Chinese Traditional Medical Journal

Machine learning for the detection of precancerous lesions and esophageal squamous cell carcinoma by endoscopic imaging

S. Venkatesh, A. Meghana

[1,2] Academy of Acupuncture and Moxibustion, Fujian University of Traditional Chinese Medicine, Fuzhou, Fujian 350122, China

Received on: 21 Jan 2025 Revised on: 20 Mar 2025 Accepted Date: 25 April 2025
Published on: 18 May 2025

Abstract: Improving patient outcomes requires early identification, rapid diagnosis, and fast treatment of esophageal squamous cell carcinoma (ESCC), a major worldwide health concern. When it comes to this, endoscopic inspection is crucial. There are a number of endoscopic procedures available, but all have limitations that may lead to overlooked or misdiagnosed ESCCs. The use of artificial intelligence (AI) in endoscopic diagnostics is currently enhancing the accuracy of detecting ESCC and precancerous lesions, therefore overcoming these constraints. We survey the present status of artificial intelligence (AI) applications for endoscopic detection of endometrial squamous cell carcinoma (ESCC) and precancerous lesions in this review, including topics such as microvascular subtype categorization, invasion depth estimate, lesion characterisation, and margin delineation. In addition, we provide predictions for the future of this area, pointing out possible developments that can improve patient prognoses via more precise diagnosis.

Keywords: Terminology: precancerous lesions; endoscopy; artificial intelligence; esophageal squamous cell carcinoma

Introduction

Rapid improvements in endoscopic procedures and the widespread adoption of screening programs have contributed to a dramatic increase in the number of people screened for esophageal cancer (EC) in recent years. Nevertheless, EC continues to be a significant worldwide health concern, since it is the third most common cancer and the sixth most common cancer overall, according to recent cancer data. [1] Both esophageal squamous cell carcinoma (ESCC) and esophageal adenocarcinoma are subtypes of esophageal cancer (EC). In Asia and several regions of sub-Saharan Africa, ESCC is still the most common kind of EC, even though its incidence is on the decline. [2] in The clinical stage upon diagnosis has a significant impact on the prognosis of ESCC patients. Even with therapies like chemotherapy and surgery, the overall 5-year survival rate is less than 30% since most patients are identified at an advanced stage. three to five] On the flip side, with the right therapy, patients with early-stage ESCC have an incredible 90% 5-year survival rate. [6] Patients with ESCC have a better prognosis when the disease is detected early, diagnosed quickly, and treated quickly. In order to detect and cure ESCC early, endoscopic inspection and therapy are crucial. White light imaging (WLI), chromoendoscopy (e.g., iodine staining), magnification endoscopy (ME), electronic staining endoscopy (e.g., narrow band imaging, BLI), and blue laser imaging are some of the endoscopic examination methods that are available. To begin a standard endoscopic esophageal examination, a physician will first look for abnormalities using wide-field imaging (WLI) or non-magnifying endoscopy with narrow band imaging (NBI/BLI). [7] Afterwards, these lesions are helped to be characterized by using iodine staining or ME-NBI/BLI. [7] Missed diagnoses of ESCC continue to occur due to specific limitations in the endoscopic technologies that are now available. Some of these limitations include the following: the influence of visual tiredness, variances in endoscopists' skill, and sub-tle microendoscopic aspects of early lesions. From 4.2% to 17.0% of ESCC diagnoses go unnoticed, according to the research. Chapters 8–13 To minimize the number of ESCC diagnoses that go unnoticed, it is critical to overcome these obstacles and make the most of limited medical resources.

Artificial intelligence (AI)-assisted diagnostic techniques based on deep learning offer a promising solution to the aforementioned challenges. The development of AI models aimed at assisting endoscopists in achieving more effective and accurate disease diagnosis has emerged as a prominent area of research. Notably, research in the realm of early diagnosis of ESCC has made substantial advancements, presenting encouraging prospects [Figure 1]. In this review, we provide an overview of the current state of utilizing AI for the endoscopic diagnosis of ESCC and precancerous lesions, as well as explore potential directions for its future development.

Deep Learning-Assisted Detection of Lesions

Single modality

WLI modality

When it comes to lesion identification, the WLI modality is crucial, as it is the fundamental testing approach. After a long period of study, scientists have created AI models using WLI and deep learning techniques to improve diagnostic skills for early identification of ESCC and precancerous lesions [Table 1]. To identify superficial ESCC and precancerous lesions, Cai et al.[14] were the first to create an AI model based on WLI. A dataset consisting of 24,282 endoscopic pictures of the esophagus was used. After that, the model was tested with 187 photos, and it showed great diagnostic performance with a sensitivity of 97.8% and a specificity of 85.4%. That is to say, this model greatly enhanced the precision of sixteen endoscopists' diagnoses, particularly those of their more junior colleagues. This research demonstrates that artificial intelligence has great promise as an ancillary diagnostic tool for endoscopists. An artificial intelligence model for the detection of superficial ESCC using WLI was developed by Feng et al.[15] using the ResNet50 network. For training, the model was fed 5892 photos; for validation, it was fed 1466 images from inside the company and 3063 images from outside sources. When compared to professional endoscopists and non-expert endoscopists, the model's diagnostic performance was excellent. Also, endoscopists' accuracy increased substantially when AI was included (75.12% vs. 84.95%, P = 0.008). By modifying the YOLO (You Only Look Once) v3 algorithm, Tang et al.[16] created an AI model that relies on WLI. A dataset with 4002 photos was used to train this model. It was then validated on three other datasets: one with 333 images, one with 700 images from an external source, and a video test dataset with fourteen movies. Accurate lesion identification in thirteen movies and adequate diagnostic performance on the image test set proved the model's resilience and generalizability, paving the way for its possible use in real-time clinical settings.

NBI modality

Non-ME NBI modality is a common technique in endoscopic examinations. The characteristic teal color tone serves as a hallmark for identifying ESCC and precancerous lesions under NBI. Li *et al*[17] undertook the development of an NBI-AI model for detecting superficial ESCC and precancerous lesions utilizing the fully convolutional network (FCN) based on the VGGNet framework.

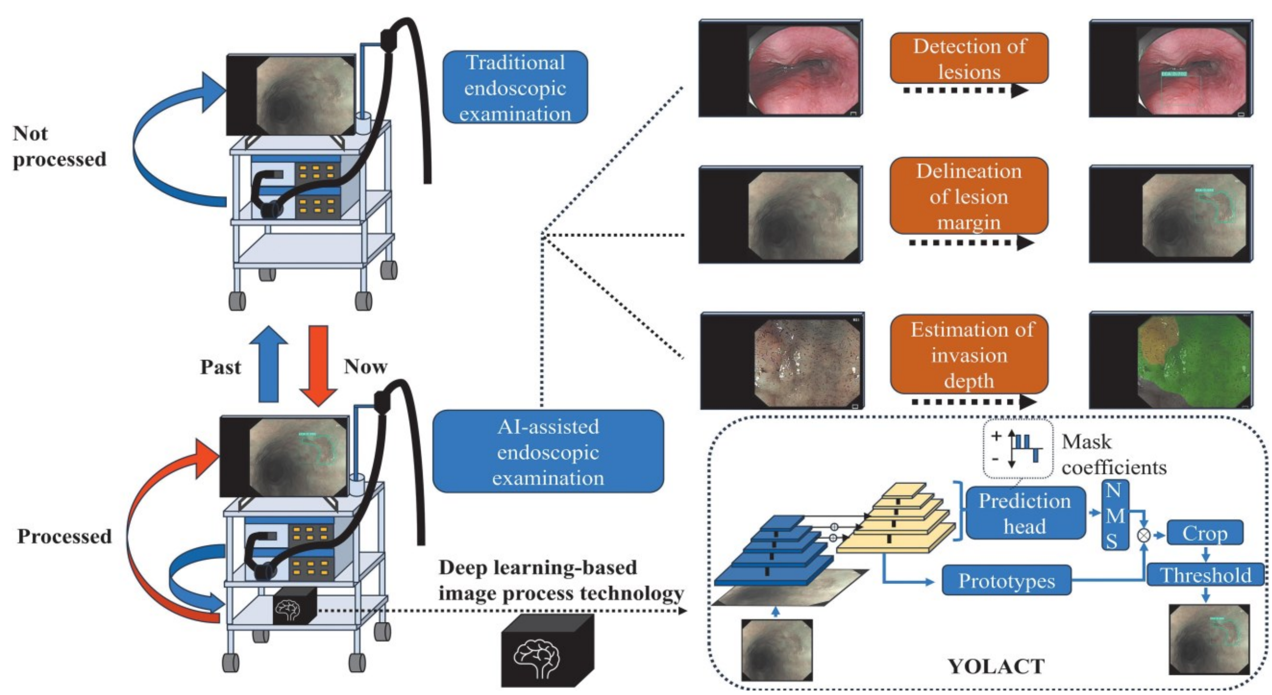

Figure 1: AI-assisted endoscopy for diagnosing ESCC and precancerous lesions, including detection of lesions, delineation of lesion margin, estimation of invasion depth, and classification of the microvessel patterns. AI: Artificial intelligence; ESCC: Esophageal squamous cell carcinoma; NMS: Nonmaximum suppression; YOLACT: You Only Look At CoefficienTs.

