

Hyperspectral imaging for tissue classification, a way toward smart laparoscopic colorectal surgery

Elisabeth J. M. Baltussen
Esther N. D. Kok
Susan G. Brouwer de Koning
Joyce Sanders
Arend G. J. Aalbers
Niels F. M. Kok
Geerard L. Beets
Claudie C. Flohil
Sjoerd C. Bruin
Koert F. D. Kuhlmann
Henricus J. C. M. Sterenborg
Theo J. M. Ruers

Elisabeth J. M. Baltussen, Esther N. D. Kok, Susan G. Brouwer de Koning, Joyce Sanders, Arend G. J. Aalbers, Niels F. M. Kok, Geerard L. Beets, Claudie C. Flohil, Sjoerd C. Bruin, Koert F. D. Kuhlmann, Henricus J. C. M. Sterenborg, Theo J. M. Ruers, "Hyperspectral imaging for tissue classification, a way toward smart laparoscopic colorectal surgery," *J. Biomed. Opt.* **24**(1), 016002 (2019), doi: 10.1117/1.JBO.24.1.016002.

Hyperspectral imaging for tissue classification, a way toward smart laparoscopic colorectal surgery

Elisabeth J. M. Baltussen,^{a,*†} Esther N. D. Kok,^{a,†} Susan G. Brouwer de Koning,^a Joyce Sanders,^b Arend G. J. Aalbers,^a Niels F. M. Kok,^a Geerard L. Beets,^a Claudie C. Flohil,^c Sjoerd C. Bruin,^d Koert F. D. Kuhlmann,^a Henricus J. C. M. Sterenberg,^{a,e} and Theo J. M. Ruers^{a,†}

^aAntoni van Leeuwenhoek Hospital, The Netherlands Cancer Institute, Department of Surgery, Amsterdam, The Netherlands

^bAntoni van Leeuwenhoek Hospital, The Netherlands Cancer Institute, Department of Pathology, Amsterdam, The Netherlands

^cSlotervaart Medical Centre, Department of Pathology, Amsterdam, The Netherlands

^dSlotervaart Medical Centre, Department of Surgery, Amsterdam, The Netherlands

^eAmsterdam University Medical Centre, University of Amsterdam, Department of Biomedical Engineering and Physics, Amsterdam, The Netherlands

[†]Technical University Twente, MIRA Institute, Enschede, The Netherlands

Abstract. In the last decades, laparoscopic surgery has become the gold standard in patients with colorectal cancer. To overcome the drawback of reduced tactile feedback, real-time tissue classification could be of great benefit. In this *ex vivo* study, hyperspectral imaging (HSI) was used to distinguish tumor tissue from healthy surrounding tissue. A sample of fat, healthy colorectal wall, and tumor tissue was collected per patient and imaged using two hyperspectral cameras, covering the wavelength range from 400 to 1700 nm. The data were randomly divided into a training (75%) and test (25%) set. After feature reduction, a quadratic classifier and support vector machine were used to distinguish the three tissue types. Tissue samples of 32 patients were imaged using both hyperspectral cameras. The accuracy to distinguish the three tissue types using both hyperspectral cameras was 0.88 (STD = 0.13) on the test dataset. When the accuracy was determined per patient, a mean accuracy of 0.93 (STD = 0.12) was obtained on the test dataset. This study shows the potential of using HSI in colorectal cancer surgery for fast tissue classification, which could improve clinical outcome. Future research should be focused on imaging entire colon/rectum specimen and the translation of the technique to an intraoperative setting. © The Authors. Published by SPIE under a Creative Commons Attribution 4.0 Unported 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.24.1.016002]

Keywords: hyperspectral imaging; colorectal cancer; margin assessment; machine learning; support vector machine.

Paper 180550RRRR received Sep. 17, 2018; accepted for publication Jan. 11, 2019; published online Jan. 30, 2019.

1 Background

Colorectal cancer is the third most commonly diagnosed cancer worldwide and the fourth cause of death due to cancer.^{1,2} For patients with colorectal cancer, surgery is the cornerstone of the treatment. In the last decades, laparoscopic surgery for colorectal cancer has become common practice. Randomized controlled trials have proven similar clinical outcomes for laparoscopic surgery as for open surgery with a decrease in hospital stay.³ One of the drawbacks of laparoscopic surgery is the reduced tactile feedback during surgery.^{4,5} The lack of tactile feedback makes tissue recognition more cumbersome, especially in areas where radical resection margins are often compromised, like in locally advanced tumors and in rectal cancer. Hence, an alternative technique that would enable the surgeon to distinguish tumor from normal tissue during laparoscopic surgery in real-time would be of great benefit to secure radical resection in difficult areas, such as in rectal cancer. We will investigate the use of hyperspectral imaging (HSI) as a tool to ensure radical margins in these circumstances, distinguishing tumor from healthy colorectal tissue.

In HSI, a broadband light source is used to illuminate an object, like, e.g., tissue. The light will interact with the tissue

by reflection, scattering, and absorption of the photons. This interaction strongly depends on the tissue type and wavelength.⁶ After several interactions within the tissue, part of the light will be reflected to the surface of the tissue and is detected by the hyperspectral camera. In the resulting hyperspectral image, the tissue specific spectral changes of the light can be analyzed. Ultimately, HSI result in 2-D images of the object obtained over several wavelengths, resulting in a 3-D datacube with two spatial dimensions and in the third dimension the wavelengths (Fig. 1).

Previous studies used HSI as a diagnostic tool in cancer of the cervix,⁷ breast,⁸ skin,⁹ tongue,¹⁰ head and neck,^{11,12} gastric,^{13,14} and colon and rectum.^{15–22} In colorectal cancer, most studies focused on HSI of the hematoxylin–eosin (H&E) pathology slides^{15–19} or tissue classification during endoscopy to distinguish tumor from healthy tissue.^{20–22} In the current *ex vivo* study, we investigated the use of HSI to differentiate normal colorectal tissue from tumor tissue in a surgical setting looking from the surface of the tissue instead of from the lumen of the colon. To this end, colorectal cancer samples obtained during surgery were imaged using HSI in the visible- to near-infrared region. The spectra obtained from these images were classified using a classification algorithm and were verified with histology. Finally, complete hyperspectral images were classified using the trained classifier. The ultimate goal is to develop a real-time technique for tissue identification in laparoscopic colorectal surgery.

*Address all correspondence to Elisabeth J. M. Baltussen, E-mail: l.baltussen@nki.nl

[†]These authors contributed equally to this work.

