Scope of Artificial Intelligence in Gastrointestinal Oncology

Hermant Goyal MD
Syed A.A. Sheraz
Rupinder Mann
Zainab Gandhi
Abhilash Perisetti MD

*Parkview Health*, abhilash.perisetti@gmail.com

See next page for additional authors

Follow this and additional works at: https://researchrepository.parkviewhealth.org/oncol

Part of the *Gastroenterology Commons*, and the *Oncology Commons*

**Recommended Citation**
Goyal, Hermant MD; Sheraz, Syed A.A.; Mann, Rupinder; Gandhi, Zainab; Perisetti, Abhilash MD; Aziz, Muhammad; Chandan, Saurabh; Kopel, Jonathan; Tharian, Benjamin MD; Sharma, Neil MD; and Thosani, Nirav, "Scope of Artificial Intelligence in Gastrointestinal Oncology" (2021). *PCI Publications and Projects*. 61.
https://researchrepository.parkviewhealth.org/oncol/61

This Article is brought to you for free and open access by the Parkview Cancer Institute at Parkview Health Research Repository. It has been accepted for inclusion in PCI Publications and Projects by an authorized administrator of Parkview Health Research Repository. For more information, please contact julie.hughbanks@parkview.com.
Authors
Hermant Goyal MD, Syed A.A. Sheraz, Rupinder Mann, Zainab Gandhi, Abhilash Perisetti MD, Muhammad Aziz, Saurabh Chandan, Jonathan Kopel, Benjamin Tharian MD, Neil Sharma MD, and Nirav Thosani

This article is available at Parkview Health Research Repository: https://researchrepository.parkviewhealth.org/oncol/
Review

Scope of Artificial Intelligence in Gastrointestinal Oncology

Hemant Goyal 1,*, Syed A. A. Sherazi 2, Rupinder Mann 3, Zainab Gandhi 4, Abhilash Perisetti 5, Muhammad Aziz 6, Saurabh Chandan 7, Jonathan Kopel 8, Benjamin Tharian 9, Neil Sharma 5 and Nirav Thosani 10

1 Department of Internal Medicine, The Wright Center for Graduate Medical Education, 501 S. Washington Avenue, Scranton, PA 18505, USA
2 Department of Medicine, John H Stroger Jr Hospital of Cook County, 1950 W Polk St, Chicago, IL 60612, USA; syedaliamir.sherazi@cookcountyhhs.org
3 Department of Medicine, Saint Agnes Medical Center, 1303 E. Herndon Ave, Fresno, CA 93720, USA; rupindrmann@yahoo.com
4 Department of Medicine, Geisinger Wyoming Valley Medical Center, 1000 E Mountain Dr, Wilkes-Barre, PA 18711, USA; drzainabgandhi@gmail.com
5 Division of Interventional Oncology & Surgical Endoscopy (IOSE), Parkview Cancer Institute, 11050 Parkview Circle, Fort Wayne, IN 46845, USA; abhilash.perisetti@gmail.com (A.P.); neil.sharma@parkview.com (N.S.)
6 Department of Gastroenterology and Hepatology, University of Toledo Medical Center, 3000 Arlington Avenue, Toledo, OH 43614, USA; marajani@hotmail.com
7 Division of Gastroenterology and Hepatology, CHI Health Creighton University Medical Center, 7500 Mercy Rd, Omaha, NE 68124, USA; saurabhchandan@gmail.com
8 Department of Medicine, Texas Tech University Health Sciences Center, 3601 4th St, Lubbock, TX 79430, USA; jonathan.kopel@ttuhsc.edu
9 Department of Gastroenterology and Hepatology, The University of Arkansas for Medical Sciences, 4301 W Markham St, Little Rock, AR 72205, USA; btharian@uams.edu
10 Division of Gastroenterology, Hepatology & Nutrition, McGovern Medical School, UTHealth, 6410 Fannin, St #1014, Houston, TX 77030, USA; nirav.thosani@uth.tmc.edu
* Correspondence: doc.hemant@yahoo.com

Simple Summary: Gastrointestinal cancers cause over 2.8 million deaths annually worldwide. Currently, the diagnosis of various gastrointestinal cancer mainly relies on manual interpretation of radiographic images by radiologists and various endoscopic images by endoscopists. Artificial intelligence (AI) may be useful in screening, diagnosing, and treating various cancers by accurately analyzing diagnostic clinical images, identifying therapeutic targets, and processing large datasets. The use of AI in endoscopic procedures is a significant breakthrough in modern medicine. Although the diagnostic accuracy of AI systems has markedly increased, it still needs collaboration with physicians. In the near future, AI-assisted systems will become a vital tool for the management of these cancer patients.

Abstract: Gastrointestinal cancers are among the leading causes of death worldwide, with over 2.8 million deaths annually. Over the last few decades, advancements in artificial intelligence technologies have led to their application in medicine. The use of artificial intelligence in endoscopic procedures is a significant breakthrough in modern medicine. Currently, the diagnosis of various gastrointestinal cancer relies on the manual interpretation of radiographic images by radiologists and various endoscopic images by endoscopists. This can lead to diagnostic variabilities as it requires concentration and clinical experience in the field. Artificial intelligence using machine or deep learning algorithms can provide automatic and accurate image analysis and thus assist in diagnosis. In the field of gastroenterology, the application of artificial intelligence can be vast from diagnosis, predicting tumor histology, polyp characterization, metastatic potential, prognosis, and treatment response. It can also provide accurate prediction models to determine the need for intervention with computer-aided diagnosis. The number of research studies on artificial intelligence in gastrointestinal cancer has been increasing rapidly over the last decade due to immense interest in the field. This review aims to review the impact, limitations, and future potentials of artificial intelligence in screening, diagnosis, tumor staging, treatment modalities, and prediction models for the prognosis of various gastrointestinal cancers.
Keywords: artificial intelligence; colorectal cancer; gastrointestinal cancer; hepatocellular cancer; pancreatico-biliary cancer; gastric cancer; esophageal cancer

1. Introduction

Artificial intelligence is described as the intelligence of machines compared to natural human intelligence. It is a computer science field dedicated to building a machine that simulates the cognitive functions of humans, such as learning and problem solving [1,2]. Recent advances in artificial intelligence technologies have been developed due to technical advances in deep learning technologies, support vector machines, and machine learning, and these advanced technologies have played a significant role in the medical field [3–5]. Virtual and physical are the two main branches of AI in the medical field. Machine learning (ML) and deep learning (DL) are two branches of the virtual branch of AI. Convolutional neural networks (CNN), an essential deep neural network, represent a multilayer artificial neural network (ANN) useful for image analyses. The physical branch of AI includes medical devices and robots [6,7].

