects in deep learning methods applied to the analysis of radiology, endoscopy and histologic whole-slide images of the gastrointestinal tract.

Key Words: Artificial intelligence; Machine learning; Radiology; Endoscopy; Pathology; Gastroenterology

©The Author(s) 2021. Published by Baishideng Publishing Group Inc. All rights reserved.
Core Tip: Artificial intelligence in general, and machine learning (ML) in particular, have great potential as supporting tools for physicians in the evaluation of neoplastic diseases and other conditions of the gastrointestinal tract. Radiology, endoscopy and pathology images can be read and interpreted using ML approaches in a wide variety of clinical scenarios. These include detection, classification and automatic segmentation of tumor lesions, tumor grading, patient stratification and prediction of treatment response.

Citation: 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(27): 4395-4412
URL: https://www.wjgnet.com/1007-9327/full/v27/i27/4395.htm
DOI: https://dx.doi.org/10.3748/wjg.v27.i27.4395

INTRODUCTION

Among all the emerging technologies that are shaping the future of medicine, artificial intelligence (AI) is arguably the one which will most alter the way care is delivered in the short and medium term. There is agreement in the field that AI will deeply impact healthcare, allowing for better diagnostics, better treatments, a more efficient use of the medical resources, and a more personalized management of patients[1,2]. In particular, diagnostic techniques based on medical images are spearheading this revolution[3], because their scope (the analysis of images, which, in general, are already digital), is highly accessible for computing systems. In comparison, other specialties such as emergency medicine or cardiac surgery, are less approachable by computers, and penetration of IA in these medical disciplines will be slower and later.

In the field of gastroenterology, several imaging modalities are used for evaluation of the digestive tract and the diagnosis of gastrointestinal (GI) tumors. These include radiology, endoscopy and histologic sections of GI specimens. In automatic analyses of GI images, a subtype of AI called machine learning (ML) is mainly involved. However, other branches of AI, such as natural language processing (NLP), have also been applied in the field of gastroenterology.

The aim of this review is to synthesize the available evidence on the application of AI to the analysis of radiology, histology and endoscopic images of the GI tract. Other applications of AI in the field of digestive disease, such as NLP, have been reviewed elsewhere[4,5] and are out of the scope of this article.

AI IN A NUTSHELL

AI is an umbrella term referring to different techniques used to solve a given problem by a computing method that mimics human intelligence. This includes automatic image identification, classification and interpretation, as well as recognition and processing of natural language. As such, AI encompasses a wide range of technologies, which vary in their complexity, versatility, and applicability to different types of problems.

Of particular interest is ML, which refers to a specific form of AI that endows computers with the ability to learn from and improve with experience, by-passing the need of being explicitly programmed to perform a specific task[6]. This learning process can be supervised or unsupervised, depending on the existence of labels in the sample dataset used to train the ML algorithms. Supervised learning consists of training an algorithm with classified or labeled examples, while unsupervised learning involves training with unstructured, unlabeled bulk data, and requires that the algorithm extracts the inherent structure thereof[7]. Typical techniques used in supervised learning approaches include support vector machine (SVM), naïve Bayes and random forest (RF). For its part, popular clustering algorithms used for unsupervised learning include k-means, k-medoids and Gaussian mixtures.

Both learning schemes can be used to train artificial neural networks (ANN), a subtype of ML algorithms which have a structure loosely inspired by the human brain.
They are made up of several node units, known as neurons, which are mostly structured in successive layers. The products of one layer go through weighted connections to the entries of the next layer, until a final output is produced. As an example, Figure 1 shows a three-layered neural network: When ANN are composed of many layers of nodes between the input and output, they are called deep neural networks (DNN), and using these models is called deep learning (DL). Thanks to their complexity, DNNs are able to model more complex relationships and execute more difficult tasks.

Convolutional neural networks (CNN) are DNN specialized in image recognition and classification. CNNs automatically extract descriptors by using a special neural network whose weights are determined by a training set. Figure 2 shows the typical structure of a CNN. The scheme shows how the image is consumed by the feature extraction network, which is a neural network composed of subsequent pairs of convolutional and pooling layers. Convolutional layers filter and analyze different image features, which are convoluted with the input data. Then, a pooling layer is used to reduce the size of the image produced by the convolutional layer. This process is repeated as per the number of layer pairs in the extraction network. Finally, the classification is done by an ANN acting as a classifier network.

The performance of CNN in the analysis of natural images has already attained similar, and even superior levels compared to that of the human eye[8]. Indeed, thanks to the rise in computational power, the reduction of hardware costs, the substantial development of efficient network architectures and the increasing wealth of data, the last years have witnessed a spectacular surge in AI-driven applications related to medicine.

Today, AI is a hot research topic in many medical fields. As a token, the number of publications found in PubMed using the search term “AI” more than doubled in only 3 years, rising from 6761 in 2016 to 15435 in 2019. The number of results per year after adding the term “gastroenterology” to that search also speaks for itself, increasing from 30 to 169 over the same time span (Figure 3).

This illustrates a growing interest among both academic and industry groups in the design and clinical validation of novel applications of AI. This task is facilitated by technology companies, which offer cloud-based services to remotely train DL models. It is expected that ML solutions will become increasingly available through their implementation in computer-aided detection (or diagnosis) (CAD) systems[9], thereby facilitating the widespread usage of these tools in clinical practice. These solutions are becoming progressively offered by the main medical technology vendors, as well as independent clinical software firms.

APPLICATIONS OF AI IN GI RADIOLOGY

Reports of the use of AI in GI radiology are abundant in the literature. Advances in ML have been applied to a wide range of clinical problems related with interpretation of images for diagnosis and decision-making in every anatomic site of the GI tract, from the esophagus to the rectum.

ML models have been trained to process quantitatively mineable features from a wide variety of radiologic modalities, such as computed tomography (CT), magnetic resonance imaging (MRI) and ultrasound (US). Within MRI, virtually any type of sequences, including structural such as T1-weighted and T2-weighted, and functional such as diffusion-weighted imaging (DWI) and apparent diffusion coefficient (ADC) maps, have been incorporated into AI-driven models. In addition, some of the most complex networks have been designed to interpret contrast-enhanced multiphasic CT and MRI images (i.e., dynamic contrast-enhanced MRI series) and even analyze, in a conjoined fashion, multimodal (CT, MR) studies.

Many AI applications focus on, or incorporate steps to automatically segment radiologic images, a task which is very time-demanding when done manually. An array of network architectures is available to provide segmented images of the same resolution of the original ones, but with pixel or voxel-based boundary delimitation of adjacent tissue types[10].

The use of ML and DL algorithms is, in many instances, accompanied by radiomic approaches. Radiomics, a term which was coined as late as 2012[11], refers to the quantitative and objective analysis of features in radiologic images by computational means, in order to gain insight into otherwise invisible, or hardly quantifiable information of potential clinical relevance. A cornerstone of radiomic assessments involves the analysis of image texture features. Radiologic textures (Figure 4) refer to
Figure 1 A neural network made up of 3 layers: An input layer, a hidden layer and an output layer. The input data is taken by the neurons (shown as blue circles) in the input layer, and produce an output which is consumed by the next layer (hidden layer). They, in turn, perform similar computations and provide an output to the output layer, which yields a final output.

Figure 2 Conceptual example of a simple convolutional neural network used for classification of a stomach tumor.

differences in the grayscale intensities of adjacent pixels or voxels within a region of interest, and have been associated with intratumoral heterogeneity[12]. Radiomics is closely linked to IA, because it often uses ML methods to discover patterns in large datasets[13], although, alternatively, statistical approaches can be used instead.

In the following sections, we review the currently reported experience in the AI-based analysis of GI radiology studies in different clinical scenarios.

