Stomach tissue classification using autofluorescence spectroscopy and machine learning

Stomach tissue classification using autofluorescence spectroscopy and machine learning

Stomach tissue classification using autofluorescence spectroscopy and machine learning

Soo Yeong Lim,[a] Hong Man Yoon,[b] Myeong-Cherl Kook, [b] Jin Il Jang,[a] Peter T. C. So,[c] Jeon Woong Kang,[c]* and Hyung Min Kim[a]*

[a] Department of Chemistry, Kookmin University, 77, Jeongneung-ro, Seongbuk-gu, Seoul, 02707, Republic of Korea

[b] Division of Convergence Technology, National Cancer Center, 323 Ilsan-ro, Ilsandong-gu, Goyang-si Gyeonggi-do, 10408, Republic of Korea

[c] Laser Biomedical Research Center, G. R. Harrison Spectroscopy Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Category: Original Articles

* Soo Yeong Lim and Hong Man Yoon contributed equally to this work as first authors.

* Jeong Woong Kang and Hyung Min Kim equally contributed to this work as co-corresponding authors.

Correspondence to:

Jeon Woong Kang, Ph.D.

Laser Biomedical Research Center, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139

Tel: 1-617-258-9404

Fax: 1-617-253-4513

E-mail: jwkang76@mit.edu

Hyung Min Kim, Prof.

Department of Chemistry, Kookmin University, 77, Jeongneung-ro, Seongbuk-gu, Seoul, 02707, Republic of Korea

Tel: +82-2-910-5883

Fax: +82-2-910-4415

E-mail: hyungkim@kookmin.ac.kr

Abstract

Background and Objectives

Determination of stomach tumor location and invasion depth requires delineation of gastric histological structure, which has hitherto been widely accomplished by histochemical staining. In recent years, alternative histochemical evaluation methods have been pursued to accelerate intraoperative diagnosis, often by bypassing the time-consuming step of dyeing. Owing to strong endogenous signals from coenzymes, metabolites, and proteins, autofluorescence spectroscopy is a favorable candidate technique to achieve this aim.

Materials and Methods

We investigated stomach tissue slices and block specimens using a fast fluorescence imaging scanner. To obtain histological information from broad and structureless fluorescence spectra, we analyzed tens of thousands of spectra with multiple machine learning algorithms and built a tissue classification model trained with dissected gastric tissues.

Results

A machine learning based spectro-histological model was built based on the autofluorescence spectra measured from stomach tissue samples with delineated and validated histological structures. The scores from a principal components analysis were employed as input features, and prediction accuracy was confirmed to be 92.0%, 90.1%, and 91.4% for mucosa, submucosa, and muscularis propria, respectively. We investigated the tissue samples in both sliced and block forms using a fast fluorescence imaging scanner.

Conclusion

We successfully demonstrated differentiation of multiple tissue layers of well-defined specimens with the guidance of a histologist. Our spectro-histology classification model is applicable to histological prediction for both tissue blocks and slices, even though only sliced samples were trained.

Keywords: Spectroscopy, Autofluorescence, Machine learning, Histology

Introduction

The histological examination of human organs has facilitated surgical disease diagnosis, as pathological stages—including tumor metastasis—are strongly correlated with the invasion depth of the primary lesion. Stomach cancer is the third commonest cancer type among Koreans (regardless of gender). Overall survival rates after cancer treatment are increasing every year with state-of-the-art diagnostics and treatment innovations. Since early tumor detection contributes greatly to extended life expectancy, it is important to understand the gastric histological concepts underpinning cancer invasion.

Stomach tissue is composed of mucosa, submucosa, muscularis propria, subserosa, and serosa from the inner surface toward the outer layer. Stomach cancer generally originates in the mucosa and infiltrates into the subsurface structures; therefore, diagnostic tools for determining tumor invasion depths are needed to specify latent tumor layers in multi-layer structures. Microscopic analysis with histochemical staining has been widely employed in the fields of histology to investigate biological tissue structures. Alternative spectroscopic methods, including infrared [1], Raman [2], diffuse reflectance [3], autofluorescence [4] methods, and combined modalities [5], have been developed to enable quick medical decisions. Spectroscopy yields vibrational (infrared and Raman spectroscopy) and electronic (absorption and fluorescence) information about biological structures without chemical or biological labeling. Recently, machine learning (ML) [6] and deep learning [7] have risen to prominence in the field of optical and spectroscopic diagnosis, outperforming traditional data processing methods.

Autofluorescence spectroscopy/imaging has been considered a superior non-destructive diagnostic method because it detects intense fluorescence signals from nascent endogenous fluorophores. Tissue autofluorescence [8] emanates from collagen [9, 10], elastin [11, 12], lipofuscin [13, 14], protoporphyrin [15, 16], and coenzymes [14, 17], meaning that autofluorescence spectroscopy does not require labeling with exogenous dyes, so the results can be quickly confirmed using simple steps. Furthermore, multiple studies have demonstrated that fluorescence lifetime can provide substantial diagnostic information, and fluorescence lifetime imaging microscopy (FLIM) has been applied for diagnosing several types of cancer, including ovarian [18], breast [19], and lung cancer [20]. Fluorescence measurement in the spectral domain takes advantage of simpler imaging

systems compared to those used with FLIM, but it requires elaborate mathematical or statistical analysis because typical autofluorescence spectra are broad and featureless, originating from multiple fluorophores. In recent years, ML algorithms have facilitated efforts to yield biological and medical information from infrared, Raman, and fluorescence spectra.