As per the WHO report, nearly 5 million new gastrointestinal, pancreatic, and hepatobiliary cancers were recorded worldwide in 2020. Gastrointestinal cancers include esophageal, colorectal (colon and rectum), and gastric cancer. Colorectal cancer (CRC) is the most common cancer of all gastrointestinal cancers. Overall, CRC is the second in terms of mortality and third in incidence after breast and lung cancers globally [8]. Although there are significant advances in diagnostics, including predictive and prognostic biomarkers and treatment approaches for gastrointestinal, pancreatic, and hepatobiliary cancers, there is still high potential to improve further for better clinical outcomes and fewer side effects [6,9]. The data from advanced imaging modalities (including advanced endoscopic techniques with the addition of AI) with high accuracy, novel biomarkers, circulating tumor DNA, and micro-RNA can be beyond human interpretation. In the clinical setting, various diagnostic methods (endoscopy, radiologic imaging, and pathologic techniques) using AI, including imaging analysis, are needed [10–15]. In this narrative review, we discussed the application of AI in diagnostic and therapeutic modalities for various gastrointestinal, pancreatic, and hepatobiliary cancers.

2. Esophageal Cancers

Esophageal cancer, comprising esophageal adenocarcinoma and esophageal squamous cell carcinoma (ESCC), is the ninth most common cancer globally by incidence and sixth by cancer mortality, with an estimated over 600,000 new cases and half a million deaths in 2020 [8]. Even though the incidence of esophageal adenocarcinoma is increasing, ESCC remains the most common histological type of cancer worldwide, with higher prevalence in East Asia and Japan [16,17]. The early diagnosis of ESCC has cure rates > 90%; however, early diagnosis remains a challenge that can be missed even on endoscopic examination [18]. Various diagnostic techniques, such as chromoendoscopy with iodine staining and narrow-band imaging (NBI), are helpful in detecting esophageal cancer at its early stages. While a diagnosis with only white light can be challenging, iodine staining can improve the sensitivity and specificity but can cause mucosal irritation, leading to retrosternal pain and discomfort to the patient [19–22]. NBI is another promising screening method for early esophageal cancer diagnosis [19,20]. Artificial intelligence can improve the sensitivity and specificity of diagnosis of esophageal cancer by improving the endoscopic and image diagnosis. Various retrospective and prospective studies have been conducted to study the role of different AI techniques in improving the diagnosis of esophageal cancer.

Retrospective studies included non-magnifying and magnifying images and real-time endoscopic videos of normal or early esophageal lesions to measure the diagnostic performance of AI models (23–24). One of the earliest retrospective studies conducted by Liu et al. with white light images used joint diagonalization principal component analysis...
(JDPCA), in which there are no approximation, iteration, or inverting procedures. Thus, JDPCA has low computational complexity and is suitable for the dimension reduction of gastrointestinal endoscopic images. A total of 400 conventional gastroscopy esophagus images were used from 131 patients, which showed an accuracy of 90.75%, with an area under the curve (AUC) of 0.9471 in detecting early esophageal cancer [23]. Another retrospective analysis was performed on ex-vivo volumetric laser endomicroscopy (VLE) images to analyze and compare various computer algorithms. Three novel clinically inspired algorithm features (“layering,” “signal intensity distribution,” and “layering and signal decay statistics”) were developed. When comparing the performance of these three clinical features and generic image analysis methods, “layering and signal decay statistics” showed better performance with sensitivity and specificity of 90% and 93%, respectively, along with an AUC of 0.81 compared to other methods tested [24].

Further retrospective analyses have been performed to analyze the role of different AI models, including supervised vector machines, convolutional neural networks on various diagnostic methods such as white light endoscopy, NBI, and real-time endoscopy videos. Cai et al. developed a computer-aided detection (CAD) using a deep neural network system (DNN) to detect early ESCC using 2428 esophageal conventional endoscopic white light images. DNN-CAD had sensitivity, specificity, and accuracy of 97.8%, 85.4%, and 91.45%, respectively, with a ROC of >96% when tested on 187 images from the validation dataset. Most importantly, the diagnostic ability of endoscopists improved significantly in sensitivity (74.2% vs. 89.2%), accuracy (81.7% vs. 91.1%), and negative predictive value (79.3% vs. 90.4%) after referring to the performance of DNN-CAD [25]. In another retrospective analysis performed by Horie et al., deep learning through convolutional neural networks were developed using 8428 training images of esophageal cancer, including conventional white light images and NBI. When tested on a set of 1118 test images, a CNN analyzed images in 27 s to accurately diagnose esophageal cancer with a sensitivity of 98%. It also showed an accuracy of 98% in differentiating superficial esophageal cancer from late-stage esophageal cancer, which can improve the prognosis of the patients and decrease the morbidity of more invasive procedures [26].

The definitive treatment of ESCC varies from endoscopic resection to surgery or chemoradiation depending on the level of invasion depth, so it is very important to determine it. In a study conducted in Japan, 1751 retrospectively collected training images of ESCC were used to develop an AI-diagnostic system of CNN using deep learning technique to detect the depth of invasion of ESCC. The AI-diagnostic system identified ESCC correctly in 95.5% of test images and estimated the invasion depth with a sensitivity of 84.1% and accuracy of 80.9% in about 6 s, which is higher than endoscopists [27]. Intrapapillary capillary loops (IPCLs) are microvessels seen visualized using magnification endoscopy. IPCLs are an endoscopic feature of early esophageal squamous cell neoplasia, and changes in their morphology are correlated with invasion depth. In this study, 7046 high-definition magnification endoscopies with NBI were used to train a CNN. As a result, CNN was able to identify abnormal from normal IPCLs patterns with an accuracy, sensitivity, and specificity of 93.7%, 89.3%, and 98%, respectively [28]. Based on various retrospective analyses, it was established that the use of AI in diagnosing esophageal cancer would prove beneficial.

Many prospective analyses have also been performed to further assess the application of AI in the diagnosis of esophageal cancer. Struyvenberg et al. conducted a prospective study to detect Barrett’s neoplasia by CAD using a multi-frame approach. A total of 3060 VLE images were analyzed using a multi-frame analysis. Multi-frame analysis achieved a much higher AUC (median level = 0.91) than a single frame one (median level = 0.83). CAD was able to analyze multi-frame images in 3.9 s, which, traditionally, is a time-consuming and complex procedure due to the subtle gray shaded VLE images [29]. Thus, on the prospective study, as well, CAD proved beneficial for analyzing VLE images. Similarly, in another prospective study, a hybrid ResNet-UNet model CAD system was developed using five different independent endoscopic datasets to improve the identification of early
neoplasm in patients with BE. When comparing the CAD system with general endoscopists, the study found higher sensitivity (93% vs. 72%), specificity (83% vs. 74%), and accuracy (88% vs. 73%) with the CAD system in the classification of images as containing neoplasm or non-dysplastic BE on dataset 5 (second external validation). CAD was also able to identify the optimal site to collect biopsy with higher accuracy of 97% and 92% of cases in datasets 4 and 5, respectively (dataset 4, 5 external validation sets, datasets 1, 2, and 3 were pre-training, training, and internal validation, respectively) [30].

With the promising results from retrospective and prospective studies conducted on endoscopic images, studies were designed to evaluate the role of AI for in vivo analysis to aid the diagnosis of Barrett’s during endoscopies. A prospective study developed and tested a CAD system to detect Barrett’s neoplasm during live endoscopic procedures. The CAD system predicted 25 of 33 neoplastic images and 96 of 111 non-dysplastic BE images correctly and thus had an image-based accuracy, sensitivity, and specificity of 84%, 76%, and 86%, respectively. Additionally, the CAD system predicted 9 of 10 neoplastic patients correctly, resulting in a sensitivity of 90%. So, this study showed high sensitivity to predict neoplastic lesions with the CAD system. However, it is in vivo single center, so further large multicenter trials are needed [31].