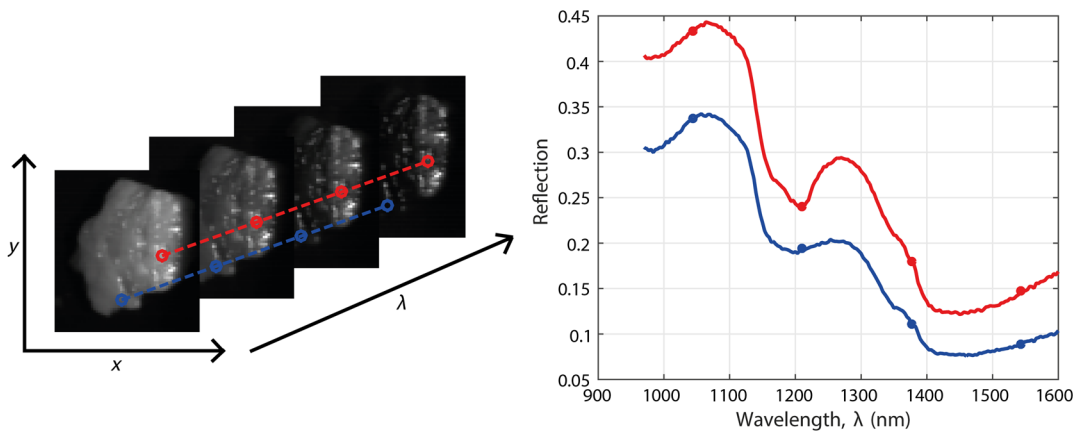


Fig. 1 (a) Hyperspectral image, with two spatial dimensions (x, y) and one spectral dimension (λ). (b) On the right side, the spectra of the two selected pixels are shown.

2 Materials and Methods

2.1 Hyperspectral Cameras

Two hyperspectral cameras were used for the measurements, one in the visual wavelength range and one in the near-infrared wavelength range. The first was a SPECIM (Spectral Imaging Ltd., Finland) spectral camera (PFD-CL-65-V10E) with a wavelength range from 400 to 1000 nm, a CMOS sensor of 1×1312 pixels, and a spectral resolution of 3.0 nm, hereafter referred to as the visual camera. The second was also a SPECIM spectral camera (VLNIR CL-350-N17E) with a wavelength range from 900 to 1700 nm, an InGaAs sensor of 1×320 pixels, and a spectral resolution of 5.0 nm, hereafter referred to as the near-infrared camera. Both cameras were push broom cameras, meaning that they image a single line (x -axis) only. To obtain a full 2-D image, the samples were placed on a translational stage and pushed underneath the camera (y -axis). All samples were illuminated using a halogen light source.

Both a dark and white reference image were taken before each measurement. The dark reference image was taken by closing the shutter of the cameras. The white reference image was taken on a Spectralon reflectance standard. The linear behavior of the visual camera allowed for a simple calibration of the camera using Eq. (1). In Eq. (1), x_{cal} is the calibrated spectrum, x is the original spectrum, D_{ref} is the dark reference, and W_{ref} is the white reference. The near-infrared camera, however, had a slight nonlinear behavior and was therefore calibrated using a fourth order polynomial, Eq. (2), instead of the linear formula.²³ In Eq. (2), b_i are variables determined using a series of five reference samples. The values of b_i differ per pixel and wavelength. Furthermore, x is the original spectrum and x_{cal} is the calibrated spectrum:²⁴

$$x_{\text{cal}} = \frac{x - D_{\text{ref}}}{W_{\text{ref}} - D_{\text{ref}}}, \quad (1)$$

$$x_{\text{cal}} = b_0 + b_1x + b_2x^2 + b_3x^3. \quad (2)$$

2.2 Study Protocol

Patients who underwent surgery for colorectal cancer in the Antoni van Leeuwenhoek—The Netherlands Cancer Institute

(Amsterdam, the Netherlands) and the Slotervaart Medical Centre (Amsterdam, The Netherlands) were included in this *ex vivo* study. The study was performed under approval of the protocol by the institutional ethics review board.

Immediately after colorectal resection, the entire resected specimen was taken to the pathology department. Cross-sections were cut by the pathologist from the specimen and three tissue samples were obtained; tumor tissue, healthy colon or rectal wall, and (pericorectal) fat. The cross-sections were placed in a pathology cassette, where they remained during the entire data acquisition. All measurements were performed within 1 h after specimen resection. Before the hyperspectral measurements, an RGB image was taken of each tissue sample. Next, hyperspectral images were obtained from the tissue samples after which the samples were taken back to the pathology department. The samples were processed according to standard protocol in the same pathology cassette, to prevent large tissue deformations. The corresponding H&E slides were examined by the pathologist, who annotated the various tissue types. For further data analysis, the digitized annotated slides were registered to the RGB image and the RGB image to the HSI in MATLAB (version 8.5, MathWorks Inc., Natick, Massachusetts, United States), using a nonrigid transformation to overcome the effects of mechanical deformation of the tissue during the standard workflow of tissue processing and staining. Finally, the registered pathology slide was registered to the registered RGB image (Fig. 2). Using these registrations, each pixel from the hyperspectral image could be given a histological classification. To create a database of hyperspectral pixels, pixels were manually selected within areas that were defined by the pathologist as absolute certain for a tissue type. About 30 pixels per tissue type per patient were selected when possible. When the surface area of a tissue type was too small to select 30 individual pixels, less pixels were selected. Pixels in the mucosal layer were not taken into account because the mucosa will not be visible during the ultimate surgical application of the technology. Hence, the pixels from the healthy colorectal wall were all in the muscular layer.

