



Artificial intelligence-assisted endoscopic ultrasound diagnosis of esophageal subepithelial lesions

Ai-meng Zhang¹ · Dai-min Jiang² · Shu-peng Wang¹ · Wen Liu¹ · Bei-bei Sun¹ · Zhe Wang³ · Guo-yi Zhou² · Yao-fu Wu⁴ · Qing-yun Cai⁵ · Jin-tao Guo^{1,6} · Si-yu Sun^{1,6}

Received: 4 December 2024 / Accepted: 17 April 2025 / Published online: 7 May 2025
© The Author(s) 2025

Abstract

Background Endoscopic ultrasound (EUS) is one of the most accurate methods for determining the originating layer of subepithelial lesions (SELs). However, the accuracy is greatly influenced by the expertise and proficiency of the endoscopist. In this study, we aimed to develop an artificial intelligence (AI) model to identify the originating layer of SELs in the esophagus and evaluate its efficacy.

Methods A total of 1445 cases of esophageal SELs were used to develop the model. An AI model stemming from YOLOv8s-seg and MobileNetv2 was developed to detect esophageal lesions and identify the originating layer. Two seniors and two junior endoscopists independently diagnosed the same test set.

Results The precision, recall, mean average precision @ 0.5, and F1-score of the AI model were 92.2%, 73.6%, 0.832, and 81.9%, respectively. The overall accuracy of the originating layer recognition model was 55.2%. The F1-scores of the second, third, and fourth layers were 47.1%, 51.7%, and 66.1%, respectively. The accuracy of the AI system in differentiating layers 2 and 3 from four was 76.5% and was similar to that of senior endoscopists (74.9–79.8%, $P = 0.585$) but higher than that of junior endoscopists (65.6–66.7%, $P = 0.045$).

Conclusions The EUS-AI model has shown high diagnostic potential for detecting esophageal SELs and identifying their originating layers. EUS-AI has the potential to enhance the diagnostic ability of junior endoscopists in clinical practice.

Keywords Artificial intelligence · Endoscopic ultrasound · Esophagus · Subepithelial lesion

Most esophageal subepithelial lesions (SELs) are benign, such as lipoma, leiomyoma, and schwannoma. However, SELs such as gastrointestinal stromal tumors (GISTs) and granular cell tumors can undergo malignant transformation

[1, 2]. In current guidelines, there is no consensus on the management of esophageal SELs [1, 3, 4]; however, some suggest that asymptomatic, hypoechoic, well-defined (no high-risk features on endoscopic ultrasound [EUS]) SELs with a diameter of < 2 cm can be managed with continuous surveillance strategy [4]. Nevertheless, extended patient

Ai-meng Zhang considered as first author.

✉ Jin-tao Guo
guojt@sj-hospital.org

Ai-meng Zhang
zhangam_cmu@163.com

¹ Department of Gastroenterology, Shengjing Hospital of China Medical University, No. 36, Sanhao Street, Shenyang 110004, Liaoning Province, China

² Research Center for Innovation, SonoScape Medical Corporation, Shenzhen 518107, Guangdong Province, China

³ Department of Pathology, Shengjing Hospital of China Medical University, Shenyang 110004, Liaoning Province, China

⁴ Digital Information Development Department, SonoScape Medical Corporation, Shenzhen 518107, Guangdong Province, China

⁵ Product Management Department, SonoScape Medical Corporation, Shenzhen 518107, Guangdong Province, China

⁶ Engineering Research Center of Ministry of Education for Minimally Invasive, Gastrointestinal Endoscopic Techniques, Shenyang 110004, Liaoning Province, China

monitoring causes apprehension regarding adverse events, cost-effectiveness, risks of repeat endoscopy, and treatment delay of potentially malignant lesions. Therefore, for the pathological diagnosis and treatment of esophageal SELs, surgical resection remains a crucial approach [5]. There are several techniques for the endoscopic resection of SELs. For example, endoscopic mucosal resection [6, 7] is commonly performed on lesions in the second and third layers. The fourth layer is treated using submucosal tunnel endoscopic resection [8, 9], endoscopic submucosal dissection [10], or endoscopic full-thickness resection [11, 12]. Therefore, the preoperative diagnosis of the originating layer is critical in selecting the best treatment approach [3, 13, 14]. Moreover, grouping the second and third layers versus isolating the fourth layer can accurately match therapeutic decision trees, preventing overtreatment or undertreatment.

Due to the rapid advancement of endoscopic technology and health consciousness among individuals, an increasing number of esophageal SELs have been detected during examination, with an increased demand for EUS diagnosis [15]; however, there is a significant shortage of endo-sonographers. To assess the competency in EUS, the American Society for Gastrointestinal Endoscopy recommends 225 cases [16], whereas the European Society of Gastrointestinal Endoscopy performing at least 250 cases [17]. However, a considerable number of endoscopists have limited practical training, and tertiary referral hospitals house skilled endo-sonographers [18–20]. For amateurs, the learning curve of EUS technology is steep, the diagnostic level varies, and the training system requires improvement [21, 22].

In recent years, artificial intelligence (AI) has made significant advances in medical science [23, 24]. Currently, the role of AI and EUS in diagnosing digestive system disorders has received considerable attention in medical research [25, 26]. Recent research has predominantly focused on GISTs, and the supplementary role of AI in diagnosing esophageal SELs has not been investigated. In this study, we aimed to develop AI models that detect and classify the originating layers of esophageal SELs and evaluate their diagnostic efficacy.

Materials and methods

EUS image selection

In this retrospective study, EUS images of esophageal SELs were collected at Shengjing Hospital of China Medical University from January 2015 to September 2023.

Inclusion criteria:

- A. Images of esophageal SELs that were either diagnosed and resected (confirmed by pathology) or unanimously

diagnosed by two specialist endo-sonographers (with over 10 years of EUS diagnosis experience and more than 5000 EUS procedures performed).

- B. Complete EUS examination data.

Exclusion criteria:

- A. Poor quality or duplicate images.
- B. Images with artificial markers.

This study was approved by the Medical Ethics Committee of Shengjing Hospital of China Medical University (No. 2023PS1355K).