Shiroma et al. conducted a study to examine AI’s ability to detect superficial ESCC using esophagogastroduodenoscopy (EGD) videos. A CNN through deep learning was developed using 8428 EGD images of esophageal cancer. The AI system performance was evaluated using two validation sets of a total of 144 videos. The AI system correctly diagnosed 100% and 85% ESCC in the first and second validation sets, respectively. Whereas endoscopists detected only 45% of ESCC and their sensitivities improved significantly with AI real-time assistance compared to those without AI assistance ($p < 0.05$) [32]. In a retrospective study, a deep learning-based AI system was developed to detect early ESCC. Magnifying and non-magnifying endoscopy images of non-dysplastic, early ESCC, and advanced esophageal cancer lesions were used to train and validate the AI system. For non-magnifying images, AI diagnosis had a per-patient accuracy, sensitivity, specificity of 99.5%, 100%, and 99.5%, respectively, for white light imaging, and for magnified images, the per-patient accuracy, sensitivity, and specificity were 88.1%, 90.9%, and 85.0%, respectively. The accuracy of AI diagnosis was similar to experienced endoscopists; however, it was better than trainees [33].

Systematic reviews and meta-analyses of 21 and 19 studies, respectively, were conducted to test the CAD algorithm’s diagnostic accuracy to detect esophageal cancer using endoscopic images. It showed that the pooled AUC, sensitivity, specificity, and diagnostic odds ratio of CAD algorithms for the diagnosis of esophageal cancer for image-based analysis were 0.97 (95% CI: 0.95–0.99), 0.94 (95% CI: 0.89–0.96), 0.88 (95% CI: 0.76–0.94), and 108 (95% CI: 43–273), respectively. The pooled AUC, sensitivity, specificity, and diagnostic odds ratio of CAD algorithms for the diagnosis of esophageal cancer depth of invasion were 0.96 (95% CI: 0.86–0.99), 0.90 (95% CI: 0.88–0.92), 0.88 (95% CI: 0.83–0.91), and 138 (95% CI: 12–1569), respectively. There was no heterogeneity or publication bias on meta-regression [34]. Table 1 summarizes recent key studies assessing the role of AI in the diagnosis of esophageal cancer and pre-cancerous lesions using imaging [32,35–46].

The available literature provides strong evidence that the utilization of CAD in esophageal cancer can prove beneficial for early diagnosis, which remains crucial to prevent the significant morbidity and mortality of the patients [18]. While limitations such as external validation, clinical application, and the need for randomized control trials remain, the evidence so far supports the use of AI and therefore necessitates the need for larger controlled trials.

<table>
<thead>
<tr>
<th>Author, Year, Reference</th>
<th>Dataset (Images Count and Lesions Type)</th>
<th>AI System</th>
<th>Modality</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shin 2015 [35]</td>
<td>375 images (esophageal squamous cell cancer)</td>
<td>Two-class linear discriminant analysis</td>
<td>HRME</td>
<td>Sensitivity 84%, specificity 95%, and AUC 0.95</td>
</tr>
<tr>
<td>Quang 2016 [36]</td>
<td>375 images (esophageal squamous cell cancer)</td>
<td>Fully automated real-time analysis algorithm</td>
<td>HRME</td>
<td>Sensitivity 95%, specificity 91%, and AUC 0.937</td>
</tr>
<tr>
<td>Van der Sommen 2016 [37]</td>
<td>100 images (60 early BE neoplasia and 40 BE)</td>
<td>SVM</td>
<td>WLI</td>
<td>Per image-sensitivity 83% and specificity 83%. For per patient sensitivity 86% and specificity 87%.</td>
</tr>
<tr>
<td>Swager 2017 [24]</td>
<td>60 images (30 early BE neoplasia and 30 BE)</td>
<td>SVM</td>
<td>VLE</td>
<td>Sensitivity 90%, specificity 93%, and AUC 0.95</td>
</tr>
<tr>
<td>Mendel 2017 [38]</td>
<td>100 (50 BE and 50 esophageal adenocarcinoma)</td>
<td>CNN</td>
<td>WLI</td>
<td>Sensitivity 94% and specificity 88%</td>
</tr>
<tr>
<td>Cai 2019 [25]</td>
<td>2615 images (early esophageal squamous cell cancer)</td>
<td>DNN-CAD</td>
<td>WLI</td>
<td>Sensitivity 97.8%, specificity 85.4%, and accuracy 91.4%</td>
</tr>
<tr>
<td>Horie 2019 [26]</td>
<td>9546 images (esophageal cancer)</td>
<td>CNN-SSD (single shot multbox detector)</td>
<td>WLI and NBI</td>
<td>Per image Sensitivity 72% (WLI) and 86% (NBI). Per case Sensitivity 79% (WLI) and 89% (NBI).</td>
</tr>
<tr>
<td>Ebigbo 2019 [39]</td>
<td>248 images. Two databases. [(a)-Augsburg dataset-148. (b) MICCAI dataset-100]</td>
<td>CNN-ResNet (residual net)</td>
<td>WLI and NBI</td>
<td>Augsburg database—WLE sensitivity 97% and specificity 88% and NBI sensitivity 94% and specificity 80%. MICCAI database—sensitivity 92% and specificity 100%</td>
</tr>
<tr>
<td>Everson 2019 [28]</td>
<td>7046 images (Intrapapillary capillary loop patterns in early esophageal squamous cell cancer)</td>
<td>CNN</td>
<td>Magnified NBI</td>
<td>Sensitivity 89.3%, specificity 98%, and accuracy 93.7%.</td>
</tr>
</tbody>
</table>
Table 1. Cont.

