High-throughput molecular imaging via deep learning enabled Raman spectroscopy

Conor C. Horgan\textsuperscript{1,5}, Magnus Jensen\textsuperscript{1}, Anika Nagelkerke\textsuperscript{2}, Jean-Philippe St-Pierre\textsuperscript{3}, Tom Vercauteren\textsuperscript{4}, Molly M. Stevens\textsuperscript{5*}, Mads S. Bergholt\textsuperscript{1*}

\textsuperscript{1}Centre for Craniofacial and Regenerative Biology, King’s College London, London SE1 9RT, UK
\textsuperscript{2}University of Groningen, Groningen Research Institute of Pharmacy, Pharmaceutical Analysis, P.O. Box 196, XB20, 9700 AD Groningen, The Netherlands
\textsuperscript{3}Department of Chemical and Biological Engineering, University of Ottawa, Ottawa, ON K1N 6N5, Canada
\textsuperscript{4}School of Biomedical Engineering and Imaging Sciences, King’s College London, WC2R 2LS, UK
\textsuperscript{5}Department of Materials, Department of Bioengineering, and Institute of Biomedical Engineering, Imperial College London, London SW7 2AZ, UK
*Corresponding authors: mads.bergholt@kcl.ac.uk, m.stevens@imperial.ac.uk

Supporting Information

The supporting information contains additional materials and methods pertaining to the preparation of cell and tissue culture samples, Raman spectroscopic imaging setup and processing, details of the calculations of image quality metrics, neural network implementation details, and supplementary figures.
Cell Culture
MDA-MB-231 breast cancer cells were originally obtained from the ATCC (Manassas, VA, USA) and authenticated via STR profiling. Cells were maintained in Dulbecco’s modified Eagle’s medium (DMEM), supplemented with 1x non-essential amino acids, 25 mM Heps, 1x penicillin/streptomycin, and 10% (v/v) foetal bovine serum (FBS), all obtained through Gibco (Thermo Fisher Scientific, Inc., Waltham, MA, USA). Cells were grown on magnesium fluoride (MgF\textsubscript{2}) windows (Global Optics Ltd, Bournemouth, UK). 20,000 cells per cm\textsuperscript{2} were seeded in serum-supplemented DMEM and allowed to adhere overnight. Prior to imaging, cells were washed with DPBS and fixed with 4% (v/v) paraformaldehyde (PFA, Sigma-Aldrich, St Louis, MO, USA) in DPBS for 30 minutes at room temperature. PFA solution was aspirated and samples washed three times with DPBS.

Tissue Culture and Sample Preparation
Articular cartilage was excised aseptically from the full thickness of metacarpal–phalangeal joints of mature cows (n = 3) aged 24 to 36 months within 48 h of death. Chondrocytes were isolated from the tissue by sequential enzymatic digestion at 37 °C (0.2% (w/v) Pronase (Roche Applied Science) in DMEM with 4.5 g/L glucose (Invitrogen); for 1 h followed by 0.04% (w/v) collagenase type I (Sigma-Aldrich) in DMEM overnight. Hydrophilic polytetrafluoroethylene (PTFE) membranes (Millipore) encased in a cell chamber were incubated with 100 µL of 0.5 mg/mL collagen type II (Sigma-Aldrich) in 0.1 N acetic acid overnight to allow full evaporation of the solution and then washed 3 times in PBS. Cell chambers were placed in 24 well plates, and 750 µL of DMEM supplemented with 5% (v/v) FBS (HyClone) was added to each well outside the cell chamber. Isolated chondrocytes were seeded on top of PTFE membrane inserts (1 × 10\textsuperscript{6} cells in 750 µL per membrane; 12 mm diameter) in either 2 mL of DMEM with 1 g/L glucose, 11 mL of DMEM with 1 g/L glucose, or 11 mL of DMEM with 4.5 g/L glucose; supplemented with 5% (v/v) FBS and incubated at 37 °C and 5% CO\textsubscript{2}. On day 2, the medium was supplemented with ascorbic acid (50 µg/mL; Sigma-Aldrich) and the FBS supplementation was increased to 10% (v/v). Culture medium was changed every 2–3 days and cultures harvested at 14, 28, or 42 days. Tissue-engineered cartilage constructs were washed 3 times for 15 min in PBS, and fixed in 4% (v/v) paraformaldehyde for 30 min. Tissue-engineered cartilage constructs were cemented onto a polystyrene surface using a small quantity of cyanoacrylate. These were then immersed in a drop of water and frozen on a cryosectioning block. Using a cryostat (Bright Instruments Ltd.), the tissues were sectioned flat at a 45° angle relative to the articular surface. Samples were stored at 4 °C in PBS until Raman spectroscopic imaging was performed. Tissue cross-section samples imaged by Raman spectroscopy were prepared to a thickness of greater than 1 mm.

