

Evaluating the identification of the extent of gastric cancer by over-1000 nm near-infrared hyperspectral imaging using surgical specimens

Tomohiro Mitsui,^{a,b} Akino Mori,^c Toshihiro Takamatsu^{①,d,e,*} Tomohiro Kadota,^a Konosuke Sato,^c Ryodai Fukushima,^f Kyohei Okubo^{①,c} Masakazu Umezawa^{①,c} Hiroshi Takemura,^{e,f} Hideo Yokota,^g Takeshi Kuwata,^h Takahiro Kinoshita,ⁱ Hiroaki Ikematsu,^{a,d} Tomonori Yano,^a Shin Maeda,^b and Kohei Soga^{c,e}

^aNational Cancer Center Hospital East, Department of Gastroenterology and Endoscopy, Kashiwa, Japan

^bYokohama City University Graduate School of Medicine, Department of Gastroenterology, Yokohama, Japan

^cTokyo University of Science, Department of Materials Science and Technology, Tokyo, Japan

^dNational Cancer Center, Exploratory Oncology Research and Clinical Trial Center, Chiba, Japan

^eTokyo University of Science, Research Institute for Biomedical Sciences, Chiba, Japan

^fTokyo University of Science, Department of Mechanical Engineering, Chiba, Japan

^gRIKEN Center for Advanced Photonics, Saitama, Japan

^hNational Cancer Center Hospital East, Department of Pathology and Clinical Laboratories, Kashiwa, Japan

ⁱNational Cancer Center Hospital East, Department of Gastric Surgery, Chiba, Japan

ABSTRACT. **Significance:** Determining the extent of gastric cancer (GC) is necessary for evaluating the gastrectomy margin for GC. Additionally, determining the extent of the GC that is not exposed to the mucosal surface remains difficult. However, near-infrared (NIR) can penetrate mucosal tissues highly efficiently.

Aim: We investigated the ability of near-infrared hyperspectral imaging (NIR-HSI) to identify GC areas, including exposed and unexposed using surgical specimens, and explored the identifiable characteristics of the GC.

Approach: Our study examined 10 patients with diagnosed GC who underwent surgery between 2020 and 2021. Specimen images were captured using NIR-HSI. For the specimens, the exposed area was defined as an area wherein the cancer was exposed on the surface, the unexposed area as an area wherein the cancer was present although the surface was covered by normal tissue, and the normal area as an area wherein the cancer was absent. We estimated the GC (including the exposed and unexposed areas) and normal areas using a support vector machine, which is a machine-learning method for classification. The prediction accuracy of the GC region in every area and normal region was evaluated. Additionally, the tumor thicknesses of the GC were pathologically measured, and their differences in identifiable and unidentifiable areas were compared using NIR-HSI.

Results: The average prediction accuracy of the GC regions combined with both areas was 77.2%; with exposed and unexposed areas was 79.7% and 68.5%, respectively; and with normal regions was 79.7%. Additionally, the areas identified as cancerous had a tumor thickness of >2 mm.

Conclusions: NIR-HSI identified the GC regions with high rates. As a feature, the exposed and unexposed areas with tumor thicknesses of >2 mm were identified using NIR-HSI.

© The Authors. Published by SPIE under a Creative Commons Attribution 4.0 International License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.JBO.28.8.086001](https://doi.org/10.1117/1.JBO.28.8.086001)]

*Address all correspondence to Toshihiro Takamatsu, takamatsu@rs.tus.ac.jp

Keywords: gastric cancer; near-infrared; hyperspectral imaging; tumor thickness; machine learning

Paper 230099GRR received Apr. 11, 2023; revised Jul. 22, 2023; accepted Aug. 2, 2023; published Aug. 22, 2023.

1 Introduction

Gastric cancer (GC) is a commonly diagnosed cancer worldwide and one of the leading causes of cancer-related mortality.¹ Gastrectomy is a curative treatment for GC.² The effect of gastric remnant volume on the postoperative quality of life and nutritional status has been recognized. Furthermore, organ-sparing surgery is increasingly preferred.^{3–5} Japanese GC treatment guidelines (5th edition) recommend a proximal resection margin of at least 3 cm for tumors \geq T2 in the case of an expansive growth pattern, whereas a proximal resection margin of at least 5 cm is required in the case of an infiltrative growth pattern.⁶ The large margin in the infiltrative growth patterns is attributed to the extensive submucosal or deeper layer invasions. Hence, if the extent of the non-exposed cancer invasion is identified, the stomach can be resected more appropriately. However, assessing the extent of the tumor and determining the appropriate resection margins during surgery remains difficult. Hence, several methods, including intraoperative endoscopy and endoscopic tattooing or clipping, that can identify the tumor extent and determine the appropriate surgical margins exist.^{7–10} However, the endoscopic determination of the extent of the GC that is not exposed to the mucosal surface remains difficult (e.g., the infiltrative growth pattern¹⁰).