<table>
<thead>
<tr>
<th>Author, Year, Reference</th>
<th>Dataset (Images Count and Lesions Type)</th>
<th>AI System</th>
<th>Modality</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zhao 2019 [40]</td>
<td>1350 images (early esophageal squamous cell cancer)</td>
<td>Double labeling fully convolutional network (FCN)</td>
<td>Magnifying endoscopy with NBI</td>
<td>Diagnostic accuracy at the lesion level 89.2% and at the pixel level 93.0%.</td>
</tr>
<tr>
<td>Nakagawa 2019 [41]</td>
<td>15252 images (early esophageal squamous cell cancer)</td>
<td>CNN</td>
<td>Magnified and non-magnified WLI, NBI, and BLI</td>
<td>Sensitivity 90.1%, specificity 95.8%, and accuracy 91%</td>
</tr>
<tr>
<td>Guo 2019 [42]</td>
<td>6473 images and 47 videos (early esophageal squamous cell cancer)</td>
<td>CNN-SegNet</td>
<td>Non-magnified and magnified NBI</td>
<td>Per image sensitivity 98.04%, specificity 95.03%, and AUC 0.989 Per frame sensitivity 91.5% and specificity 99.9%</td>
</tr>
<tr>
<td>Hashimoto 2020 [43]</td>
<td>1832 images (916 early BE neoplasia and 916 BE)</td>
<td>CNN</td>
<td>WLI and NBI</td>
<td>WLI sensitivity 98.6% and specificity 88.8%, NBI sensitivity 92.4% and specificity 99.2%</td>
</tr>
<tr>
<td>Ohmori 2020 [44]</td>
<td>23289 (superficial early esophageal squamous cell cancer)</td>
<td>CNN</td>
<td>Non-magnified WLI, NBI, and BLI. Magnified NBI and BLI</td>
<td>Non-magnified NBI/BLI—sensitivity 100%, specificity 63%, and accuracy 77%. Non-magnified WLI—sensitivity 90%, specificity 76%, and accuracy 81%. Magnified NBI—sensitivity 98%, specificity 56%, and accuracy 77%.</td>
</tr>
<tr>
<td>Tokai 2020 [27]</td>
<td>1751 (superficial early esophageal squamous cell cancer)</td>
<td>CNN</td>
<td>WLI and NBI</td>
<td>Sensitivity 84.1%, specificity 73.3%, and accuracy 80.9%</td>
</tr>
<tr>
<td>Li 2021 [45]</td>
<td>2167 images (early esophageal squamous cell cancer)</td>
<td>CAD</td>
<td>WLI and NBI</td>
<td>CAD-NBI sensitivity 91%, specificity 96.7%, and accuracy 94.3%. CAD-WLI sensitivity 98.5%, specificity 83.1%, and accuracy 89.5%</td>
</tr>
<tr>
<td>Shiroma 2021 [32]</td>
<td>8428 and 80 videos (T1 esophageal squamous cell cancer)</td>
<td>CNN</td>
<td>WLI and NBI</td>
<td>WLI sensitivity 75%, specificity 30% NBI sensitivity 55%, specificity 80%</td>
</tr>
<tr>
<td>Ebigbo 2021 [46]</td>
<td>230 WLI images (108 T1a, 122 T1b stage)</td>
<td>ANN</td>
<td>WLI</td>
<td>Sensitivity 77%, Specificity 64%, Diagnostic accuracy 71% to differentiate T1a from T1b lesions. Not significantly different from clinical experts</td>
</tr>
</tbody>
</table>
3. Gastric Cancers

Gastric cancer is the sixth most common cancer worldwide by incidence with over 1 million new cases and the third leading cause of cancer-related death in 2020 [8]. The diagnosis of gastric cancer has transitioned from histology-only samples to precise molecular analysis of cancer. With the advent and use of endoscopy in diagnosing gastric cancer, the medical field is interested in earlier diagnosis in a non-invasive manner. The level of expertise required for endoscopic diagnosis of early gastric cancer remains high, and artificial intelligence can help with more accurate diagnosis and higher efficiency in image interpretation [47,48]. Here, we discuss the role of AI in diagnosing H. Pylori infection, a precursor of gastric cancer, and various methods of diagnosing gastric cancer with the help of machine learning methods.

3.1. Use of AI in Helicobacter Pylori Detection

Chronic, untreated H. Pylori infection is strongly associated with chronic gastritis, ulceration, mucosal atrophy, intestinal metaplasia, and gastric cancer [49]. On endoscopy, H. pylori infection is diagnosed by redness and swelling, which artificial intelligence can optimize. Various retrospective studies have been conducted to compare and develop a higher efficiency model of H. pylori diagnosis. Huang et al. pioneered the application of refined feature selection with neural network (RFSNN) developed using endoscopic images with histological features from 30 patients. It was then tested on 74 patients to predict H. pylori infection and related histological features. It showed sensitivity, specificity, and accuracy of 85.4%, 90.9%, and more than 80%, respectively [50]. A retrospective study developed a two-layered CNN model in which, first, CNN identified positive or negative H. pylori infection, and second, CNN classified the images according to the anatomical locations. The sensitivity, specificity, accuracy, and diagnostic time were 81.9%, 83.4%, 83.1%, and 198 s, respectively, for the first CNN and 88.9%, 87.4%, 87.7%, and 194 s, respectively, for the secondary CNN. Compared to the board-certified endoscopists, they were 85.2%, 89.3%, 88.9%, and 252.5 ± 92.3 min, respectively. The accuracy of the secondary CNN was significantly higher than all endoscopists, including relatively experienced and board-certified endoscopists (5.3%; 95% CI: 0.3–10.2), although they had comparable sensitivity and specificity [51]. Zheng et al. developed and used ResNet-50 based on endoscopic gastric images to diagnose H. pylori infection, which was confirmed with immunohistochemistry tests on biopsy samples or urea breath tests. The sensitivity, specificity, accuracy, and AUC were 81.4%, 90.1%, 84.5%, and 0.93, respectively for a single gastric image and 91.6%, 98.6%, 93.8%, and 0.97, respectively, for multiple gastric images [52]. These studies showed the high accuracy of CNN in diagnosing H. Pylori infection based on endoscopic imaging, and it was found to be comparable to the expert endoscopist.

A single-center prospective study compared the accuracy of the AI system with endoscopy images taken with white light imaging (WLI), blue laser imaging (BLI), and linked color imaging (LCL) in 105 H. pylori-positive patients. The AUC for WLI, BLI, and LCL were 0.66, 0.96, and 0.95, respectively ($p < 0.01$). Thus, this study showed a higher accuracy of H. pylori infection diagnosis with BLI and LCL with AI systems than WLI [53]. A systematic review and meta-analysis of eight studies were performed to evaluate AI accuracy in diagnosing H. pylori infection using endoscopic images. The pooled sensitivity, specificity, AUC, and diagnostic odds ratio to predict H. pylori infection were 87%, 86%, 0.92, and 40 (95% CI 15–112), respectively. The AI had a 40 times higher probability of predicting H. pylori infection than standard methods [54].

AI systems can be considered a valuable tool in the endoscopic diagnosis of H. Pylori infection based on available data from various studies. Although most of these studies lack external validation, promising results have been observed so far.

3.2. Use of AI in Gastric Cancer

Early diagnosis of gastric cancer remains prudent to provide less invasive and more successful treatments such as endoscopic submucosal dissection, which can be offered to
patients with only intramucosal involvement [55]. AI can help by using endoscopy images for early diagnosis and thus better survival. A single-center observational study was conducted to test the efficacy of CAD for diagnosing early gastric cancer using magnifying endoscopy with narrow-band imaging. The CAD system was first pre-trained using cancerous and noncancerous images and then tested on 174 cancerous and noncancerous videos. The results showed a sensitivity, specificity, accuracy, PPV, NPV, and AUC of 87.4%, 82.8%, 85.1%, 83.5%, 86.7%, and 0.8684, respectively, for the CAD system. When comparing CAD against 11 expert endoscopists, the diagnostic performance of CAD was comparable to most expert endoscopists. Given the high sensitivity of CAD in diagnosing early gastric cancer, it can be helpful for endoscopists who are less experienced or lack endoscopic skills of ME-NBI. It can also be useful for experts with low diagnostic performance as diagnostic performance varies among experts [56]. Various CNN models have been developed to determine gastric cancer invasion depth, which can be used as a screening tool to determine the patient qualification for submucosal dissection. In another study, AI-based convolutional neural network computer-aided detection (CNN-CAD) was developed based on endoscopic images and then used to determine the invasion depth of gastric cancer. The AUC, sensitivity, specificity, and accuracy were 0.94, 76.47%, 95.56%, and 89.16%, respectively, for the CNN-CAD system.

