Received: 9 August 2024 | Revised: 18 November 2024 | Accepted: 25 November 2024

DOI: 10.1111/cas.16426

Cancer Science WILEY

ORIGINAL ARTICLE

Deep learning detected histological differences between invasive and non-invasive areas of early esophageal cancer

Akiko Urabe[1,2] | Masahiro Adachi[1] | Naoya Sakamoto[3] | Motohiro Kojima[1] | Shumpei Ishikawa[3] | Genichiro Ishii[1] | Tomonori Yano[2] | Shingo Sakashita[3]

[1]Department of Pathology and Clinical Laboratories, National Cancer Center Hospital East, Kashiwa, Chiba, Japan

[2]Department of Gastroenterology and Endoscopy, National Cancer Center Hospital East, Kashiwa, Chiba, Japan

[3]Division of Pathology, Exploratory Oncology Research & Clinical Trial Center, National Cancer Center, Kashiwa, Chiba, Japan

Correspondence
Shingo Sakashita, Division of Pathology, The Exploratory Oncology Research & Clinical Trial Center, National Cancer Center, 6-5-1, Kashiwanoha, Kashiwa, Chiba 277-8577, Japan.
Email: ssakashi@east.ncc.go.jp

Funding information
National Cancer Center Research and Development Fund, Grant/Award Number: 2021-A-07; JSPS KAKENHI, Grant/Award Number: JP20K22859 and JP21K06899

Abstract

The depth of invasion plays a critical role in predicting the prognosis of early esophageal cancer, but the reasons behind invasion and the changes occurring in invasive areas are still not well understood. This study aimed to explore the morphological differences between invasive and non-invasive areas in early esophageal cancer specimens that have undergone endoscopic submucosal dissection (ESD), using artificial intelligence (AI) to shed light on the underlying mechanisms. In this study, data from 75 patients with esophageal squamous cell carcinoma (ESCC) were analyzed and endoscopic assessments were conducted to determine submucosal (SM) invasion. An AI model, specifically a Clustering-constrained Attention Multiple Instance Learning model (CLAM), was developed to predict the depth of cancer by training on surface histological images taken from both invasive and non-invasive regions. The AI model highlighted specific image portions, or patches, which were further examined to identify morphological differences between the two types of areas. The 256-pixel AI model demonstrated an average area under the receiver operating characteristic curve (AUC) value of 0.869 and an accuracy (ACC) of 0.788. The analysis of the AI-identified patches revealed that regions with invasion (SM) exhibited greater vascularity compared with non-invasive regions (epithelial). The invasive patches were characterized by a significant increase in the number and size of blood vessels, as well as a higher count of red blood cells (all with p-values <0.001). In conclusion, this study demonstrated that AI could identify critical differences in surface histopathology between non-invasive and invasive regions, particularly highlighting a higher number and larger size of blood vessels in invasive areas.

KEYWORDS
artificial intelligence, epithelium, esophageal cancer, submucosa, surface histomorphology

Abbreviations: ACC, accuracy; AI, Artificial intelligence; AUC, area under the receiver operating characteristic curve; CLAM, Clustering-constrained Attention Multiple Instance Learning; CRT, chemoradiotherapy; EP, epithelial; ER, endoscopic resection; ESCC, Esophageal squamous cell carcinoma; ESD, endoscopic submucosal dissection; MIL, multi-instance learning; NBI, narrow-band imaging; SM, submucosal; WSI, whole slide imaging.

1 | INTRODUCTION

According to the Global Cancer Observatory 2020,[1] esophageal cancer accounts for 3.1% of new cancer cases globally and is the sixth leading cause of cancer-related deaths. ESCC or adenocarcinoma is the most common esophageal cancer worldwide.[2]

Treatment options for esophageal cancer include surgery (including endoscopic resection (ER)), chemoradiotherapy (CRT), and chemotherapy. Of these options, surgery and CRT are the most curative treatments, whereas ER is the least invasive. ER is appropriate for early-stage T1a cancers limited to the epithelium (T1a-EP) or superficial lamina propria mucosa (T1a-LPM), which have a low risk of lymph node metastasis. The indication for ER is expanded to patients with more advanced T1a lesions that extend deep into the lamina propria (T1a-MM) or T1b lesions with minor SM invasion (<200 μm). The risk of lymph node metastasis is 10%–20%. In these cases, additional CRT or surgery is performed for T1b or greater depths with pathological evaluation. The decision between endoscopic and surgical resection depends on the invasion depth, risk factors, and extent of lymph node metastasis.[3]

The rate of lymph node metastasis varies with the depth of involvement and the treatment required. In other words, the prognosis depends on the depth of invasion. However, there is limited knowledge of why invasion occurs and why it is associated with a poor prognosis. In this study, we would like to help clarify the mechanism by looking at the differences in the tissues of the non-invasive and invasive areas.

AI has become a widespread tool used also in endoscopy and pathology.[4–10] Digital pathology images are used for pathological diagnosis.[11] Particularly, WSI, which refers to the digitized images of glass-slide specimens, enables microscopic observations on a personal computer using WSI-specific analysis software. These images are also used for tissue analysis, image processing, and annotation.[12]

CLAM[13,14] is a weakly supervised machine learning model developed for histopathological image classification. The model combines MIL with an attention branch network. MIL has been applied in pathological diagnosis because it allows classification on a case-by-case basis, rather than on a patch-by-patch basis.[15–18] Attention branch network refers to the mechanism through which a neural network focuses on a specific part of the input data and assigns different levels of importance to different elements. Thus, the model not only classifies but also extracts patches that significantly influence classification. In this study, we applied the CLAM model to investigate potential differences in the histopathology of superficial tissues between areas with and without invasion in esophageal tumors. In particular, using endoscopically resected ESCC specimens at SM depth, we aimed to investigate whether the model could predict depth by training AI on superficial tissue samples from regions without and with invasion (referred to as EP and SM, respectively) and, if so, which areas to focus on.

2 | MATERIALS AND METHODS

2.1 | Cases and specimen

We included 75 patients (Table S1) who underwent ESD for ESCC with a pathological depth of SM1 or SM2 at the National Cancer Center Hospital East, between January 1, 2019, and March 15, 2023. Cases without superficial exposure and cases with slides borrowed from other hospitals ($n = 10$) were excluded. This study was approved by the Ethical Review Committee of the National Cancer Center Hospital East (2023-029).