We previously reported that near-infrared (NIR) hyperspectral imaging (HSI) could be used to identify gastrointestinal stromal tumors in the muscular layer that was not exposed to the mucosa.^{11,12} Light in the NIR region, ranging from 800 to 2500 nm, is useful for probing deep parts of the tissues owing to its low absorption and scattering. The absorption spectra in the NIR region convey fingerprint data owing to the overtone or combined vibrations of the chemical bonds.^{13–15} This permits the investigation of the distribution of the chemical composition within a sample.¹⁶ Hyperspectral imaging provides a three-dimensional dataset (two spatial and one spectral) that allows a spectral curve at each pixel in the acquired images to be obtained.¹⁷ It uses a machine-learning algorithm to acquire spectral information from each pixel and extract critical imaging data from several hyperspectral images.^{18,19} Hyperspectral imaging has the potential for non-invasive, label-free diagnosis and surgical guidance. Hence, NIR-HSI can identify the extent of the GC that is not exposed to the mucosal surface.

Studies have reported the identification of GC using NIR-HSI, although its characteristics have not been considered.^{19–21} In this study, we aim to explore whether NIR-HSI can identify the GC regions, including each exposed and unexposed area, using surgical specimens and investigate the characteristics of the identifiable GC region.

2 Materials and Methods

2.1 Patients and Surgical Specimen

This study included 10 patients with clinically diagnosed GC who underwent surgery between September 2020 and October 2021. The inclusion criteria were as follows: (i) clinical diagnosis of the GC, (ii) age \geq 20 years, and (iii) written informed consent. The exclusion criteria were as follows: (i) history of prior chemotherapy and (ii) presence of the hepatitis B virus surface antigen or hepatitis C virus antibody. Ten patients who underwent gastrectomy during the study period were enrolled. Table 1 shows the characteristics of the patients and lesions. The median tumor size was 38 mm (range: 20 to 70 mm). The clinical T stage was T1 in 1 lesion, T2 in 3 lesions, and T3–4 in 6 lesions. The macroscopic type was 0-IIc in 1 lesion, T2 in 4 lesions, and T3 in 5 lesions.

The indicators for the types of gastrectomy and a proximal resection margin were determined based on the preference of the surgeon and were primarily based on the Japanese GC treatment guidelines.⁶ All of the surgical specimens were fixed using formalin, cut into 5- or 10 mm-thick slices, stained with hematoxylin and eosin, and evaluated by experienced pathologists. Clinical and pathological outcomes included the Borrmann classification of the tumor, tumor size, depth of invasion, tumor thickness, and histology based on differentiation. The extent

Table 1 Patient and lesion characteristics.

	N = 10
Sex	
Male	7
Female	3
Age, median (range), year	73 (68 to 83)
Pathological tumor size, median (range) (mm)	38 (20 to 70)
Macroscopic type	
0-IIc	2
Type2	3
Type3	5
Pathological tumor depth	
pT1b (SM)	2
pT2 (MP)	4
pT3 (SS)	3
pT4a (SE)	1
Histological type	
tub1, tub2	3
por, sig	0
Mixed	7
Unexposed area	
Present	6
Absent	4

Note: SM, submucosa; MP, muscularis propria; SS, subserosa; SE, serosa exposed; tub1, well differentiated adenocarcinoma; tub2, moderately differentiated adenocarcinoma; por, poorly differentiated adenocarcinoma; sig, signet-ring cell carcinoma.

of the GC was evaluated separately for exposed and unexposed areas. The exposed area was defined as an area where the cancer itself was pathologically exposed on the surface and not covered by any normal tissue. The non-exposed area was defined as an area where the cancer was present although not exposed on the surface and the surface was covered by normal tissue. The normal areas were defined as areas where the cancer was not present, as shown in Fig. 1.

This study was approved by the Institutional Review Board of the National Cancer Center, Japan (Approval No. 2015-339), which conformed to the provisions of the Declaration of Helsinki and Epidemiological Study Guidelines issued by the Japan Ministry of Health, Labor, and Welfare. All patients provided a written informed consent prior to their inclusion in the study.

2.2 Near-Infrared Hyperspectral Image Capture and Data Preprocessing

An imaging system with a high-speed NIR hyperspectral camera (CompoVision, CV-N800HS; Sumitomo Electric Industries, Ltd., Osaka, Japan) was used to obtain the NIR-HSI images (wavelength: 1000 to 2350 nm, wavelength resolution: 6.3 nm, and depth resolution: 14 bit). The detector (NIR spectroscopic camera) captured the data values for each wavelength band on each pixel per line of the image in one scan. Three-dimensional HSI images (x - y - λ axes) were obtained by scanning multiple lines (by sliding the sample stage), thereby producing a

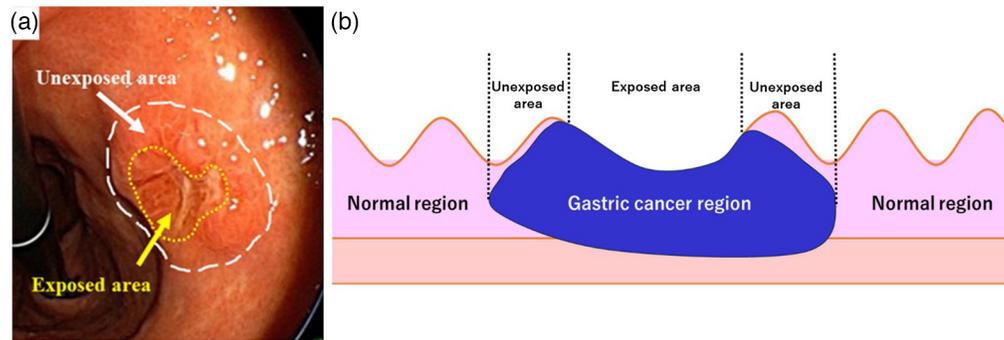


Fig. 1 (a) Endoscopic view of gastric cancer. White line is the predicted unexposed area. Yellow line is the predicted exposed area. (b) Histopathology image schema. Cancerous and normal areas were distinguished by the pathologist. Cancerous areas were divided into exposed and unexposed areas, according to their definitions.