**Esophageal neoplasia**

Esophageal cancer (EC) is a very aggressive neoplasm, with two main histological subtypes: Squamous cell carcinoma (ESCC) and adenocarcinoma. Currently, EC is preferably treated with chemoradiation (CRT), with a positive correlation between the histopathological response and overall survival[14]. However, common derived parameters from CT and positron emission tomography (PET) have shown a limited accuracy in treatment prediction and response assessment. In this scenario, radiomics and DL methods have shown initial promise in prediction of response to treatment in patients with EC. Radiomic signatures in PET images have been used to predict prognosis in patients with esophageal adenocarcinoma[15] and in patients with ESCC[16]. Also, 18F-fluorodeoxyglucose (FDG) PET/CT metrics and textural features showed utility in predicting response to induction chemotherapy followed by neoadjuvant CRT in a mixed cohort including patients with either type of EC[17].

Prognosis prediction of patients with EC has also been addressed from the viewpoint of CT. For instance, Ou et al[18] used multivariable logistic regression, RF and SVM classifiers on radiomic biomarkers extracted from CT data to predict resectability of ESCC[18]. Also, Jin et al[19] evaluated the potential of an integrated model combining radiomic analyses on CT images and dosimetric parameters in predicting response to CRT in patients with esophageal adenocarcinoma or ESCC. The combined model achieved an accuracy of over 0.7, and displayed a better prediction performance than the model using radiomic features alone[19]. Later, Hu et al[20] investigated the...
role of radiomic features and DL models in predicting response to neoadjuvant CRT in patients with ESCC eligible for surgery. Manual radiomic feature extraction and feature mapping by CNN were done on pre-treatment CT images. DL modeling based on SVM classifiers showed a greater capacity in assessing tumor heterogeneity and outperformed handcrafted radiomic markers in predicting pathologic complete responses of the tested cohort. Of note, the combination of radiomics and DL did not result in a better predictive performance than DL alone[20]. CT radiomics features have also been successfully used to identify programmed death-ligand 1 and CD8+ tumor infiltrating lymphocyte expression levels in patients with ESCC. Such analyses, performed on pretreatment CT images, allowed for better patient stratification and selection of candidates for immune checkpoint inhibitor therapy[21]. Finally, other encouraging applications of CT radiomics analysis in ESCC are pretreatment local staging[22] and assessment of lymph node status[23,24] with advantage over morphological features and size criteria.

In addition, radiomics analysis has also been applied to MRI, although this technique is challenging to perform in the esophagus due to motion artifacts. Preliminary results of radiomics analyses using pretreatment T2-weighted sequences show promise in the prediction of treatment response to CRT in patients with ESCC [25].

**Gastric adenocarcinomas and GI stromal tumors**

Radiomic analyses with different techniques have been applied to improve malignant tumor detection and characterization in the stomach. In this regard, a nomogram including multiphasic CT radiomics features was able to predict the presence of gastric adenocarcinomas[26].

---

**Figure 3 PubMed results by year using the search terms.** A: Artificial intelligence; B: Artificial intelligence gastroenterology.
The aggressiveness potential of gastric adenocarcinomas can be predicted by ML and radiomic analyses on CT and MRI of the stomach. Texture analysis performed in portal venous phase CT scans discovered feature correlates of lymphovascular and perineural invasion potential in patients with tubular gastric adenocarcinomas treated with total gastrectomy. Classifications done with eight different ML models were evaluated, among which naïve Bayes and RF exhibited the best performance in predicting the existence of lymph nodes metastasis, vascular and perineural invasion [27]. Similarly, quantitative assessment of intratumoral heterogeneity of gastric adenocarcinomas by entropy parameters extracted from ADC maps could be correlated with overall stage and prognostic factors of malignant behavior, including vascular and perineural invasion[28]. Another very recent study has shown the
potential role of a double-energy CT radiomics nomogram for prediction of lymph node metastases in gastric adenocarcinomas, with advantages over the current clinical model. In addition, the nomogram demonstrated a significant correlation with patient survival[29]. Also, AI and radiomics analyses have been applied to the prediction of treatment response of metastatic gastric adenocarcinomas. An ANN predictive model using radiomics analyses of pretreatment contrast-enhanced computed tomography (CECT) demonstrated significant differences between responders and non-responders, in a cohort of patients treated with pulsed low dose radiotherapy[30].

In addition, ML has been successfully applied to diagnosis and clinical decision-making of therapeutic strategies in patients with GI stromal tumors (GIST). As reported by Wang et al[31], ML showed the ability to differentiate between gastric schwannomas and GISTs by assessing CT images[31]. In that study, radiologists performed worse than all of the five different ML models tested, which included RF, logistic regression and decision trees. Also, analysis of CECT scans by a residual neural network has proved useful to predict the risk of recurrence after curative resection of localized primary GIST[32]. This offers new avenues for non-invasively distinguishing between patients with high risk of recurrence, for which adjuvant treatment with imatinib is recommended, and low-risk patients eligible for curative resection who are not likely to benefit from imatinib adjuvant therapy.

**Colorectal cancer**

AI has been extensively researched in the detection, characterization and staging of colorectal cancer (CRC)[33]. Several ML techniques have shown utility in the automated detection of polyps by CT colonography[34,35]. Also, radiomic analyses of colorectal CT images were successful in classifying CRC lesions according to their KRAS gene mutation status[36].

Currently, MRI is considered the most accurate test for rectal cancer staging[37]. A faster region-based CNN was trained to detect metastatic lymph nodes in T2-weighted and DWI images of the pelvis. The N staging provided by this network was very consistent with that done by radiologists, while average diagnostic time was markedly shorter (20 ms vs 600 ms, respectively)[38,39].

In addition, radiomic models aided by SVM were useful in predicting liver metastases of colon cancer. A combined model which included radiomic features of preoperative CT scans with two clinical variables (tumor site and diameter of tumor tissue) showed a higher prediction performance than either the clinical features or the radiomic signatures taken separately[40].

Radiomics analyses of CT, MRI and PET/CT images have been extensively used to predict treatment outcome and survival in patients with CRC. A recent systematic review analyzed 81 studies focused on this task, finding only 13 high-quality reports demonstrating a good performance in predicting treatment response, which mainly involved MRI studies of rectal cancer. Of note, the authors concluded that radiomics research in this field needs more clinical validation, rather than new algorithms[41]. Van Helden et al[42] discovered radiomic predictors of response to palliative systemic therapy and survival of patients with metastatic CRC, using pre-treatment 18F-FDG PET/CT images processed with semiautomatic segmentation[42].

Also, in this line, a recent report investigated the prognostic value of a ML model based on liver CT radiomics analyses to predict survival of patients with metastatic CRC[43]. Interestingly, CT-based prediction models were superior to their clinical counterparts for 1-year survival prediction.

Radiomic analyses may be useful to guide treatment of locally advanced rectal cancer (LARC). A radiomic signature comprising 10 texture features from T2-weighted images were modeled by a RF algorithm and used to evaluate treatment responses of subjects with LARC after neoadjuvant chemotherapy. The signature showed good prediction of patient response to treatment, with a good discrimination power among complete responders, partial responders and non-responders, with areas under the curve (AUC) in the 0.83-0.86 range[44].

In another study, Cui et al[45] developed and validated a radiomic signature based on 12 features extracted from multiparametric MR images [T2-weighted, dynamic contrast enhanced (DCE) and ADC maps] to predict complete response in patients with LARC after neoadjuvant CRT. The combination of all three image types showed a better predictive value than either of them alone, with an AUC of over 0.94[45]. This was hypothesized to result from the different aspects of tumor behavior reflected by each modality, such as tumor intensity, vascularization and cellularity, respectively.
**Segmentation of radiologic images**

Segmentation is a key step in the diagnosis and evaluation of tumor diseases, as well as in the treatment volume calculation for radiotherapy planning. However, manually delineating the relation between tumors and adjacent structures is a very time-consuming task for radiologists, requires a considerable level of expertise, and often lacks adequate inter-observer reproducibility.

Several DL models have been developed for automatic delimitation of tissue boundaries, many of which have been specifically tested in GI sites. Of note, automatic segmentation can be particularly challenging in the abdomen, due to the lack of clear boundaries between some organs which are displayed with similar intensities in this anatomy, such as the liver, the stomach, the spleen and the kidneys. Also, peristaltic and breathing motions make the assessment of the GI tract with US and MRI more difficult.

