Supplementary Materials for:

Spectral endoscopy enhances contrast for neoplasia in surveillance of Barrett's esophagus

Dale J Waterhouse[1], Wladyslaw Januszewicz[2], Sharib Ali[3], Rebecca C Fitzgerald[4], Massimiliano di Pietro[4]* and Sarah E Bohndiek[1]*

1. Department of Physics and Cancer Research UK Cambridge Institute, University of Cambridge, UK
2. Department of Gastroenterology, Hepatology and Clinical Oncology, Medical Centre for Postgraduate Education, Warsaw, Poland
3. Institute of Biomedical Engineering, University of Oxford, UK
4. MRC Cancer Unit, Hutchison/MRC Research Centre, University of Cambridge, UK

Supplementary Methods:

Exclusion Criteria

Exclusion criteria included history of esophageal stricture, pregnancy or breastfeeding, history of esophageal varices or liver impairment of moderate or worse severity (Child's Pugh class B or C), history of esophageal surgery except for uncomplicated fundoplication, history of coagulopathy (INR>1.3 and/or platelet count <75000) or on clopidogrel and/or anti-coagulant medication for high risk condition and unable to withhold medication temporarily.

Classification using Deep-Learning

To classify the attenuation spectra, a 1D convolutional neural network (CNN) model was trained. Classical 1D CNN convolution kernels were used to obtain feature maps and a rectified linear unit (ReLU) was used as an activation function. Each convolution was also followed by a max pooling layer with stride of 2. Dropout was introduced in the 2nd, 3rd and the 4th layers to reduce the overfitting of the network as this could be the case, for e.g., the marginal gaps between NDBE and neoplasia. A log-softmax classifier was used for improved gradient optimization and to compute the logarithmic probability of each class. The complete block diagram is shown in Figure 5 in the main text. An Adam (adaptive moment estimation) optimizer with a learning rate (LR) hyperparameter of 0.001 and weight decay of 0.0001 x LR was used to minimize negative log likelihood loss. The hyperparameters $\beta 1$ and $\beta 2$, which represent exponential decay rate for the first and second moment, respectively, are empirically chosen to be $\beta 1 = 0.90$ and $\beta 2 = 0.999$. To account for a smaller number of samples in the balanced dataset and presence of narrow variability in class instances, we have adjusted the LR to 0.0008 and weight decay to 0.00001 x LR that allowed to obtain a smoother loss computation and avoid overfitting. An open-source PyTorch platform (1) was used to implement the described model architecture. All training was performed on NVIDIA GeForce RTX 2080 Ti with batch size of 16 and 20% validation (random) split on the training data. The train-time was approximately 1 min and test-time per-spectrum was < 1 ms.

Spectral Acquisition and Processing

Dark Spectra Acquisition

During each trial, dark spectra, $D_i(\lambda, t_{\text{exp}})$, were incidentally captured when both the SOC and spectral endoscope light sources were switched off, where t_{exp} is the exposure time. t_{exp} is manually controlled to achieve good images during endoscopy.

Dark Spectra Processing

For each exposure time, the mean dark spectrum across all trials was calculated,

$$\overline{D}(\lambda, t_{\text{exp}}) = \frac{1}{n_d(t_{\text{exp}})} \sum_{i=1}^{n_d(t_{\text{exp}})} D_i(\lambda, t_{\text{exp}}) \quad \text{(S1)}$$

where $n_d(t_{\text{exp}})$ is the number of dark spectra per exposure time.

White Spectra Acquisition

At the end of each trial, a series of white reflectance spectra, $I_i'(\lambda, t_{\text{exp}})$, were captured from a diffusely reflecting white tile, where t_{exp} is the exposure time.

White Spectra Processing

1. Saturated white reflectance spectra, $\max\left(I_i'(\lambda, t_{\text{exp}})\right) > 16383$, were excluded.
2. White reflectance spectra with low signal, $\max\left(I_i'(\lambda, t_{\text{exp}})\right) \leq 500$, were excluded.
3. Using the dark spectrum with the appropriate exposure time for a given spectral data acquisition, dark subtraction was performed according to,

$$I_i'(\lambda, t_{\text{exp}}) = I_i'(\lambda, t_{\text{exp}}) - \overline{D}(\lambda, t_{\text{exp}}) \quad \text{(S2)}$$

Since the exposure time was regularly changed during each endoscopy, some trials included spectral data captured at an exposure time for which no dark spectra were captured. In this

case, the dark spectrum was linearly interpolated from the dark spectra at the other exposure times.

4. An additional background subtraction was performed by fitting a first-degree polynomial to each spectrum in the regions from 170 – 400 nm and 950 – 1100 nm, where no illumination is present, and subtracting the result.
5. White reflectance spectra were normalized to area under the curve (AUC) = 1,

$$I_i''(\lambda, t_{\text{exp}}) = \frac{I_i'(\lambda, t_{\text{exp}})}{\int I_i'(\lambda, t_{\text{exp}})\text{d}\lambda} \tag{S3}$$

6. For each exposure time, the mean white reflectance spectrum was calculated,

$$\overline{I''}(\lambda, t_{\text{exp}}) = \frac{1}{n_w(t_{\text{exp}})} \sum_{i=1}^{n_w(t_{\text{exp}})} I_i''(\lambda, t_{\text{exp}}) \tag{S4}$$

where $n_w(t_{\text{exp}})$ is the number of white reflectance spectra per exposure time. Note: since $I_i''(\lambda, t_{\text{exp}})$ was normalised to AUC=1, it should be independent of t_{exp}, but $\overline{I''}(\lambda, t_{\text{exp}})$ could be plotted prior to the final averaging step (Step 16) to check this.

7. Over all exposure times, and all trials with reliable white reflectance spectra (as determined by consistency with other trials), the mean white reflectance spectrum $\overline{I''}(\lambda)$ was calculated.
8. The mean white reflectance spectrum, $\overline{I''}(\lambda)$, was smoothed using the MATLAB® function 'smooth' using a moving average with a window size of 30, corresponding to ~15 nm.

Signal Spectra Processing

The raw acquired signal, $S(\lambda)$, was processed as follows **(Supplementary Fig. 2)**.

1. Saturated spectra, $\max(S(\lambda)) > 16383$, were excluded ($S(\lambda)$ is a 14-bit unsigned integer so it has maximum value 16384).
2. Using the dark spectrum with the appropriate exposure time for a given spectral data acquisition, dark subtraction was performed according to,

$$S'(\lambda, t_{\text{exp}}) = S(\lambda, t_{\text{exp}}) - \overline{D}(\lambda, t_{\text{exp}}) \tag{S5}$$

Since the exposure time was regularly changed during each endoscopy, some trials included spectral data captured at an exposure time for which no dark spectra were captured. In this case, the dark spectrum was linearly interpolated from the dark spectra at the other exposure times.

3. An additional background subtraction was performed by fitting a first-degree polynomial to each spectrum in the regions from 170 – 400 nm and 950 – 1100 nm, where no illumination is present, and subtracting the result.
4. Spectra with low signal, $\max\left(S'(\lambda)\right) \leq 100$, were excluded.
5. Spectra, $S'(\lambda)$, were smoothed using the MATLAB® function 'smooth' using a moving average with a window size of 30, corresponding to ~15 nm.
6. Spectra containing negative values in the range 470 – 720 nm, $S'(470-720) < 0$, were excluded.
7. Spectra containing low signal in the range 500 – 700 nm, $\min\left(S'(500-700)\right) \leq 1$, were excluded.
8. The tissue reflectance was calculated as,

$$R(\lambda) = \frac{S'(\lambda)}{\overline{I''}(\lambda)} \tag{S6}$$

9. The attenuation was calculated as,

$$A(\lambda) = -\log_e R(\lambda) \tag{S7}$$

10. Finally, attenuation spectra were normalized by subtracting the minimum attenuation from 500 – 720 nm,

$$A'(\lambda) = A(\lambda) - \min_{\lambda=500-720\ \mathrm{nm}}\left(A(\lambda)\right) \tag{S8}$$

Variance in Tissue Attenuation Spectra

The tissue attenuation spectra for each individual trial are shown in **Supplementary Fig. 5.** The differences between spectra of different pathology classes are conserved within each patient but substantial variation is clearly evidence between patients.