Image processing

Two specialist endo-sonographers employed Sonokit software (SonoScape Medical Corp., Shenzhen, China) to mark the boundary contour, short and long axes, and originating layer of the esophageal SELs (second layer/deep mucosa, third layer/submucosa, and fourth layer/lamina propria) in each image. Any two identical diagnoses were considered conclusive. However, with different diagnoses, the final diagnosis was referred to an authoritative expert in EUS (over 10 years of EUS diagnosis experience and > 5000 EUS examinations) for further discussion.

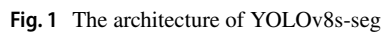
Deep learning algorithms

Lesion detection algorithm

In order to identify the best AI model, we developed three distinct AI frameworks: YOLOv8s-seg, YOLOv8, and YOLOv5s. The overall performance of YOLOv8s-seg was comparatively superior to the other convolutional neural network (CNN) architectures (Table S1). Consequently, YOLOv8s-seg was selected for the subsequent study. YOLO [27] is a real-time target detection algorithm that utilizes CNN for feature extraction and the identification of species and locations. YOLOv8 builds on the previous YOLO series by adopting its C2f structure as the new backbone network. It introduced a new anchor-free detection header that does not rely on an anchor box. It applies a loss function with a task-aligned assigner-positive sample allocation strategy to introduce the Distribution Focal Loss as a regression loss. YOLOv8s-seg provides better detection and segmentation results while maintaining a fast processing speed. The architecture of YOLOv8s-seg is illustrated in Fig. 1.

Originating layer recognition algorithm

In recognition of the originating layers, a classification network called MobileNetv2 performed better than did the other



The labeled EUS images were randomly divided into training, validation, and test sets in a ratio of 8:1:1 at the patient level, ensuring that there was no intra-patient leakage. Subsequently, we constructed the lesion detection and originating layer recognition models. The training set images were employed to learn the characteristics of the esophageal SELs.

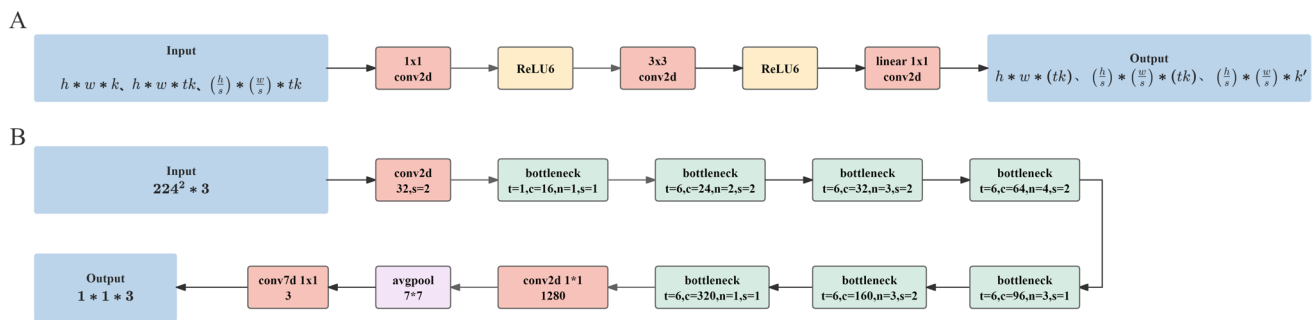


Fig. 2 The architecture of MobileNetV2. **A** Bottleneck residual block. **B** Overall network structure of MobileNetV2

The hyperparameters were optimized using the validation set images (Table S3). The test set images helped to assess the diagnostic capability of the optimal model for esophageal SELs.

Model testing and comparison of diagnostic efficacy with endoscopists

Model testing: The location and originating layer of SELs were diagnosed by AI.

Physician identification: Image boundary contour and originating layer labeling were performed by two senior endoscopists (> 5 years of EUS experience and 500 EUS examinations) and two junior endoscopists (less than 1 year of EUS experience and less than 200 EUS examinations). None of the endoscopists were involved in selecting or labeling the dataset, and the clinical characteristics and the results of endoscopic and pathological examination remained unknown.

Evaluation metrics

The performance of the AI model was measured with the precision (P), recall (R), mean average precision (mAP), and F1-score. The F1-score is the harmonic mean of precision and recall. A precision-recall curve (P-R curve) was constructed. The term “Average Precision” (AP) represents the area under the curve. The mean precision of all sample categories produced a weighted average. This metric was employed to assess the detection capabilities across all the categories. In this study, the AP was equivalent to the mAP. The receiver operating characteristic curve (ROC curve) was illustrated, and the area under the curve (AUC) was calculated. The formulas for the related concepts are as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$F1 = \frac{2TP}{2TP + FP + FN} \quad (4)$$

$$AP = \int_0^1 P(R) dR \quad (5)$$

A true positive (TP) signifies that the genuine category of the sample is positive, and the model predicts positive cases. A true negative (TN) indicates that the genuine category is positive and the prediction result is negative. A false positive (FP) means that the genuine category is negative and the prediction result is positive. A false negative (FN) signifies that the genuine category is negative and the prediction result is negative. Furthermore, the accuracies of the AI model and endo-sonographers in diagnosing esophageal SELs were appraised with 95% confidence intervals (CI). The time required to diagnose the same test set was measured by both groups.

Statistical methods

Statistical analyses were carried out using SPSS Statistics version 26.0 by the author (A-MZ). Categorical variables were compared using the two-paired McNemar’s test, with statistical significance set at $P < 0.05$. The Kappa test was used to evaluate the consistency of endoscopists’ diagnostic results: $\kappa \geq 0.75$, good agreement; $0.75 > \kappa \geq 0.4$, moderate agreement; $\kappa < 0.4$, poor agreement.

Results

General information

A total of 1445 patients diagnosed with esophageal SELs using EUS were enrolled. In addition, 1605 EUS images

were labeled by endoscopists. Table 1 presents the precise distribution of the data. The data were randomly assigned by a computer into training, validation, and test sets (Fig. 3). Table 2 shows the cohort characteristics of patients for each group.