Moreover, the CNN-CAD had an accuracy of 17.25% and a specificity of 32.21% higher than endoscopists [57]. Joo Cho et al. studied the application of a deep learning algorithm to determine the submucosal invasion of gastric cancer in endoscopic images. The mean AUC to discriminate submucosal invasion was 0.887 with external testing. Thus, deep learning algorithms may have a role in improving the prediction of submucosal invasion [58].

Ali et al. studied the application of AI on chromoendoscopy images to detect gastric abnormalities using endoscopic images. Chromoendoscopy is an advanced image-enhanced endoscopy technique that uses spraying dyes such as methylene blue to enhance gastric mucosa. This study uses a newer feature extraction method called Gabor-based gray-level co-occurrence matrix (G2LCM) for the computer-aided detection of chromoendoscopy abnormal frames. It is a hybrid approach of local and global texture descriptions. The G2LCM texture features and the support vector machine classifier were able to classify abnormal from normal frames with a sensitivity of 91%, a specificity of 87%, and an AUC of 0.91 [59]. In another study, CADx was trained with magnifying NBI and further with G2LCM-determined images from the cancerous blocks and compared to expert-identified areas. The CAD showed an accuracy of 96.3%, specificity of 95%, PPV of 98.3%, and sensitivity of 96.7%. This study showed that this CAD system could help diagnose early gastric cancer [60].

A systematic review and meta-analysis of 16 studies were performed to understand AI efficacy in endoscopic diagnosis of early gastric cancer. The use of AI in the endoscopic detection of early gastric cancer achieved an AUC of 0.96 (95% CI: 0.94–0.97), pooled sensitivity of 86% (95% CI: 77–92%), and a pooled specificity of 93% (95% CI: 89–96%). For AI-assisted depth distinction, the AUC, pooled sensitivity, and specificity were 0.82 (95% CI: 0.78–0.85), 72% (95% CI: 58–82%), and 79% (95% CI: 56–92%), respectively [61].

Most of the available literature currently focuses on AI applications in diagnosing gastric cancer rather than on the treatment response and prediction. Joo et al. constructed and studied the application of a one-dimensional convolution neural network model (DeepIC50), which showed accuracy in pan-cancer cell line prediction. This was applied to approved treatments of trastuzumab and ramucirumab, which showed promising predictions for drug responsiveness, which can be helpful in the development of newer medication [62].

While many studies have been conducted independently, there is a need for larger prospective trials studying the application of AI in the entirety of gastric cancer diagnosis and treatment to better assess its efficacy and application in clinical practice. Table 2 summarizes key studies assessing the role of AI in the diagnosis of Gastric cancer by imaging [23,63–70].

<table>
<thead>
<tr>
<th>Author, Year, Reference</th>
<th>Dataset (Images Count and Lesions Type)</th>
<th>AI System</th>
<th>Modality</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Miyaki 2015 [63]</td>
<td>587 cut out images, early gastric cancer</td>
<td>SVM</td>
<td>Magnifying endoscopy-BLI</td>
<td>SVM output 0.846 ± 0.220 for cancerous lesions and 0.219 ± 0.277 for surrounding tissues</td>
</tr>
<tr>
<td>Liu 2016 [23]</td>
<td>400 images, early gastric cancer</td>
<td>JDPCA</td>
<td>WLI</td>
<td>AUC—0.9532, accuracy—90.75%</td>
</tr>
</tbody>
</table>
| Shichijo 2017 [39]      | 32,208 images (CNN 1) and images classified based on 8 anatomic locations (CNN2), Helicobacter pylori infection | Deep CNN  | WLI      | CNN 1—Sensitivity 81.9%, specificity 83.4%, and accuracy 83.1%  
                |                                        |           |          | CNN2—Sensitivity 88.9%, specificity 87.4%, and accuracy 87.7% |
| Ali 2018 [46]           | 176 images, abnormal gastric mucosa including metaplasia and dysplasia | G2LCM     | Chromoendoscopy | Sensitivity 91%, specificity 82%, accuracy 87%, and AUC 0.91 |
| Hirasawa 2018 [64]      | 13584 endoscopic images, gastric cancer | CNNS bases single shot Multibox Detector | WLE, NBI and chromoendoscopy | Sensitivity 92.2% |
| Kanesaka 2018 [47]      | 126 images, early gastric cancer       | GLCM features, SVM | Magnifying endoscopy NBI | Sensitivity 96.7%, specificity 95%, and accuracy 96.3% |
| Liu 2018 [65]           | 1120 Magnifying endoscopy NBI images, early gastric cancer | Deep CNN  | Magnifying endoscopy NBI | Top Sensitivity 96.7%, specificity 95%, and accuracy 98.5% |
| Horiuchi 2019 [66]      | 2828 images (1643 early gastric cancer, 1185 gastritis, early gastric cancer) | CNN       | Magnifying endoscopy NBI | Sensitivity 95.4%, specificity 71%, and accuracy 85.3% |
| Zhu 2019 [45]           | 993 images, Invasive depth of gastric cancer invasion | CNN-CAD system | WLI | Sensitivity 76.47%, specificity 95.56%, accuracy 89.1%, and AUC 0.98 |
| Guimaraes 2020 [67]     | 270 images, gastric precancerous condition such as atrophic gastritis | Deep learning, CNN | WLI | AUC 0.98, sensitivity 93% |
| Yasuda 2020 [68]        | 525 images, H. pylori infection        | SVM       | LCI      | Sensitivity 90.4%, specificity 85.7%, and accuracy 87.6% |
Table 2. Cont.

<table>
<thead>
<tr>
<th>Author, Year, Reference</th>
<th>Dataset (Images Count and Lesions Type)</th>
<th>AI System</th>
<th>Modality</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wu 2021 [69]</td>
<td>1050 patients, early gastric cancer</td>
<td>ENDOANGEL- deep CNN based system</td>
<td>WLI</td>
<td>Per lesion, sensitivity 100%, specificity 84.3%, and accuracy 84.7%</td>
</tr>
<tr>
<td>Xia 2021 [70]</td>
<td>1,023,955 images, gastric lesion</td>
<td>Faster region-based convolutional neural network</td>
<td>Magnetically controlled capsule endoscopy</td>
<td>Sensitivity 96.2%, specificity 76.2%, and accuracy 77.1%</td>
</tr>
</tbody>
</table>
4. Colorectal Cancer

Colorectal cancer (CRC) is the fourth most common cancer diagnosed in the United States [71]. Its incidence has been falling steadily, mostly due to screening using colonoscopy and treatment of polyps that can effectively prevent the development of colon cancer [71,72]. The use of AI is increasingly studied to improve polyp and cancer detection in the colon [73].