To quantify the nature of the variation in the spectra, one-way analysis of variance was performed. This proceeded by calculating the within-group and between-group variance at each wavelength according to,

$$\mathrm{MS_w}(\lambda) = \frac{\sum_{i=1}^{k}(n_i - 1)s(\lambda)_i^2}{n - k} \tag{S9}$$

and

$$\mathrm{MS_b}(\lambda) = \frac{\sum_{i=1}^{k} n_i\left(\overline{y_i}(\lambda) - \bar{Y}(\lambda)\right)^2}{k - 1} \tag{S10}$$

where $i = 1 - k$ is the index of each group within the dataset, k is the total number of groups, n_i is the number of observations in group i, n is the total number of observations, $s(\lambda)_i^2$ is the variance of all observations in group i at wavelength λ, $\overline{y_i}(\lambda)$ is the mean of all observations in group i at wavelength λ and $\bar{Y}(\lambda)$ is the global mean of all observations at wavelength λ.

The variance ratio,

$$\mathrm{F}(\lambda) = \frac{\mathrm{MS_b}(\lambda)}{\mathrm{MS_w}(\lambda)} \tag{S11}$$

describes the ratio of the within-group to between-group variance.

Finally, these measures were averaged over the wavelength range of interest (470 – 720 nm) and the square root taken to reach,

$$\sqrt{\mathrm{F}} = \sqrt{\frac{1}{N}\sum_{\lambda=470}^{720}\mathrm{F}(\lambda)} \tag{S12}$$

$$\mathrm{RMS_w} = \sqrt{\mathrm{MS_w}} = \sqrt{\frac{1}{N}\sum_{\lambda=470}^{720}\mathrm{MS_w}(\lambda)} \tag{S13}$$

$$\mathrm{RMS_b} = \sqrt{\mathrm{MS_b}} = \sqrt{\frac{1}{N}\sum_{\lambda=470}^{720} \mathrm{MS_b}(\lambda)} \tag{S14}$$

where $N = 433$ is the number of wavelengths.

These metrics were calculated using two distinct groupings. First, in patients where two distinct regions of the same tissue pathology were available, the spectra were grouped by region. Thus, $\mathrm{RMS_w^{region}}$ represents the intra-region variance and $\mathrm{RMS_b^{region}}$ represents the inter-region variance and $\sqrt{\mathrm{F^{region}}}$ represents their ratio. Second, all patients were grouped by patient, thus $\mathrm{RMS_w^{patient}}$ represents the intra-patient variance (irrespective of region) and $\mathrm{RMS_b^{patient}}$ represents the inter-patient variance and $\sqrt{\mathrm{F^{patient}}}$ represents their ratio.

An F-test for equality of variances cannot be accurately applied to the acquired F-statistics as these are an average of non-independent $\mathrm{F}(\lambda)$, but an F-test was performed using $\mathrm{F}(550\ \mathrm{nm})$ to give an indication of the significance of the variation between vs. within the groups. The inter-region variance was significant in 1/4 of the patients with multiple NDBE regions (trial number 8, $p<0.001$) and in 2/3 patients with multiple neoplastic regions (trial numbers 12 and 18, $p=0.040$, $p<0.001$ respectively). The inter-patient variance was significant in all pathologies ($p<0.001$).

Simulation of Multispectral Images

Details of the procedure used to simulate images from the acquired spectra are as follows.

Data Preparation

1. The acquired tissue spectra were processed to yield reflection spectra (see Supplementary Equation 6), $R(\lambda)$, because it is the reflected light, rather than the attenuated light, that enters a detection system, which is the process we intended to model.
2. These were normalized to the maximum from 470 – 720 nm.
3. These were trimmed to the wavelength range 470 – 750 nm, the region where the reflection spectra are reliably measured.
4. For each pathology, the mean spectrum across all patients, $\bar{R}(\lambda)$, was calculated.

Data Augmentation

To enable the generation of an arbitrarily large set of noised reflection spectra, the next step was to perform principal component analysis (PCA)-based noising of $\bar{R}(\lambda)$. PCA-based noising was used as it ensures the spectral shape of the simulated noise reproduces that of the measured noise.

1. Principal components of variation in $R(\lambda)$, and the associated eigenvalues were determined;

$$PC_i(\lambda), \qquad i = 1 - N \tag{S15}$$

$$E_i, \qquad i = 1 - N \tag{S16}$$

where N is the number of principal components, which is equal to the number of wavelengths. Note that only $n - 1$ of these are non-trivial, where n is the number of spectra.

2. Next, these were used to augment $\bar{R}(\lambda)$ to generate noised spectra,

$$\tilde{R}_j(\lambda) = \bar{R}(\lambda) + D \sum_{i=1}^{N} \left(\rho \cdot PC_i(\lambda) \cdot \sqrt{s^2 \cdot \frac{E_i}{E_T}} \right) \tag{S17}$$

with,

$$E_T = \sum_{i=1}^{N} E_i \tag{S18}$$

where ρ is a random number drawn from a standard normal distribution, s^2 is the sum of variances across all wavelengths, and D is a dimensionless constant representing the degree of noise to augment. A value of $D = 1$ represents equal noise to that found in the original dataset and a value of $D = 0$ represents zero noise. For modelling purposes, a value of $D = 0.25$ was used. This was repeated M times to generate $j = 1 - M$ augmented reflection spectra $\tilde{R}_j(\lambda)$.

3. Since narrow band imaging uses a narrow band centered at 415 nm, it was necessary to extrapolate the augmented data in the low wavelength region. Since the reflection spectra are relatively flat and smooth in this region, this extrapolation was performed by fitting a first-degree polynomial to the region from 470 – 500 nm and using this to extrapolate the region 380 – 470 nm for each augmented spectrum $\tilde{R}(\lambda)$.

Spectral Band Optimization

To select the spectral bands for simulated spectral endoscopy, the following steps were taken.

1. Data augmentation steps 1 – 2 were performed as described above to generate augmented reflection spectra $\tilde{R}_j(\lambda)$ in the region 470 – 750 nm.
2. Gaussian spectral response bands with full width half maxima of 10 nm were generated according to,

$$S({\lambda^c}_k, H_l, \lambda) = H_l \cdot \exp\left(\frac{-4 \cdot \log(2) \cdot (\lambda - {\lambda^c}_k)^2}{10^2}\right), \tag{S19}$$

where the center wavelengths are,

$${\lambda^c}_k = \begin{cases} 490, & k = 1 \\ 500, & k = 2 \\ \vdots & \\ 720, & k = 24 \end{cases} \tag{S20}$$

and the sensitivities are,

$$H_l = \begin{cases} 0.2, & l = 1 \\ 0.4, & l = 2 \\ \vdots & \\ 1.0, & l = 5 \end{cases} \tag{S21}$$

3. The augmented spectra $\tilde{R}_j(\lambda)$ were propagated through the spectral response bands to generate signals,

$$I_j({\lambda^c}_k, H_l) = \int_{470}^{720} S({\lambda^c}_k, H_l, \lambda)\, \tilde{R}_j(\lambda)\, d\lambda \tag{S22}$$

4. The average signal for each band was calculated as

$$I({\lambda^c}_k, H_l) = \frac{1}{M}\sum_{j=1}^{M} I_j({\lambda^c}_k, H_l) \tag{S23}$$

5. For every combination and permutation of 3 spectral response bands (not allowing repetition of λ_0, but allowing repetition of H), an rgb color, C_{rgb}, was constructed.

