Maynooth University
National University of Ireland Maynooth

Deep Learning & Hyperspectral Imaging with Raman Spectroscopy for Histology

Tim McNamara

A thesis presented in partial fulfilment of the requirements for the degree of Master of Engineering by Research

Supervisor: Dr. Bryan Hennelly

Head of department: Prof. Gerard Lacey

Electronic Engineering

National University of Ireland – Maynooth

Maynooth, Ireland

October, 2025

Table of Contents

Acknowledgements

I would like to take this space to thank all those who supported me through this research.

Special thanks to my supervisor Dr Bryan Hennelly for his invaluable guidance, expertise, training and encouragement.

My parents and sisters, for their continued and unwavering support. To Trisha and Andy for their constant encouragement.

Finally to all my friends who supported me. This achievement would not have been possible without you.

1 Chapter 1: Introduction

1.1 Aims & Objectives

The overarching aim is to develop tools for quantitative, chemically-specific histopathology by (i) developing an integrated hyperspectral imaging platform and (ii) developing deep learning methods for the physical characterisation of biological spectra. Specifically, we define the following objectives:

1. To perform a comprehensive review of hyperspectral imaging using Raman spectroscopy and its many variants in the domain of histopathology alongside the application of advanced deep learning techniques for both style transfer and automated diagnostics. This review was published in a recent book chapter 1 and is the basis for Chapter 2.

2. To compare the spectra obtained from both Spontaneous Raman and BCARS for biological samples, specifically white blood cells (Cal-1 cell line), in order to validate the chemical consistency and identify the strengths (Raman: high quality spectra; BCARS: high speed spectra) and weaknesses (Raman: slow acquisition; BCARS: lesser quality spectra) of both techniques for single cell analysis. This comparison is detailed in Chapter 3 and published in Ref. 2.

3. To build and characterise a dual-modality system that combines the benefits of both spontaneous Raman and BCARS. This involved optical engineering an optoelectronic system such that flip mirrors could be used to easily switch between the two modalities. Spectra from a range of samples were examined for both modalities. This work is detailed in Chapter 4 and has been published in part in Ref 3, with a journal paper in preparation.

4. To develop a novel deep learning methodology using convolutional autoencoders for the automated decomposition of complex spectra into their constituent Lorentzian parameters (position, width, amplitude) to enable more precise quantitative analysis. This work is detailed in Chapter 5 and has been

published in part in Ref 4, with a journal paper in preparation. We acknowledge that the deep-learning method developed in this Chapter was applied to simulated spectra only. It will be the work of a follow-up PhD project to further investigate deep learning methods of this type and to test them on experimental data.

1.2 A brief history of Raman spectroscopy

Raman Spectroscopy is an analytical technique that probes molecular vibrations by detecting the inelastic scattering of light, a phenomenon where photons exchange energy with a molecule, causing a shift in the light's wavelength. This effect was first theoretically predicted by Adolf Smekal in 1923[5] and subsequently confirmed experimentally in 1928. Sir C.V. Raman, using filtered sunlight as a light source and a telescope, observed that light scattered from a liquid sample contained new frequencies, visible as a faint, color-shifted hue[6]. Simultaneously and independently Grigory Landsberg and Leonid Mandelstam observed a similar phenomenon while studying light scattering in quartz crystals[7]. Using a mercury arc lamp and a spectrograph, they detected new, faint spectral lines shifted from the original, demonstrating energy exchange between light and the crystal's vibrations. These complementary discoveries provided definitive proof of the Raman effect and established the foundation for the techniques discussed in this thesis.

Three decades later the theoretical underpinings for the laser were proposed by Townes and Schawlow in 1958[8] and brought to experimental fruition when Maiman invented the first laser in 1960[9]. These lasers provided a coherent and monochromatic light source that dramatically improved the signal-to-noise ratios and made it possible to spectrally filter out the elastic (Rayleigh) scattering thus recovering only the weak, frequency shifted Raman photons. Despite these improvements the scattered signal was still weak, requiring long acquisition times as they were dependent on random transitions to spontaneously occur hence the name "spontanoeus Raman spectroscopy". By 1962 E.J. Woodbury and W.K. Ng accidentally discovered Stim-

ulated Raman (SRS), a third order non linear process. As they were conducting experiments with a 694.3nm ruby laser utilizing nitrobenzene as a Q-switch, they noticed a new beam of light at a longer than expected wavelength[10].

In 1965 P.D Maker and R.W Terhune provided the first demonstration of a new third order non linear optical effect by mixing two light beams at different frequencies in a sample. This process generated a new coherent beam at a higher frequency that was resonantly enhanced by molecular vibrations[11] and was later named coherent anti Stokes Raman scattering (CARS). This experimental work was followed in 1977 by a comprehensive review from Tolles et al., which established the fundamental theory of the CARS process and described the mechanism of the non-resonant background[12]. While CARS was initially used for gas analysis in 1973 by Regnier and Taran[13] and subsequently in the study of combustion in 1975[14] it rapidly found prominence in microscopy, a notable demonstration of which occurred in 1999 when CARS was utilized to record cells in 3d[15]. A key disadvantage of the initial work in CARS with respect to Spontaneous Raman was the limitation of the method to stimulating an individual resonance at a time. This led researchers in the early 2000s to utilize "broadband coherent anti Stokes Raman spectroscopy" [16] whereby a spectrally broad Stokes beam is used to simultaneously generate resonances across a range of resonant frequencies. This marked the beginning of a rapid acquisition, large spectral width recording mechanism for micro spectroscopy.

Of particular relevance to this thesis, is the application of Raman and CARS in the area of histopathology, which has become an exciting area of research in recent years and has been demonstrated to facilitate real-time surgical guidance; we will review the state of the art of spontaneous Raman, BCARS and SRS in histopathology today in Chapter 2. In this context, we proceed to contrast the relative strengths and limitations of the two prominent single acquisition full-spectrum methods in Chapter 3, specifically spontaneous Raman and BCARS. Following this, we introduce a novel system in Chapter 4 built as part of this project which combines the rapid scanning ability of BCARS with the high quality of spontaneous Raman into a single optical

system.

1.3 A brief history of deep learning

Machine learning is a field within computer science where algorithms are developed to improve as they are exposed to data, rather than being explicitly programmed. Within this field, the "Deep learning" approach has become particularly prominent. Deep learning is inspired by the structure of the human brain, using interconnected layers of artificial neurons to form a neural network. The "deep" refers to the typically large number of layers found in these networks.

Deep learnings origins can be traced as far back as 1943 when McCulloch and Pitts introduced the McCulloch Pitts neuron[17]. A combination of algorithms and mathematics designed to utilize threshold logic inspired by human neurons. The ability to tie a series of these neurons together and learn simultaneously was made possible over several decades with the invention of back propogation, a prerequisite for which were Henry Kellers gradient calculation methods introduced to control theory for dynamic systems in 1960[18]. In 1970 Seppo Linnainmaa developed the reverse mode of automatic differentiation[19], providing an efficient algorithm for computing gradients in layered computational graphs. Finally in 1982 Paul Werbos applied this reverse differential method to training an artificial neural network, the method now dubbed "back propogation[20]".

The 1986 paper by Rumelhart, Hinton, and Williams popularised the backpropagation algorithm for training multi-layer neural networks and also introduced the concept of recurrent neural networks by describing how back propagation could be extended to networks with temporal feedback connections[21]. Building on this foundation, the 1989 paper by Williams and Zipser formalized the backpropagation through time algorithm, enabling more effective training of fully recurrent networks[22]. In 1990, Elman introduced a simplified recurrent architecture, now known as the Elman network, which demonstrated the ability of recurrent neural networks to learn temporal structures in sequential data such as natural language[23].

The first practical successful convolutional neural network was developed in 1998 by Yann Lecun for handwriting recognition[24] which utilized supervised learning end-end and became the foundational model upon which subsequent systems were built. At the same time, Recurrent Neural Networks were being developed to handle sequential data by maintaining hidden states that capture information across time steps, making them well suited to tasks like speech recognition and language modelling. However standard Recurrent Neural Networks struggled with long term dependencies due to vanishing and exploding gradients. Long Short Term Memory networks, introduced by Hochreiter and Schmidhuber in 1997, addressed this by incorporating gating mechanisms that allow the network to retain or forget information over long sequences, enabling more effective learning from temporal data[25].

Although much of the foundational methods for deep learning were established by the early 1980s, the field did not gain widespread traction until a decade into the 2000s. This delay was due in large part due to the lack of sufficient computational power and large scale datasets, both of which are necessary for training deep neural networks effectively. With the introduction of GPUs, which leveraged emerging parallel processing algorithms, alongside the development of large scale datasets following the global adoption of the internet, training deep learning networks became practically viable. Simultaneously the development of computationally efficient activation functions including the Rectified linear unit, enabled more efficient and stable training[26] and in some part addressed the issue of the vanishing gradients found with previous activation functions. The groundbreaking AlexNET was produced by Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton in 2012[27], which revolutionized the field of computer vision and was shortly followed by massive breakthroughs in natural language processing with the introduction of the attention mechanism[28] by Dzmitry Bahdanau and the transformer architecture in 2017 developed by Ashish Vaswani et al[29].

In recent years deep learning research has exploded into every discipline with advanced network architectures being applied to a wide body of problems including

recognition[30], classification[30], segmentation[31], transformation[32] and generation[33] for a wide range of signals and complex data types. With innovations in convolutional networks such as U-Net[34] and ResNET[35] for feature extraction, enabling both segmentation and classification which in turn have allowed these networks to be tailored to a wide range of problems. These convolutional networks have also been utilized in generative architectures such as GANs and diffusion models[36] for synthesis and style transfer[33]. Contemporaneously the self attention and the transformer architectures[29] have become the industry standard for natural language processing and text generation. Recent advances in the transformer architectures have enabled their application not only to discrete tokens but to continuos data[37 38]. This has enabled them to move into the computer vision space previously dominated by the convolutional approach[39]. We will review applications of state of the art deep learning methods in Chapter 2 for both virtual staining and automated diagnostics with hyperspectral datasets based on Raman spectral datasets. In Chapter 5 we present a network developed in this project for enhancing the physical characterisation of Raman spectra by decomposing them into values directly representative of the individual resonances recorded. We believe that this approach could be used to enhance the characterisation and classification of biological spectra, although we accept that this has not been proven within the scope of this Research Masters thesis.

2 Hyperspectral Imaging & Histology with Raman Spectroscopy: A Comprehensive Review

The work presented in this chapter is based on the textbook chapter titled "Hyperspectral imaging & histology with Raman Spectroscopy", which is being published by the Institute Of Physics in the upcoming book titled "Spectroscopy for Cancer"[1]

2.1 Introduction

Raman spectroscopy enables detailed visualisation of molecular compositions, providing a 'fingerprint' by which materials can be identified, a capability that is of growing interest in medical diagnostics. Hyperspectral imaging and histology with Raman spectroscopy represent important recent advancements in the medical and scientific communities, offering an innovative approach to the molecular analysis of tissues. The Raman spectrum at each point in the image can be recorded with micrometre resolution using a variety of different modalities, producing a three dimensional dataset, where the third dimension is the Raman wavenumber. Raman spectroscopy is non-destructive and label-free, thereby allowing for the recording of hyperspectral information without altering the samples properties. This chapter introduces the various techniques and methodologies involved in hyperspectral Raman imaging, focusing particularly on their application in histological studies.

One of the most transformative applications of this technology is virtual staining, a non-invasive method that mimics traditional chemical staining processes used in histology. Virtual staining does not require the physical application of dyes to tissue samples, preserving the original specimens while providing essential information about tissue morphology and pathology. This technique is finding important application in clinical settings, particularly during surgical procedures, where it facilitates rapid intra-surgical pathological diagnostics. By allowing pathologists and surgeons to visualise and assess histopathological details in real-time, virtual staining signifi-

cantly enhances the decision-making process during surgeries, potentially improving patient outcomes by enabling more precise and timely interventions.

Furthermore, the integration of machine learning with hyperspectral Raman imaging opens new avenues for automated tissue classification and diagnosis. Machine learning algorithms can analyse complex hyperspectral data to identify patterns that may not be immediately visible to the human eye, even to the trained pathologist, thereby providing an additional layer of diagnostic insight. This capability is especially valuable in scenarios where rapid and accurate tissue analysis is critical, such as during tumour resection surgeries. In Chapter 5 we take advantage of this pattern recognition capacity to develop a neural network with the aim of better quantifying the individual resonant contributions (the Lorentzian line-shapes) from a standard mixed spectrum.

This chapter is structured to provide a comprehensive overview of various Raman spectroscopy techniques, including Spontaneous Raman, Stimulated Raman Spectroscopy (SRS), and Coherent Anti-Stokes Raman Scattering (CARS). Each section will detail the specific contributions of these techniques to hyperspectral imaging, discuss their practical applications, and explore the future potential of integrating these approaches with advanced computational methods for enhanced diagnostic accuracy. While commercial development in this area place a particular emphasis on SRS owing to its high-speed and its recently proven application for real-time virtual staining and tissue classification, this work focuses on the only two modalities that can produce the full spectrum from a single acquisition, namely spontaneous Raman and BCARS. This provides the inspiration for the more detailed comparison between spontaneous Raman and BCARS in Chapter 3. This, in turn, led to the development of a new optical system capable of leveraging the strengths of both modalities, and in turn mitigating some of their inherent limitations in Chapter 4.

We begin in the next section with a review of the various opto-electronic modalities for recording and processing a hyperspectral image. This is followed by a series of sections that review of Stimulate Raman Histology, whereby the Raman hyperspec-

tral image can be processed to produce a H&E like stained image for pathological interpretation or machine learning based classification. We conclude with a section that offers a vision of the future for this exciting area of research.

2.2 Recording a Hyperspectral Raman Image

2.2.1 Overview

In this section, we briefly review the state-of-the-art in recording a hyperspectral Raman image, which is the precursor to Raman histology. We begin with a review of spontaneous Raman systems that employ a line-scanning approach. Spontaneous Raman spectroscopy remains the gold standard in the field in the sense that all of the coherent Raman systems attempt to produce the same information with higher speed, but none can match it in terms of quantitative accuracy. Next we will review the Stimulated Raman Spectroscopy (SRS) system which is the basis for most of the research that is reviewed in this chapter. SRS is very high-speed but it is typically used to record only one or two Raman wavenumbers. Finally we will review two alternative coherent Raman methods, which although not as fast as SRS, can provide broadband spectra, namely broadband and Fourier transform CARS. We note that the comparison and combination of spontaneous Raman and BCARS are the subjects of Chapters 3 4 which follow this review chapter.

2.2.2 Spontaneous Raman Scattering

Hyperspectral imaging with spontaneous Raman is not as commonly applied as SRS, but it has seen growing interest in recent years. Raman spectroscopy is a phenomenon whereby incident light interacts with molecular vibrations, resulting in scattered light with shifted wavelengths. This effect provides a unique spectral fingerprint for each molecule across the full spectral window, typically encompassing 400cm^{-1} - 3500cm^{-1}, and is primarily limited by the weak nature of the generated signal (1x10^{-6} photons scattered) and the geometric challenge of capturing photons which scatter isotropically in random directions. The setup for spontaneous Raman

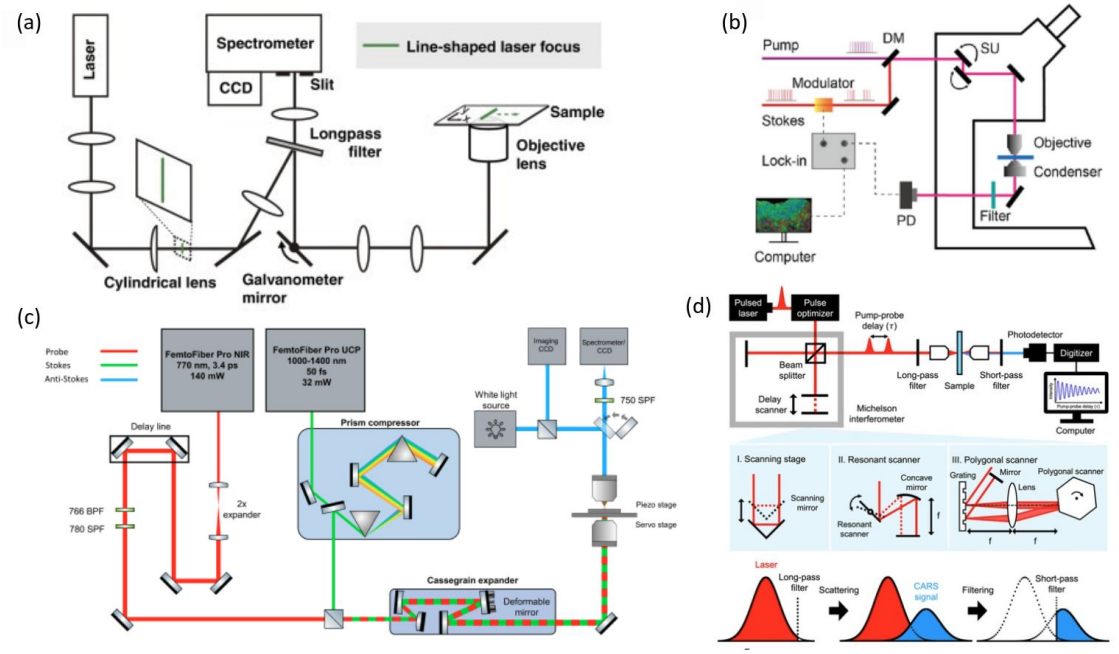

Figure 1: Diagram of the optical setup for the various Raman hyperspectral imaging modalities using (a) line scanning spontaneous Raman [40]; (b) Stimulated Raman Spectroscopy point scanning; (c) Broadband Coherent anti-Stokes Raman Spectroscopy [41]; (d) Fourier transform Coherent anti-Stokes Raman Spectroscopy [42]. These figures were reproduced from Fig. 1 [40], Fig. 1 [43], Fig. 1 [41], Fig 3 [44]

is comparably simple compared to other methods, typically requiring a single laser source that is focused on a sample. The Raman scattered light is separated from the laser using a longpass filter and is directed to a spectrograph where the spectral intensity is measured. The method can easily be coupled with a microscope by focusing the laser through the microscope objective and collecting the back-scattered photons via the same objective. This approach, termed "pi-Raman" is the most common and requires a dichroic beamsplitter to separate the laser and Raman photons since the input laster and output photons share the same path into and out of the objective. A confocal aperture is commonly included in the setup to enable 3D resolution as well as reducing the impact of laser induced luminescence from the glass substrate, which can be mitigated using background subtraction algorithms. [45] This architecture is adopted by many commercial Raman instrument companies including Horiba.

Typically recording a single spectrum from a biological sample requires an acquisition time in the order of 1 second. This is because, despite it's high spectral

resolution and relatively simple setup, spontaneous Raman is an inherently weak process with fewer than 1 in $1e^6$ photons being Raman scattered. This means that typical hyperspectral imaging, whereby the sample is point-scanned through the laser focus, becomes prohibitively time consuming. Nevertheless, because of the attractiveness of confocal Raman in terms of spectra some authors have advanced the approach toward 150 ms acquisition times, which is made possible with Singular Value Decomposition (SVD) denoising.[46]

In recent years, reducing the total time required for recording spontaneous Raman spectra has been a goal of many researchers, employing novel processing methods such as optical setups that allow for parallelised acquisition. Fig. 1 (a) is one such example developed by Fujita et al. in Ref. 40. This system, similar to the Renishaw commercial Raman instruments, projects a laser line at the sample plane that is focused by a cylindrical lens combined with the microscope objective. Raman scattering occurs in the sample at all points along this line. The back scattered photons are imaged to the line that is the spectrograph slit. A galvo mirror is placed in the system such that the illuminating line can be scanned across the sample, while the backscattered light will also be incident on the static spectrograph slit. Although this line scanning approach affords less depth resolution than confocal Raman, it has the significant advantage of enabling parallel Raman spectra to be recorded all along the line and therefore represents the state of the art for hyperspectral Raman imaging with full band. One limitation of this approach is the stronger background signal from glass, when compared with confocal. For this reason line-scanning often requires the use of quartz or expensive Raman grade CaF_2 substrates. 47 Also

worth mentioning is Ref. 48 in which the same research group have proposed an even higher-speed line scanning Raman system for hyperspectral imaging. Although this method is limited to the CH-band, it facilitates hyperspectral imaging at a speed that is comparative to BCARS, reviewed below. This technique employs multiple line illuminations for sample scanning, significantly improving the throughput compared single line scanning. This system utilises a powerful continuous-wave laser

and a multi-slit detection setup to scan the sample in parallel, reducing the overall scan time to approximately 10 minutes per megapixel. A bandpass filter narrows the detected spectral range to focus on specific Raman bands while filtering out unwanted fingerprint regions. Additionally, the system incorporates a custom-built confocal slit array, enhancing signal clarity and reducing noise. Importantly, SVD is applied to the spectral data for noise reduction and signal enhancement, making this approach highly effective for rapid and precise biomedical imaging and diagnostics.

2.2.3 Stimulated Raman Spectroscopy (SRS)

Stimulated Raman Spectroscopy (SRS) is the core technology behind many advancements in virtual staining and disease classification methodologies discussed in this review chapter. Since its innovation by the Xi group a decade ago,[49] SRS has gained widespread adoption due to its exceptionally fast acquisition speeds and enhanced signal clarity compared to traditional Raman techniques. SRS differs from spontaneous Raman spectroscopy by utilising two synchronised laser beams—a pump and a Stokes beam—to induce and amplify molecular vibrations within the sample. This enhancement is achieved through a nonlinear interaction process, where the energy difference between the two beams matches the vibrational energy of the sample's molecular bonds. The basic system used for hyperspectral imaging is shown in Fig. 1 (b), which is taken from[43].

Unlike the inherently weak inelastic scattering of spontaneous Raman, SRS amplifies the Raman signal, significantly enhancing detection sensitivity and speed. This makes it ideal for rapid and detailed molecular imaging. The implementation of SRS involves several optical components including a pulse laser system to generate the necessary beam properties, Acousto-Optic Modulators (AOMs) to precisely control beam frequencies, and Lock-In Amplifiers (LIAs) to detect the subtle changes in intensity resulting from the Raman effect.

One of the most significant milestones achieved by the Xi group in the development of SRS was the successful suppression of non-resonant background (NRB) noise, a

pervasive challenge in traditional Raman techniques. Another key innovation from the Xi group that propelled the use of SRS was the development of high-speed tuning systems for the pump laser making use of fiber lasers and optical parametric oscillators (OPOs), allowing rapid and precise adjustments across different Raman shifts to scan multiple molecular bonds efficiently. These adjustments typically require sequential tuning to cover a broad spectral range, which traditionally increasing the recording time. However, recent advancements have sought to circumvent this limitation by using broadband SRS techniques that can capture multiple wavenumbers simultaneously, thus speeding up the data acquisition process without sacrificing spectral information[50]. This approach, though more complex optically, reduces the need for multiple sequential acquisitions and integrates advanced optical modulation and signal detection strategies to manage the increased data throughput effectively.

2.2.4 Broadband Coherent Anti-Stokes Raman Spectroscopy (BCARS)

Broadband Coherent Anti-Stokes Raman Spectroscopy (BCARS) offers a significant advancement in the field of hyperspectral Raman imaging by providing full-spectrum Raman images without the need for the wavelength scanning typically employed in ultrafast Stimulated Raman Spectroscopy (SRS). This technique, which is depicted in Fig. 1 (c),[51] utilises a unique laser configuration that allows for the simultaneous probing of the entire biologically-relevant Raman spectrum, ranging from 500 cm^{-1} to 3500 cm^{-1}. Unlike ultra fast SRS, which enhances specific Raman lines by tuning the pump and Stokes beams to match molecular vibrations, BCARS generates a broadband spectrum in a single acquisition. This is achieved through the use of a supercontinuum source combined with a narrowband probe. The broadband nature of the light source in BCARS accesses a wide range of Raman shifts simultaneously, enabling the capture of comprehensive chemical information across the full fingerprint and CH regions in a single shot.

One of the primary advantages of BCARS is its ability to provide detailed chemical maps of samples without the need for sequential tuning, which is a time-consuming

aspect of SRS. However, BCARS systems typically face challenges with non-resonant background (NRB) noise, which can obscure the weaker Raman signals and distort the spectral data. Advanced numerical methods, such as the Kramers-Kronig transform or more recently using physics informed deep learning methods,[52,53] are employed to numerically remove NRB-induced distortions and retrieve the Raman-like spectrum from the raw BCARS data.

The acquisition time for BCARS, at approximately 3 milliseconds per spectrum, positions it between the ultra-fast SRS and the slower spontaneous Raman spectroscopy. This makes BCARS an ideal candidate for applications that benefit from rapid imaging but require more comprehensive spectral information than SRS can provide. For example, BCARS can be used to scan wide areas of tissue to provide a full spectral map within several minutes, which is significantly faster than spontaneous Raman but with a richer dataset than SRS. Furthermore, BCARS can complement SRS in a hybrid imaging approach, where SRS is used for initial rapid screening to identify regions of interest with potential pathological significance. Subsequently, these targeted areas can be examined in greater detail using BCARS to obtain comprehensive biochemical information. This complementary use of both technologies leverages the speed of SRS for screening and the spectral breadth of BCARS for detailed analysis, potentially providing a powerful toolset for biomedical research and diagnostic applications.

