Supplementary Material: Additional functions of the anomaly detection method.

Weakly Supervised Anomaly Detection Coupled with Fourier Transform Infrared (FT-IR) Spectroscopy for the Identification of Non-normal Tissue

Dougal Ferguson*[ab], Alex Henderson[ab], Elizabeth F. McInnes [c], and Peter Gardner [ab]

The aim of the anomaly detection paper is to highlight the ability of the technique to identify non-normal tissue components with biological relevance, such as non-neoplastic lesions or alternative tissue components. However, the method is also able to identify artefacts and interferents that are typically identified and removed at the quality control stage. The anomaly detection method can be seen to identify dust or hair particles that have made their way onto the sample, as shown in Figure 1 and Figure 2, alongside damaged tissue shown in Figure 3. Additionally, tissue edges that can be considered mixed pixels (with contribution from the paraffin wax captured in the pixel region) are also captured as shown in Figure 4 and Figure 5. While not directly relevant to the aim of the publication, these results still serve to highlight the possible application of this technique.

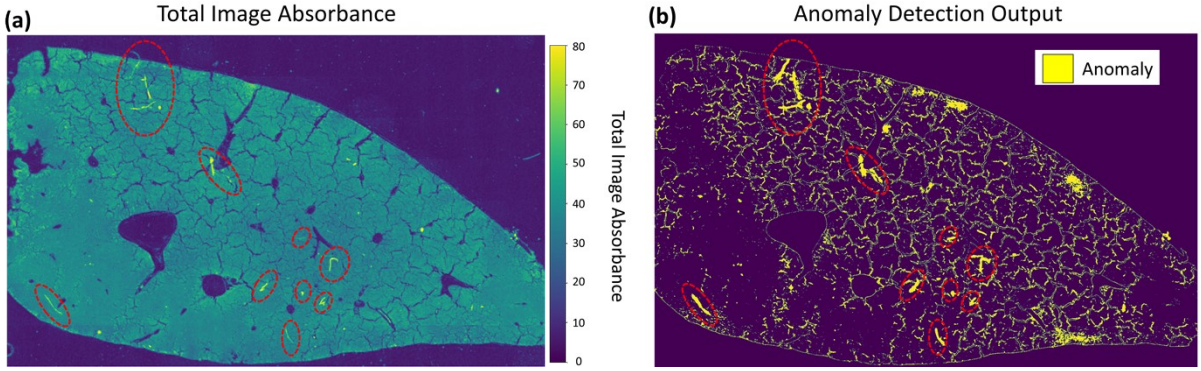

Figure 1 - Identification of interferents (dotted red ellipses) using the anomaly detection method. Suspected dust and/or hair are spotted in the total absorbance of the fingerprint region (a) correspond to those identified by the anomaly detection output (b).

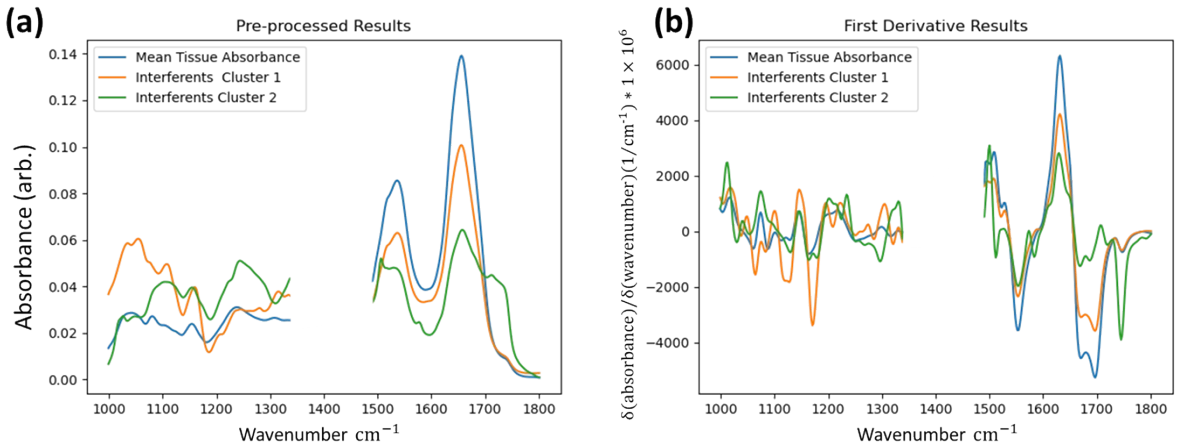

Figure 2 - Confirmation the anomaly detection method captures interferents, plotting the mean absorbance profile of pre-processed (a) and first derivative (b) spectral groups for an entire tissue section (blue) and k-means clusters of identified interferents (orange and green).