In this work, we developed an automated classification method for stomach tissue using a fluorescence scanning system and ML. Gastric tissues were obtained from the National Cancer Institute (NCI) Cooperative Human Tissue Network (CHTN) and cut into two or more slices for histochemical and spectroscopic evaluation. Tissue substructures were indicated on histochemical images according to a three-domain (mucosa, submucosa, and muscularis) classification, and autofluorescence data from each domain were collected for ML training with the guidance of a histologist. Multiple ML algorithms were applied to train and validate models to predict slight differences in ten features extracted by principal component analysis (PCA). Finally, ML-guided prediction results were compared with expert-guided classification results on tissue slices and intact tissue specimens.

Materials and methods

Stomach tissue sample processing

Stomach tissue samples were obtained from the Cooperative Human Tissue Network (CHTN) and IRB review is not required for our study because the samples were not collected specifically for the research. (https://couhes.mit.edu/definitions#Human%20subject). Fourteen frozen human stomach tissue specimens were scanned on a tissue scanner. Eleven of these were rapidly cooled and sliced using a cryotome (Cryostat HM 505 E; MICROM International GmbH, Germany). Adjacent slices were used for autofluorescence measurement and hematoxylin and eosin (H&E) staining, respectively. During frozen section preparation, we minimally applied optimal cutting temperature (OCT) compound to avoid the unwanted detection of OCT signals from sliced tissues.

The slice thickness was adjusted for optical biopsies: 100 μm for autofluorescence measurements and 20 μm for histological examination. A single section from each sliced tissue was reserved for H&E staining, and

photomicrographs were examined histologically. The remaining slice section was sampled on a quartz glass for autofluorescence mapping. Figure 1(a) illustrates the dual usage of sample species in our study. Three tissues blocks were preserved without dissection to be analyzed without sectioning. For the tissue blocks, photographs and autofluorescence images were captured without H&E staining.

Instrumental setup

We assembled a high-speed fluorescence scanner equipped with a broadband laser-driven light source (EQ-99-FC-S LDLS; Energetiq Technology, USA), a 355-nm ultraviolet laser (SNV-40F-000; Teem Photonics, France), a spectrometer (USB2000+; Ocean Optics, USA), a fiber-optic probe, and a motorized scanning stage (MS-6000-XY; Applied Scientific Instruments, USA). The overall design of the system is based on a previously described instrument. [5, 21] The fiber-optic probe utilized in this study consisted of three multi-mode fibers arranged in a triangular configuration. One fiber is used for laser excitation, another for broadband excitation, and the third for signal collection. However, the lamp commonly used for broadband excitation in diffuse reflectance measurement was not used in this study.

Figure 1 (b) demonstrates the instrumental setup developed for this study. Sample tissues were placed on glass coverslips and placed on top of a large quartz plate (t = 175 × 175 × 1.6 mm) equipped on the stage, which was controlled using LabVIEW software (LabVIEW 2017; National Instruments, USA). Excitation light at 355 nm was transferred to the bottom of a cover slip through a fiber and irradiated beneath the sample to induce fluorescence. Subsequently, the emitted fluorescence signal was collected through another fiber located adjacent to the excitation fiber and delivered to the spectrometer. For the sake of simplicity in optical configuration and robustness of measurement, fluorescence light was not passed through optical filters. Signal collection was synchronized with step-by-step stage movement along the vertical and horizontal directions. Given that the fluorescence originated from the tissue components themselves (autofluorescence), it was unnecessary to stain the stomach tissues with fluorophores.

For autofluorescence image analysis, a sample of about 10 × 10 mm was typically measured with a resolution of 100 μm and an exposure time at each point of 30 ms. The laser power was maintained at 40 μW

at the sample surface. The fiber bundle was placed directly under the quartz plate because excitation light expands at the outlet of the optical fiber. Thus, the spatial resolution was determined by the core size of the optical fibers and the distance between the fibers and sample stage. While analyzing the intrinsic autofluorescence from the tissue samples, we performed data preprocessing, including intensity correction, Savitzky-Golay smoothing, background subtraction, and cosmic-ray correction. After the fluorescence spectra were obtained, their spectral intensities were corrected to complement the performance of our system. The performance of the optical components and spectrometer were calibrated using an NIST traceable lamp (HL-3P-CAL; Ocean Insight). The optical background and electronic offset were removed by subtracting the mean spectrum from bare coverslip, corresponding to the first row of autofluorescence map.

Data and results

Autofluorescence map, photograph, and hematoxylin and eosin–stained image

Figure 2 shows the basic scheme of the autofluorescence-based tissue classification model developed in this work. In our classification study, tissue specimens excised from patients were divided into training and test sets for ML modeling and cut into two slices or more. One of the slices from a tissue block was reserved for expert-based histology and stained with H&E and the other slice was reserved for photographic and fluorescence imaging without further treatment. With guidance from a histologist, each tissue sample was divided into three sub-tissue regions (mucosa, submucosa, and muscularis propria) on the H&E images. Moreover, to label the fluorescence spectra for the training set, each region on the photographic and fluorescence images was matched with the corresponding H&E image. Fluorescence spectra were preprocessed with intensity correction and baseline subtraction to improve ML prediction accuracy. PCA was conducted to reduce the dimensions of spectral data space, and multiple ML algorithms were performed simultaneously. The most accurate algorithm was selected for the final ML prediction to improve accuracy.