2.2 | Sampling

First, ESD specimens were fixed in 10% formalin and sectioned parallel to the short axis in 4-mm slices, which were used to prepare H&E-stained sections. Next, H&E slides containing the EP and SM portions in subject cases were selected. For EP assessment, we then selected the sections with the EP portion farthest from the SM portion of the same lesion. In cases where it was difficult to determine the furthest portion at both ends, both slides were used. Then, when more than one SM portion was present, all regions were used for evaluation. In total, 156 slides were used (EP, 82; SM, 74). Finally, the specimens were scanned at ×40 magnification using NanoZoomer s360 (Hamamatsu Photonics, Shizuoka, Japan) to obtain WSIs.

2.3 | Annotation

To examine the differences in surface histology between the EP and SM sections, we annotated the surface portion of the lesion on each slide and extracted histology using QuPath version 0.4.0, a publicly available annotation tool for digital slides[19] (Figure S1). The annotation sites in this study were cancerous in the superficial layer for both the EP and SM regions. The annotation size was limited to a maximum of 4 mm, matching typical biopsy specimen dimensions; hence, the AI analysis would be relevant for clinical biopsy samples. For SM invasion, the maximum depth from the superficial mucosal layer was always annotated. For SM lesions with multiple deep portions below the MMs, each distinct area was separately annotated. The most superficial cancerous mucosal sections located furthest from areas of maximum-depth invasion were selected and annotated for non-invasive lesions. For each lesion, the superficial mucosal layers from the sections with and without invasion were extracted and utilized as training data. Two pathologists, including a specialist, double-checked the annotation sites to ensure accuracy.

2.4 | Creating AI models

The CLAM model was used for AI development. When preprocessing the slide images, CLAM segmented the image into small patches. The annotated slide images were cropped into 128×128-, 256×256-, and 512×512-pixel regions to create patches (Figure 1A). The sections with SM invasion were labeled as 1 (SM), and the EP sections were labeled as 0 for model training. For model development and evaluation, a 10-fold Monte Carlo cross-validation strategy was used, in which the training, validation, and testing subsets were randomly derived from the cohort. Each fold was divided randomly into training (80% of the cases), validation (10%), and testing (10%) sets. Performance was assessed using the area under the receiver operating characteristic curve (AUC) and ACC. The AUC and ACC values were compared to determine the size (128, 256, or 512 pixels) used to create the patches. The model was trained using the adaptive moment estimation optimizer with a learning rate of 2×10^{-4}. We used the default algorithm for other parameters and did not perform data augmentation. The training process ended at 200 epochs if validation loss did not decrease from its previous minimum for 20 consecutive epochs.

2.5 | Feature extraction

For each patch used for inference in CLAM, an attention score was calculated based on the proportion of contributions to the prediction. The patch with the highest attention score in the model and the best AUC value was extracted. They were then examined for features based on which the inference model classified them as SM or EP. In this study, the total area of the nuclei, the major and minor diameters and circumference of cancer-cell nuclei, the presence of blood vessels, and the presence or absence of distinct nucleoli in each patch were examined. In patches with blood vessels, the total area and number of erythrocytes and variability in findings related to nuclei were also evaluated. The presence of blood vessels and distinct nucleoli were assessed visually. For erythrocytes and vascular cavities, manual annotation was performed and analyzed using ImageJ.[19] The last analysis of tumor nuclear parameters was performed after manual annotation of the tumor area using QuPath.[20] Regarding the work performed by people, all tasks were conducted by two pathologists.

2.6 | Statistical analyses

The analysis was performed on an Ubuntu 20.04 Linux system with an A100 GPU (NVIDIA, Santa Clara, CA, USA). Statistical analysis was performed using R version 4.3.0 (The R Project, Vienna, Austria), with a chi-squared test for categorical variables and a non-operating characteristic (ROC) curves were equal variance *t*-test (Welch test) for continuous variables for comparison between the two counts. The *F*-test was used to test for variance. Logistic regression analysis was performed for multivariate analysis. *p*-values <0.05 were considered statistically significant.

2.7 | Inference for LPM and MM cases

We examined the possibility of EP- and SM-like features extracted by the AI used to assign EP or SM. For MM, 89 cases diagnosed by ESD between January 1, 2019, and March 15, 2023, were utilized. As with SM cases, the superficial layer of the deepest esophageal cancer portion was annotated, with multiple annotations if multiple deep spots were present. A total of 103 muscularis mucosa (MM) slides were annotated. Because there were many LPM cases, 84 LPM cases diagnosed by ESD between January 1, 2020, and March 15, 2023, were used. In total, 103 LPM slides were annotated.

3 | RESULTS

3.1 | AI model development

Based on CLAM, a prediction model for tumor depth (EP or SM) was created from surface histology. In total, 128-, 256-, and 512-pixel models were created (Figure 1A), and the test AUC and ACC values were compared. The mean test AUC values for 128, 256, and 512 pixels were 0.69, 0.87, and 0.72, respectively (Figure 1B). The mean test ACC values were 128, 256, and 512 pixels were 0.61, 0.79, and 0.64, respectively (Figure 1C). The 256-pixel model was superior to the 128- and 512-pixel models in terms of both test AUC and ACC values; therefore, we created 256-pixel models. All the 10 models trained at 256 pixels achieved AUC >0.68 and ACC >0.68 on the test set (Table 1). The models with the highest AUC (1.0) and ACC (0.938) values for the test data were selected for further analysis. ROC curves were drawn using matplotlib, a Python library.

3.2 | Attention score

For each patch, the CLAM model calculated an attention score based on its contribution to prediction. Representative patches that received the highest scores were extracted from the CLAM model with the best AUC performance (256-pixel model). These high-scoring patches (Figures 2 and S2) were assumed to contain key distinguishing features, dense vasculature, and red blood cell numbers in the SM patches (Figures 2B and S2) versus higher nuclei numbers in the EP patches (Figures 2A and S2), and were used for further histological evaluation.