   For example, for $k = 1, 6, 19$ and $l = 1, 3, 2$,

$$r = I(490, 0.2), \qquad \mathrm{g} = I(540, 0.6), \qquad \mathrm{b} = I(670, 0.4), \tag{S24}$$

$$C_{rgb} = \frac{0.8}{\max(\mathrm{r}, \mathrm{g}, \mathrm{b})}\begin{pmatrix}\mathrm{r}\\ \mathrm{g}\\ \mathrm{b}\end{pmatrix} \tag{S25}$$

   where the color is normalized to a peak value of 0.8. This normalization is applied because (i) we don't have relative intensity information for our acquired spectra and (ii) we are interested in chromaticity (color independent of brightness). The value of 0.8 is selected to produce a bright but not saturated image.

   For K center wavelengths, L sensitivities, this results in a total of N_c colors, C_{rgb}, where N_c is given by,

$$N_c = \frac{K!}{(K-3)!\,3!} \cdot 3! \cdot L^3 \tag{S26}$$

6. The above process was repeated for NDBE and neoplasia and the International Commission on Illumination Delta E 2000 (CIEDE2000) color difference (ΔE_{00}) calculated (2).
7. The set of 3 spectral response bands that resulted in the largest CIEDE2000 color difference was selected as the spectral endoscopy filter set.

Image Simulation

To simulate images, augmented reflection spectra $\tilde{R}_j(\lambda)$ were first generated as described above. For narrow band and spectral endoscopy, the color of each pixel was simulated as follows.

1. Using a novel $\tilde{R}_j(\lambda)$ for each pixel, the rgb color,

$$\begin{pmatrix} \mathrm{r}_j \\ \mathrm{g}_j \\ \mathrm{b}_j \end{pmatrix} = \int_{380}^{750} \begin{pmatrix} S_\mathrm{r}(\lambda) \\ S_\mathrm{g}(\lambda) \\ S_\mathrm{b}(\lambda) \end{pmatrix} \tilde{R}_j(\lambda)\, d\lambda \tag{S27}$$

   was calculated, where $S_{\mathrm{r/g/b}}(\lambda)$ are the red/green/blue spectral response bands of the sensor.

2. And thus the rgb color of the pixel is defined as above,

$$C_{\mathrm{rgb}_j} = \frac{0.8}{\max(\mathrm{r}_j, \mathrm{g}_j, \mathrm{b}_j)} \begin{pmatrix} \mathrm{r}_j \\ \mathrm{g}_j \\ \mathrm{b}_j \end{pmatrix} \tag{S28}$$

The r/g/b spectral response bands $S_{\mathrm{r/g/b}}(\lambda)$ for narrow band imaging are proprietary and unpublished, so they were estimated from the responses shown in the product information brochures (3) **(Supplementary Fig. 8)**. Since the spectral response bands $S_{\mathrm{r/g/b}}(\lambda)$ of standard of care HR-WLE are not published, a different approach to simulate HR-WLE color images was adopted. We assumed that the response and post-processing occurring in HR-WLE is such that the images represent the human visual response as closely as possible. So, the International Commission on Illumination (CIE) color matching functions were used to determine the rgb color as follows:

1. The xyz color coordinates were determined as,

$$\begin{pmatrix} X_j \\ Y_j \\ Z_j \end{pmatrix} = \frac{1}{\int_{380}^{750} \bar{y}(\lambda)\, d\lambda} \cdot \int_{380}^{750} \begin{pmatrix} \bar{x}(\lambda) \\ \bar{y}(\lambda) \\ \bar{z}(\lambda) \end{pmatrix} \tilde{R}_j(\lambda)\, d\lambda \tag{S29}$$

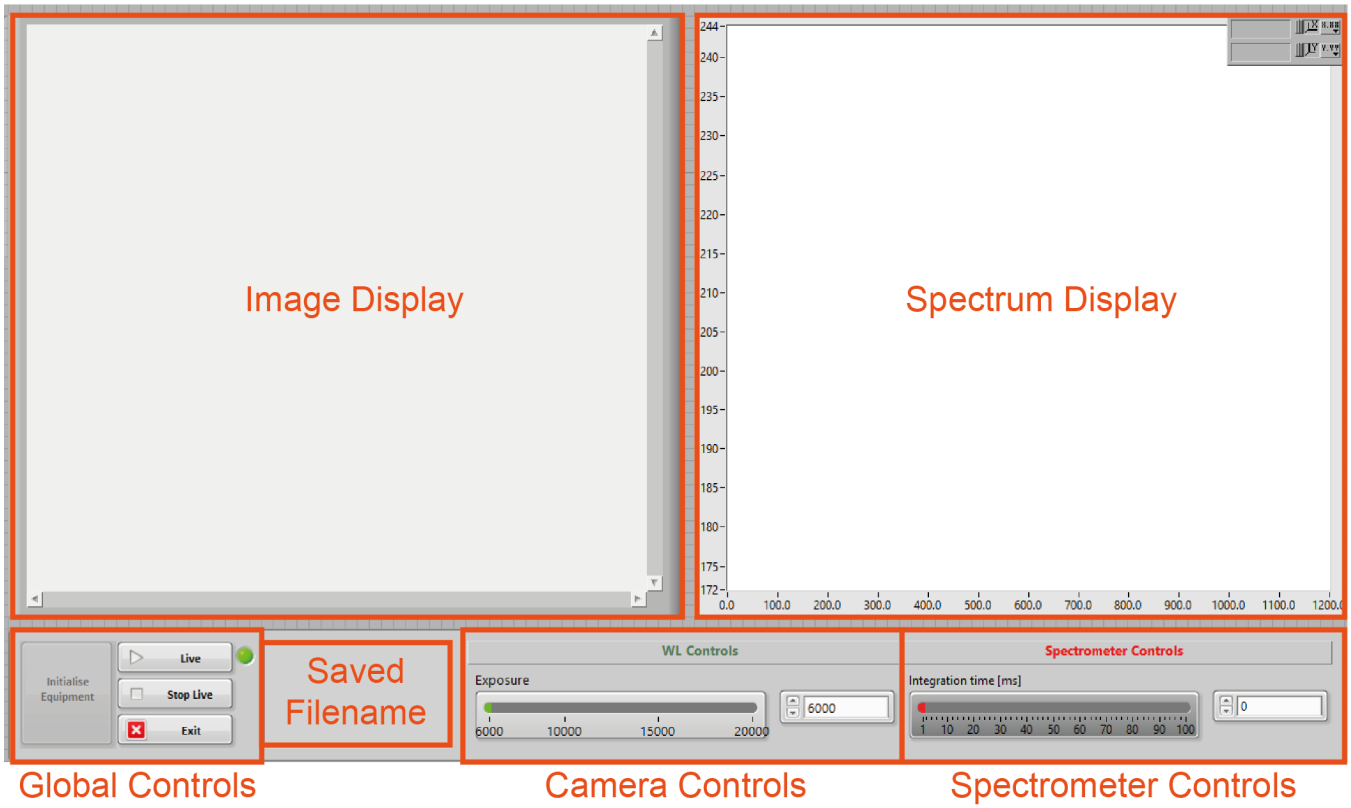
   and normalized such that,

$$\begin{pmatrix} x_j \\ y_j \\ z_j \end{pmatrix} = \frac{1}{X + Y + Z} \cdot \begin{pmatrix} X_j \\ Y_j \\ Z_j \end{pmatrix} \tag{S30}$$

