Artificial intelligence: Emerging player in the diagnosis and treatment of digestive disease

Abstract
Given the breakthroughs in key technologies, such as image recognition, deep learning and neural networks, artificial intelligence (AI) continues to be increasingly developed, leading to closer and deeper integration with an increasingly data-, knowledge- and brain labor-intensive medical industry. As society continues to advance and individuals become more aware of their health needs, the problems associated with the aging of the population are receiving increasing attention, and there is an urgent demand for improving medical technology, prolonging human life and enhancing health. Digestive system diseases are the most common clinical diseases and are characterized by complex clinical manifestations and a general lack of obvious symptoms in the early stage. Such diseases are very difficult to diagnose and treat. In recent years, the incidence of diseases of the digestive system has increased. As AI applications in the field of health care continue to be developed, AI has begun playing an important role in the diagnosis and treatment of diseases of the digestive system. In this paper, the application of AI in assisted diagnosis and the application and prospects of AI in malignant and benign digestive system diseases are reviewed.

Key Words: Artificial intelligence; Digestive disease; Convolutional neural network; Deep learning; Review

Core Tip: With the continuous development of artificial intelligence, its integration in the field of medical and health care has received increasing attention, allowing the development and application of medical expert systems and artificial neural networks in the medical field. The development of artificial intelligence has resulted in not only more accurate diagnoses of digestive system diseases but also new treatments. Further research on the progress of artificial intelligence in digestive system diseases is needed to better serve patients.

INTRODUCTION

Artificial intelligence (AI) is a branch of computer science inspired by human intelligence that operates in a different way. AI can be considered a set of computing techniques that allow machines to sense, learn, reason, act, etc. AI was first used at the Dartmouth AI Conference in the summer of 1956, which is considered the official birth of AI. In the 1990s, as technology continued to improve, hardware became more affordable, performance became more reliable, computing power, storage capacity and the ability to process data were greatly improved, and AI returned to life. In recent years, the application of AI in the medical field has reached an unprecedented height and scale. To date, great progress has been achieved in the use of AI in medicine and health, including speech recognition, medical imaging, wearable devices, risk management, pathology, etc.

The digestive system contains the largest number of organs in the human body and can be divided into the digestive tract and digestive glands. The digestive tract includes the oral cavity, pharynx, esophagus, stomach, duodenum, small intestine, large intestine, colon, etc. The digestive glands comprise the salivary glands, liver, pancreas,
etc. The digestive system digests food, takes up nutrients and eliminates waste. The coordination of gastrointestinal physiological activities depends on the completion of the normal physiological function of the digestive system. Digestive system diseases, which are the most common clinical diseases, have complicated clinical manifestations, are interrelated with each other, and form a hot spot in the medical field[1]. The incidence of digestive system diseases is high, and often, there are no obvious symptoms in the early stage. Consequently, most digestive diseases are first noted in the middle and late stages; at this point, the prognosis is poor. As AI technology has continued to be developed, its strong computing power and learning ability have been gradually applied to solve complex medical problems. In the field of digestive diseases, such as gastric, liver, and pancreatic cancer and ulcerative colitis, AI can help improve the accuracy and sensitivity of the diagnosis and provide treatment options (Figure 1). AI not only improves the diagnostic accuracy of the disease but also provides a plan for clinical treatment.

In this review, we discuss AI and related technologies, their applications in digestive diseases and their future developments. We used PubMed and Google to search for relevant articles by using the keywords “digestive system diseases”, “artificial intelligence”, “machine learning”, “deep learning”, “computer-aided diagnosis”, “convolutional neural network”, and “endoscopy and endoscope images” for this review.

**AI AUXILIARY DIAGNOSIS**

In China, the development of AI has received considerable attention, particularly in the application of AI in the medical field. The development of AI has laid a solid foundation for precision medicine. In 2017, the China State Council announced that AI application should be promoted in the health care system. Subsequently, hundreds of new companies dedicated to the application of medical AI have emerged in China, promoting the development of AI in the medical field[6]. AI effectively improves the uneven allocation of medical resources, reduces medical costs and enhances medical
efficiency. AI algorithms improve the ease of medical work and lead to not only technological innovations but also a transformation of the medical service mode. AI is applied in all fields of medicine and health, especially in the auxiliary diagnosis of diseases, and includes the following three aspects: intelligent guidance, electronic health records and risk monitoring of diagnosis and treatment.