2.3 Data Preprocessing

Preprocessing of the data was performed using a 3.40 GHz Intel Xeon E3-1240 CPU processor and 16 GB RAM and consisted of two steps. First, the spectra were normalized using standard normal variate (SNV) normalization.²⁵ SNV normalization was

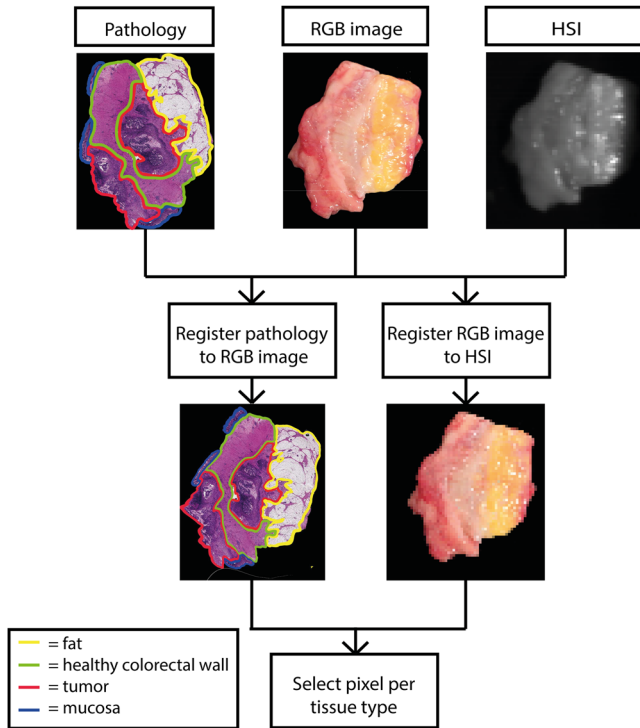


Fig. 2 Registration of the HSI, RGB, and pathology images. In the upper row from left to right, annotated pathology image (yellow = fat, green = healthy colorectal wall, red = tumor, blue = mucosa), RGB image and HSI. The second row from left to right, the annotated pathology image registered to the RGB image and the RGB image registered to the HSI.

performed for each individual spectrum. First, the mean was subtracted from the spectrum after which the spectrum was divided by the standard deviation of the spectrum, see Eq. (3). Here, x_{corr} is the normalized spectrum, x_{cal} is the calibrated spectrum as given in Eq. (1) or Eq. (2), and x_{mean} and x_{std} are the mean and standard deviation of x_{cal} , respectively. This normalization created a zero baseline and a variance equal to one for all spectra:

$$x_{\text{corr}} = \frac{x_{\text{cal}} - x_{\text{mean}}}{x_{\text{std}}}. \quad (3)$$

After combination of the visual and near-infrared images, all outliers caused by specular reflection were removed. Outliers were defined as spectra with an average distance from the mean spectrum of more than three standard deviations, determined over all wavelengths.

Next, in order to combine the spectra of the two cameras, the visual images were downsampled. Down sampling was necessary because of the higher resolution of the visual camera compared to the near-infrared camera. A rigid spatial registration was performed obtaining a pixel-to-pixel correlation between the images of the two cameras.

2.4 Data Analysis

Data analysis was performed using the PerClass toolbox (Academic version 5.0, PR Sys design, Delft, The Netherlands) in MATLAB. The data were randomly divided into a training and test set. Per patient, all spectra were assigned to either the training or test set, indicating that spectra from one

patient were not split between the training and test set. The training set contained 75% of the patients and the remaining 25% was used as a test set. The data contained a hyperspectral image from both the visual and near-infrared camera.

The development of the classification algorithm consisted of three steps. First, feature reduction was applied to the spectra to prevent overfitting of the classifier. For this purpose, k-means clustering was used to determine spectral bands. The clustering is based on the average intensity values of the spectra from the training dataset. For each cluster, if the wavelengths are not continued, the cluster will be divided into multiple spectral bands. The spectral bands were determined once on the training set and were also used on the test set. From these spectral bands, only the mean intensity was used as a feature in the classification algorithm. Second, fat was classified first using a quadratic classifier, which was optimized with a 10-fold cross-validation. The selected optimum in the ROC curve was the point with the lowest mean error. Third, a linear support vector machine (SVM) was used to distinguish the tumor spectra from the healthy colorectal wall spectra. The SVM was also optimized using a 10-fold cross-validation. Classification of the pixels was done based on the probability given by the classifiers. Pixels were assigned to the tissue type with the highest probability. The performance of the classifiers was compared using the area under the ROC curve (AUC), the accuracy, the sensitivity, specificity, and the Matthews correlation coefficient (MCC) [Eq. (4)]:

$$\text{MCC} = \frac{\text{TP} \times \text{TN} - \text{FP} \times \text{FN}}{\sqrt{(\text{TP} + \text{FP})(\text{TP} + \text{FN})(\text{TN} + \text{FP})(\text{TN} + \text{FN})}}. \quad (4)$$

In Eq. (4), TP, TN, FP, and FN are the number of true positives, true negatives, false positives, and false negatives, respectively. The MCC returns a value from -1 to $+1$, where -1 indicates a total disagreement and $+1$ indicates a perfect prediction.

The accuracy was determined in two different ways. The first accuracy was determined per tissue type and thereafter averaged. The second accuracy was determined per patient and averaged.

Finally, to assess the contribution of each camera, the classification was also trained and tested on datasets containing data from only one of the two cameras. The performance of this classification was compared with the classification of the dataset containing data from both cameras, using the ROC curves and the performance measures.

3 Results

3.1 Patients

In total, 54 patients were included in this study: 27 men (50%), 27 women (50%), with a median age of 65.5 years (IQR: 60 to 73). The samples of 32 patients were imaged with both the near-infrared and the visual camera, 22 additional patients were imaged with only the near-infrared camera. Most of the tumors were located in the colon and sigmoid. One sample showed complete pathological response on preoperative treatment and thus, no tumor tissue could be taken from the sample. Patient and tumor characteristics are described in Table 1.

Table 1 Characteristics of the group of patients measured with the near-infrared camera and the group of patients measured with the visual camera.

		NIR camera	VIS camera
Total number of patients		54	32 ^a
Gender	Male	27	14
	Female	27	18
Age	Median	65.5	66.5
	Interquartile range	60–73	60–71.5
Tumor location	Cecum	7	7
	Colon	23	12
	Sigmoid	21	11
	Rectum	3	2
Tumor type	Complete response	1	1
	Adenocarcinoma	46	25
	Mucinous adenocarcinoma	7	6
	Tumor stage		
	pT0	1	1
	pT1	3	1
	pT2	9	4
	pT3	26	16
	pT4	15	10

^aAll patients measured with the VIS camera are also included in the NIR camera measurements.

3.2 Data Acquisition and Processing Time

The obtained tissue samples were first placed under the near-infrared camera and subsequent under the visual camera. Data acquisition times were 20 and 30 s for the near-infrared and the visual camera, respectively. In total, data of one patient were acquired in 1 min. The duration of image preprocessing for each tissue sample was 60 s in total for both cameras combined for all wavelengths. After training of the classifier, classification of the test data took 2 s per patient.