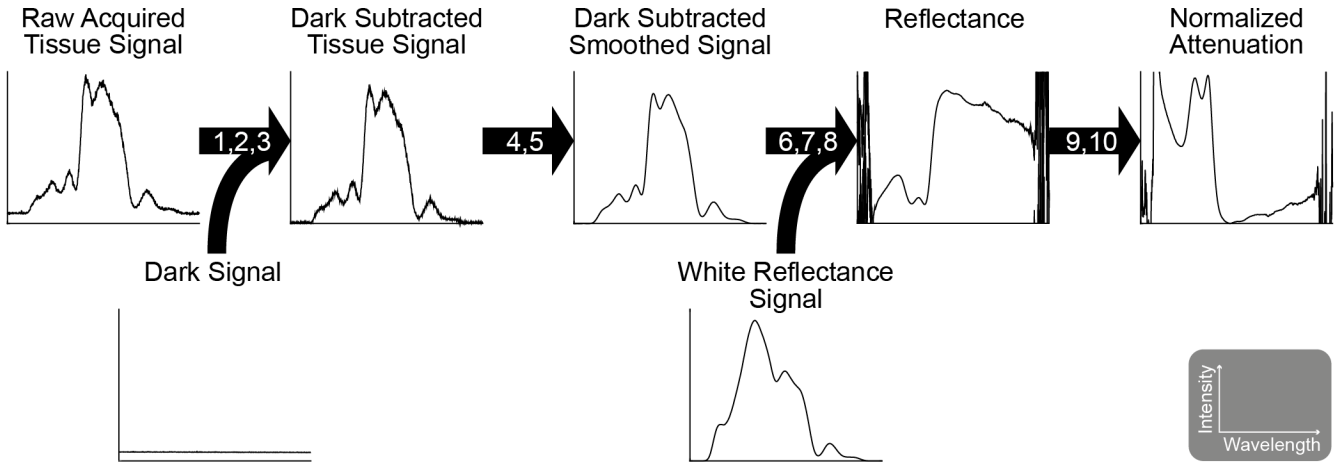
2. The xyz coordinates are transformed to rgb using the MATLAB® function ‘xyz2rgb’ using the standard CIE illuminant D65 and sRGB conversion, which applies a gamma correction.
3. The rgb coordinates are normalized as above so,

$$C_{\mathrm{rgb}_j} = \frac{0.8}{\max\left(\mathrm{r}_j, \mathrm{g}_j, \mathrm{b}_j\right)} \begin{pmatrix} \mathrm{r}_j \\ \mathrm{g}_j \\ \mathrm{b}_j \end{pmatrix} \tag{S31}$$

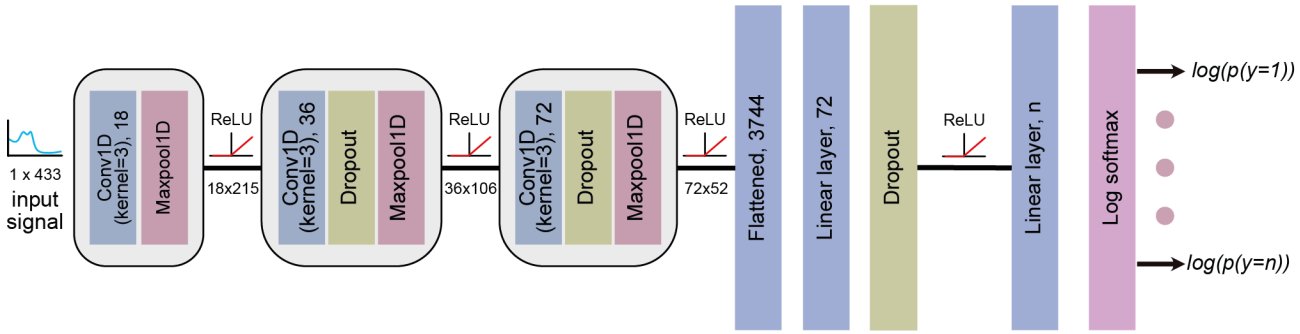
Supplementary Figures:



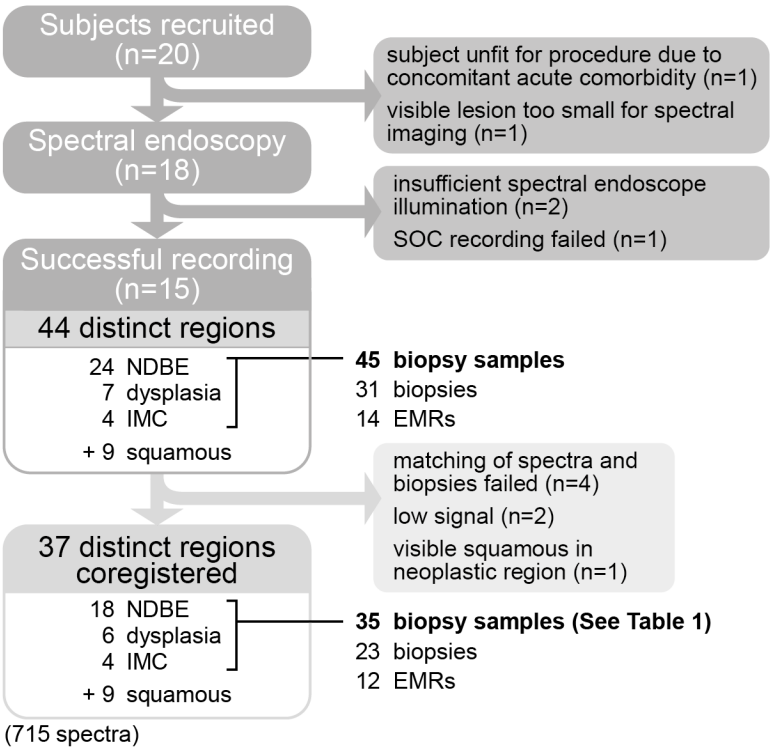
Supplementary Figure 1 | Spectral endoscope user interface. The interface used to control the spectral endoscope and display images and spectra in real time (developed in LabVIEW [National Instruments, USA], slightly modified for presentation).



Supplementary Figure 2 | Summary schematic of the spectral processing pipeline as described in "Spectral Acquisition and Processing". **1.** Saturated spectra removal; **2.** Subtraction of measured dark signal; **3.** Additional dark subtraction from linear fit to low signal regions; **4.** Low max signal removal; **5.** Smoothing; **6.** Negative value removal; **7.** Low min signal removal; **8.** Reflectance calculation by division by measured white reflectance signal; **9.** Attenuation calculation; **10.** Minimum normalization.

