Artificial intelligence in gastric cancer: a translational narrative review

Chaoran Yu[1,2], Ernest Johann Helwig[3]

[1]Department of Oncology, Shanghai Medical College, Fudan University, Shanghai, China; [2]Fudan University Shanghai Cancer Center, Shanghai, China; [3]Tongji Medical College of Huazhong University of Science and Technology, Wuhan, China

Contributions: (I) Conception and design: All authors; (II) Administrative support: All authors; (III) Provision of study materials or patients: All authors; (IV) Collection and assembly of data: All authors; (V) Data analysis and interpretation: All authors; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors

Correspondence to: Chaoran Yu, MD, PhD. Department of Oncology, Shanghai Medical College, Fudan University Shanghai Cancer Center, Fudan University, Shanghai 200025, China. Email: chaoran_yu@yeah.net; Ernest Johann Helwig, MD. Tongji Medical College of Huazhong University of Science and Technology, Wuhan, China. Email: ernjohelwig@gmail.com.

Abstract: Increasing clinical contributions and novel techniques have been made by artificial intelligence (AI) during the last decade. The role of AI is increasingly recognized in cancer research and clinical application. Cancers like gastric cancer, or stomach cancer, are ideal testing grounds to see if early undertakings of applying AI to medicine can yield valuable results. There are numerous concepts derived from AI, including machine learning (ML) and deep learning (DL). ML is defined as the ability to learn data features without being explicitly programmed. It arises at the intersection of data science and computer science and aims at the efficiency of computing algorithms. In cancer research, ML has been increasingly used in predictive prognostic models. DL is defined as a subset of ML targeting multilayer computation processes. DL is less dependent on the understanding of data features than ML. Therefore, the algorithms of DL are much more difficult to interpret than ML, even potentially impossible. This review discussed the role of AI in the diagnostic, therapeutic and prognostic advances of gastric cancer. Models like convolutional neural networks (CNNs) or artificial neural networks (ANNs) achieved significant praise in their application. There is much more to be fully covered across the clinical administration of gastric cancer. Despite growing efforts, adapting AI to improving diagnoses for gastric cancer is a worthwhile venture. The information yield can revolutionize how we approach gastric cancer problems. Though integration might be slow and labored, it can be given the ability to enhance diagnosing through visual modalities and augment treatment strategies. It can grow to become an invaluable tool for physicians. AI not only benefits diagnostic and therapeutic outcomes, but also reshapes perspectives over future medical trajectory.

Keywords: Artificial intelligence (AI); endoscope; convolutional neural networks (CNNs); gastric cancer; genomics

Introduction

While the complete incorporation of artificial intelligence (AI) into daily medical use is still decades away, its growth and potential applications for improving medicine cannot be denied (1-3). Cancers like gastric cancer, or stomach cancer, are ideal testing grounds to see if early undertakings of applying AI to medicine can yield valuable results.

There are numerous concepts derived from AI, including machine learning (ML) and deep learning (DL) (*Figure 1*) (1). In brief, ML is defined as the ability to learn data features without being explicitly programmed. It arises at the
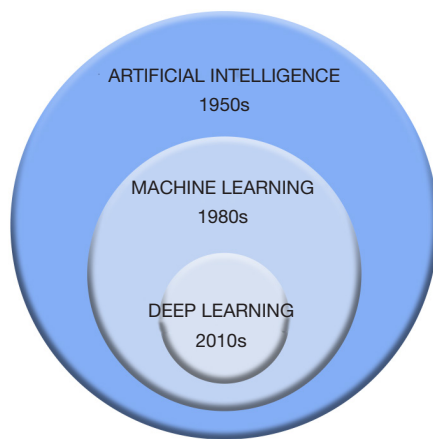
Figure 1 Schematic display of artificial intelligence (AI) and machine learning (ML) and deep learning (DL).

intersection of data science and computer science and aims at the efficiency of computing algorithms (2). Exceptional performances of ML can be achieved in a comparable short time on a dataset of comparably small size (1). It contains several steps, including data input, pre-processing, feature extraction, feature selection, classification and output interpretation (1-3). Conventionally, the subgroups of ML include supervised learning and unsupervised learning (2). To translate ML into the field of medicine, Deo Rahul suggested an additional division: learning on the subjects that physicians are advantageous at and the ones with limited success (2). In cancer research, ML has been increasingly used in predictive prognostic models (1).

DL is defined as a subset of ML targeting multilayer computation processes (1). It contains several layers, such as subsampling layer and convolutional layer (1). Of note, DL is less dependent on the understanding of data features than ML. Therefore, the algorithms of DL are much more difficult to interpret than ML, even potentially impossible (1).

To further delineate the current status of AI/ML/DL, we performed a search query regarding the number of published papers index in PubMed. The search query of AI, ML and DL was as follow: (AI[Title/Abstract]) AND (gastric cancer[Title/Abstract]), (Machine learning[Title/Abstract]) AND (gastric cancer[Title/Abstract]), (Deep learning[Title/Abstract]) AND (gastric cancer[Title/Abstract]), respectively. Interestingly, a noticeable increasing number of published papers relating to AI/ML/DL were found at the last three years (*Figure 2*). Although by definition, both

ML and DL belongs to AI, this result demonstrated AI has been increasingly popular in the field of gastric cancer.

Typical workups for individuals with suspected gastric cancer involve barium swallows, followed by upper endoscopy for visualization and biopsies (4). This opens the door for one of AI's most practical applications in the form of convolutional neural networks (CNNs) and their prospective integration into cancer detection with the aid of important instruments like the endoscope, and other imaging modalities (5).

The importance of integrating DL methods for improving detection of gastric cancer cannot be understated. Gastric cancer currently has the 5th highest total for number of new cases of all specific organs and ranks as high as 3rd for number of deaths (6). Both Western and Eastern Asia account for majority of cases worldwide, while Asian-Americans have a 5 times higher incidence, than other groups in the United States (6,7). The most common cause of gastric cancer remains *Helicobacter Pylori* infection, followed by chronic gastritis and diet (8). Pathologic sub-classification indicates that the most common types are adenocarcinomas and are also further classified into *intestinal type* associated with metaplasia and *diffuse type* associated with signet ring cells and linitis plastica (9).

Treatment options and strategies can be enhanced by DL. The quantities of specific information for the carcinogenesis of gastric cancer give it an edge over other cancer types, whose causes may not be wholly understood or completely deciphered. Knowing the most common cause, *H. pylori*, means finding ways to prevent gastric cancer are specific and allow for the predictive analytics arm of AI to be able to contribute. A more formidable challenge comes from finding therapeutic solutions in confirmed cancer cases, whether they are pharmaceutical or surgical therapy.

