



HAL
open science

Pre-cancerous stomach lesions detections with Multispectral-augmented endoscopic prototype

Alexandre Krebs, Yannick Benezeth, Thomas Bazin, Franck Marzani,
Dominique Lamarque

► **To cite this version:**

Alexandre Krebs, Yannick Benezeth, Thomas Bazin, Franck Marzani, Dominique Lamarque. Pre-cancerous stomach lesions detections with Multispectral-augmented endoscopic prototype. Applied Sciences, 2020, 10 (3), pp.795. 10.3390/app10030795 . hal-02468198

HAL Id: hal-02468198

<https://u-bourgogne.hal.science/hal-02468198>

Submitted on 5 Feb 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Article

Pre-cancerous stomach lesions detections with Multispectral-augmented endoscopic prototype

Alexandre Krebs ^{1,*} , Yannick Benezeth ¹, Thomas Bazin ², Franck Marzani ¹ and Dominique Lamarque ²

¹ ImViA EA 7535, Univ. Bourgogne Franche-Comté; firstname.lastname@u-bourgogne.fr

² Hôpital Ambroise Paré de Boulogne Billancourt; thomasbazin@club-internet.fr, dominique.lamarque@aphp.fr

* Correspondence: alexandre.krebs@u-bourgogne.fr

Version January 14, 2020 submitted to Appl. Sci.

Featured Application: The presented work has been developed to be used in operating rooms, during upper gastrointestinal exploration of the stomach. The system can distinguish between healthy mucosa, chronic gastritis and intestinal metaplasia.

Abstract: In this paper, we are interested in the *in vivo* detection of pre-cancerous stomach lesions. Pre-cancerous lesions are unfortunately rarely explored in research papers as most of them are focused on cancer detection or conducted *ex-vivo*. To this purpose, a novel prototype is introduced. It consists of a standard endoscope with multispectral cameras, an optical setup, a fiberscope, and an external light source. Reflectance spectra are acquired *in vivo* on 16 patients with a healthy stomach, chronic gastritis, or intestinal metaplasia. A specific pipeline has been designed for the classification of spectra between healthy mucosa and different pathologies. The pipeline includes a wavelength clustering algorithm, spectral features computation, and the training of a classifier in a "leave one patient out" manner. Good classification results, around 80% have been obtained, and two attractive wavelength ranges were found in the red and near-infrared ranges: [745, 755 nm] and [780, 840 nm]. The new prototype and the associated results give good arguments in favor of future common use in operating rooms, during upper gastrointestinal exploration of the stomach for the detection of stomach diseases.

Keywords: multispectral imaging, stomach lesions, gastroendoscopy

1. Introduction

Based on GLOBOCAN 2018 data [1], stomach cancer is the sixth most common and the fourth most deadly cancer. Unfortunately, the five-year survival rate of this cancer is about 30% and has not increased since the 90's century. Improving the survival rate of this cancer is important to have a better diagnosis and a better understanding of the early stages of cancer. Thus, a promising way is to focus on pre-cancerous lesions and to use new imaging modalities with supervised learning techniques.

Nowadays, practitioners mostly use upper gastrointestinal endoscopy, under white light (WL) or Narrow Band Imaging (NBI), to diagnose stomach pathologies. Recognizing inflammatory lesions is still a hard task as they are often undetectable, especially under white light. Lesions are often too small to be seen at a macroscopic level. Thus, biopsies are performed in a systematic and non-oriented way. Little pieces of the stomach are collected and analyzed *a posteriori* in a laboratory to make a diagnosis. As the location of the samples is randomly chosen, the result could be different from one location to another and depends on the endoscopist's experience.

31 NBI is a system created by Olympus, made of two or three narrow filters centered on the
32 Hemoglobin absorption peaks, thus showing vascularity darker than the surroundings. The system has
33 been successfully used to diagnose Barrett's oesophagus [2] and allows good visualization of specific
34 patterns of chronic gastritis and intestinal metaplasia [3].

35 The main alternative to NBI is the "Fuji Intelligent Chromo Endoscopy" (FICE) system by FUJI [4].
36 FICE estimates multispectral images, between 400 and 695 nm, with a 5 nm step, from a standard
37 endoscopic image. Then, the practitioner can choose to display three images among them in red, green,
38 and blue channels. Ten settings are pre-configured. FICE has been proven to be useful for the detection
39 of complete and incomplete metaplasia [5]. The significant difference between FICE and NBI is that
40 FICE proposes a software improvement with the computation of spectral images, and NBI proposes
41 an optics improvement with specific filters [6].

42 Fuji recently proposed two new medical systems, the Blue Light Imaging (BLI) and the Linked
43 Color Imaging (LCI). BLI allows better visualization of superficial vascular and mucosal patterns with
44 a sensor sensitivity focussed on the Hemoglobin absorption peak at 410 nm [7]. LCI has a specific
45 pre-processed composition of the light and applies image processing techniques to increase the color
46 contrast, thus improving also the delineation of pathologies [8].

47 Lately, endoscope manufacturers (e.g., FUJI, Olympus, Pentax) greatly improved image
48 resolutions [9]. However, fewer improvements are introduced for the detection of the lesions. A
49 better system is the magnifying NBI [10]. Uedo *et al.* used this system to link the apparition of white
50 and blue crests on the stomach's surface to metaplasia [11].

51 These systems show that non-conventional imaging, more specifically the use of spectral
52 information, is promising to improve the detection of stomach lesions. Multispectral or hyperspectral
53 systems use reflectance or fluorescence to get meaningful information on tissue composition. In
54 particular, reflectance represents the fraction of light coming back from the tissue as a function of
55 the wavelengths. Fluorescence is the emission of light by tissues in response to a shorter exciting
56 wavelength. The light response can be influenced in a different way by amount and type of each of a
57 different biomolecule acting as chromophore or fluorophore present in the tissue. The overall light
58 signal recorded from the tissue will result from the different effects of various components, and that is
59 the reason for which it depends on the tissue composition.

60 For example, the system made by Gu *et al.*, made of seven bands between 465 nm and 665 nm.
61 The three bands, 645 nm, 545 nm, and 495 nm, have been selected as the triplet that minimizes mutual
62 information. These three bands are then used to create an enhanced RGB image for better visualization.
63 The system has been tested for ulcer detection [12].

64 Visualization helps to detect pathologies, but the interpretation remains in endoscopists' hands.
65 The diagnosis thus depends on the endoscopist's experience and is not necessarily correlated to
66 histological results. It is essential to improve the detection thanks to signal and image processing
67 methods adapted to non-conventional modalities. For example, Kiyotoki *et al.* used a hyperspectral
68 camera between 400 and 800 nm to characterize the spectral response of resected tumors [13]. 14
69 patients have been studied by selecting spectra on healthy and cancerous areas. Ogihara *et al.* have
70 conducted similar works. They also used a hyperspectral camera in the range [400, 800 nm] to analyze
71 resected stomachs. They noticed a high inter-patient variability that led them to normalize all spectra
72 according to the mean spectra taken on healthy areas. These works are promising but are all conducted
73 *ex vivo*.

