Role of Artificial Intelligence in Upper Gastrointestinal Diseases: Present and Future

Jin Lee*
Division of Gastroenterology, Department of Internal Medicine, Inje University Haeundae Paik Hospital, Busan 48108, Korea

ABSTRACT
There has been extensive research and progress in the medical application of Artificial Intelligence (AI) technology. Radiologic and endoscopic diagnosis based on image analysis has been one of the main areas of focus. In the field of Gastrointestinal (GI) endoscopy, AI has been applied to gastroscopy, capsule endoscopy, and colonoscopy. There are relatively fewer studies in which AI is applied to upper GI diseases, and analysis has mainly focused on the detection of esophageal or gastric neoplasms in upper GI endoscopy. AI is also used to assess the depth of cancer invasion, detect Helicobacter pylori (H. pylori) infection, and diagnose the characteristics of subepithelial lesions. In the near future, AI is expected to expand to the field of disease treatment.

INTRODUCTION
Artificial Intelligence (AI) technology has received significant attention in the medical field, and Computer-Aided Diagnosis (CAD) is widely used to improve the accuracy of diagnosis. The use of AI is well established in lower Gastrointestinal (GI) diseases, but there are relatively fewer studies in which AI is applied in upper GI diseases. There may be several reasons for the slow progress in AI research on upper GI diseases; the most important reason is the characteristics of early upper GI neoplasms. These neoplasms are more complex than those of colo-rectal adenomas or cancers, and can appear in various forms such as depression, flatness, and uplift.

Recently, many studies on the application of AI in upper GI endoscopy have been attempted. The researchers mainly focused on the automatic detection of esophageal dysplasia and early esophageal and gastric cancer, to help endoscopists avoid missing and misdiagnosing malignant lesions and help solve the current limitations of endoscopy. AI is also used to assess cancer invasion depth, detect Helicobacter pylori (H. pylori) infection, and diagnose the characteristics of Subepithelial Lesions (SELs) such as Gastrointestinal Stromal Tumors (GISTs). In this article, we review the recently published literature and the present and future applications of AI in upper GI diseases.

OVERVIEW OF ARTIFICIAL INTELLIGENCE
AI refers to a computer or other technology that can mimic the cognitive functions of the human brain. Currently available AI is classified as “weak” AI, and mimics intelligence by using specific algorithms to perform tasks such as reading or analyzing endoscopic images according to human-directed commands. In theory, “strong” AI describes a self-aware system capable of independent problem solving. Development of this level of AI is controversial.

The field of AI has a long history, and innovative developments such as image recognition are now considered commonplace. Two factors have contributed to recent AI growth. The first is the availability and incorporation of big data. In fact, it is now more important to secure big data than to utilize artificial intelligence algorithms. The second is deep learning. Deep learning is a type of machine learning based on the human-like independent accumulation of knowledge. Deep learning uses an artificial neural network that enables computers to learn on their own like humans using multiple data. The human brain has an information processing method that classifies objects after discovering patterns in a number of data, and computers also learn to identify objects in this way. Reinforcement learning is needed in this process.

Deep learning uses Convolutional Neural Networks (CNNs), and these are most commonly applied to analyze visual images. CNNs are useful for finding patterns in order to analyze images. Images are learned directly from data and classified using patterns. At present, the deep learning method based on CNNs is widely applied to the analysis of GI endoscopic images for the detection of esophageal or gastric cancer, assessment of cancer invasion depth, detection of H. pylori, and assessment of the characteristics of subepithelial lesions.

PRESENT APPLICATIONS OF AI IN UPPER GI DISEASES

Esophageal diseases
Barrett’s esophagus: Dysplasia and adenocarcinoma: Barrett’s Esophagus (BE) is a precancerous esophageal lesion, and early detection of dysplastic and cancerous changes is important. Detecting adenocarcinoma in an early stage allows endoscopic resection instead of surgery and improves the patient’s prognosis and quality of life. Thus, automated detection of dysplastic changes using AI has been widely studied in GI endoscopy.