**Intelligent guidance**

Outpatient triage has gradually transitioned from a traditional manual system to an internet- and intelligence-based system. Most admitted patients lack an understanding of their disease and the hospital. To improve the efficiency and experience of the patients, intelligent guidance systems have been developed. Patients enter the Mini Program through screens on official hospital accounts to describe symptoms or diseases, and guidance assistants intelligently follow up and guide the patients to provide symptoms to accurately match departments, recommend doctors with the most consistent professional direction, and perform online guidance registration in one step. The research and development leading to intelligent guidance systems has led to a reduction in not only the workload of outpatient staff but also the human resource needs and costs of hospitals. These systems can guide patients to seek medical treatment quickly and accurately, improve the accuracy of patient registration and the efficiency of the physician response, increase efficiency at all steps of the patient stay, and effectively provide accurate medical treatment (Figure 2).

**Electronic health record**

Traditional health records are typically kept on paper, which tend to be scattered in hospital archives, limiting their value. The widespread use of the internet has facilitated the creation of electronic health records (EHRs), which are currently widely used worldwide[5]. EHRs can improve management processes, overcome problems with paper documentation, and ensure the quality of treatment. Both theoretically and practically, EHRs have guiding and economic value, while their applicative value
continues to improve. In addition, EHRs can improve clinical decision-making, collaboration between health care teams, and medical efficiency.

Risk monitoring of diagnosis and treatment
Diagnostic risk monitoring systems can improve the diagnostic accuracy and greatly reduce the risk of misdiagnosis, and in this regard, AI has significant advantages over clinical experience. Compared with the United States, domestic equipment is used much more frequently under the same resource allocation[6]. This is the main point of medical care in China. Additionally, the large associated workload and high repeatability result in a cumbersome manual analysis and diagnosis. An accurate diagnosis rate is important for ensuring that radiologists can provide high-quality care to patients. A report in 1999 showed that 98000 people die of misdiagnosis each year, and diagnostic errors by radiologists are a key contributor[7]. Under a heavy workload, the diagnosis results can only be judged and analyzed by subjective individual experience, leading to high misdiagnosis and missed diagnosis rates. The large market demand for pathologists and diagnostic efficiency have also led to new challenges in the medical service industry. However, digestive system diseases increasingly depend on auxiliary diagnoses based on medical images while continuing to demonstrate a yearly increase in incidence.

The emergence of AI has led to its gradual use in digestive diseases. Machine learning (ML) is a branch of AI that aims to predict results by analyzing existing data and performing calculations based on an understanding of these data. Deep learning (DL) is a relatively new branch of ML that imitates learning through machine algorithms and provides a set of possible and reasonable solutions[8,9]. In esophageal cancer, van der Sommen et al[10] developed an automatic algorithm to distinguish early neoplastic lesions. Huang et al[11] used 84 images from 30 patients to train an AI-based neural network model to predict Helicobacter pylori infection. DL can be used to share more comprehensive medical diagnosis and treatment information with doctors with lightweight methods. At a higher level, we can comprehensively manage and guide the
overall diagnosis and treatment level of the physician team to build a more scientific, accurate and friendly diagnosis and treatment risk monitoring system.

**MALIGNANT DIGESTIVE DISEASES**

The digestive system is a common site for malignant tumors, particularly gastric cancer, liver cancer, colorectal cancer and pancreatic cancer. Given the changes in lifestyle, diet and disease spectrum of the population, the incidence of malignant tumors of the digestive system continues to rise\(^\text{[12]}\). The early diagnosis of malignant tumors is critical for treatment, but patients with malignant tumors often have no obvious symptoms at the early stage, leading to difficulties in early diagnosis. This section summarizes the application of AI in the following three types of malignant tumors with high morbidity and mortality: gastric, liver and pancreatic cancer. The latest clinical studies concerning AI, gastric cancer, liver cancer, and pancreatic ductal adenocarcinoma are summarized in Table 1.

