Electronic supplementary information (ESI)

List of Tables

S1 Mean $\bar{f}$ and standard deviation $\sigma$ of the partition quality measures computed over 100 replicates for KM, and 10 for MC, GKA and GABC. For each patient, bold values represent the smallest $\bar{f}$ among the four tested clustering methods. ........................................ 3

List of Figures

S1 Comparison of raw (a) and EMSC preprocessed (b) spectra acquired on patient #1. Standard deviations (shaded area) are shown overlaying the mean spectra (black solid lines). . 4

S2 Comparison between conventional histology and pseudo-color-coded images reconstructed from KM and MC partitions for patient #2. Scale bars indicate 100 $\mu$m. The quality measure $f$ of each partition is provided above each pseudo-color-coded image. The cluster assignments are detailed below each pseudo-color-coded image. ......................... 5

S3 Comparison between conventional histology and pseudo-color-coded images reconstructed from KM and MC partitions for patient #3. Scale bars indicate 100 $\mu$m. The quality measure $f$ of each partition is provided above each pseudo-color-coded image. The cluster assignments are detailed below each pseudo-color-coded image. ......................... 6

S4 Comparison between conventional histology and pseudo-color-coded images reconstructed from KM and MC partitions for patient #4. Scale bars indicate 100 $\mu$m. The quality measure $f$ of each partition is provided above each pseudo-color-coded image. The cluster assignments are detailed below each pseudo-color-coded image. ......................... 7

S5 Comparison between conventional histology and pseudo-color-coded images reconstructed from KM and MC partitions for patient #5. Scale bars indicate 100 $\mu$m. The quality measure $f$ of each partition is provided above each pseudo-color-coded image. The cluster assignments are detailed below each pseudo-color-coded image. ......................... 8
Comparison between conventional histology and pseudo-color-coded images reconstructed from MC, GKA and GABC partitions for patient #1. Scale bars indicate 100 µm. The quality measure $f$ of each partition is provided above each pseudo-color-coded image. The cluster assignments are detailed below each pseudo-color-coded image.
Table S1: Mean $\mathcal{F}$ and standard deviation $\sigma$ of the partition quality measures computed over 100 replicates for KM, and 10 for MC, GKA and GABC. For each patient, bold values represent the smallest $\mathcal{F}$ among the four tested clustering methods.
Figure S1: Comparison of raw (a) and EMSC preprocessed (b) spectra acquired on patient #1. Standard deviations (shaded area) are shown overlaying the mean spectra (black solid lines).
<table>
<thead>
<tr>
<th>HE-stained image</th>
<th>Best KM partition $(f = 4.0043)$</th>
<th>Most frequent KM partition $(f = 4.1532)$</th>
<th>Worst KM partition $(f = 4.2897)$</th>
<th>MC partition $(f = 3.9729)$</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="HE-stained image" /></td>
<td><img src="image2.png" alt="Best KM partition" /></td>
<td><img src="image3.png" alt="Most frequent KM partition" /></td>
<td><img src="image4.png" alt="Worst KM partition" /></td>
<td><img src="image5.png" alt="MC partition" /></td>
</tr>
</tbody>
</table>

(1) Crypts  
(2) Lamina propria  
(3) Muscularis mucosae  
(4) Submucosa  
(5) Lymphoid aggregate

Figure S2: Comparison between conventional histology and pseudo-color-coded images reconstructed from KM and MC partitions for patient #2. Scale bars indicate 100 µm.

The quality measure $f$ of each partition is provided above each pseudo-color-coded image. The cluster assignments are detailed below each pseudo-color-coded image.
The quality measure $f$ of each partition is provided above each pseudo-color-coded image. The cluster assignments are detailed below each pseudo-color-coded image.

Figure S3: Comparison between conventional histology and pseudo-color-coded images reconstructed from KM and MC partitions for patient #3. Scale bars indicate 100 µm.
HE-stained image & Best KM partition \((f = 15.3063)\) & Most frequent KM partition \((f = 15.8329)\) & Worst KM partition \((f = 15.8945)\) & MC partition \((f = 15.2046)\) \\
| (1) Crypts | | | |
| (2) Lamina propria | | | |
| (3) Muscularis mucosae | | | |
| (4) Submucosa | | | |

Figure S4: Comparison between conventional histology and pseudo-color-coded images reconstructed from KM and MC partitions for patient #4. Scale bars indicate 100 \(\mu m\).

The quality measure \(f\) of each partition is provided above each pseudo-color-coded image. The cluster assignments are detailed below each pseudo-color-coded image.
<table>
<thead>
<tr>
<th>HE-stained image</th>
<th>Best KM partition ($f = 5.3728$)</th>
<th>Most frequent KM partition ($f = 5.4735$)</th>
<th>Worst KM partition ($f = 5.8176$)</th>
<th>MC partition ($f = 5.3720$)</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="HE-stained image" /></td>
<td><img src="image2.png" alt="Best KM" /></td>
<td><img src="image3.png" alt="Most frequent KM" /></td>
<td><img src="image4.png" alt="Worst KM" /></td>
<td><img src="image5.png" alt="MC" /></td>
</tr>
</tbody>
</table>

(1) Crypts  
(2) Lamina propria  
(3) Muscularis mucosae  
(4) Submucosa

Figure S5: Comparison between conventional histology and pseudo-color-coded images reconstructed from KM and MC partitions for patient #5. Scale bars indicate 100 µm.

The quality measure $f$ of each partition is provided above each pseudo-color-coded image. The cluster assignments are detailed below each pseudo-color-coded image.
Figure S6: Comparison between conventional histology and pseudo-color-coded images reconstructed from MC, GKA and GABC partitions for patient #1. Scale bars indicate 100 µm. The quality measure $f$ of each partition is provided above each pseudo-color-coded image. The cluster assignments are detailed below each pseudo-color-coded image.