74 A few works considering *in vivo* exploration with multispectral systems exist in the literature.
75 In 2017, Hohmann *et al.* demonstrated that multispectral data was useful for the detection of
76 adenocarcinoma [14]. Their system is composed of a standard endoscope and a modified light source.
77 The spectral range used in this study is [438, 628 nm]. They have tested several classifiers including
78 Random Forest, RobustBoost, Adaboost or Support-Vector Machine (SVM). Their classification
79 accuracy was around 60%. Makhlof *et al.* worked on a confocal microendoscope with two modes. One

80 with a gray-level image to observe the fluorescence response and the other to obtain a multispectral
81 image of fluorescence [15].

82 An attractive aspect of hyperspectral acquisitions is the link between spectra and biological
83 parameters. A spectrum is the sum of several chromophores' contributions. Thus, an unmixing process
84 can be done. For example, Grosberg *et al.* characterized gastric mucosa with hyperspectral two-photon
85 microscopy. They obtained images at cellular levels. Unmixing was performed to distinguish four
86 tissue types: epithelium, *lamina propria*, collagen, and lymphatic tissue [16]. Bergholt *et al.* acquired
87 spectra via Raman spectroscopy [17]. Spectra were also unmixed to find the relative contribution of
88 DNA, proteins, lipids, and glycoproteins. Relative contributions were successfully used to classify
89 four tissue types: healthy, intestinal metaplasia, dysplasia, and adenocarcinoma. They have obtained
90 a good sensitivity between 80% and 90% and a specificity higher than 90% with a Partial Least
91 Squares-Discriminant Analysis PLS-DA).

92 Martinez *et al.* developed a system with a filter wheel for acquisitions at six different wavelengths.
93 The principle drawback is that images have to be registered. A "Nearest Neighbors" classifier and an
94 SVM classifier (with a linear or a gaussian kernel) have been used on 5x5 patches to detect pre-cancerous
95 states in the stomach [18]. The best results have been obtained with the SVM classifier with a gaussian
96 kernel with an accuracy of 77%. Except for this work, we rarely find other papers related to *in vivo*
97 stomach exploration.

98 Other techniques have been developed to help the detection of pathologies with imaging systems.
99 For example, contrast agents can be used to characterize the stomach wall like chromoendoscopy [19].
100 Some agents are absorbed by cells or accumulate on peaks and valley of the mucosa, thus showing the
101 structure of the cells. The use of contrast agents is not mandatory. Some biological substances can be
102 stimulated by light, acting as fluorophores. This phenomenon is exploited by fluorescence techniques:
103 by exciting fluorophores with at a certain wavelength, they re-emit light in a higher wavelength [20].
104 In practice, collagen is the most important in sub-mucosa but we can also find Pyridoxal phosphate,
105 riboflavin, phospholipid and porphyrins I and II (PPS) [21]. To the opposite, Chromophores
106 (Hemoglobin for instance) does not re-emit light but absorb it. The result of fluorophores and
107 chromophores is usually visualized by a multispectral camera [22].

108 In the present paper, a new prototype is introduced for acquiring NBI endoscopic images and
109 multispectral images. We demonstrate how this prototype can be used for the recognition of several
110 pre-cancerous gastric lesions such as chronic gastritis or intestinal metaplasia. In section 2, we first
111 introduced a previous preclinical study on mice's stomach [23]. Then a new clinical prototype based
112 on multispectral cameras is detailed and the acquisition of data and the signal processing methods are
113 given. In section 3, the classification results are presented and discussed.

114 2. Materials and Methods

115 2.1. Preclinical study on mice's stomach

116 In a previous paper, the reflectance of the stomach has been studied on mice infected by *H. pylori*,
117 a bacteria involved in the development of gastritis on mice and humans [23]. The mice were sacrificed
118 at different time-point after infection and their stomach resected for multispectral analysis with a
119 spectrometer. The work on mice can be seen as a preliminary work in which a classification pipeline
120 has been developed. This pipeline allowed us to separate spectra between control and inflamed
121 mice with good accuracy of 98% and, more interestingly, to identify two wavelength ranges in the
122 red/infrared ranges that are discriminant for the classification: [620, 668 nm] and [668, 950 nm]. This
123 study proves that analyzing reflectance spectra is a promising way to detect pathologies. We re-use the
124 methodology and the pipeline for the processing of *in vivo* human's data as presented in the following
125 paragraphs.

126 The data that are acquired consist of a set of spectra. More precisely, two sets of spectra, one in the
127 visible range with wavelengths between 400 and 630 nm with a 10 nm step (thus, 24 points by spectra)

128 and in the near-infrared with wavelengths between 610 and 840 nm with a 10 nm step (also 24 points
 129 by spectra). It is interesting to examine the correlation between the wavelengths (*i.e.* the 24 points).
 130 Figures 1a and 1b show this correlation. The value of each point (i,j) of these correlation matrices
 131 corresponds to the correlation between reflectance spectra at wavelengths i and j . In both images, red
 132 squares can be observed along the diagonal. These squares indicate that some wavelengths are highly
 133 correlated, especially the ranges around [410, 480 nm] and [520, 580 nm] in the visible and around [610,
 134 690 nm] and [780, 840 nm] in the near-infrared. This observation motivates us to cluster the spectral
 135 bands into coherent groups. Figures 1c and 1d show the results of the band clustering algorithm. A
 136 median spectrum is presented with vertical black lines allowing the visualization of borders. For the
 137 visible (VIS) camera, eight bands are obtained, and, for the near-infrared (NIR) camera, seven bands
 138 are obtained.