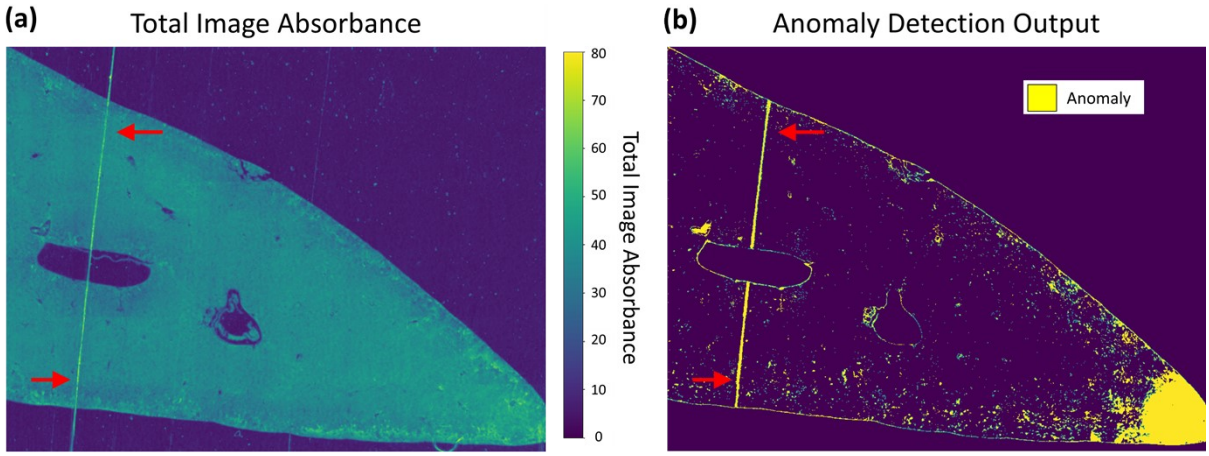

Figure 3 - Identification of damaged tissue (red arrows) as a result of tissue sectioning using the anomaly detection method. Suspected damaged tissue is spotted in the total absorbance of the fingerprint region (a) and is found in the same position as the anomaly detection output (b).

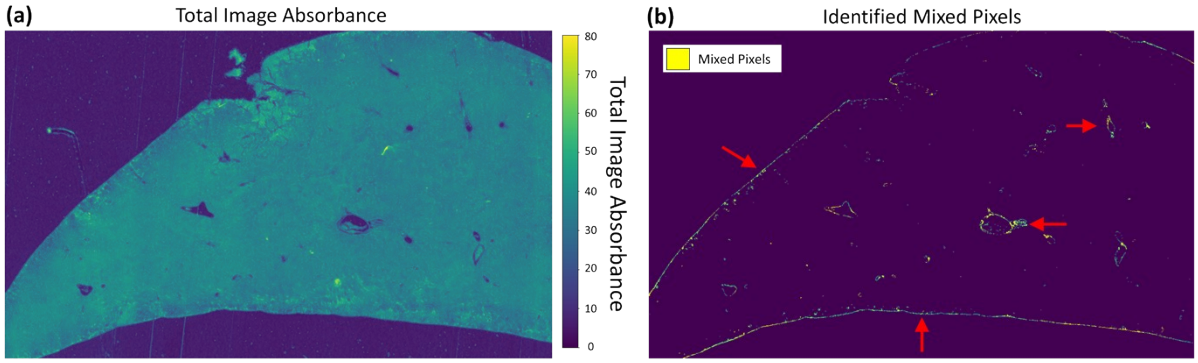

Figure 4 – Identification of mixed pixels (red arrows) using the anomaly detection method. Pixels at the tissue edge (b) can be verified through comparisons of the total absorbance of the fingerprint region (a) and the edge pixels identified using the anomaly detection algorithm (b).

