Fig. S1: Histograms showing the distribution for each tissue type at the top ranked metric for OSCC (i.e. 1562 cm^{-1}/1539 cm^{-1}; Table 2).
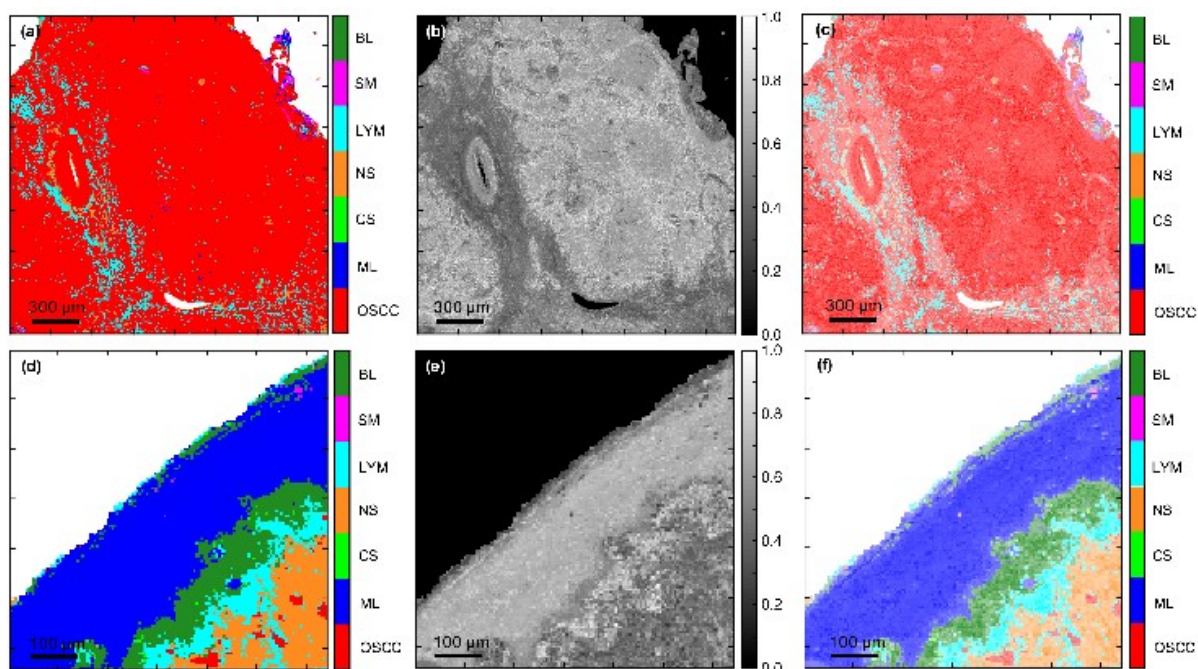

Fig. S2: Pseudo-colouring of FTIR-HS images of oral cancer (top) and oral epithelium (bottom) (a) and (d): pseudo-colour images labelled by the MLA; (b) and (e): confidence level of the MLA in labelling each pixel and (c) and (f) a combination image with the saturation of each pixel corresponding to the confidence of the MLA in identifying the tissue type. OSCC: oral squamous cell carcinoma; CS: tumour stroma with immune/inflammatory reaction; BL: progenitor layers of oral epithelium; ML: maturation layers of oral epithelium; NS: pre-existing stroma; SM: submucosal components; LYM: lymphoid node tissue.