Raman Spectroscopic Imaging
Raman imaging was performed using a confocal Raman microscope (alpha300R+, WITec, GmbH, Germany). A 532 nm laser light source at 35 mW power output was applied through a 63x/1.0 NA water-immersion microscope objective lens (W Plan-Apochromat, Zeiss, Germany). Inelastically-scattered light was collected through the objective and directed via a 100 µm diameter silica fibre, acting as a confocal pinhole, to a high-throughput imaging spectrograph (UHTS 300, WITec, GmbH, Germany) with a 600 groove/mm grating and equipped with a thermoelectrically cooled (-60 °C) back-illuminated charge-coupled device (CCD) camera. Hyperspectral Raman images were acquired with a 0.5 µm spatial resolution (1 µm spatial resolution for tissue-engineered cartilage) and a 1 second integration time, where for each pixel a Raman spectrum was acquired in the range from 0 to 3700 cm\textsuperscript{-1} with a spectral resolution of 11 cm\textsuperscript{-1}.

Hyperspectral Raman Image Processing
Preliminary hyperspectral Raman image processing was kept constant for all datasets and performed using Witec ProjectFOUR software. Spectra were cropped to the fingerprint spectral region (600 – 1800 cm\textsuperscript{-1}) and background autofluorescence subtraction performed using ProjectFOUR’s ‘shape’ background filter with α = 500. Hyperspectral images were then exported to MATLAB. Where applicable, Savitzky-Golay filtering (with varying order and window length) was performed in Python using SciPy’s savgol_filter function, PCA denoising was performed in Python using scikit-learn’s PCA function, and wavelet denoising was performed in MATLAB using the wdenoise function in MATLAB’s Wavelet Toolbox. For PCA and wavelet denoising, parameterisation was empirically determined, with selected parameters minimising the mean squared error between the denoised spectra and the target high SNR spectra.
Image Quality Metrics
The PSNR of an image, $y$, relative to an uncorrupted ground truth image, $x$, is defined as:

$$PSNR(x,y) = 10 \cdot \log_{10} \left( \frac{x_{\text{max}}^2}{MSE} \right)$$

Where $x_{\text{max}}$ is the maximum possible pixel value (across all channels) of the ground truth image, $x$, and the mean squared error (MSE), calculated across all channels, is defined as:

$$MSE = \frac{1}{m \cdot n \cdot p} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \sum_{k=0}^{p-1} [x(i,j,k) - y(i,j,k)]^2$$

The SSIM between two images, $x$ and $y$, is defined as

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

Where $\mu_x$ and $\mu_y$ are the means of $x$ and $y$, $\sigma_x^2$ and $\sigma_y^2$ are the variances of $x$ and $y$, and $c_1$ and $c_2$ are two numerical constants to stabilize the division. For hyperspectral Raman images, the SSIM is calculated independently for each channel and then averaged.

Implementation
Complete implementation details including dataset size, training/validation/testing splits, hyperparameter selection, and training times are shown in Supplementary Tables 2 and 3.

Network training and optimisation were performed using the Joint Academic Data science Endeavour (JADE) high performance computing (HPC) facility on an NVIDIA DGX-1 deep learning system with 8 NVIDIA V100 GPUs.

Final network implementation and training was performed in Python 3.7.3 using PyTorch 1.4.0 on a desktop computer with a Core i7-8700 CPU at 3.2 GHz (Intel), 32 GB of RAM, and a Titan V GPU (NVIDIA), running Windows 10 (Microsoft).

Supplementary Figure 1 | Hyperspectral Residual Channel Attention Network. Hyperspectral residual channel attention network architecture employed for super-resolution of hyperspectral Raman images. The network consists of multiple repeated sub-units including the residual groups (purple), residual channel attention blocks (orange), and the channel attention blocks (green). In the final network architecture, $n$(residual groups) = 18 and $n$(residual channel attention blocks) = 16. Spectral downsampling is used at the beginning of the network, exploiting the redundancy of information contained in the high spectral resolution Raman spectra, to reduce computational load, with spectral upsampling placed at the end of the network to return a hyperspectral image with the sample spectral resolution and length as the inputs.
Supplementary Figure 2 | VCA Endmembers for HR, High SNR Hyperspectral Raman Cell Image for Raman Spectral Denoising. (a-e) VCA endmembers for high SNR hyperspectral Raman cell image (Figure 2d) corresponding to (a) background (black), (b) nucleic acids (blue), (c) proteins (green), (d) lipids (yellow), and (e) proteins (green).
Supplementary Figure 3 | Data Augmentation for Training of Residual Channel Attention Network for Hyperspectral Image Super-Resolution. Data augmentations randomly applied to training set of hyperspectral Raman image dataset to increase effective dataset size for effective neural network training. Image data augmentations include subsampling (selecting different sub-regions of an image), horizontal and vertical flipping, rotations, and mixup (the addition of two images following a beta distribution). Additional spectral augmentations include spectral shifting and flipping.
Supplementary Figure 4 | Deep learning enabled hyperspectral image super-resolution. 2×, 3×, and 4× super-resolution of example test set hyperspectral Raman image enables a significant reduction in imaging times (shown in white) while recovering important spatial and spectral information (scale bars = 10 µm). Images shown are the result of a VCA performed on the target HR hyperspectral Raman image, which identified 5 key components (proteins [green], nucleic acids [blue], lipid signature 1 [yellow], lipid signature 2 [red], background [black]). VCA components were applied to the nearest neighbour output, bicubic output, and neural network output images via non-negatively constrained least-squares regression.
Supplementary Figure 5 | Deep learning enabled hyperspectral image super-resolution. 2×, 3×, and 4× super-resolution of example test set hyperspectral Raman image enables a significant reduction in imaging times (shown in white) while recovering important spatial and spectral information (scale bars = 10 µm). Images shown are the result of a VCA performed on the target HR hyperspectral Raman image, which identified 4 key components (nucleic acids [blue], lipids [yellow], lipid droplets [magenta], proteins [green], background [black]). VCA components were applied to the nearest neighbour output, bicubic output, and neural network output images via non-negatively constrained least-squares regression.
Supplementary Figure 6 | VCA Endmembers for HR, High SNR Hyperspectral Raman Cell Image for Hyperspectral Image Super-Resolution. (a-e) VCA endmembers for high SNR hyperspectral Raman cell image (Figure 3) corresponding to (a) background (black), (b) nucleic acids (blue), (c) proteins (green), and (d) lipids (yellow).