Table 1 Number of images for each type of data for esophageal SELs

	The second layer	The third layer	The fourth layer
Annular array EUS	526	475	462
Linear array EUS	25	45	72

SEL Subepithelial lesion, *EUS* endoscopic ultrasound

Lesion detection model

The precision of the lesion location detection was 92.2% (95%CI, 87.0–95.3%), the recall was 73.6% (95%CI, 66.1–79.5%), mAP@0.5 was 0.832, and F1-score was 81.9% (95%CI, 75.2–87.0%). Figure 4 shows an example of esophageal SEL recognition using an EUS-AI image model.

The EUS-AI model illustrated an accuracy of 73.6%, which was similar to the senior endoscopists' range of (71.5–80.6%) ($P=0.146$). In contrast, the accuracy was significantly greater than that of junior endoscopists (61.8–65.5%, $P=0.041$) (Table 3).

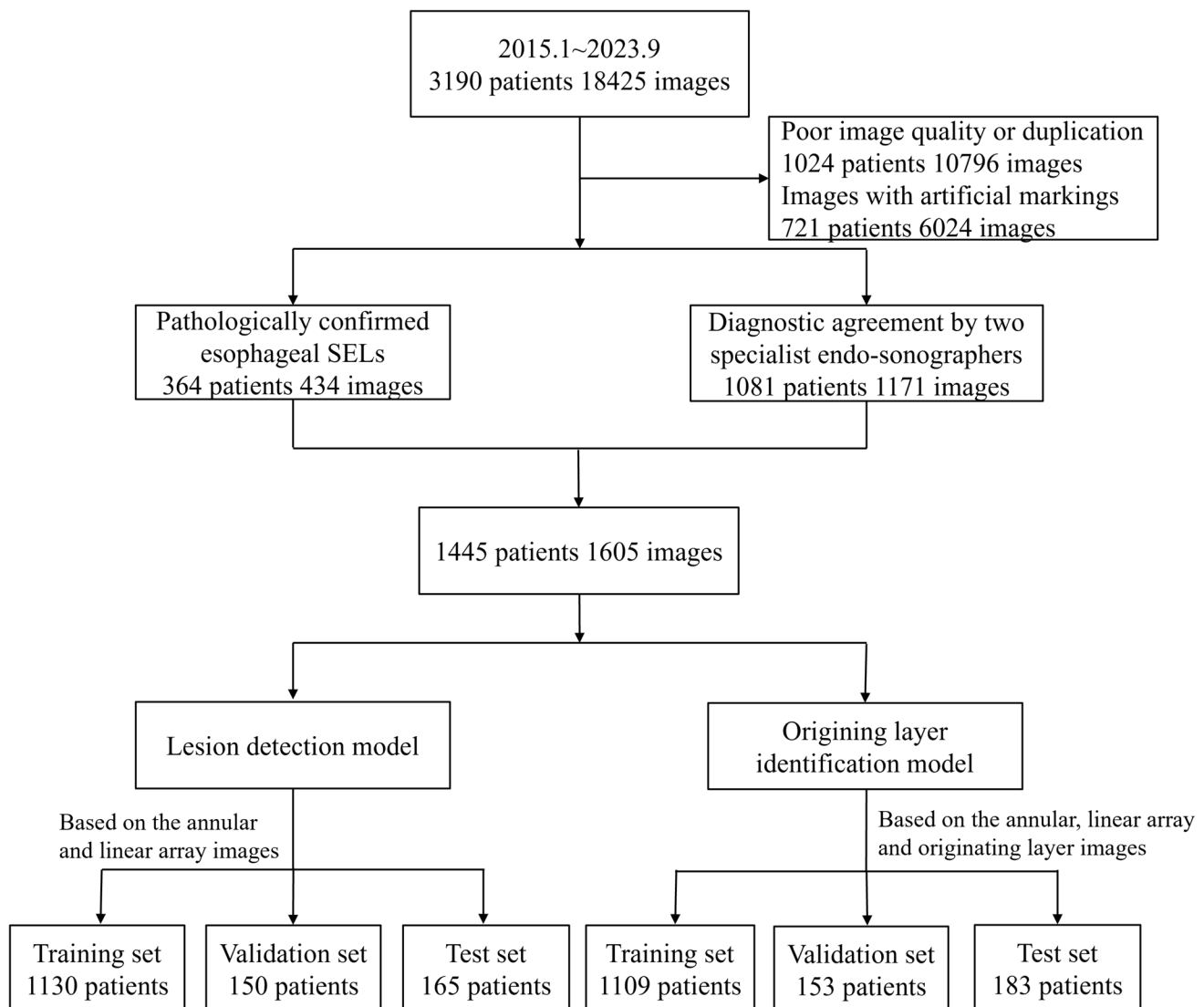


Fig. 3 Case inclusion, exclusion, and grouping flowchart

Table 2 Clinical characteristics of patients for each dataset of AI models

Characteristics	Lesion detection algorithm			Originating layer recognition algorithm		
	Training Set (<i>n</i> = 1130)	Validation Set (<i>n</i> = 150)	Test Set (<i>n</i> = 165)	Training Set (<i>n</i> = 1109)	Validation Set (<i>n</i> = 153)	Test Set (<i>n</i> = 183)
Age, Years (mean ± SD)	54.6 ± 10.8	54.2 ± 11.1	53.3 ± 11.1	54.5 ± 11.1	54.2 ± 9.7	53.9 ± 10.5
Sex						
Male	600 (53.1%)	83 (55.3%)	84 (50.9%)	594 (53.6%)	84 (54.9%)	90 (49.2%)
Female	530 (46.9%)	67 (44.7%)	81 (49.1%)	515 (46.4%)	69 (45.1%)	93 (50.8%)
Tumor size, mm (mean ± SD)	8.3 ± 7.7	9.4 ± 10.9	9.4 ± 8.4	8.5 ± 8.2	7.9 ± 7.7	9.1 ± 8.1
Tumor location						
Upper esophagus	279 (24.7%)	34 (22.7%)	46 (27.9%)	274 (24.7%)	36 (23.5%)	48 (26.2%)
Middle esophagus	278 (24.6%)	52 (34.7%)	60 (36.4%)	372 (33.5%)	54 (35.3%)	62 (33.9%)
Lower esophagus	473 (50.7%)	64 (42.6%)	59 (35.7%)	463 (41.8%)	63 (41.2%)	73 (39.9%)