3.3 | Histopathological morphological analysis

The attention score was utilized to extract patches with histological features that the AI model for depth prediction focused on. Significant

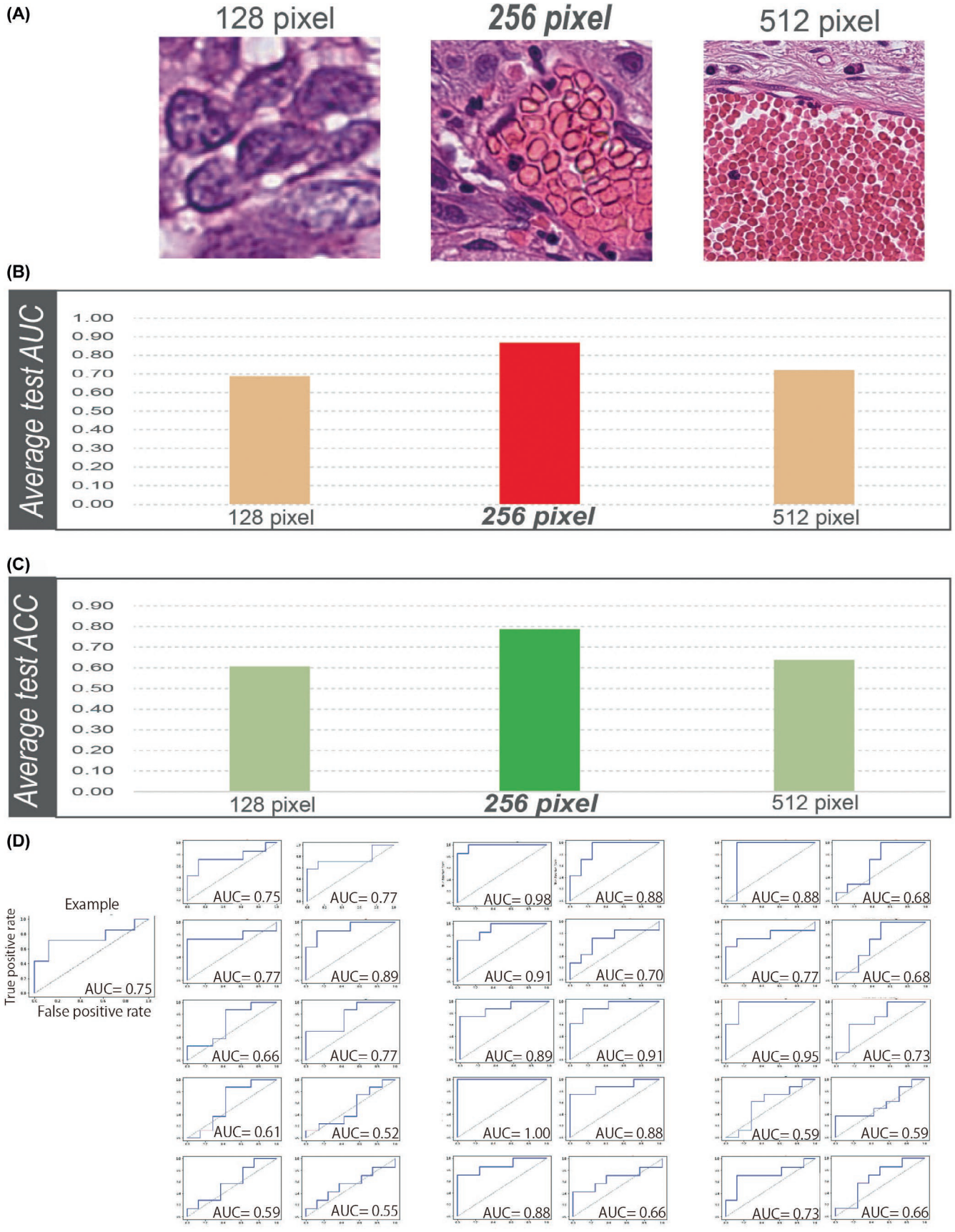

FIGURE 1 Difference in artificial intelligence (AI) accuracy by image size. (A) Three different image sizes (128, 256, and 512 pixels) were used for AI training. Comparison of AI inference using test images. (B) Average values for the area under the curve (AUC) and (C) average accuracy (ACC) estimations, using three image sizes. ROC curves generated 10 times for each image size (D).

TABLE 1 Ten models created by 10-fold Monte Carlo cross-validation using Clustering-constrained Attention Multiple Instance Learning.

Model	test_auc	val_auc	test_acc	val_acc
1	0.969	0.938	0.938	0.875
2	0.891	0.688	0.813	0.625
3	0.875	0.984	0.688	0.938
4[a]	1.000	1.000	0.938	1.000
5	0.859	0.828	0.813	0.688
6	0.906	0.984	0.688	0.875
7	0.703	0.922	0.750	0.813
8	0.906	0.875	0.688	0.813
9	0.891	0.953	0.875	0.938
10	0.688	1.000	0.688	0.938
Average	0.869	0.917	0.788	0.850

[a]Used model because of highest test AUC and ACC.

differences were observed between the EP and SM patches in terms of the presence of blood vessels. However, no significant differences were observed in nucleolus clarity. The AI model focused on patches without blood vessels in the EP region, while the proportion of blood vessels increased in the SM region (Table 2 and Figure S3).

3.4 | Analysis of blood vasculature

As the presence of blood vessels was a significant feature of the AI model, further comparisons were made between the SM and EP patches that contained vessels (Table 2). The total blood vessel area was significantly larger in SM patches than in EP patches ($p < 10^{-5}$; Figure 3A,B). Similarly, red blood cell counts were significantly higher in SM patches than in EP patches ($p < 10^{-5}$; Figure 3A,C). These findings indicate that the AI model may predict invasion depth by focusing on differences in blood vessel morphology between superficial and deep areas.

3.5 | Nuclear variability

Morphological variations in the nuclei (Figure 4A) in the EP and SM patches were compared using the AI high-interest images. *T*-tests were used to compare mean values related to the size of nuclei, including area, maximum and minimum diameter, and circumference. *F*-tests were used to assess differences in the variance. Although none of the results were significantly different with respect to the mean, variations in the area and minimum diameter of nuclei were significantly different in the EP and SM samples ($F_{81,71} = 0.542$, $p = 0.004$, and $F_{81,70} = 0.669$, $p = 0.04$, respectively; Figure 4B–E). These results suggest that regions with invasive SM characteristics may contain more diverse micronuclei and larger nuclei than EP regions.