Finally, even though a basic understanding of gastric cancer has been attained, there's still much more information to uncover when it comes to the genomics of gastric cancer. The roles of genomic alteration by FGFR2, KRAS, and others need to be further examined (10). The expansion of potential biomarkers can also help elucidate an unclear picture, which not only can look at the cancer itself, but also how the causes of gastric cancer work as in the case of *H. pylori* (11,12). This information benefits not only the ability to diagnose, but also allows for specific targeted treatment. General problems do arise though when considering the scarcity of information and research being done in these areas. We present the following article in
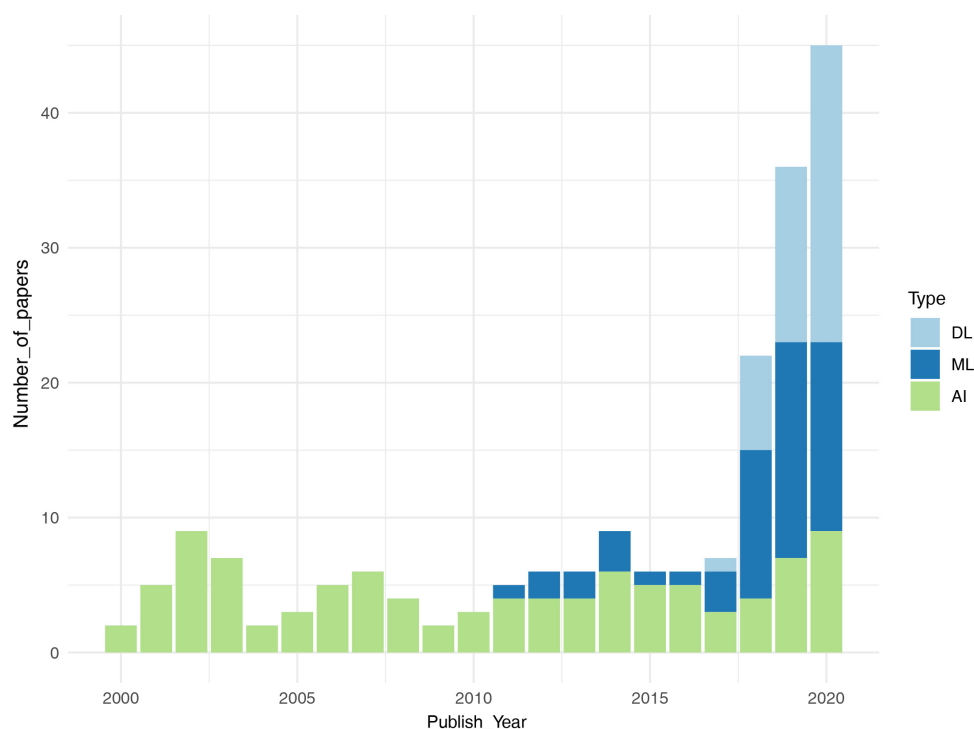
Figure 2 A search result of the number of published papers index in PubMed regarding artificial intelligence (AI) and machine learning (ML) and deep learning (DL).

accordance with the Narrative Review reporting checklist (available at http://dx.doi.org/10.21037/atm-20-6337).

The role of AI in diagnosis of gastric cancer

Important first steps are being undertaken with regards to integrating AI directly into the diagnostic process of gastric cancer. The first ventures into how we can improve existing medical technologies, namely, our visual modalities, have begun. The diagnostic process for gastric cancer is an integrated approach but ultimately synonymous with the endoscope. It remains the gold standard for diagnosing gastric cancer, thus, is the tool that has garnered most attention for integrating DL (13). Additionally, other aspects of imaging have been impacted by the CNN branch of AI. MRI and CT greatly benefit and enhance their value to the diagnostic process allowing for increased accuracy and confirmation. Lastly, the actual pathological diagnosis has shown promise that it can benefit greatly from AI. This ranges from how we understand the disease process, methods of how we reach conclusions, or even how we classify causes and cancers.

AI for the endoscopic diagnosis

When it comes to AI applications for endoscopy, Togashi believes that we are currently initiating a "Golden Era" (14). This is because there has been an explosion of research that has been initiated, which is currently setting the groundwork for the future. In the United States alone, AI-related research papers presented by members of the American Society of Gastrointestinal Endoscopy increased from 3 to 30 in one year's time (2017 to 2018), with similar trends in Japan. Research, he annotates, is still mostly confined to reading tests that analyze the sensitivity of AI detection algorithms through various image types, i.e., plain, magnified, augmented, or videos that make use of computer vision (14,15). In 2018, research topics from Japan Digestive Disease Week saw 8 presentations, 6 for colorectal cancer, 1 esophageal and 1 gastric. Sensitivity for detection ranged depending on experiments. Ranges for sensitivities in the 8 experiments were: 100% (colon N+ in T1), 90% (colon lesions), 97.6% (colon lesions), 95.7% (colon adenoma), 95.2% (colon adenoma), 82.4 to 90.4% (colon), 98% (esophageal cancer), and 93.7% (fundic gland polyp) (14). While boasting high sensitivities, specificity

remained inconsistent ranging from 43.1 to 99% (14). It should be noted that there is a severe lack of clinical trials, due mainly to the fact that this technology is fairly new, but also, ethical questions about computer-aided diagnosis (CAD) with AI have yet to be answered (16).

An important step forward in improving the endoscope as a tool has been through the improvements to video quality and video fidelity. Higher quality endoscopic vision allows for technicians to be able to identify structures and abnormalities more clearly by increasing resolution by roughly 16 times using 8K technology (7,680×4,320 pixels, or roughly 33 million pixels) compared to standard high definition (1,920×1,080 pixels, or roughly 2 million pixels) (17). Though added tools like increased zoom capability and increased field of view are helpful to surgeons, segmentation technology, powered by AI, provides an even bigger help than just image quality enhancements (18). Segmentation uses CNNs to identify and partition images or sections more clearly. Segmentation uses a metric for measuring and quantifying the accuracy of detecting an object appropriately called Intersection-over-Union (IoU), also referred to as the *Jaccard Index* (19). This is helpful when comparing the accuracy of different algorithms using the same dataset.

Ali *et al.* devised an "Endoscopic Artefact Detection challenge" to compare various algorithms to see if there was a method to objectively see how segmentation algorithms vary using a uniform dataset of video frames (20). Artefacts seen in the image dataset included objects like bubbles, image quality issues (pixel saturation), and organ debris (20). The overall goal of segmentation in a medical setting is to reduce any interference and allow for the highest quality interpretation of findings. Ideally, this leads to more specific and accurate diagnosing. Ali *et al.* postulate that technicians run into issues without segmentation technology, these include, different endoscopic modalities making tissues appear differently, physical phenomena, and variability and potential overlap of objects (20). Since segmentation works to address and recognize unique objects, it can present as a massive advantage to CAD, more so than seeing something clearly and still not being able to differentiate what is. The study found that the most important factors for identification were the size of the artefact and how much overlap with other objects is present (20).

Though the majority of endoscopy related AI research is heavily aimed towards colonoscopy, inroads have been made to emphasize the importance of upper GI applications. Hirasawa, *et al.* constructed a CNN system to automatically detect gastric cancer via endoscopic images (21). Formed around on the Single Shot Multibox Detector architecture, the algorithm used a dataset of 13, 584 endoscope images taken from 4 institutions over a span of 12 years (April 2004 to December 2016), which were inspected by a certified expert (21). To detect the accuracy of the algorithm, investigators compiled a further 2,296 images from 77 gastric lesions in 2017 to appraise and compare. The CNN featured a yellow rectangular box, which notified of "early or advanced gastric cancer" and location. Results showed that the CNN required 47 seconds to evaluate all images leading to 92.2% sensitivity (71/77), with missed lesions being superficially depressed or differentiated intra-mucosal cancers (21). False positives were caused by gastritis and anatomical variation (21). Ruffle *et al.* give insight to this type of experiment, which is referred to as *supervised learning*, which refers to a human providing the correct answer, i.e., this is *cancer*, which a model can be built around (22). The opposite of supervised learning is referred to as *unsupervised learning* and allows for specific information to be given to without specifying correct or incorrect answers allows conclusions to be drawn from patterns, more akin to *data mining* (22).