**AI and gastric cancer**

Gastric cancer (GC) is the fifth most common cancer and ranks third in cancer mortality worldwide, with more than 1 million new cases and 780000 deaths each year\(^\text{[13]}\). The stage at which GC is diagnosed can greatly affect the prognosis, with a poor prognosis typical in the late stage and a 5-year survival rate of early GC (EGC) exceeding 90\%\(^\text{[14]}\). There are two difficulties in the early diagnosis of GC. First, EGC is difficult to detect, often presenting with slight reddish bulges or depressions under endoscopy that are not easily distinguished from inflammatory lesions. The second difficulty lies in judging the depth of the invasion of GC. In general, intramucosal GC or GC invading the superficial submucosal layer should be resected endoscopically, while GC invading the deep submucosal layer should be resected with open surgery; otherwise, there could be a risk of lymph node and distant metastasis. However, it is difficult to distinguish among these three forms of GC. The rapid development of DL and ML has led to great progress in AI, whose image recognition has transcended that of humans. By reviewing
the published literature, this section discusses the diagnosis, treatment and prognosis of GC with AI.

Hirasawa et al\cite{5} constructed a diagnostic system based on a convolutional neural network (CNN). The authors used 13584 endoscopic images to train the web-based learning model and 2296 GC images from 69 patients to evaluate its performance. The model took 47 s to analyze 2296 images and correctly diagnosed 77 GC lesions with a total sensitivity of 92%. This web-based learning model for the detection of EGC processes a large number of stored endoscopic images with clinical-related diagnostic capabilities in a very short time and may play a role in real-time gastroscopy\cite{5}.

Kubota et al\cite{6} used a back propagation neural network algorithm to analyze the depth of gastric wall infiltration in endoscopic images of 344 GC patients and developed a computer-aided design (CAD) system. The cross-validation results revealed that the diagnostic accuracy of T1, T2, T3 and T4 GC was 77%, 49%, 51% and 55%, respectively. The accuracy in detecting intramucosal carcinoma (T1a) and submucosal invasive carcinoma (T1b) is 69%\cite{6}. This study may be used to screen for EGC in the near future.

**AI and liver cancer**

Liver cancer (LC) is the second leading cause of death worldwide, and its incidence continues to increase\cite{7}. In recent years, AI has been used in various aspects of LC diagnosis and treatment, especially in the field of imaging. However, individual differences in liver morphology and the internal anatomy and the highly heterogeneous molecular and pathological features of LC lead to great challenges in the implementation of AI, ML and DL for the diagnosis and treatment of this disease\cite{8}.

The following two modes of AI application are commonly used in the diagnosis of LC: DL and radiomics. Imaging plays a key role in the identification and diagnosis of hepatobiliary diseases and the establishment of surgical treatment plans. Radiologists and surgeons demarcate organs and diagnose diseases based on their experience and judgment, which can be subjective, time consuming and unrepeatable; AI can
address these limitations. DL is mainly used for the diagnosis of LC via traditional machine learning (TML) and CNNs. In TML, a well-developed anatomical model must be established in advance and manually corrected in the later stages, while the application of CNNs allows images to be directly used in the learning process, allowing the extraction and utilization of all information contained in the image for DL. Shaw Hospital of Zhejiang University developed a DL system for liver tumor diagnosis through a CNN. In addition to learning more than 30000 magnetic resonance (MR) images, the system incorporates patient clinical data (medical records, examination results, etc.) and demonstrates excellent diagnostic performance with a coincidence rate with a final pathological result of 92%.

The application of AI in the treatment of LC has led to new opportunities and challenges, promoting the standardization and intelligence of LC management. Hu et al. used 3D visualization to evaluate the spatial relationship between tumors and surrounding blood vessels before hepatectomy, which serves as a suitable guide for surgical resection. In the treatment of LC by radiofrequency ablation, fusion imaging can integrate enhanced computed tomography (CT)/MR images with real-time ultrasound imaging by electromagnetic tracking to accurately locate the target lesions. AI has been applied to all aspects of LC treatment and auxiliary diagnosis, but further prospective research is needed.

**AI and pancreatic ductal adenocarcinoma**

Pancreatic ductal adenocarcinoma (PDAC) is a rapidly progressive and highly malignant tumor of the digestive tract. Although many methods have been developed for the diagnosis and treatment of PDAC in recent years, the effective rate of treatment and prognosis remain poor (the 5-year survival rate is only 9%). The early detection and accurate treatment of PDAC are important directions requiring urgent solutions.