AI artificial intelligence, SD standard deviation

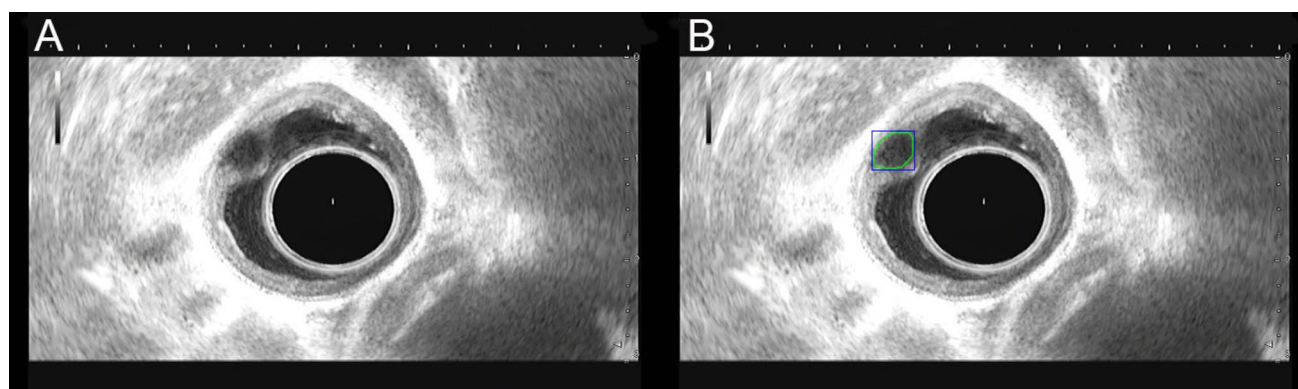


Fig. 4 Example of test set image labeling. **A** Representative image of esophageal SELs. **B** AI model recognition image. The green frame indicates the diagnosis by the AI system. SEL subepithelial lesion, EUS endoscopic ultrasound

Originating layer recognition model

Ternary-category classification performance

The originating layer recognition model's overall accuracy was 55.2% (95%CI, 48.0–62.2%). Figure 5 shows an example of the prediction results. The originating layers were 2, 3, and 4, with precisions of 50.0% (95%CI, 37.5–62.5%), 54.5% (95%CI, 41.5–67.0%), and 60.0% (95%CI, 48.3–70.7%), respectively; moreover, the recalls were 44.6% (95%CI, 33.2–56.7%), 49.2% (95%CI, 37.1–61.4%), and 73.7% (95%CI, 61.0–83.4%), respectively. The F1-score of the second, third, and fourth layers was 47.1% (95%CI, 39.9–54.2%), 51.7% (95%CI, 44.7–59.0%), and 66.1% (95%CI, 59.0–72.6%), respectively. The confusion matrix is shown in Fig. 6A. The AUC for the ternary-category classification was 0.687 (Fig. 6B).

Table 3 Comparison of diagnostic performance between the lesion detection model and endoscopists

Diagnostician	Accuracy, % (95%CI)	<i>P</i> ^a
AI model	73.6(66.1–79.5)	Reference
All senior endoscopists	76.1(69.3–81.6)	0.146
Senior endoscopist 1	71.5(64.2–77.9)	0.649
Senior endoscopist 2	80.6(73.9–85.9)	0.245
All junior endoscopists	63.6(56.2–70.0)	0.041*
Junior endoscopist 1	65.5(57.9–72.3)	0.086
Junior endoscopist 2	61.8(54.2–68.9)	0.026*

CI confidence interval, AI artificial intelligence

^aComparing statistical differences of accuracy between endoscopists and lesion detection models

**P* < 0.05

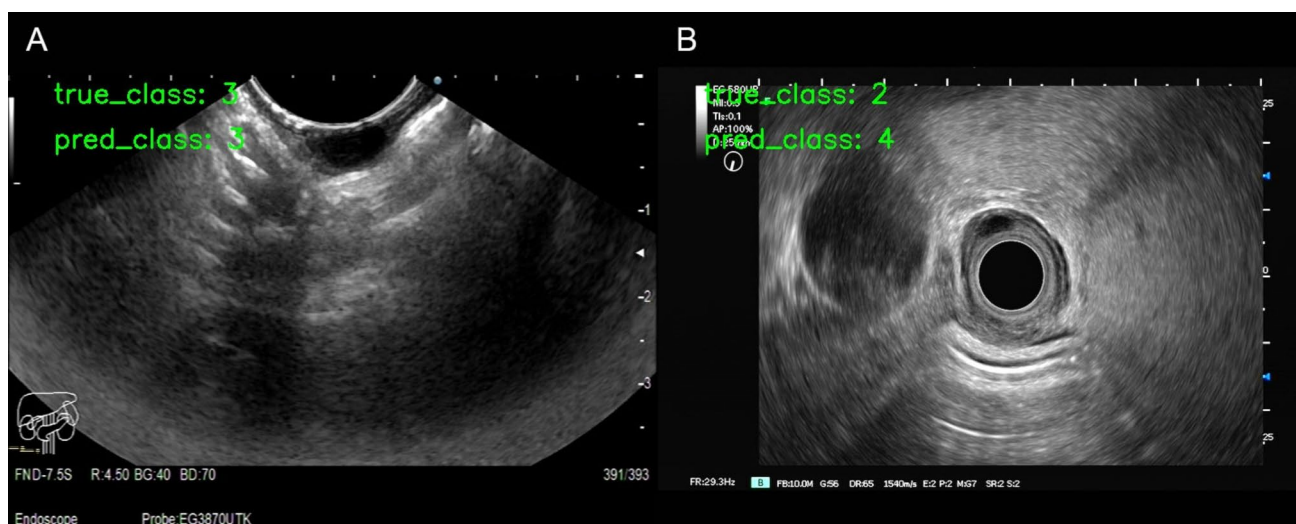


Fig. 5 Example of origin-level prediction results. **A** The AI model predicted the third layer lesion as the correct layer. **B** The AI model misdiagnosed the second layer lesion as the fourth