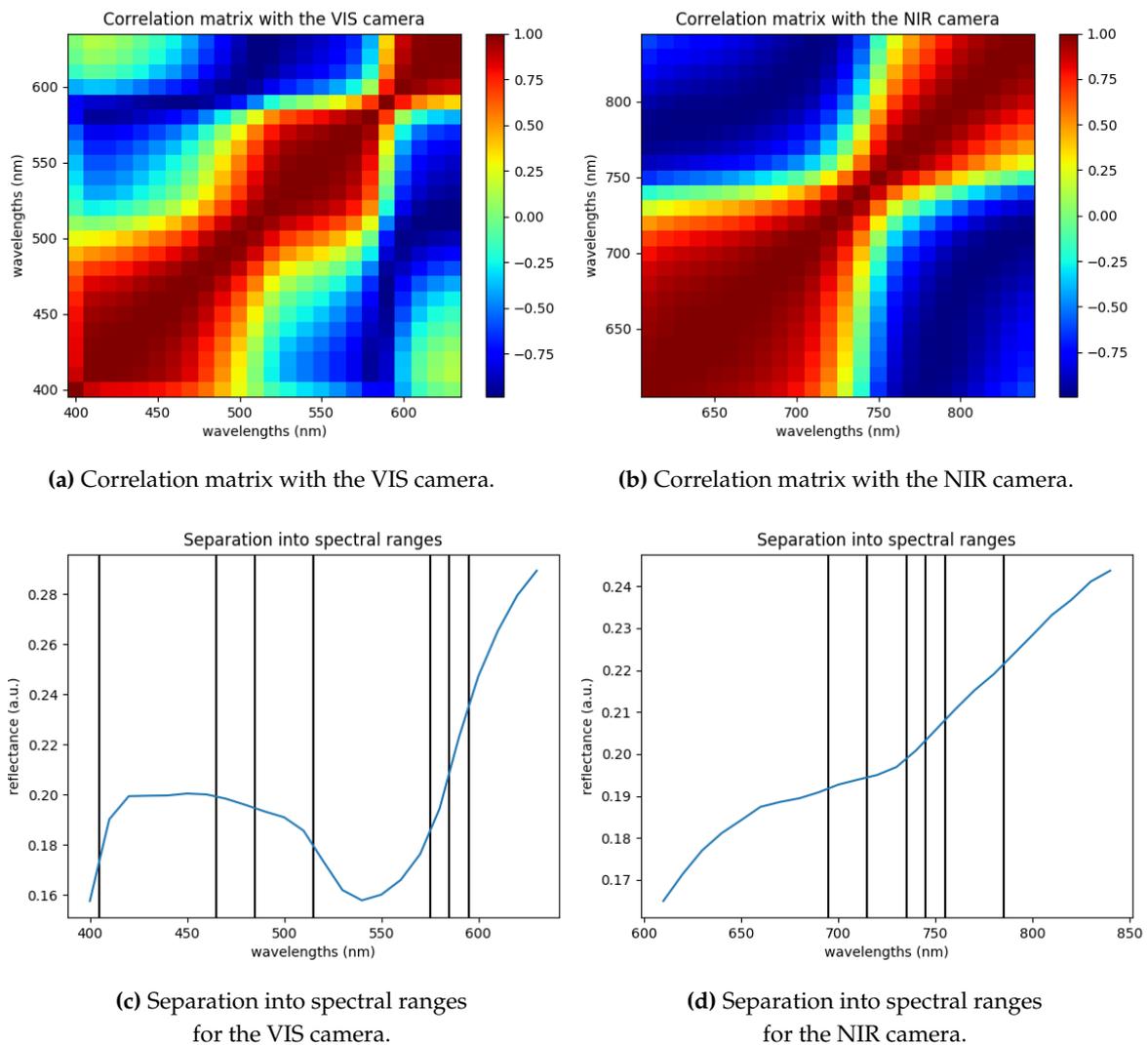


Figure 1. Correlation between wavelengths.

139 To obtain these results, we recursively merge the neighboring bands that are correlated more than
 140 90%. At every iteration, the correlation between neighboring bands is computed, and the two most
 141 correlated bands are merged. The detailed algorithm is presented in the box 1. The algorithm uses the
 142 same scheme as hierarchical clustering with an additional connectivity constraint. At the beginning of
 143 the algorithm, every wavelength is a cluster, and at each iteration, two clusters are merged. A version
 144 of this algorithm exists with connectivity constraints. In our case, only neighboring wavelengths can be

145 merged. This algorithm is unsupervised because the label (*i.e.* control, chronic gastritis, or metaplasia) is not used.

Algorithm 1: Pseudo code of the wavelengths clustering algorithm

Input : $M = a (N_{spectra}, N_{bands})$ matrix containing all spectra

$\tau = a$ constant equal to 0.9 by default

Output: $M_{reduced} = A (N_{spectra}, N_{comp})$ matrix with $N_{comp} \leq N_{bands}$

$m \leftarrow 1, 0$

$M_{reduced} \leftarrow M$

repeat

$c = \text{correlation}(M_{reduced})$, the correlation between neighboring bands

$m \leftarrow \max(c)$

$x, y \leftarrow \text{argmax}(c)$

 MERGE columns x et y in $M_{reduced}$

until $m > \tau$

146

The wavelengths that have been clustered are reduced to their spectral mean μ_i on the cluster range according to equation 1:

$$\mu_i = \frac{1}{N_i} \sum_{\lambda=\lambda_i}^{\lambda_{i+1}} S_\lambda \quad (1)$$

147 where N_i is the number of wavelengths in the range $[\lambda_i, \lambda_{i+1}[$ and S_λ is the reflectance of the spectra at
148 λ .

149 After the reduction of the number of wavelengths, features are extracted. Three types of features
150 are investigated: either the reduced bands are directly used, or the features are computed as the
151 ensemble of all possible subtractions or divisions. More formally, we denote the ratios as $R_{i,j}$ and the
152 differences as $D_{i,j}$. They are defined in the equations 2 and 3:

$$R_{i,j} = \frac{\mu_i}{\mu_j}, \quad (2)$$

$$D_{i,j} = \mu_i - \mu_j. \quad (3)$$

153 Then, a univariate selection step is performed to reduce the complexity of the classifier. Reducing
154 complexity helps us to avoid over-fitting problems. The F-test of ANalysis Of VAriance (ANOVA) [24]
155 is used to extract the k most discriminant features. It consists of the computation of a p-value, the
156 probability that two distributions are the same. The lower the p-value is, the more discriminant the
157 feature is.

158 The classification consists of a simple SVM classifier with a linear kernel. Leave one patient out
159 cross-validation is performed. At every iteration, all spectra except from one patient are used to train
160 the classifier. The complete pipeline is written in Python thanks to the Scikit-learn library [25] and is
161 summarized in the box 2.