3.3 Classification with the Combination of Visual and Near-Infrared Camera

For the classification of the combination of the visual and near-infrared camera, only the tissue samples scanned with both cameras were included. The dataset of the combined visual and near-infrared camera images contained 2194 spectra, from 32 patients. After outlier removal, 2170 spectra were present in the combined dataset, of which 857 were taken from fat, 563 from muscle, and 750 from tumor. The training set consisted of 24 patients with a total of 1726 spectra. The test set consisted

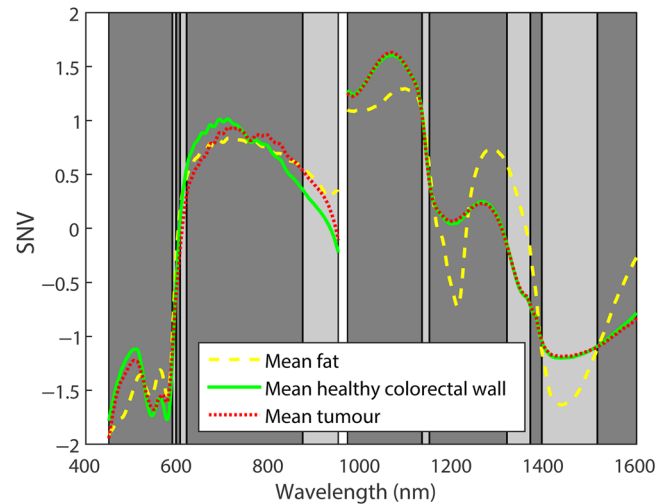


Fig. 3 Spectral bands determined for the dataset with the combination of visual and near-infrared camera images are shown together with the mean spectra of fat (yellow), healthy colorectal wall (green), and tumor (red). The 13 spectral bands all have a different gray value and are separated by black vertical lines. Between 950 and 970 nm, a gap is shown in the data. This region is not covered by the cameras.

of the remaining 8 patients and 444 spectra. Due to the presence of noise in the lower and upper wavelength range of both cameras, the wavelength ranges of 450 to 950 and 970 to 1600 were selected to analyze for the visual and near-infrared camera, respectively.

As shown in Fig. 3, the reflection spectra obtained with the visual camera and the near-infrared camera are not connected after calibration. This is related to differences in optical geometry of the two camera setups. In Fig. 3, the large difference in normalized intensity visible at 960 nm is caused by individual SNV normalization of both cameras before combining the datasets.

Based on the training set, 13 spectral bands were determined, as shown in Fig. 3. On the training data, the quadratic classifier obtained an MCC, AUC, accuracy, sensitivity, and specificity of 1.00 to separate fat from healthy and tumor. The SVM applied on the training dataset, to distinguish tumor from muscle, provided an MCC of 0.83, an AUC of 0.98, an accuracy of 0.91, and the sensitivity and specificity were 0.93 and 0.90, respectively. The accuracy of combination of the quadratic classifier and the SVM on the training dataset was 0.94 (STD = 0.04) when assessed on the tissue types. When determined per patient and averaged, a training accuracy of 0.94 (STD = 0.13) was obtained.

The results of the combination of the quadratic classifier and the SVM on the test dataset are shown in Table 2. The accuracy determined per tissue type and averaged on the test dataset was 0.88 (STD = 0.13). The accuracy calculated per patient and thereafter averaged was 0.93 (STD = 0.12).

In Fig. 4, the results of the classification of all pixels in a hyperspectral image of one patient from the test set are shown. The different colors represent different tissue types. The certainty of the classification, based on the probability, is shown by the intensity of the color. The more intense the color is, the higher the certainty of the classifier is for this classification.

Table 2 Results of the combined classifiers on the test dataset with a combination of visual and near-infrared camera images. The mean accuracy averaged over the tissue types was 0.88 (STD = 0.13) and over the patients was 0.93 (STD = 0.12).

		Decision based on hyperspectral data			Total
		Fat	Muscle	Tumor	
Gold standard	Fat	187	0	0	187
	Muscle	30	81	4	115
	Tumor	0	9	133	142
	Total	217	90	137	444

3.4 Classification with a Single Camera

The classification of fat, healthy colon or rectal wall, and tumor was also performed on the datasets including only one of the two cameras. The dataset with only spectra from the near-infrared camera contained 54 patients and 4352 spectra of which 1690 were measured in fat, 1251 in the muscular layer of healthy colon or rectal wall, and 1411 in tumor. After removal of the outliers 4309 spectra remained (1676 fat, 1232 muscle, and 1401 tumor). For the training set, 41 patients were randomly selected with a total of 3241 spectra. The test set contained 13 patients and 1068 spectra.

For the images of the visual camera, the same pixels were used as selected for the images of the near-infrared camera. A total of 2194 spectra, from 32 patients, were included in this dataset. From the spectra, 866 were measured in fat, 569

in muscle, and 759 in tumor. After removal of the outliers, 2164 spectra remained (854 fat, 560 muscle, and 750 tumor). The training set included 24 patients with 1723 spectra, and the remaining 8 patients were included in the test set, which contained 441 spectra.

After spectral bands were extracted from the datasets, the quadratic classifiers were trained and ROC curves for both datasets were obtained. In Fig. 5, the ROC curves of both classifiers are shown together with the ROC curve of the classifier created with the combined dataset. This shows a slightly worse performance for the dataset with only visual camera images compared to the dataset with only near-infrared images or the dataset including image of both cameras. This was also seen in the performance measures, with an MCC of 0.90 for the dataset with only visual camera images, and an MCC of 0.99 and 1.00 for the dataset with only near-infrared camera images and the combined dataset, respectively. All other performance measures showed the same trend.

In Fig. 6, the ROC curves of the SVMs are shown for the three training datasets. Here, again, the dataset containing only visual camera images showed the worst performance. However, there is a clear difference between the dataset containing only near-infrared camera images and the dataset containing images of both cameras, where the latter outperformed the first. A summary of the performance measures of the SVM is shown in Table 3. The same trend is shown for all performance of the three datasets; the dataset including only visual camera images performed worst and the combined dataset performed best.