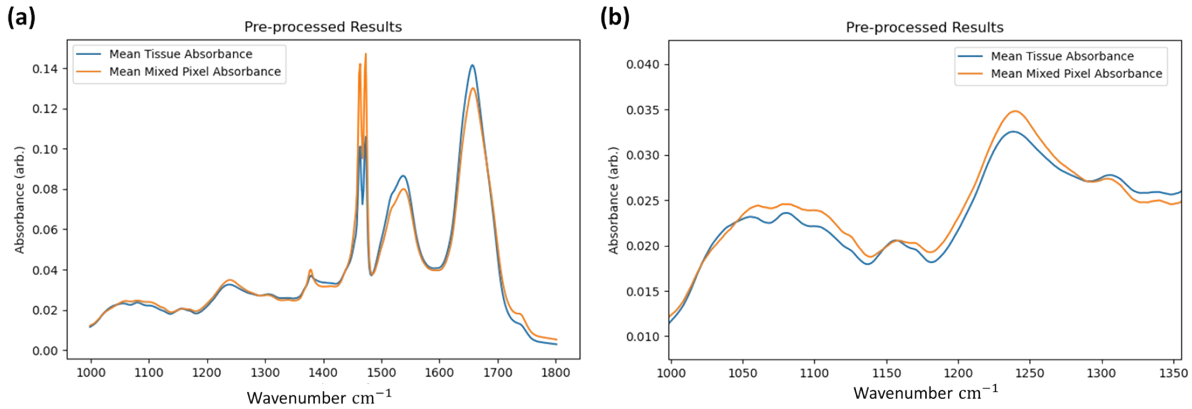

Figure 5 - Confirmation the anomaly detection method captures mixed pixels, plotting the mean absorbance profile of pre-processed spectral groups across the full fingerprint region (a) and lower wavenumber region (b) show the increased paraffin contribution in mixed pixels (orange) compared to the mean tissue spectra (blue). Differences in lower wavenumber region indicate the overall absorbance profile with the paraffin peaks removed are still distinctly different enough for discrimination.

Additionally, it is deemed important to also show the method's ability to discriminate between key biologically relevant cell subgroups that have subtle spectral differences. This is done to highlight the application of the method beyond capturing greatly different spectral groups that may simply be removed during correct spectra pre-processing steps. The term "normal" when training the model consisted of hepatocytes of control group mouse livers, which contain different subgroups within them that exist during the normal liver function. Three different subgroups of "normal" group cells are plotted in Figure 6. It is clear that there are differences within the training group that may not be clear when plotting the averages of the groups, as hypertrophy (a normal

liver cell function, however a process that predates cell death) is included within this “normal” group definition.

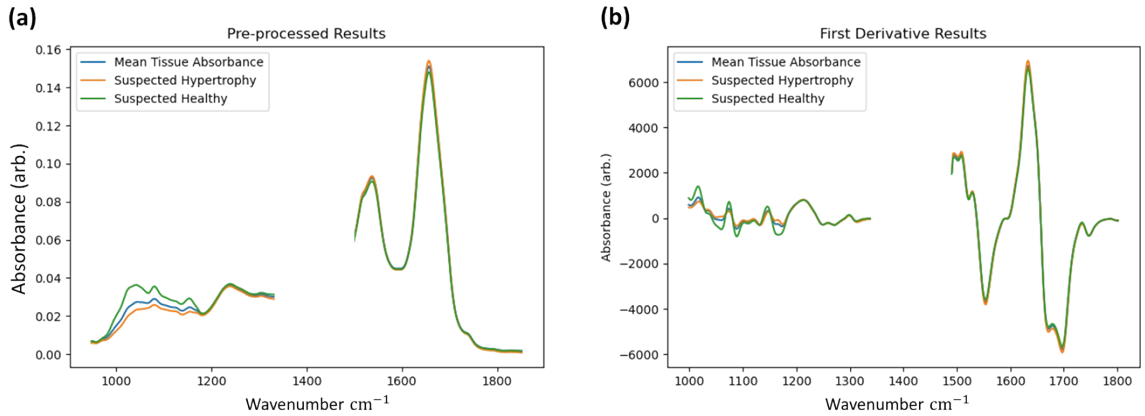

Figure 6 – Key spectral subgroups that can be found in the normal group of liver tissue spectra used in the training of the anomaly detection model. The means of pre-processed (a) and first derivative (b) spectral groups across the fingerprint region are plotted excluding the paraffin wax peaks.