Algorithm 2: Classification pipeline

Input : $M = a (N_{spectra}, N_{bands})$ matrix containing all spectra
 $y = a N_{spectra}$ elements vector corresponding to labels

Output: $r = a N_{spectra}$ elements vector corresponding to classification results
 $M_{reduced}$ = Reduction of M by algorithm 1

foreach patient p do

- M_{test} = Spectra from patient p in $M_{reduced}$
- $M_{training}$ = difference between $M_{reduced}$ and M_{test}
- $X_{training}$ = COMPUTE features of $M_{training}$
- X_{test} = COMPUTE features of M_{test}
- F = SELECTION of the k best features according to $X_{training}$
- $X_{F,training}$ = Features F de $X_{training}$
- $X_{F,test}$ = Features F de X_{test}
- clf = A SVM classifier with linear kernel
- TRAIN clf with $X_{F,training}$
- $r[p]$ = TEST clf with $X_{F,test}$

end foreach

162 2.2. Description of the prototype

163 The system that has been designed in this study can acquire NBI images and multispectral images.
164 A standard endoscope (Olympus Exera III) that already exists in most operating rooms is used. Besides,
165 there are two multispectral cameras (XIMEA based on CMOSIS CMV2000 technology [26]), one in the
166 visible range (450-620 nm) with 16 bands and one in the red and near-infrared (600-1000 nm) with 25
167 bands. The cameras are connected to a computer with USB 3.0 ports. For convenience, these cameras
168 are respectively called VIS and NIR. To acquire images inside the stomach, a fiberscope (microflex
169 fiberscope from ITConcepts, 2.5 m length, and 2.5 mm thick [27]) is used. The fiberscope contains
170 emitting fibers to illuminate the stomach and receiving fibers to catch images. It is to notice that this
171 fiberscope must be sterilized before each use, this is one of the current limitations as the multispectral
172 system cannot be used during the cleaning of the fiberscope procedure. To connect the fiberscope
173 to the cameras, a beamsplitter system is used (TwinCam from Cairn Research [28]). A dichroic
174 filter splits light rays into two according to the wavelength ranges [470-620 nm] and [600-975 nm].
175 Additionally, a filter is added to avoid Ultraviolet (UV) light. In our case, a 1000W powerful external
176 light source is needed because the fiber and the sensitivity of the camera induce a loss of intensity.
177 Thus a Mercury-Xenon light source controlled by a controller device helps to illuminate the stomach
178 (Newport reference: 66924-1000XF-R1)

179 The system has to be calibrated spatially and spectrally. The spatial and spectral calibration have
180 been presented in a previous paper [29]. Spatial calibration is needed to correct the radial distortion
181 and to register the NBI image with the multispectral images. The calibration is done by using the
182 tip of the fiberscope, which is visible in the NBI images. Spectral calibration is required because the
183 filters of the cameras are based on the Fabry-Perot structure. They have a primary response but also a
184 secondary response. Moreover, it exists noisy response out of the initial expected wavelength range
185 called crosstalk problem. The calibration has been done by learning a multilinear transformation on
186 a set of patches of the ColorChecker®. More concretely, the 24 bands ([400, 630 nm] with a step of
187 10 nm) are deduced from the 16 original bands with the multiplication of a 16×24 matrix and, in the
188 same way from the 24 bands ([610, 840 nm] with a step of 10 nm) are deduced from the 25 original
189 bands with the multiplication of a 25×24 matrix. After this transformation, all spectra are normalized
190 according to their l_2 norm, meaning that all spectra are divided by their sum of squares.

191 In a practical situation, the endoscopist inserts the fiberscope in the operating channel of the
192 endoscope, as described in image 2. This operating channel is usually used to insert tools like pincer

193 or cutting/grasping instruments. The endoscopy usually lasts 1 or 2 minutes during which the
 194 multispectral cameras acquire images every second approximatively. A specific acquisition software
 195 has been developed in C# programming language with .NET 4.5 framework and the library xiAPI that
 196 allows the control of the multispectral cameras [30]. This software manages the data saving and adapts
 197 the exposition time of the cameras automatically. The cameras have been optically calibrated such that
 198 the focus is approximately 2cm. This distance is the mean distance that exists between the fiberscope
 199 tip and the stomach wall.

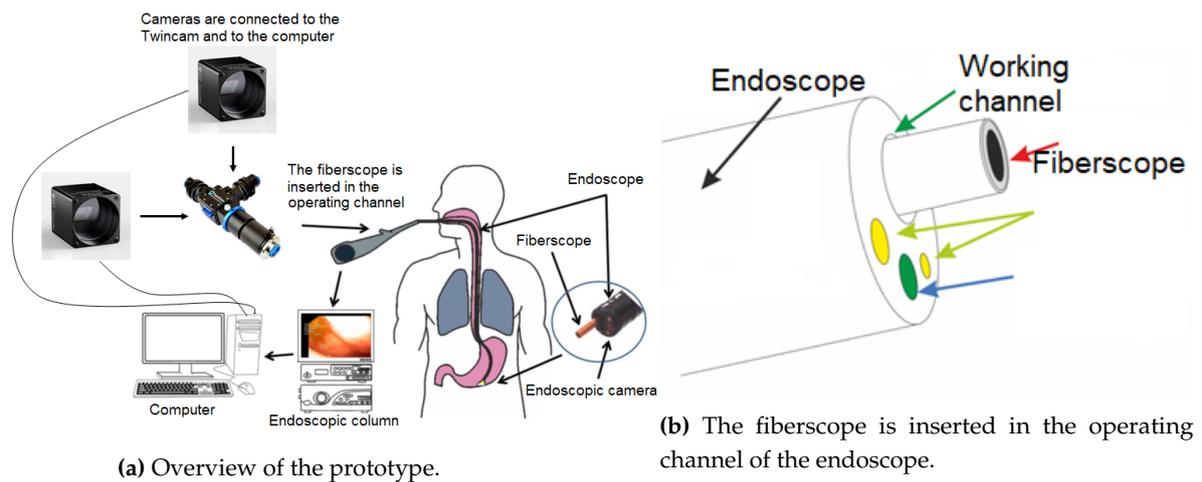
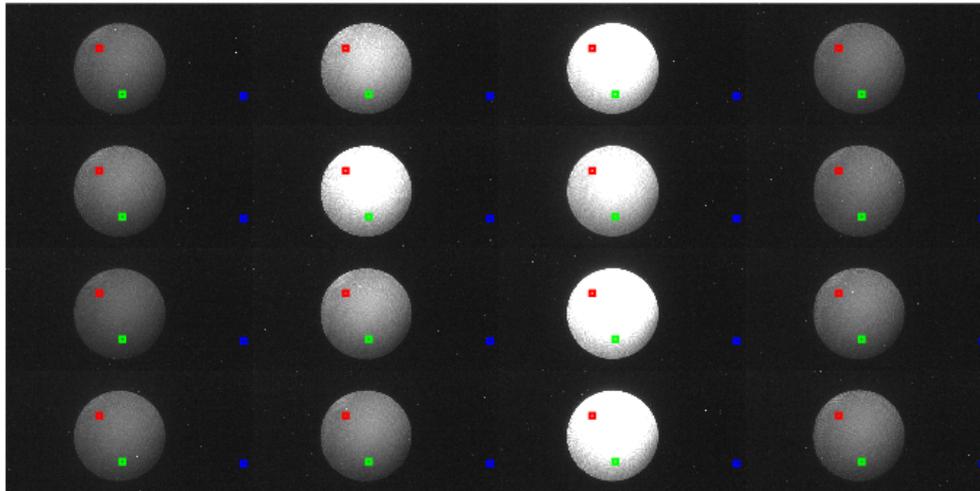


Figure 2. Functional scheme of the prototype.