2.2.5 Fourier Transform Coherent Anti-Stokes Raman Spectroscopy (FT-CARS)

For the sake of completeness we also briefly review FT-CARS as another possible method for hyperspectral Raman imaging, which has received some attention for its ability to acquire broadband Raman spectra at extraordinarily high speeds. As illustrated in Fig. 1 (d), the basic principle in FT-CARS is to split a femtosecond pulse laser beam into two paths to create a pump and a probe beam. These beams interact in a sample through a Michelson interferometer setup where one arm is fixed

and the other includes a rapidly scanning delay line. This configuration generates a CARS signal whose frequency components are down-converted into a range that can be rapidly read by conventional photodetectors. The technique's capability to rapidly scan the delay line allows it to perform spectroscopic measurements much faster than traditional methods, reaching acquisition rates of up to 50,000 spectra per second.[54] In Fig. 1 (d), the delay scan is implemented using a polygonal mirror array in its Fourier plane, which, through its rotation, scans the delay line at high speeds. This setup enables a rapid succession of Raman spectra over a wide range of Raman shifts, making FT-CARS highly suitable for very high speed spectral recording. Another advantage is that the NRB, a problem in BCARS, is mitigated by the FT recording process.

Despite these technological advancements, the application of FT-CARS in biological sampling poses specific challenges, and has yet to be fully tested to the same level as SRS or BCARS. The rapid scanning required by FT-CARS (such as the polygonal mirror array in the figure), while beneficial for speed, might simply be too fast to record useful broadband spectra from a biological sample, which are inherently weak scatterers. Biological samples often exhibit lower Raman scattering efficiencies, which means that the high-speed capabilities of FT-CARS might not translate directly to practical speeds when imaging biological tissues. The faster the scan, the lower the signal-to-noise ratio tends to be, unless compensated by high laser powers or advanced noise-reduction technologies (such as the use of LIAs in SRS). This could necessitate slower scanning to accumulate sufficient signal, balancing the inherent advantages of the method's speed. In Table 1 a comparison of the various hyperspectral Raman modalities is provided for bio-imaging.

2.2.6 Preprocessing

Preprocessing in hyperspectral Raman imaging is a necessary first step before any virtual staining or classification algorithms. This step ensures the accuracy and relevance of the spectral data obtained from various Raman spectroscopy modalities.

Table 1: Comparison of Raman Hyperspectral Modalities

Acquisition Mode	**Effective Per-Pixel Dwell Time**	**Spectral Resolution**	**Bandwidth (cm^{-1})**
Point Scan Raman	150 ms*[46]	1-10 cm^{-1}	400-4000
Line Scan Raman	24 ms*[40]	1-10 cm^{-1}	400-4000
Multi-line Scan Raman	617 μs*[48]	3-5 cm^{-1}	2800-3200
32-channel SRS	40 μs[50]	10 cm^{-1}[55]	2800-3200
Dual-channel SRS	10 μs[56]	10 cm^{-1}	10-20
BCARS	3 ms*	10-30 cm^{-1}[57]	400-4000

*Dwell times achieved using SVD denoising.

Each method comes with its own set of challenges and requirements that dictate the necessary preprocessing steps. In spontaneous Raman spectroscopy, preprocessing involves wavenumber calibration[58] to align spectral features accurately with their true positions. Intensity calibration is also important to ensure that the response from the spectrometer is consistent across different wavelengths from one instrument to another. Cosmic ray removal is also important and for hyperspectral imaging this is best done by comparing a given spectra with other recorded spectra in the dataset to indemnify and replace data affected by cosmic rays[59]. Background subtraction is another critical step, particularly to remove signals arising from the substrate or medium, which can obscure or distort the Raman signal; glass in particular can be problematic but it can be effectively removed using a least squares regression method. [45] Additionally, baseline removal is employed to correct for distortions due to autofluorescence or other broad spectral features, often using techniques like polynomial fitting or rolling ball algorithms.[45] For SRS, minimal processing is required, with wavenumber calibration and simple baseline correction being sufficient. Additionally, signal enhancement techniques, such as Lock-In Amplification, are utilised to bolster the weak Raman signal against a dominant noisy background. BCARS preprocessing focuses on wavenumber calibration for spectral accuracy. The most significant preprocessing step in BCARS is the removal of non-resonant background (NRB) noise, often using methods like the Kramers-Kronig transform or deep learning methods[52,53]. This step is crucial as the NRB can significantly obscure or distort the actual Raman signal in the collected spectra. Intensity calibration is trivial in

BCARS. FT-CARS primarily involves transforming time-domain signals into the frequency domain to extract Raman spectra using a discrete Fourier transform. Across all these modalities, Singular Value Decomposition (SVD) is a powerful tool used for denoising hyperspectral datasets. SVD works by decomposing the data set into singular vectors and values, which can be truncated to minimise noise while preserving the essential structure of the data.

2.3 Virtual Staining with Procedural Approaches

2.3.1 Overview

We use the term *procedural approach* to refer to the identification of a small subset of wave-number information that can be directly related to the molecular bonds associated with clinical H&E staining, which are then used to 'virtually stain' a hyper-spectral SR image such that it has the same appearance as a H&E image. The goal here is to make use of the label-free molecular information provided by real-time SR in order to produce histopathology images that are equivalent to an image that could be obtained by chemical staining. This procedural approach is not probabilistic as for the case of generative deep learning methods. It provides a direct chemical link between the label free imaging and the labeled imaging by exploiting the chemistry involved in chemical staining. The primary source of information for procedural generation of H&E images the procedural methods reviewed below are the CH_2 (found in cell bodies) and CH_3 (in cell nuclei) chemical bonds, relating directly to the 2845 cm-1 and 2930 cm-1 wavenumbers. These bonds are found in abundance in cell bodies and cell nuclei. Where the 2930cm^{-1} region has stronger responses to nucleic acids and similar contributions as the 2845cm^{-1} from lipids and proteins, by subtracting the 2845cm^{-1} from the 2930cm^{-1} image the contribution from proteins and lipids is weakened thereby enhancing the contrast for nucleic acid specific peaks.

There exist two research groups that have contributed to this area in recent years,

both of which have spun out start-up companies to commercialise their technologies towards real-time histopathology imaging in surgery to assist with identifying diseased tissue and tumour margin. The Xie group, which pioneered the development of SRS, have developed a high-speed low-noise SRS bioimaging system that can be tuned to 2845 and 2930 cm^{-1} and two wavenumber images are recorded sequentially.[60,61,62,63] This hyperspectral dataset is then subject to a bespoke image processing algorithm to produce a H&E like image. The Rigneault group at the university of Marseilles make use of a customised SRS system capable of recording the two wavenumbers simultaneously (via a phase) and the resultant hyper-spectral image is subject to a pseudo-automated image processing routine that can be tuned to produce convincing H&E like images[56,64,65,66]. This group also makes use of the second harmonic generation(SHG) signal that can also be generated by the same system, in order to further enhance the quality of their virtual staining.

2.3.2 Stimulated Raman Histology - The Xie Group[63]

The Xie group were the first to develop low-noise stimulated Raman scattering (SRS) bioimaging systems in 2008,[61] pioneering a technique that significantly improved the sensitivity and speed of label-free imaging for biological applications. They went on to commercialise this breakthrough through the company Invenio Imaging. A key innovation by the Xie group was the use of a fiber laser system for SRS imaging, which provided a more stable, compact, and user-friendly light source compared to traditional laser setups. This innovation allowed for better noise suppression and enhanced signal quality, particularly in live-cell and tissue imaging, by stabilising the power and phase of the laser output. Combined with their methods to minimise the non-resonant background, the fiber laser-based SRS system became a practical and powerful tool for real-time, high-contrast imaging in biomedical research and clinical diagnostics.

SRS images are captured at 2845cm^{-1} and 2930cm^{-1} using the portable set-up shown

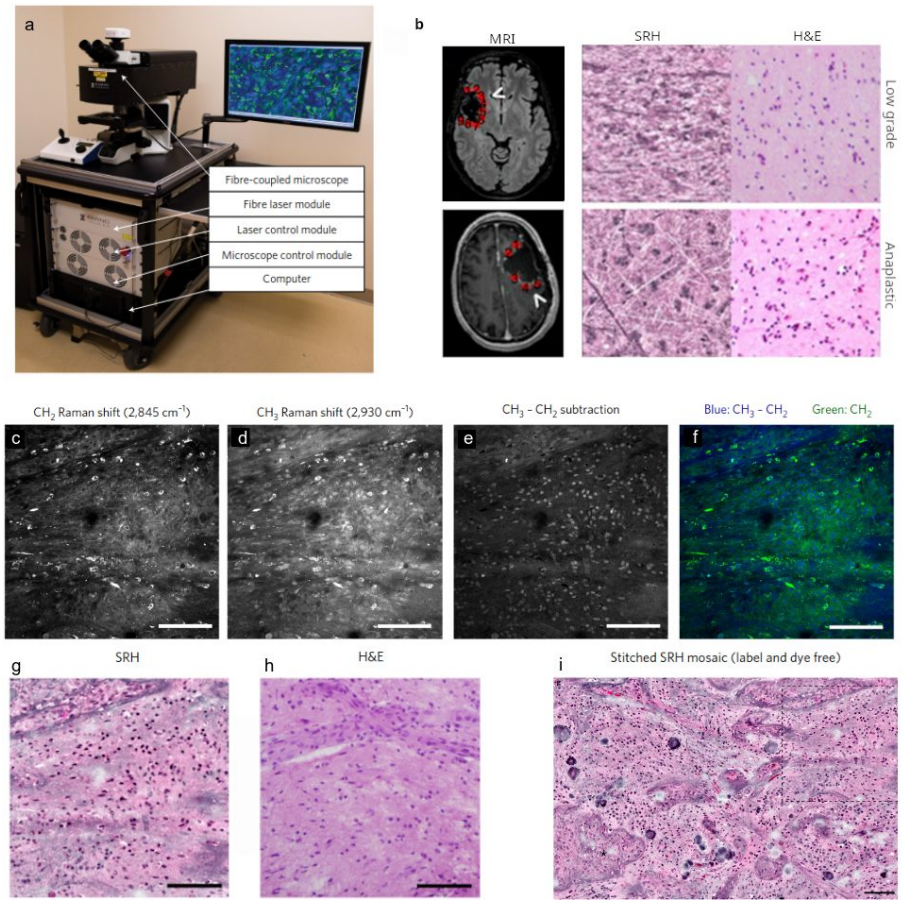

Figure 2: (a) Invenio Imaging SRS System In Operating Room; (b) MRI of Recurrent oligodendroglioma & glioblastoma side by side with comparison of virtually stained SRH and true H&E images. The remaining image parts relate to the virtual staining method: (c,d) Individual Wavenumber Rasters of 2845cm^{-}1 and 2930cm^{-}1 relating to CH_2 and CH_3 respectively; (e) Subtraction of these two images. (f) False Color Composite where the blue band is the $CH_2 - CH_3$ and Green the CH_2. (g) The false colour composite image is input to a look up table to produce the virtually stained SRH image; (h) True H&E stained image for comparison. (i) The system records a grid of NxN images which are then stitched together to form a large mosaic SRH image. This figure has been reproduced from Fig. 1 & 2 in[63], Fig. 2 in[67].

in Fig. 2 (a) relating to information from the CH_2 and CH_3 chemical bonds, respectively. In Fig. 2 (b) virtually stained SRH images are compared with corresponding H&E images for low-grade and anaplastic brain tissue sample; the MRI image on the left indicate the region from which this tissue has been excised. Brain surgery is a particularly suitable application for real-time SRH due to the need to carefully and accurately identify tumour margins during surgery, A key component in the SRH system is the image processing method that takes as input the two wavenumber images and produces as output a convincing virtually stain H&E like image. This image processing method is illustrated in Fig. 2 (c)-(i). The 2845cm^{-1} (CH_2) and 2930cm^{-1} (CH_3) recordings are shown in Fig. 2 (c) and (d). The data from each recording is placed in a separate channel of a two-channel image. A blurred version of the two-channel image is generated by performing a Gaussian blur, the initial unblurred image is then divided by the blurred image to remove artefacts in either acquisition or tissue preparation producing a new denoised two channel image. A nucleic index is generated by subtracting the channels of the new denoised two channel image (corresponding to CH_3 - (CH_2) shown in Fig. 2 (e) The next step is to create an RGB image using these two channels. A simple example of creating an RGB image from the nucleic map and the denoised 2845cm-1 denoised raster is shown in Fig. 2 (f) where in this case the two channels are mapped to a blue scale and a green scale respectively. Creating a convincing virtual H&E image requires a more nuanced approach in mapping to the RGB colours. The virtual H&E image is generated by taking each channel independently and mapping these to two different RGB colour schemes, followed by combining the two RGB images. The 2845cm^{-1} CH_2 image pixel intensity is linearly related to a white to pink colour scale. Separately, the CH_2 - CH_3 image pixel intensity is linearly mapped to a white to purple colour scale. The combined RGB image is shown in Fig. 2 (g) and compared with the real H&E image in Fig. 2 (h). Each of these images contains 1024x1024 pixels. Given the high-speed of data aquisition for SRH, it is possible to scan the slide capturing a large grid and generating mosaic image of a wide field of tissue as shown in

Fig. 2 (i).

The Xie Group has utilised this method in a variety of studies[67,68,69,70,71] demonstrating the ability of Stimulated Raman histology to be used as a diagnostic tool both by histopathologists by providing real-time H&E images in surgery as well and in conjunction with computational methods for automated diagnosis.

2.3.3 Virtual HES coloring - The Rigneault Group[56]

Separately the Rigneault group has also focused on SRH using a bespoke SRS system and using more complex image processing methods.[56,64,66] The key methods for virtual staining proposed by the Rigneault group are shown in Fig. 3 below. This includes a novel optical system as shown in Fig. 3 (a), which combines multiple imaging modalities, including both SRS and Second Harmonic Generation (SHG) in their virtual staining method. The SRS setup uses a mode-locked ytterbium (Yb) fiber laser operating at 1031 nm as a Stokes beam. Two optical parametric oscillators (OPOs) are tuned to different wavelengths (797.3 nm and 792.2 nm) to probe the CH_2 and CH_3 vibrations making use of the same two wavenumbers used by the Xie group. These beams are recombined and scanned across tissue samples to create high-resolution images. There are, however, a number of advancements over the methods used by the Xie group. Firstly, the authors introduce a method called frequency modulated SRS (FM-SRS); whereas traditional SRS requires the two wavenumbers (2845 cm^{-1} for CH_2 and 2930 cm^{-1} for CH_3) images to be recorded sequentially and the difference to be calculated, FM-SRS enables the difference to be recorded directly using a single-channel lock-in amplifier, thereby speeding up the process. This is achieved by setting a 180^o phase between AOM_1 and AOM_2 shown in the figure. As well as speed another advantage of this approach is lower noise. Noise can accumulate during the separate sequential imaging processes, therefore the capturing of two wavenumber signals in a single acquisition may reduce noise by minimising time between measurements and reducing the likelihood of noise interference from external factors or sample movement. In using the FM-SRS modality

it is possible to simultaneously record a CARS signal, TPEF, SHG and SRS, the latter two being utilised for virtual staining.

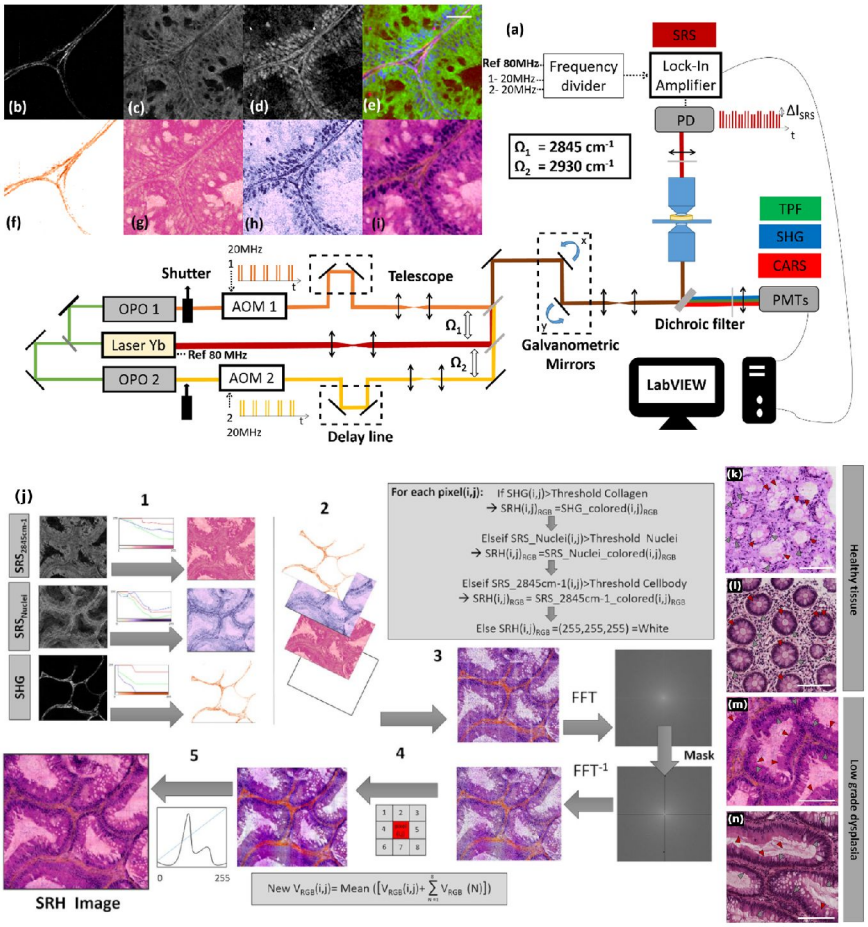

Figure 3: [56] (a) Optical diagram of FM-SRS, TPEF/CARS & SHG capable system. (b-d) Grayscale images relating to SHG, 2845cm^{-1} and (2930cm^{-1}-2845cm^{-1}) (e) A simple histologically relevant visualisation of collagen, lipids/proteins and nuclei by combining the three grayscale images previously mentioned as the channels in an RGB image. (f-h) Colour images generated from the previously described grayscale images, mapped to orange (collagen), pink and violet (lipids/proteins and nuclei). (i) A virtual HES image generated by hierarchically superimposing the three coloured images described previously. (j 1-5) Illustration of the algorithm used to generate a large virtually stained HES image; see text for further details. This figure has been reproduced from Fig. 1 2 3[56]

The algorithm demonstrated in Fig. 3 J illustrates the virtual staining algorithm in λ-switch mode as described in Ref. 56 and will be discussed in that context in the next paragraph. Fig. 3 (b-i) show the Virtual Hematoxylin, Eosin and Saffron (HES) staining method being applied to standard λ-switch SR in which the 2845cm^{-1} and 2930cm^{-1} are recorded sequentially by alternating between OPO 1 and OPO 2.

Fig. 3(b) showing the Second Harmonic Generation (SHG) image, which is particularly useful for visualising non-centrosymmetric and highly ordered molecular arrangements such as those found in collagen[72]. Fig. 3(c) shows the 2845cm^{-1} a signal in the CH_2 region relating to proteins/lipids. Finally Fig. 3(d) shows the subtraction: 2930cm^{-1}-2845cm^{-1}. Fig. 3 (e) demonstrates a false colour composite built by combining these three grayscale images into a standard RGB image. Fig. 3 (f-g) illustrates the colour layers generated by applying color maps to the grayscale images (b,c,d): orange for the SHG, pink for the 2845cm^{-1} (CH_2) and dark blue for the 2930cm^{-1}-2845cm^{-1} (nucleic map). A virtually stained patch is the combination of these images as shown in Fig. 3 (i). The method for producing a full-scale virtual label free HES-like image is described below. (k) Shows a virtually stained healthy tissue, contrasted with (l) which is a true HES stained image. (m) Similarly shows a virtual HES image of low grade dysplasia contrasted with with (n) a true HES image.

In order for the methods to be clinically relevant, it is necessary to produce widefield virtually stained images for the clinician. This is achieved by recording and processing a grid of smaller virtually stained images, followed by stitching these images together into a large mosaic. However, this approach can result in discontinuities at the borders and to mitigate this problem further image processing is required. The full pipeline is illustrated in Fig. 3 (j). Starting with Step 1, where a color gradient from white to a targeted colour is generated: orange to mimic a saffron dye for collagen, pink to mimic eosin, and dark blue to mimic hematoxylin. In Step 2, we see the hierarchically exclusive composition of the color maps applied to each individual pixel. For each channel, there is a user-adjustable threshold, of the form K times the mean of the given channel with the constraint $0 < K < 2$. The user may adjust the value of K to increase or decrease sensitivity of the given channel. Should the pixel value be above a defined threshold, the colour corresponding to the channel will be set. If the pixel is not accepted by the first threshold it will go on to the next. Should none of the channels values be accepted by the corresponding thresh-

olds the pixel will default to white. This sequence of steps is applied to each of the smaller images, and the results are stitched accordingly. Step 3 is the stitching step. This involves a filter in the spatial frequency domain, whereby the Discrete Fourier Transform (DFT) of the image is calculated using an FFT algorithm, followed by multiplication with a cross-hair mask and an inverse DFT. This process helps to remove obvious linear discontinuities between neighbouring $100\mu m^2$ images. Step 4 involves a standard box blur being applied to smooth the image and a final user guided colour correction may be applied via histogram to better match a standard HES image in Step 5.

Fig. 3 (k) shows a virtual HES image of a healthy tissue sample contrasted with Fig. 3 (l) a ground truth image from a similar region. Similarly Fig. 3 (m) shows a virtual HES image of a low grade dysplasia contrasted with Fig. 3 (n), the corresponding ground truth physically stained image. In both cases the virtually generated label-free images bear a convincing similarity to their labelled counterparts, albeit with some subtle variations in colour intensity and contrast. Variations of these types can be effectively addressed using methods of style transfer and deep learning, which are reviewed in the next section.

2.4 Virtual Staining with Style Transfer

2.4.1 Overview

In the previous section we reviewed methods for mapping from a SRS label free image to an image that is visually similar to a labelled H&E or HES histological image using procedural algorithms that map from specific Raman wavenumbers to various colour maps. While these methods are advantageous in terms of computational time as well as the existence of a clear chemical link between the label and label free images, there is, however, the disadvantage of variable results in terms of subtle variations in colour intensity and contrast when comparing the label-free and labelled images. This can be mitigated to an extent using manual tuning of the procedural algorithms. To better address these issues, deep learning based

image-to-image translation methods been investigated, which can be collectively described as Style Transfer methods. Style transfer[73] is a field within machine learning involving the algorithmic manipulation of an image's semantic content to embody the stylistic characteristics of another. While many style transfer techniques originated in image synthesis and have traditionally been associated with artistic, design, and entertainment applications, their utility has expanded into the realm of data augmentation and analysis. Notably, in the domain of histology, researchers have begun to leverage style transfer methodologies to transform between a variety of sensing modalities, such as brightfield microscopy with staining agents. Raman Spectroscopy is no exception to this, and novel approaches hold significant promise for enhancing data interpretability and potentially uncovering new insights within histological studies.

Early style transfer networks utilised a dual loss function based on minimising (1) the difference between extracted features of a given input image with a newly generated image and (2) the differences in the co-occurrences of feature matrices from the generated image to the target style image.[74] While there were significant successes with this methodology, it often lead to local artefacts that detracted from the verisimilitude of the generated images. Adversarial approaches quickly emerged that were less prone to this issue in the form of Generator-Adversary Network (GAN) architectures. In the GAN architecture a generator would produce novel images and place it within a collection of original images. i.e. a set of true images and generated images. The generator architecture works in tandem with a discriminator architecture; The loss applied to the generator is based on how well the discriminator is fooled into misclassifing the generated image as a true image. The end result is a greater emphasis on global contexts within an image. Unfortunately these GAN networks even in the context of style transfer were more likely to hallucinate features locally in order to maintain the global context. In an attempt to maintain the increased global context and simultaneously

bolster local accuracy the Cycle Consistent GAN architecture emerged.[75] The core idea in the CycleGAN architecture is to devise a two-way transfer between domains with dual-generators and discriminators where a mapping is only successful if it can translate from domain A to domain B, and independently be translated from domain B back to domain A in such a way that the image $Ax_i \mapsto Bx_i$ and $Bx_i \mapsto Ax_i$. Recently these methods have been applied to produce virtual staining using SRS (and other forms of non-linear spectroscopy) images, and these contributions are described in the subsequent sections. While the procedural methods explored earlier have no need of co-registered label-free and matching labelled data, traditional generative deep learning methods often do have this requirement. One of the significant advantages of the CycleGAN is it's potential to circumvent this need by utilising the full cyclical mapping from the original domain to the transformed and back to original. The total differences between the original image and it's regenerated counterpart provides reasonable ground truth, ensuring no information necessary to reproduce itself was lost while translating the image style between the two domains.