3.6 | Inference for LPM and MM cases

In the previous study of EP and SM lesions, the probability of being determined as EP was 97.6% and 8.6%, respectively, whereas, in the present study, 57.3% of LPM and 37.9% of MM were classified as EP (Figure 5A). EP and SM patches are presented in Figure 5B. SM patches contained more blood vessels than EP patches as well as when SM and EP were compared.

4 | DISCUSSION

We demonstrated that AI could predict the depth of invasion from surface histological findings alone. We also obtained an understanding of the structural differences between infiltrated and non-infiltrated areas.

This study has three major implications. First, AI was able to not only classify, but also identify differences in morphology that were difficult for pathologists to notice. For instance, the AI model focused on patches of increased vascular density in the SM region; however, this is not a point of particular interest to many pathologists when diagnosing biopsy material.

Second, the ability of AI to infer SM invasion from the evaluation of surface tissues suggests that detectable changes have already occurred in these layers. Although endoscopic observation is limited to the surface epithelium, invasion depth can potentially be predicted through careful examination of surface histological features. Thus, superficial tissues may contain subtle evidence indicative of deep SM spread that AI models can leverage for diagnosis and staging.

Third, a detailed analysis of the patches with high attention scores on which the AI model focused revealed an emphasis on blood vessel features rather than cancer-cell morphology. This study used CLAM to determine the patches that contributed the most to model predictions. CLAM learns by extracting only the most informative patches with high attention scores. Examination of these patches provided insight into the histological features that the model used to predict invasion depth. The prominence of blood vessel differences suggests that the AI model may predict depth by focusing on vascular changes between areas with and without invasion. This is a remarkable finding because pathologists traditionally do not emphasize vascular morphology in any cancer staging system. However, pronounced vascular changes, visible even in superficial tissues, may serve as key signals for predicting underlying invasion. AI revealed overlooked histological differences with clear diagnostic significance. The key difference between areas with and without invasion that enabled AI-based depth prediction was quantitative information about the tumor stroma, particularly vascular morphology.

Although these observations regarding vascular architecture are interesting pathological findings, endoscopists routinely use vascular patterns to predict invasion depth based on NBI or other image-enhanced endoscopy with magnification.[21,22] The current results suggest that endoscopic visualization of vascular changes may reflect underlying histopathology.

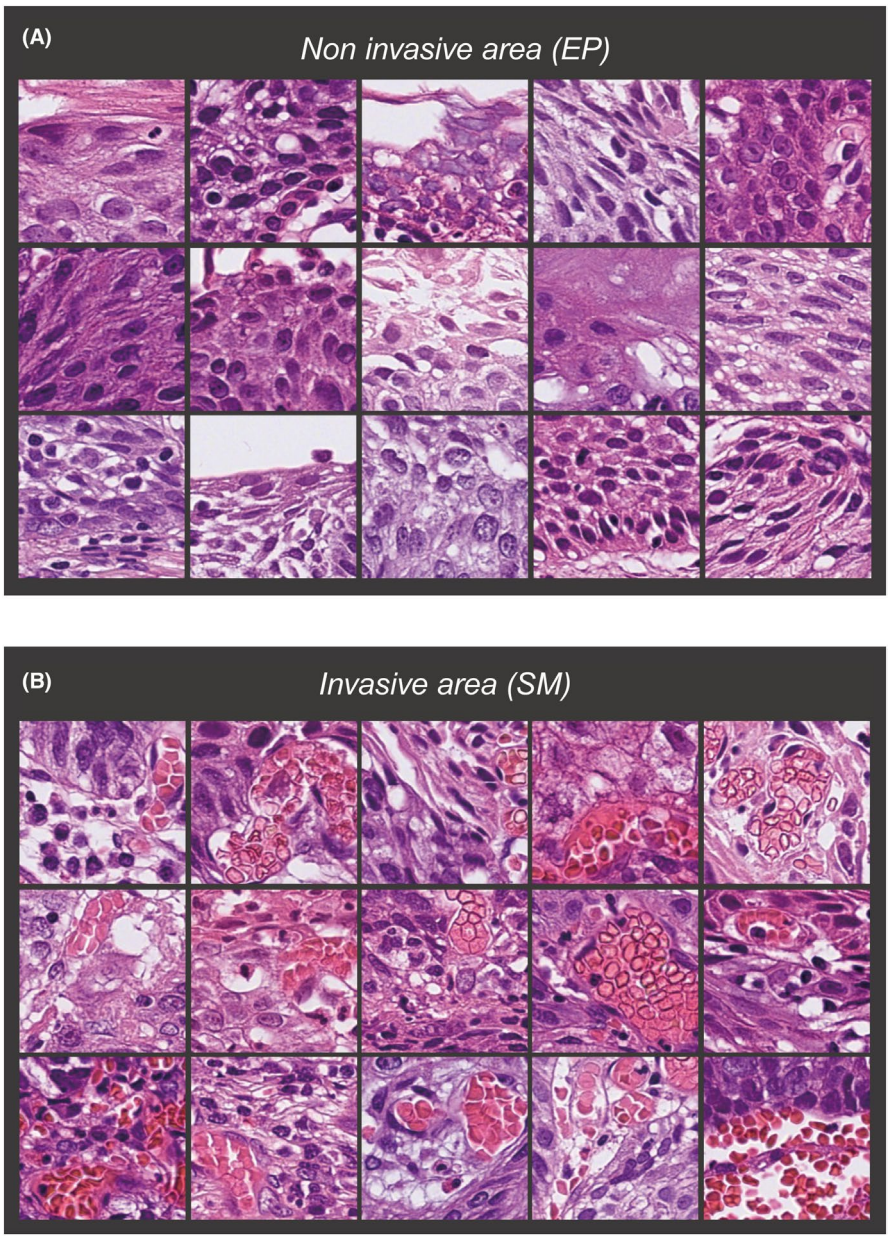

FIGURE 2 Representative examples of patch images extracted by artificial intelligence. (A) Non-invasive area. (B) Invasive area.