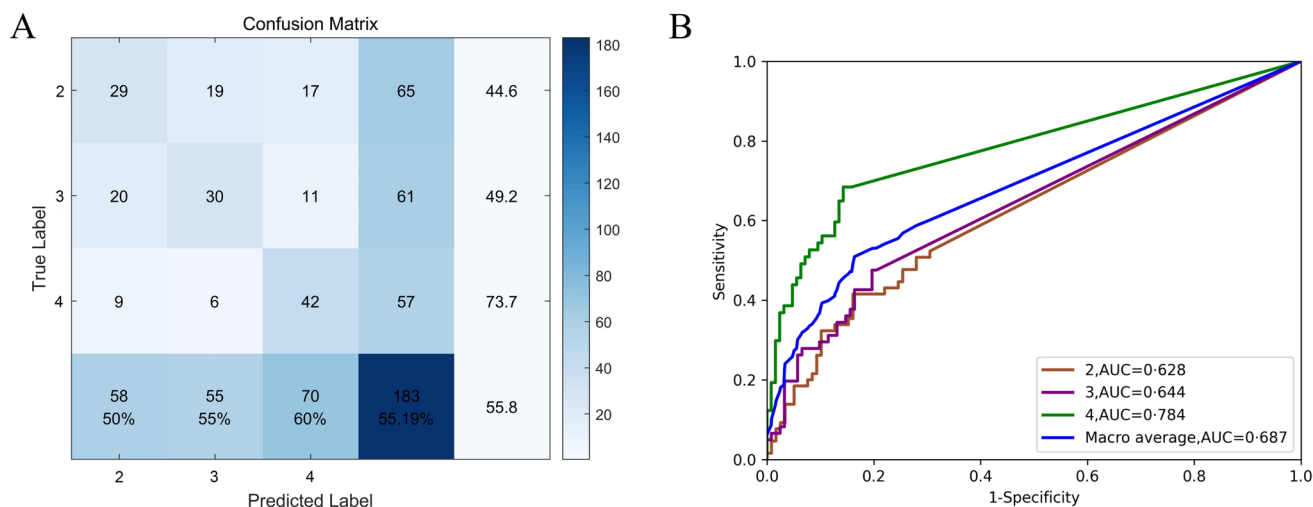


Fig. 6 Diagnostic performance of the originating layer recognition model. **A** The confusion matrix of the originating layer recognition model. The originating layers were 2, 3, and 4, with a precision of 50.0, 54.5, and 60.0%, respectively, and a recall of 44.6, 49.2, and

73.7%, respectively. **B** ROC curves of the originating layer recognition model. The AUC for the second, third, and fourth layers was 0.628, 0.644, and 0.784, respectively

Binary category classification performance

Endoscopic resection of esophageal SELs originating from the muscularis propria (fourth layer) differs from the lesions in the second and third layers. Resection of these lesions can be difficult. Consequently, we categorized the identification of the originating layer into the second or third layer and the fourth layer. The accuracy of the EUS-AI system in distinguishing the fourth layer from the second or third layers was 76.5% (95%CI, 69.9–82.1%) with an AUC of 0.784 (Fig. 5B).

Comparison of diagnostic performance between AI model and endoscopists

The accuracy of the AI model (55.2%) in the ternary-category did not show a significant difference when compared with senior endoscopists (53.6–65.0%, $P=0.057$) and was superior to that of junior endoscopists (42.1–48.1%, $P=0.043$). In the binary category classification, the accuracy of the AI model (76.5%) was similar to that of the senior endoscopists (74.9–79.8%, $P=0.585$) but

Table 4 Comparison of efficacy between the originating layer identification model and endoscopists

	Diagnostician	Accuracy, % (95%CI)	<i>P</i> ^a
Ternary	AI model	55.2(48.0–62.2)	Reference
	All senior endoscopists	59.3(51.8–65.9)	0.057
	Senior endoscopist 1	53.6(46.3–60.6)	0.255
	Senior endoscopist 2	65.0(57.9–71.6)	0.058
	All junior endoscopists	45.1(37.8–52.0)	0.043*
	Junior endoscopist 1	48.1(41.0–55.3)	<0.001***
	Junior endoscopist 2	42.1(35.2–49.3)	0.098
Binary	AI model	76.5(69.9–82.1)	Reference
	All senior endoscopists	77.4(70.4–82.5)	0.585
	Senior endoscopist 1	74.9(68.1–80.6)	0.306
	Senior endoscopist 2	79.8(73.4–85.0)	>0.999
	All junior endoscopists	66.2(59.0–72.6)	0.045*
	Junior endoscopist 1	66.7(59.6–73.1)	0.047*
	Junior endoscopist 2	65.6(58.4–72.1)	0.018*

CI confidence interval, AI artificial intelligence

^aStatistical differences in accuracy of comparing endoscopists to the originating layer identification model in tertiary and binary classification

P*<0.05, *P*<0.01, ****P*<0.001

higher than that of the junior endoscopists (65.6–66.7%, *P*=0.045) (Table 4).

Interobserver agreement among endoscopists

In ternary categorization, interobserver agreement in distinguishing the originating layer was higher for senior endoscopists (κ =0.430) than for junior endoscopists (κ =0.227). In the binary categorization, the interobserver agreement between senior endoscopists distinguishing between the second or third layer with the fourth layer (κ =0.419) was also higher than the interobserver agreement between junior endoscopists (κ =0.257).

Diagnostic time comparison of artificial intelligence models and endoscopists

When comparing the time required to complete the initial layer diagnosis, the junior endoscopists took an average of 6.78 s (ranging from 6.13 to 7.42 s) per image, whereas the senior endoscopists had an average of 4.03 s (ranging from 3.64 to 4.41 s). Conversely, the EUS-AI model required only 0.01 s,

indicating that the model can achieve a diagnosis of more than 100 images per second.