2.4.2 Computational Tissue Staining - The Blocklitz Group[76]

In their work on Computational Tissue Staining, the Blocklitz Group were the first to transfer non-linear microscopy (NLM) images to histopathologically stained H&E image in an unsupervised manner. They contrasted the semi-supervised conditional GAN (CondGAN) model with the unsupervised cycle consistent conditional GAN (CycleCGAN) model to explore the potential for label-free style transfer methods with non-linear microscopy.

Initially a series of NLM images are recorded where signals relating to CARS, TPEF and SHG are used to generate a False Colour Composite by setting each as the channels in an RGB image. An example of these images is seen in Fig. 4 (A) & (H). In order to aid the training using colour styles close to that of a H&E style, these false colour composites are inverted, moving from red and green hues, to purples,

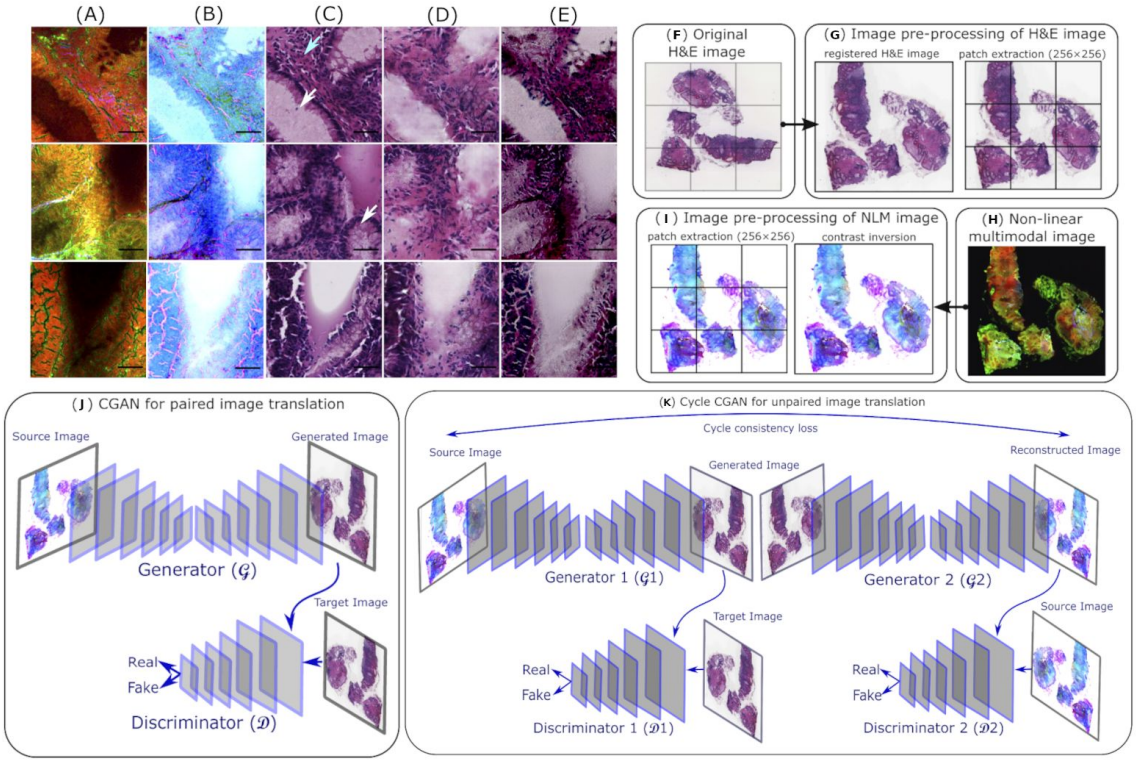

Figure 4: (A) False colour composites relating the CARS,TPEF and SHG to RGB channels. (B) Inverted false colour composites to better match a H&E colour distribution. (C) True H&E images recorded via brightfield microscopy. (D) Virtual H&E stain generated via conditional GAN and (E) virtual H&E generated via CycleGAN. (F-H) Illustration of the preprocessing steps undertaken. (F) Shows the raw H&E input image. (G) Illustrates the image registration to NLM data and 256x256 patch extraction. (H) The initial false colour composite of CARS, TPEF & SHG before contrast inversion. (J) Illustrates the conditional GAN architecture. (K) Illustrates the CycleGAN architecture Figure Reproduced from Fig. 2 Fig. 3 in Ref.[76].

and blues. Ground truth images of H&E stains are also recorded via brightfield microscopy, examples of which can be seen in Fig. 4 (C). The inverted false colour composite images are passed through both a CondGAN and a CycleGAN where H&E resembling virtually stained images are produced. Examples of this can be seen in Fig. 4 (D) for the conditional GAN and (E) for the Cyclegan. As is visually apparent between the images generated by both networks they achieve comparable results despite the lack of paired images for the CycleGAN.

Before discussing the networks themselves we will briefly discuss the preprocessing methods as shown in Fig. 4 (F-H) The images were prepared such that both inputs should match in terms of both pixel count and pixel size. The NLM images

were recorded with a pixel size of $0.227\mu m^2$, and the brightfield microscopy images $0.219\mu m^2$. The H&E images were registered to align with the NLM data, an example of this translation can be seen in Fig. 4 (F-G). Non-overlapping image patches of pixel count 256x256 were extracted from both datasets as shown in Fig. 4 (G) for H&E and Fig. 4 (I) for NLM images. The inversion of the false colour composites discussed earlier is implemented before patch extraction in order to better match the initial input image to the expected color distribution of the histopathologically stained H&E image.

The conditional GAN while taking significantly less time to train, requires strongly supervised input data with images that share spatial dimensions, resolution and are co-registered to one another. As shown in Eq. 1 and Fig. 4 (J) there are two sources of loss in the CondGAN, the conditional loss and the discriminator loss.

$$L_{\text{Generator}} = \underbrace{\lambda MAE(z_{generated}, z_{target})}_{\text{Conditional Loss}} + \underbrace{(1 - D(z_{generated}))}_{\text{Discriminator Loss}} \tag{1}$$

$$L_{\text{Discriminator}} = \underbrace{D(z_{generated})^2}_{\text{Generated input mislabelled}} + \underbrace{(1 - D(z_{target}))^2}_{\text{Real input mislabelled}} \tag{2}$$

The conditional loss is given by the mean absolute error between the image generated and the known paired target image for the given input. This is followed by the inversion of the discriminator loss. This is because the discriminator component of the generator loss is determined by how frequently the generator fools the discriminator. The architectures used for the generator and discriminator were selected from the available literature. In the case of the discriminator it was based on the PatchGAN architecture first proposed by Isola et al[33]. The generator was inspired by the U-NET model, an autoencoder with skip connections between encoder and decoder layers of corresponding dimensions[34].

The Blocklitz group compared this CondGAN with CycleCGAN, the latter having the significant advantage of not requiring directly paired data. This can be de-

scribed by the loss function given in Eq. 3 and Fig. 4 (K). The structure of the discriminator for the cycle was kept mainly the same, a discriminator based on the PatchGAN architecture such that the final layer is not fully connected allowing for back-propagation to select areas. A generator architecture that favoured a RESNET model over the previous U-NET. In this RESNET model, residual connections are used in the bottleneck but no skip connections exist between the encoder/decoder layers of equivalent dimensions[35].

The CycleCGAN total loss can be given by the following

$$\begin{aligned} L_{\text{total}} = & \underbrace{L_{(D1)}(G_1, D_1, X, Y)}_{\text{Discriminator 1 loss}} + \underbrace{L_{(id1)}(G_1, D_1, X)}_{\text{Id loss}} \\ & + \underbrace{\lambda L_{(cycf)}(G_1, G_2)}_{\text{Forward Cycle Consistency}} + \underbrace{\lambda L_{(cycb)}(G_1, G_2)}_{\text{Backwards Cycle Consistency}} \end{aligned} \tag{3}$$

In this formulation the H&E domain discriminator loss is once again defined as the mean absolute error between the discriminator predictions of authenticity and their corresponding inverted accuracy. In addition an id loss term is added which calculates the MAE between a generators input and target image, thus guarding against dramatic changes of the image content. In particular this id loss is used in the training process where images from a generators target domain are given as an input. This penalizes any deviations from the original input, and encourages preservation of essential image content. The forward Cycle Consistency measures the mean absolute difference between an input image and a cyclical generation of the same domain. That is to say a non-linear microscopy image will be compared with the translation of the H&E generated by the first generator after it has been regenerated by the second generator back into a non-linear microscopy image, and the same for the backwards Cycle consistency with the domains inverted. These cycle consistency terms ensure that the information required to translate from one domain to another is preserved.

2.4.3 Stimulated Raman CycleGan (SRC-GAN) - The Ji Group[77]

While the previous methods used frozen H&E samples for training and validation, formalin-fixed and paraffin embedded (FFPE) H&E staining is currently considered the gold-standard for histopathology. This is because it is demonstrably better at preserving tissue architecture and cellular detail, it also allows for thinner samples that are more in line with the vertical focal area of non linear microscopy systems. However, frozen samples are often prefferred in surgical situations due to the foreshortened time frame for analysis. The Frozen H&E staining process takes in the proximity of 45 minutes, while FFPE can take longer than 36 hours. The Ji Group developed a routine to produce FFPE appearing images using SRS in shorter time than either the frozen H&E or the FFPE staining methods. The acquisition time and subsequent style transfer of a non-linear microscopy image taking in the vicinity of 3 minutes according to the author's estimates.

Image-to-image translation between stimulated Raman scattering (SRS) microscopy and formalin-fixed paraffin-embedded (FFPE) hematoxylin and eosin (H&E) staining presents a significant challenge for paired networks. This difficulty arises from the interference of the paraffin Raman signal, which obscures the signals of interest and hinders the establishment of accurate ground truth for supervised training. To address these limitations, Ji et al. developed a Stimulated Raman CycleGAN (SRC-GAN) 5 (a) based on a modified CycleGAN architecture. Their key innovation was the incorporation of a dynamic training domain that includes both strongly and weakly supervised elements.

As SRC-GAN is designed to map between two image domains, corresponding datasets were required for training. The SRS dataset was acquired using a dual-phase lock-in detection SRS system. A series of brain tissue lipid and protein (L&P) false-color composite images were generated using two-color SRS microscopy at two Raman frequencies: 2845 cm^{-1} (CH_2 stretching vibration) and 2930 cm^{-1} (CH_3 stretching vibration). The H&E dataset was generated through two approaches. First, true FFPE H&E images were acquired using a bright-field mi-

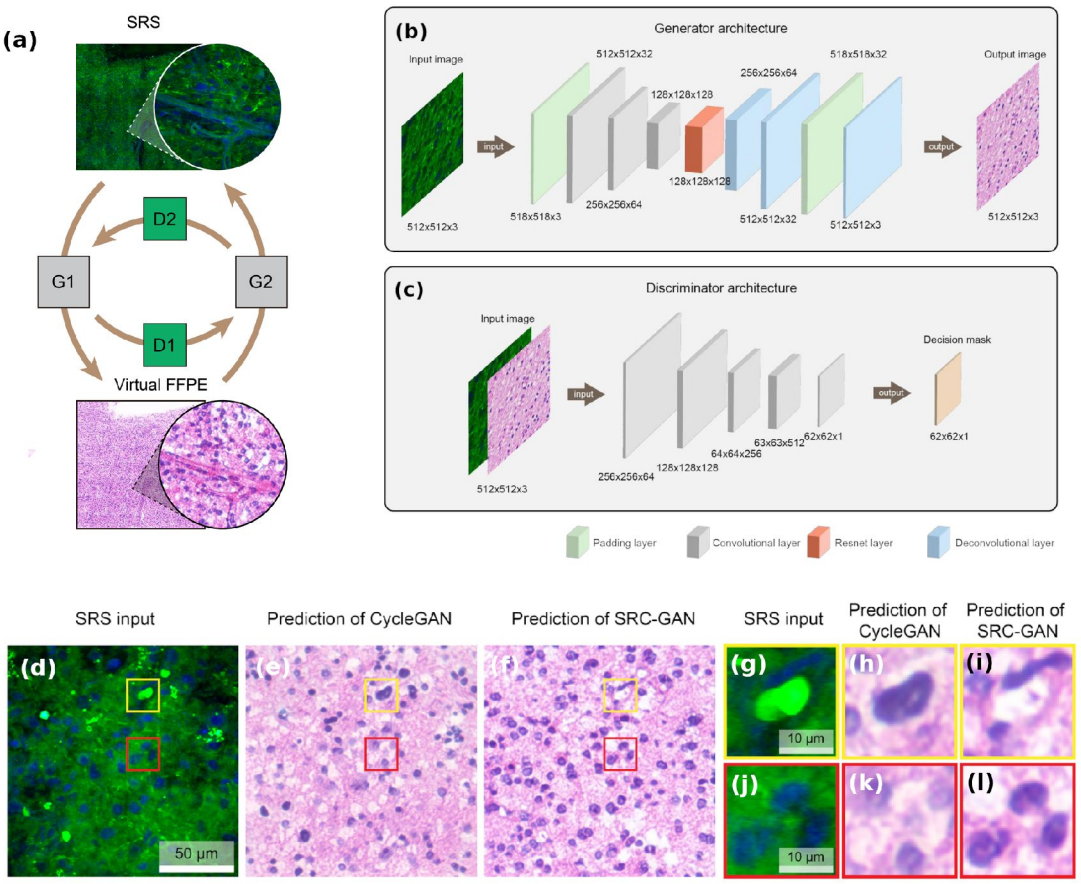

Figure 5: (a) An illustration of the CycleGAN structure between the SRS image domain and Virtual FFPE domain. (b) A diagram of the generator CNN architecture. (c) Diagram of the discriminator architecture. (d) Lipid and Protein False Colour Composite. (e) Traditional CycleGAN generated image. (f) SRC-GAN generated image. (g-l) Close-up of two mismap errors highlighted between source SRS image, CycleGan Prediction and SRC-GAN prediction. Figure reproduced from Fig 1 and Fig.2 in Ref.[77].

croscope equipped with a CCD. Mosaic images were captured and stitched together to create large whole-slide images of the stained tissue. Second, virtual H&E images were generated by applying H&E look-up tables to the L&P false-color composites, providing strongly supervised training data.

Fig. 5 (a) Shows the basic structure of the CycleGAN network with the SRS Lipid and Protein false colour composites as the first domain, and the virtual FFPE images as the second domain. In the standard CycleGAN form there are two pairs of generators and discriminators. The first for mapping from the SRS domain to the H&E domain, the second to map the inverse. For each input H&E image the

total loss is that of the generators ability to overcome the discriminator and the total difference between the original image and the twice transformed image. Real H&E data from the second domain are treated similarly, translated to the SRS lipid and protein false colour composites and tested by the second discriminator, and cycled back through G1 where a cycle consistency loss is used to compare the newly generated H&E image with the initial input. As such the loss can be formulated as below in Eq. 4.

$$Loss_{Generator_A} = \underbrace{MSE(D_B(G_A(X)), real\_mask)}_{\text{Generated Image Classified as Original}} + \underbrace{L1\_loss(X, G_B(G_A(X)))}_{\text{Cycle Consistency Loss}} \quad (4)$$

$$Loss_{Discriminator_A} = \underbrace{MSE(D_a(G_B(Y)), fake\_mask)}_{\text{Failure to Identify generated image}} + \underbrace{MSE(X, real\_mask)}_{\text{Failure to identify real image}} \quad (5)$$

Where the ability of the generator is measured by the minimisation of the mean squared error between the real mask and the discriminators predicted mask for generated images. This is followed by a cycle consistency term, in this case the rectilinear distance between the generated image and it's initial source. The discriminators loss is given as the mean squared error between the discriminators prediction of the generated images and the false mask summed with the mean squared error of the original images and the true mask. Generator B follows the same logic with the opposite domains and generator/discriminator pairs. Unlike for the GAN developed by the Blocklitz group, no id loss term was used for regularisation.

The structure of the generator and discriminator architectures are illustrated in Fig. 5 (b) and (c), where the generator is an autoencoder comprised of an initial padded layer followed by three convolutional layers, each with a 3x3 kernel. The padding layers were incorporated to assist with poor edge mapping. A residual architectural layer was used before deconvolution to enable the network to map both linearly and non-linearly. This is particularly useful in the context of style transfer where the non-linear transformations are likely to capture complex relationships and

the linear mappings allow for preservation of image content. The discriminator is a classifier based on PatchGAN[33]. This discriminator is fully convolutional with 4x4 kernels. Decision labels and output dimensions are kept the same to allow varying dimensions to the input images. This discriminator provides an output array of values rather than a single scalar, enabling the discriminator to pay closer attention to more refined segments.

To aid the network in learning correct mapping despite the lack of real ground truth, a dynamic training set was used to set the initial direction of the networks learning in a strongly supervised manner. These strongly supervised samples are slowly dominated by additional weakly supervised data until eventually they are completely phased out. The strongly supervised samples used are SRS images and virtually stained H&E images. The weakly supervised samples contain true SRS false colour composites and FFPE H&E images. The authors highlight that this dynamic training set was utilised to address a common pitfall of traditional CycleGANs whereby incorrect mappings can be learned due to the significant difference in color styles and contrast mechanisms between the two modalities. Two common examples of these mapping errors are highlighted in Fig. 5 (g-l); lipid droplets being mismapped to cell nuclei and nuclei mismapped to common voids.

As can be seen in Fig. 5 (g-l), the SRC-GAN predictions (i)&(l) appear to be significantly better than the traditional CycleGAN predictions (h)&(k). While true FFPE were diagnosed equally accurately or more accurately by histologists in the study, the Stimulated Raman virtual histology generated from the SRC-GAN significantly outperformed the Frozen H&E in terms of accurate diagnosis and approached close to that of the FFPE.

2.5 Automated Diagnosis with SRH

2.5.1 Overview

The virtual staining methods explored earlier hold great utility for enabling intra-operative pathological tissue imaging. This is achieved by mitigating time-

consuming bottlenecks and providing FFPE-H&E (like) gold standard images using frozen tissue. However, the pathologist is still required to interpret these virtually stained images and there currently exists a global shortages in the number of histopathologists. In intra-operative scenarios it is also preferable to avoid the need for an additional on-site clinician. Another concern is that it is well known that pathologists can produce variable diagnostics. For these reasons, a growing corpus of work has been built to not only reduce the time required to generate usable imagery for diagnosis, but also to automate the diagnosis process itself based on these images. In this section we review several recently proposed methods to do this. The question of automated diagnosis can generally be viewed as a classification problem and as such there are many viable machine learning methods that can be utilised to address it, including, multi-layer perceptrons, random forests, and support vector machines. These methods can be applied to the virtually stained imagery generated with the previously explored methods and/or the relevant spectra themselves without the need to produce RGB images. Research has been carried out in a wide number of examples.

2.5.2 Multi-Layer Perceptrons[78]

The intra-operative image data required for diagnosis and clinical decision making can be produced using Raman based virtual staining methods as described in earlier sections. However the interpretation of this data can still be costly in terms of time and prone to variability of diagnosis by different histopathologists. To produce more quantitative analysis of tissue attributes, the Orringer group proposed the use of a multilayer perceptron (MLP) for interpreting and classifying the SRS microscopy images automatically with an appropriate diagnosis. Multi layer perceptrons are feed forward neural networks with a minimum of three layers, an input layer, one or more hidden layers and an output layer. Each node except for those in the input layer, are neurons that utilise a nonlinear activation function. The inputs to the activation functions are updated with weights and biases throughout training,

where the weights control how sensitive each neuron is to changes in the input and the biases control the threshold for activation. In an MLP, attempting to classify between more than two classes, the output layer has as many nodes as there are classes. The neurons in this layer utilise a softmax activation function (eq 6) to transform the raw output values into a probability distribution, ensuring that all neurons have values between 0 and 1, and their sum adds to 1. In a supervised manner the network is trained to produce these probability values for each of the final nodes based on the loss between the output probabilities and the known labels for each input by back-propagation the difference between the supervised label of an input with the predicted output neuron.

$$\sigma(\mathbf{z})_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}} \quad \text{for } i = 1, ..., N \tag{6}$$

The Orringer group generated a labelled dataset of SRH images. This dataset was generated by imaging biopsies from neurosurgical patients undergoing tumour resection, imaged immediately following removal, via an SRS microscopy system at the 2845cm^{-1} and 2930cm^{-1} wavenumbers. SRH images were generated following the procedural methods developed by the Xie group for virtual staining explored in section 2.3.2. The data for training the MLP was a subset of the total 101 neurosurgical patients with a total of 12,879 $400\mu \times 400\mu$ SRH fields of view (FOV). Of these 30 patients were used for training the MLP. A sample of a mosaic with visible FOV grid lines can be seen in Fig. 6 (a).

The Orringer group utilised a well-known image processing library "WND-CHRM" to extract a total of 2919 image attributes from each FOV. These features (included statistical features such as the mean & variance), texture features calculated from co-occurence matrices (including contrast, correlation, energy and homogeneity), geometric features (such as area, perimeter and compactness of an image region), and wavelet features (that capture frequency information at different scales). Normalised image attributes were fed to the multi-layer perceptron for training which utilised the leave-one-out approach, thus maximizing the size of the training set and mitigating

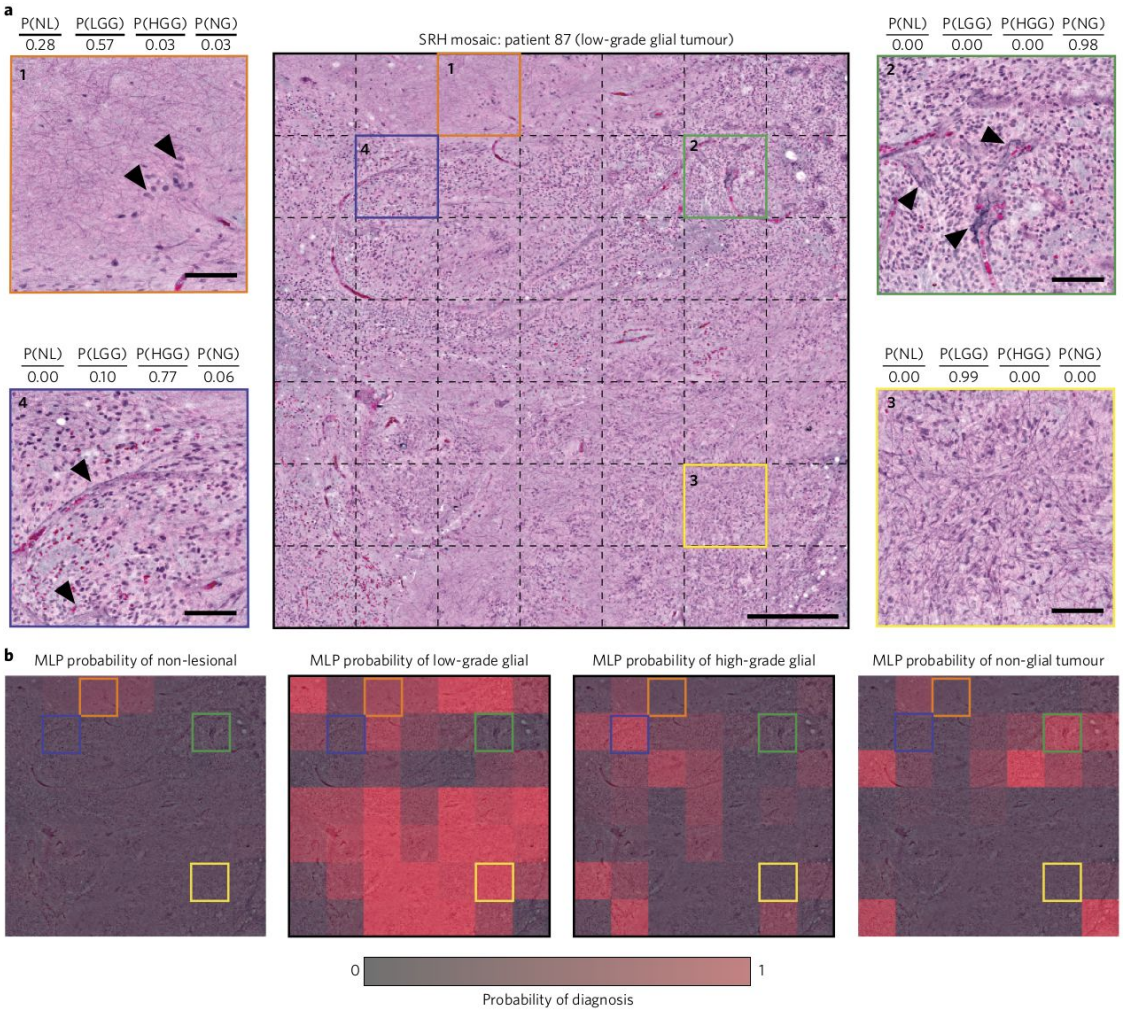

Figure 6: (a) An SRH mosaic with its corresponding field of view (FOV) grid-lines. Four enlarged, color-coded sample FOVs (1-4) are shown either side of the mosaic, highlighting specific regions of interest. Each enlarged FOV is accompanied by the diagnostic class probabilities: P(NL) for non-lesional, P(LGG) for low-grade glial, P(HGG) for high-grade glial, and P(NG) for non-glial tumors. (b) The four heatmaps represents the overall diagnostic classification probabilities across the entire SRH mosaic. The four color-coded FOVs from (a) are also indicated on each heatmap. In this particular case, the modal diagnosis can be seen by eye and is correct as low-grade glial tumor. This figure was reproduced from Fig 6[78].

risks of overfitting. The final output of the multi-layer perceptron was a probability vector where each neuron in the output layer gave a value for the learned probability that the input data belonged to the diagnostic classes: non-lesional, low-grade glial, high-grade glial or non-glial tumour. The authors highlight these four classes in particular as being particularly significant in decision making during brain tumour surgeries.