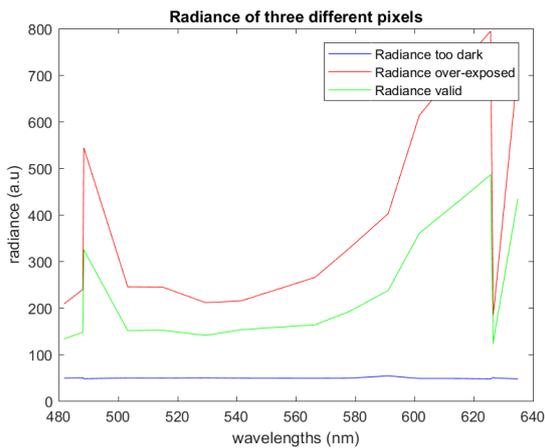
200 2.3. Data acquisition and preprocessing

201 The presented bimodal system can acquire multispectral images at approximately one image per
 202 second. This exposure time is due to the low power of light that comes back in the fiber. Figure 3a
 203 shows an example of a multispectral image acquired in the stomach. The image is rearranged in 16
 204 small images to show the 16 bands of the VIS camera. We can see the disk in the center, which is the
 205 area covered by the fiberscope. All images do not necessarily contain useful information. In general,
 206 multispectral images are often blurry. This phenomenon is due to the movements of the endoscope,
 207 the poor spatial resolution, the fiberscope, and the high exposure time. Then, it is challenging to use
 208 textural information as some parts of the image could be too dark or over-exposed. Therefore, we make
 209 sure to keep only pixels whose maximum is under 512. Indeed, the camera's technical documentation
 210 specifies that, if the 10th bit is set to 1, then the pixel is considered saturated. In the same manner, pixels
 211 whose spectral mean does not exceed 128 are considered too dark to be exploited. To understand this
 212 choice, we present an example in figure 3b. Three pixels are compared: in blue, a pixel which is too
 213 dark, in green, a valid pixel, and in red, a saturated pixel.

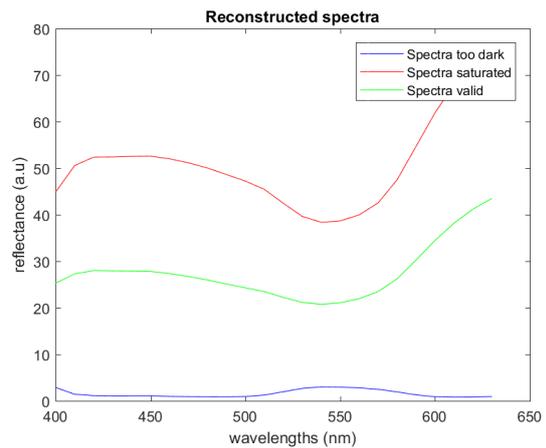
214 A linear transformation described in a previous paper [29] is applied to the spectral radiance to
 215 get estimated reflectance spectra (figure 3c) *i.e.* to estimate the 24 points in the VIS and the 24 points in
 216 the NIR from the original 16 bands of the VIS camera and the 25 bands of the NIR camera.



(a) Multispectral image acquired inside the stomach.



(b) Raw data obtained from the pixels of the multispectral image.



(c) Corresponding estimated spectra.

Figure 3. Example of the extraction of spectra from the multispectral image.

217 As the multispectral images are difficult to interpret directly, an experienced gastro-endoscopist
 218 selected NBI endoscopic images in which the fiber is focused on a significative zone of the pathology.
 219 Then we can select the temporally nearest multispectral image for the analysis.

The image should then be filtered with a low-pass filter to reduce the noise. Although it exists a various number of filtering like the use of wavelets [31]. In our case, we have chosen to use a simple pyramidal filter:

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 3 & 2 & 1 \\ 2 & 4 & 6 & 8 & 6 & 4 & 2 \\ 3 & 6 & 9 & 12 & 9 & 6 & 3 \\ 4 & 8 & 12 & 16 & 12 & 8 & 4 \\ 3 & 6 & 9 & 12 & 9 & 6 & 3 \\ 2 & 4 & 6 & 8 & 6 & 4 & 2 \\ 1 & 2 & 3 & 4 & 3 & 2 & 1 \end{bmatrix} .$$

220 Then, spectra are taken on a grid by considering 1 pixel over 7 (vertically and horizontally). We
 221 gathered data for 16 patients. The table 1 classifies them according to their histological results. For
 222 all these patients, 200 valid spectra were randomly taken among the valid spectra to have the same
 223 number of spectra for each patient.

Table 1. Distribution of multispectral data according to histology.

Healthy	Chronic Gastritis	Intestinal Metaplasia
4	7	5

224 Figures and results from the two cameras are intentionally separated. Cameras acquire images
 225 independently, and in practice, if no pixel in an image is valid (*i.e.* either too dark or over-exposed),
 226 the image is removed.

227 The two figures 4 show median spectra of the 16 patients after calibration, respectively for VIS
 228 camera (figure 4a) and NIR (figure 4b). The color indicates histological results: blue for control, green
 229 for chronic gastritis, and red for intestinal metaplasia. We do not see specific behavior in the visible
 230 range, but control spectra seem clustered for the NIR camera, they have a lower derivative. On the
 231 spectral range [610, 730 nm], plots corresponding to metaplasia are under the plots of control class and
 232 inversely for the spectral range [730, 840 nm].

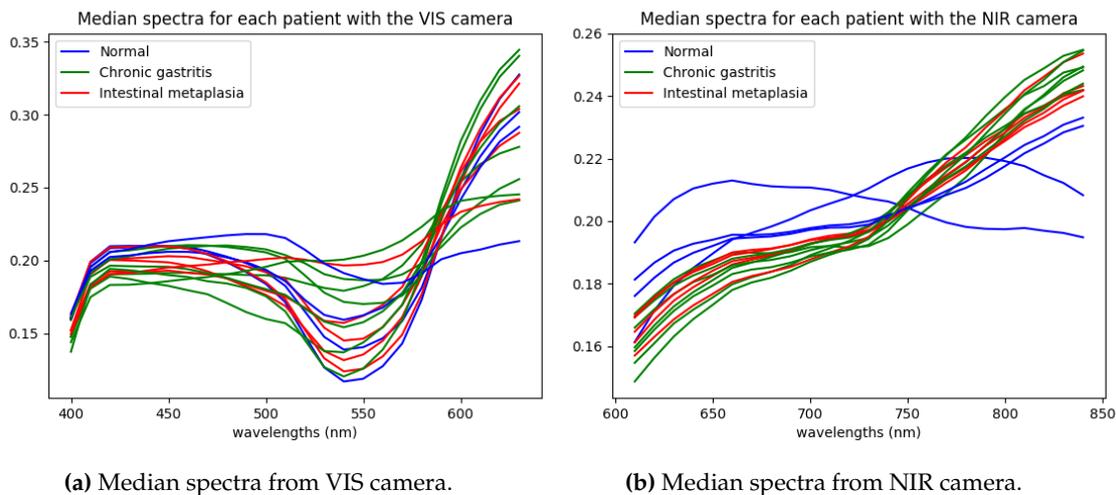


Figure 4. Acquisitions with VIS and NIR cameras.