Supplementary Table 1 | Performance Comparison for Hyperspectral Image Super-Resolution. ↑ indicates higher values are better, ↓ indicates lower values are better, bold text indicates best performance.

|               | 2x                | 3x                | 4x                |
|---------------|-------------------|-------------------|-------------------|
|               | PSNR ↑ | SSIM ↑ | MSE ↓ | PSNR ↑ | SSIM ↑ | MSE ↓ | PSNR ↑ | SSIM ↑ | MSE ↓ |
| Nearest Neighbours | 39.90  | 0.6345 | 0.000129 | 38.13  | 0.6261 | 0.000201 | 36.57  | 0.5258 | 0.000309 |
| Bicubic       | 40.39  | 0.6419 | 0.000114 | 39.28  | 0.6308 | 0.000150 | 37.59  | 0.5491 | 0.000239 |
| HyRISR        | **41.99** | **0.8448** | **0.000078** | **40.61** | **0.8424** | **0.000110** | **39.27** | **0.8281** | **0.000159** |
Supplementary Figure 7 | VCA Endmembers for HR, High SNR Hyperspectral Raman Cell Image for Combined Raman Spectral Denoising and Hyperspectral Image Super-Resolution. (a-d) VCA endmembers for high SNR hyperspectral Raman cell image (Figure 4) corresponding to (a) nucleic acids (blue), (b) background (black), (c) proteins (green), and (d) lipids (yellow).
Supplementary Figure 8 | VCA Endmembers for HR, High SNR Hyperspectral Raman Image of Tissue-Engineered Cartilage for Transfer Learning. (a-e) VCA endmembers for high SNR hyperspectral Raman cell image (Figure 5) corresponding to (a) substrate (blue), (b) background (black), (c) dense extracellular matrix region (green), (d) cells (red), and (e) sparse extracellular matrix region (yellow).
Supplementary Table 2 | Hyperparameter Details for Training of 1D Residual UNet for Raman Spectral Denoising.

| Hyperparameter                  | Details                                                                 |
|---------------------------------|-------------------------------------------------------------------------|
| Dataset size                    | 172,312 Raman spectra from 11 hyperspectral Raman images                |
| Training/validation/testing split| 11-fold leave one image out cross-validation                             |
|                                 | Training/validation: Raman spectra from 10 images (90:10 split)          |
|                                 | Testing: Raman spectra from 1 remaining image                            |
| Optimizer                       | Adam                                                                    |
| Maximum learning rate           | 5x10^{-4}                                                               |
| Scheduler                       | One Cycle LR                                                            |
| Epochs                          | 500                                                                     |
| Batch size                      | 256                                                                     |
| Batch norm                      | Yes                                                                     |
| Dropout                         | None                                                                    |
| Training Time                   | ~26 hours on 1x Titan V GPU                                             |

Supplementary Table 3 | Hyperparameter Details for Training of Residual Channel Attention Network for Hyperspectral Image Super-Resolution.

| Hyperparameter                  | Details                                                                 |
|---------------------------------|-------------------------------------------------------------------------|
| Dataset size                    | 169 hyperspectral Raman images                                         |
| Training/validation/testing split| 85:10:5                                                                 |
| No. Residual groups             | 18                                                                      |
| No. Residual blocks             | 16                                                                      |
| Optimizer                       | Adam                                                                    |
| Maximum learning rate           | 1x10^{-5}                                                               |
| Scheduler                       | Constant LR                                                            |
| Epochs                          | 600                                                                     |
| Batch size                      | 2                                                                       |
| Batch norm                      | No                                                                      |
| Dropout                         | None                                                                    |
| Training Time                   | ~18 hours on 1x Titan V GPU                                            |