The complexity of networks involved in image segmentation varies. Today, the more recent and sophisticated networks are based on fully convolutional networks (FCN), which are optimal frameworks for semantic image segmentation. Popular FCN-based architectures include SegNet [46] and U-net, a CNN specialized in fast and accurate segmentation of biomedical images [47]. U-net networks can be trained with a relatively limited set of data and are a usual choice for many automatic segmentation approaches.

Automatic segmentation of rectal tumors has been attained using T2-weighted [48, 49], DCE [50] and multiparametric (T2, DWI) MRI [51]. Multi-organ, and even full-abdomen segmentation are feasible, although they usually rely on multi-atlas label fusion [52]. This strategy requires registration and fusion of images acquired at different levels of the same patient and is prone to error due to inefficient image registration. Nonetheless, Gibson et al. [53] recently proposed a registration-free method for multi-organ segmentation of CT images, based on a FCN called DenseVNet [53].

A useful application of automatic segmentation is the calculation of the clinical target volume of tumors in planning CT images for radiotherapy purposes. This approach has the advantage of yielding more reproducible measurements of treatment areas, leading to more consistent radiation doses, and has been explored in the assessment of EC [54] as well as rectal tumor lesions [55, 56], among others.

### APPLICATIONS OF AI IN ENDOSCOPIC TECHNIQUES

AI in GI endoscopy holds tremendous promise to augment clinical performance, establish better treatment plans, and improve patient outcomes.

In September 2019, the first multidisciplinary global gastroenterology AI meeting (First Global AI in Gastroenterology and Endoscopy Summit) was held in Washington, DC (United States), with a mandate to discuss and deliberate on practice, policy, ethics, data security, and patient care issues related to identification and implementation of appropriate use cases in gastroenterology [57]. Among the many challenges related to the use of AI for gastroenterology applications, Summit attendees highlighted 7 future needs to advance the field of AI in this discipline (Table 1).

Today, a wide variety of AI applications in GI endoscopy are being proposed, developed, and, in some cases, validated in clinical trials. The most current AI algorithms for gastroenterology are in the field of computer vision, which refers to technologies that can “see and interpret” visual data, such as a live video stream from the endoscope.

**Polyp detection**

Computer-aided polyp detection (CADe) and computer-aided polyp diagnosis (CADx) are two applications of AI-assisted computer vision for colonoscopy, which are already extensively studied [58, 59]. CADe and CADx have been the areas of most rapid progress so far in applying AI/computer vision to GI endoscopy. Given the central importance of screening and surveillance colonoscopy for CRC prevention, continued development and validation of CADe and CADx remains a top priority.

AI-assisted polyp detection studies must apply validated outcome parameters, such as adenoma detection rate, adenomas per colonoscopy, or adenoma miss rate, among others. Most of these studies are early-stage, but show promising clinical relevance. For instance, a recent study showed that an AI system based on DL and its real-time performance led to significant increases in polyp detection rate [60].
Table 1 Future needs to advance the field of artificial intelligence in gastroenterology

| No. | Future needs |
|-----|--------------|
| 1   | Identification of relevant and well-defined use cases |
| 2   | Development of high-quality metrics |
| 3   | Creation of large-scale imaging datasets |
| 4   | Clarity on the regulatory path to market |
| 5   | Use of appropriate AI methods |
| 6   | Clarification of patient privacy issues |
| 7   | Education of gastroenterologists on the risks and benefits of AI |

AI: Artificial intelligence.

Computer vision and image classification

Prioritization of additional use cases for computer vision in GI endoscopy must consider the prevalence of the targeted disease state, the possible clinical impact of the proposed algorithm, and the potential solvability of the clinical problem by AI.

The most common GI cancers share a common natural history. In this manner, CRC, gastric cancers, and ECs have precursor lesions that can be diagnosed by traditional endoscopic modalities. Additionally, inflammatory bowel disease (IBD) represents a relatively high prevalence group of conditions with an elevated risk of CRC but with dysplastic precursor lesions that are difficult to recognize endoscopically.

A major barrier to progress in ML in GI endoscopy is the relative absence of high-quality, labeled images for training AI algorithms. To overcome this, there is an urgent need to develop rules and recommendations regarding appropriate formats and quality standards for endoscopy image and video, along with accepted protocols for categorizing such images, storing metadata, and transferring and storing images while protecting private health information.

Image classification using AI has been developed to detect gastric cancer in endoscopic images with an overall sensitivity of 92.2%[61]. Another study applied a CNN to quantify the invasion depth of gastric cancer based on endoscopy data[62]. Similarly, a CNN was constructed to characterize EC, achieving a sensitivity as high as 98%[63].

GI bleeding

DL is especially relevant to video capsule endoscopy, given the large amount of data (around 8 h of video) and low efficiency of manual review by the physician. AI tools have also been shown to detect small-bowel bleeding and other relevant images most likely to be clinically relevant. The first studies involving CAD of bleeding from video capsule endoscopy images used color and texture feature extraction to help distinguish areas of bleeding from area of nonbleeding[64]. More recent studies, including those published by Jia and Meng[65], and Hassan et al[66], used DL-based features to achieve sensitivities and specificities as high as 99% for the detection of GI bleeding.

Shung et al[67] recently published a systematic review summarizing the use of ML techniques to predict outcomes in patients with acute GI bleeding. Fourteen studies with 30 assessments of ML models were included in their analysis. They predicted that ML performed better than clinical risk scores for mortality in upper GI bleeding (UGIB). Overall, the AUC for ANNs (median, 0.93; range, 0.78-0.98) was higher than other ML models (median, 0.81; range, 0.40-0.92)[67].

The emergence of automated lesion recognition endoscopy software combined with the boost in robotic techniques in surgery[68] has captured the attention of GI endoscopists and surgeons. For example, clinical risk-scoring systems offer invaluable, but not always practical, help for better stratification of patients at risk of UGIB and hemodynamic instability. Seo et al[69] developed an ML algorithm that predicts adverse events in patients with initially stable non-variceal UGIB[69]. Primary outcomes analyzed in this study included adverse events such as mortality, low blood pressure and rebleeding within 7 d. The authors compared four ML algorithms (logistic regression with regularization, RF classifier, gradient boosting classifier and voting classifier) with clinical Glasgow–Blatchford and Rockall scores. They found that the RF model achieved the highest accuracy and offered significant improvement over
conventional methods used for mortality prediction.

**IBD and AI**

Recently, AI solutions have been explored to better define mucosal healing in IBD\(^70\). With the incidence and global burden of IBD still on the rise, involving a large number of young patients with normal life expectancy, there is a constant need for more accurate ways to stratify risks and predict prognosis in these patients. Because the concept of mucosal healing is a rather new concept, the authors expect that AI tools for healing evaluation through endoscopic monitoring, which have recently been developed, could play a key role in better standardizing it\(^70\).

Waljee et al\(^71\) used RF methods to develop and validate prediction models of remission in patients with moderate to severe Crohn’s disease. In 401 participants, they showed an AUC of 0.78 at week 8 and an AUC of 0.76 at week 6. Also, Klang et al\(^72\) reported on the training of a CNN to detect Crohn’s disease ulcers, using 17,640 capsule endoscopy images from 49 patients\(^72\).

Assessment of ulcerative colitis (UC) can sometimes be very difficult. In early stages of the disease, erythema can be attributed to a number of other conditions, delaying UC diagnosis. On the other hand, foci of dysplasia in advanced disease are often missed due to the small amount of colonic mucosa sampled in surveillance colonoscopy. Thus, the development of AI solutions aimed at helping assess UC activity is a hot research topic.

For instance, Gutierrez et al\(^73\) proposed an automated end-to-end system using DL to binary predict the Mayo Clinic Endoscopic subscore, which showed a high degree of precision and robustness, with an AUC of 0.84\(^73\). In this connection, Kirchberger-Tolstik et al\(^74\) recently reported a non-destructive biospectroscopy technique assisted by neural networks to assess the severity of the disease according to the Endoscopic Mayo Score, with a mean sensitivity of 78% and a mean specificity of 93%\(^74\).