Applying DL to various endoscope modalities is another method by which researchers have improved the standard endoscope. Narrow band imaging (NBI) operates by altering standard red, green and blue optical filters (23). This allows for differing wavelengths and penetration leading to less red light being applied, and more blue and green wavelengths to penetrate, resulting in greater detail and visualization of surface mucosa and microvasculature (23). Li *et al.* created a system that incorporates CNNs for analyzing NBI images (24). The study involved observation by 2 experienced endoscopists diagnosing gastric mucosal lesions from patients using NBI to classify lesions as non-cancerous or early gastric based on (I) Vessels + surface classification system and (II) magnifying endoscopy simple diagnostic algorithm for early gastric cancer (MESDA-G) (24). This was followed by 2 pathologists reviewing histopathology of the lesions using the revised Vienna classification. Category 1 to 3 lesions were considered non-cancerous, while category 4 to 5 was considered early gastric cancer (EGC).

The study initially began by collecting images from 4 facilities resulting in 386 non-cancerous lesion images and 1,702 images of cancerous lesions, but led to issues of generalization so sample size needed to be increased to 20,000 images to appropriately train the algorithm (10,000

non-cancerous and 10,000 cancerous) (24). Algorithm architecture used the Inception v3 model, with changes made to boost accuracy by increasing image input type from 299×299 pixels to 512×512 pixels. Additionally, 4 individuals (2 experts and 2 non-experts) were utilized to interpret images and to divide them into non-cancerous or cancerous to be compared with the algorithm. 171 cases of non-cancerous lesions and 170 of cancerous lesions were used to see the diagnostic ability of the CNN algorithm. The CNN correctly interpreted 155 of 170 cancerous lesions (sensitivity of 91.1%) and 155 of 171 non-cancerous lesions (sensitivity of 90.6%), whilst 2 experts respectively identified non-cancer or cancer 78.2 and 81.1% of the time and the 2 non-experts 77.6 and 74.1% of the time. Lan, *et al.* note how much more sensitive and specific NBI is for diagnosing is compared to standard white light (WL) and the importance of incorporating AI into existing technologies, as early detection is important to patient outcome (24).

Additionally, researchers have been incorporating AI as an aid to diagnosing for certain endoscopic examinations. Ogawa *et al.* used a support vector machine (SVM) for diagnosing and classifying EGC (25). They attempted to try and reduce subjectivity of diagnosis by integrating 3D vectors with RGB values to help show color differences for areas believed to be cancerous or non-cancerous, which was assessed by an SVM. SVM is a practical AI tool, whose strength allows for categorizing and classifying information. Standard endoscopy is usually done under WL, whilst indigo carmine chromoendoscopy (indigo) allows for general cancer diagnosing, and acetic acid-indigo carmine chromoendoscopy (AIM) allows mucus differentiation and boasts a high diagnosability (25). The SVM was incorporated to allow for discrimination in using different endoscopic techniques.

18 subjects were admitted for endoscopy using WL, Indigo, and AIM followed by resection. Images were identified histopathologically and diagnosed as EGC. Imaging was performed and captured in stepwise order, first by standard WL endoscopy, then indigo carmine was introduced and distributed for 12 seconds, then finally, acetic acid was distributed and captured (25). 54 images from the 18 patients' lesions were assessed according to macroscopic and histopathology to determine an area where cancer was found. The researchers extracted pixels randomly from cancerous and non-cancerous images, which were further evaluated by having each pixel calculated using RGB values in order to ensure there was not a major

difference in luminance. In total, 2,000 pixels with equal luminance from both cancerous and non-cancerous images were taken (4,000 total), which are then represented as 3D vectors that have an RGB value. The 2,000-pixel samples for cancerous segments allowed for the calculations of *mean vector* and *covariance matrix*, and the same was done with noncancerous. Furthermore, the measure of Mahalanobis distance, which is a clear delineation via color differences of cancerous or noncancerous areas, was done for WL, indigo, and AIM. The mean distance between different types of modalities were compared amongst themselves (25).

One hundred samples from cancerous and 100 from noncancerous were used for training the LIBSVM and the remaining 3,800 for testing, with the main goal of determining whether the area was cancerous or noncancerous (25). A value was created to determine diagnosability, which was composed of the F1 measure by finding the harmonic mean of both sensitivity and PPR. A total of 54 images from the 3 compared modalities were randomly sorted and given to endoscopists who were tasked with diagnosing gastric cancer and then were a ratio from each type was created to compare with. Mahalanobis distances and F1 measures from the 3 modalities were: (I) WL – MD 1.52 and F1 0.636, (II) indigo – MD 1.32 and F1 0.618, and (III) AIM – MD 2.53 and F1 0.687. AIM was determined as having the highest diagnosability based on enhanced color difference. AIM images were also determined to be the best source for subjective interpretation from endoscopists, as mean rates for diagnosing gastric cancer correctly for AIM were 83.3%, when compared to WL at 50.0% and indigo at 52.2% (25).

The benefits of early cancer detection, especially gastric cancer, cannot be overlooked. Since *H. pylori* remains the most common cause, it is important to focus on early detection of *H. pylori* with the endoscope. Zheng *et al.* used CNN modeling accuracy in the setting of detecting *H. pylori* via endoscope (26). Initial sampling of patients excluded those who had previous cancer, submucosal tumors, peptic or endoscopic ulcers, masses or strictures, in addition to those who had been given PPIs within 2 weeks or antibiotics within a month. Those included needed to have immunohistochemistry staining on all samples for *H. pylori* and if no evidence was found on the stains, then a positive *H. pylori* breath test was needed within 1 month, before or after the sample was collected. Image collection was done by standard endoscope, with images screened by experts for quality assurance and augmented to help improve accuracy for model training. The foundation for the algorithm comes

from ResNet-50, a CAD support system made up of up to 50 layers, and was trained using PyTorch, DL framework and further incorporated ImageNet to assist in overfitting issues (26).With the model constructed, a validation cohort was used for determining the precision of the CNN and endpoints were defined as sensitivity, specificity, and accuracy of the trained CNN, with even more detail supplemented by the addition of subgroups added based on anatomical location and mode of biopsy.

A total of 1,959 patients were found to be fitting of requirements and included in the study. The most common characteristics in the patient population were male, 48.5 years old, outpatient, and *H. pylori* documented on breath test or biopsy. In the end, 15,484 images were used, and comparisons made for CNN using a single gastric image or multiple gastric images. For single gastric image, AUC was 0.93 (95% CI: 0.92 to 0.94), while sensitivity was 81.4% (95% CI: 79.8% to 82.9%), specificity was 90.1% (95% CI: 88.4% to 91.7%), and accuracy was 84.5% (95% CI: 83.3% to 85.7%) at an optimal cutoff value of 0.28 (23). For multiple images, an average of 8.3 images were used and AUC for ROC was calculated at 0.97 (95% CI: 0.96 to 0.99), furthermore, using an optimal cutoff value of 0.4, sensitivity was 91.6% (95% CI: 88.0% to 94.4%), specificity was 98.6% (95% CI: 95% to 99.8%) and accuracy was 93.8% (95% CI: 91.2% to 95.8%) (26). This reinforces the idea that multiple images are more helpful when evaluating with CNNs. Subgrouping for anatomical location showed that in multiple images from antrum biopsies alone (398 patients) showed sensitivity at 91.7% (95% CI: 87.9% to 94.6%), specificity at 98.2% (95% CI: 93.5% to 99.8%), and accuracy at 93.5% (95% CI: 90.6% to 95.7%), while antrum and body biopsies (54 patients) showed sensitivity at 90.5% (95% CI: 69.6% to 98.8%), specificity at 100% (95% CI: 89.4% to 100%), and accuracy at 96.3% (95% CI: 87.3% to 99.5%) (26). There proved to be no difference in the 3 parameters when comparing antrum alone to antrum and body biopsies. Zheng *et al.* postulate that CNNs may eventually be used to evaluate for *H. pylori* infections without the use of biopsies through its image recognition ability, with the caveat of having to use multiple gastric images to do so.