virtual “data cube” for processing and analysis.¹⁷ Each fresh specimen that, prior to the formalin fixation was resected from the stomach, was placed on the sliding stage without trimming. Each NIR-HSI image was acquired once from the mucosal side under illumination from a halogen lamp (0.96 W/cm^2). To analyze the HSI data, it is necessary to calibrate the images using dark noise and a white standard for each pixel(i, j). Each reflectance is expressed as follows:

$$\mathbf{R}(i, j) = \frac{\mathbf{I}_r(i, j) - \mathbf{I}_d(i, j)}{\mathbf{I}_w(i, j) - \mathbf{I}_d(i, j)}, \quad (1)$$

where $\mathbf{R}(i, j)$ is a row vector of the reflectance spectrum of the obtained image and $\mathbf{I}_r(i, j)$, $\mathbf{I}_w(i, j)$, and $\mathbf{I}_d(i, j)$ are the row vectors of raw, white standard, and dark noise data, respectively. Spectra with wavelengths $>1400 \text{ nm}$ were removed from the analysis because of the high absorption by the water in those bands and the lower sensitivity of the NIR camera, unlike the transmission spectroscopy.²⁰ Additionally, the reflectance rates of more than 70% as highlights and below 10% as shadows were defined in the 1300-nm spectra. Further, the pixels of the necrosis were removed from the dataset and analysis. Normal, exposed, and unexposed GCs were marked by a pathologist to create three regions showing the pixels of the boundary lines. The 5-pixel margin around the boundary lines was excluded while extracting the spectra from each region because the boundary line was drawn freehand by the pathologist. Hence, it was susceptible to error.

NIR spectral measurements reported the variance and the non-specific scattering at the surface of the sample.²² The standard normal variate (SNV) of absorbance was used for the baseline correction of the spectrum to reduce the variance and is as expressed as follows:

$$\mathbf{A}(i, j) = -\log_{10} \left(\frac{\mathbf{I}_r(i, j) - \mathbf{I}_d(i, j)}{\mathbf{I}_w(i, j) - \mathbf{I}_d(i, j)} \right), \quad (2)$$

$$\mathbf{Z}(x) = \frac{x - \text{mean}(x)}{\text{std}(x)}, \quad (3)$$

where $\mathbf{A}(i, j)$ is a row vector of the absorbance spectrum, x is a row vector containing the original spectrum, $\text{mean}(x)$ is the mean of x , $\text{std}(x)$ is the standard deviation of x , and $\mathbf{Z}(x)$ is the SNV-transformed spectrum.

2.3 Spectral Data Analysis

A support vector machine (SVM) was used to perform a two-class classification of the normal tissues and entire GC lesions, excluding the 5-pixel margin of the boundary line and necrosis area [inside of the pink line in Fig. 3(B)].

During this process, the pixels to be trained were randomly extracted from a specimen to avoid overfitting. The number of pixels was aligned with normal 200 px and GC 200 px from the confirmed area by pathology. Further, leave-one-out cross-validation was employed, wherein a specimen is classified by a training dataset (each 200 px of normal and GC from nine specimens;

total 3600 spectra) that excludes the specimen pixels because the procedure provides an approximately unbiased estimate of the generalization ability.²³ This algorithm solves an optimization problem expressed as follows:²⁴

$$\begin{aligned} \text{minimize } t(w_n, \xi_i) &= \frac{1}{2} \sum_{n=1}^k \|w_n\|^2 + \frac{C}{m} \sum_{i=1}^m \xi_i, \\ \text{subject to } \langle x_i, w_{y_i} \rangle - \langle x_i, w_n \rangle &\geq b_i^n - \xi_i \quad (i = 1, \dots, m), \\ b_i^n &= 1 - \delta_{y_i, n}, \end{aligned} \quad (4)$$

where w_n is a weight vector, C is the cost, and ξ_i is the slack variable. Further, the decision function is expressed as

$$g \max_{m=1, \dots, k} \langle x_i, w_n \rangle. \quad (6)$$

The optimization problem was solved using the decomposition method.²⁵ This study used $C = 1$, and the RBF kernel is expressed as

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{\sigma^2}\right), \quad (7)$$

where the optimal values of the hyperparameter σ^2 are estimated as follows:

$$\sigma^2 := \text{median}\{\|x_i - x_j\|^2 | i < j\}. \quad (8)$$

The SNV-transformed spectra of the normal tissues, except for the test sample, were averaged. Further, the mean SNV-transformed spectra of the exposed and unexposed regions in the tumor of the test sample were calculated. Hence, the difference in the spectra was obtained by comparing the spectra of the normal and tumor tissues.