Figure 3 demonstrates the four image types typically obtained in this study: H&E–stained histological images, photographic images, and the maps of peak intensity and peak shift of the highest autofluorescence band. H&E staining is a reference method for viewing cellular and tissue structure in histology in which H&E

dye stains the nuclei of a cell dark blue, and eosin stains the cytoplasm pink. The structure of gastric tissue can be completely classified based on H&E–stained preparations, and the red and blue dotted lines in Figure 3 (a) and (b) outline the mucosal and submucosal domains in tissue S−1 and S−2, respectively. The photograph of each sample was taken while the tissue was placed on the fluorescence scanning stage and accordingly matching autofluorescence map, as shown in Figure 2 (c)–(h). The peak intensity maps in Figure 3 (e) and (f) represent the integral values of autofluorescence from 400 to 600 nm at each point, and the peak shifts in Figure 3 (g) and (h) correspond to the center wavelengths of the highest peaks.

Figure 4 (a) shows the images of eleven H&E–stained tissue slices and photographs of three tissue blocks. Samples sliced to a thickness of 100 μm are denoted as 'S−', and samples not sliced are denoted as 'B−'. In the histological images, the regions outlined in red, blue, and green dotted lines represent the mucosal, submucosal, and muscularis propria domains, respectively. Among the tissue slices, four samples (S−5, S−7, S−9, and S−10) had single-layer structures of mucosa or muscularis propria, six samples (S−1, S−2, S−3, S−4, S−6, and S−8) had double-layer structures of mucosa/submucosa, and one sample (S−11) had a triple-layer structure of mucosa/submucosa/muscularis propria, as shown in the figure. The size and aspect ratio of the images was adjusted for comparison, and the scales for each image were different.

Figure 4 (b) illustrates the autofluorescence intensity images of eleven tissue slices and three tissue blocks. All spectral intensities are shown in different scales for each image, and the scale bar for each image is shown in Figure S1 in the Supporting Information. For tissue blocks, the fluorescence intensity depends not only on the tissue composition but also on the contact environment on quartz glass. The voids and leakage of tissue fluid between tissue and glass influence changes in refractive index, affecting the collection efficiency of autofluorescence signals or generation of oscillating interference patterns. Fluorescence signals were collected through a fixed fiber probe, while moving the sample position at intervals of 100 μm in whisk–broom scanning mode. Impurities, such as dust collected during sample preparation, were distinguished from the signals of interest because they had distinct fluorescence spectra.

Feature extraction using PCA

Since we recorded fluorescence with signal point-by-point mapping of tissue samples, the spectrum at each point was used for ML-guided tissue classification, and many spectra labeled by the histologist were trained for the modeling. For instance, when we set the spectra of S–1, S–2, S–3, S–4, and S–5 as the training set for ML modeling, a total of 18,000 spectra were included in the training set. To reduce the dimensions of numerous spectra, we extracted characteristic features using PCA. PCA scores were employed as training and validation input data, and the number of principal components (PCs) was important for improving modeling efficiency.

The scree plot is a general method for determining the number of PCs to retain in an analysis by plotting the eigenvalue with respect to the number of PCs. Figure 5 (a) represents the eigenvalues for 30 principal factors, with the plot flattening out over four principal factors, representing when we performed PCA analysis on S–1, S–2, S–3, S–4, and S–5. The cumulative variance plot in Figure 5 (b) shows that four factors were sufficient for dimensionality reduction for PCA, which covers over 99.999% of eigenvalue space. Additionally, Figure 5 demonstrates the loading vectors of (c) PC 1–5 and (d) PC 6–10, and the common spectral structures are shown in the plot. The first five loading vectors in Figure 5 (c) represent the main spectral shape of autofluorescence, and the sixth to tenth vectors in Figure 5 (d) incorporate the periodic oscillations resulting from the multiple reflections at the tissue–substrate interface.

ML-based spectro-histology modeling

During the experiments, we recorded fluorescence signals using point-by-point mapping in whisk–broom scanning mode. All the spectra were used for ML-guided tissue classification. A total of 60,457 autofluorescence spectra were analyzed, with 18,000 in the training set and 42,457 in the test set. Since the ML classification was affected by training set selection, feature extraction, and the classification algorithm, we attempted to build a classification model using multiple ML algorithms with varying number of PCs. The accuracy of each ML model was compared using decision-tree, logistic-regression, Bayesian-classification, support-vector-machine, and k-nearest-neighbors algorithms. The bagging tree (random-forest algorithm) among the tested algorithms yielded the most accurate results for our dataset. The bagging tree model is a prominent ensemble learning algorithm that effectively combines the predictions from multiple decision tree

models to enhance the accuracy and robustness of the overall model. These models operate by classifying input data into respective categories.

Figure 6 shows the classification results of fourteen tissues in parallel with images of H&E–stained preparations. Figure 6 (a)–(c) show the H&E–staining images in the left column and classification maps in the right column, with the training set and test set of (b) tissue slices and (c) tissue blocks. Red, blue, green, gray, and black pixels represent mucosa, submucosa, muscularis propria, impurities, and empty background, respectively. The gray part of the noise originates from foreign substances, such as dust. Given that the prediction accuracy of ML varies depending on the data used for training, we attempted the learning with varying training set data. Finally, we divided all samples into training vs. test set, where the samples with clear histological structures were selected as the training set. For the classification results shown in Figure 6, five tissue samples (S–1 to S–5) were employed as the training set, and the other samples were used to test the prediction. Figure 6 (b) and (c) confirm that the most of the ML-based classification results were in concordance with the histological examinations in most tissues.