Table 1: Deep learning-assisted detection of ESCC and precancerous lesions.

							Outcomes			Compared to endoscopists
Ref.	Year	Aim	Endoscopic modality	Algorithm	Training dataset	Test dataset	Accuracy	Sensitivity	Specificity	
Cai *et al*[14]	2019	Detection of superficial ESCC and precancerous lesions	WLI	DNN	2428 images	187 images	91.40%	97.80%	85.40%	Superior to endoscopists
Kumagai *et al*[18]	2019	Diagnosis of ESCC	Endocytoscopy	GoogleNet	4715 images	1520 images	90.00%	92.60%	89.30%	NA
Horie *et al*[20]	2019	Detection of EC	WLI and NBI	SSD	8428 images	1118 images	99% (ESCC)	97% (ESCC, per-person)	79% (overall)	NA
Ohmori *et al*[24]	2020	Detection of superficial ESCC	WLI, NBI, BLI, and ME	SSD	22,562 images	727 images	81% (WLI), 77% (non-ME NBI/BLI), 77% (ME-NBI/BLI)	90% (WLI), 100% (non-ME NBI/BLI), 98% (ME-NBI/BLI)	76% (WLI) 63% (non-ME-NBI/BLI), 56% (ME-NBI/BLI)	Comparable to experienced endoscopists
Fukuda *et al*[25]	2020	Identification and characterization of superficial ESCC	NBI, BLI, and ME	SSD	28,333 images	144 specimens	63% (detection), 88% (characterization)	91% (detection), 86% (characterization)	51% (detection), 89% (characterization)	Inferior to expert (detection) Superior to expert (characterization)
Tang *et al*[16]	2021	Identification of superficial ESCC	WLI	YOLO v3	4002 images	1033 images and 14 videos	91.3% (internal validation), 90.1% (external validation 1), 91.8% (external validation 2)	97.9% (internal validation), 96.8% (external validation 1), 100% (external validation 2)	88.6% (internal validation), 86.8% (external validation 1), 88.9% (external validation 2)	Superior to endoscopists
Li *et al*[17]	2021	Detection of superficial ESCC and precancerous lesions	NBI	FCN network	4735 images	632 images	94.30%	91.00%	96.70%	Comparable to experienced endoscopists
Ikenoyama *et al*[23]	2021	Detection of multiple Lugol-voiding lesions	WLI and NBI	GoogleNet	6634 images	667 images	76.40%	84.40%	70.00%	Comparable to experienced endoscopists
Shiroma *et al*[29]	2021	Detection of superficial ESCC	WLI and NBI	SSD	8428 images	144 videos	52.5% (WLI), 67.5% (NBI)	75% (WLI), 55% (NBI)	30% (WLI), 80% (NBI)	Superior to endoscopists
Waki *et al*[30]	2021	Identification of superficial ESCC	WLI, NBI, and BLI	BiSeNet	18,797 images	100 videos	65.5% (NBI/BLI), 55.8% (WLI)	85.7% (NBI/BLI), 77.8% (WLI)	40% (NBI/BLI), 28% (WLI)	Superior to endoscopists
Yang *et al*[34]	2021	Identification of ESCC and high-grade EP neoplasia	WLI, BLI, iodine staining, and ME	YOLO v3 and ResNet V2	10,988 images	2309 images and 104 videos	99.2% (per-patient, non-ME, early ESCC), 88.1% (ME, early ESCC)	97.4% (per-patient, non-ME, early ESCC), 90.9% (ME, early ESCC)	99.4% (per-patient, non-ME, early ESCC), 85.0% (ME, early ESCC)	Comparable to expert; superior to non-expert
Kumagai *et al*[19]	2022	Classification of esophageal lesions (ESCC and other non-malignant lesions)	endocytoscopy	DeiT	7983 images	114 images	91.2% (per-image), 94.7% (per-patient)	84.8% (per-image), 90.9% (per-patient)	93.9% (per-image), 96.3% (per-patient)	Superior to endoscopists
Meng *et al*[21]	2022	Detection of superficial ESCC and high-grade EP neoplasia	WLI and NBI	YOLO v5	4447 images	1695 images	92.90%	91.90%	94.70%	Comparable to expert; superior to non-expert
Tajiri *et al*[26]	2022	Classification of esophageal lesion	WLI, NBI, BLI, and ME	ResNet-101 × 1	29,794 images	147 videos	80.90%	85.50%	75.00%	Superior to endoscopists

(*continued*)

Table 1

(Continued)

Ref.	Year	Aim	Endoscopic modality	Algorithm	Training dataset	Test dataset	Outcomes			Compared to endoscopists
							Accuracy	Sensitivity	Specificity	
Aoyama *et al*[28]	2022	Detection of superficial ESCC	NBI	NA	200 cases	97 cases	49.50%	87.50%	22.80%	NA
Yuan *et al*[35]	2022	Detection of superficial ESCC and precancerous lesions	WLI, NBI, iodine staining, and ME-NBI	YOLO v3	45,770 images	8163 images and 142 videos	91.1% (internal validation), 91.2% (external validation), 90.8% (video validation)	96.9% (internal validation), 97.2% (external validation), 97.4% (video validation)	83.1% (internal validation), 82.3% (external validation), 83.3% (video validation)	Comparable to experienced endoscopists
Feng *et al*[15]	2023	Identification of superficial ESCC	WLI	ResNet50	5892 images	4529 images	90.60% (internal validation); 90.39% (external validation)	99.23% (internal validation), 91.60% (external validation)	84.66% (internal validation), 88.89% (external validation)	Comparable to experts; superior to non-experts
Chou *et al*[22]	2023	Identification of ESCC and high-grade dysplasia	WLI and NBI	EfficientNet and Vision Transformer networks	1002 images		accuracy: 96.32%, precision: 96.44%			NA
Tani *et al*[27]	2023	Diagnosis of superficial ESCC	Non-ME and ME	ResNet-101 × 1	29,794 images	237 lesions	80.60%	68.20%	83.40%	Non-inferiority not proven

BLI: Blue laser imaging; DNN: Deep neural network; EC: Esophageal cancer; ESCC: Esophageal squamous cell carcinoma; FCN: Fully convolutional network; ME: Magnifying endoscopy; NA: Not available; NBI: Narrow band imaging; Ref.: Reference; SSD: Single Shot MultiBox Detector; WLI: White light imaging; YOLO: You Only Look Once.

A total of 4735 images were amassed for model training. The performance of this NBI-AI model was evaluated and compared with the WLI-AI model introduced by Cai *et al*[14] using a dataset of 316 paired WLI and NBI images. The comparative analysis revealed that the NBI-AI model exhibited higher accuracy and specificity than its WLI-AI counterpart, albeit with a slightly lower sensitivity. Intriguingly, when evaluating 20 endoscopists with varying levels of expertise, the use of both the WLI-AI model and the NBI-AI model significantly improved their diagnostic performance, surpassing the outcomes achieved with either model in isolation.