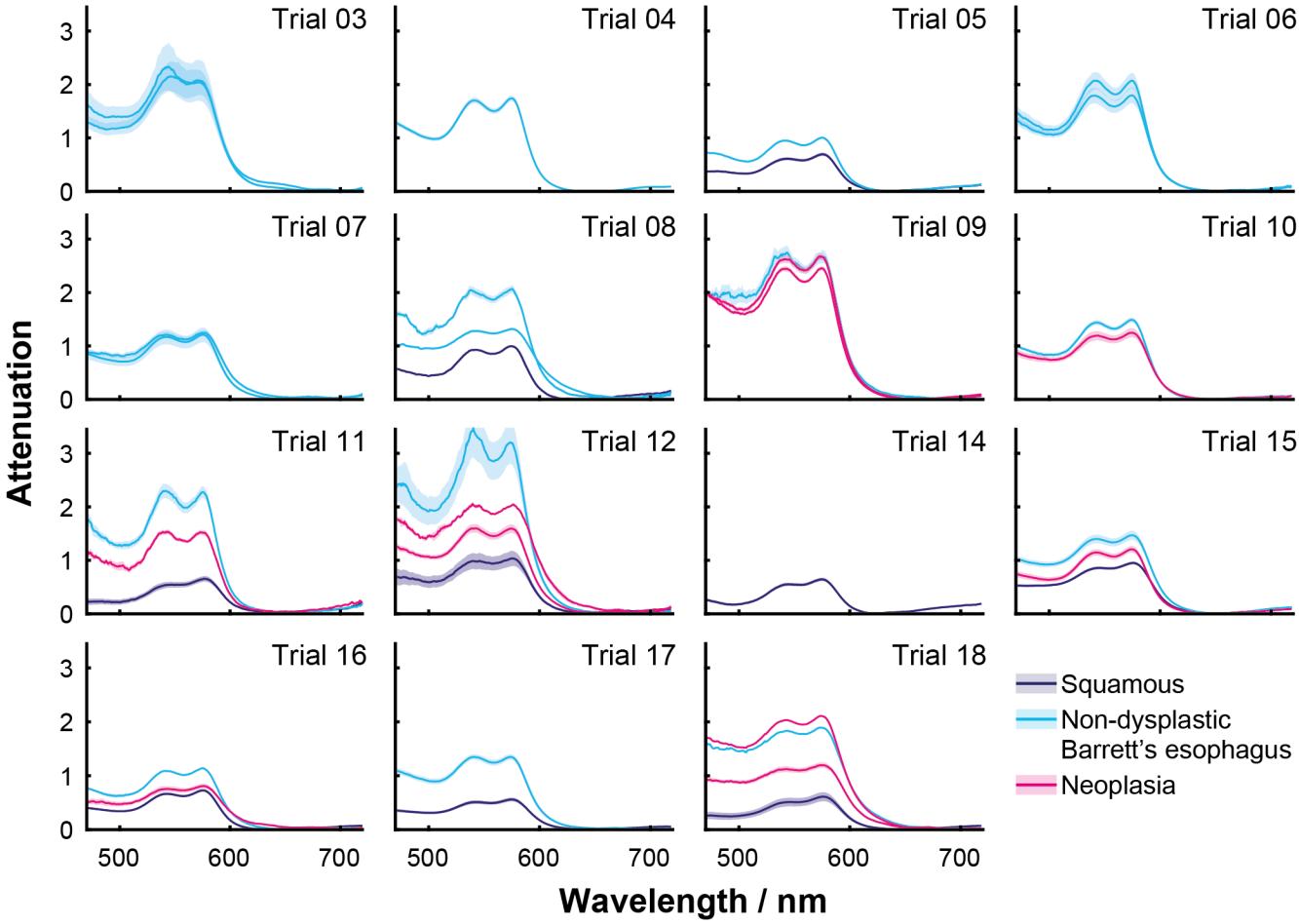


Supplementary Figure 3 | 1D convolutional neural network (CNN)-model used for classification of attenuation spectra in this study. Classical 1D CNN convolution kernels were used to obtain feature maps and a rectified linear unit (ReLU) was used as an activation function. Each convolution was also followed by a max pooling layer with stride of 2. Dropout was introduced in the 2nd, 3rd and the 4th layers to reduce the overfitting of the network as this could be the case for the marginal gaps between NDBE and neoplasia. A log-softmax classifier was used for improved gradient optimization and to compute the logarithmic probability of each class. N, number of classes; p, probability.

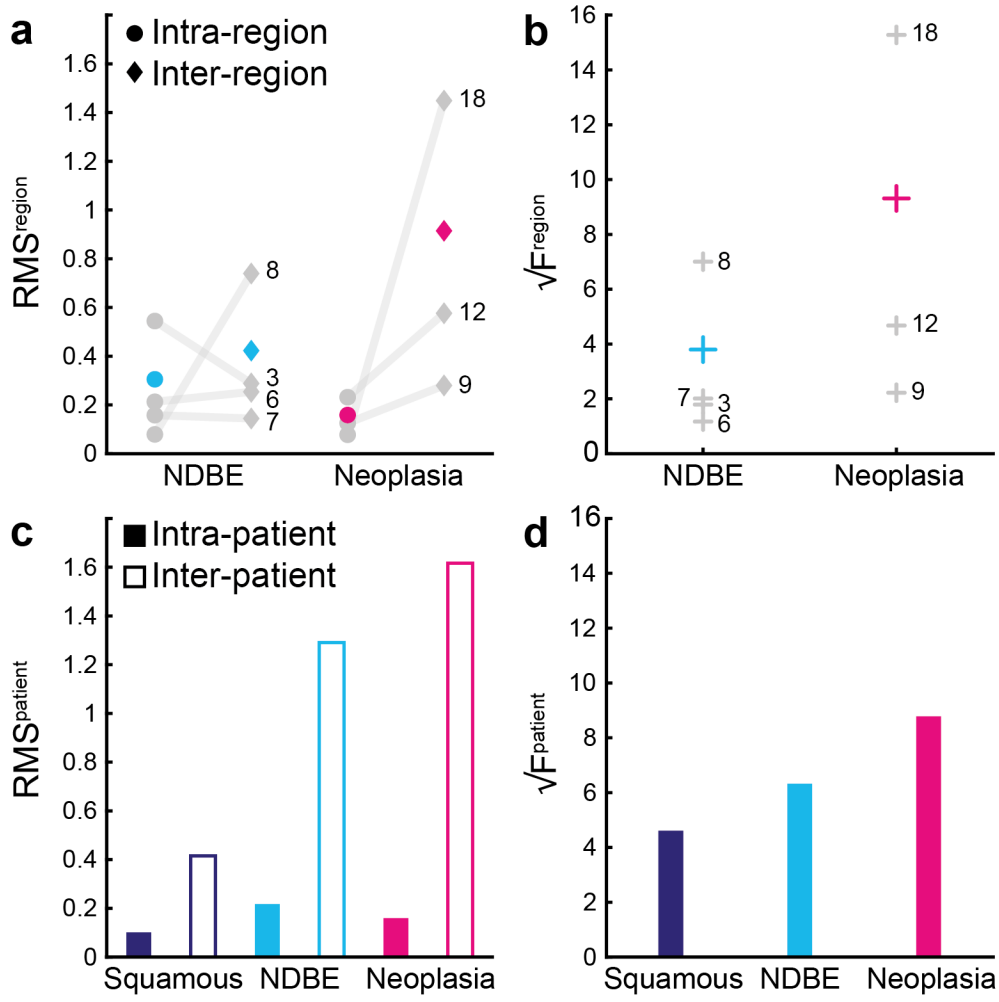


Supplementary Figure 4 | Recruitment flowchart.