2.4 Calculation of Prediction Accuracy

The coordinates of the classification pixels were compared to the line drawing images by a pathologist to evaluate the classification accuracy. The area of analysis in this study was limited to the pathologically confirmed area (green area) because only the green area in Fig. 4 is the pathologically evaluated area. This is because the pathologically evaluated sites are the cancer and surrounding normal tissue. The pixels were classified into four groups: GC is predicted as GC [true positive (TP)], GC is predicted as normal [false negative (FN)], normal is predicted as GC [false positive (FP)], and normal is predicted as normal [true negative (TN)]. The false-positive rate and false-negative rate were calculated from the classified pixels. The specificity, sensitivity, and accuracy were expressed as follows:

$$\text{specificity (\%)} = \frac{\text{TN}}{\text{FP} + \text{TN}} \times 100, \quad (9)$$

$$\text{sensitivity (\%)} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100, \quad (10)$$

$$\text{accuracy (\%)} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \times 100. \quad (11)$$

2.5 Comparison of Histopathologic Tumor Thickness

Histopathological tumor thicknesses were compared to determine the factors that contributed to the identified areas (TP pixels) and unidentified areas (FN pixels) of the GC using NIR-HSI. The criteria for the identified and unidentified areas of each specimen were as follows. If the sensitivity was $>80\%$, the pathological sections in the identified areas could be accurately recognized, but the unidentified areas were difficult to recognize. Therefore, tumor thickness was measured only in the identified areas. However, when the sensitivity was $<80\%$, it was possible to recognize pathological sections in the identified and unidentified areas; therefore, tumor thickness was measured in the identified and unidentified areas.

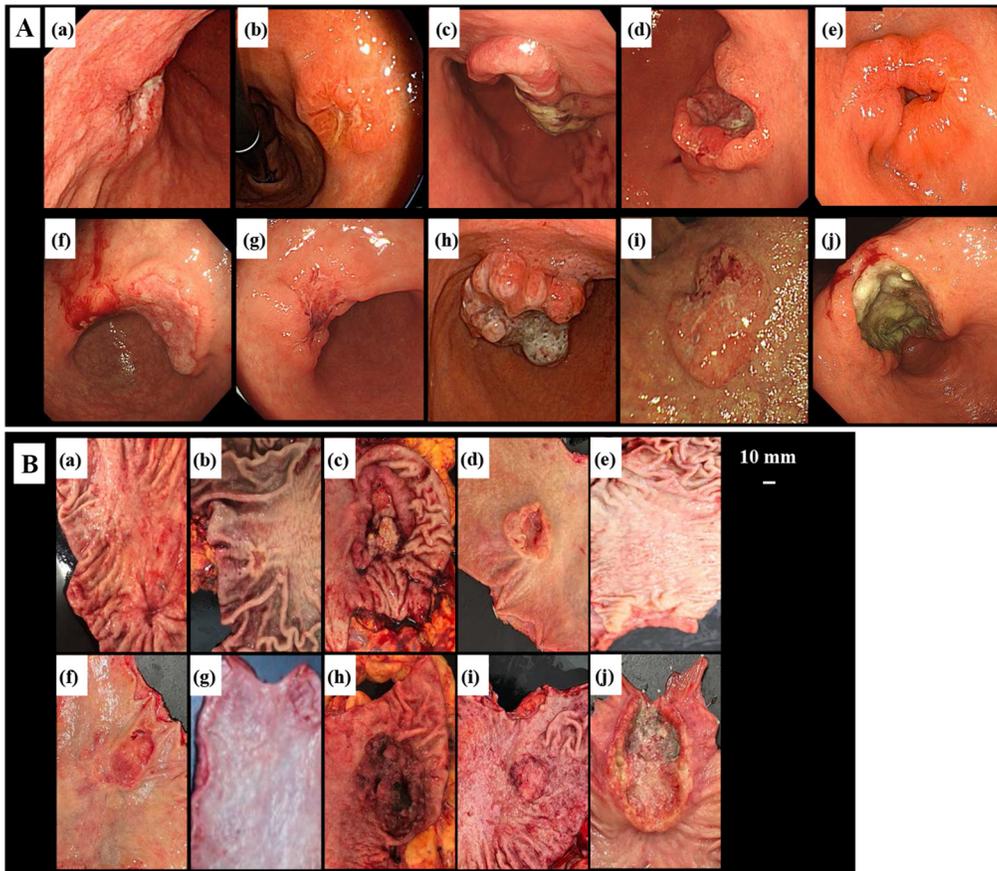


Fig. 2 (A), (a)–(j) Endoscopic view of GC. (B), (a)–(j) Visible light photograph of specimens captured using a digital camera.

3 Results

3.1 Endoscopic Images

The endoscopic images of the lesions and pictures of the 10 excised specimens are shown in Figs. 2(A) and 2(B), respectively. Four specimens [d, e, g, and i in Fig. 2(B)] were exposed to the mucosal surface of all regions for cancer. Six specimens [a–c, f, h, and i in Fig. 2(B)] had unexposed cancer in some regions.

Figure 3(A) shows the pictures of the 10 specimens captured using the NIR-HSI. The GC and normal regions in the NIR-HSI images were marked by a pathologist [Fig. 3(B)]. The white lines represent the extent of the GC, and the yellow ones represent the borderline between the exposed and unexposed areas of the GC. To visualize the trend of the spectral regions contributing to the cancer classification, the difference between the SNV-transformed spectra of the normal tissues (leave-one-out training data) and mean SNV-transformed spectra of the exposed and unexposed regions in the tumor of the test sample is shown in Fig. 3(C).