Van der Sommen et al.[1] developed a computer algorithm that could detect early neoplastic lesions using 100 images from 44 patients with BE. The system adopted algorithms based on
specific texture, color filters, and machine learning for conventional White-Light Endoscopic (WLE) images, and detected early neoplastic lesions with a sensitivity and specificity of 83% (per-image analysis). Furthermore, Swager et al.[2] developed a computer algorithm for detecting neoplastic lesions in patients with BE based on Volumetric Laser Endomicroscopy (VLE) images in 2017. The algorithm improved the detection rate of esophageal neoplastic lesions with a sensitivity of 90% and a specificity of 93%.

Recent studies used CNN analysis for detecting neoplastic lesions in patients with BE. Mendel et al.[3] conducted a CNN analysis using 50 WLE images of neoplastic lesions and 50 images of non-neoplastic lesions, which showed a sensitivity of 94% and a specificity of 88%. From the same research group, Ebigbo et al.[4] performed a CNN analysis using high-definition WLE and Narrow Band Imaging (NBI) and showed reasonable accuracy in detecting neoplastic lesions. They used 71 high-definition WLE and NBI images of neoplastic BE and non-dysplastic BE, which reached a sensitivity of 97% and specificity of 88% (WLE) and a sensitivity of 94% and specificity of 80% (NBI). In 2020, Hashimoto et al. assessed the CNN analysis of BE using 225 images of dysplasia and 233 images of non-dysplasia, achieving a sensitivity of 96% and a specificity of 94%.

**Squamous cell carcinoma:** Lugol's solution and narrow band imaging have been widely used to detect and identify Squamous Cell Carcinoma (SCC); however, this method is limited by low specificity. Recently, various advanced diagnostic tools have been developed (Table 1). These include an Image-Enhanced Endoscopy (IEE) system, Endocytoscopic System (ECS), and Intrapapillary Capillary Loops (IPCL). Several CAD-based studies using these advanced diagnostic tools have been reported. Table 2 shows the clinical studies of AI applied to esophageal diseases.

**Endocytoscopic** system: Endocytoscopy is a magnifying tool that enables *in vivo* visualization of superficial epithelial cells in real-time. Several studies have assessed the efficacy of CAD for distinguishing malignant and non-malignant lesions. Kodashima et al.[6] developed a computer algorithm using endocytoscopic images to detect malignant lesions, which enabled microscopic discrimination of normal and malignant tissues. In 2015, Shin et al.[7] performed an image analysis using High-resolution Microendoscopy (HRME) to identify squamous dysplasia with a sensitivity of 87% and a specificity of 97%. The same research group[8] reported an improved algorithm to allow real-time analysis in a tablet-interfaced HRME, and it showed reasonable diagnostic accuracy with lower cost compared with previous HRME.

**Intrapapillary capillary loops:** With advances in magnification endoscopy, esophageal IPCLs were identified on the surface of the esophagus and used as a marker of pathological atypia and the invasion depth of early squamous carcinoma. Zhao et al. reported a CAD model for classifying IPCL to identify early squamous cell carcinoma [9]. A total of 1,383 images were analyzed using magnification NBI, and this system showed good diagnostic accuracy of 89.2% at the lesion and 93% at pixel levels, which was better than that of endoscopists.

**CNN using conventional images:** Endocytoscopy, HRME, and IPCL have been used to differentiate squamous dysplasia from normal tissue. Although these advanced tools showed reasonable results for detecting malignant lesions, their application in clinical practice has not been widespread because of limited availability. To overcome this limitation, Horie et al. [10] recently developed a CNN-based AI algorithm using conventional endoscopic images such as WLE and NBI images. They used 8,428 conventional endoscopic images of esophageal cancer (SCC or adenocarcinoma) to detect esophageal cancer, and the CNN system achieved a sensitivity of 98% in distinguishing superficial from advanced cancer.