When metrics related to the size of nuclei were compared, no differences were found in size or diameter, but variation analysis revealed differences in heterogeneity, whereby regions with invasion showed a greater diversity of small and large nuclei.

A discussion of biological mechanisms is given here as an addition. It is said that the malignant transformation of cancer is essentially caused by changes in cancer cells, such as the accumulation of genetic mutations. In the present results, although the variation in size was observed in cancer cells between non-invasive and invasive areas, there was no significant difference in nuclear size, long diameter, or roundness, and a difference was observed in blood vessels. This suggests that vascular changes may be involved in cancer invasion. Although it is unclear whether this is the cause or the result, the fact that similar characteristics are also observed to some extent in LPM and MM cancers, even if not SM, suggests that changes in the surrounding area, such as blood vessels, may be responsible for the invasion.[23–26]

Ultra-high magnification endocytoscopy[27,28] allows up to 580-fold magnification compared with the previous maxima of ×80–100 to evaluate histological findings without biopsy. Images captured by the endocytoscopic system have been used to evaluate the differential diagnosis of ESCC and benign lesions using AI, and favorable preliminary results have been reported.[29] While evaluating the invasion depth in ESCC using this system, patterns regarding blood vessels and individual cells should be resolved, even by detecting previously invisible nuclei. The histological features highlighted by AI in the present study, such as vascular and nuclear patterns, may be observed using endoscopy in the future. Further studies will be required, as ultra-high-resolution endoscopy develops, to determine the optimal diagnostic criteria based on histopathological correlations, such as those demonstrated in this study.

The detection of differences that pathologists did not observe using CLAM is significant. The CLAM model learns only from patch

TABLE 2 Comparison of the presence of blood vessels, nucleolus clarity and tumor nucleus in the patch images that artificial intelligence focused on the most.

		EP($n=82$)	SM($n=74$)	p
Vessel	Positive, *n* (%)	20 (32.3)	58 (78.4)	<0.001*
	Negative, *n* (%)	62 (67.7)	16 (21.6)	
Nucleolus	Clear, *n* (%)	27 (32.9)	17 (23.0)	0.167
	Non-clear, *n* (%)	55 (67.1)	57 (77.0)	
Tumor nuclear	Tumor nuclear area, mean (Range), pixel	636.017 (287–1296)	589.857 (296–1695)	0.426
	Max diameter, mean (Range), pixel	42.632 (26.6–61.0)	40.416 (27.8–60.1)	0.181
	Minimum diameter, mean (Range), pixel	20.370 (9.8–31.1)	20.320 (14.6–38.0)	0.448
	Circumference, mean (Range), pixel	105.068 (67.7–150)	99.514 (70.1–154.8)	0.280
	The number of nuclear, mean (Range), pixel	31.500 (8–72)	18.000 (3–73)	<0.001*

Note: Values are presented as *n* (%). *Statistical significance ($p<0.05$).

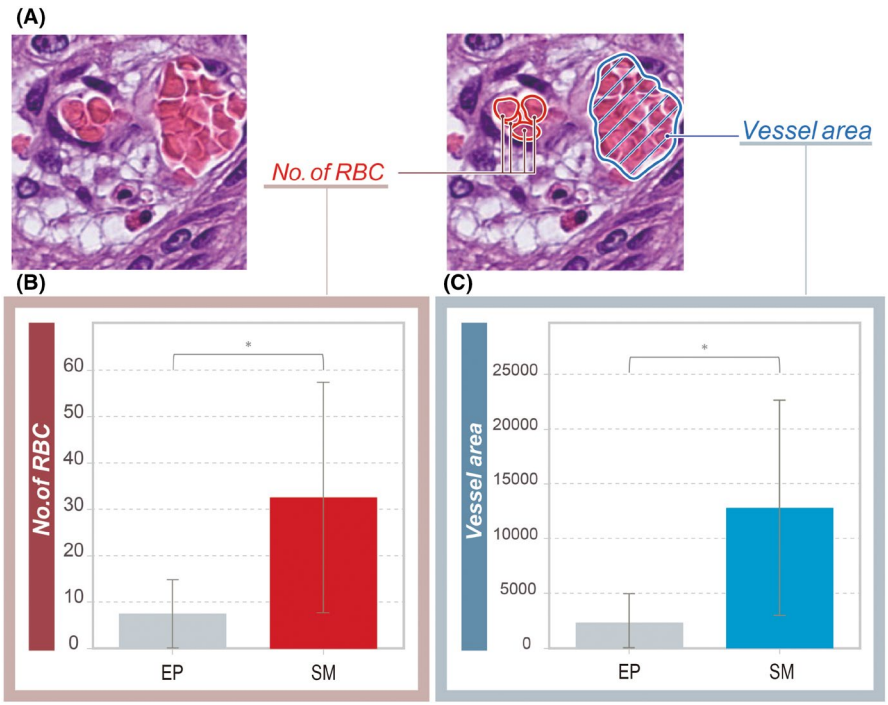

FIGURE 3 Morphological analysis of blood vessels. (A) Estimation of the number of red blood cells (RBCs) (left) and blood vessel area (right). Comparison of (B) the number of RBCs and (C) blood vessel area in regions with (submucosal, SM) and without (epithelial, EP) tumor invasion.

images, which improves accuracy, and incorrect patches do not advance learning. By analyzing the patches with high attention scores in an accurately trained model, the histological differences between EP and SM samples that drive model predictions are revealed. Thus, we used AI's patch-level focus as a tool to uncover distinguishing features that a pathologist may not be aware of. AI guides us to subtle but predictive vascular differences between areas with and without invasion that pathologists do not routinely examine but that can aid in diagnosis. This demonstrates that the interpretation of an optimized AI model can reveal new perspectives on diagnostic tissue morphology that are invisible to the human eye.

This study demonstrates the potential of biopsy histopathology to quantitatively diagnose the depth of invasion and provide a qualitative diagnosis of cancer. Although biopsies typically only provide information on whether cancer is present, the ability to determine depth from biopsy analysis can better guide clinical decision-making and treatment planning. However, challenges remain in translating these findings into clinical biopsy interpretation, such as uncertainty regarding whether the biopsy sample adequately captures the invasive tumor front. In this study, the AI model assessed the ability to accurately diagnose the depth of tumor invasion in the histology of biopsy specimens but was not applicable to biopsy material.