Recently, rapid developments have been achieved in the AI-based screening and diagnosis of PDAC, leading to acceptable effects. Ansari et al. applied artificial neural
networks (ANNs) to a long-term survival prediction of patients after radical surgery for PDAC for the first time and achieved a consistency index of 0.79, indicating that their ANN could predict the survival rate of PDAC patients after radical surgery. Fine needle aspiration (FNA) biopsy is an elaborate approach used to diagnose solid masses of the pancreas. Due to the small number of cases, its development is limited. Momeni-Boroujeni et al[28] used a multilayer-perceptron neural network (MNN) to evaluate 277 cell mass images from 75 pancreatic FNA patients and achieved an accuracy of 100% in the classification of benign and malignant tumors. In particular, the sensitivity and specificity of MNN in identifying atypical cases were 80% and 75%, respectively. Intraductal papillary mucinous neoplasm (IPMN), a papillary mucinous tumor with many subtypes, has particular malignant potential. Endoscopic ultrasonography (EUS) can reveal the structure of the pancreas and the dilation of the pancreatic duct and can be used to evaluate the degree of malignancy of IPMN. Indeed, Kuwahara et al[29] used AI to build a system to perform this function. The authors uploaded 3970 EUS images of 50 patients to the system for DL analysis. The endpoint of their system was the accuracy in diagnosing the malignancy degree of IPMNs by AI. The results showed that the AI system had significantly better value in diagnosing malignant IPMNs than benign IPMNs, with an area under the receiver operating characteristic curve of 0.91 in diagnosing malignant IPMNs. In addition, a three-dimensional depth supervised segmentation network has been developed to classify CT images from PDAC patients and control groups. Chu et al[30] standardized the number of cases, the sizes of the segmented images, and the selection of organs and then segmented the boundaries of the abdominal organs and pancreatic tumors in 575 control subjects and 750 patients with PDAC. The results showed that the sensitivity and specificity of the algorithm in detecting PDAC in patients were 94.1% and 98.5%, respectively. The early diagnosis of PDAC is a problem that many physicians find difficult. Appelbaum et al[31] developed a risk model for the prediction of PDAC based on patient EHRs that could identify high-risk groups one year in advance, suggesting that their risk scores could also be used as an initial filter for PDAC.
BENIGN DIGESTIVE DISEASES
In addition to gastric cancer, liver cancer and other malignant tumors, other digestive system diseases include ulcerative colitis, pancreatitis, gastric ulcer, gastritis, intestinal polyp and other benign digestive system diseases. We summarize the application of AI for the benign digestive system diseases discussed below. The latest applications of AI in ulcerative colitis, severe acute pancreatitis, gastric ulcers, gastritis and intestinal polyps are summarized in Table 2.

AI and ulcerative colitis
Ulcerative colitis (UC) is a common chronic intestinal disease; however, there is no single gold standard for the diagnosis of this disease. Combining clinical, laboratory, imaging, endoscopic and histopathological findings, the diagnosis is made by excluding infectious and other noninfectious forms of colitis[32]. Endoscopy is essential, but the accuracy and precision of the diagnosis are limited by doctors’ subjectivity. Yao et al[33] established a fully automatic endoscopic disease grading video analysis system through AI. These authors used a CNN to train models to predict the amount of information in still images and, thus, determine the severity of the disease. Then, the authors examined 51 and 264 videos from developed high-resolution and multicenter clinical trial test sets, respectively. The results showed that the static image information classification system had good performance, with a sensitivity of 0.902 and specificity of 0.870. For the high-resolution videos, the fully automated methods correctly predicted the Mayo endoscopic score (MES) in 78%. The automated MES grading of the clinical trial videos (often low resolution) correctly distinguished remitted vs active disease in 83.7%. To analyze endoscopic images of UC more accurately, Takenaka et al[34] developed a deep neural network system and tested 40758 colonoscopy images and 6885 biopsy results from 2012 UC patients. The results identified endoscopic remission with 90.1% accuracy and histologic remission with 92.9% accuracy. In addition, Maeda et al[35] developed a CAD system to predict persistent histologic inflammation using endocytoscopy and
evaluated its accuracy using 12900 endoscopy images. The results showed that the CAD system provided a diagnostic sensitivity, specificity, and accuracy of 74%, 97% and 91%, respectively. In recent years, great progress has been achieved in the treatment of UC, but surgery remains important[36]. These results show that AI has great potential in the diagnosis and treatment of UC.