The results of the test set of the two datasets containing data of only one of the two cameras resulted in an accuracy for determining the tissue types of 0.67 (STD = 0.19) and 0.83 (STD = 0.12) for the visual camera data and near-infrared camera data, respectively. The accuracy calculated per patient and

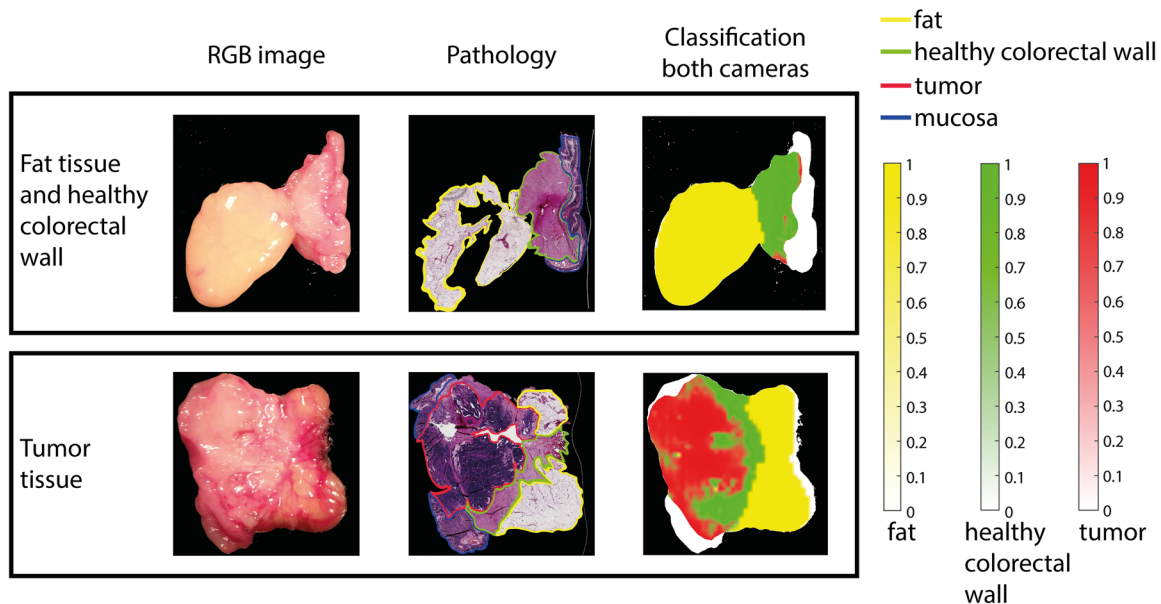


Fig. 4 Classification of the tissue samples of one patient from the test set of the combined data set of the visual and near-infrared camera. In the first column, the RGB image of each tissue sample is shown. The second column shows the registered annotated pathology image (yellow = fat, green = muscle or healthy colorectal wall, red = tumor, blue = mucosa). In the third column, the classification based on the visual and near-infrared spectra is shown projected on the binary mask of the RGB image (yellow = fat, green = muscle of healthy colorectal wall, red = tumor). The first row shows the healthy tissue including fat and healthy colorectal wall. The second row shows the tumor tissue sample. Tissue annotated by the pathologist as mucosa (blue) is not classified and is shown a white in the third column.

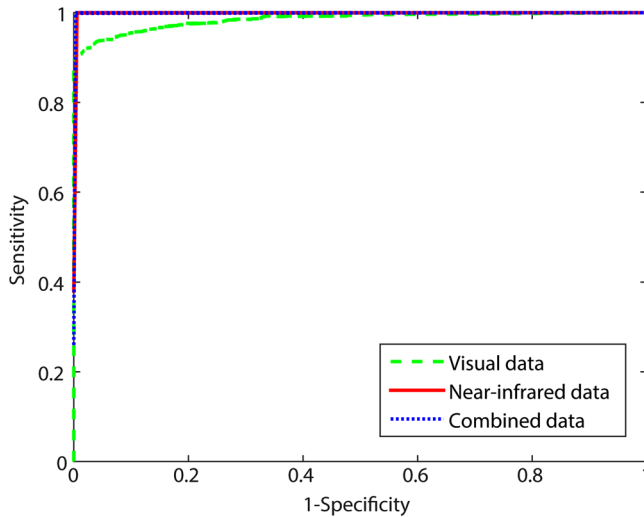


Fig. 5 ROC curves of the training results of the quadratic classifier distinguishing fat from all other tissue types. The three datasets are shown as the visual camera (green), near-infrared camera (red), and the combination of the visual and near-infrared camera (blue).

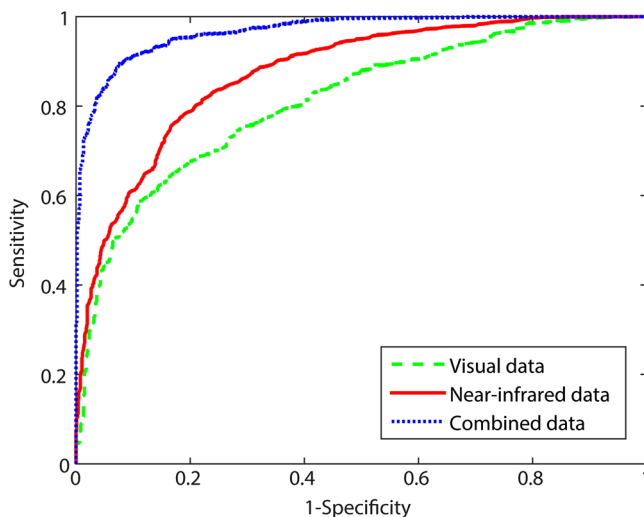


Fig. 6 ROC curves of the training results of the SVMs distinguishing tumor from healthy colorectal tissue. The three datasets are shown as the visual camera (green), near-infrared camera (red), and the combination of the visual and near-infrared camera (blue).

Table 3 Performance measures for the SVM from the three training datasets.

Performance measure	Visual camera	Near-infrared camera	Combined dataset
MCC	0.50	0.59	0.83
AUC	0.81	0.87	0.98
Accuracy	0.74	0.80	0.91
Sensitivity	0.77	0.78	0.93
Specificity	0.74	0.81	0.90

averaged was 0.71 (STD = 0.19) for the visual camera data and 0.83 (STD = 0.14) for the near-infrared camera data.

The classifiers created were also used to classify the spectra from each pixel of entire hyperspectral images of one of the patients from the test set that was imaged by both cameras. In Fig. 7, the result of this classification is shown for all three classifications.

4 Discussion

In this study, the potential added value of HSI for fast tissue classification during colorectal cancer surgery was examined. As a first step, an *ex vivo* study was designed in which tissue samples from colorectal cancer surgery were imaged with two hyperspectral cameras. One camera obtained images in the visual wavelength range (400 to 1000 nm), and the second camera obtained images in the near-infrared wavelength range (900 to 1700 nm). HSI allowed accurate discrimination of fat, healthy colon or rectal wall, and tumor tissue, with an accuracy of 0.88 (STD = 0.13) for the combination of the visual and near-infrared camera images.