The prediction accuracy for our model is summarized in Figure 7 (a), with accuracy values of 92.0%, 90.1%, and 91.4% for mucosa, submucosa, and muscularis propria, respectively. Additionally, we investigated the association between model prediction accuracy and the number of principal factors used for input features. Figure 7 (b) is the plot of prediction accuracies with respect to the number of PCs. The out-of-bag (OOB) observations were the observations not included in the bootstrap samples, and OOB classification errors in Figure 7 (c) were used to estimate the prediction error in the bagging tree algorithm. The plot in the figure represents the OOB error of our model with respect to the number of grown trees, and the optimal number of trees were determined based on OOB error calculations.

Spectral analysis of autofluorescence spectra

Autofluorescence spectra of individual tissue layers are shown in Figure 8. We present the spectra of S–1 and S–5 as prototypical spectra of mucosal, submucosal, and muscularis propria layers. S–1 tissue had submucosa and mucosa, and S–5 had only muscle structure. In Figure 8 (b), (c) and (e), it is noted that

fluorescence intensity strength decreased from the submucosal layer to the mucosal layer to the muscularis propria.

Discussion

The mucosal and muscularis propria layers are composed of muscle fibers and epithelial cell line that emit relatively weak fluorescence, while the submucosal layer is composed of collagen fibers that emit strong fluorescence [22]. For colon tissue, it has been reported that superficial submucosal fluorescence originates mainly from collagen [23], and Izuishi et al. attribute the unexpected contribution of collagen to glycosylation and the relatively large amount in colon tissue. Considering the similar composition of gastric tissue and colon, we surmised that the main contributions of the mucosa and submucosa come from collagen components. The relative intensity of the muscularis propria fluorescence differed from that emitted by submucosa for tissues measured in our study.

In the case of S−1 and S−2 in the figures, the photographs and autofluorescence maps match well with the images of H&E–stained preparations in terms of shape; however, the H&E images and photographic/fluorescence images did not always match well for all tissues. There were two reasons for dissimilarities between H&E–staining images and optical photographs. First, there were slight variations between sample compositions and structures were after the slice thickness of 100 μm was achieved because gastric tissue structure varies continuously along the depth. Second, the cryotome slicing process is delicate work and is cutter dependent. For instance, the pressure and speed of the cutting blade affect the shape and size of each slice. Even when the appearance of a photograph substantially differed from an H&E–staining image, the histologist could compare the H&E preparation and the photograph to match the inner structure.

The prediction accuracy of ML modeling depends on the abundance of information, which is determined by the number of PCs in this study because the PCs were used as input features. To find the optimal number of features, we used the scree plot of cumulative variance shown in Figure 5, which depicts the accuracy of the PCA and reduction of dimensionality; four PCs seem to be sufficient for the reduction. Meanwhile, the accuracy of ML modeling increases in accordance with the number of PCs and remain constant over 6 PCs although 4

PCs for dimensionality reduction. To assure the model accuracy of multiple algorithms performed, we used 10 PCs as an input feature for the training and validation.

The prediction of our model is over 90%, as described in the confusion matrix in Figure 7 (a). This accuracy can provide the clinical decision on the depth of invasion of cancer across mucosa and submucosa. Unlike tissue slice, tissue blocks show the gradual structural transition among tissue layers because of their thickness. Depth-resolved spectroscopic study is expected to shed light on resolving the layer of block tissue in three dimensions; previously, we investigated the fat structure of swine tissue with depth-resolved Raman spectroscopy [24].

Our tissue scanner could acquire a spectral image of a 10 × 10 mm^2 tissue section with a resolution of 100 μm in just 15 minutes; it could be significantly improved by using a high-power light source or faster stages. Compared to conventional biopsies, this method is faster as it only requires the analysis of the sliced and non-sliced sample without the time-consuming preparation process. We have been developing the tissue scanner as an intraoperative cancer margin assessment tool. The goal is to reduce the secondary surgery rate caused by remaining cancer tissues by providing real-time feedback of surgical margins in the OR, especially during tissue-conserving surgery. As the system is based on intrinsic tissue signal, it does not require any additional tissue processing. Surgically removed tissues can be measured on the tissue scanner before transferring to the pathologist.

With recent improvement of ML algorithms, not only cancer-noncancer differentiation but also differentiating subtypes of tissues have become possible. Whether ML-based spectro-histology can replace the current standard procedure is controversial because the discussion involves many non-technical aspects of the procedures. By utilizing accumulated database, pathology information on the freshly excised tissues will be available without in-house pathologists. Moreover, we also expect that the method can augment the current function of pathologists by providing a second opinion or supporting difficult cases. However, it is important to note that the use of machine learning technology in the operating room or endoscopy suite would require rigorous testing and validation to ensure its safety and effectiveness. Additionally, the technology would need to be integrated into existing clinical workflows, and trained clinicians should interpret and act on the results.

Conclusion

In this study, we developed a method for rapidly diagnosing and classifying gastric cancer using autofluorescence and ML. We trained the ML algorithm using well-controlled specimen tissue and successfully classified the tissue blocks. We expect that our method will pave the way for label-free in situ tissue classification and optical diagnosis via endoscopy.

Acknowledgement

This research was funded by an NIH National Institute of Biomedical Imaging and Bioengineering grant (41EB015871) and by the National Cancer Center of Korea (2010360). S.Y. Lim, J. I. Kang, and H. M. Kim were supported by a National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (NRF-2020R1A2C2010170)

Disclosure

Soo Yeong Lim, Hong Man Yoon, Myeong-Cherl Kook, Jin Il Jang, Peter T. C. So, Jeon Woong Kang, and Hyung Min Kim have no conflicts of interest or financial ties to disclosure.