AI in the MRI/CT imaging

While a plethora of research has been undertaken to enhance the endoscope, the same can't be said about MRI or CT imaging. Though it is understandable that the focus should be squarely on improving our current gold standard, there have been examples of how MRI/CT imaging can prove to be helpful in other aspects of gastric cancer detection. Yu *et al.* show that enhanced CT may have a place in predicting T stage for patients with gastric cancer (27). By using a group of 40 patients, all with confirmed gastric cancer, enhanced CT imaging was undertaken that included enhanced CT scanning, ingestion of anisodamine (to slow peristalsis), and a gas-generating agent was given to help inflate stomach. Images yielded were independently observed and a double-blind diagnosis was made by 2 senior associate professors to determine T staging. Subsequently, surgery was undertaken to remove the neoplasms from participating individuals and the samples collected were handled by using the pathology standard of cutting them into sections and applying H&E staining so that a pathologist can observer microscopically, and classify them by appropriate TNM staging criteria (27). Of the 40 cases, 5 were T1, 9 were T2, 20 were T3, and 6 were T4, according to pathology. CT staging alone was found to be 75.00% accurate (85.00% when combined with doppler ultrasound with echo-type contrast agents). Yu, *et al.* conclude that there can be a space for CT imaging use when applying to preoperative T staging. This begs the question of why so little research is currently being done to try to incorporate other imaging modalities in the general process.

In other branches of oncology, CT and MRI imaging remain essential components. The use of mpMRI (multiparametric magnetic resonance imaging) helps with diagnosing prostate cancer. Early incursions into applying AI to MRI modalities concentrate on detecting suspicious zones or classifying findings. By using T2-weighted imaging, diffusion weighted imaging, and dynamic contrast enhanced imaging of the prostate and nearby sectors, urologists have been aided in their biopsy techniques allowing for better disease sampling and in the surgical setting for aiding in radical organ removal (28). Breast cancer also has attempted to incorporate mpMRI for detection by trying to combine DL machines into the formula. The previously mentioned T2-weighted imaging and dynamic contrast have been powered by CNNs and integrated into the CAD process, allowing for improvement in the ways benign and malignant growths are distinguished (29). Models for incorporating low-dose CT scans for detecting pulmonary nodules for lung cancer have also proved insightful. Taking specific features from image sets, a model was made to help predict malignancy in pulmonary nodules (30). The same underlying idea has been applied

to reduced-dose PET for lung cancer (general), which showed a potential to drastically reduce radiation doses given to patients (31). Of the top 5 most common cancer types world-wide, both gastrointestinal cancers (gastric and colorectal) seem far removed from the current zeitgeist and still leave a lot to be desired with regards to their lack of incorporation of CT or MRI imaging.

AI in the pathological diagnosis

The understanding of gastric cancer at a pathological level has seen many advances through the incorporations of genetic and histopathological components. These promising areas of pathological diagnosis would be some of the most benefited if the appropriate strategies for DL integration was undertaken. The current feeling is optimistic as to how diagnoses will be impacted. This includes a shift from simple microscopic imaging to digitized microscopy, the reproducibility of certain biomarkers for consistency, digital image analysis, new insight on causes of disease, and even how certain nuclear characteristics can predict survival outcomes (32).

Thought it is impossible to replicate human thinking and problem solving, attempts are being made into trying to replicate a pathologist's ability to perceive information and to draw worthwhile conclusions even from small datasets. Using CNNs, Qu *et al.* suggest that a big contributor for classification failures in cancer can be due to histological shape extraction, or parameters for generalized texture images, while computer vision, though initially burdensome to train and calibrate, can relieve the time and financial burden and essentially imitate pathologist thinking and even find new pathology-relevant information (33). CNN foundation is derived from neocognitron, which simulates the visual primary cortex allowing for pattern recognition in the living brain (34). Standard fine-tuning procedures were done to train for discovering image features, though the task of finding the appropriately attunement for distributional difference between standard pre-training data sets and actual targets proves difficult. This, coupled with the idea of *pathologist's intuition*, shows the large effort needed when trying to bring both ML and human thought processes together.

The stepwise process involved the important first step of understanding basic tissue structures before differentiating between benign and malignant. This included understanding spreading status, density of cells, degree of nucleus distortion, nucleus size and nuclear to cytoplasm

ratio (33). This basic understanding allowed for the CNN to acquire pathological information rather than being given it as preset that moved it from large-scale pretraining set interpretation to higher, or "middle-level", datasets with more specificity and information. 2 fine-tuning steps are needed to advance from "low-level" to "middle-level" to "high-level". Stroma and Epithelium-based datasets are thought to be the most cost-effective yet useful, as it is seen in every organ and can be easily categorized in cancer scenarios (33). Acquiring "middle-level" status was done by indexing morphological features of the nucleus. Color deconvolution was done to separate H&E stain into their own vectors H Stain and E Stain since most nuclei appear on the H stain. Other algorithms are added to enable the separation of the channels, separating conjoined or overlapped nuclei, and contour detection. A total of 48,000 images were separated into the channels of background, stroma and epithelium, while cell data about the nuclei was made up of 7,672 and 6,457 patches that were selected due to consistency. Using tissue data, AUC value in small data set groups rose by 0.035, 0.016, and 0.039 when rival architectures VGG-16, AlexNet and InceptionV3 were added, while in large date set groups AUC rose by 0.027, 0.053, and 0.053. Using nuclei data, AUC value in small data sets rose 0.023, 0.024, and 0.034, still using VGG-16, AlexNet and InceptionV3 and rose again in large data set groups to 0.029, 0.048, and 0.052 (30). This showed that the stepwise method for both tissue and cell-based data sets were able to enhance the pretrained neural networks for classification purposes.

While replacing pathologists for CNN-based systems is not currently plausible, what is plausible is pathologists using their experience to try and understand metastasis at a deeper level. The most common sites for metastasis in gastric cancer are liver, peritoneum, lung and bone (35). Gastric metastases rates have increased vastly in certain countries, with a median survival of 16 months, and give a further unclear picture due to the practice of M stage being a binary function (yes or no) with no detail into origin (35). An incursion into detecting metastatic lymph node cancer using neural networks was undertaken by Gao and associates. Gao *et al.* began with 750 patients who were diagnosed with gastric cancer, underwent CT imaging and after 2 DL sessions were able to extract 20,151 images from 313 patients in the 1st session and 12,344 images from 189 patients in the 2nd session (36). Three experienced radiologists labeled 1,371 images of metastatic lymph nodes, then another radiologist labeled node shape

(round, irregular or spiculated), sizes (greater than 10 mm), and enhancement density difference. Labeling was further evaluated by 3 radiologists and 1 surgeon and then submitted for use in DL for establishing a database and algorithm training. A total of 1,004 images were later added to improve accuracy in identification.

Characteristics of the patients were 72.4% male with the highest incidence being poorly differentiated gastric adenocarcinoma of lower gastric stomach. Precision and recall rates were generated for both training and validation set. AUC was found to be 0.5019 (AP =0.5019). In addition, to show regression and classification true and false positives were counted under different probability thresholds and plotted (ROC curves). Results showed an AUC of 0.8995. Training set and validation set precision and recall rate was also plotted. AUC of 0.7801 indicated that CNN training had advanced. True and false positives were again looked at in the test set and calculated and plotted. AUC was 0.9541 (compared to first AUC of 0.8995) (36). Researchers believe that this application of CNNs is more accurate and specific, will aid heavily pre-operative neoadjuvant chemotherapy, and improve methods and strategies of lymph node dissections.