Endocytoscopy

The endocytoscopic system is an advanced magnifying endoscopic examination technique with ultra-high resolution capabilities, providing magnification of up to 1000 times, effectively simulating the effects of histological biopsy. By using *in vivo* staining agents such as methylene blue, the endocytoscopic approach allows real-time observation of the cellular structure within the superficial mucosal layer of the esophagus, facilitating dynamic lesion classification. Kumagai *et al*[18] harnessed a dataset consisting of 4715 esophageal endocytoscopic images for model training and an additional set of 1520 endocytoscopic images for model testing. Their AI model, designed for diagnosing ESCC and developed based on the GoogleNet architecture, demonstrated commendable performance metrics, with a sensitivity, specificity, and accuracy of 92.6%, 89.3%, and 90.0%, respectively. Moreover, Kumagai *et al*[19] introduced an innovative AI model founded on the DeiT framework. This model underwent training using an extensive dataset comprising 7983 endocytoscopy images from 55 cases of ESCC and 81 cases of non-malignant lesions. Subsequent testing involved 114 images from 11 ESCC cases and 27 non-malignant lesions. To assess its diagnostic efficacy, the AI model's performance was benchmarked against that of an expert pathologist and non-expert endoscopists. In per-image analysis, the AI model achieved an overall accuracy equal to that of the pathologist (91.2% *vs.* 91.2%) and outperformed the two endoscopists (91.2% *vs.* 85.9% and 91.2% *vs.* 83.3%, respectively). In per-patient analysis, the AI model exhibited higher overall accuracy compared to both the pathologist and the two endoscopists, with accuracies of 94.7%, 92.1%, 86.8%, and 89.5%, respectively. These findings highlight the potential of AI as a valuable tool in assisting endoscopists, particularly those without specialized pathological expertise, in the non-invasive diagnosis of ESCC and precancerous lesions through endocytoscopy, obviating the need for invasive histological biopsies.

Bimodality or multimodality

In clinical practice, endoscopic examinations often entail the use of multiple endoscopic modalities. Therefore, the integration of two or more modalities into bimodal or multimodal AI models holds greater clinical application value compared to AI models based on a single modality.

Horie *et al*[20] firstly introduced a bimodal AI model that integrated WLI and NBI employing the Single Shot MultiBox Detector (SSD) for diagnosing ESCC and esophageal adenocarcinoma. This model was trained on 8428 images from 397 ESCC lesions and 32 EAC lesions. The AI model displayed exceptional performance by detecting 98% of ESCC lesions within a mere 27 s. The sensitivities of the AI model for diagnosing ESCC under WLI and NBI were 79% and 89%, respectively. Meng *et al*[21] developed an AI model utilizing the YOLO v5 algorithm, training it on 4447 images containing both WLI and NBI images. The model was then tested on 1695 images. Remarkably, the accuracy of the model in the diagnosis of superficial ESCC and high-grade intraepithelial (EP) neoplasia was comparable to that of expert endoscopists (92.9% *vs.* 91.0%, P = 0.129) and significantly higher than that of non-expert endoscopists (92.9% *vs.* 78.3%, P <0.001). Moreover, it contributed to a significant enhancement in diagnostic accuracy among non-expert endoscopists (88.2% *vs.* 78.3%, P <0.001). Chou *et al*[22] constructed an AI model based on EfficientNet and Vision Transformer networks for the identification of ESCC and high-grade dysplasia. The training dataset consisted of 650 WLI images and 352 NBI images. The model achieved an accuracy and precision of 96.32% and 96.44%, respectively. Ikenoyama *et al*[23] developed an AI model based on GoogleNet with the unique ability to detect multiple Lugol-voiding lesions without the need for iodine staining. Training this model entailed a dataset of 6634 WLI and NBI images, followed by testing on 667 images. Intriguingly, the model demonstrated significantly higher sensitivity compared to endoscopists (84.4% *vs.* 46.9%). This innovation has the potential to reduce patient discomfort associated with iodine staining while effectively identifying high-risk ESCC patients. Ohmori *et al*[24] introduced a groundbreaking multimodal AI model that combined WLI, NBI, BLI, and ME modalities for the detection of the superficial ESCC. This model was trained on an extensive dataset comprising 22,562 images and underwent rigorous testing on 255 WLI images, 268 non-ME NBI/BLI images, and 204 ME NBI/BLI images. It exhibited commendable sensitivity and specificity under various modalities: 90% and 76% under WLI, 100% and 63% under non-ME NBI/BLI, and 98% and 56% under ME-NBI/BLI, respectively. Importantly, this multimodal AI model's performance was on par with that of 15 experienced endoscopists, underscoring the potential for AI models combined with multiple modalities to more accurately replicate real-world clinical scenarios compared to a single WLI approach [Table 1].

Real-time detection

The advancements achieved in converting AI models to multi-mode operation are praiseworthy. Keep in mind that these AI models have mostly been tested and trained on high-quality static photos, thus their results in actual clinical settings may not be completely accurate. As a result, dynamic movies that include complicated environmental aspects should be a part of AI model development and evaluation. The results will be more precise and

useful understanding of these models' capacities in a therapeutic setting. Fukuda et al.[25] were the first to use 28,333 non-ME NBI/BLI and ME-NBI/BLI pictures retrieved from 862 movies to build an artificial intelligence model based on SSD for the detection and characterization of superficial ESCC. Two modes of operation were included into the AI model: a non-ME mode for lesion detection and a ME mode for lesion characterisation. The model was evaluated using 5- to 9-second video clips obtained from 144 patients diagnosed with superficial ESCC. The outcomes were then compared to those of thirteen endoscopists. Notably, when compared to the experts, the AI model showed far higher sensitivity and specificity in lesion characterisation (86.7% vs. 74.4%, 89.8% vs. 76.5%), and much higher sensitivity in lesion identification (91% vs. 79%). In order to distinguish between ESCC and other benign esophageal tumors, Tajiri et al.[26] built a tailored AI model. An extensive dataset was used to train the model. It included 25,048 photos from 1433 instances of superficial ESCC and 4,746 images from 410 non-cancerous lesions recorded using ME, WLI, NBI, and BLI. The following testing was carried out using 147 NBI videos, and the results were quite impressive: sensitivity of 85.5%, specificity of 75.0%, and accuracy of 80.9%. The AI model developed by Tajiri et al. was used by Tani et al. [27]. [26]

The training dataset comprised of 29,794 images acquired under non-ME and ME modalities. In the validation phase, the AI model and five endoscopists analyzed 237 lesions in 380 patients. The AI model demonstrated an accuracy of 80.6%, sensitivity of 68.2%, and specificity of 83.4%, while the endos- copists achieved figures of 85.7%, 61.4%, and 91.2%, respectively. It is noteworthy that the AI model did not establish non-inferiority when compared to endoscopists in real-time ESCC diagnosis, potentially due to factors such as compressed images in the validation dataset or an imbalanced proportion of ESCC cases among the target lesions. Aoyama *et al*[28] harnessed endoscopic videos to develop an AI model for the detection of superficial ESCC. The model was trained on NBI images sourced from endoscopic examination videos of 200 patients under NBI, which were dissected into individual frames. The test dataset consisted of images from 97 patients. The model achieved sensitivity, specificity, and accuracy figures of 87.5%, 22.8%, and 49.5%, respectively. Shiroma *et al*[29] introduced a bimodal AI model based on SSD for real-time detection of superficial ESCC. The training dataset contained 8428 WLI and NBI images, while the testing dataset featured 64 slow-speed videos (each lasting 5–15 s) and 80 high-speed videos at 2 cm/s. The model processed 30 frames of images per second and accurately identified all superficial ESCC lesions in slow-speed videos and 85% of lesions in high-speed videos. Furthermore, the introduction of AI led to enhanced detection sensitivity in 13 out of 18 endoscopists. Waki *et al*[30] developed a multimodal AI model designed for identifying superficial ESCC. This AI model was trained using 18,797 WLI and NBI/BLI images and was subsequently tested with 100 video clips simulating scenarios where lesions had been overlooked. Each video clip ranged from 20 s to 30 s in duration. The AI model exhibited sensitivities of 85.7%

under NBI/BLI and 77.8% under WLI, surpassing the performance of 21 endoscopists, whose sensitivities stood at 75% and 56.1%, respectively. Additionally, the AI-assisted sensitivity of endoscopists in detecting lesions under NBI/BLI improved by 2.7%.