**Depth of esophageal cancer:** In early esophageal cancer, accurate es-

---

**Table 1: Endoscopic technologies applied to computer-aided diagnosis.**

<table>
<thead>
<tr>
<th>Conventional technologies</th>
<th>Advanced technologies</th>
<th>Endocytoscopic system</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electronic chromoendoscopy: NBI</td>
<td>Endoscopy: VLE, HRME</td>
<td></td>
</tr>
<tr>
<td>Electronic chromoendoscopy: i-SCAN, FICE (BLI, LCI)</td>
<td>Endocytoscopy</td>
<td></td>
</tr>
<tr>
<td>Endomicroscopy: Ex-vivo</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 2: Clinical studies of AI in esophageal diseases.**

<table>
<thead>
<tr>
<th>Reference</th>
<th>Year</th>
<th>Disease</th>
<th>Design</th>
<th>Modality</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Van der Sommen et al. [1]</td>
<td>2016</td>
<td>Barret esophagus</td>
<td>Retrospective</td>
<td>WLE</td>
<td>Sensitivity 83% / Specificity 83%</td>
</tr>
<tr>
<td>Swager et al. [2]</td>
<td>2017</td>
<td>Barret esophagus</td>
<td>Retrospective</td>
<td>VLE</td>
<td>AUC 0.95</td>
</tr>
<tr>
<td>Mendel et al. [3]</td>
<td>2017</td>
<td>Barret esophagus</td>
<td>Retrospective</td>
<td>WLE</td>
<td>Sensitivity 94% / Specificity 88%</td>
</tr>
<tr>
<td>Ebigbo et al. [4]</td>
<td>2019</td>
<td>Barret esophagus</td>
<td>Retrospective</td>
<td>WLE/NBI</td>
<td>Accuracy: 89.9%</td>
</tr>
<tr>
<td>Hashimoto et al. [5]</td>
<td>2020</td>
<td>Barret esophagus</td>
<td>Retrospective</td>
<td>WLE/NBI</td>
<td>Accuracy: 95%</td>
</tr>
<tr>
<td>Kodashima et al. [6]</td>
<td>2007</td>
<td>SCC</td>
<td>Prospective Ex-vivo</td>
<td>Endocytoscopy</td>
<td>the mean ratio of total nuclei: 6.4%(normal) vs. 25%(malignant samples)</td>
</tr>
</tbody>
</table>
timation of the invasion depth is important for determining the treatment strategy. Several studies have used CNNs to assess the invasion depth of esophageal cancer [11-13]. Nakagawa et al. [11] reported the first study of AI for evaluating the invasion depth of SCC and showed favorable results in patients with early esophageal squamous cell carcinoma. In 2020, the CNN-based diagnostic system developed by Tokai et al. detected early SCC with an accuracy of 95%, and estimated the invasion depth with an accuracy of 81% [12]. In the same year, Shimamoto et al. analyzed WLE, NBI, and BLI images of early esophageal SCC to assess the invasion depth of SCC using AI, which showed that most diagnostic parameters were higher than expert endoscopists [13].

**Stomach diseases**

**Gastric cancer:** Early gastric cancer is difficult to detect because it often presents as a flat or slightly elevated or depressed lesion. There have been several studies on the application of AI in detecting and characterizing early gastric cancer (Table 3).

**Detection of gastric cancer:** Hirasawa et al. [14] developed a deep learning-based CNN system with 13,584 high-resolution WLE, NBI, and indigo carmine-stained images of gastric cancer. On a validation set of 2,296 images, the sensitivity was 92.2%. Although this system achieved a high detection rate (overall: 92.2%, diameter >6 mm: 98.6%), the positive predictive value was only 30.6%, showing that non-cancerous lesions were misdiagnosed as cancer. The same research group applied their CNN system to endoscopic video images and showed a diagnosis rate of 94.1% [15]. In a further study, Kanesaka et al. reported that the CNN system showed excellent results with an accuracy of 96% for identifying gastric cancer on magnified NBI images [16].