Genetic pathology is a discipline that potentially has the most to gain from information that can potentially be yielded when DL data interpretation is applied. While known determinants involved in gastric cancer like *H. pylori*, natural environment, and diet play a large role in carcinogenesis, the role of IL-1β, IL-10, TFF2, and CDH1 cannot be ignored (37). The impact of genes, biomarkers and their interpretation hold the key to understanding and treating gastric cancer. Jiang *et al.* used a method of tabulating gene data from patients with gastric cancer and normal patient to try and identify certain genes or roles they play by analyzing multiple transcription datasets. Datasets came from research done on the topic and uniform datasets were ranked and analyzed using RankProd and INMEX. Gene Expression Omnibus database was used to find gene expression data then research, literature analysis, and bioinformatics data was searched to find promising and related information. Geno Ontology and pathway analysis were added to deepen the function. After rounds of exclusion, 3 microarray datasets were refined to find 1,153 differentially expressed genes (787 for downregulation and 366 for upregulation). Genes were then ranked by most upregulated or most downregulated yielding Progastricsin (PGC) and collagen type VI alpha 3 chain (COL6A3), respectively. PGC and COL6A3 can potentially serve as a biomarker for gastric cancer going forward, along with a slew of others that that deal with translational elongation, respond to chemical or external stimuli, generate energy, form part of extracellular matrix or cytoplasm, protein binding, cell-cell adhesion, and many more (37). Applications like this are essential for time saving and to provide researchers with potential starting points or genes of interest.

Predictive models can help inform or define pathology and can be especially useful when understanding pathogenesis or oncogenesis. Screening methods and data collection can play an important role to elucidate EGC pathology. This varies from simple ideas like age, gender, diet to more complex like symptoms, history, or genetics. Incorporating datamining techniques, Liu *et al.* recruited 618 patients and devised a questionnaire that included demographic characteristics, eating habits, symptoms and family history and additionally serological examination and gastroscopy (38). Keywords and topics of interest were eventually whittled down to make a list of 34 standout factors that could be used to make predictive models. Random sampling allowed for the creation of training and data sets that could be used to create and evaluate models so future models would yield relevant information. Since there was heavy skew in favor of low risk EGC, synthetic minority oversampling technique was employed which can help with SVMs, decision trees and network types. Results of data collection yielded useful information on high risk EGC including main symptoms of belching, reflux, and postprandial discomfort, increased family history of hyperlipidemia, while also yielding potentially less useful information for pathology purposes such as speed of eat or spoken language of individuals (38). The standard issues of information sorting that always plague datamining of any type are painfully apparent.

While limiting the amount of radiation exposure to patients or having the ability to run thousands of calculations at once are major benefits of the technology, there are various ways AI can provide an advantage to medical experts and patients (21,31). Vollmer *et al.* pose 20 questions ranging from implementation, statistical methods, and reproducibility, amongst others, to try and more clearly find the benefits for patients (39). The belief is that if there is a workable framework with a common technical language and a reliance on empirical evidence, then there will be benefits to patients and health systems. Additionally, we can see that, if properly constructed, computer-based systems can help reduce error rates, work more efficiently, do not

overexert human resources, and have the ability to improve itself by altering the algorithm. Ahuja believes an area that might see the most growth is personalization of patient care and treatment, which is more tailored to individual patients to allow for individual care plans that feature improved efficiency and accuracy, due to the amounts of data that can be drawn or calculated by ML (39). One potential "advantage" that we overlook, especially in computer-aided diagnoses, is the ability of AI to only observe facts and remove the human or emotional element from important decision making.

The role of AI from endoscopic diagnosis to treatment

A practical guidance of adapting ML into clinical endoscopy for an accurate diagnosis of gastrointestinal disease has been proposed by van der Sommen et al. (40). To fully appreciate the weight of ML upon the challenge of gastrointestinal diagnosis by endoscopy, solid technique basis is required, particularly for each clinician (40). Anatomically, the stomach is distinct from other gastrointestinal organs, such as colon and esophagus (41). It is characterized by wider bent lumen with potential blind spots that require extra laborious observations. Clinicians therefore require multiple distant sights to avoid possible omitting (41). Additionally, helicobacter pylori infection may also hide some early appearance of EGC (41). These difficulties lead to variation of endoscopic diagnosis as reported (42,43). Moreover, this also reminds us that simply translating the AI of colon cancer into gastric cancer may not be the right scenario.

In fact, endoscopy also serves as a therapeutic option for gastric cancer, such as endoscopic mucosal resection (EMR) of EGC (44). Marked by minimally invasiveness, EMR is a standard treatment in gastric cancer with low risk of lymph node metastasis in Japan and has been increasingly popular in the West world (44-46). Notably, with local lesions >15 mm, EMR may lead to difficulty in the assessment of tumor depth and an increase of local recurrence (44). Meanwhile, another endoscopic dissection technique, endoscopic submucosal dissection (ESD), has been introduced for en bloc resection of larger lesions (44). To some degree, ESD has been a strong competitor to conventional open/ laparoscopic surgery for EGC (44).

To further translate AI into the clinical practice of endoscopic resection, Zhu et al. constructed a convolutional neural network computer-aided detection (CNN-CAD) system with high accuracy and specificity (47). But, the specific role of AI in the endoscopic resection process remains sparse. Although AI-based detection system can reduce unnecessary gastrectomy with precisely prediction of the depth of tumor invasion, it may not navigate the endoscopic resection procedures, or set up alarms in high risks of complications. Indeed, ESD-related complications, including perforation, peritonitis and bleeding, remain a major challenge in gastric cancer with 3.5% rate (48).

Of note, Odagiri and Yasunaga also indicated a linear association between a lower frequency of ESD-related complications and a higher hospital volume (48). As both ESD and EMR require a comparably high level of operational skills, it is reasonable and reliable that both design and training of AI-based techniques should be assigned to hospitals with high volume in the first place, especially the techniques containing ML or DL that require sufficient data training.

In an endoscopic clinical guide of AI, Namikawa et al. summarized the usage of AI in the field of stomach, including clinical detection, classification, and blind spot monitoring (49). They also predicted in perspective that future AI could be fully trained to distinguish from gastric neoplastic and non-plastic lesions and make more significant contribution (49).

However, the use of AI in the treatment of gastric cancer remains sparse. Unlike the endoscopic diagnosis of gastric cancer that heavily relies on the interpretation of pictures, the implication of AI in chemoradiotherapy may require multidimensional data interpretation, including genomic characterization, immunohistochemistry results, mutation analysis or insensitivity prediction. Drug responsiveness, another topic featured by clinical heterogeneity and discrepancies, is now open to AI. Joo et al. constructed a 1-dimensional convolution neural network model, DeepIC50, to deliver a reliable prediction of drug responsiveness in gastric cancer (50). They validated the tool both in cell lines and a real gastric cancer patients' dataset and achieved a comparable result (50).