The endoscopic method of iodine staining is used for the screening of ESCC and precancerous lesions. Its sensitivity ranges from 92% to 100%. Chapters 31–33 Because malignant or precancerous areas can seem unstained or very barely stained, endoscopists depend on the intensity of staining to identify lesions. Because of this, iodine staining is a powerful modality that AI models may use. A multimodal AI model was introduced by Yang et al. [34] for the purpose of identifying ESCC and high-grade EP neoplasia using WLI, BLI, iodine staining, and ME modalities. There were two main parts to this model. The first part, built on the YOLO v3 framework, was concerned with lesion identification in the non-ME modality, while the second part, based on ResNet V2, was concerned with lesion characterisation in the ME-BLI modality. Testing included 2309 photos and 104 video clips, whereas the model's training used 10,988 images. In terms of diagnostic accuracy, the results showed that the AI model was comparable to two expert endoscopists ($P = 0.205$) and much better than four non-expert endoscopists ($P <0.05$). Also, in films that weren't ME, the AI model could analyze 10 images/second with a latency of under 100 ms. Yuan et al. [35] developed an AI model that uses WLI, non-ME NBI, iodine staining, and ME NBI to identify superficial ESCC and precancerous lesions. The model is based on YOLO v3. The training dataset consists of 45,770 pictures in total. Separate from the external validation dataset, which included 2088 photos sourced from various hospitals, the internal validation dataset had 6075 photographs that were homologous to patients in the training dataset. In addition, the video test set included 142 video clips. The AI model's diagnostic ability in the picture collection was on par with eleven seasoned endoscopists. Most notably, as compared to WLI, the model was much more sensitive in identifying epithelial-bound superficial ESCC (90.8% vs. 82.5%). Crucially, the AI model ran very efficiently, with no latency whatsoever, and a scan time of just 17 ms per picture for diagnosis. The ability of AI models to quickly analyze movies and aid endoscopists in diagnosing lesions in real-life clinical settings is highlighted by this [Table 1]. Artificial intelligence's efficacy in actual healthcare contexts requires assessment. To evaluate the adjunctive diagnostic efficacy of an AI system in identifying superficial ESCC and precancerous lesions in actual clinical settings, Yuan et al.[36] performed the biggest multi-center randomized controlled experiment. In order to aid endoscopists in identifying lesions under WLI and NBI, the AI system was directly incorporated into the endoscopic equipment. This research had 11,846 patients in total. With a per-lesion risk ratio of 0.25 and a per-patient risk ratio of 0.37 and a p-value of 0.40, respectively, the AI-first group had lower miss rates than the routine-first group. We may infer important information about the further diagnostic utility of the AI system in actual healthcare environments.

Deep Learning-Assisted Delineation of Lesion Margin

Endoscopy examination combined with histopathology examination represents the gold standard for diagnosing ESCC and precancerous lesions. Ensuring the biopsy of lesions within the designated area is crucial for accurate diagnosis, preventing potential misses or misdiagnoses. Consequently, the accurate delineation of lesion margins holds paramount importance as it facilitates the selection of appropriate biopsy sites and contributes to the early detection of ESCC. In terms of treatment, compared with traditional surgery, endoscopic resection (ER) offers superior preservation of the physiological function of the digestive tract while effectively resecting lesions. ER has received strong recommendations in guidelines as the first-line treatment for early ESCC. However, for lesions exceeding 3/4 of the esophageal circumference, surgery is recommended given the high incidence of esophageal stenosis (exceeding 60%) following ER.[37] Therefore, achieving accurate preoperative delineation of lesion margins plays a pivotal role in treatment selection and the successful completion of ER.

Accurate detection and delineation of lesion margins can be challenging when using WLI. Liu *et al*[38] constructed an AI model based on YOLACT (You Only Look At CoefficienTs) using 10,467 images capable of accurately detecting and delineating the margins of superficial ESCC and precancerous lesions under WLI. The model's performance was evaluated using 2616 internal validation images and 563 external validation images. The results indicated that the delineation accuracy of the model was comparable to that of seven endoscopic experts (98.1% *vs.* 95.3%) and superior to that of seven experienced endoscopists (98.1% *vs.* 78.6%). Remarkably, this model achieved a rapid diagnostic speed of only 17 ms per image, suggesting its potential for real-time endoscopic diagnosis. NBI is a commonly used clinical modality for determining the boundaries of esophageal lesions. Guo *et al*[39] reported an AI model designed for diagnosing superficial ESCC and precancerous lesions while generating probability heat maps on endoscopic images to indicate lesion margins. The AI model, constructed based on SegNet, underwent training with 6473 images and was subsequently tested on 6671 images and 80 video clips. The results showed that regions of lesions indicated by this model closely aligned with those indicated by endoscopists. Yuan *et al*[40] complied 10,047 still images and 140 videos under NBI to create an AI model capable of detecting and delineating superficial ESCC and precancerous lesions. The AI model was seamlessly integrated into the endoscopy equipment. Comparative assessments were conducted between the model and 11 endoscopists. The model exhibited strong performance, achieving delineation accuracies of 88.9% in internal tests and 87.0% in external tests. Impressively, the model completed the delineation process in a mere 12 ms per image, outpacing both senior and junior endoscopists who required about 21.4 s and 33.6 s per image, respectively. The model's delineation performance was superior to that of junior endoscopists (89.6% *vs.* 74.7%,

P <0.001) and comparable to that of senior endoscopists (89.6% *vs.* 89.6%, P = 1.000). Furthermore, Yuan *et al*[41] were the first to report a multimodal AI model capable of real-time detection and delineation of small flat-type superficial ESCC. The AI model was seamlessly integrated into an endoscopy equipment and successfully detected and delineated a 3-mm flat-type ESCC under WLI, NBI, iodine staining, and ME. This suggests that AI models hold great promise in facilitating the real-time detection and delineation of margins for ESCC and precancerous lesions in clinical practice [Table 2].

Table 2: Deep learning-assisted delineation of margin of ESCC and precancerous lesions.

Ref.	Year	Aim	Endoscopic modality	Algorithm	Training dataset	Validation/test dataset	Outcome			Compared to endoscopists
							Accuracy	Sensitivity	Specificity	
Guo *et al*[39]	2020	Detection and delineation of the margins of superficial ESCC	NBI	SegNet	6473 images	6671 images and 80 videos	95.70%	98.04%	95.03%	NA
Liu *et al*[38]	2022	Detection and delineation of the margins of superficial ESCC	WLI	YOLACT	10,467 images	3179 images	98.1% (delineation)	88.5% (delineation)	85.8% (delineation)	Comparable to experts; superior to experienced endoscopists
Yuan *et al*[40]	2023	Detection and delineation of the margins of superficial ESCC	NBI	YOLACT	7530 images	2517 images and 140 videos	88.9% (delineation, internal test), 87.0% (delineation, external test), 85.9% (delineation, prospective)	80.7% (delineation, internal test), 80.6% (delineation, external test), 79.4% (delineation, prospective)	86.5% (detection, internal test), 84.5% (detection, external test), 84.3% (detection, prospective)	Comparable to senior endoscopists; superior to junior endoscopists
Yuan *et al*[41]	2023	Detection and delineation of the margins of superficial ESCC	WLI, NBI, iodine staining, and ME	YOLACT	NA	NA	Detected and delineated a 3 mm flat-type ESCC			NA

ESCC: Esophageal squamous cell carcinoma; ME: Magnifying endoscopy; NA: Not available; NBI: Narrow band imaging; Ref: Reference; WLI: White light imaging; YOLACT: You Only Look At CoefficienTs.

Deep Learning-Assisted Estimation of Invasion Depth and Classification of the Microvessel Patterns of the Lesions

Estimation of invasion depth of lesions

The invasion depth of ESCC is closely associated with the risk of lymph node metastasis. For precancerous lesions, lesions confined to the epithelium (EP), and lesions limited to the lamina propria (LPM), the risk of lymph node metastasis is minimal. However, the risk significantly increases for lesions extending into the mucosal muscular (MM), lesions involving mucosal muscular (M3), superficial submucosa (SM1), and the submucosa that is 200 mm or deeper (SM2–3), with risk ranges of 8–18%, 11–53%, and 30–54%, respectively.[42–44] Precancerous lesions and EP/LPM lesions are absolute indications for ER, MM-SM1 lesions are relative indications for ER, and SM2–3 lesions and lesions with deeper invasion require surgery or chemoradiotherapy.[37] Hence, the accurate assessment of the invasion depth within a lesion assumes paramount importance in guiding precise treatment decisions.