Discussion

EUS is an important method for identifying the location and originating layer of SELs [3]. However, the diagnostic accuracy predominantly depends on the expertise and experience of the endo-sonographers. CEUS encounters certain problems, such as unbalanced technical development, subpar diagnosis and treatment quality, and limited training facilities [16, 17, 29]. Therefore, in regions with a low EUS prevalence, the extensive development, diagnosis, and treatment of SEL diseases are limited. However, AI has the potential to address the demand for EUS by improving the diagnostic performance of junior endoscopists [30].

Previous studies have established AI models for the diagnoses of SELs [31, 32]. Hirai et al. [26] constructed a CNN model that classified five categories of SELs with an accuracy of 86.1%. For non-gastric SELs, Minoda et al. [33] developed an AI system to differentiate GISTs from non-GISTs, with an accuracy of 94.4%. However, these studies focused solely on the nature of the tumors. For classifying the origin of mural layers in the upper gastrointestinal tract, Li et al. [34] utilized data from 313 patients and developed an AI system with an 80.68% accuracy. However, the number of patients included in the study was limited.

Unlike previous studies, in this study, we developed an AI model using the YOLOv8s-seg algorithm for detecting esophageal SELs, used for analyzing esophageal SELs and constructed an AI model with the MobileNetv2 algorithm to differentiate the originating layer of the lesions, used for analyzing the accuracy. This represents a novel study in which AI was applied for lesion detection and layer identification in esophageal SELs, incorporating a substantial number of cases. This study demonstrates that the AI-based EUS diagnostic method exhibits high accuracy in identifying esophageal SELs, which is crucial for enhancing early diagnostic capabilities. With the assistance of the AI model, clinicians can more accurately assess the characteristics of lesions, thereby providing preliminary data support for the implementation of the AI model in real clinical settings. In the future, we anticipate that the model will be validated in larger and more diverse clinical samples to ensure its universality and effectiveness across various medical environments.

The AI model exhibited the highest AUC (0.832) and excellent diagnostic performance among the lesion

models. Misdiagnosis of esophageal SELs may lead to the omission of malignant lesions and unnecessary treatment or follow-up. Consequently, using accuracy as the primary evaluation metric, we compared the diagnostic capabilities of the AI model with the endoscopists. The results indicated that the diagnostic performance of the AI model was comparable to that of senior endoscopists but significantly surpassed that of their junior counterparts. Hence, our AI model could be valuable for novice endoscopists in detecting esophageal SELs during EUS examinations.

The accuracy achieved by the endoscopists and AI system in the original layer of the recognition model was not as high as anticipated. In this study, only static images of the esophageal SELs were used. However, during endoscopy, endoscopists can enhance their view of the location of worrisome lesions by modifying or moving the sweep area. This aids in gathering additional information, ultimately leading to a more precise clinical diagnosis. Therefore, the test settings used in this study differ from the actual conditions encountered during endoscopy. Depending on the originating layer, size, and location of the esophageal SELs, appropriate treatment can be selected [3]. The accuracy of the EUS-AI model improved to 76.5% after the originating layers (second or third and fourth layers) were further binary-categorized, which was on par with the senior endoscopists but higher than that of the junior endoscopists. These results demonstrate the reliable diagnostic capability of the EUS-AI model.

Due to the subjective nature of the endoscopists' diagnosis of EUS images, a lack of high interobserver agreement was revealed in this study, which emphasizes the necessity of standardized diagnostic tools. The measure of experience performing EUS is particularly important for diagnostic accuracy. Our AI model may effectively enhance the diagnostic accuracy of less-experienced endoscopists and reduce their learning duration. In addition, compared with endoscopists, the AI model has an extremely fast diagnostic speed (10 ms per image). It can provide rapid diagnosis using EUS, therefore significantly reducing diagnostic time and relieving the burden on endoscopists. In conclusion, the AI model is expected to function as a "secondary observer" during EUS examinations, serving as a backup resource to fully assist endoscopists in diagnosing and managing esophageal SELs.

There were some limitations in this study. First, this was a single-center retrospective study conducted at Shengjing Hospital, which excluded lower-quality EUS images and, therefore, may have been subject to selection bias. The diagnostic accuracy was not satisfactory, highlighting the necessity for future multicenter prospective studies to broaden the database for AI model training. Second, the limited inclusion of pathologically confirmed SELs in this study stems from

the following reasons: (1) several esophageal SELs, such as cysts and lipomas, exhibit characteristic EUS features (e.g., anechoic cystic structures, homogeneous hyperechoic patterns) that enable reliable non-pathological diagnosis through follow-up observations; (2) a significant number of patients opt for surveillance over intervention; and (3) our endoscopic team (performing > 3000 annual EUS examinations) possess extensive experience and ensure diagnostic reliability. Although potential bias exists, we maintain that the direct utilization of EUS diagnostic outcomes does not undermine the credibility of model identification. Third, only image recognition of esophageal SELs was analyzed in this study, whereas video recognition and real-time experiments were not. The AI model accurately diagnosed EUS images at a rate of 10 ms/image. This is adequate for analyzing video images, indicating that real-time analysis will soon be achievable. Finally, one of the challenges with AI is the "black box problem." This implies that only the input and output layers are visible, whereas the processes and recognition in the hidden layers remain unknown.

In conclusion, the EUS-AI model has high diagnostic performance for detecting lesions and identifying the originating layer in esophageal SELs. This model has the potential to enhance the proficiency of amateur endoscopists in diagnosing esophageal SELs during clinical practice, optimize diagnostic efficiency, and mitigate the impact of subjective bias on diagnostic accuracy. In the future, it will be essential to conduct prospective multicenter studies to develop a more precise and thorough AI system that may further enhance diagnostic performance.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00464-025-11767-5>.

Acknowledgements We express our gratitude to all the doctors involved in this project. We thank the Liaoning Province Applied Basic Research Program Joint Program Project (2022 JH2/101500076), Shenyang Young, the Middle-aged Science and Technology Innovation Talent Support Program (Grant No. RC200438), and the Tree Planting Program of Shengjing Hospital (M1595) for their financial support. We thank SonoScape Medical Corp., Shenzhen, China, for the technical support. We would like to thank Editage (www.editage.cn) for English language editing.