4.1. AI and Colon Polyp Detection

A recent systematic review and meta-analysis reported a pooled sensitivity and specificity of 92% and 93%, respectively, with 92% accuracy using still and video images from colonoscopy in over 17,400 patients. When using video frames or images, the pooled sensitivity and specificity were higher (92% and 89%, respectively) compared to studies that used still images alone, reporting a pooled sensitivity and specificity of 84% and 87%, respectively, for AI in the detection of colon polyps. Most of the studies were retrospective in design [74]. The types of AI used ranged from SVM, ANN, and CNN to several modifications of deep learning methods. A meta-analysis of seven randomized controlled trials (RCTs) showed a significant increase in the rate of polyp detection when AI was used with colonoscopy compared to colonoscopy alone, with an odds ratio of 1.75 (95% CI: 1.56–1.96, \( p < 0.001 \)). All studies had a higher polyp detection rate in the AI group than the standard colonoscopy alone group [74]. Recent studies have also shown considerable promise in the use of AI, especially CNN-based systems in colon capsule endoscopy, to improve rates of colon polyp detection [75,76]. Laiz et al. developed a CNN model to detect polyps of all sizes and morphologies using capsule endoscopy images and reported specificity of over 90% for small to large size polyps as well as pedunculated or sessile polyps [76].

4.2. AI and Colon Polyp Characterization

A pooled analysis of 20 studies that used AI to predict the histology of a polyp detected on colonoscopy revealed a sensitivity and specificity of 94% and 87% with 91% accuracy in detecting an adenomatous polyp. SVM was the most common AI in these studies, followed by DNN, CNN, and other DL methods [74]. A meta-analysis of seven RCTs also showed an increased detection rate of adenomas when AI was combined with colonoscopy compared to colonoscopy alone with an OR 1.53 (95% CI: 1.32–1.77, \( p < 0.001 \)). The absolute improvement in adenoma detection ranged from 6% to 15.2% for AI compared to colonoscopy alone in the studies. All RCTs utilized ANN- or CNN-based AI systems using video streams from colonoscopy [74]. Table 3 summarizes select studies assessing role of AI in diagnosis/characterization of polyps and colon cancer by imaging [77–99].

<table>
<thead>
<tr>
<th>Author, Year, and Reference</th>
<th>Dataset</th>
<th>AI System</th>
<th>Modality</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tischendorf 2010 [77]</td>
<td>209 polyps, colorectal polyps</td>
<td>Region growing algorithm</td>
<td>Magnifying NBI</td>
<td>Sensitivity 90% and specificity 70%.</td>
</tr>
<tr>
<td>Takemura 2012 [78]</td>
<td>371 colorectal lesions</td>
<td>SVM</td>
<td>Magnification chromoendoscopy</td>
<td>Sensitivity 97.8%, specificity 97.9%, and accuracy 97.8%</td>
</tr>
<tr>
<td>Mori 2015 [79]</td>
<td>176 colorectal lesions</td>
<td>SVM, EC-CAD</td>
<td>WLI, endocytoscopy</td>
<td>Sensitivity 92.0%, specificity 79.5%, and accuracy 89.2%</td>
</tr>
<tr>
<td>Fernandez 2016 [80]</td>
<td>31 colorectal polyps from 24 videos</td>
<td>WM-DOVA energy maps</td>
<td>WLI</td>
<td>Sensitivity 70.4% and specificity 72.4%</td>
</tr>
<tr>
<td>Kominami 2016 [81]</td>
<td>118 colorectal lesions</td>
<td>Real-time CAD and SVM</td>
<td>Magnifying NBI</td>
<td>Sensitivity 93.0%, specificity 93.3%, and accuracy 93.2%</td>
</tr>
<tr>
<td>Park and Sargent 2016 [82]</td>
<td>11802 images patches</td>
<td>CNN</td>
<td>WLI and NBI</td>
<td>Sensitivity 86% and specificity 85%</td>
</tr>
<tr>
<td>Tamai 2017 [83]</td>
<td>121 colorectal lesions</td>
<td>CAD</td>
<td>Magnifying NBI</td>
<td>Sensitivity 83.9%, specificity 82.6%, and accuracy 82.8%</td>
</tr>
<tr>
<td>Zhang 2017 [84]</td>
<td>215 colorectal polyps</td>
<td>CNN</td>
<td>WLI and NBI</td>
<td>Precision 87.3 and accuracy 85.9%</td>
</tr>
<tr>
<td>Misawa 2018 [85]</td>
<td>155 polyps from 73 colonoscopy videos</td>
<td>CNN</td>
<td>WLI</td>
<td>Per frame: Sensitivity 90%, specificity 63.3%, and accuracy 76.5%</td>
</tr>
<tr>
<td>Mori 2018 [86]</td>
<td>466 diminutive colorectal polyps from 325 patient</td>
<td>CAD, SVM</td>
<td>NBI</td>
<td>Pathologic prediction rate 98.1%</td>
</tr>
<tr>
<td>Urban 2018 [87]</td>
<td>8,641 images from screening colonoscopy containing 4088 colorectal polyp detection</td>
<td>CNN</td>
<td>WLI</td>
<td>Accuracy 96.4%, AUC 0.991</td>
</tr>
<tr>
<td>Bryne 2019 [88]</td>
<td>125 diminutive colorectal polyp videos</td>
<td>Deep CNN</td>
<td>NBI</td>
<td>Sensitivity 98%, specificity 83%, and accuracy 94%</td>
</tr>
<tr>
<td>Figueiredo 2019 [89]</td>
<td>1680 frames with polyps and 1360 frames without polyps from 42 patients</td>
<td>SVM binary classifiers</td>
<td>WLI</td>
<td>Sensitivity 99.7%, specificity 84.9%, and accuracy 91.1%</td>
</tr>
<tr>
<td>Author, Year, and Reference</td>
<td>Dataset</td>
<td>AI System</td>
<td>Modality</td>
<td>Results</td>
</tr>
<tr>
<td>----------------------------</td>
<td>---------</td>
<td>-----------</td>
<td>----------</td>
<td>---------</td>
</tr>
<tr>
<td>Horiuchi 2019 [90]</td>
<td>429 diminutive colorectal polyps (258 rectosigmoid and 171 non-rectosigmoid polyps)</td>
<td>Color intensity analysis software</td>
<td>Autofluorescence imaging</td>
<td>Sensitivity 80.0 %, specificity 95.3%, and accuracy 91.5%</td>
</tr>
<tr>
<td>Ito 2019 [91]</td>
<td>190 images of colon lesions, stage 1b colon cancer</td>
<td>CNN</td>
<td>WLI</td>
<td>Sensitivity 67.5%, specificity 89%, accuracy 81.2%, and AUC 0.871</td>
</tr>
<tr>
<td>Wang 2019 [92]</td>
<td>1058 patients (536 randomized to standard colonoscopy and 522 to colonoscopy with computer aided diagnosis)</td>
<td>CNN</td>
<td>WLI</td>
<td>Adenoma detection rate 29.1% for standard colonoscopy and 20.3% for colonoscopy with computer-aided diagnosis group</td>
</tr>
<tr>
<td>Jin 2020 [93]</td>
<td>300 images of colorectal polyps (180 adenomatous polyps and 120 hyperplastic polyps)</td>
<td>CNN</td>
<td>NBI</td>
<td>Sensitivity 83.3 %, specificity 91.7%, and accuracy 86.7%</td>
</tr>
<tr>
<td>Kudo 2020 [94]</td>
<td>2000 images, colorectal polyps</td>
<td>Endocytoscopy with NBI and methylene blue staining modes</td>
<td>EndoBRAIN</td>
<td>Sensitivity 96.9%, specificity 100%, and accuracy 98%</td>
</tr>
<tr>
<td>Nakajima 2020 [95]</td>
<td>78 images, Colorectal cancer with deep submucosal invasion</td>
<td>CAD</td>
<td>Non magnified WLI</td>
<td>Sensitivity 81%, specificity 87%, and accuracy 84%</td>
</tr>
<tr>
<td>Ozawa 2020 [96]</td>
<td>1172 colorectal polyp images from 309 polyps.</td>
<td>CNN</td>
<td>NBI</td>
<td>Trained CNN detection—Sensitivity 92% and positive predictive value 86%. Colorectal polyp characterization with NBI—sensitivity 97% and positive predictive value 98%.</td>
</tr>
<tr>
<td>Repici 2020 [97]</td>
<td>685 patients who underwent screening colonoscopy</td>
<td>CADe</td>
<td>WLI</td>
<td>Adenoma detection rate for CADe (54.8%) higher than standard colonoscopy (40.4%) with relative risk 1.30, 95% CI: 1.14–1.45</td>
</tr>
<tr>
<td>Lai 2021 [98]</td>
<td>16 patients, colorectal polyps</td>
<td>DNN</td>
<td>WLI and NBI</td>
<td>Sensitivity 100%, specificity 100%, and accuracy 74–95%</td>
</tr>
<tr>
<td>Luo 2021 [99]</td>
<td>150 patients who underwent screening colonoscopy</td>
<td>CNN</td>
<td>WLI</td>
<td>Polyp detection rate for AI-assisted colonoscopy group (38.7%) higher than standard colonoscopy group (34.0%), p &lt;0.001.</td>
</tr>
</tbody>
</table>
4.3. AI and Colorectal Cancer