The results showed that the absorbance difference in the wavelength range of 1050 to 1100 nm and 1380 to 1400 nm was negative and in the wavelength range of 1250 to 1350 nm was positive, except for the unexposed and exposed areas of the specimens (f and g, respectively).

3.2 Identification Results of Gastric Cancer and Normal Regions using NIR-HSI

The analysis of the spectra from the HSI images was performed based on the training data using the SVM algorithm. Further, the GC and normal regions were identified. The pixels that were predicted as GC were assigned a red color, and the normal tissues were assigned a green color. The color-coded pixels that were predicted as GC and normal tissues are shown in the upper

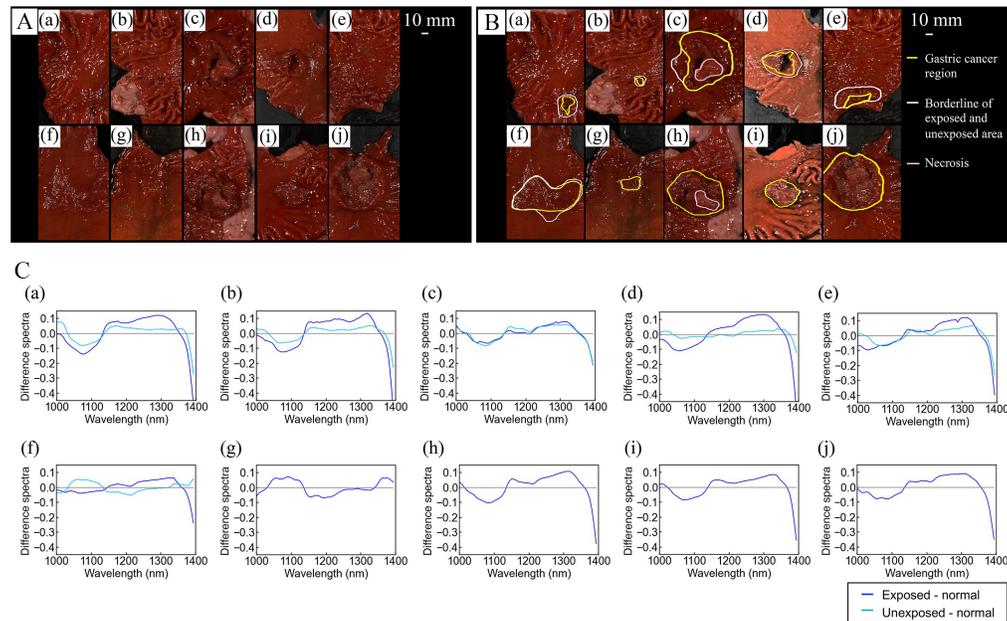


Fig. 3 (A), (a)–(j) Pseudocolored pictures of the specimens captured using a CompoVision (NIR camera, R: 1065 nm, G: 1280 nm, B: 1981 nm). (B), (a)–(j) Gastric cancer area (white line), exposed and unexposed areas of the GC (yellow line), and necrosis area (pink line) marked by a pathologist. (C), (a)–(j) The difference in the spectra between the SNV-transformed spectra of normal tissues (leave-one-out training data) and mean SNV-transformed spectra of the exposed and unexposed regions in the tumor of the test sample.

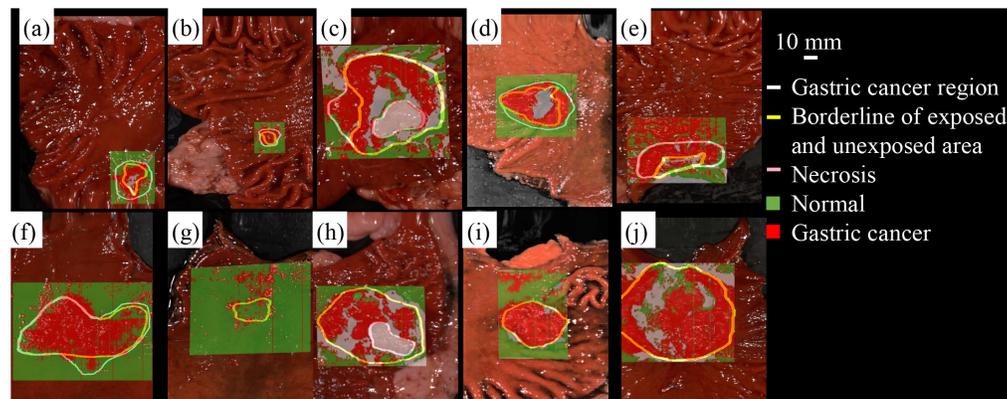


Fig. 4 (a)–(j) Gastric cancer region prediction analyzed using machine learning. The color-coded pixels represent the prediction of GC (red) and normal (green).

images in Figs. 4(a)–4(j). The lower images in Figs. 4(a)–4(j) were merged to include the boundary line of the pathologist. Furthermore, the pixel areas used for calculating the prediction accuracy were also shown. The pixels (459,688 px) were classified as follows: TP: 160,693 px, FN: 35,517 px, FP: 55,567 px, and TN: 207,911 px. The specificity, sensitivity, and accuracy of 74.8%, 77.2%, and 79.7%, respectively, were evaluated from the classified pixels.