The role of AI in surgery

The implication of AI in surgery may rely on additional clinical and physical data for surgical training and improvement of performance. A review by Jin et al. indicated that the future application of AI in surgery will likely be focused on surgical training, skill assessment and guidance (51). In robotic surgery, Andras et al. indicated that ML approaches enhanced the improvements of surgical skills, quality of surgical process and guidance with favored

Table 1 Studies of artificial neural network (ANN)-based models for the prognostic prediction of GC patients

Studies	Objectives/nation	Methods	Major findings
Korhani Kangi and Bahrampour, 2018 (56)	339 GC [2001–2015]/Iran	ANN versus BNN	Sensitivity and specificity of ANN: 0.882, 0.903; sensitivity and specificity of BNN: 0.954, 0.909; prediction accuracy and AUC of ANN: 0.891, 0.944; prediction accuracy and AUC of BNN: 0.935, 0.961
Biglarian et al., 2011 (57)	436 GC [2002–2007]/Iran	ANN versus CPH	True prediction of ANN: 83.1%; true prediction of CPH: 75%
Amiri et al., 2013 (58)	330 GC/Iran	ANN versus CPH/KM	No significant difference of the ratios of SE between ANN and CPH; no significant difference of the ratios of SE between ANN and KM
Nilsaz-Dezfouli et al., 2017 (59)	452 GC/Iran	ANN	1-year prediction of sensitivity, specificity, accuracy of ANN: 0.707, 0.962, 0.903; 2-year prediction of sensitivity, specificity, accuracy of ANN: 0.844, 0.932, 0.889; 3-year prediction of sensitivity, specificity, accuracy of ANN: 0.896, 0.890, 0.894
Oh et al., 2018 (60)	1,243 GC [2007–2010]/Korea	ANN (SRN)	AUC of SRN (5th year): 0.81
Que et al., 2019 (61)	1,608 GC [2011–2015]/China	Preope-ANN	C-index, likelihood ratio chi-square and AUC of the preope-ANN were all superior to clinical TNM. Prediction efficiency of the preope-ANN was similar to pathological TNM
Yazdani Charati et al., 2018 (62)	430 GC [2006–2013]/Iran	ANN	AUC of selected ANN: 0.94; prediction of selected ANN: 0.92

GC, gastric cancer; ANN, artificial neural network; BNN, Bayesian neural network; AUC, area under the curve; KM, Kaplan-Meier; SE, standard errors; SRN, survival recurrent network.

postoperative outcomes (52). Also, surgical experience and movements can be optimized with tension-sensitive robotic arms and the integration of augmented reality methods (52). To further automatically classify surgeons' expertise, Fard et al. built a ML-based classification framework with movement trajectory to demonstrate whether a surgeon is novice or expert (53). Meanwhile, AI also has the power to revolutionize how surgery will be taught in future (54). Interestingly, O'Sullivan et al. envisioned that a routine surgery can be performed independently by a robot supervised by a human surgeon (55).

The role of AI models in the prognostic prediction

Next, AI has also been applied to the prognostic prediction of gastric cancer aiming for individualized administration. Conventionally, cox proportional hazard (CPH) model is used for the prediction of risk scores. A nomogram model is established based on CPH to further set up individualized prediction. Nonetheless, all these algorithms are merely based on linear equation. When taking non-linear and biological complexity into consideration, it is both reproducible and applicable

to introduce AI into the establishment of non-linear statistical models. Up to now, artificial neural network (ANN) is one of the most used models for prognostic prediction. In fact, a series of ANN-based models have been published with distinct clinical values (56-62) (Table 1). Several studies have demonstrated the values of AI models in the prediction of gastric cancer prognosis, such as ANN and Bayesian neural networks (BNN) (56,57,62). Amiri et al. reported that no significant difference in the ratios of standard errors between ANN versus CPH, or ANN versus Kaplan-Meier (KM) method (58). Specifically, ANN with predefined nodes (3,5,7 nodes) in the hidden layer showed no significant better prognostic prediction than conventional methods like KM (58). However, authors also repeated that sample-size could negatively impact the accuracy of prediction (58). Moreover, Nilsaz-Dezfouli et al. highlighted a distinct predictive value of ANN model for the outcome of gastric cancer after 1–5 years, however, without comparison to other conventional methods (59). Similarly, Oh et al. reported a more powerful tool for prognostic prediction using a survival recurrent network (SRN) (60). Interestingly, Que et al. established a preoperative ANN (preope-ANN) model to preoperatively

predict the prognosis of gastric cancer (61). It provided a reliable prediction for long term survival of gastric cancer prior to surgery with distinct statistical superiority to both clinical and pathological TNM (cTNM, pTNM) (61). Without a doubt, much more data is required for further upgrading of ANN models.

Another particularly important thing we should not overlook are the potential risks involved with incorporating AI models into the prognostication process. Though obviously there are massive benefits to eventually perfecting this new technology, it is not without its shortcomings and many questions still need to be addressed regarding its true implementation. Challen *et al.* point to 3 potential timeframes in which issues can arise: short term issues, medium term issues and long term issues (63). Macrae summarizes all timeframes and states that short term issues come from "reliability and interpretability of predictions" and how the systems used for training may not necessarily reflect actual real-world scenarios (64). In the medium term, the problems are thought to be an overreliance on ML and the further need of to have massive amounts of information to feed algorithms. Finally, in the long term the issue of autonomy with systems finding ways to independently get to a certain desired conclusion but in an alternative way leading to potential adverse effects. Hamid also points towards the veracity of issues arising short and medium term and adds that accuracy and security are big questions we need answered while this technology is still in the infant stage (65).

Naturally, we must be vigilant and try to heed the advice of AI ethicists, though these positions of authority are hardly fleshed out now. For short term issues, we can plausibly continue to improve testing and hold off full implementation until real life scenarios can be achieved. The more tests and simulation, the more information we can retrieve and the better the algorithms will be. Stepwise is most wise. Real issues arise in addressing the challenges of the medium term, as it is only natural for human behavior to completely replace old methods with new ones, especially, when dealing with convenience and time saving. Trends of this behavior have been apparent for centuries, for example, abacus to calculator or telegraph to video calls. Prevention of overindulging and becoming overly dependent on AI lays squarely with our behavior and choices. As for long term issues, we can only aim to include safeguards in algorithms to prevent any independent activity or creating a system to oversee any changes the ML modality may be implementing (63,65).

Clinicians are the key roles to oversight the improvement of performance of AI in both diagnosis and treatment area. The rapid development of AI is a reciprocal benefit for both patients and clinicians. Initial training of AI is mainly based on the data fed by each clinician, which may be overburden, but the long-term benefits far exceed the initial cost. AI provides a standardized and united health care system for diagnosis and treatment and offers an improvement opportunity for all the rest clinicians with comparable weak experiences (66).

Future

For this technology to be truly commonplace, much growth is still needed. Incorporation of this magnitude requires time to calibrate and make systems that are not only accurate, but also user-friendly and accessible (67). The biggest challenge for AI and other forms of DL is that they require information to learn and improve as it essentially nourishes itself from the quantities of data it is given (68). Issues arrive due to lack of availability of datasets, which in turn, potentially limit its growth (68). Tools like data mining can help fill the void to draw meaningful conclusions from available datasets but also suffer some drawbacks since more datasets yield more specific and meaningful data (69).

The underlying importance of accessibility cannot be overstated, especially in geographic regions where resources are more limited. Currently in the United States alone, the number of CT scanners increased from 25.9 scanners per million to roughly 45 scanners per million in 2019 (70). This is in stark contrast to Nigeria, Africa's most populous country, which boasted a population of 195.9 million in 2018, and in the same year had a total of 183 machines for the whole country, or slightly less than 1 per million (71). Naturally, the question arises of how AI strategies can be implemented to help. Data mining strategies seem like a natural fit to yield useful information. The simplicity of being able to draw conclusions from any information fed to the algorithm such as age, local diet, gender, tribe/ethnic group or seemingly unrelated variables like height or economic status will always yield patterns, highlight abnormalities and show trends that seemingly go unnoticed. This serves as a perfect starting point to gauge the health status of a geographic group. Since the population is vast (almost 200 million), there is a massive benefit of having a lot of raw data to draw from, with the real drawback being limited ways of getting that information (183 machines). Using strategies like this, local physicians can be better

informed about their population and make more specific and informed decisions to benefit their local communities. A real push for access and education is needed from developed countries to ensure that people in other regions can have access and the tools they need.