References

1. Schnell M, Mittal S, Falahkheirkhah K, Mittal A, Yeh K, Kenkel S, Kajdacsy-Balla A, Carney PS, Bhargava R (2020) All-digital histopathology by infrared-optical hybrid microscopy. Proceedings of the National Academy of Sciences 117:3388–3396. https://doi.org/10.1073/pnas.1912400117

2. Anderson TA, Kang JW, Gubin T, Dasari RR, So PTC (2016) Raman Spectroscopy Differentiates Each Tissue from the Skin to the Spinal Cord: A Novel Method for Epidural Needle Placement? Anesthesiology 125:793–804. https://doi.org/10.1097/ALN.0000000000001249

3. Nazarian S, Gkouzionis I, Kawka M, Jamroziak M, Lloyd J, Darzi A, Patel N, Elson DS, Peters CJ (2022) Real-time Tracking and Classification of Tumor and Nontumor Tissue in Upper Gastrointestinal Cancers Using Diffuse Reflectance Spectroscopy for Resection Margin Assessment. JAMA Surgery e223899. https://doi.org/10.1001/jamasurg.2022.3899

4. Rivenson Y, Wang H, Wei Z, de Haan K, Zhang Y, Wu Y, Günaydın H, Zuckerman JE, Chong T, Sisk AE, Westbrook LM, Wallace WD, Ozcan A (2019) Virtual histological staining of unlabelled tissue-autofluorescence images via deep learning. Nat Biomed Eng 3:466–477. https://doi.org/10.1038/s41551-019-0362-y

5. Yoon HM, Kim H, Sohn DK, Park SC, Chang HJ, Oh JH, Dasari RR, So PTC, Kang JW (2021) Dual modal spectroscopic tissue scanner for colorectal cancer diagnosis. Surg Endosc 35:4363–4370. https://doi.org/10.1007/s00464-020-07929-2

6. Myszczynska MA, Ojamies PN, Lacoste AMB, Neil D, Saffari A, Mead R, Hautbergue GM, Holbrook JD, Ferraiuolo L (2020) Applications of machine learning to diagnosis and treatment of neurodegenerative diseases. Nat Rev Neurol 16:440–456. https://doi.org/10.1038/s41582-020-0377-8

7. Liu Y, Jain A, Eng C, Way DH, Lee K, Bui P, Kanada K, de Oliveira Marinho G, Gallegos J, Gabriele S, Gupta V, Singh N, Natarajan V, Hofmann-Wellenhof R, Corrado GS, Peng LH, Webster DR, Ai D, Huang SJ, Liu Y, Dunn RC, Coz D (2020) A deep learning system for differential diagnosis of skin diseases. Nat Med 26:900–908. https://doi.org/10.1038/s41591-020-0842-3

8. Croce AC, Bottiroli G (2014) Autofluorescence spectroscopy and imaging: a tool for biomedical research and diagnosis. European Journal of Histochemistry 58:. https://doi.org/10.4081/ejh.2014.2461

9. Kuehn A, Graf A, Wenzel U, Princz S, Mantz H, Hessling M (2015) Development of a highly sensitive spectral camera for cartilage monitoring using fluorescence spectroscopy. Journal of Sensors and Sensor Systems 4:289–294. https://doi.org/10.5194/jsss-4-289-2015

10. Wu Y, Qu JY (2006) Autofluorescence spectroscopy of epithelial tissues. JBO 11:054023. https://doi.org/10.1117/1.2362741

11. Sun Y, Sun Y, Stephens D, Xie H, Phipps J, Saroufeem R, Southard J, Elson DS, Marcu L (2011) Dynamic tissue analysis using time- and wavelength-resolved fluorescence spectroscopy for atherosclerosis diagnosis. Opt Express, OE 19:3890–3901. https://doi.org/10.1364/OE.19.003890

12. Wu Y, Xi P, Qu JY, Cheung T-H, Yu M-Y (2004) Depth-resolved fluorescence spectroscopy reveals layered structure of tissue. Opt Express, OE 12:3218–3223. https://doi.org/10.1364/OPEX.12.003218

13. Rice WL, Kaplan DL, Georgakoudi I (2010) Two-Photon Microscopy for Non-Invasive, Quantitative Monitoring of Stem Cell Differentiation. PLOS ONE 5:e10075. https://doi.org/10.1371/journal.pone.0010075

14. Schaefer PM, Kalinina S, Rueck A, von Arnim CAF, von Einem B (2019) NADH Autofluorescence—A Marker on its Way to Boost Bioenergetic Research. Cytometry Part A 95:34–46. https://doi.org/10.1002/cyto.a.23597

15. Seo I, Tseng SH, Cula GO, Bargo PR, Kollias N (2009) Fluorescence spectroscopy for endogenous porphyrins in human facial skin. In: Photonic Therapeutics and Diagnostics V. SPIE, pp 10–15

16. Minamikawa T, Matsuo H, Kato Y, Harada Y, Otsuji E, Yanagisawa A, Tanaka H, Takamatsu T (2016) Simplified and optimized multispectral imaging for 5-ALA-based fluorescence diagnosis of malignant lesions. Sci Rep 6:25530. https://doi.org/10.1038/srep25530

17. Islam MS, Honma M, Nakabayashi T, Kinjo M, Ohta N (2013) pH Dependence of the Fluorescence Lifetime of FAD in Solution and in Cells. International Journal of Molecular Sciences 14:1952–1963. https://doi.org/10.3390/ijms14011952