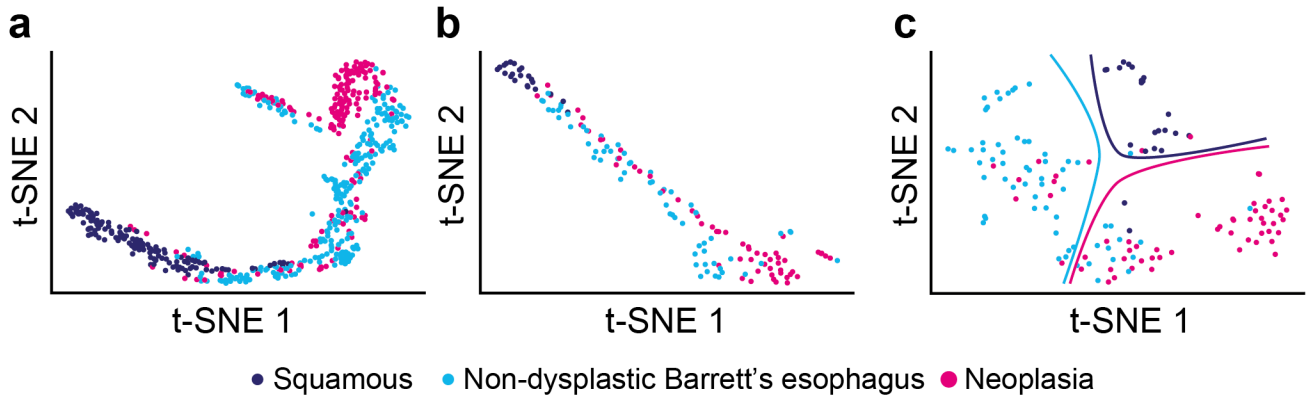
EMR, endoscopic mucosal resection; IMC, intramucosal carcinoma; NDBE, non-dysplastic Barrett's esophagus; SOC, standard of care.



Supplementary Figure 5 | Per-patient per-region attenuation spectra captured with the spectral endoscope across all patients. The shaded region represents the standard error.

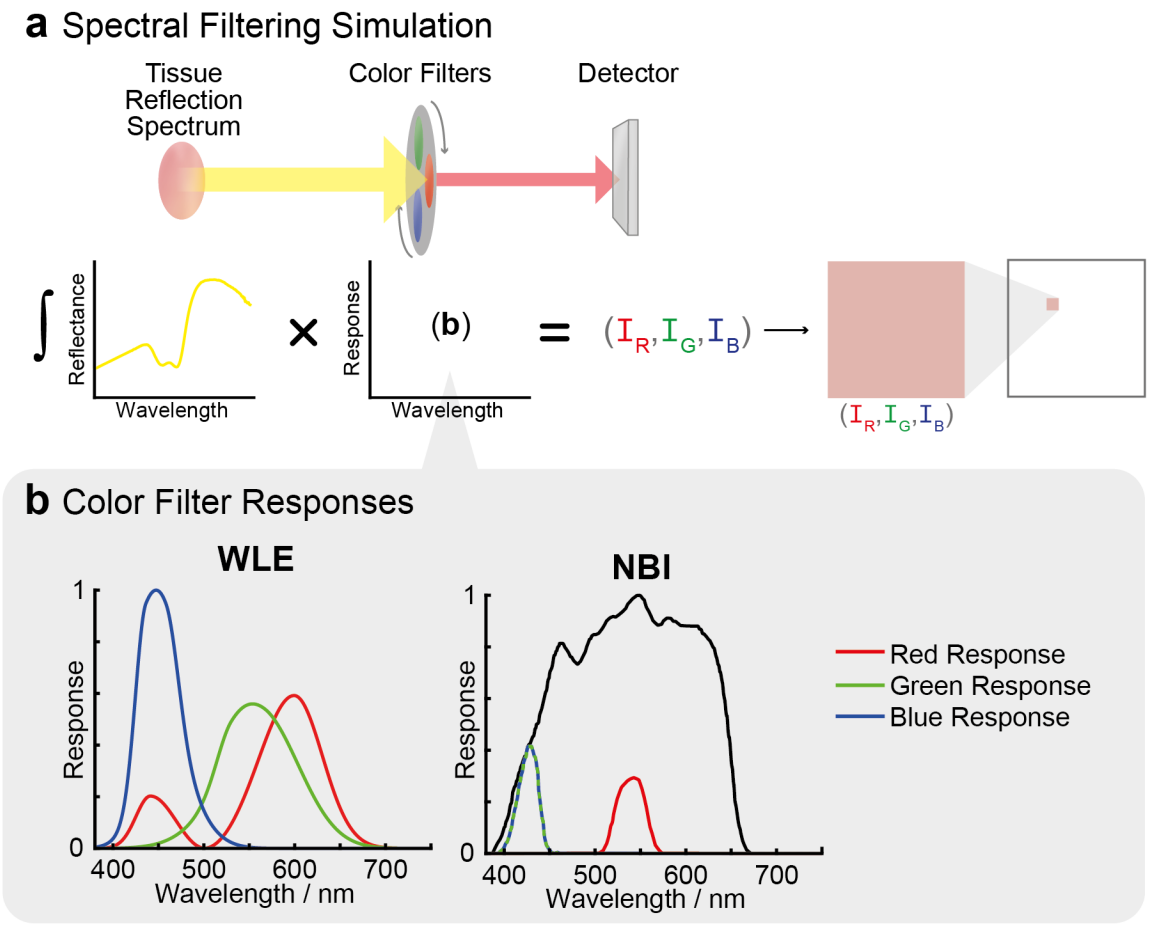


Supplementary Figure 6 | Variance of acquired attenuation spectra. (a) Intra-region vs inter-region variance (RMS^{region}). Grey lines connect points representing different regions captured from the same patient. Trial numbers are shown to the right-hand side of points for cross reference to Table 1. Colored points are the mean across patients (n=4 NDBE, n=3 neoplasia). **(b)** The variance ratio $\sqrt{F^{region}}$, describing the ratio of inter-region to intra-region variance. **(c)** Intra-patient vs inter-patient variance ($RMS^{patient}$). **(d)** The root mean variance ratio $\sqrt{F^{patient}}$, describing the ratio of inter-patient to intra-patient variance. NDBE, non-dysplastic Barrett's esophagus.



Supplementary Figure 7 | t-distributed Stochastic Neighbor Embedding (t-SNE). Components for 3 classes are plotted for **(a)** training, **(b)** test, and **(c)** predicted log probability from our trained CNN-model. t-SNE is a non-linear projection that allows us to project a high-dimensional feature vector (433-point spectrum in this case) to a 2-dimensional point.

t-SNE, t-distributed Stochastic Neighbor Embedding.



Supplementary Figure 8 | (a) Schematic representation of the process for simulating the propagation of the tissue reflectance spectrum through color filters to calculate detected intensities in the red, green and blue channels, I_R, I_G and I_B, thus generating a 3-channel color swatch. **(b)** The color filter responses used to simulate WLE and NBI images. The International Commission on Illumination (CIE) color matching functions were used to simulate the WLE color images from the acquired reflection spectra. The NBI response bands were calculated by multiplying the Olympus light source spectrum (3) (shown in black), with the narrow band response spectra shown in the product information brochures (3). The narrow band centered at 540 nm is applied to the red color channel and the narrow band centered at 415 nm is applied to the green and the blue channel. WLE, white light endoscopy; NBI, narrow band imaging.