Author contributions AZ and JG participated in the conception and design of the work; AZ collected the clinical data and drafted the article; DJ was involved in the development of the artificial models; AZ, GZ, YW and QC participated in data analysis and interpretation; SW, WL and BS were involved in the endoscopic procedures; ZW was involved in the pathological diagnosis; JG and SS participated in the critical revision and final approval of the article.

Funding Liaoning Province Applied Basic Research Program Joint Program Project (2022 JH2/101500076); Shenyang Young and Middle-aged Science and Technology Innovation Talent Support Program (Grant No. RC200438); Tree planting program of Shengjing Hospital (M1595).

Data availability The model source code can be accessed through this link: <https://github.com/GXDD666/SMT-E>.

Declarations

Disclosures Ai-meng Zhang, Dai-min Jiang, Shu-peng Wang, Wen Liu, Bei-bei Sun, Zhe Wang, Guo-yi Zhou, Yao-fu Wu, Qing-yun Cai, Jin-tao Guo, and Si-yu Sun have no conflicts of interest or financial ties to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Jacobson BC, Bhatt A, Greer KB, Lee LS, Park WG, Sauer BG, Shami VM (2023) ACG clinical guideline: diagnosis and management of gastrointestinal subepithelial lesions. *Am J Gastroenterol* 118:46–58
- Meng Y, Wang L, Yao J (2024) EUS diagnosis and rescue of gastrointestinal stromal tumor rupture and massive hemorrhage (with video). *Endosc Ultrasound* 13:123–125
- Sharzei K, Sethi A, Savides T (2022) AGA clinical practice update on management of subepithelial lesions encountered during routine endoscopy: expert review. *Clin Gastroenterol Hepatol* 20:2435–2443.e4
- Deprez PH, Moons LMG, O'Toole D, Gincul R, Seicean A, Pimentel-Nunes P, Fernández-Esparrach G, Polkowski M, Vieth M, Borbath I, Moreels TG, Nieveen van Dijkum E, Blay J-Y, van Hooft JE (2022) Endoscopic management of subepithelial lesions including neuroendocrine neoplasms: European society of gastrointestinal endoscopy (ESGE) guideline. *Endoscopy* 54:412–429
- Ono H, Yao K, Fujishiro M, Oda I, Uedo N, Nimura S, Yahagi N, Iishi H, Oka M, Ajioka Y, Fujimoto K (2021) Guidelines for endoscopic submucosal dissection and endoscopic mucosal resection for early gastric cancer (second edition). *Dig Endosc*. <https://doi.org/10.1111/den.13883>
- Kajiyama T, Sakai M, Torii A, Kishimoto H, Kin G, Uose S, Ueda S, Okuma M, Inoue K (1995) Endoscopic aspiration lumpectomy of esophageal leiomyomas derived from the muscularis mucosae. *Am J Gastroenterol* 90:417–422
- Guo J, Liu Z, Sun S, Wang S, Ge N, Liu X, Wang G, Liu W (2013) Endosonography-assisted diagnosis and therapy of gastrointestinal submucosal tumors. *Endosc Ultrasound* 2:125–133
- Du C, Ma L, Chai N, Gao Y, Niu X, Zhai Y, Li Z, Meng J, Tang P, Linghu E (2018) Factors affecting the effectiveness and safety of submucosal tunneling endoscopic resection for esophageal submucosal tumors originating from the muscularis propria layer. *Surg Endosc* 32:1255–1264
- Dalal I, Andalib I (2022) Advances in endoscopic resection: a review of endoscopic submucosal dissection (ESD), endoscopic full thickness resection (EFTR) and submucosal tunneling endoscopic resection (STER). *Transl Gastroenterol Hepatol* 7:19
- Draganov PV, Wang AY, Othman MO, Fukami N (2019) AGA institute clinical practice update: endoscopic submucosal dissection in the United States. *Clin Gastroenterol Hepatol* 17:16–25. e1
- Wang K, Gao P, Cai M, Song B, Zhou P (2023) Endoscopic full-thickness resection, indication, methods and perspectives. *Dig Endosc* 35:195–205
- Liu S, Zhou X, Yao Y, Shi K, Yu M, Ji F (2020) Resection of the gastric submucosal tumor (G-SMT) originating from the muscularis propria layer: comparison of efficacy, patients' tolerability, and clinical outcomes between endoscopic full-thickness resection and surgical resection. *Surg Endosc* 34:4053–4064
- Gupta S, Gauci J, O'Sullivan T, Cronin O, Whitfield A, Craciun A, Awadie H, Yang J, Kwan V, Lee EYT, Burgess NG, Bourke MJ (2024) A sub-epithelial lesion algorithm for endoscopic (SAFE) resection in the upper gastrointestinal tract. *Endoscopy*. <https://doi.org/10.1016/j.gie.2024.04.1827>
- Standards of Practice Committee, Faulx AL, Kothari S, Acosta RD, Agrawal D, Bruining DH, Chandrasekhara V, Eloubeidi MA, Fanelli RD, Gurudu SR, Khashab MA, Lightdale JR, Muthusamy VR, Shaikat A, Qumseya BJ, Wang A, Wani SB, Yang J, DeWitt JM (2017) The role of endoscopy in subepithelial lesions of the GI tract. *Gastrointest Endosc* 85:1117–1132
- Jiao Y, Cheng Z, Gao Y, Wang T, Xin L, Lin H, Cai M, Ma X, Li Z, Wang L (2024) Development and status quo of digestive endoscopy in China: an analysis based on the national census in 2013 and 2020. *J Transl Int Med* 12:177–187
- Faulx AL, Lightdale JR, Acosta RD, Agrawal D, Bruining DH, Chandrasekhara V, Eloubeidi MA, Gurudu SR, Kelsey L, Khashab MA, Kothari S, Muthusamy VR, Qumseya BJ, Shaikat A, Wang A, Wani SB, Yang J, DeWitt JM (2017) Guidelines for privileging, credentialing, and proctoring to perform GI endoscopy. *Gastrointest Endosc* 85:273–281
- Badaoui A, Teles de Campos S, Fusaroli P, Gincul R, Kahaleh M, Poley J-W, Sosa Valencia L, Czako L, Gines A, Hucl T, Kalaitzakis E, Petrone MC, Sadik R, van Driel L, Vandeputte L, Tham T (2024) Curriculum for diagnostic endoscopic ultrasound training in Europe: European society of gastrointestinal endoscopy (ESGE) position statement. *Endoscopy* 56:222–240
- Xin L, Gao Y, Wang T-J, Meng Q-Q, Jin Z-D, Fu Z-J, Wang Y-L, Lin H, Li Z-S, Wang L-W (2023) EUS development in China: results from national surveys in 2013 and 2020. *Endosc Ultrasound* 12:90–95
- Lekkerkerker SJ, Voermans RP (2023) EUS and ERCP training in Europe: time for simulation, optimization, and standardization. *United Eur Gastroenterol J* 11:407–409
- Wong JYY, Kongkam P, Ho KY (2017) Training in endoscopic ultrasonography: an Asian perspective. *Dig Endosc* 29:512–516
- Hedenström P, Marasco G, Eusebi LH, Lindkvist B, Sadik R (2021) GAPS-EUS: a new and reliable tool for the assessment of basic skills and performance in EUS among endosonography trainees. *BMJ Open Gastroenterol*. <https://doi.org/10.1136/bmjga-2021-000660>
- Wani S, Han S, Simon V, Hall M, Early D, Aagaard E, Abidi WM, Banerjee S, Baron TH, Bartel M, Bowman E, Brauer BC, Buscaglia JM, Carlin L, Chak A, Chatrath H, Choudhary A, Confer B, Coté GA, Das KK, DiMaio CJ, Dries AM, Edmundowicz SA, El Chafic AH, El Hajj I, Ellert S, Ferreira J, Gamboa A, Gan IS, Gangarosa L, Gannavarapu B, Gordon SR, Guda NM, Hammad HT, Harris C, Jalaj S, Jowell P, Kenshil S, Klapman J, Kochman ML, Komanduri S, Lang G, Lee LS, Loren DE, Lukens FJ, Mullady D, Muthusamy RV, Nett AS, Olyae MS, Paksresht K, Perera P, Pfau P, Piraka C, Poneros JM, Rastogi A, Razzak A, Riff B, Saligram S, Scheiman JM, Schuster I, Shah RJ, Sharma R, Spaete JP, Singh A, Sohail M, Sreenarasimhaiah J, Stevens T,