18. Lin F, Zhang C, Li Y, Yan J, Xu Y, Pan Y, Hu R, Liu L, Liu L, Qu J, Qu J (2021) Human serum albumin gradient in serous ovarian cancer cryosections measured by fluorescence lifetime. Biomed Opt Express, BOE 12:1195–1204. https://doi.org/10.1364/BOE.415456

19. Unger J, Unger J, Unger J, Hebisch C, Phipps JE, Lagarto JL, Kim H, Darrow MA, Bold RJ, Marcu L, Marcu L, Marcu L (2020) Real-time diagnosis and visualization of tumor margins in excised breast specimens using fluorescence lifetime imaging and machine learning. Biomed Opt Express, BOE 11:1216–1230. https://doi.org/10.1364/BOE.381358

20. Awasthi K, Chang F-L, Hsieh P-Y, Hsu H-Y, Ohta N (2020) Characterization of endogenous fluorescence in nonsmall lung cancerous cells: A comparison with nonmalignant lung normal cells. Journal of Biophotonics 13:e201960210. https://doi.org/10.1002/jbio.201960210

21. Lue N, Kang JW, Yu C-C, Barman I, Dingari NC, Feld MS, Dasari RR, Fitzmaurice M (2012) Portable Optical Fiber Probe-Based Spectroscopic Scanner for Rapid Cancer Diagnosis: A New Tool for Intraoperative Margin Assessment. PLOS ONE 7:e30887. https://doi.org/10.1371/journal.pone.0030887

22. DaCosta RS, Lilge LD, Kost J, Cirroco M, Hassaram S, M.d NEM, Wilson BC (1997) Confocal fluorescence microscopy, microspectrofluorimetry, and modeling studies of laser-induced fluorescence endoscopy (LIFE) of human colon tissue. In: Laser-Tissue Interaction VIII. SPIE, pp 98–107

23. Izuishi K, Tajiri H, Fujii T, Boku N, Ohtsu A, Ohnishi T, Ryu M, Kinoshita T, Yoshida S (1999) The Histological Basis of Detection of Adenoma and Cancer in the Colon by Autofluorescence Endoscopic Imaging. Endoscopy 31:511–516. https://doi.org/10.1055/s-1999-57

24. Kang JW, Lim SY, Galindo LH, Yoon H, Dasari RR, So PTC, Kim HM (2020) Analysis of subcutaneous swine fat via deep Raman spectroscopy using a fiber-optic probe. Analyst 145:4421–4426. https://doi.org/10.1039/D0AN00707B

Figure legends

Figure 1. (a) Sampling of human tissues for histological and spectroscopic examination. (b) Schematic diagram of instrumental setup.

Figure 2. The scheme of the machine-learning based spectro-histology classification model. The upper black dotted box corresponds to the training step in modeling, and the bottom red box corresponds to the test of our model with tissue blocks and slices.

Figure 3. (a) and (b) H&E staining image, (c) and (d) optical photographic image, (e) and (f) autofluorescence intensity maps, and (g) and (h) peak shift maps for tissue S−1 and (e)-(h) tissue S−2, respectively. Red and blue dotted lines correspond to mucosal and submucosal regions, respectively.

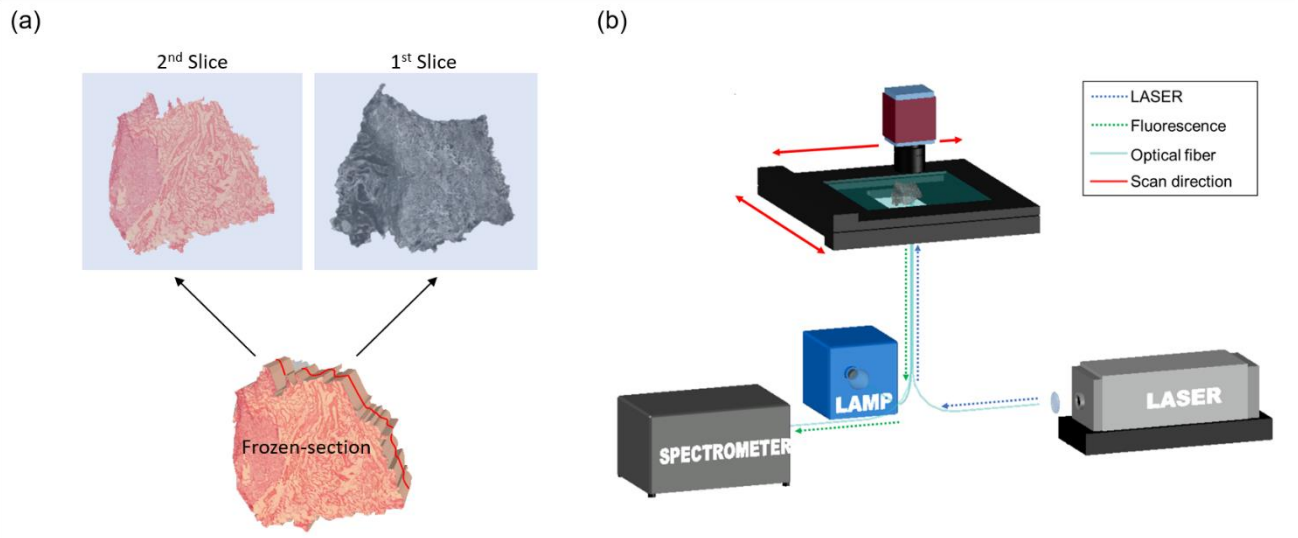
Figure 4. (a) Histological diagnosis of tissue slices (S−1 to S–11) and tissue blocks (B−1 to B–3): the regions outlined in red, blue, and green dotted lines represent the mucosal, submucosal, and muscularis domains, respectively. The classification was determined by a histologist. (b) Fluorescence intensity