Maeda et al\(^75\) developed and evaluated a CAD system for predicting persistent histologic inflammation using endocytoscopy. To do so, they classified the endoscopic studies according to the histological findings of the corresponding biopsies, obtaining a predictive model with a high specificity and sensitivity\(^75\). In addition, Bossuyt et al\(^76\) constructed an algorithm tracking mucosal redness density in the red channel of endoscopic images along with vascular patterns. The results were accurately correlated with the activity of the disease at the endoscopic and histological level\(^76\).

**APPLICATIONS OF AI IN GI PATHOLOGY**

Today, most pathology labs still rely on fully analogic workflows, including the use of optical microscopes. However, with the availability of digital slide scanners clearly on the rise, there is a growing interest in the development of AI applications for histopathological studies.

Computational analyses of whole-slide images (WSI) of histology sections are regarded as a very promising form of improving diagnostic accuracy, reducing turnaround times and increasing interobserver agreement. Some studies regarding ML applications in pathology are discussed in this section.

**Esophageal neoplasia**

AI can be used for detection and grading of tumor lesions in the esophagus, and in particular of Barrett’s esophagus (Figure 5). Early detection of this condition is essential to improve patient prognosis. Histopathology is currently the gold standard for the diagnosis of this entity, but its study is characterized by a low degree of interobserver agreement in dysplasia grading.

To overcome this problem, Sabo et al\(^77\) developed two computerized morphometry models based on size, shape, texture, symmetry and distribution analyses of the epithelial nuclei of the esophagus. The first neural network model showed an accuracy of 89% in the differentiation of normal esophagus vs esophagus with low-grade dysplasia. The second model distinguished between low-grade and high-grade dysplasia with an accuracy of 86%, demonstrating potential for assisting pathologists in the differential diagnosis of indistinguishable lesions\(^77\). Van Sandick et al\(^78\) and Polkowski et al\(^79\) also addressed this issue, by combining morphometric analyses in hematoxylin and eosin (H&E) stains with immunohistochemistry data for p53 and Ki67. This hybrid approach improved the accuracy of the differentiation between low-grade and high-grade dysplasia up to 94%. Lastly, Salit et al\(^80\) compared the performance of three different models in classifying precursor lesions of Barrett’s
esophagus, which differed in their training approach (supervised, weakly supervised and unsupervised). The CNN trained with an unsupervised approach extracted the most relevant image features for identifying and classifying the cancer precursors[80].

**Inflammatory and infectious lesions of the stomach**

Chronic gastritis is a very prevalent condition. Its diagnosis is established by evaluating the degree of chronic and active inflammation, presence of *Helicobacter pylori* (*H. pylori*), atrophy of the mucosa and intestinal metaplasia. Steinbuss *et al*[81] used the Xception CNN architecture for automatic classification of three types of gastritis, namely type A (autoimmune), type B (bacterial), and type C (chemical) gastritis, in histological sections of antrum and corpus biopsies, with an overall accuracy of 84%[81].

*H. pylori* cells can be visualized in histology sections of gastric biopsy samples using different staining techniques, such as H&E, Giemsa or Warthin-Starry silver stains (Figure 6). Klein *et al*[82] published a DL algorithm for automatic screening in Giemsa stains, with a sensitivity of 100% and a specificity of 66%[82]. In parallel, Zhou *et al*[83] used a CNN to assist pathologists in the detection of *H. pylori* cells in H&E-stained WSI, but failed to demonstrate significant improvements in diagnostic accuracy and turnaround times in comparison with unassisted case studies[83].

**Gastric cancer**

Early detection and histopathologic characterization of gastric tumors are essential to improve treatment outcomes. A number of recent studies have paid attention to lesion detection, classification and characterization in this anatomic site. For instance, Song *et al*[84] developed and trained a deep CNN to differentiate between benign and malignant gastric tumors, with a sensitivity of 100% and a specificity of 80.6%[84].

Also, a network developed by Sharma *et al*[85] classified gastric cancer cases according to immunohistochemical response and presence of necrosis, with accuracy rates of 0.699 and 0.814, respectively[85].

Mori and Miwa[86] used a 6-layer CNN to study tumoral invasion depth in gastric adenocarcinoma images, with an accuracy of 85%, a sensitivity of 90% and a specificity of 81%[86]. Later, Iizuka *et al*[87] trained CNN and recurrent neural networks to classify stomach and colon WSI into adenocarcinoma, adenoma and non-neoplastic, with AUC values up to 0.97 and 0.99 for adenocarcinoma and gastric adenoma, respectively[87].

Lastly, Kather *et al*[88] trained a DL model to predict microsatellite instability from H&E-stained gastric cancer images, using a pathomics approach. This represents a very innovative approach, which extracts molecular and histochemical features solely based on H&E images and may circumvent the need to conduct genetic and/or immunohistochemical tests[88].

**Colonic inflammatory disease, polyps and CRC**

Klein *et al*[89] evaluated the histomorphometric features of colon biopsies from patients with Crohn’s disease. These analyses revealed that differences in the number
of inflammatory cells, lymphocytic aggregates and collagen density can be used as predictors of clinical phenotypes with an accuracy of 94%[89].

CRC is among the most common malignancies and a major cause of cancer-related death worldwide. Today, it is known that the vast majority of CRC cases arise from the adenoma-carcinoma sequence[90], and early detection of these lesions is considered of utmost importance to reduce CRC incidence rates. Despite this, evidence concerning the application of DL techniques to detect and characterize precancerous lesions in WSI of this anatomic site is still scarce.

Rodriguez-Diaz et al[91] developed a DL model to locate areas of malignant transformation inside polyps using semantic segmentation, distinguishing between neoplastic and non-neoplastic polyps, with a sensitivity of 0.96 and a specificity of 0.84[91]. Haj-Hassan et al[92] used a CNN to classify segmented regions of interest into three tissue types related to CRC progression (benign hyperplasia, intraepithelial neoplasia and carcinoma), with an accuracy of 99.17%[92]. Korbar et al[93] applied a residual network architecture to classify five polyp types (hyperplastic, sessile serrated, traditional serrated, tubular, and tubulovillous/villous) on WSI with an overall prediction accuracy of over 93%[93]. Also, several studies have proposed DL predictive models of survival for patients with CRC based on the extraction of prognostic markers in H&E images[94,95].

CONCLUSION

In this review, we have summarized the current evidence on the application of AI to the interpretation of radiology, endoscopy and histological images of the GI tract. At this moment, data on the use of AI in the assessment of radiology images of the GI are far more abundant than those from endoscopy and pathology studies. This difference is mainly a consequence of the asymmetrical availability of data and degree of digital transformation of each specialty. In any case, taken together, the available body of knowledge allows us to anticipate the central role that can be expected for AI in the personalized management of patients with high-risk GI tumors and other GI conditions.

On the downside, a large proportion of papers published to date are proof of concept approaches, based on retrospective analyses and single-center studies, usually involving limited data. As a result, and although there is growing evidence on its applicability in many clinical scenarios, the actual contribution of AI to clinical care in gastroenterology is still very limited.

The next several years will likely be a period of rapid development for AI tools in gastroenterology imaging. It is crucial that, as this field advances, there is a focus on technologies that provide real clinical benefits that have been validated in high-quality clinical trials. To this end, prospective, multicenter studies involving large sets of real-world data that reflect the high variability of image quality among institutions, will be key to assess the actual clinical applicability of AI solutions in this field.
As the development and uptake of AI in gastroenterology continues to grow, we will likely see a shift in the way diagnoses of GI conditions rely on AI. Evidence on the superior performance of computers over human experts in most diagnostic scenarios involving image interpretation will soon be notorious. This might ultimately lead to a deskilling of specialists, who would potentially suffer from too great a reliance on AI for their diagnoses, a situation which might be aggravated by the “black-box” nature of many AI applications. indeed, radiomics and pathomics approaches, which use imaging features that are otherwise invisible to the eye, or highly subjective, have shown a clear value in identifying patient profiles with higher risk, even before clinical symptoms appear. This can facilitate personalized treatment and improve the prognosis of patients with GI conditions, but efforts should be made to increase the “explainability” of AI-rendered results to both physicians and patients, in order to facilitate broad readership and understandability of diagnoses.