- Tabibian JH, Tzimas D, Uppal DS, Urayama S, Vitterbo D, Wang AY, Wassef W, Yachinski P, Zepeda-Gomez S, Zuchelli T, Keswani RN (2019) Setting minimum standards for training in EUS and ERCP: results from a prospective multicenter study evaluating learning curves and competence among advanced endoscopy trainees. *Gastrointest Endosc*. <https://doi.org/10.1016/j.gie.2019.01.030>
23. Yao L, Zhang C, Xu B, Yi S, Li J, Ding X, Yu H (2023) A deep learning-based system for mediastinum station localization in linear EUS (with video). *Endosc Ultrasound* 12:417–423
 24. Cui H, Zhao Y, Xiong S, Feng Y, Li P, Lv Y, Chen Q, Wang R, Xie P, Luo Z, Cheng S, Wang W, Li X, Xiong D, Cao X, Bai S, Yang A, Cheng B (2024) Diagnosing solid lesions in the pancreas with multimodal artificial intelligence: a randomized crossover trial. *JAMA Netw Open* 7:e2422454
 25. Weigel A, Caldas C, Meyer A, Morris SA (2022) The impact of AI on research. *Cell* 185:2621–2622
 26. Hirai K, Kuwahara T, Furukawa K, Kakushima N, Furune S, Yamamoto H, Marukawa T, Asai H, Matsui K, Sasaki Y, Sakai D, Yamada K, Nishikawa T, Hayashi D, Obayashi T, Komiyama T, Ishikawa E, Sawada T, Maeda K, Yamamura T, Ishikawa T, Ohno E, Nakamura M, Kawashima H, Ishigami M, Fujishiro M (2022) Artificial intelligence-based diagnosis of upper gastrointestinal subepithelial lesions on endoscopic ultrasonography images. *Gastric Cancer* 25:382–391
 27. Redmon J, Divvala S, Girshick R, Farhadi A (2016) You only look once: Unified, real-time object detection. 2016 IEEE conference on computer vision and pattern recognition (CVPR), pp 779–788
 28. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen L-C (2018) MobileNetV2: inverted residuals and linear bottlenecks. 2018 IEEE conference on computer vision and pattern recognition (CVPR)
 29. Yang D, Wagh MS, Draganov PV (2020) The status of training in new technologies in advanced endoscopy: from defining competence to credentialing and privileging. *Gastrointest Endosc* 92:1016–1025
 30. Zhang D, Wu C, Yang Z, Yin H, Liu Y, Li W, Huang H, Jin Z (2024) The application of artificial intelligence in EUS. *Endosc Ultrasound* 13:65–75
 31. Ye XH, Zhao LL, Wang L (2022) Diagnostic accuracy of endoscopic ultrasound with artificial intelligence for gastrointestinal stromal tumors: a meta-analysis. *J Dig Dis* 23:253–261
 32. Goto O, Kaise M, Iwakiri K (2022) Advancements in the diagnosis of gastric subepithelial tumors. *Gut Liver* 16:321–330
 33. Minoda Y, Ihara E, Fujimori N, Nagatomo S, Esaki M, Hata Y, Bai X, Tanaka Y, Ogino H, Chinen T, Hu Q, Oki E, Yamamoto H, Ogawa Y (2022) Efficacy of ultrasound endoscopy with artificial intelligence for the differential diagnosis of non-gastric gastrointestinal stromal tumors. *Sci Rep* 12:16640
 34. Li X, Zhang C, Yao L, Zhang J, Zhang K, Feng H, Yu H (2023) A deep learning-based system to identify originating mural layer of upper gastrointestinal submucosal tumors under EUS. *Endosc Ultrasound* 12:465–471

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.