*AI and pancreatitis*

Acute pancreatitis (AP) is a common clinical acute abdomen disease with a complex and variable presentation. Severe AP (SAP) is often life threatening, with multiple organ dysfunction induced by SAP being a particularly difficult clinical problem[37]. Qiu et al[38] developed three ML algorithms to predict the risk of organ failure in moderate SAP (MSAP) and SAP. These authors inputted 16 parameters from 263 patients with MSAP and SAP into a support vector machine (SVM), logical regression analysis (LRA) and ANN models. The results showed that the SVM, LRA and ANN models could effectively predict the risk of multiorgan failure in MSAP and SAP. Respiratory failure is the main cause of early death in SAP. The early prediction of acute lung injury (ALI) is beneficial for reducing mortality from SAP. The ANN model established by Fei et al[39] can effectively predict the occurrence of ALI induced by SAP, with the accuracy of predicting severe acute respiratory distress syndrome reaching 82.8%. Abdominal infection is an important risk factor for organ failure induced by SAP. Compared with logistic regression analyses, the ANN model developed by Qiu et al[40] can more accurately predict abdominal infection in MSAP and SAP, with a sensitivity and specificity of 80.99% and 89.44%, respectively. Fei et al[41] constructed and verified an ANN model based on clinical and laboratory data from 72 patients with AP. Notably, the sensitivity, specificity, positive predictive value and negative predictive value of the ANN model for portosplenicomesenteric venous thrombosis (PSMVT) were 80%, 85.7%, 77.6% and 90.7%, respectively. Thus, this model provides a good reference for the clinical prediction of the occurrence of PSMVT. The recurrence rate after the first attack of AP is approximately 22%, but few indicators can predict this recurrence. Chen et al[42]
developed a quantitative radiology model based on contrast-enhanced CT. The results of the verification analysis showed that the model can predict the recurrence of AP well in the main and verification populations, with accuracies of 87.1% and 89.0%, respectively. This finding suggests that their model may be a good quantitative method for the prediction of recurrence in patients with AP. Chronic pancreatitis (CP), which is another common clinical digestive system disease, is an important factor in inducing pancreatic cancer. Although traditional pancreaticoduodenectomy, distal pancreatectomy, and the Frey and Beger procedures have achieved considerable success in relieving CP pain, the approaches are often accompanied by serious complications. Therefore, minimally invasive surgery and robot platforms are gradually being applied in the surgical treatment of CP. A retrospective study based on 812 cases of robotic pancreatectomy and reconstruction showed that the robotic platform is safe and feasible in performing complex pancreatectomy to alleviate the sequelae of CP\(^{[43]}\). Autoimmune pancreatitis (AIP) is a rare and special type of chronic pancreatitis, and its diagnosis is clinically challenging. Marya et al\(^{[44]}\) developed and validated a CNN model based on a database containing 174461 EUS images and videos from 583 patients. After training, the model could accurately distinguish AIP from PDAC and other benign pancreatic diseases with a sensitivity and specificity of 90% and 85%, respectively.

**AI and other diseases**

AI has been used in numerous applications for other diseases, such as gastric ulcers, gastritis and colorectal polyps. Namikawa et al\(^{[45]}\) developed an advanced CNN (A-CNN) based on the original convolutional neural network and evaluated its applicability in the classification of gastric ulcers. These authors used 720 images from 120 patients with gastric ulcers to evaluate the diagnostic performance of the A-CNN. The results showed that the sensitivity, specificity, and positive predictive value of the A-CNN in classifying gastric ulcers were 93.3%, 99.0%, and 99.1%, respectively. There are several subtypes of gastritis, including autoimmune, bacterial, and chemical. Steinbuss et al\(^{[46]}\) evaluated the ability of a CNN to classify the common gastritis
subtypes using a small data set of 1230 gastric antrum and corpus biopsy images. The results showed that the overall accuracy in the test set was 84%, and the sensitivity and specificity for bacterial gastritis were 100% and 93%, respectively. Ozawa et al. constructed a deep CNN architecture and trained the CNN using 16418 images from 4752 colorectal polyps and 4013 normal colorectal images. Then, the performance of the trained CNN was verified with 7077 additional colonoscopy images. The results showed that the sensitivity of the CNN was 92%, and the positive predictive value was 86%, proving that it has excellent potential as a colonoscopy CP-diagnosis support system based on AI.