REFERENCES

1. Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, Wang Y, Dong Q, Shen H. Artificial intelligence in healthcare: past, present and future. Stroke Vasc Neurol 2017; 2: 230-243 [PMID: 29507784 DOI: 10.1136/svn-2017-000101]
2. Esteva A, Robicquet A, Ramsundar B, Kuleshov V, DePristo M, Chou K, Cui C, Corrado G, Thrun S, Dean J. A guide to deep learning in healthcare. Nat Med 2019; 25: 24-29 [PMID: 30617335 DOI: 10.1038/s41591-018-0316-z]
3. Martín Noguerol T. Paulano-Godino F, Martín-Valdivia MT, Menias CO, Luna A. Strengths, Weaknesses, Opportunities, and Threats Analysis of Artificial Intelligence and Machine Learning Applications in Radiology. J Am Coll Radiol 2019; 16: 1239-1247 [PMID: 31942401 DOI: 10.1016/j.jacr.2019.05.047]
4. Hou JK, Imler TD, Imperiale TF. Current and future applications of natural language processing in the field of digestive diseases. Clin Gastroenterol Hepatol 2014; 12: 1257-1261 [PMID: 24858706 DOI: 10.1016/j.cgh.2014.05.013]
5. Wang Y, He X, Nie H, Zhou J, Cao P, Ou C. Application of artificial intelligence to the diagnosis and therapy of colorectal cancer. Am J Cancer Res 2020; 10: 3575-3598 [PMID: 33294256]
6. House of Lords-AI in the UK: ready, willing and able?- Artificial Intelligence Committee [Internet]. [cited 2 January 2021]. Available from: https://publications.parliament.uk/pa/Ld201719/Ldselect/Ldai/100/10005.htm
7. Alloghani M, Al-Jumeily D, Mustafina J, Hussain A, Aljaaf AJ: A Systematic Review on Supervised and Unsupervised Machine Learning Algorithms for Data Science. 2020; 3:21
8. Schelp K, Kohl S, Radtke JP, Wiesenfarth M, Kickingereder P, Bickelhaupt S, Kuder TA, Stenzinger A, Hohenfellner M, Schlemmer HP, Bonekamp D. Classification of Cancer at Prostate MRI: Deep Learning versus Clinical PI-RADS Assessment. Radiology 2019; 293: 607-617 [PMID: 31592731 DOI: 10.1148/radiol.2019190938]
9. Fujita H. AI-based computer-aided diagnosis (AI-CAD): the latest review to read first. Radiol Phys Technol 2020; 13: 6-19 [PMID: 31898014 DOI: 10.1007/s12194-019-00524-4]
10. Cardenas CE, Yang J, Anderson BM, Court LE, Brock KB. Advances in Auto-Segmentation. Semin Radiat Oncol 2019; 29: 185-197 [PMID: 31027636 DOI: 10.1016/j.sradonc.2019.02.001]
11. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, Zegers CM, Gillies R, Boellard R, Dekker A, Aerts HJ. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer 2012; 48: 441-446 [PMID: 22257792 DOI: 10.1016/j.ejca.2011.11.036]
12. Varghese BA, Cen SY, Duddalwar VA. Texture Analysis of Imaging: What Radiologists Need to Know. AJR Am J Roentgenol 2019; 212: 520-528 [PMID: 30645163 DOI: 10.2214/AJR.18.20624]
13. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. Radiology 2016; 278: 563-577 [PMID: 26579733 DOI: 10.1148/放射学.2015151169]
14. Hammoud ZT, Kesler KA, Ferguson MK, Battafaranno RJ, Bhogaraju A, Hanna N, Govindan R, Mauer AA, Yu M, Einhorn LH. Survival outcomes of resected patients who demonstrate a pathologic complete response after neoadjuvant chemoradiation therapy for locally advanced esophageal cancer. Dis Esophagus 2006; 19: [PMID: 16643172 DOI: 10.1111/j.1442-2050.2006.00542.x]
15. Foley KG, Hills RK, Berthon B, Marshall C, Parkinson C, Lewis WG, Crosby TDL, Spezi E, Roberts SA. Development and validation of a prognostic model incorporating texture analysis derived from standardised segmentation of PET in patients with esophageal cancer. Eur Radiol 2018; 28: 428-436 [PMID: 28770406 DOI: 10.1007/s00330-017-4973-y]
16. Li Y, Beck M, Pääller T, Lili C, Hua W, Mai HD, Anthauer H, Biebl M, Thuss-Patience PC, Berger J, Stromberger C, Tinphofer J, Krupp J, Budach V, Hofheinz F, Lin Q, Zscheack S. A FDG-PET radiomics signature detects esophageal squamous cell carcinoma patients who do not benefit from chemoradiation. Sci Rep 2020; 10: 17671 [PMID: 33077841 DOI: 10.1038/s41598-020-74701-w]
17. Simoni N, Rossi G, Benetti G, Zuffante M, Micra R, Pavarana M, Guariglia S, Zivelonghi E, Mengardo V, Weindelmayr J, Giacone Luigi S, de Manzoni G, Cavedon C, Mazzarotto R. 18F-FDG
Berbis MA et al. AI in gastrointestinal imaging diagnosis

PET/CT Metrics Are Correlated to the Pathological Response in Esophageal Cancer Patients Treated With Induction Chemotherapy Followed by Neoadjuvant Chemo-Radiotherapy. Front Oncol 2020; 10: 599907 [PMID: 33330977 DOI: 10.3389/fonc.2020.599907]

18 Ou J, Li R, Zeng R, Wu CQ, Chen Y, Chen TW, Zhang XM, Wu L, Jiang Y, Yang JQ, Cao JM, Tang S, Tang MJ, Hu J. CT radiomic features for predicting resectability of oesophageal squamous cell carcinoma as given by feature analysis: a case control study. Cancer Imaging 2019; 19: 66 [PMID: 31619229 DOI: 10.1186/s40644-019-0254-0]

19 Jin X, Zheng X, Chen D, Jin J, Zhu G, Deng X, Han C, Gong C, Zhou Y, Liu C, Xie C. Prediction of response after chemoradiation for esophageal cancer using a combination of dosimetry and CT radiomics. Eur Radiol 2019; 29: 6080-6088 [PMID: 31028447 DOI: 10.1007/s00330-019-06193-w]

20 Hu Y, Xie C, Yang H, Ho JWK, Wen J, Han L, Lam KO, Wong IYH, Law SYK, Chiu KWH, Vardhanaabuthi V, Fu J. Computed tomography-based deep-learning prediction of neoadjuvant chemoradiotherapy treatment response in esophageal squamous cell carcinoma. Radiother Oncol 2021; 154: 6-13 [PMID: 32941954 DOI: 10.1016/j.radonc.2020.09.014]

21 Wen Q, Yang Z, Zhu J, Qiu Q, Dai H, Feng A, Xing L. 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 [PMID: 33242424 DOI: 10.21147/OTT.S261068]

22 Wu L, Wang C, Tan X, Cheng Z, Zhao K, Yan L, Liang Y, Liu Z, Liang C. Radiomics approach for preoperative identification of stages I–II and III–IV of esophageal cancer. Chinese J Cancer Res 2018; 30: 396-405 [DOI: 10.21147/j.issn.1000-9604.2018.04.02]

23 Wu L, Yang X, Cao W, Zhao K, Li W, Ye W, Chen X, Zhou Z, Liu Z, Liang C. Multiple Level CT Radiomics Features Preoperatively Predict Lymph Node Metastasis in Esophageal Cancer: A Multicentre Retrospective Study. Front Oncol 2019; 9: 1548 [PMID: 33039021 DOI: 10.3389/fonc.2019.01548]

24 Tan X, Ma Z, Yan L, Ye W, Liu Z, Liang C. Radiomics nomogram outperforms size criteria in discriminating lymph node metastasis in resectable esophageal squamous cell carcinoma. Eur Radiol 2019; 29: 392-400 [PMID: 29922924 DOI: 10.1007/s00330-018-5581-1]