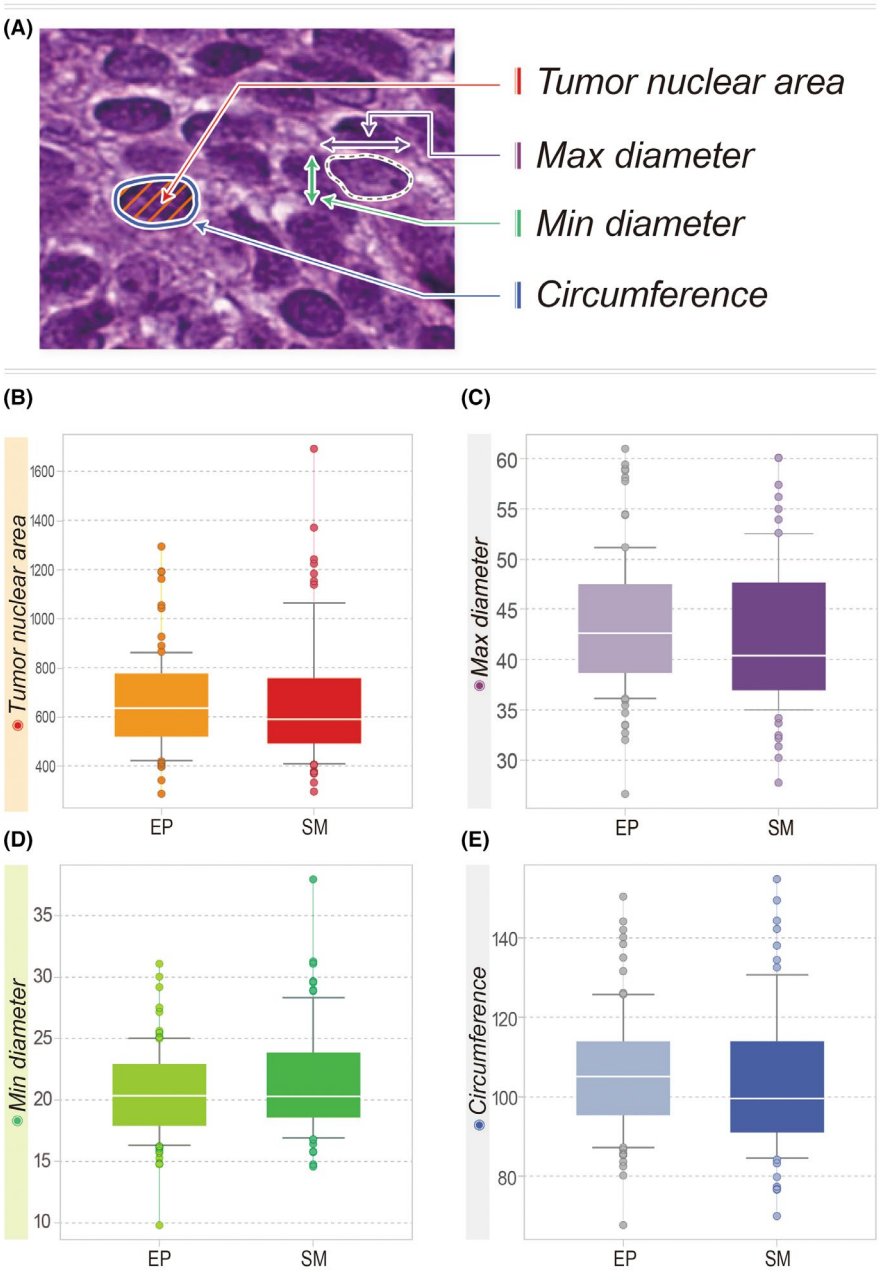

FIGURE 4 Morphological analysis of tumor-cell nuclei. (A) Description of the measurements used to calculate the area of tumor nuclei area, maximum and minimum diameters, and circumference. Comparison of characteristics in cases without (EP) and with (SM) invasion. (B) Area of tumor nuclei. (C) Maximum diameter. (D) Minimum diameter. (E) Circumference.

Specifically, it was tested using pathology images from 89 cases of biopsy material collected between 2020 and 2023, and their depth was determined through subsequent ESD and surgery. However, limited findings were obtained (Figure S4), with an AUC value of 0.49 indicating that the model's performance was comparable with random guessing and could not effectively distinguish between the two classes. This is due to differences in how ESD and biopsy specimens are viewed, potentially affecting the model's adaptability. Nevertheless, this study represents an important step toward leveraging standard biopsy procedures for more precise and quantitative staging through AI-assisted pathological assessment of surface tissues.

This study has limitations. First, it aimed to identify histological differences and did not validate a system that can predict depth using AI. Hence, multicenter studies and the creation of independent test datasets are necessary. Second, we focused on blood vessels and found corresponding differences but we could not exclude the possibility that the pathologist may have missed other differences. Last, although we used a portion of the tissue, we did not evaluate the entire cancer superficially, and the possibility that selection bias due to the annotation site may have affected the results cannot be ruled out.

In conclusion, this study demonstrated the potential of AI-based analysis of superficial histology to predict tumor invasion depth. The most important predictive features were the pronounced vascular differences between invasive and non-invasive stromal areas, rather than epithelial tumor-cell features. Therefore, standard biopsy pathology can be used for quantitative