A recent study developed a CNN-based system for the analysis of gastric mucosal lesions using 341 magnifying NBI images (171 noncancerous lesions and 170 early gastric cancers), and the sensitivity, specificity, and accuracy in the detection of early gastric cancer were 91%, 90%, and 90%, respectively [17]. The sensitivity of the CNN system was significantly higher than that of expert endoscopists.

There are few studies on the efficacy of AI in the classification of gastric

<table>
<thead>
<tr>
<th>Reference</th>
<th>Year</th>
<th>Disease</th>
<th>Design</th>
<th>Modality</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hirasawa et al.[14]</td>
<td>2018</td>
<td>Gastric cancer</td>
<td>Retrospective</td>
<td>WLE/ NBI/chromoendoscopy</td>
<td>Sensitivity: 92.2%, 98.6%(diameter&gt;6 mm)</td>
</tr>
<tr>
<td>Ishioka et al.[15]</td>
<td>2019</td>
<td>Gastric cancer</td>
<td>Prospective</td>
<td>WLE</td>
<td>Accuracy: 94.1%</td>
</tr>
<tr>
<td>Kanesaka et al.[16]</td>
<td>2018</td>
<td>EGC</td>
<td>Retrospective</td>
<td>ME-NBI</td>
<td>Accuracy: 96.3%</td>
</tr>
<tr>
<td>Li et al.[17]</td>
<td>2020</td>
<td>EGC</td>
<td>Retrospective</td>
<td>ME-NBI</td>
<td>Accuracy: 90.9%</td>
</tr>
<tr>
<td>Cho et al.[18]</td>
<td>2019</td>
<td>Gastric neoplasm</td>
<td>Retrospective/Prospective</td>
<td>WLE</td>
<td>Accuracy: 84.6%</td>
</tr>
<tr>
<td>Kubota et al.[19]</td>
<td>2012</td>
<td>Invasion depth of gastric cancer</td>
<td>Retrospective</td>
<td>WLE</td>
<td>Accuracy: 64.7%</td>
</tr>
<tr>
<td>Zhu et al.[20]</td>
<td>2019</td>
<td>Invasion depth of gastric cancer</td>
<td>Retrospective</td>
<td>WLE</td>
<td>Accuracy: 89.1%</td>
</tr>
<tr>
<td>Shichijo et al.[21]</td>
<td>2017</td>
<td>H.pylori infection</td>
<td>Retrospective</td>
<td>WLE</td>
<td>Accuracy: 87.7%</td>
</tr>
<tr>
<td>Itoh et al.[22]</td>
<td>2018</td>
<td>H.pylori infection</td>
<td>Prospective</td>
<td>WLE</td>
<td>AUROC: 0.956; Sensitivity: 86.7%; Specificity: 86.7%</td>
</tr>
<tr>
<td>Nakashima et al.[23]</td>
<td>2018</td>
<td>H.pylori infection</td>
<td>Prospective</td>
<td>Image enhanced endoscopy (BLI/LCI)</td>
<td>AUROC: 0.96 (BLI) AUROC: 0.95 (LCI)</td>
</tr>
<tr>
<td>Yasuda et al.[24]</td>
<td>2020</td>
<td>H.pylori infection</td>
<td>Retrospective</td>
<td>Image enhanced endoscopy (LCI)</td>
<td>Accuracy: 87.6%</td>
</tr>
</tbody>
</table>

H. pylori can cause gastric mucosal atrophy and intestinal metaplasia, which increases the risk of gastric cancer. Gastroscopy is useful to detect H. pylori infection, and endoscopic findings include mucosal atrophy or swelling, rugal hyperplasia, nodularity, and spotty redness. However, diagnostic accuracy is not high with endoscopy, and there have been several attempts to apply AI to evaluate the invasion depth of gastric cancer.