233 The obtained spectra are then processed by the pipeline detailed in section 2. The next sections
 234 give the classification results obtained separately with VIS and NIR data.

235 3. Results

236 In this section, the classification results are presented. A support vector machine (SVM) with
 237 a linear kernel is trained in a "leave one patient out" manner. The three types of features are tested,
 238 *i.e.* the reduced bands, the subtractions features, or the divisions' features. It is to note that two
 239 hyperparameters can be tuned: the number of features retained in the feature selection step k and the
 240 regularization parameter c that controls the balance between misclassification rate and the size of the
 241 margin. A lower value of c leads to a better misclassification rate but also a smaller margin, thus being
 242 less robust to over-fitting.

243 In this section, data acquired by the two cameras are processed separately. The table 2 presents
 244 the results obtained with the VIS camera and the table 3, the one obtained with the NIR camera.

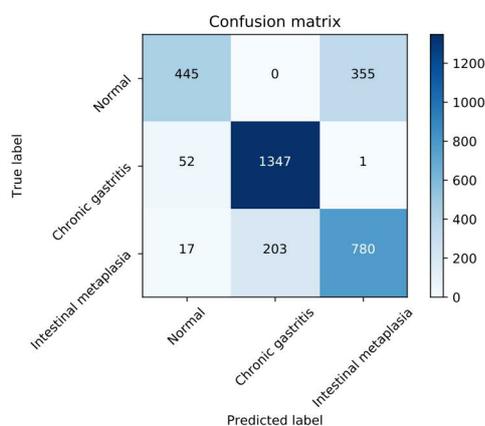
Table 2. Classification results with the VIS camera.

	Optimal hyperparameter	Class	Precision	Recall	F1-Score
Reduced Bands	c = 0,0001 k = 1	Healthy	0.98	1.00	0.99
		Chronic gastritis	0.58	0.99	0.73
		Intestinal metaplasia	0.00	0.00	0.00
		Average	0.52	0.66	0.57
Divisions	c = 0,0001 k = 1	Healthy	1.00	1.00	1.00
		Chronic gastritis	0.58	1.00	0.74
		Intestinal metaplasia	0.00	0.00	0.00
		Average	0.53	0.67	0.58
Subtractions	c = 0,0001 k = 1	Healthy	1.00	1.00	1.00
		Chronic gastritis	0.58	1.00	0.74
		Intestinal metaplasia	0.00	0.00	0.00
		Average	0.53	0.67	0.58

Table 3. Classification results with the NIR camera.

	Optimal hyperparameter	Class	Precision	Recall	F1-Score
Reduced Bands	c = 0,001 k = 1	Healthy	0.88	0.51	0.64
		Chronic gastritis	0.59	0.94	0.73
		Intestinal metaplasia	0.73	0.37	0.49
		Average	0.73	0.61	0.62
Divisions	c = 0,001 k = 1	Healthy	0.87	0.56	0.68
		Chronic gastritis	0.87	0.96	0.91
		Intestinal metaplasia	0.69	0.78	0.73
		Average	0.80	0.80	0.80
Subtractions	c = 0,001 k = 1	Healthy	0.86	0.57	0.69
		Chronic gastritis	0.86	0.96	0.91
		Intestinal metaplasia	0.69	0.76	0.73
		Average	0.80	0.77	0.77

245 In the table 2, we observe that intestinal metaplasia is confounded with chronic gastritis. For
 246 the three types of features, the results are close to 53% of precision and 67% recall. Better results are
 247 obtained in table 3, especially with divisions and subtractions' features. The average of the three
 248 classes is 80% for the precision and 80% for the recall.

**Figure 5.** Confusion matrix.

249 It can be seen in the confusion matrix figure 5 that the healthy class is sometimes confounded with
 250 intestinal metaplasia and that chronic gastritis is sometimes confounded with intestinal metaplasia.

251 This second point is more acceptable in the sense that intestinal metaplasia usually appears as little
 252 patches in the middle of chronic gastritis. Thus a patient with intestinal metaplasia has always chronic
 253 gastritis too.

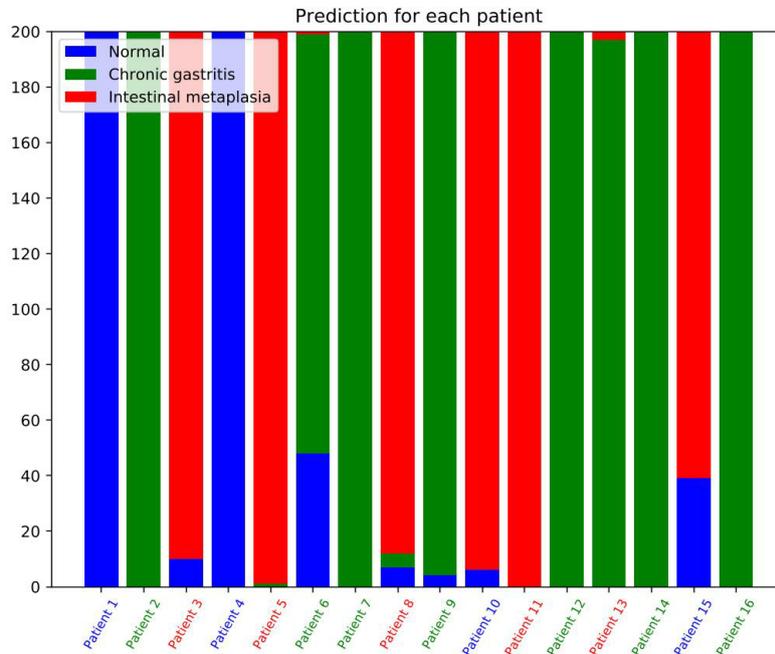
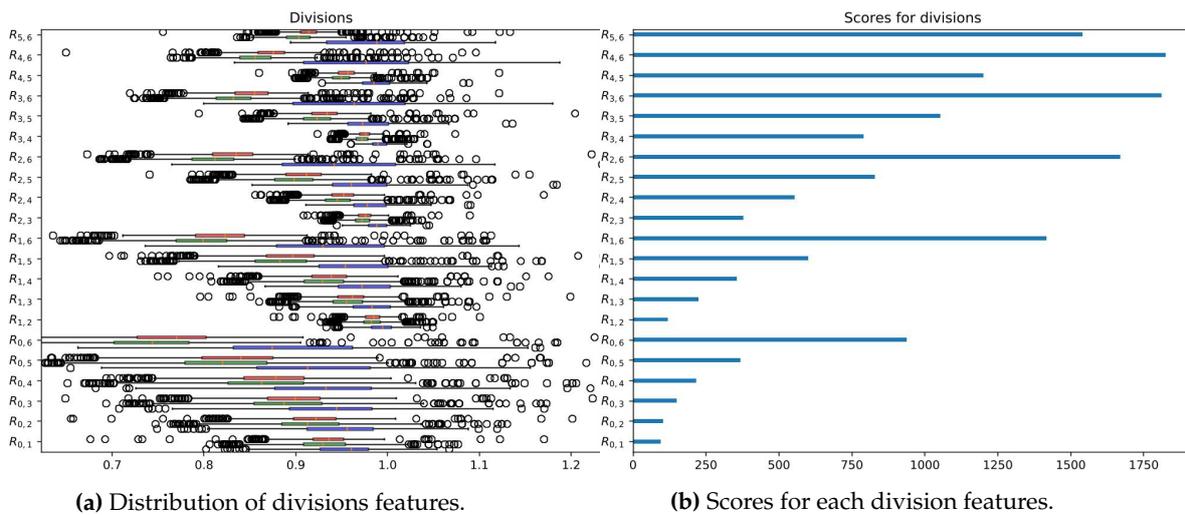