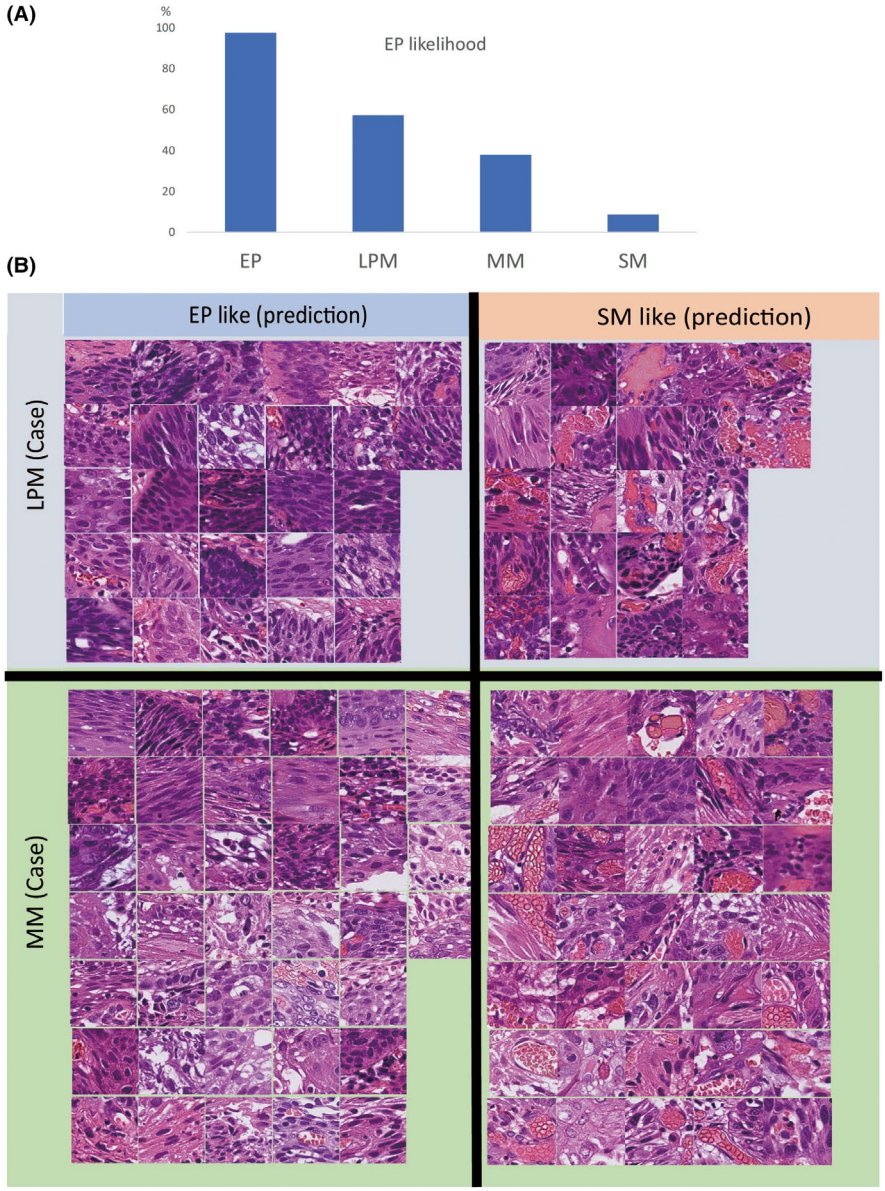

FIGURE 5 Inference for lamina propria mucosa (LPM) and muscularis mucosa (MM) cases using the learning model with the highest accuracy. (A) The probability of being determined as epithelial (EP, without invasion) in each depth. (B) Most noteworthy patches are divided between submucosal (SM) and epithelial (EP).

assessments of the invasion status. More broadly, this study highlights the underappreciated stromal vascular changes associated with invasion that may reflect important biological differences in tumor aggressiveness.

AUTHOR CONTRIBUTIONS

Akiko Urabe: Conceptualization; data curation; formal analysis; investigation; methodology; resources; software; supervision; validation; visualization; writing – original draft; writing – review and editing. **Masahiro Adachi:** Data curation; formal analysis; methodology; writing – review and editing. **Naoya Sakamoto:** Writing – review and editing. **Motohiro Kojima:** Writing – review and editing. **Shumpei Ishikawa:** Methodology; writing – review and editing. **Genichiro Ishii:** Writing – review and editing. **Tomonori Yano:** Writing – review and editing. **Shingo Sakashita:** Conceptualization; funding acquisition; methodology; project administration; software; validation; visualization; writing – review and editing.

ACKNOWLEDGMENTS

We would like to thank Editage (www.editage.jp) for English language editing.

FUNDING INFORMATION

This work was supported by JSPS KAKENHI (grant nos. JP20K22859 and JP21K06899). A part of this study was supported by the National Cancer Center Research and Development Fund (grant no. 2021-A-07).

CONFLICT OF INTEREST STATEMENT

Dr. Genichiro Ishii is an Editorial Board member of *Cancer Science*. Other authors do not have any COI to declare. Ishii Genichiro received research grants from Daiichi Sankyo, Inc., Ono Pharmaceutical Co., Ltd., Noile-Immune Biotech, Takeda Pharmaceutical Company Limited, Sumitomo Dainippon Pharma Co., Ltd., Nihon Medi-Physics Co., Ltd., and Indivumed GmbH, H.U. Group Research Institute

and consulting fee from Takeda Pharmaceutical Company Limited. Tomonori Yano received lecture fees and research grants from Olympus.

ETHICS STATEMENT

Approval of the research protocol by an Institutional Reviewer Board: This study was approved by the Ethical Review Committee of the National Cancer Center Hospital East (2023-029) and conforms to the provisions of the Declaration of Helsinki.

Informed Consent: N/A.

Registry and the Registration No. of the study/trial: N/A.

Animal Studies: N/A.

ORCID

Naoya Sakamoto https://orcid.org/0000-0001-6273-0189
Motohiro Kojima https://orcid.org/0000-0002-6150-6545
Genichiro Ishii https://orcid.org/0000-0001-8862-5811
Tomonori Yano https://orcid.org/0000-0002-8030-1449
Shingo Sakashita https://orcid.org/0000-0003-1133-3313