Special considerations also must be taken into the quality of data that is being given. As image quality is essential for training recognition, CNN architectures are still being upgraded. Improvements are being seen with textures to allow for increased accuracy and segmentation technology that help clearly differentiate objects (72,73). DL machines must also be taught and standardized by diagnostic professionals, which means there is an added effort and strain of having to train and provide appropriate calibrations for the machines by experts (74). This also does not take into consideration money needed by healthcare providers and researchers to build and maintain these systems to eventually yield relevant findings (75).

In summary, the present studies of AI in gastric cancer are mainly focused on the diagnosis. The excellence of AI, at this period, is more attractive than widely clinically applicable. Perhaps a scarce amount of AI researchers was originated as clinicians if any at all. Perhaps the developmental trajectory of minimally invasive laparoscopic surgery in clinical practice could be translated into AI and all computer-aided techniques. It is perceivable that, in near future, almost all clinical challenges may be able to be completely transformed as AI-based questions. This may suggest that, although imbalanced global medical resources affect the widespread of AI, advanced regions with adequate AI and clinical resources may have considerably accelerated the development of AI. Clinicians with AI background may serve as the key player across all AI progression. They can largely discount the negative impact of accident cases in their professional lives by AI-based regurgitating and digesting and deliver standardized clinical administration.

Conclusions

Despite growing efforts, adapting AI to improving diagnoses for gastric cancer is a worthwhile venture. The information yield can revolutionize how we approach gastric cancer problems. Though integration might be slow and labored, it can be given the ability to enhance diagnosing through visual modalities and augment treatment strategies. It can grow to become an invaluable tool for physicians, but this is entirely dependent on people to adapt and teach it.

References

1. Huang S, Yang J, Fong S, et al. Artificial intelligence in cancer diagnosis and prognosis: Opportunities and challenges. Cancer Lett 2020;471:61-71.
2. Deo RC. Machine Learning in Medicine. Circulation 2015;132:1920-30.
3. Wong D, Yip S. Machine learning classifies cancer. Nature 2018;555:446-7.
4. Layke JC, Lopez PP. Gastric cancer: diagnosis and treatment options. Am Fam Physician 2004;69:1133-40.
5. Zhu R, Zhang R, Xue D. Lesion detection of endoscopy images based on convolutional neural network features. 2015 8th International Congress on Image and Signal Processing (CISP), Shenyang, 2015:372-6.
6. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence

and mortality worldwide for 36 cancers in 185 countries [published correction appears in CA Cancer J Clin 2020;70:313]. CA Cancer J Clin 2018;68:394-424.

7. Taylor VM, Ko LK, Hwang JH, et al. Gastric cancer in asian american populations: a neglected health disparity. Asian Pac J Cancer Prev 2014;15:10565-71.

8. Venerito M, Ford AC, Rokkas T, et al. Review: Prevention and management of gastric cancer. Helicobacter 2020;25:e12740.

9. Sitarz R, Skierucha M, Mielko J, et al. Gastric cancer: epidemiology, prevention, classification, and treatment. Cancer Manag Res 2018;10:239-48.

10. Deng N, Goh LK, Wang H, et al. A comprehensive survey of genomic alterations in gastric cancer reveals systematic patterns of molecular exclusivity and co-occurrence among distinct therapeutic targets. Gut 2012;61:673-84.

11. Nagarajan N, Bertrand D, Hillmer AM, et al. Whole-genome reconstruction and mutational signatures in gastric cancer. Genome Biol 2012;13:R115.

12. McClain MS, Shaffer CL, Israel DA, et al. Genome sequence analysis of Helicobacter pylori strains associated with gastric ulceration and gastric cancer. BMC Genomics 2009;10:3.

13. Karimi P, Islami F, Anandasabapathy S, et al. Gastric cancer: descriptive epidemiology, risk factors, screening, and prevention. Cancer Epidemiol Biomarkers Prev 2014;23:700-13.

14. Togashi K. Applications of artificial intelligence to endoscopy practice: The view from Japan Digestive Disease Week 2018. Dig Endosc 2019;31:270-2.

15. Szeliski R. Computer vision: algorithms and applications. Springer Science & Business Media. 2010.

16. Pesapane F, Volonté C, Codari M, et al. Artificial intelligence as a medical device in radiology: ethical and regulatory issues in Europe and the United States. Insights Imaging 2018;9:745-53.

17. Tanioka K. 8K imaging systems and their medical applications. In Proc. Int. Image Sensor Workshop 2017.

18. Yamashita H, Aoki H, Tanioka K, et al. Ultra-high definition (8K UHD) endoscope: our first clinical success. Springerplus 2016;5:1445.

19. Eelbode T, Bertels J, Berman M, et al. Optimization for Medical Image Segmentation: Theory and Practice When Evaluating With Dice Score or Jaccard Index. IEEE Trans Med Imaging 2020;39:3679-90.

20. Ali S, Zhou F, Braden B, et al. An objective comparison of detection and segmentation algorithms for artefacts in clinical endoscopy. Sci Rep 2020;10:2748.

21. 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;21:653-60.

22. Ruffle JK, Farmer AD, Aziz Q. Artificial Intelligence-Assisted Gastroenterology- Promises and Pitfalls. Am J Gastroenterol 2019;114:422-8.

23. Singh R, Owen V, Shonde A, et al. White light endoscopy, narrow band imaging and chromoendoscopy with magnification in diagnosing colorectal neoplasia. World J Gastrointest Endosc 2009;1:45-50.

24. Li L, Chen Y, Shen Z, et al. Convolutional neural network for the diagnosis of early gastric cancer based on magnifying narrow band imaging. Gastric Cancer 2020;23:126-32.

25. Ogawa R, Nishikawa J, Hideura E, et al. Objective Assessment of the Utility of Chromoendoscopy with a Support Vector Machine. J Gastrointest Cancer 2019;50:386-91.

26. Zheng W, Zhang X, Kim JJ, et al. High Accuracy of Convolutional Neural Network for Evaluation of Helicobacter pylori Infection Based on Endoscopic Images: Preliminary Experience. Clin Transl Gastroenterol 2019;10:e00109.

27. Yu T, Wang X, Zhao Z, et al. Prediction of T stage in gastric carcinoma by enhanced CT and oral contrast-enhanced ultrasonography. World J Surg Oncol 2015;13:184.

28. Harmon SA, Tuncer S, Sanford T, et al. Artificial intelligence at the intersection of pathology and radiology in prostate cancer. Diagn Interv Radiol 2019;25:183-8.

29. Hu Q, Whitney HM, Giger ML. A deep learning methodology for improved breast cancer diagnosis using multiparametric MRI. Sci Rep 2020;10:10536.

30. Choi W, Oh JH, Riyahi S, et al. Radiomics analysis of pulmonary nodules in low-dose CT for early detection of lung cancer. Med Phys 2018;45:1537-49.

31. Schwyzer M, Ferraro DA, Muehlematter UJ, et al. Automated detection of lung cancer at ultralow dose PET/CT by deep neural networks - Initial results. Lung Cancer 2018;126:170-3.

32. Acs B, Rantalainen M, Hartman J. Artificial intelligence as the next step towards precision pathology. J Intern Med 2020;288:62-81.

33. Qu J, Hiruta N, Terai K, et al. Gastric Pathology Image Classification Using Stepwise Fine-Tuning for Deep Neural Networks. J Healthc Eng 2018;2018:8961781.