Supplementary Tables

Supplementary Table 1 | Patient Demographics and Clinical Characteristics

Trial #	Trial Date	Collection of Matched Spectra and Histopathology	Comments	Sex	Age / years	Barrett's Length (C = Circumferential, M = Maximum Extent) / cm	Lesion Morphology (Paris Classification)
1	30/05/2018	N	Insufficient spectral endoscope illumination	M	71	C3M7	0-IIb 5 mm
2	25/09/2018	N	Insufficient spectral endoscope illumination	M	68	C3M5	0-IIb 5 mm
3	12/03/2019	Y		M	65	C0M2	0-IIb 5 mm
4	18/06/2019	Y		M	74	C4M6	0-IIa 10 mm
5	02/07/2019	Y		M	74	C5M7	no visible lesion
6	09/07/2019	Y		M	71	C11M14	no visible lesion
7	09/07/2019	Y		M	86	C11M12	no visible lesion
8	23/07/2019	Y		M	62	C6M7	no visible lesion
9	13/08/2019	Y		M	72	C5M7	0-IIb 15 mm
10	10/09/2019	Y		F	78	C14M14	0-IIb 10 mm
11	10/09/2019	Y		M	55	C2M5	0-IIb 4 mm
12	17/09/2019	Y		M	75	C2M4	0-IIb 5 mm
13	24/09/2019	N	Visible lesion too small for spectral imaging	M	59	C1M5	0-IIb 1 mm
14	24/09/2019	Y		M	80	C0M3	0-IIb 2 mm
15	01/10/2019	Y		M	74	C3M4	0-IIb 25 mm
16	15/10/2019	Y		M	67	C0M4	0-IIa 20 mm
17	15/10/2019	Y		F	67	C0M2	0-IIb 2 mm
18	12/11/2019	Y		F	72	C5M7	0-IIb 20 mm
19	19/11/2019	N	Subject was considered unfit for endoscopic procedure due to concomitant acute comorbidity	M	NR	NA	NA
20	17/12/2019	N	SOC recording failed	M	81	C8M9	0-IIb 10 mm

Supplementary Table 2 | Performance of k-Nearest-Neighbor-(KNN)-Based Classification of Tissue Attenuation Spectra.

Comparison	Classification Performance				
	Accuracy % (n/total)	Sensitivity % (n/total)	Specificity % (n/total)	Positive Predictive Value % (n/total)	Negative Predicative Value % (n/total)
3-way Classification					
Squamous	97.9 (140/143)	90.9 (20/22)	99.2 (120/121)	95.2 (20/21)	98.4 (120/122)
NDBE	82.5 (118/143)	88.9 (56/63)	77.5 (62/80)	75.7 (56/74)	89.9 (62/69)
Neoplasia	83.2 (119/143)	70.7 (41/58)	91.8 (78/85)	85.4 (41/48)	82.1 (78/95)
2-way Classification[a]					
NDBE vs. Squamous	97.9 (94/96)	98.3 (59/60)	97.2 (35/36)	98.3 (59/60)	97.2 (35/36)
Neoplasia vs. Squamous	97.5 (77/79)	97.9 (47/48)	96.8 (30/31)	97.9 (47/48)	96.8 (30/31)
Neoplasia vs. NDBE	81.2 (91/112)	74.4 (32/43)	85.5 (59/69)	76.2 (32/42)	84.3 (59/70)

k-fold cross-validation (k=5) was used during training

NDBE, non-dysplastic Barrett's esophagus.

[a] In 2-way comparisons, the class in *italics* is the target for purposes of classification performance metric.

Supplementary Table 3 | Performance of Support-Vector-Machine-(SVM)-Based Classification of Tissue Attenuation Spectra

Comparison	Classification Performance				
	Accuracy % (n/total)	Sensitivity % (n/total)	Specificity % (n/total)	Positive Predictive Value % (n/total)	Negative Predicative Value % (n/total)
3-way Classification					
Squamous	98.6 (141/143)	95.5 (21/22)	99.2 (120/121)	95.5 (21/22)	99.2 (120/121)
NDBE	83.2 (119/143)	95.2 (60/63)	73.8 (59/80)	74.1 (60/81)	95.2 (59/62)
Neoplasia	81.8 (117/143)	62.1 (36/58)	95.3 (81/85)	90.0 (36/40)	78.6 (81/103)
2-way Classification[a]					
NDBE vs. Squamous	99.0 (95/96)	100.0 (60/60)	97.2 (35/36)	98.4 (60/61)	100.0 (35/35)
Neoplasia vs. Squamous	97.5 (77/79)	97.9 (47/48)	96.8 (30/31)	97.9 (47/48)	96.8 (30/31)
Neoplasia vs. NDBE	87.5 (98/112)	76.7 (33/43)	94.2 (65/69)	89.2 (33/37)	86.7 (65/75)

k-fold cross-validation (k=5) was used during training

NDBE, non-dysplastic Barrett's esophagus.

[a] In 2-way comparisons, the class in *italics* is the target for purposes of classification performance metric.

Supplementary Table 4 | Performance of Convolutional Neural Network (CNN)-based Classification of Tissue Attenuation Spectra in Training Set

Results from the training set used to train the CNN for which classification results are presented in the main manuscript.

	Classification Performance				
Comparison	Accuracy % (n/total)	Sensitivity % (n/total)	Specificity % (n/total)	Positive Predictive Value % (n/total)	Negative Predicative Value % (n/total)
3-way Classification					
Squamous	98.8 (565/572)	99.3 (136/137)	98.6 (429/435)	95.8 (136/142)	99.8 (429/430)
NDBE	97.7 (559/572)	98.1 (252/257)	97.5 (307/315)	96.9 (252/260)	98.4 (307/312)
Neoplasia	97.6 (558/572)	93.8 (167/178)	99.2 (391/394)	98.2 (167/170)	97.3 (391/402)

NDBE, non-dysplastic Barrett's esophagus.

Supplementary Table 5 | Performance of Convolutional Neural Network (CNN)-based Classification of Tissue Attenuation Spectra for Different k-fold Cross Validation