Wang *et al*[45] reported an AI model based on SSD for the differentiation of squamous low-grade dysplasia, squamous high-grade dysplasia, and squamous cell carcinoma under WLI/NBI. The model was trained using 936 images and tested on 264 images, capable of analyzing the test images within 10 s. It achieved sensitivities of 96.2% for detecting lesions, 83.4% for diagnosing low-grade dysplasia, 87.2% for diagnosing high-grade dysplasia, and 98.9% for diagnosing ESCC. Tokai *et al*[46] constructed an AI model based on GoogleNet specifically for distinguishing between EP-SM1 and SM2–3 lesions under WLI/NBI. Their dataset included 1751 training images and 291 test images. The overall accuracy of the model for estimating invasion depth reached 80.9%. It also demonstrated a sensitivity of 84.1% and a specificity of 73.3% for EP-SM1 lesions, surpassing the performance of 13 endoscopists. Nakagawa *et al*[47] introduced a multimodal AI model based on SSD for discriminating between EP-SM1 and SM2-3 lesions under WLI, NBI, iodine staining, and ME. Their dataset comprised 14,338 training images and 914 test images. The model exhibited rapid analysis capability, processing 914 images in just 29 s. It also demonstrated high sensitivity (90.1%) and specificity (95.8%). Importantly, the model's performance was comparable to that of 16 experienced endoscopists.

The AI models discussed in the mentioned studies predominantly relied on static images, which may not fully capture the dynamic nature of real clinical scenarios. Therefore, it is crucial to include dynamic videos in both the development and validation of AI models. Shimamoto *et al*[48] pioneered the development of an AI model for invasion depth estimation using endoscopic examination video data. They extracted one still image from every 30 frames (equivalent to 30 images per second of video), resulting in a total of 23,977 WLI and NBI/BLI images in the training dataset. To assess its diagnostic performance, 102 video clips were employed to compare the AI model against evaluations by 14 endoscopic experts. Remarkably, the model demonstrated superior sensitivity and specificity in distinguishing EP-SM1 lesions from SM2–3 lesions under both non-ME (50% and 99%) and ME (71% and 95%), outperforming the experts. Zhang *et al*[49] developed an AI model utilizing 14 different algorithms to differentiate SM2-3 lesions under ME-NBI. The AI model was trained with 5119 ME-NBI images from 581 ESCC patients and tested on 196 images and 33 video clips. To assess its performance, the model was compared with a deep learning model and evaluated by seven endoscopists. In the image validation, the AI model exhibited impressive sensitivity, specificity, and accuracy at 85.7%, 86.3%, and 86.2%, respectively. These figures remained notably high in video validation, with sensitivity, specificity, and accuracy values reaching 87.5%, 84.0%, and 84.9%, respectively. In contrast, the deep learning model displayed inferior performance, with a sensitivity of 83.7%, specificity of 52.1%, and accuracy of 60.0%. Notably, the assistance of the AI model led to a statistically significant improvement (P = 0.03) in the accuracy of the endoscopists. These findings suggest that AI models enhanced with real video data hold promise for practical integration into clinical practice [Table 3].

Classification of microvessel patterns of the lesions

Lesion characterization and invasion depth estimation rely heavily on ME evaluation of the esophageal Intrapapillary Capillary Loop (IPCL) shape. Classification of IPCL by the Japan Esophageal Society (JES)[50] is as follows: type A, type B, and subtypes B1, B2, and B3 of type B. In Type A, you may see normal IPCL or microvessels that are slightly aberrant, suggesting things like inflammation, low-grade EP neoplasia, or normal epithelium. Microvessels with a loop-like structure, characteristic of EP-LPM lesions, describe type B1. Microvessels of type B2 do not have a loop-like arrangement, indicating MM-SM1 lesions. Surprisingly, type B3 indicates lesions at SM2 or deeper levels and is characterized by dilated vessels with calibers more than three times larger than normal B2 vessels. Keep in mind that IPCL categorization is subjective and involves a lot of moving parts. Endoscopists need extensive training and knowledge of endoscopy to accurately diagnose IPCL and classify cases accurately.

Endoscopists may now assess lesion invasion depth using deep learning and IPCL categorization. In their study, Everson et al.[51] collected 7046 ME-NBI photos in a row from 17 patients' films who had superficial ESCC. We used this massive dataset with the primary goal of classifying microvessels into type A and type B throughout the AI model's training and validation processes. An impressive 93.7% overall accuracy was shown by the AI model after a rigorous 5-fold cross-validation procedure. Its sensitivity to detect type B vasculature was an amazing 89.3%, and its specificity was 98%. It was also appropriate for application in real-time clinical settings since its operating efficiency was maintained between 26.17 ms and 37.38 ms. Using ResNet-18 and an enormously bigger dataset of 67,742 consecutive ME-NBI pictures from 114 patients' films, Everson et al.[52] further enhanced the AI model. Showing promise as a clinical tool, the diagnostic performance was on par with nine endoscopic specialists. Although the AI models mentioned earlier have proven to be very good at differentiating between type A and type B vessels—laying the groundwork for AI-assisted IPCL diagnosis—they fail miserably when it comes to subclassifying type B vessels, which is essential for making accurate invasion depth predictions. The first artificial intelligence model to categorize microvessels of types A, B1, and B2 was published by Zhao et al. [53]. Built using a double-labeling FCN and trained on 1350 ME-NBI pictures, this model outperformed six mid-level and junior endoscopists with sensitivity levels of 71.5% for type A microvessels, 91.1% for type B1 microvessels, and 83.0% for type B2 microvessels. An artificial intelligence model for localizing and recognizing type B1 and type B2 IPCLs was trained and tested using 2887 ME-BLI pictures and 493 ME-NBI images, as reported by Wang et al. [54]. Using ME-NBI, the model was able to accurately identify type B1 IPCLs with a precision rate of 70.74% and type B2 IPCLs with a precision rate of 57.18%. When trying to estimate how deep an ESCC invasion will be, type B3 microvessels are crucial. Using cropped ME-NBI pictures, Uema et al.[55] trained an AI model based on ResNeXt-101 to distinguish between type B1, type B2, and type B3 microvessels. In under 20.3 seconds, our model was able to diagnose all 747 validation photos from a training dataset of 1,777 images. Compared to the combined performance of eight endoscopists, the sensitivity for identifying type B1, type B2, and type B3 microvessels was 96.4%, 66.8%, and 83.3%, respectively. An AI model that can distinguish between type A, type B1, type B2, and type B3 microvessels was created by Yuan et al. [56]. Using 5505 ME-NBI pictures for training and 1589 for testing, this model was able to obtain sensitivity levels of 94.1% for identifying type A microvessels, 91.6% for type B1 microvessels, 87.7% for type B2 microvessels, and 77.8% for type B3 microvessels. Most importantly, when it came to helping identify IPCL subtypes, the AI model vastly enhanced the diagnostic capabilities of seven rookie endoscopists. In order to help endoscopists properly estimate the invasion depth of ESCC, our results highlight the potential of AI models to discriminate not only kinds of IPCL but also subtypes [Table 3].

Limitations and Future Perspectives

AI models have made significant strides in various domains of ESCC and precancerous lesions. Notably, their diagnostic performance is on par with that of experienced endoscopists. However, current AI models confront limitations and challenges when it comes to

Table 3: Deep learning-assisted estimation of invasion depth and classification of the microvessel patterns of ESCC and precancerous lesions.