25 Hou Z, Li S, Ren W, Liu J, Yan J, Wan S. Radiomic analysis in T2W and SPAIR T2W MRI: predict treatment response to chemoradiotherapy in esophageal squamous cell carcinoma. J Thorac Dis 2018; 10: 2256-2267 [PMID: 29581530 DOI: 10.21037/jtd.2018.03.123]

26 Wang Y, Liu W, Yu Y, Han W, Liu JJ, Xue HD, Lei J, Jin ZY, Yu JC. Potential value of CT radiomics in the distinction of intestinal-type gastric adenocarcinomas. Eur Radiol 2020; 30: 2934-2944 [PMID: 33204040 DOI: 10.1007/s00330-019-06629-3]

27 Yardımcı AH, Koçak B, Turan Bektas C, Sel İ, Yarıkaya E, Dursun N, Bektas H, Usul Aşfar Ç, Gürsu RU, Kılıçkesmez Ö. Tubular gastric adenocarcinoma: machine learning-based CT texture analysis for predicting lymphovascular and perineural invasion. Diagn Interv Radiol 2020; 26: 515-522 [PMID: 32990246 DOI: 10.15128/diir.2020.19507]

28 Liu S, Zheng H, Zhang Y, Chen L, Guan W, Guan Y, Ge Y, He J, Zhou Z. Whole-volume apparent diffusion coefficient-based entropy parameters for assessment of gastric cancer aggressiveness. J Magn Reson Imaging 2018; 47: 168-175 [PMID: 28471511 DOI: 10.1002/jmri.25752]

29 Li J, Dong D, Fang M, Wang R, Tian J, Li H, Gao J. Dual-energy CT-based deep learning radiomics can improve lymph node metastasis risk prediction for gastric cancer. Eur Radiol 2020; 30: 2324-2333 [PMID: 31955608 DOI: 10.1007/s00330-019-06621-1]

30 Hou Z, Yang Y, Li S, Yan J, Ren W, Liu J, Wang K, Liu B, Wan S. Radiomic analysis using contrast-enhanced CT: predict treatment response to pulsed low dose rate radiotherapy in gastric carcinoma with abdominal cavity metastasis. Quant Imaging Med Surg 2018; 8: 410-420 [PMID: 29928606 DOI: 10.21037/qims.2018.05.01]

31 Wang J, Xie Z, Zhu X, Niu Z, Ji H, He L, Hu Q, Zhang C. Differentiation of gastric schwannomas from gastrointestinal stromal tumors by CT using machine learning. Abdom Radiol (NY) 2021; 46: 1773-1782 [PMID: 33083871 DOI: 10.1007/s00261-020-02797-9]

32 Chen T, Liu S, Li Y, Feng X, Xiong W, Zhao X, Yang Y, Zhang C, Hu Y, Chen H, Lin T, Zhao M, Liu H, Yu J, Xu Y, Zhang Y, Li G. 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 [PMID: 30587460 DOI: 10.1016/j.ebiom.2018.12.028]

33 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]

34 Godkhindi AM, Gowda RM. Automated detection of polyps in CT colonography images using deep learning algorithms in colon cancer diagnosis. In: 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing, IECFDS 2017. Institute of Electrical and Electronics Engineers Inc.; 2018: 1722–1728

35 Xu JW, Suzuki K. Computer-aided detection of polyps in CT colonography with pixel-based machine learning techniques. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Springer, Berlin, Heidelberg; 2011. [cited 7 January 2021]. Available from: https://Link.springer.com/chapter/10.1007/978-3-642-24319-6_44

36 González-Castro V, Cernadas E, Huelga E, Fernández-Delgado M, Porto J, Antunez JR, Souto-Bayarri M. CT radiomics in colorectal cancer: Detection of KRAS mutation using texture analysis and
Beets-Tan RGH, Lambregts DMJ, Maas M, Bipaet S, Barbaro B, Curvo-Semedo L, Fenlon HM, Gollub MJ, Gourtsoyianni S, Halligan S, Hoefll C, Kim SH, Laghi A, Maier A, Rafaelsen SR, Stoker J, Taylor SA, Torkzad MR, Blomqvist L. Magnetic resonance imaging for clinical management of rectal cancer: Updated recommendations from the 2016 European Society of Gastrointestinal and Abdominal Radiology (ESGAR) consensus meeting. *Eur Radiol* 2018; 28: 1465-1475 [PMID: 29043428 DOI: 10.1007/s00330-017-5026-2]

Ding L, Liu GW, Zhao BC, Zhou YP, Li S, Zhang ZD, Guo YT, Li AQ, Lu Y, Yao HW, Yuan WT, Wang GY, Zhang DL, Wang L. Artificial intelligence system of faster region-based convolutional neural network surpassing senior radiologists in evaluation of metastatic lymph nodes of rectal cancer. *Chin Med J (Eng)* 2019; 132: 379-387 [PMID: 30701777 DOI: 10.1097/CMA.0000000000000957]

Lu Y, Yu Q, Gao Y, Zhou Y, Liu G, Dong Q, Ma J, Ding L, Yao H, Zhang Z, Xiao G, An Q, Wang G, Xi J, Yuan W, Lian Y, Zhang D, Zhao C, Yao Q, Liu W, Zhou X, Liu S, Wu Q, Xu W, Zhang J, Wang D, Sun Z, Zhang X, Hu J, Zhang M, Zheng X, Wang L, Zhao J, Yang S. Identification of Metastatic Lymph Nodes in MR Imaging with Faster Region-Based Convolutional Neural Networks. *Cancer Res* 2018; 78: 5135-5143 [PMID: 30026330 DOI: 10.1158/0008-5472.CAN-18-0494]

Li Y, Eresen A, Shangguan J, Yang J, Lu Y, Chen D, Wang J, Velichko Y, Yaghmai V, Zhang Z. Establishment of a new non-invasive imaging prediction model for liver metastasis in colon cancer. *Am J Cancer Res* 2019; 9: 2482-2492 [PMID: 31815048]

Staal FCR, van der Reijd DJ, Taghavi M, Lambregts DMJ, Beets-Tan RGH, Maas M. Radiomics for the Prediction of Treatment Outcome and Survival in Patients With Colorectal Cancer: A Systematic Review [Internet]. *Clin Colorectal Cancer* 2021; 20: 52-71 [PMID: 33349519 DOI: 10.1016/j.ccc.2020.11.001]

van Helden EJ, Vacher YJL, van Wieringen WN, van Velden FHP, Verheul HMW, Hoevstra OS, Boelhaar R, Menke-van der Hooven van Oordt CW. Radiomics analysis of pre-treatment [18F]FDG PET/CT for patients with metastatic colorectal cancer undergoing palliative systemic treatment. *Eur J Nucl Med Mol Imaging* 2018; 45: 2307-2317 [PMID: 30094460 DOI: 10.1007/s00259-018-4100-6]

Mühlberg A, Holch JW, Heinemann V, Huber T, Moltz J, Maurus S, Jäger N, Liu L, Froelich MF, Katzmann A, Gresser E, Taubmann O, Stühling M, Nörenberg D. The relevance of CT-based geometric and radiomics analysis of whole liver tumor burden to predict survival of patients with metastatic colorectal cancer. *Eur Radiol* 2021; 31: 834-846 [PMID: 32851450 DOI: 10.1007/s00330-020-01792-y]

Ferrari R, Mancini-Terracciano C, Voena C, Rengo M, Zerunian M, Ciardiello A, Grasso S, Mare’ V, Paramatti R, Russomando A, Santacesaria R, Satta A, Solfaroli Camillocci E, Facinelli R, Laghi A. MR-based artificial intelligence model to assess response to therapy in locally advanced rectal cancer. *Eur J Radiol* 2019; 118: 1-9 [PMID: 31439226 DOI: 10.1016/j.ejrad.2019.06.013]

Cui Y, Yang X, Shi Z, Yang Z, Du X, Zhao Z, Cheng X. 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 [PMID: 30128616 DOI: 10.1007/s00330-018-5883-9]

Badrinarayanan V, Kendall A, Cipolla R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Trans Pattern Anal Mach Intell* 2017; 39: 2481-2495 [PMID: 28060704 DOI: 10.1109/TPAMI.2016.2644615]

Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. [cited 2 January 2021]. Available from: http://lmb.informatik.uni-freiburg.de/people/ronneber/u-net

Wang J, Lu J, Qin G, Shen L, Sun Y, Ying H, Zhang Z, Hu W. Technical Note: A deep learning-based autosegmentation of rectal tumors in MR images. *Med Phys* 2018; 45: 2560-2564 [PMID: 29663417 DOI: 10.1002/mp.12918]

Wang M, Xie P, Ran Z, Jian J, Zhang R, Xia W, Yu T, Ni C, Gu J, Gao X, Meng X. Full convolutional network based multiple side-output fusion architecture for the segmentation of rectal tumors in magnetic resonance images: A multi-vendor study. *Med Phys* 2019; 46: 2659-2668 [PMID: 30972763 DOI: 10.1002/mp.13541]

Irving B, Franklin JM, Papiez BW, Anderson EM, Sharma RA, Gleeson VF, Brady SM, Schnabel MA. Pieces-of-parts for supervoxel segmentation with global context: Application to DCE-MRI tumor delineation. *Med Image Anal* 2016; 32: 69-83 [PMID: 27054278 DOI: 10.1016/j.media.2016.03.002]

Trebesch S, van Griethuysen JJM, Lambregts DMJ, Lahaye MJ, Parmar C, Bakers FCH, Peters NHGM, Beets-Tan RGH, Aerts HJWL. Deep Learning for Fully-Automated Localization and Segmentation of Rectal Cancer on Multiparametric MR. *Sci Rep* 2017; 7: 5301 [PMID: 28706185 DOI: 10.1038/s41598-017-05728-9]

Okada T, Linguraru MG, Horii M, Summers RM, Tomiyama N, Sato Y. Abdominal multi-organ segmentation from CT images using conditional shape-location and unsupervised intensity priors. *Med Image Anal* 2015; 26: 1-18 [PMID: 26277022 DOI: 10.1016/j.media.2015.06.009]

Gibson E, Giganti F, Hu Y, Bonnatni E, Bandula S, Gurusamy K, Davidson B, Pereira SP, Clarkson MJ, Barratt DC. Automatic Multi-Organ Segmentation on Abdominal CT With Dense V-Networks. *IEEE Trans Med Imaging* 2018; 37: 1822-1834 [PMID: 29994628 DOI: 10.1109/TMI.2018.2806309]

Jin D, Guo D, Ho TY, Harrison AP, Xiao J, Tseng CK, Lu L. DeepTarget: Gross tumor and clinical target volume segmentation in esophageal cancer radiotherapy. *Med Image Anal* 2021; 68: 101909
AI in gastrointestinal imaging diagnosis

55 Larsson R, Xiong JF, Song Y, Ling-Fu, Chen YZ, Xiaowei X, Zhang P, Zhao J. Automatic Delineation of the Clinical Target Volume in Rectal Cancer for Radiation Therapy using Three-dimensional Fully Convolutional Neural Networks. [cited 5 January 2021]. Available from: https://pubmed.ncbi.nlm.nih.gov/33341494/

56 Men K, Dai J, Li Y. Automatic segmentation of the clinical target volume and organs at risk in the planning CT for rectal cancer using deep dilated convolutional neural networks. Med Phys 2017; 44: 6377-6389 [PMID: 28963779 DOI: 10.1002/mp.12602]

57 Parsa S, Wallace M, Bagci U, Antonino M, Berzin T, Byrne M, Celik H, Farahani K, Golding M, Gross S, Jamali V, Mendonca P, Mori Y, Ninh A, Repici A, Rex D, Skrinak K, Thakkar SJ, van Hooft JE, Vargo J, Yu H, Xu Z, Sharma P. Proceedings from the First Global Artificial Intelligence in Gastroenterology and Endoscopy Summit. Gastrointest Endosc 2020; 92: 938-945. e1 [PMID: 32343975 DOI: 10.1016/j.gie.2020.04.044]

58 Byrne MF, Chapados N, Soudan F, Oertel C, Linares Pérez M, Kelly R, Ishqul N, Chandelier F, Rok D. 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: 30966576 DOI: 10.1136/gutjnl-2017-315457]

59 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]

60 Luo Y, Zhang Y, Liu M, Lai Y, Liu P, Wang Z, Xing T, Huang Y, Li Y, Li A, Wang Y, Luo X, Liu S, Han Z. Artificial Intelligence-Assisted Colonoscopy for Detection of Colon Polyps: a Prospective, Randomized Cohort Study. J Gastrointest Surg 2020 [PMID: 32988933 DOI: 10.1007/s11605-020-04802-4]

61 Hirasawa T, Aoyama K, Tanimoto T, Ishihara S, Schichio 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]

62 Zhu Y, Wang Q, Xu MD, Zhang Z, Cheng J, Zhong YS, Zhang YQ, Chen WF, Yao LQ, Zhou PH, Li QL. Application of convolutional neural network in the diagnosis of the invasion depth of gastric cancer based on conventional endoscopy. Gastrointest Endosc 2019; 89: 806-815. e1 [PMID: 30452916 DOI: 10.1016/j.gie.2018.11.011]

63 Horie Y, Yoshito T, Aoyama K, Yoshimizu S, Horiiuchi Y, Ishiyama A, Hirasawa T, Tsuchida T, Ozawa T, Ishihara S, Kumagai Y, Fujishiro M, Maetani I, Fujisaki J, Tada T. Diagnostic outcomes of esophageal cancer by artificial intelligence using convolutional neural networks. Gastrointest Endosc 2019; 89: 25-32 [PMID: 30120958 DOI: 10.1016/j.gie.2018.07.037]

64 Pan G, Yan G, Song X, Qiu X. BP neural network classification for bleeding detection in wireless capsule endoscopy. J Med Eng Technol 2009; 33: 575-581 [PMID: 19639509 DOI: 10.1080/0309190090311974]

65 Jia X, Meng MOH. A deep convolutional neural network for bleeding detection in Wireless Capsule Endoscopy images. [cited 21 January 2021]. Available from: https://pubmed.ncbi.nlm.nih.gov/28268409/

66 Hassan AR, Haque MA. Computer-aided gastrointestinal hemorrhage detection in wireless capsule endoscopy videos. Comput Methods Programs Biomed 2015; 122: 341-353 [PMID: 26390947 DOI: 10.1016/j.cmpb.2015.09.005]

67 Shung D, Simonov M, Gentry M, Au B, Laine L. Machine Learning to Predict Outcomes in Patients with Acute Gastrointestinal Bleeding: A Systematic Review. Dig Dis Sci 2019; 64: 2078-2087 [PMID: 31055722 DOI: 10.1007/s10620-019-05645-z]

68 Ciuti G, Skonieczna-Żydecka K, Marlicz W, Iacovacci V, Liu H, Stoyanov D, Arezzo A, Chiurazzi S, Marlicz W, Iacovacci V, Liu H, Stoyanov D, Arezzo A, Chiurazzi S. Convolutional Neural Network for Detecting Gastric Cancer in Endoscopic Images. J Med Devices 2019; 13(2): 021005 DOI: 10.1115/1.4042877

69 Seo DW, Yi H, Park B, Kim YJ, Jung DH, Woo I, Sohn CH, Ko BS, Kim N, Kim WY. Prediction of Adverse Events in Stable Non-Variceal Gastrointestinal Bleeding Using Machine Learning. J Clin Med 2020; 9 [PMID: 32301619 DOI: 10.3390/jcm9061632]

70 Nakase H, Hirano T, Wagatsuma K, Ichimiya T, Yamakawa T, Yokoyama Y, Hayashi Y, Hiyamary D, Kazama T, Yoshii S, Yamano HO. Artificial Intelligence-assisted endoscopy changes the definition of mucosal healing in ulcerative colitis. Gastro Endosc 2020 [PMID: 32990283 DOI: 10.1016/j.gie.2020.10.038]

71 Waljee AK, Wallace BI, Cohen-Mekelburg S, Liu Y, Liu B, Sauder K, Stilham 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]