Current literature of HSI in colorectal cancer was mainly focused on the classification of H&E pathology slides^{15–19} or classification of tissue during endoscopy.^{20–22} For the first application, hyperspectral images were made of H&E pathology slides to obtain objective classification into healthy or malignant tissue of colon biopsies.^{15–19} This application is far from the goal of the current study as the current study is focused on near real-time imaging during surgery. For the second application, a hyperspectral camera was combined with an endoscope, to obtain hyperspectral images during endoscopy. The study by Kumashiro et al.²² obtained *in vivo* hyperspectral data during colonoscopy and was able to distinguish tumor from healthy mucosa with a sensitivity of 0.73 and a specificity of 0.82, using Pearson correlation analysis. The study by Han et al.²¹ obtained better results with an accuracy of 0.94 and a sensitivity and specificity of 0.97 and 0.91, respectively. This study used hyperspectral images in the spectral range from 405 to 665 nm, so only the visual wavelength range. These results are better than the results shown in the current study for the dataset including only the visual camera images. The main difference between the two studies and the current study is the location of the measurements. Han et al.²¹ and Kumashiro et al.²² performed measurements of the lumen of the colon during endoscopy, whereas the current study focused on the surgical application and performed measurements from the surface of the colon.

In a previous study, we showed the possibility to distinguish the three tissue types—fat, healthy colon or rectal wall, and tumor—using fiberoptic diffuse reflectance spectroscopy (DRS) in a surgical setting.²⁶ The information obtained in fiberoptic DRS is very similar to the information obtained with a hyperspectral camera system. Therefore, it is not surprising that the results obtained in the current study are comparable to the results obtained in the previous study using DRS. However, the accuracy of the current study is slightly less compared to the accuracy obtained with the DRS measurements (accuracy = 0.95, STD = 0.03). An explanation for this difference might be due to the correlation with the gold standard, histology, which is less challenging for the fiberoptic point measurements performed in the previous study. In the testing of the combination of the two cameras, 30 muscle spectra of one particular patient were classified as fat. These misclassifications are most likely due to a fault in the registration between

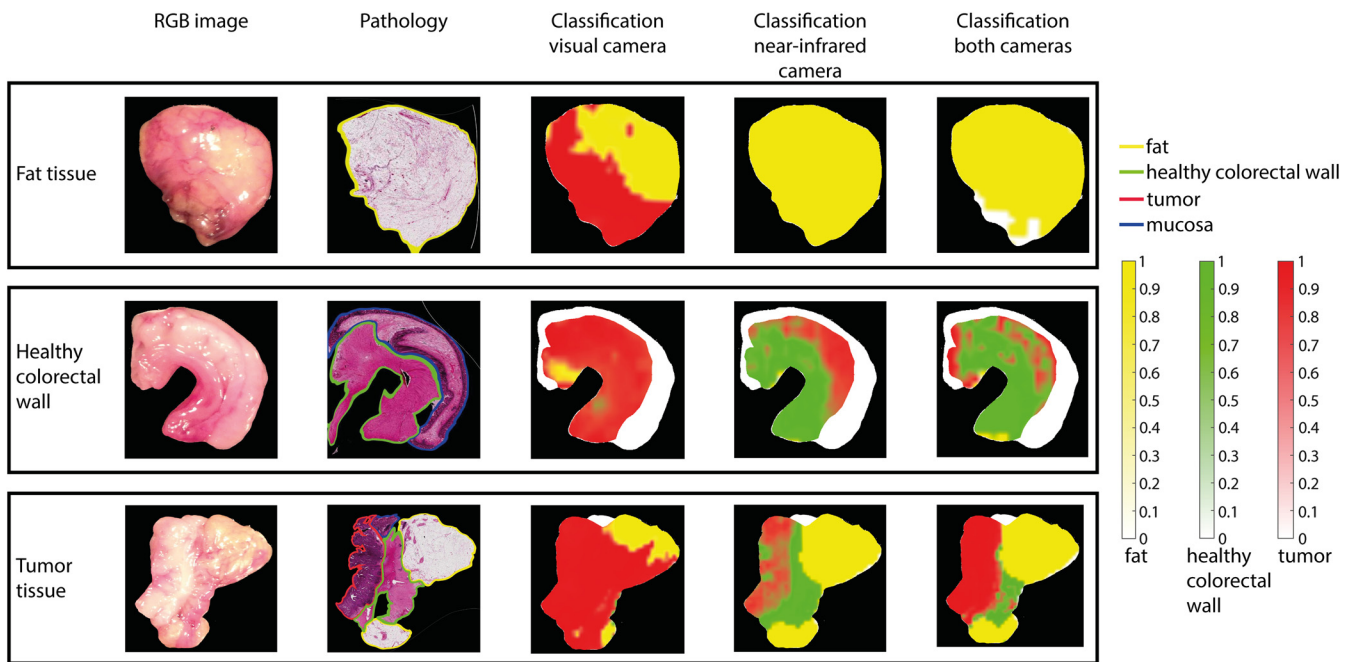


Fig. 7 Classification of the tissue samples of one patient from the test dataset of the combined dataset. In the first column, the RGB image of each tissue sample is shown. The second column shows the registered annotated pathology H&E image (yellow = fat, green = muscle or healthy colorectal wall, red = tumor, blue = mucosa). In the third to fifth column from the classification based on the visual image only, the classification based on the near-infrared image only and the classification based on the combined visual and near-infrared image are shown, respectively, projected on the binary mask of the RGB image (yellow = fat, green = muscle of healthy colorectal wall, red = tumor). From top to bottom, fat, healthy colorectal wall, and tumor tissue are shown. Tissue annotated as mucosa (blue) by the pathologist is not classified and shown as white in the third to fifth column.

the hyperspectral images and the histology in this specific patient. The influence of this one patient can be seen in the accuracy determined per patient, which is 0.93 and similar to the accuracy obtained in the DRS study.

For the evaluation of the performance of the classifiers, six different performance measures were used. Of these parameters, the MCC and AUC are most accurate in the current study, because these parameters are relatively insensitive to the effect of an imbalanced dataset.²⁷ However, for the combination of the quadratic classifier and the SVM, the MCC and AUC cannot be used, because both measures only account for a two class problem. For the quadratic classifier, no large differences are shown between the performance measures for the three datasets. However, the performance measures of the SVM did show a difference. The combination of the two cameras clearly outperforms the datasets using only one of the two cameras (Table 3). For the combination of the quadratic classifier and the SVM, the accuracies can be compared. Here, the combination of the two cameras outperforms the datasets with data from only one of the two cameras.