REFERENCES

1. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71(3):209-249. doi:10.3322/caac.21660
2. Napier KJ, Scheerer M, Misra S. Esophageal cancer: a review of epidemiology, pathogenesis, staging workup and treatment modalities. *World J Gastrointest Oncol*. 2014;6(5):112. doi:10.4251/wjgo.v6.i5.112
3. Ishihara R, Arima M, Iizuka T, et al. Endoscopic submucosal dissection/endoscopic mucosal resection guidelines for esophageal cancer. *Dig Endosc*. 2020;32:452-493. doi:10.1111/den.13654
4. Shiroma S, Yoshio T, Kato 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(1):7759. doi:10.1038/s41598-021-87405-6
5. Horie Y, Yoshio T, Aoyama K, et al. Diagnostic outcomes of esophageal cancer by artificial intelligence using convolutional neural networks. *Gastrointest Endosc*. 2019;89(1):25-32. doi:10.1016/j.gie.2018.07.037
6. Tajiri A, Ishihara R, Kato Y, et al. Utility of an artificial intelligence system for classification of esophageal lesions when simulating its clinical use. *Sci Rep*. 2022;12(1):6677. doi:10.1038/s41598-022-10739-2
7. Visaggi P, Barberio B, Gregori D, et al. Systematic review with meta-analysis: artificial intelligence in the diagnosis of oesophageal diseases. *Aliment Pharmacol Ther*. 2022;55(5):528-540. doi:10.1111/apt.16778
8. Tokai Y, Yoshio T, Fujisaki J. Development of artificial intelligence for the detection and staging of esophageal cancer. *Ann Esophagus*. 2023;6:3.
9. Schüffler P, Steiger K, Weichert W. How to use AI in pathology. *Genes Chromosomes Cancer*. 2023;62(9):564-567. doi:10.1002/gcc.23178
10. Berbís MA, McClintock DS, Bychkov A, et al. Computational pathology in 2030: a Delphi study forecasting the role of AI in pathology within the next decade. *EBioMedicine*. 2023;88:104427. doi:10.1016/j.ebiom.2022.104427
11. Nakagawa K, Moukheiber L, Celi LA, et al. AI in pathology: what could possibly go wrong? *Semin Diagn Pathol*. 2023;40(2):100-108. doi:10.1053/j.semdp.2023.02.006
12. Wu Y, Cheng M, Huang S, et al. Recent advances of deep learning for computational histopathology: principles and applications. *Cancers (Basel)*. 2022;14(5):1199. doi:10.3390/cancers14051199
13. Zheng T, Chen W, Li S, et al. A deep reinforcement learning framework for rapid diagnosis of whole slide pathological images. arXiv:220502850 [preprint] 2022 Available from: https://arxiv.org/abs/2205.02850 (Accessed month/date/year)
14. Lu MY, Williamson DFK, Chen TY, Chen RJ, Barbieri M, Mahmood F. Data-efficient and weakly supervised computational pathology on whole-slide images. *Nat Biomed Eng*. 2021;5(6):555-570. doi:10.1038/s41551-020-00682-w
15. Zeng Q, Klein C, Caruso S, et al. Artificial intelligence predicts immune and inflammatory gene signatures directly from hepatocellular carcinoma histology. *J Hepatol*. 2022;77(1):116-127. doi:10.1016/j.jhep.2022.01.018
16. Ye Q, Wan F, Liu C, Huang Q, Ji X. Continuation multiple instance learning for weakly and fully supervised object detection. *IEEE Trans Neural Netw Learn Syst*. 2022;33(10):5452-5466. doi:10.1109/TNNLS.2021.3070801
17. Javed SA, Juyal D, Padigela H, Taylor-Weiner A, Yu L, Prakash A. Additive MIL: intrinsically interpretable multiple instance learning for pathology. *Adv Neural Inf Proces Syst*. 2022;35:20689-20702.
18. Jiang S, Suriawinata AA, Hassanpour S. MHAttnSurv: multi-head attention for survival prediction using whole-slide pathology images. *Comput Biol Med*. 2023;158:106883. doi:10.1016/j.compbiomed.2023.10688319
19. Schneider CA, Rasband WS, Eliceiri KW. NIH image to ImageJ: 25 years of image analysis. *Nat Methods*. 2012;9:671-675. doi:10.1038/nmeth.2089
20. QuPath: the global impact of an open source digital pathology system. Comput Struct. *Biotechnol J*. 2021;19:852-859. doi:10.1016/j.csbj.2021.01.022
21. Oyama T, Inoue H, Arima M, 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(2):105-112. doi:10.1007/s10388-016-0527-7
22. Rha EY, Kim JM, Yoo G. Volume measurement of various tissues using the image J software. *J Craniofac Surg*. 2015;26(6):e505-e506. doi:10.1097/SCS.0000000000002022
23. Cao S, Wang X, Liu X, et al. Integrative analysis of angiogenesis-related long non-coding RNA and identification of a six-DEARlncRNA signature associated with prognosis and therapeutic response in esophageal squamous cell carcinoma. *Cancers (Basel)*. 2022;14(17):4195. doi:10.3390/cancers14174195
24. Kumagai Y, Sobajima J, Higashi M, et al. Angiogenesis in superficial esophageal squamous cell carcinoma: assessment of microvessel density based on immunostaining for CD34 and CD105. *Jpn J Clin Oncol*. 2014;44(6):526-533. doi:10.1093/jjco/hyu039
25. Kumagai Y, Tachikawa T, Higashi M, et al. Chondromodulin-1 and vascular endothelial growth factor-a expression in esophageal squamous cell carcinoma: accelerator and brake theory for angiogenesis at the early stage of cancer progression. *Esophagus*. 2020;17(2):159-167. doi:10.1007/s10388-019-00695-8
26. Schiffmann LM, Plum PS, Fuchs HF, Babic B, Bruns CJ, Schmidt T. Tumor microenvironment of esophageal cancer. *Cancer*. 2021;13(18):4678.
27. Tsai CL, Mukundan A, Chung CS, et al. Hyperspectral imaging combined with artificial intelligence in the early detection of esophageal cancer. *Cancers (Basel)*. 2021;13(18):4593. doi:10.3390/cancers13184593

28. Kubo K, Fujino MA. Ultra-high magnification endoscopy of the normal esophageal mucosa. *Gastrointest Endosc*. 1997;46(1):96-97.
29. Kumagai Y, Takubo K, Kawada K, et al. Diagnosis using deep-learning artificial intelligence based on the endocytoscopic observation of the esophagus. *Esophagus*. 2019;16(2):180-187. doi:10.1007/s10388-018-0651-7

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Urabe A, Adachi M, Sakamoto N, et al. Deep learning detected histological differences between invasive and non-invasive areas of early esophageal cancer. *Cancer Sci*. 2025;116:824-834. doi:10.1111/cas.16426