Fig. 6 (a) Displays the probabilities of 4 FOVs extracted from a larger stimulated

Raman histology mosaic with associated probability values for each of the four key diagnoses. In this instance the SRH is taken from a sample with low-grade glial tumour. Where some patches show significant probabilities of the low-grade glial tumour diagnosis, such as Fig. 6 (a-3), some of the individual patches show strong confidences to the contrary, for example Fig. 6 (a-2), indicating a strong probability of non-glial tumour, while Fig. 6 (a-4) indicates a strong probability of a high-grade glial tumour. To ensure accurate diagnoses after all patches are classified by the multi layer perceptron those with a probability of less than 25% are discarded, following which the modal diagnosis is used to determine the diagnosis for the overall mosaic. Fig. 6 (b) Shows the heatmap for a low-grade glial tumour sample with respect to each of the four diagnostic classes. While there are patches classified with high probabilities of being non-glial or high-grade and some lower probabilities of non-lesional, it is apparent by the heatmap that the low-grade glial is the dominant diagnosis in this case.

Overall this study provided a compelling proof of concept on the applicability of MLPs for automated diagnosis, with an accuracy of 100% in classifying lesional from non-lesional samples, and 90% in classifying glial from non-glial tumors. The predictions of the MLP were well in accordance with that of the pathologists survey with a 90% correspondence. The MLP did misclassify a number of cases due to subtle differences in tissue, or limited tumor presence in the analysed FOVs. These examples highlight the challenge in accurately classifying cases with minimal abnormalities, or where there is an uneven distribution of diagnostically relevant features across all FOVs. Nevertheles this paper represents a significant first step in automated SRH diagnosis, with further machine learning methods to follow.

2.5.3 Random Forests Decision Trees[79]

While multi-layer perceptrons have proven valuable in diagnostic applications, their inherent "black-box" nature presents a challenge for interpretability. The complex interplay of interconnected nodes and weighted connections within MLPs often ob-

scures the underlying decision-making process, making it difficult to understand how specific inputs contribute to the 'diagnosis'. In contrast to this, decision trees offer a transparent model of the decision making process. They achieve this by providing a visual representation of the tree structure clearly delineating the decision nodes and the specific conditions employed at each step to arrive at a classification. Decision trees are supervised learning algorithms often used for either classification or regression. They work by partitioning data into subsets based on the attribute values, with the end goal of homogeneity within each subset. The partitioning process is represented as a tree with nodes and branches, each branch representing the outcome of a decision and the final leaf nodes representing the predicted outcome. While simpler decision trees are readily interpretable, random forest comprises multiple decision trees in which each tree is trained on a subset of the data, generated by using a method called boot-strapping, in which a new dataset is constructed by randomly sampling the initial dataset, thus allowing for duplicates and missing data points. The average output value from all trees in a random forest is used to calculate the final prediction (modal for discrete predictions and arithmetic mean for continuous). This can lead to intricate interactions between the trees, potentially obscuring the overall decision logic. However, random forests still offer insights into the model's behaviour by revealing the relative importance of different features and the general patterns used for classification. Recently random forest decision trees have been applied to SRH images.

Orringer et al. acquired 25 SRH images at two wavenumbers (2845 cm^{-1} and 2930 cm^{-1}). Computational staining was performed using the methods described in Section 2.3.2. "CellProfiler" software was employed to segment individual nuclei and tumor-associated macrophages (TAMs) in the SRH images based on morphological features such as size, shape, and intensity. Following segmentation, CellProfiler was further utilised to extract key histological features, examples of which are illustrated in Fig.6(a)[1] of[79] . These features included nuclear density, TAM density, and a vari-

[1]Could not be reproduced due to copyright concerns

ety of nuclear morphology descriptors: area, perimeter, eccentricity, minimum and maximum Feret diameter, compactness, solidity, form factor, extent, orientation, and maximum radius.

The authors used these extracted features to train two random forest classifiers; The first for classifying whether a sample is normal or lesional tissue, and the second for classifying between high-grade and low-grade tissue. In total 25 full SRH images were used for the model development and validation using 10-fold cross validation, whereby the data is divided into 10 equal parts and a model is trained on 9 parts of these and tested on the remaining part. This process is repeated 10 times, with each part used for testing once. The "caret" package in R was used for model training and the 10-fold cross-validation. The "mtry" parameter in a random forest controls how many features are randomly considered at each node while building a decision tree. It is tuned by evaluating the out-of-bag accuracy, i.e. by testing each tree's prediction on the data that was not included in the bootstrapped dataset used to train it. Of the two finalised random forest models, the normal/lesional classifier had a 93.8%±2.2% accuracy, and the low-grade/high-grade classifier had a 89.4%±1.9%. Once each patch was classified, the predicted diagnosis for a whole slide was calculated based on the modal diagnosis across all FOVs. An example of this is given in Fig.7 (a)[2] where the heatmaps for both predictive decisions are given with 53.1% of FOVs classified as lesional tissue, correctly predicting the given slide.

While the classifications themselves have utility for differentiating diagnostic categories, the authors also exploited the interpretability of the random forest models. By analysing feature importance, they identified the most significant features contributing to the classification process. This information can guide further analysis or the development of new machine learning methods. Gini impurity was employed as the metric to calculate node impurity. The most salient features identified were nuclear density, TAM density, nuclear compactness, and maximum radius.

[2]Could not be reproduced due to copyright concerns

2.5.4 Deep Convolutional Networks[80]

Similar to the Blocklitz condGAN generator discussed in Section 2.4.2, which facilitated the transformation of 2-3 channel Raman/SHG data into histologically relevant H&E-appearing images, Ji et al. employed a similarly inspired U-net autoencoder architecture. Their objective was to translate between two Raman modalities, specifically converting single-shot femto-SRS spectral images to dual-channel pico-SRS images. This conversion effectively recovers chemical contrast without necessitating complex optical engineering modifications. This process is illustrated in Figure 7(a)&(b), which presents amplitude graphs and side-by-side comparisons of ground truth and predicted values.

It is important to understand the distinction between these two modalities. Femto-SRS, employing a faster laser pulse, captures a broad range of wavenumbers and compresses them into a single intensity value. This provides a composite but less specific view of the sample's chemical composition. Conversely, pico-SRS acquires hyperspectral images by recording intensities at specific wavenumbers (e.g., 2845cm^1 and 2930cm^{-1}). These wavenumbers correspond to distinct molecular vibrations, enabling the extraction of detailed chemical information and the generation of separate images for each channel. The U-net model developed by Ji et al. was trained on paired femto-SRS and pico-SRS images of live HeLa cells and fresh gastric tissues. This training enabled the model to effectively recover chemical contrast from the femto-SRS images.

The U-net generated 2 channel femto-srs predicted image, along with a separately recorded SHG image, was used to procedurally generate an SRH image employing methods similar to those described in Section 2.3.3 & 2. This SRH image was primarily intended for validation by histopathologists, as the U-net outputs were used without this transformation in the training of the deep convolutional neural networks.

The Ji group employed the Inception-ResNet-v2 architecture for classification tasks.

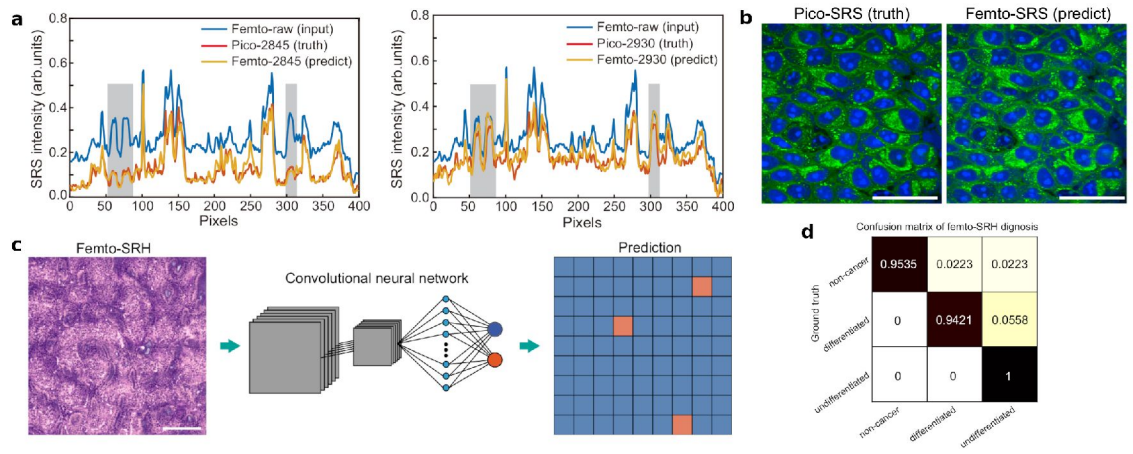

Figure 7: (a) Comparison of pixel intensity profiles for a single row: 2845 cm^{-1} (left) and 2930 cm^{-1} (right). The profiles correspond the raw femtosecond SRS image (blue), the picosecond SRS ground truth (red), and the predicted values from the femtosecond SRS (yellow). (b) Comparison of the picosecond SRS false-color composite ground truth (left) and the image generated from the autoencoder-predicted values for the femtosecond SRS image (right). (c) Schematic illustration of the classification process, where a SRH image is fed into a deep CNN ending in a softmax dense layer to generate predictions. (d) Confusion matrix summarising the performance of the classification model. This figure was reproduced from Fig 2 5.[80].

This architecture represents an advancement over the Inception family by integrating residual connections within the parallel Inception modules, inspired by the ResNet architecture. These residual connections effectively address the vanishing gradient problem commonly encountered in training deep neural networks by facilitating smoother gradient flow throughout the network. This characteristic is particularly beneficial for the Inception-ResNet-v2, which consists of 572 layers, organised into distinct blocks. The input layer initially receives the raw data, which is subsequently processed by the stem layer performing initial convolution operations for feature extraction. The first significant block encountered is the Inception-ResNet-A, designed to conduct a series of parallel convolution operations, the outputs of which are concatenated. A reduction layer follows, downsampling the extensive feature set via pooling layers. This sequence is succeeded by the Inception-ResNet-B and a reduction block, leading into the Inception-ResNet-C block. Subsequently, an average pooling layer is applied to aggregate the learned features spatially before feeding these values to the fully connected layer, which comprises 1536 neurons. Although

not explicitly stated by the authors, it is plausible that the final layer consisted of a single neuron with a sigmoid activation function for binary classification or two neurons with a softmax function for the two classes.

For the training step, the Ji group trained two separate classification networks on femto-SRS images from 279 patients. These images were split into 300x300 FOVs and fed to the network returning a classification at the FOV level as illustrated in Fig. 7 (c). The image dataset was split into training/validation and testing with a 7:3 ratio, where the training/validation set was further split into a 4:1 ratio for training to validation. The model employed five-fold cross-validation for training and validation with cross entropy loss as the loss function. Whole slide images were classified based on thresholds from the predictions on the FOVs, with a threshold of 0.476 for cancer/non-cancer and 0.505 for differentiated/undifferentiated each determined by Youdens index.

The final confusion matrix for these networks is given in Fig. 7 (d), showing that no false-negatives were returned for the cancer diagnosis and a respectable false positives at a rate of 2.23%. Similarly for differentiated/undifferentiated cancer, the more aggressive undifferentiated cancer was never labelled as differentiated and a small proportion of the differentiated were misclassified at a rate of 5.58%. The CNNs were near-perfect in determining cancerous lesions with an accuracy of 96.4%, and of particular note were capable of outperforming the pathologists at differentiating subtypes with a 94.1% accuracy compared to an 85.3% mean accuracy for the pathologists.

2.6 The Future is Broadband

As we look ahead, this section presents a vision for the future of hyperspectral Raman imaging and histology. Reflecting on previous chapters, most discussions have centred around Stimulated Raman Histology, focusing on high-speed imaging using one or two SRS wavenumbers to generate H&E like images. These images were either interpreted directly by pathologists or used as inputs for machine learning

classifiers, bypassing the need for a traditional histopathological review. This significant capability, derived from just two wavenumbers, sets the stage for a broader application spectrum.

There have been some recent advancements in hyperspectral Raman imaging that enable more chemical information to be recorded, extending beyond the two wavenumber model. This additional information could potentially enhance the SRH approach for improved automated diagnostics. One method is to scan the SRS pump wavelength during acquisition enabling multiple Raman wavenumbers to gather a broad spectral dataset quickly. This approach allows for the rapid acquisition of a hyperspectral cube where each pixel contains a full spectrum. The Graham group have applied phasor analysis to simplify this complex data into a more manageable form by transforming the time consuming multi-dimensional spectral data into a two dimensional graphical representation.[81] This transformation aids in the quick identification of chemical components within the tissue, enhancing drug and disease diagnosis by pinpointing specific cellular interactions and metabolic processes.

While the pump scanning approach is innovative it has the disadvantage of requiring both scanning in wavelength and space, which significantly limits the overall scan time. Recently, the introduction of parallel lock-in amplifiers (LIAs), as pioneered by the Polli group,[50] marks another breakthrough in the field. By employing 32 parallel LIAs, the system can simultaneously acquire signals across multiple channels, dramatically reducing scan times while covering a broad bandwidth. This setup not only speeds up the imaging process but also broadens the range of chemical information that can be captured in a single scan. The ability to process multiple wavelengths concurrently allows for a more detailed and comprehensive analysis, facilitating a deeper understanding of tissue chemistry at a much faster rate than previously possible.

Raman-based diagnostics have been explored for decades across various diseases but have thus far failed to penetrate into the clinic[82,83,84,85,86]. In this context the recent success of SRH in terms of clinical application cannot be overstated. The principle

of traditional Raman diagnostics involves recording a broad Raman spectrum from one or more points on known diseased and healthy tissues or cells, typically spanning both the fingerprint and the high-wavenumber CH-band. Due to slow acquisition times, only a few spectra are recorded per sample and these spectra have traditionally been used to train machine learning models—commonly PCA/LDA, PLS, or SVMs—to create multivariate classification systems that can predict diseases from a single spectrum. Despite the successful demonstration of this approach in numerous studies, it has yet to gain clinical traction, primarily due to slow acquisition times, the expense of the equipment, and difficulties in cross-instrument calibration. However, recent advancements have significantly mitigated these issues, particularly in calibration and cost.

This brings us to an important distinction: traditional H&E pathology—and by extension, current SRH methods as reviewed above—rely primarily on spatial, morphological classification, whereas the full spectrum Raman approach offers a chemical classification of disease. The potential to combine these methodologies opens up exciting new possibilities for integrated diagnostic platforms. The focus of Chapters 3 and 4 which follow this one, is to contrast and combine the slower but higher quality spontaneous Raman platforms with their much higher speed BCARS system.

Looking forward, we envision a hybrid approach where single or dual wavenumber SRH is used to rapidly produce a virtually stained image. Alternatively high-speed full bandwidth BCARS could be employed as the first step. Simple spatial-based algorithms (such as random forests, MLP, or CNN as described earlier) could then highlight areas likely affected by disease, and these regions could subsequently be analysed using point-based Raman spectroscopy for detailed chemical classification. This method would leverage the speed of SRH/BCARS for initial screening and the chemical specificity of Raman spectroscopy for precise diagnostics. Given that the SRH/BCARS system could easily be adapted to include the spontaneous Raman modality (see Chapter 4), this approach seems particularly viable.

Similarly, another forward-looking approach could involve using SRH for initial high-

speed imaging followed by BCARS hyperspectral scanning of suspect areas. As described in Section 2.2 BCARS (ms acquisition time) offers a compromise between the slow acquisition speeds of traditional Raman spectroscopy BCARS (s acquisition time) and the rapid imaging capabilities of SRS microscopy BCARS (us acquisition time). With acquisition times measured in milliseconds and the ability to record a full broadband spectrum ($0 - 4000cm^{-1}$), BCARS is well-suited for detailed hyperspectral imaging of smaller regions of interest which could first be identified using SRH. These regions could then be analysed using advanced multivariate classification models, potentially utilising 3D Convolutional Neural Networks (CNNs), to provide a comprehensive diagnostic assessment that incorporates both morphological and chemical insights.

2.7 Conclusion

Hyperspectral imaging and histology with Raman spectroscopy is a transformative technology in medical diagnostics, which enables non-invasive, rapid pathological assessments. The advancements in this field underscore a significant shift towards precise, molecular-level medical diagnostics. The advent of virtual staining techniques in particular stands out as a pivotal innovation, eliminating the need for chemical stains in tissue analysis. This method preserves tissue integrity while providing critical morphological insights. It has proven to be a valuable tool in surgical settings, offering real-time guidance to surgeons and pathologists, thus enhancing the precision of surgical interventions and potentially improving patient outcomes.

The integration of machine learning algorithms with hyperspectral imaging technologies further advances this field forward. These algorithms can analyse complex datasets, identifying patterns and features that are not readily visible, thereby augmenting the diagnostic capabilities of medical professionals and supporting the rapid decision-making required in clinical environments. Future directions in this field are likely to focus on refining these technologies and exploring their integration into routine clinical practice. This could involve the development of more sophisticated

and less expensive imaging systems that offer even higher resolution and deeper tissue penetration, or algorithms that can provide more nuanced interpretations of hyperspectral data.

It is the opinion of these authors that the future of hyperspectral imaging and histology lies in the seamless integration of morphological and chemical diagnostics, harnessing the strengths of high-speed coherent Raman Histology and spontaneous Raman spectroscopy. This integrated approach could revolutionise pathology by providing rapid, accurate, and detailed diagnostic information, paving the way for more personalised and effective medical treatments. Although the combination of SRS and spontaneous Raman may offer the optimal pathway in terms of speed, the combination of BCARS with Raman offers the unique ability to perform *multivariate spectral segmentation* of tissue regions for follow on Raman probing, which is the motivation of the coming chapters.

3 Direct Comparison of BCARS with Spontaneous Raman

The work presented in this chapter is based on the conference publication titled "Comparison of broadband-CARS and spontaneous Raman scattering for blood cell analysis"[2], which was presented at the SPIE 2024 conference held in Strasbourg, France.

3.1 Overview

The objective of this chapter is to contrast the spectra obtained from Spontaneous Raman (SR) Spectroscopy with Broadband Coherent Anti Stokes Raman Spectroscopy (BCARS) in relation to biospectra. To this end a particular white blood cell, specifically a plasmacytoid dendritic cell line known as CAL1, is employed. While this comparison will ultimately focus on the spectrum obtained form both methods, and the common chemical components that can be identified by both methods, there will also be an emphasis on contrasting their optical implementations and processing methodologies. This work aims to describe the processing techniques employed in both spectroscopic methods, including data acquisition, spectral pre-processing and data analysis algorithms. Building on this, this work also seeks to distinguish the distinctive features of the spectra obtained from SR and BCARS, highlighting the differences in signal-to-noise ratio (SNR) alongside spectral and spatial resolutions. This study endeavours to expose the strengths and limitations of each technique and the resulting spectra. Additionally included in this work is obtaining a single high-SNR representative BCARS spectrum from a blood cell, by applying an automated segmentation of the hyperspectral image obtained following BCARS raster-scanning, followed by averaging the spectra corresponding to the area that represents the cell segment.

3.2 Introduction

In the previous chapter we reviewed a series of Raman stimulating and sensing methods. We noted that of these methods SR and BCARS were the only two methods that were capable of viably recording the full spectral window in a single acquisition for biological contexts. In this chapter, we will compare the use of BCARS and SR for cellomics. An important application of cellomics is the sorting of cells and their quantification, which is central to immunology. Conventional sorting methods use a fluorescent label and morphological information such as cell shape, size and density[87]. The introduction of exogenous fluorescent or other labels often perturbs the biological function to some extent such as modifying local protein activity or precludes follow-on spectroscopic or other analysis. Therefore, a control study is often required to study the effect of labelling[88]. It is, therefore, attractive to obtain a classification of a cell through unlabelled means.

The classification of cells based only on differences in chemistry has been proven to be possible using the molecular fingerprint obtained through Raman micro-spectroscopy[89,90]. In this method, the cells are probed using a Raman spectrometer coupled to a microscope system (Raman micro-spectrometer), whereby the sample is moved (raster scan) or the focus is moved (focus scan) over the sample. The vibrational information from the cells biochemistry in the form of a Raman fingerprint is, therefore, enough to perform label-free cytometry, in principal. In practice, this method requires on the order of a few seconds per single acquisition due to the low probability of inelastic scattering present in SR[91] and thus is limited in terms of clinical application, where on the order of thousands of cells are studied per test. To overcome the slow nature of SR micro-spectroscopy, BCARS can be used. In BCARS, the vibrational modes are probed using a nonlinear process which allows coherent generation of the Raman signal over a volume, thus enhancing the signals by several orders of magnitude[92], compared to conventional Raman spectroscopy. BCARS as its name suggests probes the whole vibrationally resonant region (500-3500 cm^{-1}) in a sample simultaneously using broadband laser pulses.

The broadband and extremely high signal strength obtained with BCARS (albeit with lower spectral SNR than SR) is thus ideally suited to enhance cell identification and sorting speeds.

BCARS as a spectroscopy has been shown to provide tissue imaging capabilities at speeds of 3.5 ms per spectrum[93,94]. Its application to single-cell studies thus far however has been limited. In terms of BCARS, there are two main differences between tissues and cells, molecular density and chemical similarity. To date, BCARS has performed remarkably at identifying different structures within histological samples, such as collagen, protein and lipids. This is due to the fact that often a distinct vibrational mode is abundant in each structure, and its intensity can be used to infer the concentration of each group. In single cell spectroscopy, there can be significantly less material to probe due to the thin morphology which can be more challenging[95]. While many BCARS studies have focused on thicker tissue samples (30-60μm) which contain high concentrations of oscillators, single cells provide a more challenging composition for spectral recording. This is due in part to the fact that BCARS is a non linear effect which scales non linearly with the concentration of oscillators, leading to a potentially much weaker signal in thin samples compared to SR or SRS. In addition, the interference between the resonant signal and the non resonant background produce lineshapes that do not directly reflect the clean spectra typically obtained from SR and SRS. As part of our comparison we use CAL-1 blood cells, an example of a thin cell sample where the SNR is expected be low and contrast it directly against the SR spectra. The sample preparation details for which are given below.

3.3 Sample prep

We acknowledge the contribution of Dr. Marion Butler's team in preparing the samples for this study in the Biology Department of Maynooth University. Further we acknowledge the work of Dr. Ryan Muddiman and others in the construction of the SR and BCARS systems which were utilised and which are described in brief

detail below.

Suspended blood cells were centrifuged at 500xg and supernatant removed. The pellet was resuspended in 5mL of cRPMI, followed by cell counting via dilution of 50μL in 200μL of cRPMI. A total volume of 1mL in PBS containing 10 x 10^5 cells in centrifuge tubes is then prepared. Centrifugation was performed at 500xg at 4°C for 5 minutes, followed by decantation of the supernatant. The resulting pellets were resuspended in 1mL of 0.2μm filtered 10% formalin and allowed to incubate for 10 minutes at room temperature. Further centrifugation at 500xg at 4°C for 5 minutes, with subsequent supernatant decantation and resuspension in 2mL of PBS. A second centrifugation step was then executed, followed by supernatant decantation and addition of 300μL of molecular-grade water. Finally, 30μL of the cell suspension was deposited onto a coverslip and air-dried at room temperature for 2 days and a brightfield image of this sample is shown in Fig. 8

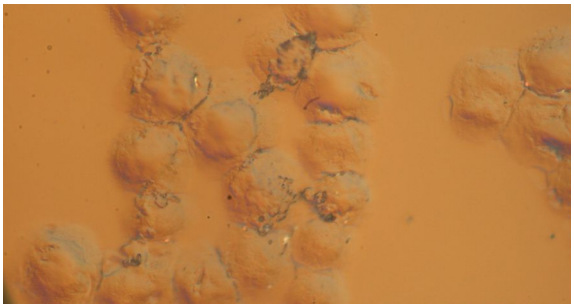

Figure 8: Brightfield image of Cal-1 plasmacytoid dendritic cells on a glass coverslip.