Ref.	Year	Aim	Endoscopic modality	Algorithm	Training dataset	Validation/test dataset	Outcome			Compared to endoscopists
							Accuracy	Sensitivity	Specificity	
Nakagawa *et al*[47]	2019	Differentiation of EP-SM1 and SM2-3 lesions	WLI, NBI iodine staining, and ME	SSD	14,338 images	914 images	91.00%	90.10%	95.80%	Comparable to experienced endoscopists
Everson *et al*[51]	2019	Classification of type A and type B microvessels	ME-NBI	CNN	7046 images		93.70%	89.30%	98%	NA
Zhao *et al*[53]	2019	Classification of type A, type B1, and type B2 microvessels	ME-NBI	FCN	1350 images		92.5% (type A), 87.6% (type B1), 93.9% (type B2)	71.5% (type A), 91.1% (type B1), 83.0% (type B2)	96.3% (type A), 79.1% (type B1), 95.7% (type B2)	Superior to endoscopists
Tokai *et al*[46]	2020	Differentiation of EP-SM1 and SM2-3 lesions	WLI and NBI	GoogleNet	1752 images	291 images	80.90%	84.10%	73.30%	Superior to endoscopists
Shimamoto *et al*[48]	2020	Estimation of invasion depth	WLI, NBI, and BLI	SSD	23,977 images	102 videos	87% (non-ME), 89% (ME)	50% (non-ME), 71% (ME)	99% (non-ME), 95% (ME)	Superior to endoscopists
Wang *et al*[45]	2021	Differentiation of squamous low-grade dysplasia, squamous high-grade dysplasia, and squamous cell carcinoma	WLI and BLI	SSD	936 images	264 images	90.9% (detection), 92% (differentiation)	96.2% (detection), 98.9% (differentiation)	70.40%	NA
Everson *et al*[52]	2021	Classification of type A and type B microvessels	ME-NBI	ResNet-18	67,742 images		91.70%	93.70%	92.40%	Comparable to experts
Uema *et al*[55]	2021	Classification of type B1, type B2, and type B3 microvessels	ME-NBI	ResNeXt-101	1777 images	747 images	88.6% (type B1), 84.3% (type B2), 95.4% (type B3)	96.4% (type B1), 66.8% (type B2), 83.3% (type B3)	78.7% (type B1), 95.6% (type B2), 96.1% (type B3)	Superior to endoscopists
Yuan *et al*[56]	2022	Classification of type A, type B1, type B2, and type B3 microvessels	ME-NBI	HRNet + OCR	5505 images	1589 images	93.2% (type A), 91.6% (type B1), 98.3% (type B2), 99.7% (type B3)	94.1% (type A), 91.6% (type B1), 87.7% (type B2), 77.8% (type B3)	93.1% (type A), 91.6% (type B1), 99.3% (type B2), 100% (type B3)	Superior to endoscopists
Zhang *et al*[49]	2023	Differentiation SM2-3 lesions	ME-NBI	14 models	5119 images	196 images and 33 videos	86.2% (image validation), 84.9% (video validation)	85.7% (image validation), 87.5% (video validation)	86.3% (image validation), 84% (video validation)	Superior to endoscopists
Wang *et al*[54]	2023	Localization and identification of type B1 and type B2 IPCLs	ME-NBI and ME-BLI	PSA-HRNetV2p	3380 images		*precision*: 70.74% (type B1), 57.18% (type B2)			NA

BLI: Blue laser imaging; CNN: Convolutional neural network; EP: Intraepithelial; ESCC: Esophageal squamous cell carcinoma; FCN: Fully convolutional network; IPCL: Intrapapillary Capillary Loop; ME: Magnifying endoscopy; NA: Not available; NBI: Narrow band imaging; Ref: Reference; SM1: Superficial submucosa cancer; SM2-3: Lesions invading the submucosa at a depth of 200 mm or deeper; SSD: Single Shot MultiBox Detector; WLI: White light imaging; HRNet: High-resolution network; OCR: Object Contextual Representation; PSA-HRNetV2p: polarized self-attention-HRNetV2p.

clinical application.

First, there is a paucity of real-world scenario representations in the endoscopic data used to build AI models, which leads to overfitting and reduces their therapeutic value. At the moment, the majority of AI models are educated using high-quality, one-time datasets that only include photographs and videos from a single location. The given references are [14,17,20,22,28]. Images of low quality, abnormal lesions, and the complex environmental details seen in the actual world are not represented in these databases. In addition, most of the current datasets only include endoscopic pictures of non-cancerous lesions or early-stage esophageal squamous cell carcinoma (ESCC), which leaves out a wide variety of esophageal lesion types include inflammation, advanced malignancy, and ulcers. Major medical institutions must combine their resources, create standardized large-scale medical picture archives, and use methods like unsupervised learning to solve these problems. Secondly, AI models still don't do a great job of diagnosing multifunctional diseases, even when they have incorporated clinical endo- scopic modalities. Multifunctional AI models that include tasks like lesion characterisation, margin delineation, invasion depth estimate, and subtype categorization should be prioritized for future developments. Third, AI model performance evaluation in the actual world requires further prospective multicenter clinical studies. [36] The AI models for ESCC need further proof before they can be used in clinical settings. The efficacy of AI application in real-world environments has been confirmed by several prospective research on AI models for colorectal polyps. pages 57 to 63 Class III medical device permits have been approved by the Chinese National Medical Products Administration for products such as EndoScreener (Wision AI, Shanghai, China). Finally, the importance of AI models as supplementary diagnostic tools must be acknowledged. The next generation of doctors and nurses might be held back by an over-reliance on AI. There are a number of important factors to think about, including the safety of patients' personal health information, the preservation of intellectual property, ethical concerns, and culpability for incorrect decisions made using AI models. To tackle these intricate issues, appropriate rules and regulations must be drafted.

In summary, AI has made great strides in several areas, such as estimating invasion depth, defining margins, and classifying ESCC and precancerous lesion subtypes. We expect the domain of endometrial squamous cell carcinoma (ESCC) and precancerous lesions will see an increase in the use of artificial intelligence (AI) technologies such as deep learning as a result of continuing advances and depth of relevant research across several areas. Both endoscopists and patients will benefit from this extension, which will include aided diagnosis, treatment assistance, and telemedicine. It will contribute to the early diagnosis and treatment of ESCC.

Contributions The research in this paper was made possible by awards from the Chinese National Natural Science Foundation (No. 62406205),

Projects for Artificial Intelligence ZYAI24006, West China Hospital, Sichuan University, and Sichuan Province Cadres Health Research (No. ZH2024-102)

.

References

1. Wu Y, He S, Cao M, Teng Y, Li Q, Tan N, *et al*. Comparative analysis of cancer statistics in China and the United States in 2024. Chin Med J 2024;137:3093–3100. doi: 10.1097/CM9.0000000000003442.
2. Abnet CC, Arnold M, Wei WQ. Epidemiology of esophageal squamous cell carcinoma. Gastroenterology 2018;154:360–373. doi: 10.1053/j.gastro.2017.08.023.
3. Zeng H, Chen W, Zheng R, Zhang S, Ji JS, Zou X, *et al*. Changing cancer survival in China during 2003-15: A pooled analysis of 17 population-based cancer registries. Lancet Glob Health 2018;6:e555–e567. doi: 10.1016/s2214-109x(18)30127-x.
4. Allemani C, Matsuda T, Di Carlo V, Harewood R, Matz M, Nikšić M, *et al*. Global surveillance of trends in cancer survival 2000-14 (CONCORD-3): Analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. Lancet 2018;391:1023–1075. doi: 10.1016/s0140-6736(17)33326-3.
5. Zhang Q, Wang F, Feng H, Xing J, Zhu S, Zhang H, *et al*. Modifiable risk factors for esophageal cancer in endoscopic screening population: A modeling study. Chin Med J 2024;137:350–352. doi: 10.1097/CM9.0000000000002878.
6. Rice TW, Blackstone EH, Rusch VW. 7th edition of the AJCC cancer staging manual: Esophagus and esophagogastric junction. Ann Surg Oncol 2010;17:1721–1724. doi: 10.1245/s10434-010-1024-1.
7. Pimentel-Nunes P, Libânio D, Bastiaansen BAJ, Bhandari P, Bisschops R, Bourke MJ, *et al*. Endoscopic submucosal dissection for superficial gastrointestinal lesions: European Society of Gastrointestinal Endoscopy (ESGE) Guideline—update 2022. Endoscopy 2022;54:591–622. doi: 10.1055/a-1811-7025.
8. Zhang H. A systematic review and meta–analysis of missed squamous cell esophageal carcinoma after esophagogastroduodenoscopy. Gastrointest Endosc 2017;85:AB568. doi: 10.1016/j.gie.2017.03.1309.
9. Chadwick G, Groene O, Hoare J, Hardwick RH, Riley S, Crosby TD, *et al*. A population-based, retrospective, cohort study of esophageal cancer missed at endoscopy. Endoscopy 2014;46:553–560. doi: 10.1055/s-0034-1365646.
10. Gavric A, Hanzel J, Zagar T, Zadnik V, Plut S, Stabuc B. Survival outcomes and rate of missed upper gastrointestinal cancers at routine endoscopy: A single centre retrospective cohort study. Eur J Gastroenterol Hepatol 2020;32:1312–1321. doi: 10.1097/meg.0000000000001863.
11. Januszewicz W, Witczak K, Wieszczy P, Socha M, Turkot MH, Wojciechowska U, *et al*. Prevalence and risk factors of upper gastrointestinal cancers missed during endoscopy: A nationwide registry-based study. Endoscopy 2022;54:653–660. doi: 10.1055/a-1675-4136.
12. Tai FWD, Wray N, Sidhu R, Hopper A, McAlindon M. Factors associated with oesophagogastric cancers missed by gastroscopy: A case-control study. Frontline Gastroenterol 2020;11:194–201. doi: 10.1136/flgastro-2019-101217.
13. Rodríguez de Santiago E, Hernanz N, Marcos-Prieto HM, De-Jorge-Turrión MÁ, Barreiro-Alonso E, Rodríguez-Escaja C, *et al*. Rate of missed oesophageal cancer at routine endoscopy and survival outcomes: A multicentric cohort study. United European Gastroenterol J 2019;7:189–198. doi: 10.1177/2050640618811477.
14. Cai SL, Li B, Tan WM, Niu XJ, Yu HH, Yao LQ, *et al*. Using a deep learning system in endoscopy for screening of early esophageal squamous cell carcinoma (with video). Gastrointest Endosc 2019;90:745–753.e2. doi: 10.1016/j.gie.2019.06.044.
15. Feng Y, Liang Y, Li P, Long Q, Song J, Li M, *et al*. Artificial intelligence assisted detection of superficial esophageal squamous cell carcinoma in white-light endoscopic images by using a generalized