Comparing the two cameras, the near-infrared camera slightly outperforms the visual camera, with a MCC of 0.59 and 0.50, respectively. In Fig. 3, only the average curves are shown per tissue type. Although the difference between the average spectra of colon and tumor tissue is bigger in the visual part of the spectrum, also the standard deviation (not shown) is higher in the visual part compared to the near-infrared part. Furthermore, the SVM used for the classification of healthy colon and tumor does not take into account each feature individually but uses the combination of features to find the optimal

hyperplane, which differentiates the two classes. In accordance, the combination of the features from the near-infrared part of the spectrum might give a better result than the combination of the features from the visual part of the spectrum. It is hard to visualize these distinctive differences in the near-infrared part of the spectrum. In line with the explanation above, as mentioned before, the MCC value for the discrimination between healthy and tumor was slightly higher in the near-infrared part of the spectrum, the combination of the near-infrared and visual part of the spectrum gives the best results because of the combination of the features that can be made. This increases the accuracy with 0.21 to 0.88.

For the translation of the technique to an *in vivo* setting, where it will be used during surgery, the large dependence on the near-infrared wavelength ranges is favorable. The main difference between an *ex vivo* setting and an *in vivo* setting is the presence of blood. For oxygenated and deoxygenated blood, the main absorption bands are located in the visual wavelength range. Therefore, blood absorption will have no influence in the near-infrared wavelength range.²⁸ The influence of blood in the translation to an *in vivo* setting using a classification method based mostly on the near-infrared wavelength range will thus be small.²⁹

Previous work in our group showed good results in combining a quadratic classifier and a linear SVM in tissue identification in colorectal cancer using fiberoptic DRS.²⁶ In the current study also, two different classifiers to distinguish between fat, healthy colorectal wall, and tumor are used. First, a quadratic classifier was used to classify fat; second, a linear SVM was used to distinguish healthy colorectal wall from tumor.

Because fat was first classified with the quadratic classifier, a binary task was left for the SVM. Therefore, a simple linear SVM could be used to distinguish healthy colorectal wall from tumor tissue. To perform a classification of three tissue types using only SVMs, a one-against-one or one-against-all classification should be performed. This will result in a combination of at least three SVMs, which will be a more complex classifier compared to the combination currently used. The more complex the classifier is, the more prone the classifier is to overfitting. Because the classification of fat was easy to perform, a more simple approach could be used in the form of a two-step classification, combining a quadratic classifier with a single SVM.

The classifiers created in this study were only based on the spectral features. However, because 2-D images were obtained with the hyperspectral cameras, there is also an option to use spatial and textural properties of the images to classify the pixels. In the current study, this option was not used because the spatial and textural properties during surgery will be very different compared to the properties, which would have been obtained in this study. However, for future studies, textural properties may be taken into account and could be used to further improve the current classification results.

In this *ex-vivo* study, the pathologist cut cross-section slices of the tumor and colorectal wall provide a large surface area of tumor and healthy tissue. This method was chosen to obtain a sufficient amount of data to create a reliable classification. In a surgical setting, these large surfaces of tumor will not be seen. Rectal tumors start developing in the mucosa and grow through the muscle layer into the surrounding mesorectal fat when becoming more advanced. In contrast to the large volume of tumor present in the lumen and wall of the rectum, smaller volumes of the tumor will be present in the mesorectal fat and possibly in the resection surface created by the surgeon. So, the main question for future research should be whether the current classification will still be able to detect an area of tumor tissue, which is much smaller compared to the cross-section slices and mainly surrounded by healthy tissue. So, as a next step toward *in vivo* use, the entire resected specimens should be imaged with HSI to validate the current accuracy in a more realistic setting.

To be able to perform HSI during surgery, some technical changes need to be made. The currently used set-up is a push-broom camera, where the samples are scanned by moving through the imaging line of the camera. In an *in vivo* setting, especially during laparoscopic surgery, this will not be possible. Therefore, a snapshot multispectral camera should be used, which can be attached to a laparoscopic system. In a multispectral camera, a limited number of wavelengths can be measured. These wavelengths should be chosen based on previous research. Therefore, further research should be performed on the selection of the most important wavelengths to distinguish tumor from healthy surrounding tissue. Using a snapshot camera will reduce the data acquisition time compared to the current set-up, from 1 min to 1 s. Moreover, the preprocessing time of the data will decrease because of the limited number of wavelengths acquired. Therefore, when the current set-up is transformed into a set-up that can be used *in vivo*, real-time tissue classification will be possible. Furthermore, in the *in vivo* setting, the measurements will be less controlled compared to the current study. For example, the illumination of the tissue will be variable during surgery. Moreover, specular reflection and glare will be present. These issues should be taken into consideration before starting

an *in vivo* study. Finally, when performing measurements *in vivo*, a real-time classification should be available. The current classification method would allow such real-time use.

5 Conclusion

In this *ex vivo* study, fat, healthy colorectal wall, and tumor tissue could be distinguished using HSI with an accuracy of 0.88 (STD = 0.13). When the accuracy is determined per patient, a mean accuracy of 0.93 (STD = 0.12) was obtained. Two hyperspectral cameras were used, one in the visual wavelength range and one in the near-infrared wavelength range. Using only one of the two cameras decreased the accuracy for the visual and near-infrared camera. The results of this study show the potential of using HSI during colorectal surgery to increase the number of radical resections. Future research should be focused on imaging of entire specimen and the translation of the technique to an intraoperative setting. This should result in a technique that provides accurate real-time tissue classification during laparoscopic colorectal cancer surgery.

Disclosures

None of the authors have relevant financial interests in the manuscript or other potential conflicts of interest to disclose.

Acknowledgments

The authors thank B. Dashtbozorg for reviewing of the manuscript. This work was supported by Stichting Stop Darmkanker. We are grateful for their support. E.J.M.B and E.N.D.K. together with S.G.B. performed all measurements done for this research. A.G.J.A., N.F.M.K., G.L.B., K.F.D.K., S.C.B., and T.J.M.R. assisted with the data acquisition. J.S. and C.C.F. performed the histological examination of all pathology slides used in this research. E.J.M.B. and E.N.D.K. did the analysis and interpretation of the data with the support of H.J.C.M.S. and T.J.M.R. The writing of this manuscript was done by E.J.M.B. and E.N.D.K., with major contributions of S.G.B., H.J.C.M.S., and T.J.M.R. and revisions of A.G.J.A., N.F.M.K., K.F.D.K., G.L.B., C.C.F. and S.C.B. All authors read and approved the final manuscript. This study was funded by Stichting Stop Darmkanker.