3.4 Raman Spectroscopy

Raman spectra from the cells were recorded using a custom-built Raman micro-spectroscopy system, as described in Ref. 96 which was built as part of a previous study. In summary, the system features a 150 mW laser with a 532 nm wavelength, a spectrograph (Kaiser, Holospec f/1.8i) alongside a low-noise cooled CCD camera for spectrum recording operating at -80°C offering a bandwidth of -34-2517 cm^{-1} and an average resolution of 5.48 cm^{-1}. The system integrates with a fully automated inverted microscope (Olympus IX81) which utilizes an Olympus UMplanFl

40/0.75NA objective and a confocal aperture to isolate signals from the glass substrate minimizing background noise. Following deposition of the blood cells onto the coverslip, the cells were individually subject to Raman spectroscopy. All spectra were recorded with approximately 70 mW of laser power and with an acquisition time of 30s. One spectrum was recorded per cell.

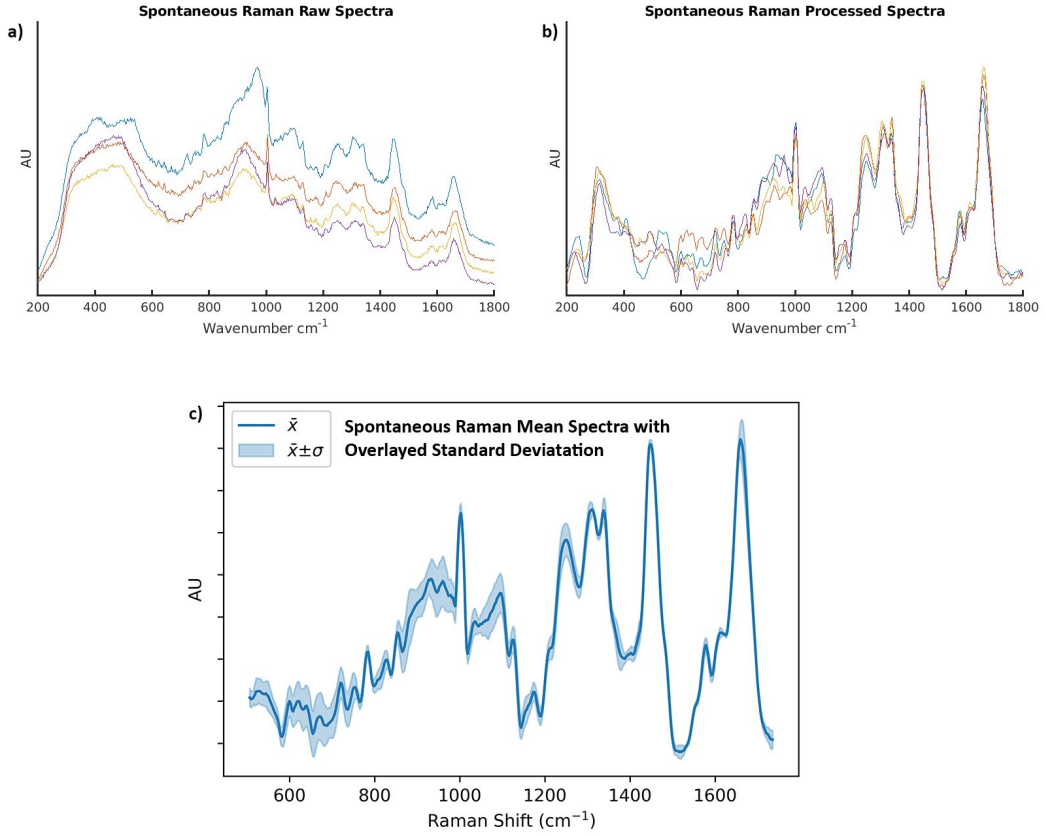

Figure 9: Diagram showing spontaneous Raman spectra (a) raw spectra (b) the same spectra following preprocessing to remove the glass spectrum and to denoise as described in the main text, and (c) the mean spectrum and the standard deviation of the dataset.

Cosmic rays were removed from all recorded spectra using the method described in Ref. 59. Raman spectra were intensity calibrated using the method described in Ref. 97, which first involves wavelength calibration using the method described in Ref. 98 followed by determining the wavelength dependent sensitivity of the Raman spectrometer by recording the spectrum of a NIST-calibrated white light and comparing the resultant spectrum with the spectrum provided by NIST. Each

recorded spectrum is intensity calibrated by correcting the intensity values using this sensitivity. Wavenumber calibration was implemented using the method described in Ref. 99, having first recorded a spectrum of 4-acetamidophenol. Denoising of spectra was performed using a Savitsky-Golay smoother with a polynomial order of 3 and a window size of 7^{100}. Finally baseline subtraction was achieved by removing the glass contaminant background spectrum and a three order polynomial using the method described in Ref. 45. The raw and processed spectra are shown in Fig. 9.

3.5 Hyperspectral - BCARS

BCARS is a four-wave mixing process whereby a pulse pair termed the pump and stokes at frequencies ω_{P} and ω_{S} excites a vibrational coherence in the molecule at the resonant frequency $\Omega = \omega_{\mathrm{P}} - \omega_{\mathrm{S}}$. A third pulse termed the probe at frequency ω_{pr} scatters from this coherence and blue-shifts to the anti-Stokes frequency $\omega_{\mathrm{as}} = \omega_{\mathrm{P}} - \omega_{\mathrm{S}} + \omega_{\mathrm{pr}}$. In BCARS, the third-order susceptibility $\chi^{(3)}$ mediates the nonlinear effect, and contains the vibrational information of the sample. As a coherent nonlinear process, efficient signal generation requires the interacting beams to spatially, temporally and geometrically align.

Additionally, the input fields can mix to produce non-resonant light at all frequencies governed by momentum conservation due to the instantaneous electronic polarization. This produces a background signal that slowly varies and which interferes with the anti-Stokes signal which has a phase variation that responds to the resonances within the sample as shown in eq11. The total signal measured can therefore be described by the total third order susceptibility which is broken into two parts. (i) The resonant signal, a complex quantity that contains molecular vibrational information wherein the imaginary part, corresponds directly to the spontaneous Raman spectrum. (ii) The NRB, a chemically ambivalent signal primarily originating from the electronic response of the material described by the purely real value.

$$I_{CARS} \propto \left|x^{(3)}\right|^2 = \left|x_{NR}^{(3)} + x_R^{(3)}\right|^2 \tag{7}$$

Expanding this into it's real and imaginary components gives.

$$I_{CARS} \propto |x_{NR}^{(3)} + Re(x_R^{(3)}) + i.Im(x_R^{(3)})|^2 \tag{8}$$

$$I_{CARS} \propto (x_{NR}^{(3)} + Re(x_R^{(3)}))^2 + Im(x_R^{(3)})|^2 \tag{9}$$

Given the nonresonant signal is much stronger than the resonant signal, the resonant components that are not scaled by the nonresonant become negligibly small in comparison to those that are.

$$I_{CARS} \propto (\chi_{NR}^{(3)})^2 + 2 \cdot \chi_{NR}^{(3)} \cdot \mathrm{Re}(\chi_R^{(3)}) + \underbrace{(\mathrm{Re}(\chi_R^{(3)}))^2 + (\mathrm{Im}(\chi_R^{(3)}))^2}_{\text{Non multiplicative component ignored}} \tag{10}$$

As such the final signal can be given in two components, the resonant signal multiplicatively amplified by the nonresonant alongside the purely non resonant strong background.

$$I_{CARS} \approx \underbrace{(\chi_{NR}^{(3)})^2}_{\text{Strong NR background}} + \underbrace{2 \cdot \chi_{NR}^{(3)} \cdot \mathrm{Re}(\chi_R^{(3)})}_{\text{The amplified signal}} \tag{11}$$

This demonstrates how the nonresonant component of the BCARS process simultaneously generates a significantly stronger resonant signal while obscuring that resonant signal with a multiplicative background. This background, being generated with the signal and being coherent with it, make it impossible to remove by baseline subtraction and therefore requires advanced numerical post processing methods such as the Kramers Kronig technique which we employ.

Hypersepctral BCARS images were obtained using the opto-electronic system described in Ref. 53 and illustrated below in Fig. 10. The BCARS system is briefly described as follows: The source is a FemtoFiber PRO MOPA mode-locked laser system at 1550 nm which generates a narrowband probe pulse. The seed pulse is coupled into a highly nonlinear fiber generating the Stokes pulse covering 900–1400

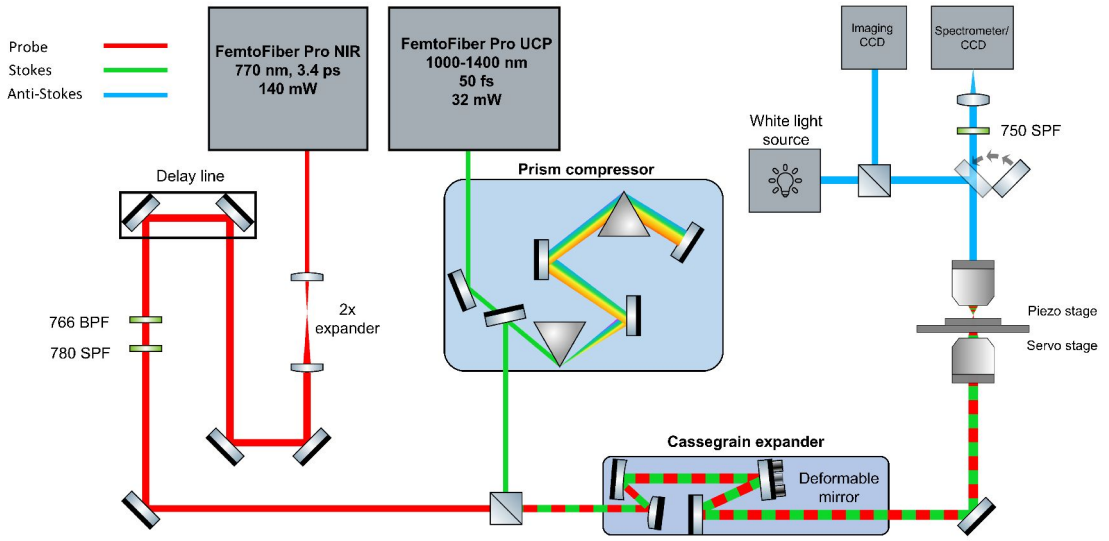

Figure 10: Diagram of BCARS setup. See Ref. 53 for more details.

nm. Group velocity dispersion is compensated by an external prism pair pulse compressor constructed with two SF10 prisms. The microscope setup includes an electronically controlled piezo based XYZ stage (Precision Instruments), facilitating precise sample manipulation at nanometer accuracy and at high scanning rates. Overlapping of the beams is achieved using a dichroic mirror and both beams were focused into the sample using a high numerical aperture (NA) objective, while the transmitted light was captured with a second objective. The filtered anti-Stokes light is directed into an Andor Shamrock 500i-A spectrograph and captured by a Peltier-cooled CCD camera. The spectrograph is equipped with a 300 lines per millimeter grating and utilizes full vertical binning to cover a broad recording range from 443 to 4489 cm^{-1}. NRB removal from the BCARS spectra was implemented using the method described below. All BCARS images were denoised using the truncated SVD prior to NRB removal.

3.6 BCARS Preprocessing

The piezo stage was used to achieve micron accuracy when raster scanning the sample through the pump and stokes lasers to capture molecular vibrational information

within the CAL-1 cells. Each spatial location was exposed for approximately 3.5ms, orders of magnitude faster than the spontaneous Raman. The stage moved through its full range of motion (200x200μm) at steps of $\frac{1}{2}\mu$m, constructing a hyperspectral datacube of dimensions (400,400,1000) where the 1000 samples are distributed approximately uniformly over the bandwidth defined in the previous subsection. In total we acquired spectra from an effective 400^2 independent spatial locations, each with a spectrum of 1000 values producing the 3D dataset.

While BCARS typically provides a stronger signal due to it's coherent excitation, in particular in contexts with fewer oscillators the NRB becomes increasingly problematic as it continues to act as an amplifier however the contribution from noise increases over signal. As such a distinct preprocessing workflow was applied. Beginning with a generalized Anscombe transform to account for Poisson noise. This transform was used to map the noise to an approximate Gaussian distribution which is more amenable to traditional noise reduction techniques. The Anscombe transform is defined below where x is the raw photon count, g is the sensor gain, μ is the read noise and σ is the standard deviation of the read noise.

$$\mathrm{A}(x, \mu, \sigma, g) = \frac{2}{g}\sqrt{\max\left(gx + \left(\frac{g^2 3}{8}\right) + \sigma^2 - g\mu, 0\right)} \tag{12}$$

where $(gx + \left(\frac{g^2 3}{8}\right) + \sigma^2 - g\mu)$ represents the expected value of the Poisson noise.

Singular value decomposition (SVD) is a a technique that breaks down a high dimensional dataset into it's most fundamental rank ordered components. Each component describes a fundamental pattern contained within the data, with the top ranking (highest singular values) describing the most variance. This is effective at removing noise due to the fact that noise components tend to be random rather than consistent, therefore as the sample size increases, the pattern (non random components) repeat more frequently than any noise source, thereby becoming statistically stronger (higher ranked) and easier to distinguish from the random, inconsistent fluctuations. This ranking system makes SVD particularly suited for denoising in

situations with many samples. The signal can be captured by the first N singular values and corresponding vectors. While the noise tends to be relegated to smaller components. Reconstructing the BCARS spectra from the first N singular values has understandably shown particular efficacy with hyperspectral BCARS data[101] where consistent resonances reoccur.

We employed the SVD method, decomposing the dataset into three matrices U, Σ, V^T. Here, V describes the principal directions in the data and is found by solving the eigendecomposition of $X^T X$ where X is the matrix containing the original data. Similarly U is found by solving the eigendecomposition of XX^T. The square roots of the eigenvalues from either decomposition are placed along the diagonal of Σ thus representing the the significance of each of the directions (Singular Values). The original dataset can be reconstructed by the product of these three matrices, however to exclude noise a threshold is applied to the singular values, retaining only the more dominant directions which correspond to more frequent patterns in data. The datacube is reconstructed using this reduced set of singular values and corresponding vectors, which effectively suppresses noise and improves the signal to noise ratio. Finally the inverse Anscombe is applied as follows where x is the previously transformed data, μ is the mean of the read noise, σ is the standard deviation of the read noise and g is the gain of the sensor. Thus returning the data to it's original physical scale from the stabilized domain used for the SVD.

$$A^{-1}(x, \mu, \sigma, \mathrm{g}) = \mathrm{g}\left(\left(\frac{\max(x,1)}{2}\right)^2 + \frac{1}{4}\sqrt{\frac{3}{2}}\left(\frac{\max(x,1)}{x}\right) - \frac{11}{8}\left(\frac{\max(x,1)}{x}\right)^2 + \frac{5}{8}\sqrt{\frac{3}{2}}\left(\frac{\max(x,1)}{x}\right)^3 - \frac{1}{8} - \sigma^2\right) + \mu \quad (13)$$

The SVD method excels at removing random noise in the recorded spectrum caused by shot noise and dark current in the detector (both governed by a Poisson distribution) as well as other noise sources such as read noise (governed by a Gaussian distribution). The recovered signal can now be subject to NRB removal methods. We recall that the NRB is a consistent slowly varying signal resulting from the third order susceptibility as described earlier in eq 11. To address this we employ

a well-known method, the Kramers-Kronig technique[102]. The Kramers-Kronig relations provide a mathematical connection between the real and imaginary parts of a complex, causal function $F(\omega)$ (where ω represents the angular frequency and $\mathcal{P}$ denotes the Cauchy principal value of the integral):

$$\mathrm{Re}[F(\omega)] = \frac{2}{\pi}\mathcal{P}\int_0^\infty \frac{\omega'\mathrm{Im}[F(\omega')]}{\omega'^2 - \omega^2}d\omega' \tag{14}$$

$$\mathrm{Im}[F(\omega)] = -\frac{2\omega}{\pi}\mathcal{P}\int_0^\infty \frac{\mathrm{Re}[F(\omega')]}{\omega'^2 - \omega^2}d\omega' \tag{15}$$

In the BCARS signal, the measured spectrum I_{CARS} is proportional to the squared modulus of the total complex susceptibility. Enabling retrieval of the phase from the modulus squared as shown by Liu et al[102].

$$\phi(\omega) = -\frac{P}{\pi}\int_{-\infty}^{+\infty} \frac{\ln|\chi(\omega'')|}{\omega'' - \omega}\mathrm{d}\omega'' \tag{16}$$

Application of a windowed Hilbert transform in combination with phase error estimation from a reference non-resonant sample can be therefore used to retrieve the complex susceptibility. This reference is taken from the glass spectra in the regions around the cells.

$$X^{(3)}(\omega) = \sqrt{\mathrm{ICARS}}\exp(i\phi(\omega)) \tag{17}$$

While this provides a reasonable analytic method, it is prone to errors due to the variability in the NRB across time and sample. In the samples presented here, the Kramers-Kronig technique was applied from the CriKit2 package developed by Camp et al[103] which use the above method as an initial guess, followed by incorporating additional corrections. This includes a phase error correction which iteratively removes baselines from the retrieved spectrum, and scale error correction to adjust the amplitudes of the Raman peaks to appropriate intensities. While these methods significantly minimize error due to the NRB varying from sample to sample, some residual errors still persist and therefore, some error can be expected in the recovered

spectrum.

3.7 Segmentation for Single Cell - Single Spectrum BCARS

To generate a mask seperating signal areas from the background a threshold was applied to the mean intensity of the signals. Each spectrum in the 3d datacube was flattened to a single scalar by taking the mean $\frac{\sum_{i=0}^{n}(x_i)}{n}$ where x_i is the i'th amplitude of the spectrum and n is the number of wavenumbers recorded in the spectrum. A threshold was applied to isolate cell areas from the background. This provided a 2d binary mask for clustering to operate on the basis of location. To segment individual cells from the mask generated previously, a kMeans clustering algorithm was used. The location values of all pixels that contained a cell were used for the clustering and can be denoted as

$$A = \{a_1, a_2, \ldots, a_n\} \tag{18}$$

where each a_i represents an (x,y) value with a positive return in the mask. The kMeans algorithm was initialized with K=77 centroids, denoted as $\mu_1, \mu_2, \ldots, \mu_K$ which were determined based on an initial estimate of the number of cells present in the image. This was estimated by simple division of the number of active cells with the average area of a cell. Each datapoint a_i was assigned to the nearest centroid μ_j based on Euclidean distance:

$$j = \arg\min_k \|a_i - \mu_k\|_2^2 \tag{19}$$

Each centroids location μ_k was then updated to be the mean of the location of the datapoints assigned to the cluster k as follows:

$$\mu_k = \frac{1}{|C_k|} \sum_{a_i \in C_k} a_i \tag{20}$$

where C_k denotes the set of data points assigned to a cluster k. The convergence criteria were set to a maximum of 300 iterations, with an early stopping condition

of cluster centres updating by less than 10^{-4} per iteration. Finally the average spectrum was calculated per cluster. Where for the given cluster C_k the average spectrum S_{avg}, k is the element-wise mean of all spectra $S(a_j)$ corresponding to the pixels a_j assigned to that cluster:

$$S_{avg,k} = \frac{1}{|C_k|} \sum_{a_j \in C_k} S(a_j) \tag{21}$$

In Fig. 11 the results of a segmentation process are illustrated. In Fig. 11 (a) the mean CARS intensity image is shown and in Fig. 11 (b) we show the corresponding image following the segmentation of cells using the aforementioned threshold. In Fig. 11 (a) the result of instance segmentation is shown using kMeans. In most cases the cell area is accurately segmented. These segments are each taken as clusters of spectra which are then averaged to produce a single representative spectrum for a given cell.

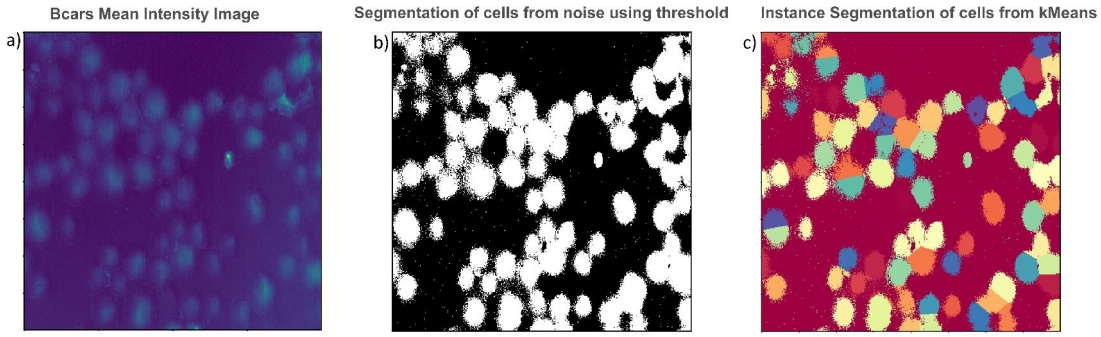

Figure 11: Diagram showing; Mean Intensity BCARS spectrum image(Left), masked image of cells(Centre), Image showing the clustered cells by colour contrasts(Right)

3.8 Results

Fig. 12 (a) displays the final individual cell spectra generated from the averaging of each spectrum in a given cluster. Where 12 (b) shows the standard deviation overlayed on the average of all the individual cell spectra. It is apparent both in the spontaneous Raman Figure: Fig 12 (c) and in the BCARS Fig 12 (b) that there

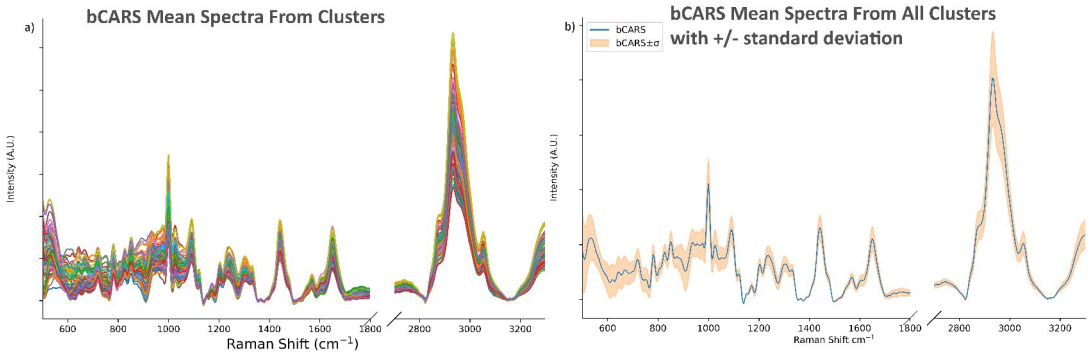

Figure 12: Diagram showing average spectrum per cell from BCARS, and overlay of BCARS Mean signal with the standard deviation

is more variability in the lower wavenumbers in the fingerprint region with respect to the relative signal intensities. The higher end of the fingerprint region, ranging from the CH2 deform peak and amide I peaks at 1443cm^{-}1 and 1655cm^{-}1 appear to have the least overall variance relative to signal intensity. This may be a result of the consistency of the respective chemicals that respond at these wavenumbers.

While there was evident differences in the amplitudes of the two spectrums (BCARS and Spontaneous), this is likely due to error in baseline correction most likely due to the Kramers Kronig method used for removing the NRB. As one of the main drawbacks of Kramers Kronig is the requirement of NRB reference spectra. Such reference spectra may not be purely real or perfectly flat due to absorbance or scattering, leading to errors in the retrieved Raman signal. The Kramers Kronig method may also introduce error in the baseline detrending parameters as part of the CriKit2 routine[103]. This requires manual optimization, as the error term varies between samples. Recent work has explored novel methods using Convolutional Neural Networks[105] and Convolutional Autoencoders[106] to improve upon the Kramers Kronigs method and better automate the overall process. These are of particular interest due to their unsupervised nature.