Lymph nodal metastases in colon cancer confer a lower survival and prognosis than patients with no nodal metastases in early-stage colon cancer, with 5-year survival dropping from over 90% to 72% in the presence of nodal metastases in early-stage colon cancer [71]. Kudo et al. studied an ANN-based system in over 3000 patients with T1 colorectal cancers and found a significantly higher nodal metastasis rate identified (AUC 0.83) compared to standard guidelines alone (AUC 0.73, $p < 0.001$). Analyzing patients who underwent the endoscopic resection of T1 tumors, the LN metastases rate was higher for the AI system compared to guidelines (AUC 0.84 vs. AUC 0.77, $p = 0.005$), signifying a potential role for AI in detecting patients who may need further nodal sampling after the endoscopic resection of early-stage colon cancer [100]. Ichimasa et al. studied AI in 590 resected patients with T1 colorectal cancer and found that AI was 100% sensitive for LN metastases detection with a specificity and accuracy over 20% points higher than US clinical guidelines alone [101].

Overall, AI has shown considerable promise to be an excellent addition to a gastroenterologist’s armamentarium to not only detect small polyps with higher accuracy but also to be able to differentiate adenomas from other histologies and identify patients with early-stage T1 colon cancer who may benefit from or may be able to avoid additional nodal surgery.

5. Pancreatic Cancer

Pancreatic cancer is an aggressive cancer with a dismal 5-year overall survival of <10%. The incidence is rising globally, and it has the highest mortality rate amongst all major cancers [102]. Surgical resection at an early stage remains the only chance of cure for patients, hence signifying the importance of early detection of these malignancies. Artificial intelligence has been studied to augment the current diagnostic armamentarium to help identify pancreatic lesions and differentiate benign from malignant disease [103]. Here, we will discuss the evidence so far about the role of AI in pancreatic cancers.

5.1. AI and Radiologic Diagnosis of Pancreatic Cancer

Liu et al. using a faster region-based CNN model reported an AUC of 0.96 for the recognition of pancreatic cancer based on contrast-enhanced CT images. In contrast, Li et al. reported an accuracy of only 72.8% of differentiating various pancreatic cysts based on CT images using densely connected convolutional networks [104,105]. Using the ML approach, Chu et al. reported a sensitivity of 100%, a specificity of 98.5%, and a 99% accuracy [106]. Wei et al. studied the SVM model and reported a much lower sensitivity and specificity of 66.7 and 81.8%, respectively [107]. Li et al. studied a method combining SVM and random forest technology using PET/CT images and showed excellent sensitivity, specificity, and accuracy of 95%, 97.5%, and 96%, respectively [108]. All these studies had small sample sizes of a few hundred patients.

AI application has been challenging in MR images for the recognition of pancreatic neoplasms. Kaisiss et al., using the ML model, reported an AUC of 0.93 with the sensitivity of 84% and specificity of 92% for the recognition of pancreatic cancers. In comparison, Corral et al. reported an efficacy of 78% using a DL model with a sensitivity of 92% and a specificity of 52% for recognizing malignant IPMNs [109,110].

5.2. AI and Endoscopic Diagnosis of Pancreatic Cancer

Several studies have reported excellent accuracy of detection of pancreatic cancer from the normal pancreas. Ozkan et al. studied ANN and reported an accuracy of 91.66% in patients over 60 years old, while Mayra et al. and Tonozuka et al. studied CNN with a reported ROC AUC of 0.94 and 0.957, respectively, to identify pancreatic cancer from the noncancerous pancreas [111–113]. All studies still used EUS images. Saftoiu et al. used video images collected prospectively and fed in an ANN model with a reported accuracy of 89.7% to differentiate malignant from benign patterns [114].
The presence of chronic pancreatitis complicates the diagnosis of pancreatic cancer, where standard EUS has low specificity, and cytology remains the mainstay of diagnosis. AI has been shown to improve the diagnostic capability of EUS to differentiate cancer from chronic pancreatitis compared to cytology. Zhang et al. showed an accuracy of 94% with 93% specificity to differentiate cancer from chronic pancreatitis using a support vector machine model of AI in a retrospective analysis of 388 patients using still EUS images [115]. Three prospective studies were conducted by Saitou et al. using video images of EUS fed in an ANN model also reported a high accuracy (90%) with sensitivities 88–95% and specificities 83–94%, much higher than traditional EUS alone [114,116,117].

Kuwahara et al. showed that the CNN model of AI had a 94% accuracy in differentiating malignant from benign IPMNs. They also reported very high sensitivity, specificity, PPV, and NPV of 96%, 93%, 92%, and 96%, respectively. However, this study was a small retrospective analysis of still EUS images from 50 patients (23 malignant and 27 benign IPMNs) [118].

5.3. AI and microRNA (miRNA) for Diagnosis of Pancreatic Cancer

Non-invasive biomarkers such as CA19-9 are non-specific in the diagnosis of pancreatic cancer. MicroRNA expressions are found to be unique in different gastrointestinal cancers, and their small size, stability, and easy detection in serum, make them potentially attractive for the early diagnosis of pancreatic cancer [119,120]. Yan et al. studied 100 pancreatic cancer patients and 150 controls to identify 13 miRNA expressions unique to pancreatic cancer with sensitivity, specificity, and accuracy of 80%, 97.6%, and 91.6%, respectively, and it was better than conventional biomarkers (CEA and CA19-9) [121]. Savareh et al. combined five unique miRNAs with an AI model consisting of an ANN with Particle Swarm Optimization (PSO) and Neighborhood Component Analysis (NCA) and reported a sensitivity, specificity, and accuracy of 93%, 92%, and 93%, respectively [122]. Sinkala et al. showed that combining AI with mRNA, miRNA, and DNA methylation patterns can recognize two distinct molecular subtypes of pancreatic cancer (one aggressive and one not so aggressive) that may have therapeutic implications [123].