Figure 6. Results by patient.

254 It is interesting to detail the classification results for each patient. Figure 6 presents the number
 255 of spectra classified in each category for each patient. It can be seen that the classification is usually
 256 good except for patients 10 and 15, which are healthy and have mostly "intestinal metaplasia"-like
 257 spectra. Patient 13 has intestinal metaplasia, but its spectra are mostly classified as chronic gastritis.
 258 More generally, 13 patients out of 16 patients have most of their spectra classified in the right class.



(a) Distribution of divisions features.

(b) Scores for each division features.

Figure 7. Analysis of divisions features.

259 Moreover, the interesting point is that only one ratio is sufficient to reach good classification
 260 results. The best results are obtained with $k = 1$, the best ratio is the one with the highest score in

261 figure 7b *i.e.* $R_{4,6}$. $R_{4,6}$ is the division of the range [745, 755 nm] by the range [780, 840 nm]. It can
262 be observed in figure 7a that the distribution of the three classes (blue for healthy, green for chronic
263 gastritis, and red for intestinal metaplasia) for this particular ratio shows a good separability, especially
264 between healthy and the two pathologies. It is interesting to also note the correlation between the
265 preclinical study on mice. In both cases, the most discriminant wavelength range is located in the red
266 and near-infrared part of the spectra.

267 4. Conclusion

268 In this paper, a new prototype is introduced to acquire endoscopic NBI videos and multispectral
269 videos during *in vivo* stomach exploration. The prototype is used to acquire spectra on 16 patients. For
270 each of these patients, 200 spectra are selected, leading to a dataset of 3200 spectra. Spectra obtained
271 with multispectral cameras are processed with a classification pipeline containing a wavelength
272 reduction algorithm, the computation of three types of features (reduced bands, subtractions and
273 divisions) and a leave one patient out classification with an SVM classifier with a linear kernel. The
274 classifier can recognize a healthy stomach wall from chronic gastritis and intestinal metaplasia with
275 a good precision of 80%. Moreover, two exciting wavelength ranges are identified in the red and
276 near-infrared ranges: [745, 755 nm] and [780, 840 nm]. To the best of our knowledge, the analysis of
277 this ensemble of three types of tissues (*i.e.* chronic gastritis, intestinal metaplasia, and healthy) has
278 never been considered in a previous paper.

279 These promising results have to be confirmed on a larger dataset. The prototype can also be
280 improved to have better quality for the images and, thus, to be able to exploit textural information
281 caught by the multispectral cameras. Finally, unmixing algorithms can be used to get information on
282 the relative concentration of chromophores and thus having a better understanding of the biological
283 changes that appear with pre-cancerous lesions.

284 **Author Contributions:** Investigation, Alexandre Krebs; Methodology, Alexandre Krebs; Project administration,
285 Franck Marzani; Resources, Dominique Lamarque and Thomas Bazin; Supervision, Yannick Benezeth, Franck
286 Marzani and Dominique Lamarque; Writing – original draft, Alexandre Krebs; Writing – review & editing,
287 Alexandre Krebs, Yannick Benezeth and Franck Marzani.

288 **Funding:** This study was supported by the French Research National Agency (ANR) program EMMIE under the
289 grant agreement 15-CE17-0015

290 **Conflicts of Interest:** The authors declare no conflict of interest.

291 References

- 292 1. Ferlay, J.; Colombet, M.; Soerjomataram, I.; Mathers, C.; Parkin, D.; Piñeros, M.; Znaor, A.; Bray,
293 F. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods.
294 *International journal of cancer* **2019**, *144*, 1941–1953.
- 295 2. Sharma, P.; Bergman, J.J.; Goda, K.; Kato, M.; Messmann, H.; Alsop, B.R.; Gupta, N.; Vennalaganti,
296 P.; Hall, M.; Konda, V.; others. Development and validation of a classification system to identify
297 high-grade dysplasia and esophageal adenocarcinoma in Barrett's esophagus using narrow-band imaging.
298 *Gastroenterology* **2016**, *150*, 591–598.
- 299 3. Bansal, A.; Ulusarac, O.; Mathur, S.; Sharma, P. Correlation between narrow band imaging and
300 nonneoplastic gastric pathology: a pilot feasibility trial. *Gastrointestinal Endoscopy* **2008**, *67*, 210–216.
- 301 4. FICE Fuji Intelligent Chromo Endoscopy. http://www.sfed.org/files/documents_sfed/partenaires/fujinon/pdf/fice.pdf.
302 Accessed: 2019-05-27.
- 303 5. Burgos, H.; Porras, M.; Brenes, F.; Izquierdo, E. Fujinon FICE electronic chromovideoendoscopy helps
304 differentiate the type of metaplasia in patients with chronic atrophic gastritis. *Gastrointestinal Endoscopy*
305 **2007**, *65*, AB353.
- 306 6. Song, L.M.W.K.; Adler, D.G.; Conway, J.D.; Diehl, D.L.; Farraye, F.A.; Kantsevov, S.V.; Kwon, R.; Mamula,
307 P.; Rodriguez, B.; Shah, R.J.; others. Narrow band imaging and multiband imaging. *Gastrointestinal*
308 *endoscopy* **2008**, *67*, 581–589.