Alongside the baseline errors there is a noticeable drift in the wavenumbers position of the spectral lineshapes with respect to spontaneous Raman - see wavenumber region $>$ 1200 wavenumbers in Fig. 13. We posit two reasonable arguments for this

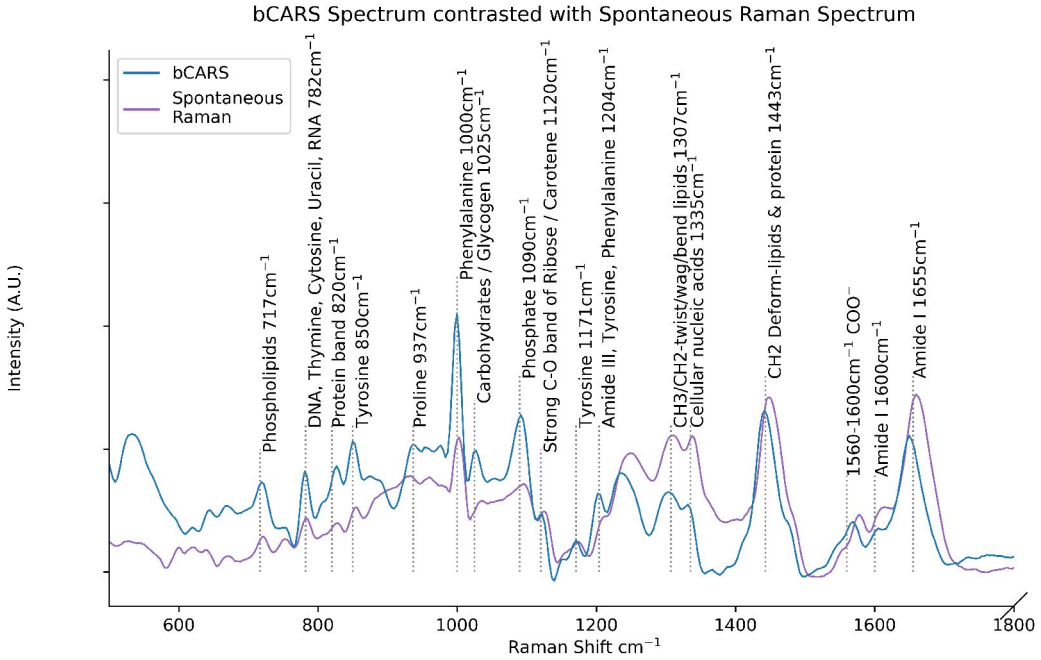

Figure 13: Diagram showing spontaneous raman, BCARS and previously identified peaks

drift. The first is that there may exist wavenumber miscalibration on the part of the Raman spectrum owing to error in wavenumber assignments outside the limits of the peaks of the wavenumber calibration reference material 4-acetaminophen. The second is that the shape of the BCARS spectrum may be inherently erroneous in this region, which corresponds to the weakest part of our NRB amplifier. Therefore the inverse problem that the KK-algorithms attempts to solve is more complex in this region and one could envisage that there are multiple acceptable solutions with different line-shapes, that accurately reproduce the very small undulation in the BCARS spectrum itself. This needs further study in order to ascertain which of these arguments, if not both, is true.

The congruence between the peaks identified in BCARS and spontaneous Raman, demonstrated in Fig. 13 underscores a significant correspondence of spectral features across both techniques. These findings are supported by the existing literature in terms of peak identification[104] shown in Table.2. The consistent observation of peaks

Material	**Wavenumber**
Phospholipids	717
DNA, Thymine, Cytosine, Uracil, RNA	782
Protein band	820
Tyrosine	850
Proline	937
Phenylalanine	1000
Carbohydrates / Glycogen	1025
Phosphate	1090
Strong C-O band of Ribose / Carotene	1120
Tyrosine	1171
Amide III, Tyrosine, Phenylalanine	1204
CH3/CH2-twist/wag/bend lipids	1307
Cellular nucleic acids	1335
CH2 Deform-lipids and protein	1443
COO^-	1560
Amide I	1600
Amide I	1655

Table 2: Wavenumber of peaks and their corresponding assignments[104]

in both datasets highlights the ability of BCARS to characterize the molecular composition of the sample under study, but the mismatch in amplitude and in some case precise wavenumber position of the line-shapes does call into question the quantitative accuracy and usefulness of BCARS.

3.9 Conclusion

The first conclusion is that segmentation approach such as kMeans, provides for enhanced bio-spectral analysis with BCARS hyperspectral images. This approach isolates cell-specific regions and can generate either (i) a single high quality representative spectrum from the aggregate of spectra within the region, therefore offering the potential for multivariate cell classification using single cell spectra much like applied in spontaneous Raman; and (ii) the potential to use the bounded regions to obtain large datasets of spectra that can be subsequently used to train classifiers.

Both BCARS and spontaneous Raman are powerful techniques in spectroscopic analysis, each with it's advantages and drawbacks. BCARS offers the advantage of fast imaging capabilities, with the capacity to produce hyperspectral images in rea-

sonable time. Due to the broadband stimulation, simultaneous detection of multiple molecular vibrations is possible, providing detailed spectral information. However BCARS must contend with nonresonant background due to the nature of the excitation interaction, in future works deep learning NRB removal techniques may be employed to correct the baseline offsets between BCARS and spontaneous Raman to determine if this can provide for closer fidelity in the direct comparison of BCARS with respect to spontaneous Raman. However, it is unlikely that such approaches will yield correspondence and we must therefore accept that spontaneous Raman will remain the gold standard for quantitative and multivariate analysis of cells. Spontaneous Raman spectroscopy is a simpler technique with less error prone noise removal strategies. While the noise profile may be simpler to correct, spontaneous Raman suffers from lower signal intensity, resulting in much longer acquisition times and reduced sensitivity. This makes hyperspectral imaging with spontaneous Raman impractical due to it's prohibitively long acquisition time. In the next chapter we attempt to develop a system that combines the strengths of both BCARS and spontaneous Raman.

4 Switchable Spontaneous Raman and BCARS Hyperspectral imaging system

The work presented in this chapter is based on the conference publication titled "Hyperspectral imaging system with switchable spontaneous Raman spectroscopy and broadband CARS"[3], which was presented at the SPIE 2025 conference held in Munich, Germany. In this work, I was required to rebuild the existing BCARS optoelectronic system independently and to integrate a new Raman modality.

4.1 Overview

In the previous section we demonstrated the ability of BCARS to record biological data in the context of cellomics with Cal-1 blood cells. The objective of this chapter is to present a novel, switchable hyperspectral imaging system that integrates both BCARS and confocal spontaneous Raman spectroscopy within a single optical platform. The system is validated by recording spectra from buccal epethelial cells, which are in the order of 2-3μm thick when deposited on substrate. This sample is thinner than the Cal-1 blood cell lines, making these epithelial cells more challenging to record in a hyperspectral mode. The dual-modality system is designed to enable full-spectrum acquisition across a broad wavenumber range (400–4000cm^{-1}) for both modalities. It facilitates direct, pixel-to-pixel comparison of Raman and BCARS modalities under identical experimental conditions for the first time. Combining both modalities allows the speed of BCARS to be married with the spectral fidelity of SR for chemically specific hyperspectral imaging and quantitative analysis. This hybrid approach provides a flexible and powerful tool for high-throughput hyperspectral Raman imaging with applications in biomedical diagnostics, materials analysis, and beyond.

4.2 Introduction

As described in the previous chapters, Raman spectroscopy is a label-free non-destructive sensing method, which has shown significant promise in a variety of disciplines for the chemical analysis of materials[107], plant tissue[108], and biological tissue[109]. In Chapter 2 we reviewed the application of hyperspectral Raman imaging (HRI), an emerging area of research in histopathology, whereby Raman spectra are captured from a grid of points, and these spectra can be processed to provide a pixelated image that highlights regions of spatially related chemical composition. An example of this is the development of Stimulated Raman Scattering (SRS) bedside platforms with sufficiently fast (μs) acquisition times for in-surgery label-free pathology[77][80]. However, typically SRS systems provide only a few discrete wavenumbers, and therefore the potential for hyperspectral tuning of different chemicals is limited.

Despite the high speed of SRS and other coherent modalities, spontaneous Raman remains the gold standard for quantitive comparison and multivariate statistical analyses including for multicomponent analysis and diagnostics. However the low probability of a Raman scattering event (only a small fraction of the incident laser photons - 1 in 10^6 to 10^8 - are in-elastically scattered) results in prohibitively long acquisition times for widefield hyperspectral imaging. One could envisage the combination of extremely high speed SRS (μs) - low bandwidth (just two wavenumbers) modality of SRS and the low speed (s) broadband (4000cm^{-1}) of spontaneous Raman. In such case the SRS wavenumbers could be selected to identify a particular chemical of interest such as the 2845cm^{-1} and2930cm^{-1} peaks relating to the nucleic acid, lipids and proteins. Thereby cell nuclei could be identified and then these regions could be targeted by the slower Raman for multivariate classification of disease type, or for multicomponent analysis.

Another approach, which is investigated here, is to use broadband Coherent anti-Stokes Raman Spectroscopy (BCARS) as the first step, followed by the targeted Raman. Admittedly, the acquisition time of BCARS ($\approx ms$) is approximately a 1000 times slower than SRS but it is still on the order of 100-1000 times faster than Raman

as discussed in Chapter 2. The advantage of using BCARS over SRS as the first step lies in the bandwidth of the single acquisition; BCARS can provide equivalent bandwidth to Raman, albeit with lower spectral quality. While this spectral quality may complicate precise quantitative measurements and comparisons, it is sufficient to apply a basic multicomponent analysis such that regions of related chemistry can be identified. Therefore, we can progress beyond the Raman targeting of one specific chemical boundary towards a multitude of different types.

As described in the previous chapters, BCARS is a nonlinear optical technique that generates a strong, coherent signal by employing pump and broadband Stokes pulses to induce a coherent molecular oscillation, which is then probed to produce an anti-Stokes signal. This four-wave mixing process offers a middle ground between SRS and spontaneous Raman, allowing for more rapid (ms) acquisition of a full spectral window. Furthermore the directional, laser like CARS signal with its anti-Stokes nature provide inherent rejection of fluorescence. However a significant challenge of BCARS is the introduction of the nonresonant background (NRB), which as discussed in the previous chapter is the source of the amplification in BCARS but also the source of the non-linear distortion which renders the BCARS spectrum non-quantitative and difficult to process. This NRB prevents direct analytical decomposition of the recorded spectrum and necessitates sophisticated postprocessing[110 52] techniques to suppress in order to accurately retrieve the Raman-like vibrational spectrum. In this paper, we describe an optical system that combines the modalities of BCARS and spontaneous Raman.

The core element of the system described in this chapter is a Toptica fiber laser system with a narrowband 770nm flat-top pulse with a synchronized broadband pulse of 1550nm, coupled with a highly nonlinear fiber for generating the supercontinuum spanning 900-1400nm at < 32mw. These two pulses are combined in space and time to generate a four wave mixing process in the sample, which underpins BCARS. A flip mirror is introduced in the delivery path (between the lasers and the microscope objective focusing into the sample) to switch the system into the spontaneous Ra-

man modality whereby the narrowband laser alone provides excitation. Similarly, a second flip mirror is introduced in the collection path to switch between colection of Raman scattered photons (Stokes) and BCARS photons (anti-Stokes), and the spectrograph as a motorised turret to enable the gratings to rotate to provide the relevant wavelength bands corresponding to both. This bi-modal HRI system has the benefit of providing rapid acquisitions of full spectral windows (400-4000cm^{-1}) for both modalities, which is a significant benefit over the limited wavenumber ranges of SRS systems. This system offers the capability to directly compare Raman and BCARS spectra over large spectra datasets, and may be an important milestone in the development of high-speed HRI systems.

4.3 Optical System

4.3.1 BCARS Modality

The BCARS micro-spectroscopy system is adapted from the design previously reported by Muddiman et al.[111]. Figure 14 depicts the optical system set in BCARS configuration when all three flip mirrors (FM1, FM2, and FM3) are in the down position. The probe beam, at a wavelength of 770 nm, is generated through second-harmonic generation (SHG) of a 1550 nm fundamental source utilizing a periodically poled lithium niobate (PPLN) crystal. This process yields narrowband pulses with an approximate temporal duration of 3.4 ps. The resultant beam subsequently traverses a 2x beam expander, which increases its diameter to approximately 4 mm. Precise temporal overlap with the Stokes beam (described subsequently) is achieved using a variable optical delay line, composed of two 45-degree mirrors mounted on a motorised linear translation stage. A Semrock line pass filter (Semrock: LL01-780) is incorporated to spectrally purify the probe beam; this needs to be rotated to tune to the 770 nm probe line. When FM1 is in the downward position, the 770 nm beam is directed to dichroic mirror DM2, where it is collinearly combined with the Stokes beam.

The broadband Stokes beam is produced via supercontinuum generation within a

highly nonlinear optical fiber. This pulse is then passed through a prism compressor, consisting of two SF10 glass prisms, which introduces negative dispersion to temporally compress the pulse and manage its chirp prior to reaching the sample. The probe and Stokes beams are combined by a 950nm shortpass dichroic (Thorlabs: DMSP950). The two beams are carefully aligned spatially by beam walking of mirrors and aligned temporally using the motorized delay line. The co-aligned beams are focused onto the sample by a high NA microscope objective (Olympus:UPlanSApo 60x/1.2 NA), satisfying the phase matching requirements. Within the sample, a four-wave mixing process occurs, generating a coherent Raman spectrum. These BCARS photons are collected by a second objective (Olympus: 20x/0.46 NA UmPlanFl). Higer NA microscope objective did not offer a significant improvement in photon colection because the phase matching condition enforces a low angle of emission. It should also be noted that for the BCARS delivery, the NA of the focused beams was only half that of the microscope objective (the $1/e^2$ width of the beams was approximately half the width of back aperture of the MO). A Casse-Grain telescope is currently being added to the system, but at time of writing this was not usable. Therefore the focus spot size that was used was twice that possible with the 1.2NA MO and significant gains for BCARS intensity can be had over the results presented later in this chapter owing to the third order nonlinear effect.

The anti-Stokes signal generated at the sample propagates co-linearly with the excitation beams in the forward direction. A 750 nm shortpass filter(Thorlabs: FESH0750) is positioned in the collection path prior to the spectrograph slit in order to block the high-intensity excitation beams, selectively transmitting only the blue-shifted coherent Raman signal. With FM3 in the downward position, this isolated coherent signal is focused by a 50 mm focal length lens onto the entrance slit (25 μm) of the Andor 500i spectrograph (focal length 500mm). Spectra were collected with a thermoelectrically cooled CCD detector (Newton DU920P Bx-DD, Andor), maintained at $-80\,^{\circ}$C. The spectrograph utilised a 300 lines/mm grating, providing a spectral resolution of approximately 10 cm^{-1}. Care was taken to ensure

that the solid angle of the input light never exceed the f-number of the spectrograph.

Sample positioning was handled using an MS-2000 XY servo stage (ASI) for coarse alignment together with a P-545.xR8S PInano XYZ piezo stage (PI) for high-precision raster scanning. Owing to the broadband spectral density of the super-continuum source, both two- and three-color excitation contributed to the measured signals, which covered the 500–4554 cm^{-1} wavenumber range. No features appear between 1800 and 2800 cm^{-1} because, for our laser configuration, the frequency-domain four-wave-mixing intensity is negligible in this region. Instrument control was carried out using a MATLAB-based custom script.

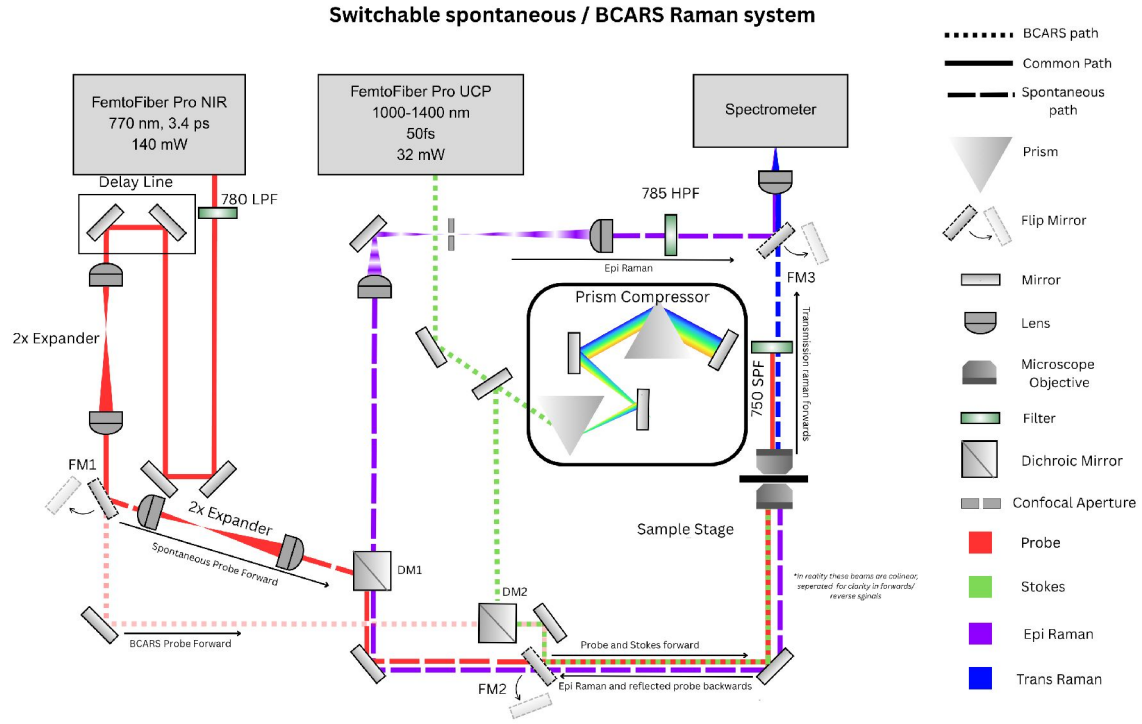

Figure 14: Illustrating the switchable optical system with both the Spontaneous and BCARS available for recording, which can be switched between using three flip mirrors.

4.3.2 Raman Modality

The spontaneous Raman spectroscopy pathway diverges from the BCARS configuration subsequent to the first 2x expander, through the actuation of three flip

mirrors (FM1, FM2, and FM3) as illustrated in Figure 14. Specifically, for spontaneous Raman measurements, FM1 is engaged to redirect the 770 nm excitation beam through a second beam expander consisting of a f50mm and f100mm plano convex B-coated lens (Thorlabs), increasing the size of the beam to a final 8mm to maximally fill the back aperture (60x) MO and achieve full NA for focusing (unlike the BCARS configuration, which achieved only 50% NA as described earlier). This beam then reflects from DM1, a dichroic mirror (Thorlabs: LPD02-785RU) to a shared path with BCARS. FM2 is positioned such that in it's upright position it reflects the BCARS path and into the MO, in it's down position the spontaneous beam is directed along the same path as the BCARS.

As an aside, initial design passed the spontaneous probe through DM2 responsible for joining the Stokes and Probe beams in the BCARS path. While this has the benefit of allowing the use of the four-wave mixing process to validate the co-alignment of the spontaneous and BCARS path, it also prevents the recording of higher Raman shifts as values above 2460cm^{-1} as these would be above the 950nm wavelengths reflected by DM2. The excitation beam finally is focused onto the sample by the main 1.2NA microscope objective. The Stokes shifted Raman signal generated within the sample, occurring at wavelengths greater than 770 nm, is collected in the backscattered direction, and therfore the same high NA MO collects the scattered photons over wide angle. DM1 is specifically chosen for its high transmissivity in this spectral region, allowing the Stokes signal to pass through, while reflecting the 770 nm laser. A telescopic arrangement of two 200 mm focal length lenses is used to maintain a collimated beam on the collection path, which is then matched to the same focusing 50 mm lens as used in the BCARS collection path above. This focusing lens must remain common to both Raman and BCARS modalities. A confocal aperture of 200 μm is inserted in between the two 200 mm lenses mentioned above, and mounted on a 3d translation stage to facilitate alignment with the laser spot in the sample plane.

In the Raman collection path, the Stokes signal passes through a 785 nm longpass

filter (Semrock: LP02-785RU), which effectively removes any residual Rayleigh scattered laser light at 770 nm. Finally, similar to the BCARS detection pathway, the filtered spontaneous Raman signal is focused by the same 50 mm focal length lens onto the entrance slit of the spectrograph for dispersion and detection. Additional alignment mirrors, not depicted in Figure 14, are incorporated to ensure (via beam walking) precise co-alignment of the spontaneous Raman signal path with the previously described BCARS signal path, thereby enabling rapid switching between the two detection modalities.

4.4 Alignment

The alignment requirements for BCARS are significantly more demanding than for spontaneous Raman spectroscopy, whereby alignment of the beams is necessary in both space and time.

Spatial alignment involves collinearly combining the pump and Stokes beams using a dichroic mirror, which is designed to transmit the 770 nm pump beam while reflecting the longer-wavelength Stokes beam. This alignment is an iterative procedure. The pump and Stokes beams are directed to impact the same location on the dichroic mirror to achieve spatial overlap. Next, the angle of the Stokes beam relative to the dichroic is adjusted to ensure the reflected Stokes beam and the transmitted pump beam are collinear. The quality of this collinearity is verified by confirming that the beams remain fully overlapped at multiple planes along the propagation path. This is achieved by observing the combined beam profile on a camera due at multiple points in the optical path. This camera is employed due to the difficulty in observing the longer wavelengths of the stokes beam.

Temporal alignment, or the synchronisation of the pulse arrival times, is controlled by a mechanical delay line mounted on a single-axis translation stage. For practical use it is necessary that adjustments to the temporal alignment do not introduce any spatial misalignment. This is ensured by the incoming beam being aligned parallel to the translation stage's axis of motion. To achieve this, a camera was mounted on the

moving stage, positioned blocking the initial mirror. The stage was then translated through its full range of motion, and the incoming beam's pointing was adjusted until its x/y position remained perfectly stationary on the moving camera sensor throughout the translation. Following this, the delay line returning mirror must be aligned to return a beam that is perfectly parallel to the input beam, regardless of the stage's position. For this, the camera was placed in a fixed position after the delay line. The stage was translated again, and the delay line's optics were adjusted until the beam spot remained stationary on the fixed camera sensor. This procedure guarantees that translating the stage to vary the temporal delay does not affect the downstream spatial alignment of the beams. Included in this temporal alignment is the compression of the stokes beam to match all wavelengths into the same time step by way of the prism compressor.

Due to the more complex alignment requirements of the BCARS system the switchable system was designed with the primary objective of maintaining the established BCARS alignments while integrating the spontaneous Raman capabilities. This was achieved by modifications to the optical path as outlined in Fig 14, including the introduction of a single flip mirror (FM3) on the signal collection side alongside the addition of two flip mirrors on the excitation side (FM1 and FM2) which were incorporated to redirect the 770 nm probe beam towards a dichroic mirror (DM1). This DM is designed for 45 degree reflection at 785 nm; therefore the angle must be increased to achieve optimal reflection of 770 nm. The angle of incidence was increased from the normal by approximately 8 degrees. This specific angular orientation, facilitated the placement of FM1 and the rotation of DM1, enabled an epi-collection, accommodating both the reflection of the probe beam towards the microscope objective and the subsequent transmission of the backscattered Raman signal towards the spectrometer. The inclusion of FM1 and FM2 also introduced a degree of redundancy for the initial temporal and spatial alignment of the BCARS probe beam, allowing the established BCARS alignment to serve as a reference for the spontaneous Raman probe path by placing FM1 in the down position and FM2

in the upright position. Future system iterations could potentially simplify this by utilising a singular, consolidated path for the probe beam to the microscope, thereby eliminating the need for two flip mechanisms and one mirror dedicated to this redirection albeit sacrificing higher wavelengths that mediate Raman shifts from the CH band.

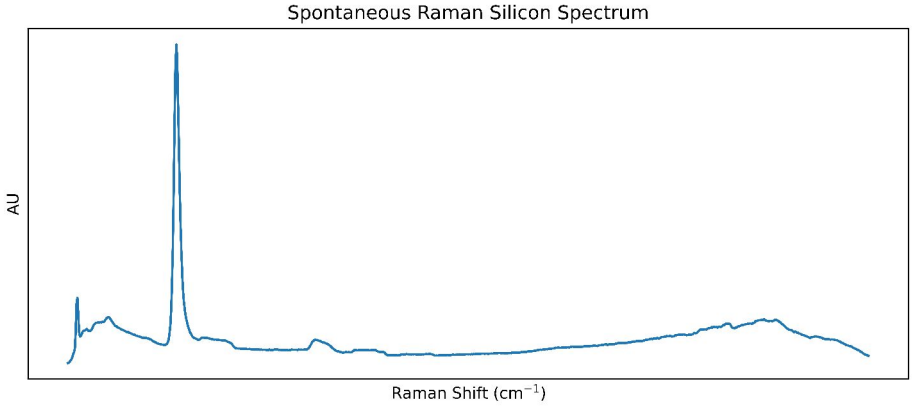

Figure 15: Shows an example of the silicon spectrum used for alignment purposes with the characteristic sharp Lorentzian signature of it's dominant peak nominally centred at 520cm^{-1} in established literature.

The alignment protocol for the spontaneous path begins with the co-alignment of the spontaneous Raman probe beam with the pre-existing BCARS probe beam path. This step is accomplished using beam walking based on a combination of visual alignment with fluorescent cards for coarse adjustments, real-time imaging feedback from a Thorlabs camera interface for fine-tuning, and quantitative beam spot measurements at multiple points along the optical train. Following initial alignment into the microscope, a reflective silicon reference sample was used for aligning the epi-Raman signal collection pathway. Initial guidance of the reflected/backscattered Raman signal towards the spectrometer was facilitated by fluorescent cards and adjustable mirrors. The beam was then "walked" into the 50 mm focusing lens coupled to the spectrometer entrance slit. With the spectrometer slit maximally open, the beam was centred on the CCD detector, and its position was iteratively adjusted to maximize signal intensity. This alignment was refined by acquiring the Raman spectrum of the silicon sample and optimizing the signal corresponding to

its characteristic spectral shape. Following this, two 200 mm plano-convex lenses, arranged in a telescopic configuration, were introduced and aligned to ensure the beam maintained its optimized path to the spectrometer while preserving signal quality. Finally, a confocal aperture was inserted and precisely adjusted using an x/y/z translation stage to maximize the collected Raman signal from the Silicon a sample of which is given in Fig 16. This was once again measured using the intensity of the signal from the spectroscope.