system. Discov Oncol 2023;14:73. doi: 10.1007/s12672-023-00694-3.
16. Tang D, Wang L, Jiang J, Liu Y, Ni M, Fu Y, *et al*. A novel deep learning system for diagnosing early esophageal squamous cell carcinoma: A multicenter diagnostic study. Clin Transl Gastroenterol 2021;12:e00393. doi: 10.14309/ctg.0000000000000393.
17. Li B, Cai SL, Tan WM, Li JC, Yalikong A, Feng XS, *et al*. Comparative study on artificial intelligence systems for detecting early esophageal squamous cell carcinoma between narrow-band and white-light imaging. World J Gastroenterol 2021;27:281–293. doi: 10.3748/wjg.v27.i3.281.
18. Kumagai Y, Takubo K, Kawada K, Aoyama K, Endo Y, Ozawa T, *et al*. Diagnosis using deep-learning artificial intelligence based on the endocytoscopic observation of the esophagus. Esophagus 2019;16:180–187. doi: 10.1007/s10388-018-0651-7.
19. Kumagai Y, Takubo K, Sato T, Ishikawa H, Yamamoto E, Ishiguro T, *et al*. AI analysis and modified type classification for endocytoscopic observation of esophageal lesions. Dis Esophagus 2022;35:doac010. doi: 10.1093/dote/doac010.
20. Horie Y, Yoshio T, Aoyama K, Yoshimizu S, Horiuchi Y, Ishiyama A, *et al*. Diagnostic outcomes of esophageal cancer by artificial intelligence using convolutional neural networks. Gastrointest Endosc 2019;89:25–32. doi: 10.1016/j.gie.2018.07.037.
21. Meng QQ, Gao Y, Lin H, Wang TJ, Zhang YR, Feng J, *et al*. Application of an artificial intelligence system for endoscopic diagnosis of superficial esophageal squamous cell carcinoma. World J Gastroenterol 2022;28:5483–5493. doi: 10.3748/wjg.v28.i37.5483.
22. Chou CK, Nguyen HT, Wang YK, Chen TH, Wu IC, Huang CW, *et al*. Preparing well for esophageal endoscopic detection using a hybrid model and transfer learning. Cancers (Basel) 2023;15:3783. doi: 10.3390/cancers15153783.
23. Ikenoyama Y, Yoshio T, Tokura J, Naito S, Namikawa K, Tokai Y, *et al*. Artificial intelligence diagnostic system predicts multiple Lugol-voiding lesions in the esophagus and patients at high risk for esophageal squamous cell carcinoma. Endoscopy 2021;53:1105–1113. doi: 10.1055/a-1334-4053.
24. Ohmori M, Ishihara R, Aoyama K, Nakagawa K, Iwagami H, Matsuura N, *et al*. Endoscopic detection and differentiation of esophageal lesions using a deep neural network. Gastrointest Endosc 2020;91:301–309.e1. doi: 10.1016/j.gie.2019.09.034.
25. Fukuda H, Ishihara R, Kato Y, Matsunaga T, Nishida T, Yamada T, *et al*. Comparison of performances of artificial intelligence versus expert endoscopists for real-time assisted diagnosis of esophageal squamous cell carcinoma (with video). Gastrointest Endosc 2020;92:848–855. doi: 10.1016/j.gie.2020.05.043.
26. Tajiri A, Ishihara R, Kato Y, Inoue T, Matsueda K, Miyake M, *et al*. Utility of an artificial intelligence system for classification of esophageal lesions when simulating its clinical use. Sci Rep 2022;12:6677. doi: 10.1038/s41598-022-10739-2.
27. Tani Y, Ishihara R, Inoue T, Okubo Y, Kawakami Y, Matsueda K, *et al*. A single-center prospective study evaluating the usefulness of artificial intelligence for the diagnosis of esophageal squamous cell carcinoma in a real-time setting. BMC Gastroenterol 2023;23:184. doi: 10.1186/s12876-023-02788-2.
28. Aoyama N, Nakajo K, Inaba A, Takashima K, Kadota T, Yoda Y, *et al*. Real-time detection of superficial esophageal squamous cell carcinoma by artificial intelligence using endoscopic videos. Gastrointest Endosco 2022;95:AB251–AB252. doi: 10.1016/j.gie.2022.04.653.
29. Shiroma S, Yoshio T, Kato Y, Horie Y, Namikawa K, Tokai Y, *et al*. Ability of artificial intelligence to detect T1 esophageal squamous cell carcinoma from endoscopic videos and the effects of real-time assistance. Sci Rep 2021;11:7759. doi: 10.1038/s41598-021-87405-6.
30. Waki K, Ishihara R, Kato Y, Shoji A, Inoue T, Matsueda K, *et al*. Usefulness of an artificial intelligence system for the detection of esophageal squamous cell carcinoma evaluated with videos simulating overlooking situation. Dig Endosc 2021;33:1101–1109. doi: 10.1111/den.13934.
31. Gruner M, Denis A, Masliah C, Amil M, Metivier-Cesbron E, Luet D, *et al*. Narrow-band imaging versus Lugol chromoendoscopy for esophageal squamous cell cancer screening in normal endoscopic practice: Randomized controlled trial. Endoscopy 2021;53:674–682. doi: 10.1055/a-1224-6822.
32. Morita FH, Bernardo WM, Ide E, Rocha RS, Aquino JC, Minata MK, *et al*. Narrow band imaging versus Lugol chromoendoscopy to diagnose squamous cell carcinoma of the esophagus: A systematic review and meta-analysis. BMC Cancer 2017;17:54. doi: 10.1186/s12885-016-3011-9.
33. Nagami Y, Tominaga K, Machida H, Nakatani M, Kameda N, Sugimori S, *et al*. Usefulness of non-magnifying narrow-band imaging in screening of early esophageal squamous cell carcinoma: A prospective comparative study using propensity score matching. Am J Gastroenterol 2014;109:845–854. doi: 10.1038/ajg.2014.94.
34. Yang XX, Li Z, Shao XJ, Ji R, Qu JY, Zheng MQ, *et al*. Real-time artificial intelligence for endoscopic diagnosis of early esophageal squamous cell cancer (with video). Dig Endosc 2021;33:1075–1084. doi: 10.1111/den.13908.
35. Yuan XL, Guo LJ, Liu W, Zeng XH, Mou Y, Bai S, *et al*. Artificial intelligence for detecting superficial esophageal squamous cell carcinoma under multiple endoscopic imaging modalities: A multicenter study. J Gastroenterol Hepatol 2022;37:169–178. doi: 10.1111/jgh.15689.
36. Yuan XL, Liu W, Lin YX, Deng QY, Gao YP, Wan L, *et al*. Effect of an artificial intelligence-assisted system on endoscopic diagnosis of superficial oesophageal squamous cell carcinoma and precancerous lesions: A multicentre, tandem, double-blind, randomised controlled trial. Lancet Gastroenterol Hepatol 2023;9:34–44. doi: 10.1016/s2468-1253(23)00276-5.
37. Ishihara R, Arima M, Iizuka T, Oyama T, Katada C, Kato M, *et al*. Endoscopic submucosal dissection/endoscopic mucosal resection guidelines for esophageal cancer. Dig Endosc 2020;32:452–493. doi: 10.1111/den.13654.
38. Liu W, Yuan X, Guo L, Pan F, Wu C, Sun Z, *et al*. Artificial intelligence for detecting and delineating margins of early ESCC under WLI endoscopy. Clin Transl Gastroenterol 2022;13:e00433. doi: 10.14309/ctg.0000000000000433.
39. Guo L, Xiao X, Wu C, Zeng X, Zhang Y, Du J, *et al*. Real-time automated diagnosis of precancerous lesions and early esophageal squamous cell carcinoma using a deep learning model (with videos). Gastrointest Endosc 2020;91:41–51. doi: 10.1016/j.gie.2019.08.018.
40. Yuan XL, Zeng XH, Liu W, Mou Y, Zhang WH, Zhou ZD, *et al*. Artificial intelligence for detecting and delineating the extentofsuperficial esophageal squamous cell carcinoma andprecancerous lesions under narrow-band imaging (with video). Gastrointest Endosc 2023;97:664–672.e4. doi: 10.1016/j.gie.2022.12.003.
41. Yuan X, Zeng X, He L, Ye L, Liu W, Hu Y, *et al*. Artificial intelligence for detecting and delineating a small flat-type early esophageal squamous cell carcinoma under multimodal imaging. Endoscopy 2023;55:E141–E142. doi: 10.1055/a-1956-0569.
42. Bollschweiler E, Baldus SE, Schröder W, Prenzel K, Gutschow C, Schneider PM, *et al*. High rate of lymph-node metastasis in submucosal esophageal squamous-cell carcinomas and adenocarcinomas. Endoscopy 2006;38:149–156. doi: 10.1055/s-2006-924993.
43. Natsugoe S, Baba M, Yoshinaka H, Kijima F, Shimada M, Shirao K, *et al*. Mucosal squamous cell carcinoma of the esophagus: A clinicopathologic study of 30 cases. Oncology 1998;55:235–241. doi: 10.1159/000011857.
44. Tajima Y, Nakanishi Y, Tachimori Y, Kato H, Watanabe H, Yamaguchi H, *et al*. Significance of involvement by squamous cell carcinoma of the ducts of esophageal submucosal glands. Analysis of 201 surgically resected superficial squamous cell carcinomas. Cancer 2000;89:248–254. doi: 10.1002/1097-0142(20000715)89:2<248:aid-cncr7>3.0.co;2-q.
45. Wang YK, Syu HY, Chen YH, Chung CS, Tseng YS, Ho SY, *et al*. Endoscopic images by a single-shot multibox detector for the identification of early cancerous lesions in the esophagus: A pilot study. Cancers (Basel) 2021;13:321. doi: 10.3390/cancers13020321.
46. Tokai Y, Yoshio T, Aoyama K, Horie Y, Yoshimizu S, Horiuchi Y, *et al*. Application of artificial intelligence using convolutional neural networks in determining the invasion depth of esophageal squamous cell carcinoma. Esophagus 2020;17:250–256. doi: 10.1007/s10388-020-00716-x.
47. Nakagawa K, Ishihara R, Aoyama K, Ohmori M, Nakahira H, Matsuura N, *et al*. Classification for invasion depth of esophageal squamous cell carcinoma using a deep neural network compared with experienced endoscopists. Gastrointest Endosc 2019;90:407–414. doi: 10.1016/j.gie.2019.04.245.
48. Shimamoto Y, Ishihara R, Kato Y, Shoji A, Inoue T, Matsueda K, *et al*. Real-time assessment of video images for esophageal squamous cell carcinoma invasion depth using artificial intelligence. J Gastroenterol 2020;55:1037–1045. doi: 10.1007/s00535-020-01716-5.