Table 2 presents the results. The prediction results of the exposed area were classified as follows: TP: 138,545 px and FN: 27,971 px. The sensitivity of 79.7% was calculated from the classified pixels. Additionally, the prediction results for the unexposed area were classified as follows: TP: 22,148 px and FN: 7545 px. Furthermore, the sensitivity of 68.5% was calculated from the classified pixels. The complete results are presented in Table 3.

Table 2 Prediction results of the NIR-HSI analysis. The GC region [the inside of the white line of Fig 2(b)] was defined as “positive,” and the outside of the GC was defined “negative” in the pathologically confirmed area. The 5 px margin of the white line and necrosis area were excluded from the analysis.

No.	TP (pixel)	FN (pixel)	FP (pixel)	TN (pixel)	Specificity (%)	Sensitivity (%)	Accuracy (%)
(a)	4070	2041	1310	7607	85.3	66.6	77.7
(b)	1520	379	777	3891	83.4	80.0	82.4
(c)	30,443	5506	7007	19,267	73.3	84.7	79.9
(d)	10,952	1804	1075	17,886	94.3	85.9	90.9
(e)	10,095	1387	15712	11,450	42.2	87.9	55.8
(f)	31,940	10354	6487	49,284	88.4	75.5	82.8
(g)	1153	3326	7624	64,705	89.5	25.7	85.7
(h)	17,642	3438	2143	10,518	83.1	83.7	83.5
(i)	13,460	485	6760	19,821	74.6	96.5	82.1
(j)	39,418	6797	6672	3482	34.3	85.3	76.1
Total	160,693	35,517	55,567	207,911	74.8 ± 19.3	77.2 ± 18.7	79.7 ± 8.9

Note: TP, true positive; FN, false negative; FP, false positive; and TN, true negative.

Table 3 Prediction results in the exposed and unexposed areas during the NIR-HSI analysis. Among the pixels analyzed in Table 2, the TP and FN of the area inside the yellow line in Fig. 3(B) were defined as “exposed,” and the TP and FN of the area between the white and yellow lines were defined as “unexposed”.

No.	Exposed			Unexposed		
	TP (pixel)	FN (pixel)	Sensitivity (%)	TP (pixel)	FN (pixel)	Sensitivity (%)
(a)	2096	300	87.5	1974	1741	53.1
(b)	901	121	88.2	619	258	70.6
(c)	24236	4855	83.3	6207	651	90.5
(d)	7816	414	95.0	3136	1390	69.3
(e)	1387	535	72.2	8708	852	91.1
(f)	30436	7701	79.8	1504	2653	36.2
(g)	1153	3326	25.7	—	—	—
(h)	17,642	3438	83.7	—	—	—
(i)	13,460	485	96.5	—	—	—
(j)	39,418	6797	85.3	—	—	—
Total	138,545	27,971	79.7 ± 19.1	22,148	7545	68.5 ± 19.5

Note: TP, true positive; FN, false negative.

3.3 Relationship between Tumor Thickness and Identifiable Area using NIR-HSI

Table 4 shows the tumor thickness in the GC region. NIR hyperspectral imaging detected cancer areas with pathologic thicknesses of >2 mm regardless of whether the cancer was exposed on the mucosal surface or not (Fig. 5).

Table 4 Relationship between the tumor thickness and identifiable area using NIR-HSI.

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)
Tumor thickness, range (mm)										
Identifiable area	2.4 to 8.4	2.0 to 4.0	4.2 to 16.0	2.2 to 11.4	2.2 to 5.5	2.1 to 10.4	—	4.8 to 17.4	3.2 to 4.8	3.4 to 13.0
Unidentifiable area	0.2 to 1.3	0.6 to 1.4	—	1.2 to 1.4	1.3 to 1.5	0.6 to 1.2	0.4 to 1.3	—	—	—

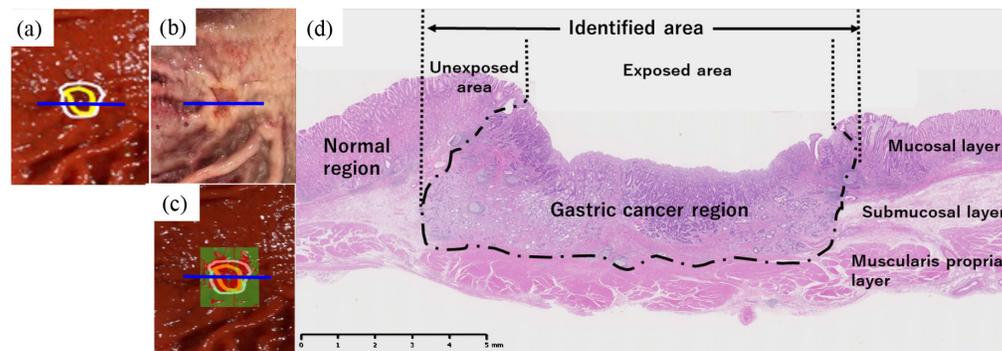


Fig. 5 For the case in Fig. 2A(b), (a) NIR-HSI image showing the gastric cancer region (white line) and borderline between the exposed and unexposed area (red line), (b) resected specimen after fixation, and (c) image showing the predicted GC and normal tissue, colored red and green, respectively. (d) The histological image revealing the area of the cut along the blue line in (a)–(c). The histology from the surgical resection revealed adenocarcinoma in the deep submucosal layer. The GC was identified using NIR-HSI regardless of the exposed or unexposed area.