References

1. J. Ferlay et al., "Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012," *Int. J. Cancer* **136**(5), E359–E386 (2015).
2. L. Rabeneck et al., "Cancer: disease control priorities," in *Disease Control Priorities*, 3rd ed., pp. 101–119, World Bank, Washington, DC (2015).
3. M. Pascual, S. Salvans, and M. Pera, "Laparoscopic colorectal surgery: current status and implementation of the latest technological innovations," *World J. Gastroenterol.* **22**(2), 704–717 (2016).
4. M. V. Ottermo et al., "The role of tactile feedback in laparoscopic surgery," *Surg. Laparosc. Endosc. Percutan. Tech.* **16**(6), 390–400 (2006).
5. S. Schostek, M. O. Schurr, and G. F. Buess, "Review on aspects of artificial tactile feedback in laparoscopic surgery," *Med. Eng. Phys.* **31**(8), 887–898 (2009).
6. G. Lu and B. Fei, "Medical hyperspectral imaging: a review," *J. Biomed. Opt.* **19**(1), 10901 (2014).
7. J. M. Benavides et al., "Multispectral digital colposcopy for in vivo detection of cervical cancer," *Opt. Express* **11**(10), 1223–1236 (2003).
8. L. E. Boucheron et al., "Utility of multispectral imaging for nuclear classification of routine clinical histopathology imagery," *BMC Cell Biol.* **8**(Suppl 1), S8–S8 (2007).

9. D. Hattery et al., "Hyperspectral imaging of Kaposi's Sarcoma for disease assessment and treatment monitoring," in *Proc. Appl. Imagery Pattern Recognit. Workshop*, pp. 124–130 (2002).
10. Z. Liu, H. Wang, and Q. Li, "Tongue tumor detection in medical hyperspectral images," *Sensors (Basel)* **12**(1), 162–174 (2012).
11. B. Fei et al., "Label-free reflectance hyperspectral imaging for tumor margin assessment: a pilot study on surgical specimens of cancer patients," *J. Biomed. Opt.* **22**(8), 086009 (2017).
12. G. Lu et al., "Detection of head and neck cancer in surgical specimens using quantitative hyperspectral imaging," *Clin. Cancer Res.* **23**(18), 5426–5436 (2017).
13. H. Akbari et al., "Cancer detection using infrared hyperspectral imaging," *Cancer Sci.* **102**(4), 852–857 (2011).
14. S. Kiyotoki et al., "New method for detection of gastric cancer by hyperspectral imaging: a pilot study," *J. Biomed. Opt.* **18**, 26010–26017 (2013).
15. K. Masood, N. M. Rajpoot, and M. Nasir, "Spatial analysis for colon biopsy classification from hyperspectral imagery," 2008, <http://wrap.warwick.ac.uk/37082/#.WmdDFQM7u8w.mendeley> (23 January 2018).
16. K. Masood, N. M. Rajpoot, and M. Nasir, "Classification of colon biopsy samples by spatial analysis of a single spectral band from its hyperspectral cube," 2007, <http://wrap.warwick.ac.uk/61638/#.WmdFQdBrhPY.mendeley> (23 January 2018).
17. K. Rajpoot and N. Rajpoot, "SVM optimization for hyperspectral colon tissue cell classification," *Lect. Notes Comput. Sci.* **3217**, 829–837 (2004).
18. S. Rathore et al., "A recent survey on colon cancer detection techniques," *IEEE/ACM Trans. Comput. Biol. Bioinf.* **10**(3), 545–563 (2013).
19. M. Maggioni et al., "Hyperspectral microscopic analysis of normal, benign and carcinoma microarray tissue sections," *Proc. SPIE* **6091**, 60910I (2006).
20. E. Claridge and D. Hidović-Rowe, "Model based inversion for deriving maps of histological parameters characteristic of cancer from ex-vivo multispectral images of the colon," *IEEE Trans. Med. Imaging* **33**(4), 822–835 (2014).
21. Z. Han et al., "In vivo use of hyperspectral imaging to develop a non-contact endoscopic diagnosis support system for malignant colorectal tumors," *Proc. SPIE* **21**(1), 016001 (2016).
22. R. Kumashiro et al., "Integrated endoscopic system based on optical imaging and hyperspectral data analysis for colorectal cancer detection," *Anticancer Res.* **36**(8), 3925–3932 (2016).
23. Ocean Optics, "OOINLCorrect loading non-linearity correction coefficients instructions," *Dunedin*, 2012, <https://oceanoptics.com/wp-content/uploads/OOINLCorrect-Linearity-Coeff-Proc.pdf>.
24. J. Burger and P. Geladi, "Hyperspectral NIR image regression part I: calibration and correction," *J. Chemom.* **19**(5–7), 355–363 (2005).
25. Å. Rinnan, F. van den Berg, and S. B. Engelsen, "Review of the most common pre-processing techniques for near-infrared spectra," *TRAC Trends Anal. Chem.* **28**(10), 1201–1222 (2009).
26. E. J. M. Baltussen et al., "Diffuse reflectance spectroscopy as a tool for real-time tissue assessment during colorectal cancer surgery," *J. Biomed. Opt.* **22**, 106014 (2017).
27. S. Boughorbel, F. Jarray, and M. El-Anbari, "Optimal classifier for imbalanced data using Matthews correlation coefficient metric," *PLoS One* **12**(6), e0177678 (2017).
28. T. M. Bydlon et al., "Chromophore based analyses of steady-state diffuse reflectance spectroscopy: current status and perspectives for clinical adoption," *J Biophotonics* **8**(1–2), 9–24 (2015).
29. J. W. Spliethoff et al., "Real-time in vivo tissue characterization with diffuse reflectance spectroscopy during transthoracic lung biopsy: a clinical feasibility study," *Clin Cancer Res.* **22**(2), 357–365 (2016).

Elisabeth J. M. Baltussen is a PhD student at The Netherlands Cancer Institute. She received her master's degree in Technical Medicine from the University of Twente, The Netherlands, in 2014. Since 2015, she is working as a PhD student at The Netherlands Cancer Institute—Antoni van Leeuwenhoek, Amsterdam, The Netherlands, at the Department of Surgery. Her research is focused on the implementation of optical technology during colorectal surgery to improve the surgical outcome.

Esther N. D. Kok is a PhD student at The Netherlands Cancer Institute. She studied medicine at Vrije Universiteit in Amsterdam and received her master's degree in 2017. Subsequently, she started working at the Department of Surgery as a PhD student at The Netherlands Cancer Institute—Antoni van Leeuwenhoek, Amsterdam. The focus of her research is the treatment of patients with colorectal cancer and the possible additional liver metastases.

Biographies of the other authors are not available.