4.5 Methods

4.5.1 Sample Preparation

Three reference samples were prepared for system alignment and performance assessment. The first was a polished silicon wafer used as a Raman calibration standard as described in section 4.4. To ensure compatibility with the water immersion objective, this substrate was placed inverted on a clean glass coverslip. The assembly was then placed on the microscope stage, and deionised water was applied as the immersion medium to couple the coverslip to the objective lens. The second sample, a polystyrene standard, was prepared by drop-casting a water solution of 5 µm polystyrene micro-beads onto a clean glass coverslip, which was then allowed to air-dry completely, leaving a sparse layer of beads. For analysis, this coverslip was placed directly atop the water from the immersion objective, with the sample side facing away from it. The final sample was a combination of buccal epithelial cells and polystyrene micro-beads; epithelial cells were collected from a donor using a sterile swab, transferred to a CaF2 coverslip (Crystran UK: Raman Grade CaF_2), and a water solution of 5 µm polystyrene beads was pipetted onto the same area before the entire sample was allowed to air-dry. This sample was similarly placed, sample-side away from the objective and coupled directly to it by water.

4.5.2 Recording and Processing Methods

System validation and comparison of spontaneous Raman with BCARS were initially performed using the spectra of a polystyrene sample with only minor processing. Three spectra were acquired from this film under both spontaneous Raman (15s acquisition time) and BCARS configurations. The aim of this sample was to confirm the functioning of both modalities and to illustrate the difference in raw spectra recorded from the system with a particular emphasis on the BCARS NRB. Neither spectra were processed as per the methods described in Chapter 3 and only the recorded irradiance was used for initial comparison. However to aid the comparison both spectra were subject to a median filter to mitigate cosmic rays which were not of particular interest, and were z-score normalized and scaled to standardize their scales for comparison.

Following this a recording of the second sample, a combination of buccal epithelial cells and polystyrene fixed to the CaF2 slide were recorded. With the BCARS raster scanning through the full travel of the piezo stage in increments of 1μm, giving a total 200x200 spatial locations with a spectrum comprising of 1000 wavenumbers. The BCARS was subjected to the full suite of the pre-processing methods as described in Chapter 3, including the Anscombe transform, SVD denoising process, Inverse Anscombe, and the Savitsky-Golay filter. In this case the wavelength calibration was not performed beyond the approximate axis returned by the spectroscope software (Andor Solis). This wavelength axis was then converted to Raman shift based on the excitation wavelength without rigorous wavenumber calibration. Following the recording and processing of the BCARS spectra, the spontaneous Raman was targeted in areas that contained evidence of CARS as indicated by an increase in total BCARS intensity. Five spectra per sample were recorded and preprocessed with median filtering, baseline removal and a Savitsky-Golay filter as described in Chapter 3. Similar to the BCARS the Raman shift was calculated based on the wavelength axis provided by the spectrometer relative and the known wavelength of the probe for wavenumber conversion.

4.6 Results

A visual comparison of the spectra in Fig 16 highlights the key differences between the raw outputs of the two modalities. Both the spontaneous Raman (blue) and BCARS (orange) spectra successfully identify the characteristic polystyrene peaks at the same Raman shifts, such as the strong resonance just below 1000 cm^{-1}. However, the BCARS spectrum is dominated by a significant non-resonant background which appears as a strong ramp, particularly in the lower wavenumber region (400-800 cm^{-1}). This sloping background is absent in the spontaneous Raman spectrum, which sits on a flat baseline. Consequently, while the primary peaks are visible in both, the weaker polystyrene features in the higher wavenumber region (e.g., 1200 cm^{-1}) are difficult to discern in the raw BCARS spectrum as the overall signal (both resonant and the non-resonant amplification) diminishes in this region. Nevertheless it is possible to recover reasonable lineshapes in this region using the processing steps described in the previous chapter.

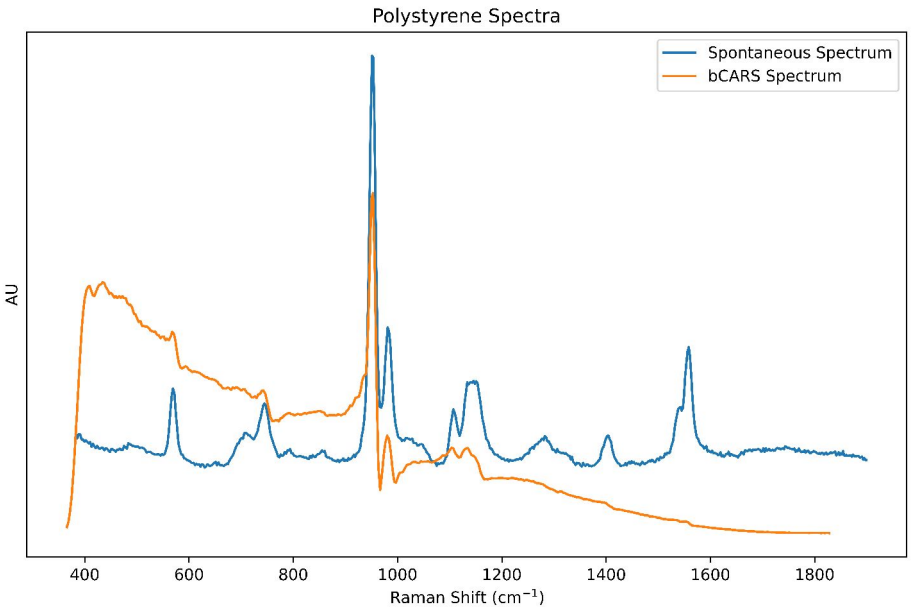

Figure 16: Shows shows a comparison of unprocessed polystyrene spectrum recorded from Spontaneous Raman (Blue) and BCARS (orange).

Following this initial validation, the combined buccal cell and polystyrene bead sample was mapped, as shown in the hyperspectral mean intensity image and cor-

responding extracted spectra in Fig. 17. The performance of the BCARS modality in this instance was unfortunately significantly poorer than expected. This is attributed to a combination of factors, including an overly conservative low probe laser power (approximately 10-40%) to avoid burning, a potential focal offset from the optimal cell depth, lower than optical numerical aperture (see earlier discussion) as well as possible misalignment of the beams and the two microscope objectives. The resulting low signal-to-noise ratio (SNR) in the BCARS results therefore should not be taken to be systematic, but rather the result of a preliminary set of suboptimal experimental parameters. This low SNR inevitably leads to the introduction of erroneous features during processing, either from the SVD denoising or the Kramers-Kronig correction. A notable artifact is the appearance of a small, sharp peak near 1000 cm^{-1} in the CaF2 spectrum. Nevertheless, the results do confirm the success in obtaining BCARS biospectra and combining the two modalities for the first time for application to cell imaging. Future work will certainly improve the BCARS modality such that multicomponent analysis will be achievable as the first step.

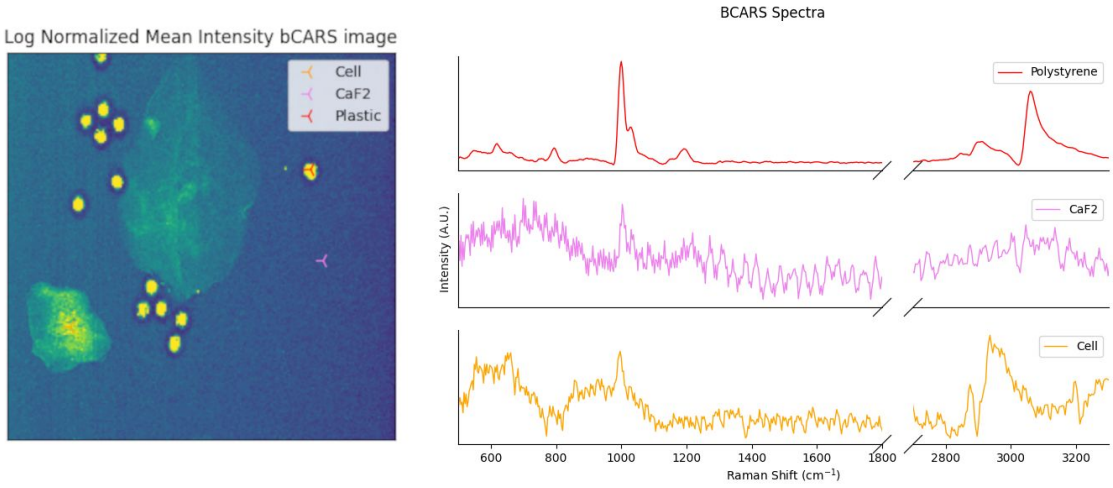

Figure 17: Shows a mean intensity log normalized hyperspectral image of a buccal epethelial cell spiked with 5μm polystyrene beads recorded via a BCARS raster scan with a 5ms dwell time per pixel. A sample of Polystyrene, a Buccal epethelial cell and a sample of the background signal are marked on the image and Corresponding colour coded processed BCARS spectra for each highlighted sample are shown on the right

Despite the low SNR and poor quality of the BCARS data, the hyperspectral image was still sufficient to guide a targeted follow-up analysis. By using the BCARS image

to identify key areas of interest (the cell body), spontaneous Raman spectra were subsequently acquired. As illustrated in the follow-up spontaneous Raman results in Fig 18, these targeted spectra show high agreement with the established literature for the peak locations within buccal epithelial cells. This demonstrates that even a significantly low-quality BCARS map can still serve as a rapid and effective tool for locating and targeting specific regions for subsequent, slower, spontaneous Raman acquisition.

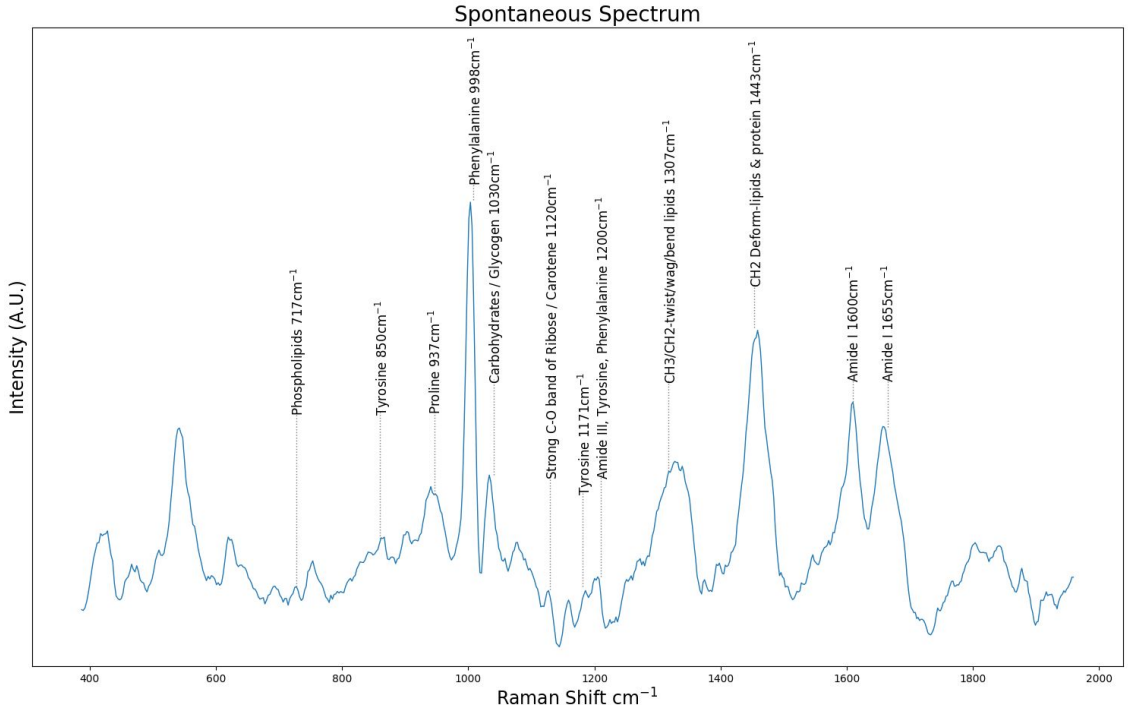

Figure 18: Shows a spontaneous Raman spectrum of the buccel epethelial cell in the fingerprint region with highlighted biological peaks in line with existing literature[104].

4.7 Conclusion

In this work, we have presented the first demonstration of a unified optical platform capable of acquiring both broadband Coherent Anti-Stokes Raman Scattering (BCARS) and Spontaneous Raman (SR) spectra. While our proof-of-concept experiments on polystyrene and buccal epithelial cells verified the system's dual-mode capability, they also highlighted a current limitation: the BCARS spectra, in their unoptimized state, yielded low signal-to-noise and limited chemical specificity which

in turn resulted in erroneous features from the denoising techniques applied.

Despite this poor BCARS SNR, the system still demonstrated that even these low-fidelity BCARS spectra were capable of identifying areas of interest by total intensity which were subsequently used to target the slower, high-fidelity SR system to specific regions of interest. This allowed us to recover meaningful chemical spectra from targeted areas of the buccal cell, proving the core value of this integrated approach—something unachievable with separate systems.

The path forward will focus on transforming this hybrid potential into a high-throughput workflow. First, by addressing the signal quality, systematically optimising the BCARS hardware, including fully utilising the delivery objective's NA and increasing laser power. We aim to incorporate an off-axis Cassegrain telescope to provide the necessary beam expansion to match both the Stokes and the spontaneous path to the full NA of the delivery objective. Corrections of beam abberations, such as defocus and astigmatism can be carried out with a deformable mirror. We predict that with these hardware optimisations, the resulting higher-quality data will allow SVD denoising to perform far more effectively. Furthermore the application of the SVD to the spontaneous Raman spectrum may further reduce acquisition times, potentially achieving sub-second SR acquisition per pixel.

5 Convolutional Auto-Encoders for Spectral Line Lorentz Decomposition

The work presented in this chapter is based on the conference publication titled "Characterizing spectral peaks with deep-learning auto encoders"[4], which was presented at the SPIE 2025 conference held in Munich, Germany.

5.1 Overview

In the previous chapters we explored the comparison of BCARS with spontaneous Raman and the subsequent development of a novel optical system and its preliminary validation. In this chapter, we move away from the optical system and towards the development of a new spectral processing tool. Specifically, we look to improve upon the utility of chemical information encoded in the hyperspectral data by developing an algorithm capable of extracting the specific set of line-shapes that make up the spectrum. Therefore, we attempt to extract more meaningful physical information from a standard mixed spectrum. In this study we have focused on the Lorentzian lineshape which is appropriate for spontaneous Raman and the resonant component of the third-order susceptibility of BCARS, but the resulting algorithm could easily be generalised to Gaussian line-shapes.

The objective of this chapter therefore is to demonstrate a machine learning method for decomposing a single recorded spectrum into a sum of constituent Lorentzian functions. This could be especially useful for precise characterisation of individual spectral features, enabling detailed analysis of properties like chemical composition, temperature, and density. While methods exist for peak position identification, the automatic and accurate retrieval of peak width and amplitude remains a significant challenge. This chapter introduces a deep learning methodology to estimate all three Lorentzian parameters, position, width, and amplitude for each peak within a given spectrum. A key challenge in this task is the inherent sparsity of the desired output

(a set of three-parameter tuples for each peak) relative to the dense 1D spectral input. To address this, our approach utilises masking techniques. This targeted learning approach enhances computational efficiency and model focus, enabling the robust extraction of Lorentzian parameters from synthetically generated spectral data.

5.2 Introduction

Spectroscopy, in many forms including but not limited to Raman, serves as a cornerstone of analytical science, providing significant insights into the fundamental properties of matter by analysing its interaction with various forms of energy. The resultant spectra, rich in information, typically manifest as a superposition of characteristic spectral lines or peaks. Decomposing this mixed spectrum into its constituent functions, frequently modelled as Lorentzian profiles (and sometimes as Gaussian or Voigt), is a critical step for the precise characterisation of individual spectral features. This detailed decomposition has utility across diverse scientific and engineering disciplines, enabling the identification and quantification of chemical compositions[112], the determination of physical conditions such as temperature and density[113,114], and extending to applications in material science for stress analysis[115], in astronomy for stellar classification[116], and in biomedical optics for tissue characterisation[117].

Each spectral peak, when modelled as a Lorentzian function, is typically characterised by three fundamental parameters: the central position (x_0), the full width at half maximum (FWHM, related to γ), and the amplitude (or intensity). While various computational methods, including iterative algorithms and simpler peak-finding routines, have been proposed and are widely used to identify the positions of these peaks in a spectrum[118,119], the automatic, robust, and accurate retrieval of peak width and amplitude, remains a more significant and often underexplored challenge. Traditional curve-fitting techniques, for instance, can be susceptible to issues like convergence to local minima, sensitivity to initial parameter guesses, and high

computational cost for complex spectra, sometimes requiring considerable manual intervention or sophisticated baseline correction steps.

Here, we investigate a deep-learning inspired method to retrieve all three Lorentzian parameters—position, width, and amplitude—for each constituent peak within a given input spectrum. Auto-encoders, a class of artificial neural networks, are particularly well-suited for tasks involving representation learning[32]. They excel at dimensionality reduction by learning efficient, compressed representations of input data through their characteristic bottleneck layer[120]. This layer captures the most salient features of the data while ideally discarding noise and redundancy. This intrinsic capability not only helps in simplifying complex datasets into more manageable forms but also significantly improves computational efficiency, making them attractive for analysing large spectral datasets[121]. Furthermore, autoencoders can learn complex, non-linear relationships within the data, which is crucial for accurately modelling the subtle interplay of overlapping spectral features that prove challenging for linear methods.

Generating the complete set of Lorentzian parameters (position, width, and amplitude for multiple peaks) from a raw 1D spectrum presents a particularly challenging problem for standard auto-encoder architectures. This is primarily due to the sparse nature of the desired output—a structured set of three-parameter tuples for each identified peak—relative to the dense, continuous 1D spectral input. For example, a spectrum might have thousands of data points, but only a few peaks, meaning the target parameter vectors will be mostly zeros. This sparsity poses a significant hurdle for auto-encoders, as the majority of the information in the output cannot be easily or effectively applied to common loss functions. Such a disconnect makes it difficult for the model to receive meaningful gradients and thus to optimise its parameters effectively towards the desired sparse parameter extraction. To address this, our proposed methodology strategically incorporates masking techniques at the output stage. This focused approach directs the model's learning towards the regions of interest (i.e., an identified peak location) within the network, thereby providing

more meaningful gradients for optimisation and enhancing the robustness of parameter extraction. This work, therefore, focuses on leveraging the representational power of autoencoders while tailoring their training to overcome the challenges of sparse parameter regression from dense spectral data.

5.3 Theory

Spectroscopic methods attempt to measure transitions between different energy levels in molecules and atoms, including electronic, vibrational and rotational states. The spectral linewidth of these transitions is influenced by the finite lifetime of the energy states involved. In the context of Raman this is done by analysing the inelastic scattering of light from molecules as they enter or exit virtual states during the interaction with incident photons. This is directly related to the Heisenberg uncertainty principle in the time energy form (eq 22), which dictates that any state with a finite lifetime (t) must have uncertainty in its energy (ΔE)

$$\Delta E.\Delta t \geq \frac{\hbar}{2} \tag{22}$$

This can be modelled in the time domain. Where after excitation, the coherent vibration of a molecule oscillates at its natural frequency (ω_0), while it's amplitude decays exponentially over time. This process is mathematically representable as a damped harmonic oscillator and is described by the function ($h(t)$).

$$h(t) = Ae^{-\Gamma t}\cos(\omega_0 t) \tag{23}$$

Of course, spectrometers measure intensity as a function of frequency, not time. We can convert from the time domain $h(t)$ to the frequency domain $H(\omega)$ by way of a Fourier transform.

$$H(\omega) = \int_{-\infty}^{\infty} h(t)\, e^{-i\omega t}\, dt \tag{24}$$

The measured intensity, $I(W)$ is proportional to the squared magnitude of this

complex result. The Lorentz function can thus be derived. Since $h(t)$ is zero for $t < 0$, the integral is taken from 0 to ∞. By substituting $h(t)$ with Euler's formula.

$$H(\omega) = \frac{A_0}{2} \int_0^\infty e^{-\Gamma t}(e^{i\omega_0 t} + e^{-i\omega_0 t})e^{-i\omega t}dt \tag{25}$$

This can be separated into two parts.

$$H(\omega) = \frac{A_0}{2} \left[\int_0^\infty e^{-(\Gamma - i(\omega_0 - \omega))t}dt + \int_0^\infty e^{-(\Gamma + i(\omega_0 + \omega))t}dt \right] \tag{26}$$

Solving each of these integral gives:

$$H(\omega) = \frac{A_0}{2} \left[\frac{1}{\Gamma - i(\omega_0 - \omega)} + \frac{1}{\Gamma + i(\omega_0 + \omega)} \right] \tag{27}$$

In spectroscopy we are interested in frequencies near the peak where $\omega \approx \omega_0$. In this area, the first term's denominator is very small, making the term large. The second term's denominator approaches $2\omega_0$, leading to the second term becoming negligibly small and therefore can be discarded to produce:

$$H(\omega) \approx \frac{A_0}{2} \frac{1}{\Gamma + i(\omega - \omega_0)} \tag{28}$$

The observable intensity is proportional to the squared magnitude of $H(w)$, found by multiplying with it's complex conjugate:

$$I(\omega) \propto |H(\omega)|^2 = \left(\frac{A_0}{2} \right)^2 \frac{1}{\Gamma + i(\omega - \omega_0)} \frac{1}{\Gamma - i(\omega - \omega_0)} \tag{29}$$

$$= \frac{A_0^2}{4} \frac{1}{\Gamma^2 + (\omega - \omega_0)^2} \tag{30}$$

This can be further normalised into it's standard form such that the area of beneath the function is 1, this is done by computing the overall integral of the function and multiplying by its reciprocal giving the final Lorentzian function:

$$L(\omega) = \frac{1}{\pi} \cdot \frac{\Gamma}{(\omega - \omega_0)^2 + \Gamma^2} \tag{31}$$

5.4 Data Generation

To generate precise labelled data, we created mixed spectra by summing individual Lorentz functions. In what follows, for simplicity we apply integer units for "recorded" wavenumber position x_0; note ω_0 and x_0 are used interchangeably. The parameters for each individual Lorentz function were randomly generated: widths (γ) as real numbers such that $0 < \gamma < 10$, and initial amplitudes also as real numbers between 0 and 1. The centrepoints (x_0) were randomly generated such that $150 < x_0 < 850$. This number was floored to acquire the pixel location in the discrete input, and the subpixel offset $x_0 - \lfloor(x_0)$ was stored in that pixel location. The Lorentz centrepoints were selected such that each was assigned to a unique integer pixel index within the spectral recording size (K, typically 1024 to match our spectroscope); this was controlled to ensure that no two generated Lorentzian centres could exist in the same integer pixel location prior to considering their sub-pixel displacements.

We make two modifications to the standard equation (eq 31) derived earlier, in order to better match the traditional interpretation of recorded Raman spectra. Firstly we scale the function such that our amplitude is in terms of the maximum value of the Lorentz function not the area. Secondly we scale our width term, such that the Γ relates to the FWHM (full width at half maximum) rather than the HWHM (half width at half maximum). The final form for generating our lineshapes and parameters is given below in eq 32. This allows us to record our parameters as they are for ground truth purposes, while simultaneously using them to generate our spectral shapes.

$$Lorentz(A, \Gamma, x_0) = \frac{A\Gamma^2}{4(x - x_0)^2 + \Gamma^2} \tag{32}$$

After summation of the individual functions, the full mixed spectrum was normalised such that its maximum amplitude was 1. This scaling factor was consistently applied to both the generated spectrum and the amplitude component of the constituent Lorentzians parameters.