**CONCLUSION**

With the steady development of AI and its heavy integration in the medical industry, AI has received increasing attention in the fields of medicine and health. Digestive system AI technology differs from other AI technologies because of the nonstandard nature of most digestive endoscopic images and the large number of influencing factors, resulting in difficulty in developing such technology. Currently, AI plays an important role in digestive system diseases, particularly in its gradual application in diagnosis and treatment in the clinic. AI has broad prospects in the field of medicine and health, but the cooperation of medical workers and computer experts is needed to develop more intelligent applications and higher-quality services for patients. In summary, as an emerging player, AI will continue to play an increasingly important role in future development. The close combination of AI and medicine can address problems, such as the insufficient supply and uneven distribution of medical resources and low degree of medical efficiency. More in-depth research should be performed to investigate the application of AI in digestive system diseases from various perspectives to obtain additional diagnosis and treatment measures, promote diagnosis, optimize treatment, improve patient prognosis, and gradually achieve intelligent and accurate medical treatment.
REFERENCES


56 Popa IV, Burlacu A, Mihai C, Prelipcean CC. A Machine Learning Model Accurately Predicts Ulcerative Colitis Activity at One Year in Patients Treated with Anti-Tumour Necrosis Factor α Agents. *Medicina (Kaunas)* 2020; 56 [PMID: 33233514 DOI: 10.3390/medicina56110628]


**Figure 1** Artificial intelligence and digestive diseases. AI: Artificial intelligence; GC: Gastric cancer; LC: Liver cancer; PDAC: Pancreatic ductal adenocarcinoma; UC: Ulcerative colitis; AP: Acute pancreatitis; GU: Gastric ulcer.

**Figure 2** Artificial intelligence and Intelligent guidance. CNN: Convolutional neural network; DL: Deep learning; ML: machine learning; CAD: Computer-aided design.
<table>
<thead>
<tr>
<th>Ref.</th>
<th>Study design</th>
<th>Study population</th>
<th>Year</th>
<th>Disease</th>
<th>Country /Region</th>
<th>Number of cases</th>
<th>Methods</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tang et al[49]</td>
<td>Retrospective</td>
<td>Hospital</td>
<td>2020</td>
<td>GC</td>
<td>China</td>
<td>45240 Images</td>
<td>DCNN</td>
<td>Accuracy: 85.1%-91.2%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sensitivity: 85.9%-95.5%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Specificity: 81.7%-90.3%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>AUC: 0.887-0.940</td>
</tr>
<tr>
<td>Zhang et al[50]</td>
<td>Retrospective</td>
<td>Hospital</td>
<td>2020</td>
<td>GC</td>
<td>China</td>
<td>21,217 Images</td>
<td>CNN</td>
<td>Accuracy: 78.7%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Accuracy: WLI (94.5%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Accuracy: NBI (94.3%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Accuracy: Indigo (95.5%)</td>
</tr>
<tr>
<td>Song et al[51]</td>
<td>Retrospective</td>
<td>Hospital</td>
<td>2020</td>
<td>GC</td>
<td>China</td>
<td>3212 Images</td>
<td>DL</td>
<td>Accuracy: 0.873</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sensitivity: 0.996</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Specificity: 0.843; AUC 0.986</td>
</tr>
<tr>
<td>Ma et al[52]</td>
<td>Retrospective</td>
<td>Hospital</td>
<td>2020</td>
<td>GC</td>
<td>China</td>
<td>763 Images</td>
<td>CNN</td>
<td>Accuracy: 98.4%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Specificity: 98.9%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sensitivity: 98.0%;</td>
</tr>
<tr>
<td>Zhen et al[53]</td>
<td>Retrospective</td>
<td>Hospital</td>
<td>2020</td>
<td>LC</td>
<td>China</td>
<td>31608 Images</td>
<td>CNN</td>
<td>AUC: 0.946</td>
</tr>
<tr>
<td>Giordano et al[54]</td>
<td>Retrospective</td>
<td>Hospital</td>
<td>2020</td>
<td>LC</td>
<td>Italy</td>
<td>167 Cases</td>
<td>SVM, RF</td>
<td>Accuracy: Exceeded 94%</td>
</tr>
<tr>
<td>Jeong et al[55]</td>
<td>Retrospective</td>
<td>Cancer Institute</td>
<td>2020</td>
<td>LC</td>
<td>China</td>
<td>1421 Cases</td>
<td>DL</td>
<td>AUC: 0.84</td>
</tr>
<tr>
<td>Appelbaum et al[56]</td>
<td>Retrospective</td>
<td>Hospital</td>
<td>2020</td>
<td>PDAC</td>
<td>America</td>
<td>594 Cases</td>
<td>LR</td>
<td>AUC: 0.71</td>
</tr>
<tr>
<td>Marya et al[57]</td>
<td>Retrospective</td>
<td>Hospital</td>
<td>2020</td>
<td>PDAC</td>
<td>America</td>
<td>1174461 Images</td>
<td>EU3 EUS-CNN</td>
<td>Sensitivity: 99%</td>
</tr>
<tr>
<td>Tonozuka et al[58]</td>
<td>Retrospective</td>
<td>Hospital</td>
<td>2020</td>
<td>PDAC</td>
<td>Japan</td>
<td>920 Cases</td>
<td>EUS-CAD</td>
<td>Specificity: 98%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>AUC: 0.94</td>
</tr>
</tbody>
</table>