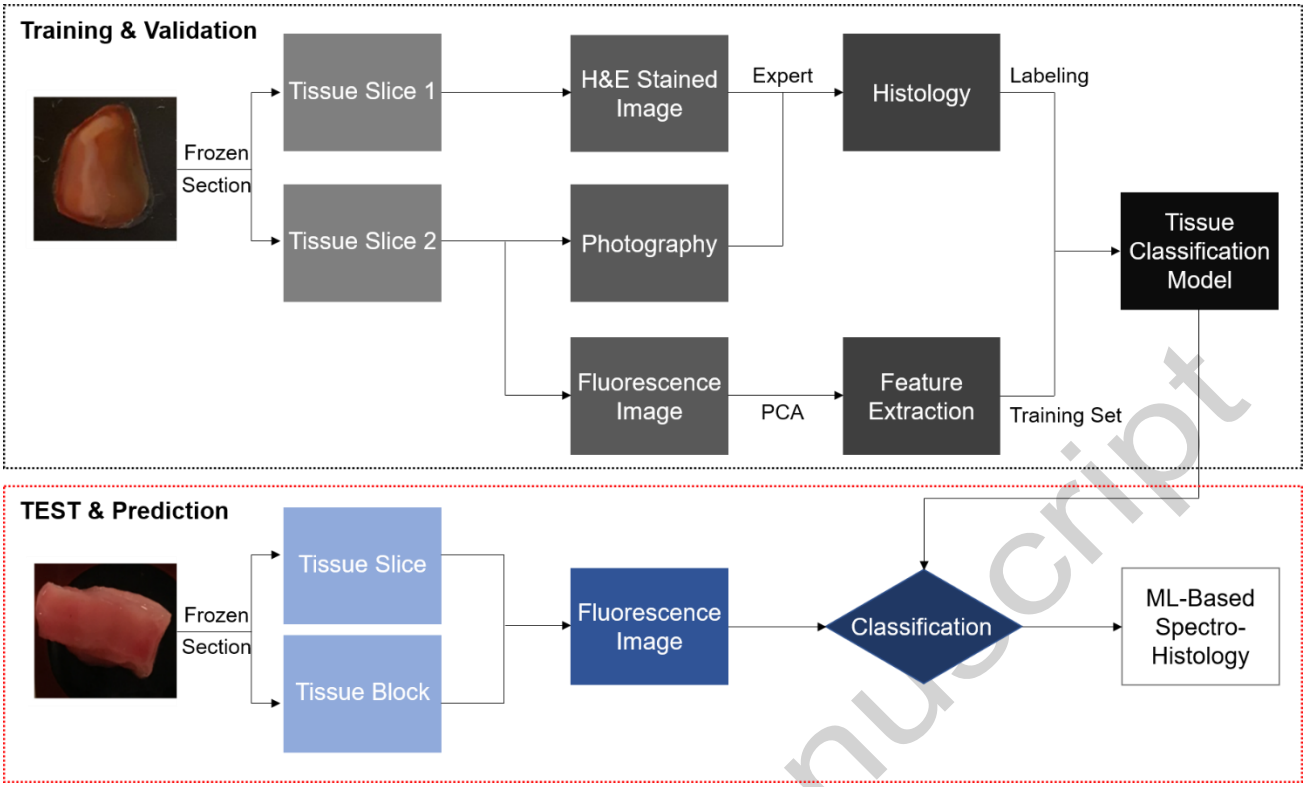
Figure 5. (a) The Scree plot and (b) cumulative variance with respect to principal components for machine learning modeling. (c)–(d) Calculated loading vectors of PC1 to PC5 of the fluorescence spectra training set tissues

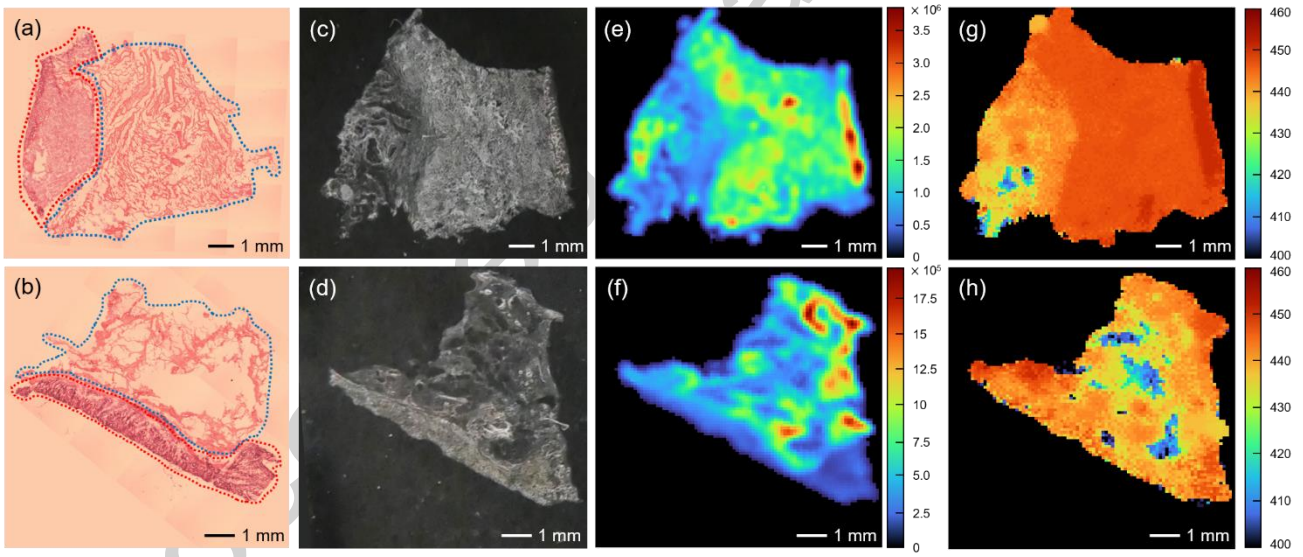
Figure 6. H&E–staining images (left column) and machine learning-based classification maps (right columns) of tissue slices belonging to the (a) training set and (b) test set. (c) Photographs and classfication maps of tissue blocks.