34. Fukushima K. Artificial vision by multi-layered neural

networks: neocognitron and its advances. Neural Netw 2013;37:103-19.

35. Riihimäki M, Hemminki A, Sundquist K, et al. Metastatic spread in patients with gastric cancer. Oncotarget 2016;7:52307-16.

36. Gao Y, Zhang ZD, Li S, et al. Deep neural network-assisted computed tomography diagnosis of metastatic lymph nodes from gastric cancer. Chin Med J (Engl) 2019;132:2804-11.

37. Jiang B, Li S, Jiang Z, et al. Gastric Cancer Associated Genes Identified by an Integrative Analysis of Gene Expression Data. Biomed Res Int 2017;2017:7259097.

38. Liu MM, Wen L, Liu YJ, et al. Application of data mining methods to improve screening for the risk of early gastric cancer. BMC Med Inform Decis Mak 2018;18:121.

39. Vollmer S, Mateen BA, Bohner G, et al. Machine learning and artificial intelligence research for patient benefit: 20 critical questions on transparency, replicability, ethics, and effectiveness [published correction appears in BMJ 2020 Apr 1;369:m1312]. BMJ 2020;368:l6927.

40. van der Sommen F, de Groof J, Struyvenberg M, et al. Machine learning in GI endoscopy: practical guidance in how to interpret a novel field. Gut 2020;69:2035-45.

41. Murakami D, Yamato M, Amano Y, et al. Challenging detection of hard-to-find gastric cancers with artificial intelligence-assisted endoscopy. Gut 2020;gutjnl-2020-322453.

42. Menon S, Trudgill N. How commonly is upper gastrointestinal cancer missed at endoscopy? A meta-analysis. Endosc Int Open 2014;2:E46-50.

43. Lee HL, Eun CS, Lee OY, et al. When do we miss synchronous gastric neoplasms with endoscopy? Gastrointest Endosc 2010;71:1159-65.

44. Gotoda T. Endoscopic resection of early gastric cancer. Gastric Cancer 2007;10:1-11.

45. Probst A, Schneider A, Schaller T, et al. Endoscopic submucosal dissection for early gastric cancer: are expanded resection criteria safe for Western patients?. Endoscopy 2017;49:855-65.

46. Abe S, Oda I, Minagawa T, et al. Metachronous Gastric Cancer Following Curative Endoscopic Resection of Early Gastric Cancer. Clin Endosc 2018;51:253-9.

47. Zhu Y, Wang QC, Xu MD, et al. 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.

48. Odagiri H, Yasunaga H. Complications following endoscopic submucosal dissection for gastric, esophageal, and colorectal cancer: a review of studies based on nationwide large-scale databases. Ann Transl Med 2017;5:189.

49. Namikawa K, Hirasawa T, Yoshio T, et al. Utilizing artificial intelligence in endoscopy: a clinician's guide. Expert Rev Gastroenterol Hepatol 2020;14:689-706.

50. Joo M, Park A, Kim K, et al. A Deep Learning Model for Cell Growth Inhibition IC50 Prediction and Its Application for Gastric Cancer Patients. Int J Mol Sci 2019;20:6276.

51. Jin P, Ji X, Kang W, et al. Artificial intelligence in gastric cancer: a systematic review. J Cancer Res Clin Oncol 2020;146:2339-2350.

52. Andras I, Mazzone E, van Leeuwen FWB, et al. Artificial intelligence and robotics: a combination that is changing the operating room. World J Urol 2020;38:2359-66.

53. Fard MJ, Ameri S, Darin Ellis R, et al. Automated robot-assisted surgical skill evaluation: Predictive analytics approach. Int J Med Robot 2018;14:10.1002/rcs.1850.

54. Hashimoto DA, Rosman G, Rus D, et al. Artificial Intelligence in Surgery: Promises and Perils. Ann Surg 2018;268:70-6.

55. O'Sullivan S, Nevejans N, Allen C, et al. Legal, regulatory, and ethical frameworks for development of standards in artificial intelligence (AI) and autonomous robotic surgery. Int J Med Robot 2019;15:e1968.

56. Korhani Kangi A, Bahrampour A. Predicting the Survival of Gastric Cancer Patients Using Artificial and Bayesian Neural Networks. Asian Pac J Cancer Prev 2018;19:487-90.

57. Biglarian A, Hajizadeh E, Kazemnejad A, et al. Application of artificial neural network in predicting the survival rate of gastric cancer patients. Iran J Public Health 2011;40:80-6.

58. Amiri Z, Mohammad K, Mahmoudi M, et al. Assessing the effect of quantitative and qualitative predictors on gastric cancer individuals survival using hierarchical artificial neural network models. Iran Red Crescent Med J 2013;15:42-8.

59. Nilsaz-Dezfouli H, Abu-Bakar MR, Arasan J, et al. Improving Gastric Cancer Outcome Prediction Using Single Time-Point Artificial Neural Network Models. Cancer Inform 2017;16:1176935116686062.

60. Oh SE, Seo SW, Choi MG, et al. Prediction of Overall Survival and Novel Classification of Patients with Gastric Cancer Using the Survival Recurrent Network. Ann Surg Oncol 2018;25:1153-9.

61. Que SJ, Chen QY, Zhong Q, et al. Application of preoperative artificial neural network based on blood biomarkers and clinicopathological parameters for

predicting long-term survival of patients with gastric cancer. World J Gastroenterol 2019;25:6451-64.

62. Yazdani Charati J, Janbabaei G, et al. Survival prediction of gastric cancer patients by Artificial Neural Network model. Gastroenterol Hepatol Bed Bench 2018;11:110-7.

63. Challen R, Denny J, Pitt M, et al. Artificial intelligence, bias and clinical safety. BMJ Qual Saf 2019;28:231-7.

64. Macrae C. Governing the safety of artificial intelligence in healthcare. BMJ Qual Saf 2019;28:495-8.

65. Hamid S. The opportunities and risks of artificial intelligence in medicine and healthcare. 2016.

66. Ahuja AS. The impact of artificial intelligence in medicine on the future role of the physician. Peer J 2019;7:e7702.

67. Li Z, Keel S, Liu C, et al. Can Artificial Intelligence Make Screening Faster, More Accurate, and More Accessible?. Asia Pac J Ophthalmol (Phila) 2018;7:436-41.

68. Murphy, Kevin P. Machine Learning: A Probabilistic Perspective. The MIT Press, 2012.

69. Hand, David & Mannila, Heikki & Smyth, Padhraic. Principles of Data Mining. 2001.

70. Number of Magnetic Resonance Imaging (MRI) Units and Computed Tomography (CT) Scanners: Selected Countries, Selected Years 1990-2009. Center for Disease Control - Health, United States, 2011,

71. Idowu BM, Okedere TA. Diagnostic Radiology in Nigeria: A Country Report. The Journal of Global Radiology, June 2020.

72. Bianco, S., Cusano, C., Napoletano, P., et al. Improving cnn-based texture classification by color balancing. Journal of Imaging 2017;3:33.

73. Fu Y, Mazur TR, Wu X, et al. A novel MRI segmentation method using CNN-based correction network for MRI-guided adaptive radiotherapy. Med Phys 2018;45:5129-37.

74. Martín Noguerol T, Paulano-Godino F, Martín-Valdivia MT, et al. Strengths, Weaknesses, Opportunities, and Threats Analysis of Artificial Intelligence and Machine Learning Applications in Radiology. J Am Coll Radiol 2019;16:1239-47.

75. Tulabandhula T, Rudin C. Machine learning with operational costs. The Journal of Machine Learning Research 2013;14:1989-2028.