Table 2 Recent researches on artificial intelligence in benign digestive system diseases

<table>
<thead>
<tr>
<th>Ref.</th>
<th>Study design</th>
<th>Study population</th>
<th>Year</th>
<th>Disease</th>
<th>Country/Region</th>
<th>Number of cases</th>
<th>Methods</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maeda et al[88]</td>
<td>Retrospective</td>
<td>Hospital</td>
<td>2019</td>
<td>UC</td>
<td>Japan</td>
<td>12900 Images</td>
<td>CAD</td>
<td>Sensitivity: 74%; Specificity: 97%; Accuracy: 91%</td>
</tr>
<tr>
<td>Popa et al[56]</td>
<td>Retrospective</td>
<td>Hospital</td>
<td>2020</td>
<td>UC</td>
<td>Romania</td>
<td>55 Cases</td>
<td>ML</td>
<td>Accuracy: 96%; AUC: 0.92</td>
</tr>
<tr>
<td>Tong et al[57]</td>
<td>Retrospective</td>
<td>Hospital</td>
<td>2020</td>
<td>UC</td>
<td>China</td>
<td>6399 Cases</td>
<td>RF, CNN (0.89); Sensitivity: 91.00%</td>
<td></td>
</tr>
<tr>
<td>Qiu et al[10]</td>
<td>Retrospective</td>
<td>Hospital</td>
<td>2019</td>
<td>SAP</td>
<td>China</td>
<td>263 Cases</td>
<td>ANN</td>
<td>Sensitivity: 80.99%; Specificity: 89.44%</td>
</tr>
<tr>
<td>Chen et al[44]</td>
<td>Retrospective</td>
<td>Hospital</td>
<td>2019</td>
<td>SAP</td>
<td>China</td>
<td>389 Cases</td>
<td>LR</td>
<td>Accuracy: 87.1%</td>
</tr>
<tr>
<td>Namikawa et al[45]</td>
<td>Retrospective</td>
<td>Hospital</td>
<td>2020</td>
<td>Gastric ulcer</td>
<td>Japan</td>
<td>720 Images</td>
<td>A-CNN</td>
<td>Sensitivity: 93.3%; Specificity: 99.0%; PPV: 99.1%</td>
</tr>
<tr>
<td>Steinbuss et al[46]</td>
<td>Retrospective</td>
<td>Hospital</td>
<td>2020</td>
<td>Gastritis</td>
<td>Germany</td>
<td>1230 Images</td>
<td>CNN</td>
<td>Accuracy: 84%; Sensitivity: 100%; Specificity: 93%</td>
</tr>
<tr>
<td>Ozawa et al[47]</td>
<td>Retrospective</td>
<td>Hospital</td>
<td>2020</td>
<td>Colorectal polyps</td>
<td>Japan</td>
<td>16418 Images</td>
<td>CNN</td>
<td>Sensitivity: 92%; PPV: 86%</td>
</tr>
</tbody>
</table>