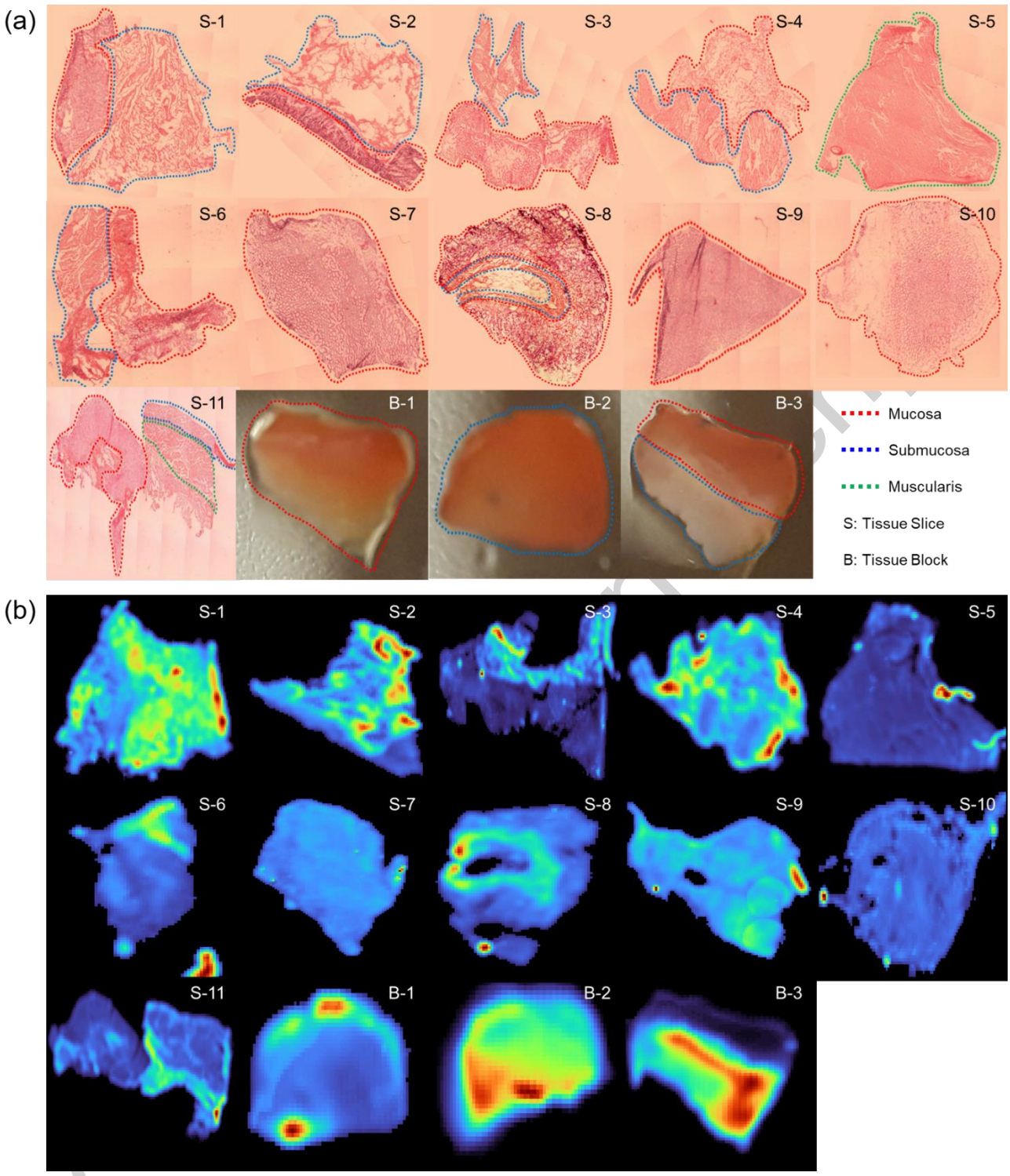
Figure 7. The accuracy of our machine learning–based model with respect to the number of principal components selected for dimensionality reduction and the table of true-positive rate (TPR) and false-negative rate (FNR), (b) Out-of-bag observation plot for the bagging tree model.

Figure 8. Spatially resolved autofluorescence image of (a) S1 and (b) S-5 tissue slices and observed spectra obtained from the (c) mucosal, (d) submucosal, and (e) muscularis regions.








(a)
S-1
S-2
S-3
S-4
S-5
S-6
S-7
S-8
S-9
S-10
S-11
B-1
B-2
B-3
Mucosa
Submucosa
Muscularis
S: Tissue Slice
B: Tissue Block
(b)
S-1
S-2
S-3
S-4
S-5
S-6
S-7
S-8
S-9
S-10
S-11
B-1
B-2
B-3