Overall, AI has shown considerable promise in the early detection of pancreatic malignancies and may become a useful adjunct to interventional gastroenterologists in the near future. Studies with CT scans and MRI showed a range of sensitivities and specificities, while a custom model of AI with PET/CT images showed promise. Endoscopic studies using AI in EUS imaging analysis showed higher sensitivities and specificities than traditional EUS to recognize pancreatic malignancies, and AI combined with miRNA also shows promise. If successful, they may help establish a diagnosis of pancreatic cancer in a non-invasive manner without cytological confirmation.

6. Hepatocellular Cancer

Hepatocellular cancer (HCC) is a common hepatic tumor that can be diagnosed with specific radiological criteria without the need for tissue biopsy in the appropriate patient population [124]. Less than half of the patients present with the local disease that is potentially curable with a 5-year OS of 35%, while patients with distant metastases have a dismal 5-year OS survival of <3% [79]. AI has the potential to augment the accuracy of diagnosis of early HCC and help increase the likelihood of early diagnosis and treatment. Here, we discuss the evidence supporting the role of AI in the diagnosis of HCC.

6.1. AI and Ultrasound Diagnosis of HCC

Virmani et al. used the SVM technique to obtain a classification accuracy of 88.8% [125]. Wu et al. used a fused feature model comprising of k-NN, fuzzy-NN, PNN, and SVM to obtain a diagnostic accuracy of 96.6% to identify HCC, while Lee et al. and Bharti et al. used multiple AI techniques (k-NN, fuzzy-NN, PNN, SVM) and combined them using CNN-DL techniques (ensemble model) to obtain a classification accuracy of 95.7% and 96.6%, respectively, to differentiate HCC from the normal and cirrhotic liver [126–128].
Schmauch et al. used a DL algorithm on 177 patients to obtain the ROC-AUC of 0.931 for the characterization of HCC amongst focal liver lesions [129].

6.2. AI and CT-Scan Diagnosis of HCC

Cao et al. utilized a deep neural network (DNN)—automated multiphase convolutional dense network (MP-CDN) in 375 patients with 517 lesions to classify liver lesions into four groups (A-HCC, B-liver metastases, C-benign non-inflammatory lesions such as cysts, adenomas, and D-liver abscesses). They reported that the AUCs for differentiating each category from the others were 0.92, 0.99, 0.88, and 0.96, respectively, for HCC, metastases, benign non-inflammatory lesions, and abscesses [130]. Yasaka et al. applied a CNN model to CT-scan images of 100 patients with liver lesions and divided them into five groups (A-classic HCC, B-non-HCC tumors, C-indeterminate lesions, or benign solid masses, D-hemangiomas, E-cysts). They reported an overall accuracy of 84% with an ROC curve of 0.92 to differentiate A-B from C-E lesions [131].

6.3. AI and MRI Diagnosis of HCC

Hamm et al. applied the CNN model to multiphasic MRI images in 494 hepatic lesions and compared the performance of the AI model to radiologists. The CNN model showed a sensitivity and specificity of 92% and 98%, respectively, compared to 82.5% sensitivity and 96.5% specificity for radiologists reading the same cases. The overall accuracy of AI was 92%, with a false positive rate of 1.6% and an ROC of 0.992 [132]. Oestmann et al. also studied a CNN model in 118 patients with 150 pathologically confirmed liver lesions to differentiate HCC from non-HCC lesions. The model showed a sensitivity and specificity of 92.7% and 82%, respectively, to identify HCC with an overall accuracy of 87.3% and ROC of 0.912 [133].

Overall, AI has shown considerable progress in application with US, CT-scans, and MRIs to increase HCC diagnosis accuracy and early detection, with some studies showing better sensitivities and specificities than radiologists. AI has the potential to become an excellent adjunct to radiology techniques in diagnosing HCC in the future.

7. AI in Histopathologic Diagnosis of GI Malignancies

Several studies have explored the role of AI in assisting with pathological diagnosis of GI cancers. Of all AI methodologies, CNN has been the most widely studied with regards to applications in the pathological diagnosis of cancer. If well built, it has the potential to ease the pathologist’s workload and increase the efficiency of pathological diagnosis [9,134]. While Sharma et al. reported an accuracy of 0.699 for the classification of cancer with DL methodology applied to digital images of H&E-stained whole slide, Iizuka et al. reported AUCs of 0.924 and 0.982 for the histological classification of gastric adenocarcinoma and colon adenocarcinoma, respectively, using DL methodology [135,136]. Kuntz et al. reported a systematic review of 16 studies using CNN in making histologic diagnosis or assessing the prognosis of CRC and GC. All studies assessing molecular characteristics were conducted in CRC, and no study on the use of AI in assessing the molecular characteristics of gastric or esophageal cancer was found. There was a wide range of sensitivities and specificities of different studies with sensitivities ranging from 52 to 100% and specificities from 57 to 100%. The performance of AI was as good as or better than pathologists. However, all studies were retrospective and there was considerable variability in the types of CNN methodologies applied, with 14 different CNN techniques in 16 studies [9]. Momemi-Boroujeni et al. studied multilayer perception neural network to recognize benign, malignant, and atypical pancreatic lesions. Although AI was 100% accurate at differentiating benign from malignant pathology, the accuracy fell to 77% for atypia [137]. Due to the large amount of variability in reported data, AI is not yet ready for real time application in pathology practice to diagnose GI malignancies, but the ongoing work is extremely promising [138]. Overall, there is a need to identify an AI model for pathological
diagnosis that is generalizable on a large scale and can be commercially developed for applicability in clinical practice.

8. Conclusions

This review outlined the current published literature on the AI application in gastrointestinal, pancreatic, and hepatocellular cancers. AI is considered an instrumental tool in changing the future of healthcare, especially oncology. AI may become a useful tool in screening, diagnosing, and treating various cancers by accurately analyzing diagnostic clinical images, identifying therapeutic targets, and processing large datasets. Although the diagnostic accuracy of AI systems has markedly increased, it still needs collaboration with physicians. Robertson et al. proposed a five-step process that they foresee can help integrate AI with clinical practice, namely, Quality improvement, Productivity improvement, and Performance improvement, followed by Evaluation (step 4) where AI replaces human analysis and the results can be reviewed by the gastroenterologist before being reported to the final step 5, Diagnostic, where AI may replace the diagnostican for simple pathologic results releasing them to the patients without review [75]. However, before we can follow these steps, there is a need to identify one or two methodologies from numerous choices that can be developed, generalized, and commercialized. Most of the data on AI use are not prospective, hence large and multicenter clinical trials are needed to further validate this system in real-time clinical settings. If successful, the AI-assisted systems have the potential to become a vital tool for the management of these cancer patients.

Author Contributions: Conceptualization, H.G.; methodology, S.A.A.S. and R.M.; writing—original draft preparation, S.A.A.S., R.M., and Z.G.; writing—review and editing, A.P., M.A., S.C., J.K., B.T., N.S. and N.T.; supervision, H.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References