Comparison	Classification Performance				
	Accuracy % (n/total)	Sensitivity % (n/total)	Specificity % (n/total)	Positive Predictive Value % (n/total)	Negative Predicative Value % (n/total)
			k = 2		
train					
Squamous	96.2 (550/572)	93.4 (128/137)	97.0 (422/435)	90.8 (128/141)	97.9 (422/431)
NDBE	86.9 (497/572)	88.3 (227/257)	85.7 (270/315)	83.5 (227/272)	90.0 (270/300)
Neoplasia	86.5 (495/572)	73.0 (130/178)	92.6 (365/394)	81.8 (130/159)	88.4 (365/413)
test					
Squamous	97.2 (139/143)	95.5 (21/22)	97.5 (118/121)	87.5 (21/24)	99.2 (118/119)
NDBE	83.2 (119/143)	90.5 (57/63)	77.5 (62/80)	76.0 (57/75)	91.2 (62/68)
Neoplasia	81.8 (117/143)	65.5 (38/58)	92.9 (79/85)	86.4 (38/44)	79.8 (79/99)
			k = 3		
train					
Squamous	97.9 (560/572)	97.1 (133/137)	98.2 (427/435)	94.3 (133/141)	99.1 (427/431)
NDBE	89.7 (513/572)	93.8 (241/257)	86.3 (272/315)	84.9 (241/284)	94.4 (272/288)
Neoplasia	89.3 (511/572)	74.2 (132/178)	96.2 (379/394)	89.8 (132/147)	89.2 (379/425)
test					
Squamous	97.2 (139/143)	95.5 (21/22)	97.5 (118/121)	87.5 (21/24)	99.2 (118/119)
NDBE	85.3 (122/143)	93.7 (59/63)	78.8 (63/80)	77.6 (59/76)	94.0 (63/67)
Neoplasia	85.3 (122/143)	69.0 (40/58)	96.5 (82/85)	93.0 (40/43)	82.0 (82/100)
			k = 4		
train					
Squamous	97.9 (560/572)	95.6 (131/137)	98.6 (429/435)	95.6 (131/137)	98.6 (429/435)
NDBE	92.0 (526/572)	93.4 (240/257)	90.8 (286/315)	89.2 (240/269)	94.4 (286/303)
Neoplasia	92.0 (526/572)	83.7 (149/178)	95.7 (377/394)	89.8 (149/166)	92.9 (377/406)
test					
Squamous	96.5 (138/143)	90.9 (20/22)	97.5 (118/121)	87.0 (20/23)	98.3 (118/120)
NDBE	83.2 (119/143)	88.9 (56/63)	78.8 (63/80)	76.7 (56/73)	90.0 (63/70)
Neoplasia	82.5 (118/143)	69.0 (40/58)	91.8 (78/85)	85.1 (40/47)	81.2 (78/96)
			k = 5		
train					
Squamous	98.3 (562/572)	93.4 (128/137)	99.8 (434/435)	99.2 (128/129)	98.0 (434/443)
NDBE	93.9 (537/572)	89.1 (229/257)	97.8 (308/315)	97.0 (229/236)	91.7 (308/336)
Neoplasia	92.5 (529/572)	96.1 (171/178)	90.9 (358/394)	82.6 (171/207)	98.1 (358/365)
test					
Squamous	97.2 (139/143)	90.9 (20/22)	98.3 (119/121)	90.9 (20/22)	98.3 (119/121)
NDBE	85.3 (122/143)	85.7 (54/63)	85.0 (68/80)	81.8 (54/66)	88.3 (68/77)
Neoplasia	85.3 (122/143)	79.3 (46/58)	89.4 (76/85)	83.6 (46/55)	86.4 (76/88)
			k = 6		
train					
Squamous	99.3 (568/572)	99.3 (136/137)	99.3 (432/435)	97.8 (136/139)	99.8 (432/433)
NDBE	97.2 (556/572)	96.9 (249/257)	97.5 (307/315)	96.9 (249/257)	97.5 (307/315)
Neoplasia	96.9 (554/572)	94.4 (168/178)	98.0 (386/394)	95.5 (168/176)	97.5 (386/396)
test					
Squamous	97.9 (140/143)	95.5 (21/22)	98.3 (119/121)	91.3 (21/23)	99.2 (119/120)
NDBE	86.0 (123/143)	90.5 (57/63)	82.5 (66/80)	80.3 (57/71)	91.7 (66/72)
Neoplasia	85.3 (122/143)	74.1 (43/58)	92.9 (79/85)	87.8 (43/49)	84.0 (79/94)
			k = 7		
train					
Squamous	99.8 (571/572)	99.3 (136/137)	100.0 (435/435)	100.0 (136/136)	99.8 (435/436)
NDBE	94.8 (542/572)	96.9 (249/257)	93.0 (293/315)	91.9 (249/271)	97.3 (293/301)
Neoplasia	94.6 (541/572)	87.6 (156/178)	97.7 (385/394)	94.5 (156/165)	94.6 (385/407)
test					
Squamous	97.9 (140/143)	95.5 (21/22)	98.3 (119/121)	91.3 (21/23)	99.2 (119/120)
NDBE	79.7 (114/143)	88.9 (56/63)	72.5 (58/80)	71.8 (56/78)	89.2 (58/65)
Neoplasia	79.0 (113/143)	60.3 (35/58)	91.8 (78/85)	83.3 (35/42)	77.2 (78/101)

NDBE, non-dysplastic Barrett's esophagus.

Supplementary Table 6 | Performance of Convolutional Neural Network (CNN)-based Classification of Tissue Attenuation Spectra (for balanced dataset)

Comparison	Classification Performance				
	Accuracy % (n/total)	Sensitivity % (n/total)	Specificity % (n/total)	Positive Predictive Value % (n/total)	Negative Predicative Value % (n/total)
3-way Classification					
Squamous	96.7 (87/90)	96.9 (31/32)	96.6 (56/58)	93.9 (31/33)	98.2 (56/57)
NDBE	87.8 (79/90)	82.1 (32/39)	92.2 (47/51)	88.9 (32/36)	87.0 (47/54)
Neoplasia	84.4 (76/90)	68.4 (13/19)	88.7 (63/71)	61.9 (13/21)	91.3 (63/69)
2-way Classification[a]					
NDBE vs. Squamous	97.1 (66/68)	97.3 (36/37)	96.8 (30/31)	97.3 (36/37)	96.8 (30/31)
Neoplasia vs. Squamous	98.2 (54/55)	100.0 (18/18)	97.3 (36/37)	94.7 (18/19)	100.0 (36/36)
Neoplasia vs. NDBE	72.4 (42/58)	50.0 (10/20)	84.2 (32/38)	62.5 (10/16)	76.2 (32/42)

NDBE, non-dysplastic Barrett's esophagus.

[a] In 2-way comparisons, the class in italics is the target for purposes of classification performance metric.

Supplementary Table 7 | Performance of k-Nearest-Neighbor-(KNN)-Based Classification of Tissue Attenuation Spectra (for balanced dataset)

Comparison	Classification Performance				
	Accuracy % (n/total)	Sensitivity % (n/total)	Specificity % (n/total)	Positive Predictive Value % (n/total)	Negative Predicative Value % (n/total)
3-way Classification					
Squamous	95.6 (86/90)	90.6 (29/32)	98.3 (57/58)	96.7 (29/30)	95.0 (57/60)
NDBE	82.2 (74/90)	76.9 (30/39)	86.3 (44/51)	81.1 (30/37)	83.0 (44/53)
Neoplasia	82.2 (74/90)	68.4 (13/19)	85.9 (61/71)	56.5 (13/23)	91.0 (61/67)
2-way Classification[a]					
NDBE vs. Squamous	95.6 (65/68)	97.3 (36/37)	93.5 (29/31)	94.7 (36/38)	96.7 (29/30)
Neoplasia vs. Squamous	90.9 (50/55)	88.9 (16/18)	91.9 (34/37)	84.2 (16/19)	94.4 (34/36)
Neoplasia vs. NDBE	67.2 (39/58)	25.0 (5/20)	89.5 (34/38)	55.6 (5/9)	69.4 (34/49)

k-fold cross-validation (k=5) was used during training

NDBE, non-dysplastic Barrett's esophagus.

[a] In 2-way comparisons, the class in *italics* is the target for purposes of classification performance metric.

Supplementary Table 8 | Performance of Support-Vector-Machine-(SVM)-Based Classification of Tissue Attenuation Spectra (for balanced dataset)

Comparison	Classification Performance				
	Accuracy % (n/total)	Sensitivity % (n/total)	Specificity % (n/total)	Positive Predictive Value % (n/total)	Negative Predicative Value % (n/total)
3-way Classification					
Squamous	96.7 (87/90)	100.0 (32/32)	94.8 (55/58)	91.4 (32/35)	100.0 (55/55)
NDBE	82.2 (74/90)	76.9 (30/39)	86.3 (44/51)	81.1 (30/37)	83.0 (44/53)
Neoplasia	81.1 (73/90)	52.6 (10/19)	88.7 (63/71)	55.6 (10/18)	87.5 (63/72)
2-way Classification[a]					
NDBE vs. Squamous	98.5 (67/68)	100.0 (37/37)	96.8 (30/31)	97.4 (37/38)	100.0 (30/30)
Neoplasia vs. Squamous	92.7 (51/55)	94.4 (17/18)	91.9 (34/37)	85.0 (17/20)	97.1 (34/35)
Neoplasia vs. NDBE	69.0 (40/58)	50.0 (10/20)	78.9 (30/38)	55.6 (10/18)	75.0 (30/40)

k-fold cross-validation (k=5) was used during training

NDBE, non-dysplastic Barrett's esophagus.

[a] In 2-way comparisons, the class in *italics* is the target for purposes of classification performance metric.

References

1. Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, et al. PyTorch: An Imperative Style, High-Performance Deep Learning Library. 2019;arXiv:1912.01703.
2. Luo MR, Cui G, Rigg B. The development of the CIE 2000 colour-difference formula: CIEDE2000. Color Res Appl. 2001;26:340–50.
3. Lambert R, Kuznetsov K, Rey JF. Narrow-band imaging in digestive endoscopy. ScientificWorldJournal. 2007;7:449–65.