72 Klange E, Barash Y, Margalit RY, Soffer S, Shimon O, Albshesh A, Ben-Horin S, Amitai MM, Eliakim R, Kopylov U. Deep learning algorithms for automated detection of Crohn's disease ulcers by video capsule endoscopy. Gastrointest Endosc 2020; 91: 606-613.e2 [PMID: 31743689 DOI: 10.1016/j.gie.2019.11.012]

73 Gutierrez Becker B, Arcadu F, Thalhammer A, Gamez Serna C, Feehan O, Drawnel F, Oh YS,
Prunotto M. Training and deploying a deep learning model for endoscopic severity grading in ulcerative colitis using multicenter clinical trial data. *Ther Adv Gastroenterol* 2021; 14: 2631774521990623 [PMID: 33718871] DOI: 10.1177/2631774521990623

Kirchberger-Tolstik T, Pradhan P, Vieth M, Grunert P, Popp J, Bocklitz TW, Stallmach A. Towards an Interpretable Classifier for Characterization of Endoscopic Mayo Scores in Ulcerative Colitis Using Raman Spectroscopy. *Anal Chem* 2020; 92: 13776-13784 [PMID: 32965101] DOI: 10.1021/acs.analchem.0b02163

Maeda Y, Kudo SE, Mori Y, Misawa M, Ogata N, Sasamura S, Wakamura K, Oda M, Mori K, Ohltsuka K. Fully automated diagnostic system with artificial intelligence using endoscopy 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]

Bossuyt P, Nakase H, Vermeire S, de Hertogh G, Ebbode T, Ferrante M, Hasegawa T, Willekens H, Ikemoto Y, Makino T, Bischops R. Automatic, computer-aided determination of endoscopic and histological inflammation in patients with mild to moderate ulcerative colitis based on red density. *Gut* 2020; 69: 1778-1786 [PMID: 31915237 DOI: 10.1136/gutjnl-2019-320056]

Sabo E, Beck AH, Montgomery EA, Bhattacharya B, Meitner P, Wang JY, Resnick MB. Computerized morphometry as an aid in determining the grade of dysplasia and progression to adenocarcinoma in Barrett's oesophagus. *Lab Invest* 2006; 86: 1261-1271 [PMID: 17075582 DOI: 10.1038/LabInvest.3700481]

van Sandick JW, Baak JP, van Lanschot JJ, Polkowski W, ten Kate FJ, Obertop H, Offerhaus GJ. Computerized quantitative pathology for the grading of dysplasia in surveillance biopsies of Barrett's oesophagus. *J Pathol* 2000; 190: 177-183 [PMID: 10657015 DOI: 10.1002/(SICI)1096-9896(200002)190:2<177::AID-PATH470-3.0.CO;2-X]

Polkowski W, Baak JP, van Lanschot JJ, Meijer GA, Schuurmans LT, Ten Kate FJ, Obertop H, Offerhaus GJ. Clinical decision making in Barrett's oesophagus can be supported by computerized immunouquantification and morphometry of features associated with proliferation and differentiation. *J Pathol* 1998; 184: 161-168 [PMID: 9602707 DOI: 10.1002/(SICI)1096-9896(199802)184:2<161::AID-PATH761-3.0.CO;2-X]

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 Oesophagus. *J Pers Med* 2020; 10 [PMID: 32977465 DOI: 10.3390/jpm10040141]

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]

Klein S, Gildenblat J, Ihle MA, Merkelbach-Bruse S, Noh KW, Peifer M, Quaas A, Büttner R. Deep learning for sensitive detection of Helicobacter Pylori in gastric biopsies. *BMH Gastroenterol* 2020; 61: 417 [PMID: 33308189 DOI: 10.1186/s12876-020-01494-7]

Zhou S, Marklund H, Blahu O, Desai M, Martin B, Bingham D, Berry GJ, Geula E, Ng AY, Shen J. Deep learning assistance for the histopathologic diagnosis of Helicobacter pylori. *Intell Med* 2020; 1:2 [DOI: 10.1106/ijbmed.2020.100004]

Song Z, Zou S, Zhou W, Huang Y, Shao L, Yuan J, Gou X, Jin W, Wang Z, Chen X, Ding X, Liu J, Yu C, Ku C, Liu C, Sun Z, Xu G, Wang Y, Zhang X, Wang D, Wang S, Xu W, Davis RC, Shi H. Clinically applicable histopathological diagnosis system for gastric cancer detection using deep learning. *Nat Commun* 2020; 11: 4294 [PMID: 32855423 DOI: 10.1038/s41467-020-18147-8]

Sharma H, Zerbe N, Klemper I, Hellwich O, Hufnagl P. Deep convolutional neural networks for automatic classification of gastric carcinoma using whole slide images in digital histopathology. *Comput Med Imaging Graph* 2017; 61: 2-13 [PMID: 28676295 DOI: 10.1016/j.compmedimag.2017.06.001]

Mori H, Miwa H. A histopathologic feature of the behavior of gastric signet-ring cell carcinoma; an image analysis study with deep learning. *Pathol Int* 2019; 69: 437-439 [PMID: 31276267 DOI: 10.1111/pin.12628]

Iizuka O, Kanavati F, Kato K, Rambeau M, Arike 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]

Kather JN, Pearson AT, Halama N, Jäger D, Krause J, Loosen SH, Marx A, Boor P, Tack F, Neumann UP, Grabisch H, Yoshikawa T, Brenner H, Chang-Claude J, Hoffmeister M, Trautwein C, Luedde T. Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. *Nat Med* 2019; 25: 1054-1056 [PMID: 31160815 DOI: 10.1038/s41591-019-0462-y]

Klein A, Mazor Y, Karban A, Ben-Itzhak O, Chowers Y, Sabo E. Early histological findings may predict the clinical phenotype in Crohn's colitis. *United European Gastroenterol J* 2017; 5: 694-701 [PMID: 28815033 DOI: 10.1177/2056646616676435]

Leslie A, Carey FA, Pratt NR, Steele RJ. The colorectal adenoma-carcinoma sequence. *Br J Surg* 2002; 89: 845-860 [PMID: 12081733 DOI: 10.1046/j.1365-2168.2002.02120.x]

Rodriguez-Diaz E, Baffy G, Lovk W, Mashimo H, Vidyarthi G, Mohapatra SS, Singh SK. Real-time artificial intelligence–based histologic classification of colorectal polyps with augmented visualization. [cited 23 January 2021]. Available from: https://pubmed.ncbi.nlm.nih.gov/32949567/

Haj-Hassan H, Chaddad A, Harkous Y, Derosiers C, Toews M, Tanougast C. Classifications of
Multispectral Colorectal Cancer Tissues Using Convolution Neural Network. *J Pathol Inform* 2017; 8: 1 [PMID: 28400990 DOI: 10.4103/jpi.jpi_47_16]

93 Korbar B, Olofson AM, Miraflor AP, Nicka CM, Suriawinata MA, Torresani L, Suriawinata AA, Hassanpour S. Deep Learning for Classification of Colorectal Polyps on Whole-slide Images. *J Pathol Inform* 2017; 8: 30 [PMID: 28828201 DOI: 10.4103/jpi.jpi_34_17]

94 Bychkov D, Linder N, Turkki R, Nording S, Kovanen PE, Verrill C, Wallander M, Lundin M, Haglund C, Lundin J. Deep learning based tissue analysis predicts outcome in colorectal cancer. *Sci Rep* 2018; 8: 3395 [PMID: 29467373 DOI: 10.1038/s41598-018-21758-3]

95 Kather JN, Krisam I, Chaoroentong P, Luedde T, Herpel E, Weis CA, Gaiser T, Marx A, Valous NA, Ferber D, Jansen L, Reyes-Aldasoro CC, Zörnig I, Jäger D, Brenner H, Chang-Claude J, Hoffmeister M, Halama N. Predicting survival from colorectal cancer histology slides using deep learning: A retrospective multicenter study. *PLoS Med* 2019; 16: e1002730 [PMID: 30677016 DOI: 10.1371/journal.pmed.1002730]