49. Zhang L, Luo R, Tang D, Zhang J, Su Y, Mao X, *et al*. Human-like artificial intelligent system for predicting invasion depth of esophageal squamous cell carcinoma using magnifying narrow-band imaging endoscopy: A retrospective multicenter study. Clin Transl Gastroenterol 2023;14:e00606. doi: 10.14309/ctg.0000000000000606.
50. Oyama T, Inoue H, Arima M, Momma K, Omori T, Ishihara R, *et al*. Prediction of the invasion depth of superficial squamous cell carcinoma based on microvessel morphology: Magnifying endoscopic classification of the Japan esophageal society. Esophagus 2017;14:105–112. doi: 10.1007/s10388-016-0527-7.
51. Everson M, Herrera L, Li W, Luengo IM, Ahmad O, Banks M, *et al*. Artificial intelligence for the real-time classification of intrapapillary capillary loop patterns in the endoscopic diagnosis of early oesophageal squamous cell carcinoma: A proof-of-concept study. United European Gastroenterol J 2019;7:297–306. doi: 10.1177/2050640618821800.
52. Everson MA, Garcia-Peraza-Herrera L, Wang HP, Lee CT, Chung CS, Hsieh PH, *et al*. A clinically interpretable convolutional neural network for the real-time prediction of early squamous cell cancer of the esophagus: Comparing diagnostic performance with a panel of expert European and Asian endoscopists. Gastrointest Endosc 2021;94:273–281. doi: 10.1016/j.gie.2021.01.043.
53. Zhao YY, Xue DX, Wang YL, Zhang R, Sun B, Cai YP, *et al*. Computer-assisted diagnosis of early esophageal squamous cell carcinoma using narrow-band imaging magnifying endoscopy. Endoscopy 2019;51:333–341. doi: 10.1055/a-0756-8754.
54. Wang J, Long Q, Liang Y, Song J, Feng Y, Li P, *et al*. AI-assisted identification of intrapapillary capillary loops in magnification endoscopy for diagnosing early-stage esophageal squamous cell carcinoma: A preliminary study. Med Biol Eng Comput 2023;61:1631–1648. doi: 10.1007/s11517-023-02777-3.
55. Uema R, Hayashi Y, Tashiro T, Saiki H, Kato M, Amano T, *et al*. Use of a convolutional neural network for classifying microvessels of superficial esophageal squamous cell carcinomas. J Gastroenterol Hepatol 2021;36:2239–2246. doi: 10.1111/jgh.15479.
56. Yuan XL, Liu W, Liu Y, Zeng XH, Mou Y, Wu CC, *et al*. Artificial intelligence for diagnosing microvessels of precancerous lesions and superficial esophageal squamous cell carcinomas: A multicenter study. Surg Endosc 2022;36:8651–8662. doi: 10.1007/s00464-022-09353-0.
57. Glissen Brown JR, Mansour NM, Wang P, Chuchuca MA, Minchenberg SB, Chandnani M, *et al*. Deep learning computer-aided polyp detection reduces adenoma miss rate: A united states multi-center randomized tandem colonoscopy study (CADeT-CS Trial). Clin Gastroenterol Hepatol 2022;20:1499–1507.e7. doi: 10.1016/j.cgh.2021.09.009.
58. Li JW, Wu CCH, Lee JWJ, Liang R, Soon GST, Wang LM, *et al*. Real-world validation of a computer-aided diagnosis system for prediction of polyp histology in colonoscopy: A prospective multicenter study. Am J Gastroenterol 2023;118:1353–1364. doi: 10.14309/ajg.0000000000002282.
59. Minegishi Y, Kudo SE, Miyata Y, Nemoto T, Mori K, Misawa M. Comprehensive diagnostic performance of real-time characterization of colorectal lesions using an artificial intelligence-assisted system: A prospective study. Gastroenterology 2022;163:323–325.e3. doi: 10.1053/j.gastro.2022.03.053.
60. Shaukat A, Lichtenstein DR, Somers SC, Chung DC, Perdue DG, Gopal M, *et al*. Computer-aided detection improves adenomas per colonoscopy for screening and surveillance colonoscopy: A randomized trial. Gastroenterology 2022;163:732–741. doi: 10.1053/j.gastro.2022.05.028.
61. Su JR, Li Z, Shao XJ, Ji CR, Ji R, Zhou RC, *et al*. Impact of a real-time automatic quality control system on colorectal polyp and adenoma detection: A prospective randomized controlled study (with videos). Gastrointest Endosc 2020;91:415–424.e4. doi: 10.1016/j.gie.2019.08.026.
62. Wang P, Berzin TM, Glissen Brown JR, Bharadwaj S, Becq A, Xiao X, *et al*. Real-time automatic detection system increases colonoscopic polyp and adenoma detection rates: A prospective randomised controlled study. Gut 2019;68:1813–1819. doi: 10.1136/gutjnl-2018-317500.
63. Zippelius C, Alqahtani SA, Schedel J, Brookman-Amissah D, Muehlenberg K, Federle C, *et al*. Diagnostic accuracy of a novel artificial intelligence system for adenoma detection in daily practice: A prospective nonrandomized comparative study. Endoscopy 2022;54:465–472. doi: 10.1055/a-1556-5984.