Kubota et al. [19] developed a CAD system for assessing the invasion depth of gastric cancer. They achieved an overall diagnostic accuracy of 64% and accuracy of 77%, 49%, 51%, and 55% in T1, T2, T3, and T4 staging, respectively. In another study [20], Zhu et al. applied a CNN to evaluate the invasion depth of gastric cancer on conventional WLEs, which achieved an overall accuracy of 89%, sensitivity of 76%, and specificity of 95%. This model distinguished early gastric cancer from cancers with deeper submucosal invasion, which could minimize the over-diagnosis of invasion depth and reduce unnecessary surgery for mucosal and superficial submucosal cancer.

Helicobacter pylori infection: H. pylori can cause gastric mucosal atrophy and intestinal metaplasia, which increases the risk of gastric cancer. Gastroscopy is useful to detect H. pylori infection, and endoscopic findings include mucosal atrophy or swelling, rugal hyperplasia, nodularity, and spotty redness. However, diagnostic accuracy is not high with endoscopy, and there have been several attempts to apply AI in the diagnosis of H. pylori infection.

Shichijo et al. applied deep learning to detect H. pylori infection, and this retrospective study showed a reasonable accuracy with a sensitivity and specificity above 87% [21]. In another study, researchers developed a CNN-algorithm based on WLEs for the diagnosis of H. pylori infection, which reached a sensitivity and specificity above 86% [22].

In an attempt to improve the accuracy of H. pylori detection, Nakashima et al. applied AI using a new image-enhanced endoscopy, including Blue-Laser Imaging (BLI) and Linked Color Imaging (LCI). The results were excellent, with the AUCs of 0.96 for BLI and 0.95 for LCI [23]. Yasuda et al. developed a machine-learning based algorithm for the detection of H. pylori using LCI images, achieving significant results with an accuracy of 88%, a sensitivity of 90%, and a specificity of 85% [24].

Gastrointestinal stromal tumor: GIST is a common type of gastric SEL. EUS has been widely used for evaluating the characteristics of SEL, and it remains difficult to distinguish GIST from other diseases. Recently, the application of AI diagnostic systems to detect SELs has been attempted. In 2020, Minoda et al. [25] reported that CNN-based AI had a reasonable diagnostic yield for SELs ≥ 2 cm with an accuracy of 90%, sensitivity of 91%, and specificity of 83%, although it had a low diagnostic yield for SELs <2 cm. In the same year, a CNN-based CAD system developed by Kim et al. also showed a high accuracy for differentiating GISTs from other mesenchymal tumors using EUS images [26]. They achieved a sensitivity of 83%, specificity of 75% and accuracy of 79%, which were all significantly higher than those of endoscopists.

Limitations of Present Studies

Although many studies have shown the usefulness of AI, there are several limitations. First, the data analyzed in most studies were retrospective endoscopic images, and most of the data were obtained from single centers. Unfortunately, many hospitals do not yet have access to most image-enhanced endoscopic images using advanced modalities.

FUTURE APPLICATIONS

It is important to appropriately apply rapidly developing AI technology in the medical field. Technology development is only the first step in medical applications. Researchers must determine the optimal field of application, verify the effectiveness of the technology, and obtain the necessary approvals for widespread use. The value of the technology must be assessed, and its performance must be monitored and improved. In addition, there are limitations to applying AI in clinical practice, including various social and ethical issues, such as patient consent for use of medical data for the development of AI technology, and liability in medical accidents of actions applied with AI technology. Considering the current state of AI development and the surrounding social issues, AI should serve to assist rather than replace the doctor’s diagnostic ability. In the near future, AI is expected to expand to the field of disease treatment.

CONCLUSION

AI has been widely applied in the field of upper GI endoscopy and shows excellent accuracy for the diagnosis of upper GI cancer. Furthermore, AI is useful to assess cancer invasion depth, detect H. pylori infection, and diagnose the characteristics of SEL. In the near future, AI is expected to expand to the field of disease treatment.

REFERENCES