However, thin-tumor areas, including those exposed on the mucosal surface, were unidentified as cancerous (Fig. 6).

Additionally, in the prediction results of the NIR-HSI analysis as shown in Table 3, specimens g and f had the lowest sensitivity in the exposed and unexposed regions, respectively. As shown in Fig. 3(C), the spectral trends of both specimens were different from those of the other highly sensitive specimens. This may be because the lesion tissue was thin (Table 4), whereas the molecular structures of the samples were different from those of the other samples.

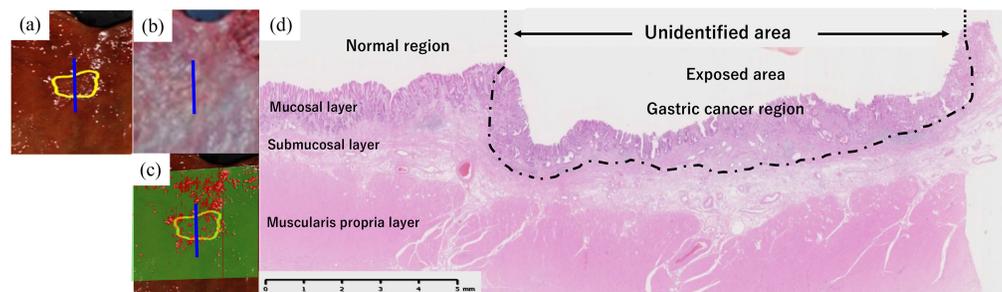


Fig. 6 For the case in Fig. 2A(g), (a) NIR-HSI image marked with GC, (b) resected specimen after fixation, and (c) image showing the predicted GC and normal tissue, colored red and green, respectively. (d) The histological image revealing the area of the cut along the blue line in (a)–(c). The histology from the surgical resection revealed adenocarcinoma in the submucosal slight layer. The GC was unidentified using NIR-HSI.

4 Discussion

This study evaluated the extent of the GC, including the unexposed areas that were difficult to recognize under white light or by machine learning using the NIR-HSI data. The findings of this study are novel because previous NIR-HSI studies on GC did not discuss the identification of cancerous areas under normal mucosa.

This study identified the exposed area of the GC as cancerous with high sensitivity, that is, with 79.7% and 68.5% sensitivities in the exposed and unexposed areas, respectively. However, the overall sensitivity was lower in the unexposed area than in the exposed area. Further, several FPs were detected in thin cancerous areas, which is possibly owing to the characteristics of the NIR absorption spectroscopy that uses reflected light: (i) the thicker the normal mucosa covering cancer, the more attenuated the characteristic absorption spectrum of cancer and (ii) the thinner the cancer, the sparser the information in the absorption spectrum. As shown in Fig. 3(C), the absorbance difference between the normal area spectra in the training data and that of the unexposed area was smaller than that of the exposed area. Areas with tumor thicknesses of >2 mm were identified using the NIR-HSI, even in unexposed areas (Table 4). However, areas with thin tumors remained unidentified in either the exposed or unexposed areas. This study shows that NIR-HSI can be used to accurately recognize unexposed areas that are difficult to recognize under white light when the thickness is 2 mm or more.

This study was limited by the small number of specimens. However, it classified NIR-HSI-captured pixels (total pixel: 459,688 px) using the leave-one-out cross-validation procedure. Hence, the sample size was sufficient, and its generalizability was demonstrated. The results for shallow-depth areas were inadequate because numerous specimens in the study had advanced cancer, and sufficient data on the areas with shallow depths of cancer invasion were lacking. Gathering more data on shallow lesions was necessary, as represented by the specimen in Fig. 2(B) [case (g)]. Hence, the shallow lesions were unidentified as cancers. The optimization of the measurement conditions, such as changes in the illumination method of the NIR-HSI system, specimen placement, and design focal length, improved the detection performance. Additionally, although an SVM was adopted as a machine learning method in this study, other methods such as neural network and principal correlation analysis have been proposed to identify lesions. The identification accuracy may be improved by investigating the optimal algorithm and parameter. Moreover, significantly increasing the number of specimens would broaden the range of options for popular deep learning techniques and allow for the exploration of more robust classifiers.

However, imaging conditions in the stomach using an endoscope with NIR-HSI differed from those of our study. For example, the postoperative specimen was imaged from the front in this study, though it may not be possible in the case of an endoscope in the stomach. Additionally, the absorption spectra may change before gastrectomy because of the blood flow and other factors. Hence, the NIR-HSI devices for endoscopic use must be developed, and the optimization of the measurement conditions must be considered in the future.

Our group is currently developing an NIR-HSI device that can be used with endoscopes. We expect to perform an endoscopic NIR-HSI analysis in the near future. If this technology is used with an endoscope, it may be possible to intraoperatively or preoperatively identify the extent of the unexposed area and define a more appropriate line for the resection of the stomach.

5 Conclusion

This study accurately identified the exposed and unexposed GC areas with thicknesses >2 mm. In the future, we expect to develop an NIR-HSI system that will be used with endoscopes to enhance its practicality.

Disclosures

The authors have no relevant financial interests in the manuscript and no other potential conflicts of interest to disclose.

Data availability

The datasets generated and analyzed during the current study are not publicly available owing to containing personal patient information but are available from the corresponding author on reasonable request.