Within the results section of this paper it is shown that necrotic cells (a non-neoplastic lesion that does not occur in normal liver function) are correctly identified by the method. While necrotic cells are clearly different from the tissue average (Figure 6 (f) in the main text), spectrally the necrotic subgroup are not so different from the hypertrophic cells identified in Figure 6, as hypertrophy will occur before necrosis. This serves to show that the anomaly detection method is able to discriminate between subtly different spectral groups, as while similar to necrotic cells spectrally, hypertrophy is still considered a normal subgroup.

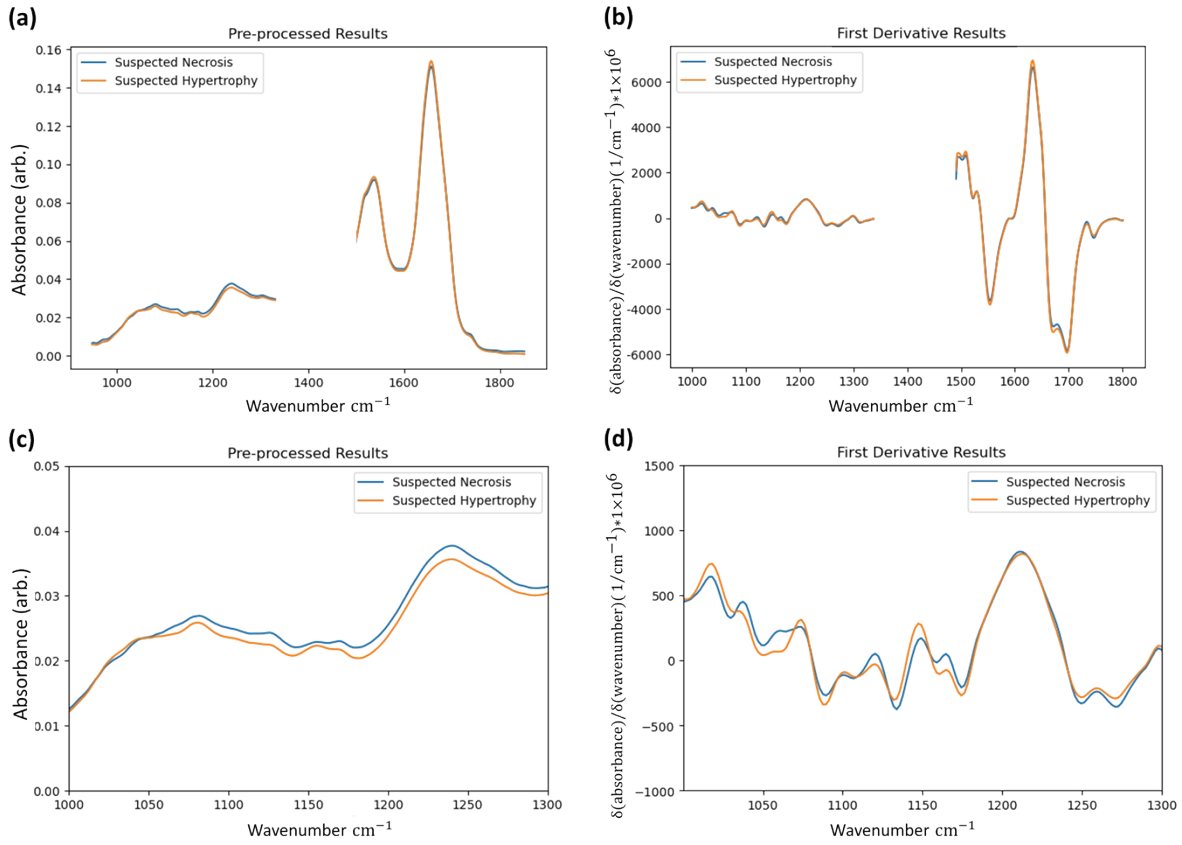

Figure 7 – Confirmation the anomaly detection method discriminating between subtly different spectral groups. The mean absorbance profiles of suspected hypertrophy (orange) and necrosis (blue) are plotted as pre-processed (a, c) and first derivative (b, d) profiles in the full fingerprint region (a, b) and lower wavenumber region (c, d).