- 309 7. Weigt, J.; Malferttheiner, P.; Canbay, A.; Haybaeck, J.; Bird-Lieberman, E.; Link, A. Blue Light Imaging and
310 Linked Color Imaging for the Characterization of Mucosal Changes in Chronic Gastritis: A Clinicians View
311 and Brief Technical Report. *Digestive Diseases* **2019**, pp. 1–6.
- 312 8. Dohi, O.; Yagi, N.; Onozawa, Y.; Kimura-Tsuchiya, R.; Majima, A.; Kitaichi, T.; Horii, Y.; Suzuki, K.; Tomie,
313 A.; Okayama, T.; others. Linked color imaging improves endoscopic diagnosis of active *Helicobacter pylori*
314 infection. *Endoscopy international open* **2016**, *4*, E800–E805.
- 315 9. Gono, K. An Introduction to High-Resolution Endoscopy and Narrowband Imaging. *Comprehensive Atlas*
316 *of High-Resolution Endoscopy and Narrowband Imaging* **2017**, pp. 7–15.
- 317 10. Bruno, M. Magnification endoscopy, high resolution endoscopy, and chromoscopy; towards a better optical
318 diagnosis. *Gut* **2003**, *52*, iv7–iv11.
- 319 11. Uedo, N.; Ishihara, R.; Iishi, H.; Yamamoto, S.; Yamada, T.; Imanaka, K.; Takeuchi, Y.; Higashino, K.;
320 Ishiguro, S.; Tatsuta, M. A new method of diagnosing gastric intestinal metaplasia: narrow-band imaging
321 with magnifying endoscopy. *Endoscopy* **2006**, *38*, 819–824.
- 322 12. Gu, X.; Yao, L.; Zhong, Y.; Han, Z.; Shi, Q.; Fu, Y.; Liu, C.; Wang, X.; Xie, T. Image enhancement based on in
323 vivo hyperspectral gastroscopic images: a case study. *Journal of biomedical optics* **2016**, *21*, 101412.
- 324 13. Kiyotoki, S.; Nishikawa, J.; Okamoto, T.; Hamabe, K.; Saito, M.; Goto, A.; Fujita, Y.; Hamamoto, Y.; Takeuchi,
325 Y.; Satori, S.; others. New method for detection of gastric cancer by hyperspectral imaging: a pilot study.
326 *Journal of biomedical optics* **2013**, *18*, 026010.
- 327 14. Hohmann, M.; Kanawade, R.; Klämpfl, F.; Douplik, A.; Mudter, J.; Neurath, M.; Albrecht, H. In-vivo
328 multispectral video endoscopy towards in-vivo hyperspectral video endoscopy. *Journal of biophotonics*
329 **2017**, *10*, 553–564.
- 330 15. Makhlof, H.; Gmitro, A.F.; Tanbakuchi, A.A.; Udovich, J.A.; Rouse, A.R. Multispectral confocal
331 microendoscope for in vivo and in situ imaging. *Journal of biomedical optics* **2008**, *13*, 044016.
- 332 16. Grosberg, L.E.; Radosevich, A.J.; Asfaha, S.; Wang, T.C.; Hillman, E.M. Spectral characterization and
333 unmixing of intrinsic contrast in intact normal and diseased gastric tissues using hyperspectral two-photon
334 microscopy. *PLoS one* **2011**, *6*, e19925.
- 335 17. Bergholt, M.S.; Zheng, W.; Ho, K.Y.; Teh, M.; Yeoh, K.G.; So, J.B.Y.; Shabbir, A.; Huang, Z. Fiber-optic Raman
336 spectroscopy probes gastric carcinogenesis in vivo at endoscopy. *Journal of biophotonics* **2013**, *6*, 49–59.
- 337 18. Martinez-Herrera, S.E.; Benezeth, Y.; Boffety, M.; Emile, J.F.; Marzani, F.; Lamarque, D.; Goudail, F.
338 Multispectral endoscopy to identify precancerous lesions in gastric mucosa. *International Conference on*
339 *Image and Signal Processing*, Springer, 2014, pp. 43–51.
- 340 19. Reddymasu, S.C.; Sharma, P. Advances in endoscopic imaging of the esophagus. *Gastroenterology Clinics*
341 **2008**, *37*, 763–774.
- 342 20. Namikawa, T.; Sato, T.; Hanazaki, K. Recent advances in near-infrared fluorescence-guided imaging
343 surgery using indocyanine green. *Surgery today* **2015**, *45*, 1467–1474.
- 344 21. Li, Y.; Xie, X.; Yang, X.; Guo, L.; Liu, Z.; Zhao, X.; Luo, Y.; Jia, W.; Huang, F.; Zhu, S.; others.
345 Diagnosis of early gastric cancer based on fluorescence hyperspectral imaging technology combined
346 with partial-least-square discriminant analysis and support vector machine. *Journal of biophotonics* **2019**,
347 *12*, e201800324.
- 348 22. Kobayashi, M.; Tajiri, H.; Seike, E.; Shitaya, M.; Tounou, S.; Mine, M.; Oba, K. Detection of early gastric
349 cancer by a real-time autofluorescence imaging system. *Cancer letters* **2001**, *165*, 155–159.
- 350 23. Krebs, A.; Camilo, V.; Touati, E.; Benezeth, Y.; Michel, V.; Jouvion, G.; Yang, F.; Lamarque, D.; Marzani, F.
351 Detection of *H. pylori* Induced Gastric Inflammation by Diffuse Reflectance Analysis. *18th International*
352 *Conference on Bioinformatics and Bioengineering (BIBE)*. IEEE, 2018, pp. 287–292.
- 353 24. Neuhauser, M. *Nonparametric statistical tests: A computational approach*; Chapman and Hall/CRC, 2011.
- 354 25. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer,
355 P.; Weiss, R.; Dubourg, V.; others. Scikit-learn: Machine learning in Python. *Journal of Machine Learning*
356 *Research* **2011**, *12*, 2825–2830.
- 357 26. USB3 Vision compliant cameras - xiQ. <https://www.ximea.com/en/products/usb3-vision-standard-designed-cameras-xiq>.
358 Accessed: 2019-05-27.
- 359 27. ITConcepts microFlex. <http://www.itconceptsworld.com/index.html>. Accessed: 2019-05-27.
- 360 28. TwinCam. <https://www.cairn-research.co.uk/product/twincam/>. Accessed: 2019-05-25.

- 361 29. Zenteno, O.; Krebs, A.; Treuillet, S.; Lucas, Y.; Benezeth, Y.; Marzani, F. Spatial and Spectral Calibration of
362 a Multispectral-Augmented Endoscopic Prototype. *International Joint Conference on Computer Vision,
363 Imaging and Computer Graphics*. Springer, 2018, pp. 262–280.
- 364 30. Ximea API. <https://www.ximea.com/support/wiki/apis/xiAPINET>.
- 365 31. Cetin, A.E.; Tofighi, M. Projection-based wavelet denoising [lecture notes]. *Signal Processing Magazine* **2015**,
366 32, 120–124.

367 **Sample Availability:** Samples of the compounds are available from the authors.

368 © 2020 by the authors. Submitted to *Appl. Sci.* for possible open access publication under the terms and conditions
369 of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).