Acknowledgments

The authors would like to thank the investigators and patients for their participation in the study. This study was supported by the Japan Society for the Promotion of Science KAKENHI (Grant No. 21H038440) and National Cancer Center Research and Development Fund (Grant No. 31-A-11).

References

1. H. Sung et al., "Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA Cancer J. Clin.* **71**(3), 209–249 (2021).
2. R. Glynne-Jones et al., "Anal cancer: ESMO-ESSO-ESTRO clinical practice guidelines for diagnosis, treatment and follow-up," *Radiother. Oncol.* **111**(3), 330–339 (2014).
3. H. J. F. Brenkman et al., "Factors influencing health-related quality of life after gastrectomy for cancer," *Gastric Cancer* **21**, 524–532 (2018).
4. M. Takahashi et al., "Quality of life after total vs distal gastrectomy with Roux-en-Y reconstruction: use of the Postgastrectomy Syndrome Assessment Scale-45," *World J. Gastroenterol.* **23**(11), 2068–2076 (2017).
5. R. Rosania et al., "Nutrition in patients with gastric cancer: an update," *Gastrointest. Tumors* **2**(4), 178–187 (2016).
6. Japanese Gastric Cancer Association, "Japanese gastric cancer treatment guidelines 2018 (5th edition)," *Gastric Cancer* **24**, 1–21 (2021).
7. D. J. Park et al., "Intraoperative gastroscopy for gastric surgery," *Surg. Endosc.* **19**, 1358–1361 (2005).
8. H. Matsui et al., "Endoscopy-assisted gastric resection: a safe and reliable procedure for tumor clearance during laparoscopic high distal or proximal gastrectomy," *Surg. Endosc.* **23**, 1146–1149 (2009).
9. R. I. Beretvas and J. Ponsky, "Endoscopic marking: an adjunct to laparoscopic gastrointestinal surgery," *Surg. Endosc.* **15**, 1202–1203 (2001).
10. S. Kawakatsu et al., "Use of endoscopy to determine the resection margin during laparoscopic gastrectomy for cancer," *Br. J. Surg.* **104**(13), 1829–1836 (2017).
11. D. Sato et al., "Distinction of surgically resected gastrointestinal stromal tumor by near-infrared hyperspectral imaging," *Sci. Rep.* **10**, 21852 (2020).
12. K. Akimoto et al., "Wavelength bands reduction method in near-infrared hyperspectral image based on deep neural network for tumor lesion classification," *AIVP* **9**(1), 273–281 (2021).
13. R. R. Anderson and J. A. Parrish, "The optics of human skin," *J. Invest. Dermatol.* **77**(1), 13–19 (1981).
14. K. Okubo et al., "Visualization of quantitative lipid distribution in mouse liver through near-infrared hyperspectral imaging," *Biomed. Opt. Express* **12**(2), 823–835 (2021).
15. T. Takamatsu et al., "Over 1000 nm near-infrared multispectral imaging system for laparoscopic in vivo imaging," *Sensors* **21**(8), 2649 (2021).
16. A. M. Smith, M. C. Mancini, and S. Nie, "Second window for in vivo imaging," *Nat. Nanotechnol.* **4**, 710–711 (2009).
17. Q. Li et al., "Review of spectral imaging technology in biomedical engineering: achievements and challenges," *J. Biomed. Opt.* **18**(10), 100901 (2013).
18. Z. Liu, H. Wang, and Q. Li, "Tongue tumor detection in medical hyperspectral images," *Sensors* **12**(1), 162–174 (2012).
19. H. Akbari et al., "Cancer detection using infrared hyperspectral imaging," *Cancer Sci.* **102**(4), 852–857 (2011).
20. A. Yahata et al., "Wavelength selection of near-infrared hyperspectral imaging for gastric cancer detection," in *6th Int. Conf. Intell. Inf. and Biomed. Sci. (ICIIBMS)*, Oita, pp. 219–223 (2021).
21. N. Liu et al., "Gastric cancer diagnosis using hyperspectral imaging with principal component analysis and spectral angle mapper," *J. Biomed. Opt.* **25**(6), 066005 (2020).
22. Q. Guo, W. Wu, and D. L. Massart, "The robust normal variate transform for pattern recognition with near-infrared data," *Anal. Chim. Acta* **382**(1–2), 87–103 (1999).
23. G. C. Cawley and N. L. C. Talbot, "Fast exact leave-one-out cross-validation of sparse least-squares support vector machines," *Neural Networks* **17**(10), 1467–1475 (2004).
24. A. Karatzoglou et al., "kernlab—an S4 package for kernel methods in R," *J. Stat. Software* **11**(9), 1–20 (2004).
25. C. W. Hsu and C. J. Lin, "A simple decomposition method for support vector machines," *Mach. Learn.* **46**, 291–314 (2002).

Toshihiro Takamatsu is an associate professor at Tokyo University of Science, Chiba, Japan, and a researcher at the National Cancer Center, Chiba, Japan. He received his BS degree in physics from Tokyo University of Science in 2010 and his MS degree and PhD in engineering from Tokyo Institute of Technology, Yokohama, Japan, in 2011 and 2014, respectively. He was a postdoctoral researcher at Kobe University, Hyogo, Japan, from 2014 to 2018.

Biographies of the other authors are not available.