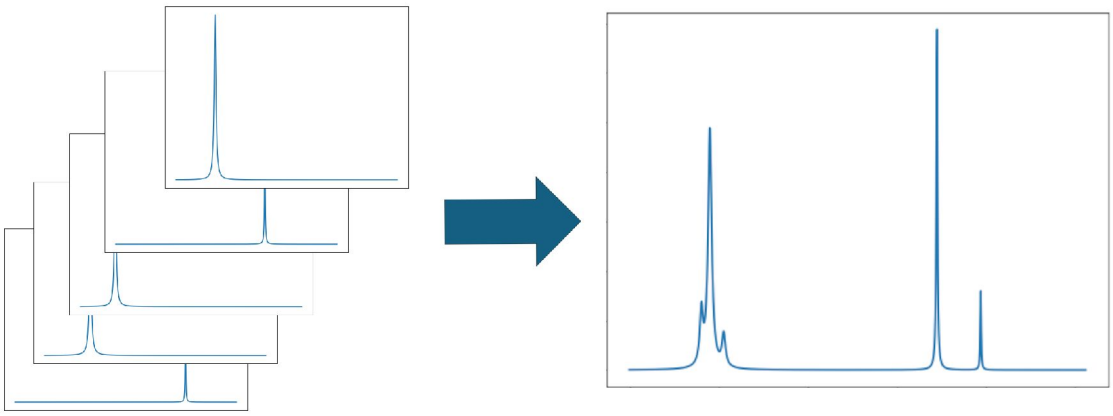

Figure 19: Lorentz functions generated from eq. (32) (left), then summed and normalized (right)

These Lorentz parameters (sub-pixel location, width, and scaled amplitude) for each peak became the targets for the respective autoencoders. For each type of parameter (e.g., amplitude, width, sub-pixel shift), the target values were stored in a 1-D vector of length K. These target vectors were sparse, meaning the majority of entries were zero. For instance, the target vector for amplitude would contain the specific amplitude value (e.g., a value between 0 and 1) at the integer index corresponding to a peak's centre, and zero in all non-peak locations. Similarly, separate K-length sparse vectors stored the subpixel positions and widths at their respective peak locations.

5.5 Model Architecture

The network architecture employed is an adaptation of a previously reported model used for BCARS denoising[52], a 1D Convolutional Autoencoder (CAE) with a[122]U-Net-like structure. Each model in the sequence features a symmetric encoder-decoder path. The encoder comprises 7 convolutional blocks. The first encoder

block utilises a 1-D kernel of size 8 with a stride of 1, transforming the input from input channels to 64. Subsequent encoder blocks (encoder 2 through 6) also use a kernel size of 8 but employ a stride of 2 for progressive down-sampling, with channel dimensions increasing as follows: 64→128→256→512→1024→2048. The final encoder block (encoder 7, the bottleneck) uses a stride of 1, maintaining the feature map length while operating with 2048 channels.

Correspondingly, the decoder path consists of 7 blocks that mirror this structure to reconstruct the sequence. The first decoder block (decoder 1, operating on the bottleneck output) uses a stride of 1 (2048 channels in and out). Decoder blocks 2 through 6 utilise a kernel size of 8 and a stride of 2 for up-sampling, progressively reducing channel dimensions: 2048→1024→512→256→128→64. The final decoder layer (decoder7) is a 1D transposed convolution with a stride of 1 and dynamically calculated padding to ensure the output sequence length matches the original input, transforming the 64 channels to the specified output dimension. Skip connections are implemented between analogous encoder and decoder layers (e.g., the output of encoder6 is added to the input of decoder2, encoder5 to decoder3, etc.), facilitating the propagation of high-resolution feature information and aiding in fine-grained reconstruction. The final output is passed through a Sigmoid activation function for mask generation.

In this sequential framework, the same model processes a single-channel 1-D input to generate an initial mask. Subsequent models in the sequence then receive a two-channel input, comprising the original 1-D data and the mask generated by the preceding model as illustrated in Fig 20.

5.6 Model Training

The Mask Generation Model is trained first. Its objective is to produce a spectral mask identifying potential peak regions. This model is trained using an L1 norm (Mean Absolute Error, Eq. 33) between its sigmoid output and the ground-truth binary mask. The continuous sigmoid outputs (m_i) from the trained Mask Gener-

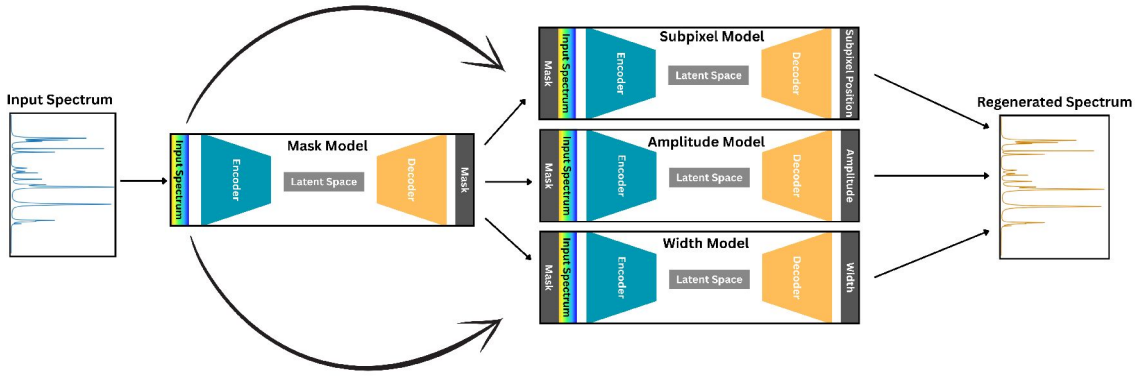

Figure 20: Network structure illustrating the total model sequence, beginning with an input spectrum on the left, being fed to the "Mask Model" extracting a mask for peak centre locations, the same input spectrum and output mask are then fed to subsequent models where each one predicts one of three parameters, subpixel, amplitude and width respectively. Finally a new spectrum is generated from the extracted parameters to validate the performance

ation Model are then converted to a binary mask ($m'i$) by thresholding at a value (e.g., 0.05, as in Eq. 34), this threshold was optimised based on precision and recall values.

$$L_{\text{mask}} = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{33}$$

$$m'_i = \begin{cases} 1 & \text{if } m_i > 0.05 \\ 0 & \text{otherwise} \end{cases} \tag{34}$$

The application of this mask in the loss function reduces noise in gradient calculations by restricting the networks learning to areas where the target feature is present thus preventing the network from trying to predict the background noise or zero values. The same mask is also fed to the input of the models where it provides explicit guidance to the network, highlighting areas of importance. The output of the network therefore, is directly related to only the prediction of the key values rather than forcing the model to predict zero values for non-parameter locations in the vector.

Following the training of the mask model the three parameter models are trained, a

custom loss is used for each parameter. The loss for the amplitude model $L_{amplitude}$, calculates a masked Mean Squared Error. This loss focuses on the squared differences between true and predicted values specifically where an associated input mask is active (values > threshold) eq 34, normalised by the number of active mask elements eq 35:

$$L_{amplitude} = \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \cdot m'_i}{\sum_{i=1}^{N} m'_i} \tag{35}$$

The loss for the subpixel model $L_{subpixel}$ also uses a masked Mean Squared Error, this time it introduces a -0.5 offset to the prediction before error calculation, allowing for the zero value to describe a peak position directly in the centre of a pixel, a negative value indicating a negative subpixel offset, and a positive value a positive subpixel offset.

$$L_{subpixel} = \frac{\sum_{i=1}^{N}(y_i - (\hat{y}_i - 0.5))^2 \cdot m'_i}{\sum_{i=1}^{N} m'_i} \tag{36}$$

Finally, a width loss, again utilises a masked Mean Squared Error, this loss introduces a scalar multiplier to the prediction before computing the error, designed to penalise discrepancies in the predicted width of features within the masked areas: where S is the scalar given by the max possible value thereby producing a bounded output from the final sigmoid layer between 0 and S.

$$L_{\text{widths}} = \frac{\sum_{i=1}^{N}(y_i - S \cdot \hat{y}_i)^2 \cdot m'_i}{\sum_{i=1}^{N} m'_i} \tag{37}$$

Key hyperparameters were optimised through a series of preliminary short training runs. These initial optimisations were conducted on smaller representative datasets of 30,000 simulated spectra for training and 10,000 for validation. The learning rate was systematically evaluated , with values range tested[123] including 0.05, 0.025, 0.01, and 0.005, to identify an optimal starting point for full-scale training, furthermore learning schedulers were employed to dynamically update the learning rate. Momentum values were trialled at 85, 90, and 95. With the optimal per model being

selected. For the full scale-training the dataset was increased by a factor of 10 to 300,000 simulated spectra for training and 100,000 for validation. Checkpoints of the model weights were recorded at the lowest validation loss per training run.

5.7 Results

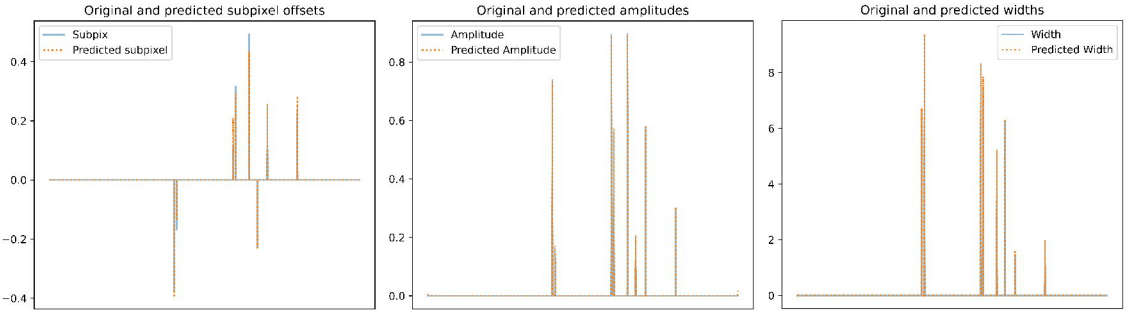

Figure 21: Network parameter predictions are overlayed against their true values with subpixel (left), amplitudes (centre) and widths (right) all demonstrating excellent agreement with the ground truths

The success of the proposed deep learning auto-encoder is demonstrated by the high accuracy with which the model extracts the Lorentzian parameters. As illustrated in Figure 22, the spectrum reconstructed using the parameters predicted by the autoencoder sequence shows a strong visual correspondence to the original input spectrum, indicating that the essential characteristics of the functions have been successfully captured. Confirming this, a direct comparison of the predicted parameters—position, width, and amplitude—against their ground truth values, as depicted in Figure 21, reveals a close alignment. Notably the sub-pixel position appear to be experience a slightly higher error than either the amplitudes or widths. This confirms the model's capability to accurately retrieve the individual Lorentzian parameters for each peak within the a spectrum. While promising, future work will be necessary to investigate similar methods can also be used on data with inherent noise or more complex spectral shapes, while Fig 21 clearly shows that these networks can accommodate overlapping Lorentz functions, increases in the number of functions and maximum values for their width, alongside the introduction of shot noise and convolved Gaussians as is typical in real-world recording scenarios, will

most likely require additional innovation in the model architecture and training.

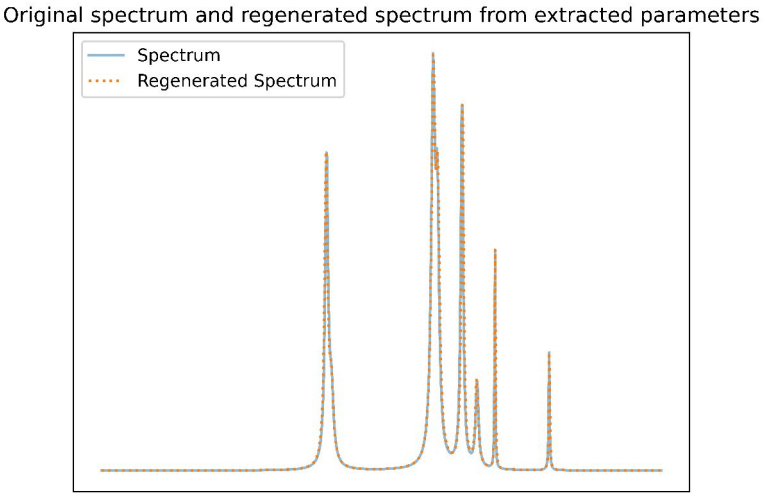

Figure 22: A sample input spectrum is overlayed with a spectrum that has been regenerated from the lorentz parameters output by the auto-encoders showing near perfect agreement

5.8 Conclusions

This chapter has successfully implemented a novel deep learning framework for the decomposition of a spectrum into their constituent Lorentzian parameters: position, width, and amplitude. The primary challenge addressed was the sparse nature of this regression task — transforming a dense 1D spectral input into a sparse set of three-parameter tuples. Our approach is the sequential application of 1D Convolutional Auto-Encoders. With the initial autoencoder generating a peak identification mask which is then utilised in the parameter extraction. The custom masked loss functions constrain the loss calculation only to the active (sparse) regions defined by the mask, thereby reducing significant penalties for near-correct pixel shifted values and thus provides a smoother loss topology.

The performance of this framework was validated on a large, synthetically generated dataset. The results were promising, showing a near-perfect reconstruction of the input spectra from the parameters predicted by the network. This is further vali-

dated by a direct comparison between the predicted parameters and their ground truths. Despite this success, in it's current form it serves only as a proof-of-concept on idealised, noise free data. The next step is to evaluate this model's robustness against real-world imperfections. Future work will therefore focus on enhancing the model to handle signal noise by introducing typical noise models (e.g., shot noise, readout noise) into the training data to develop a noise-resilient architecture. Furthermore, in real world scenarios the signal is subject to instrumental broadening due to spectrometer slit functions. The model should be capable of fitting to these more realistic Voigt profiles. The culmination of which will be the testing of the network on real-world data.

In conclusion, this work establishes the first steps towards a viable deep learning strategy for automated spectral analysis for the deconstruction of the spectrum into its underlying line shapes. By overcoming the sparse output challenge through a sequential masking technique, we have laid the groundwork for a tool that can serve as a robust initial parameter estimator. This method could provide a highly accurate first guess for the Lorentzian parameters thereby effectively removing the requirement for manual initial guesses enabling for fully automated decomposition methods.

6 Conclusions

6.1 Overview

This thesis set out to develop tools for quantitative, chemically-specific histopathology by developing an integrated hyperspectral imaging platform and novel deep learning methods for spectral characterisation. This work has successfully established a framework to achieve this goal, making contributions in both optical system design and computational spectral analysis. The research was motivated by the findings of the comprehensive review in Chapter 2, which noted that while current high-speed methods like Stimulated Raman Histology (SRH) are clinically promising, they typically rely on a small subset of the available chemical information (e.g., two wavenumbers) for virtual staining.

This led to our comparison of two single acquisition full spectrum modalities, namely BCARS and spontaneous Raman, which have complementary strengths and weaknesses in terms of speed and SNR. We posited therefore, that combining both modalities could be a useful strategy. To evaluate the validity of this approach, in Chapter 3, we presented, a direct comparison between BCARS and Spontaneous Raman where we used Cal-1 white blood cells. This work validated the similar spectral shapes and chemical co-localisation of spectral features between the two modalities despite drifts in the wavenumbers and differences in amplitudes. BCARS, despite the contribution of the NRB was capable of capturing similar fundamental chemical information as Spontaneous Raman. This was further underscored with the trade-offs previously outlined in the beginning of Chapter 2. BCARS provides hyperspectral data at speeds orders of magnitude faster making it viable for hyperspectral imaging. In contrast, spontaneous Raman offers a less polluted, NRB-free signal with higher spectral fidelity. This comparison provided the justification for a hybrid system leveraging the strength of both techniques. Additionally we developed a rudimentary segmentation workflow in this chapter to demonstrate the ability of BCARS to aggregate spectra over areas has the potential to improve future cellular

analysis, specifically in the area of flow cytometry.

Building directly on these foundations, Chapter 4 presented the optical design and evaluation of a novel, switchable hyperspectral imaging system. This system integrates BCARS and spontaneous Raman into a single system. Allowing for the direct pixel to pixel comparison of spontaneous Raman with BCARS. It's efficacy was also demonstrated on buccal epethelial cells. Despite significant underperformance of the BCARS for various reasons that can easily be corrected in future iterations, this chapter demonstrated the use of BCARS as an exploratory tool for rapid, wide-field hyperspectral mapping, capable of capturing morphological features and utilising the relatively weak resonant response to target the higher SNR slow acquisition spontaneous Raman for recording the higher fidelity spectral data necessary for quantitative chemical analysis.

Finally in chapter 5 we presented a deep learning approach, moving beyond spectral pattern matching toward quantitative analysis. In this chapter, we presented a preliminary deep learning methodology designed to decompose spectra into their constituent Lorentzian parameters which are directly reflective of the individual molecular resonances. By training this model on physically informed synthetic data and employing strategic masking techniques to overcome the sparse output of the auto encoder, we demonstrate an approach that can retrieve these fundamental parameters when no noise is present in the sample. The natural evolution of this approach will be to include noise in the training datasets and extend the application to real-world Raman and BCARS data.

6.2 Future perspectives

The preliminary works developed in this thesis provide clear avenues for future development;

The significant underperformance of the BCARS system in Chapter 4 provide a clear pathway for further work. The immediate priority being the optimization of the system with a focus on alignment, fully utilising the delivery objective's 1.2 NA

and increasing the laser power to dramatically improve the signal-to-noise ratio. A deformable mirror will also be included in that path to mitigate any astigmatism introduced by the Cassegrain telescope. Once optimized, the switchable system can be applied to its intended target: relevant histological samples. We expect a significantly higher SNR when working with tissue samples, which will be the next focus of this system. This will allow for a demonstration of its full potential, generating rapid, wide-field BCARS maps to identify regions of interest, followed by targeted, high-fidelity SR acquisition in those regions. This workflow could provide detailed chemical targeting beyond that capable of SRS.

The Lorentzian decomposition model from Chapter 5 was highly successful on synthetic data but requires significant innovation to handle real world scenarios. Future work must focus on making this model robust by training it on data that includes typical real world issues, including signal noise, instrumental broadening and complex baselines. A logical progression would be to first apply the network to high-SNR long-acquisition spontaneous Raman data. The deconstruction of the spectrum into a sum of constituent Lorentzian functions could be especially useful for precise characterisation of individual spectral features, enabling detailed analysis of properties like chemical composition, temperature, and density. Further this could lead to an improved approach for the characterisation of disease, multicomponent analysis, or multivariate diagnostic classification models. Multivariate statistical analysis of spectra often relies on projecting the spectrum into an orthogonal set of 'spectra' components as a first step. However these components are often not interpretable in terms of chemistry and lack the fundamental Lorentzian lineshapes. Our approach could offer a means to constrain multivariate spectral analysis to operate within the domain of physically meaningful components, and offer an entirely new research pathway.

The final goal is the complete integration of the hardware and software systems developed in Chapters 4 and 5. The integration of which would enable a novel, high-throughput workflow for extracting physically rich chemical information from

rapidly scanned tissue samples.

Bibliography

[1] Mcnamara, T.; Hennelly, B. In *Spectroscopy For Cancer*; Henelly, B., Ed.; Institute of Physics, 2025; Forthcoming.

[2] Muddiman, R.; McNamara, T.; Harkin, S.; Butler, M.; Hennelly, B. Comparison of broadband-CARS and spontaneous Raman scattering for blood cell analysis. Biomedical Spectroscopy, Microscopy, and Imaging III. 2024; pp 232–241.

[3] McNamara, T.; Hennelly, B. M. Hyperspectral imaging system with switchable spontaneous Raman spectroscopy and broadband CARS. Multimodal Sensing and Artificial Intelligence for Sustainable Future. 2025; pp 346–351.

[4] McNamara, T.; Hennelly, B. M. Characterizing spectral peaks with deep-learning auto encoders. Optical Methods for Inspection, Characterization, and Imaging of Biomaterials VII. 2025; pp 87–92.

[5] Smekal, A. Grundlagen der theoretischen Optik und des allgemeinen dynamischen Prinzips der Lichtstreuung. *Die Naturwissenschaften* **1923**, *11*, 873–875.

[6] Raman, C. V.; Krishnan, K. S. A new type of secondary radiation. *Nature* **1928**, *121*, 501–502.

[7] Landsberg, G. S.; Mandelstam, L. I. Uber die Lichtzerstreuung in Kristallen. *Naturwissenschaften* **1928**, *16*, 557.

[8] Schawlow, A. L.; Townes, C. H. Infrared and optical masers. *Physical Review* **1958**, *112*, 1940–1949.

[9] Maiman, T. H. Stimulated Optical Radiation in Ruby. *Nature* **1960**, *187*, 493–494.

[10] Woodbury, E. J.; Hercher, M. A. A New Type of Light-Scattering From Solids. *Physical Review Letters* **1962**, *9*, 221–222.

[11] Maker, P. D.; Terhune, R. W. Study of Optical Effects Due to an Induced Polarization Third Order in the Electric Field Strength. *Phys. Rev.* **1965**, *137*, A801–A818.

[12] Tolles, W. M.; Nibler, J. W.; McDonald, J. R.; Harvey, A. B. A Review of the Theory and Application of Coherent Anti-Stokes Raman Spectroscopy (CARS). *Applied Spectroscopy* **1977**, *31*, 253–271.

[13] Regnier, P. R.; Taran, J. P. E. On the possibility of measuring gas concentrations by stimulated anti-Stokes scattering. *Applied Physics Letters* **1973**, *23*, 240–242.

[14] Moya, F.; Druet, S. A. J.; Taran, J. P. E. Gas spectroscopy and temperature measurement by coherent Raman anti-Stokes scattering. *Optics Communications* **1975**, *13*, 169–174.

[15] Zumbusch, A.; Holtom, G. R.; Xie, X. S. Three-Dimensional Vibrational Imaging by Coherent Anti-Stokes Raman Scattering. *Physical Review Letters* **1999**, *82*, 4142–4145.

[16] Cheng, J. X.; Volkmer, A.; Book, L. D.; Xie, X. S. Multiplex Coherent Anti-Stokes Raman Scattering Microspectroscopy and Study of Lipid Vesicles. *The Journal of Physical Chemistry B* **2002**, *106*, 8493–8498.

[17] McCulloch, W. S.; Pitts, W. A Logical Calculus of the Ideas Immanent in Nervous Activity. *The Bulletin of Mathematical Biophysics* **1943**, *5*, 115–133.

[18] Kelley, H. J. Gradient Theory of Optimal Flight Paths. *ARS Journal* **1960**, *30*, 947–954, Presented at ARS Semi-Annual Meeting, Los Angeles, May 9–12, 1960.

[19] Seppo, L. Algoritmin kumulatiivinen pyöristysvirhe yksittäisten pyöristysvirheiden Taylor-kehitelmänä. Master's thesis, University of Helsinki, Helsinki, Finland, 1970; In Finnish.

[20] Paul, J. W. Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences. Ph.D. thesis, Harvard University, Cambridge, Massachusetts, 1974.

[21] Rumelhart, D. E.; Hinton, G. E.; Williams, R. J. Learning representations by back-propagating errors. *Nature* **1986**, *323*, 533–536.

[22] Williams, R. J.; Zipser, D. A Learning Algorithm for Continually Running Fully Recurrent Neural Networks. *Neural Computation* **1989**, *1*, 270–280.

[23] Elman, J. L. Finding Structure in Time. *Cognitive Science* **1990**, *14*, 179–211.

[24] LeCun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-Based Learning Applied to Document Recognition. *Proceedings of the IEEE* **1998**, *86*, 2278–2324.

[25] Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Computation* **1997**, *9*, 1735–1780.

[26] Nair, V.; Hinton, G. E. Rectified Linear Units Improve Restricted Boltzmann Machines. Proceedings of the 27th International Conference on Machine Learning (ICML-10). 2010; pp 807–814.

[27] Krizhevsky, A.; Sutskever, I.; Hinton, G. E. ImageNet Classification with Deep Convolutional Neural Networks. Proceedings of the 25th International Conference on Neural Information Processing Systems (NeurIPS 2012). Red Hook, NY, USA, 2012; pp 1097–1105.

[28] Bahdanau, D.; Cho, K.; Bengio, Y. Neural Machine Translation by Jointly Learning to Align and Translate. *arXiv preprint arXiv:1409.0473* **2015**,

[29] Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, ; Polosukhin, I. Attention Is All You Need. Advances in Neural Information Processing Systems. 2017.

[30] Wu, M.; Zhou, J.; Peng, Y.; Wang, S.; Zhang, Y. *Deep Learning for Image Classification: A Review*; 2024; pp 352–362.

[31] Minaee, S.; Boykov, Y.; Porikli, F.; Plaza, A.; Kehtarnavaz, N.; Terzopoulos, D. Image Segmentation Using Deep Learning: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2022**, *44*, 3523–3542.

[32] Zhuang, F.; Cheng, X.; Luo, P.; Pan, S. J.; He, Q. Supervised representation learning: Transfer learning with deep autoencoders. IJCAI. 2015; pp 4119–4125.

[33] Isola, P.; Zhu, J.-Y.; Zhou, T.; Efros, A. A. Image-to-image translation with conditional adversarial networks. Proceedings of the IEEE conference on computer vision and pattern recognition. 2017; pp 1125–1134.

[34] Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18. 2015; pp 234–241.

[35] He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition. 2016; pp 770–778.

[36] Ho, J.; Jain, A.; Abbeel, P. Denoising Diffusion Probabilistic Models. Advances in Neural Information Processing Systems (NeurIPS) 33. 2020; pp 6840–6851.

[37] Schneider, S.; Baevski, A.; Collobert, R.; Auli, M. wav2vec: Unsupervised Pre-training for Speech Recognition. *CoRR* **2019**, *abs/1904.05862*.

[38] Lim, B.; Arık, S. ; Loeff, N.; Pfister, T. Temporal Fusion Transformers for interpretable multi-horizon time series forecasting. *International Journal of Forecasting* **2021**, *37*, 1748–1764.

[39] Dosovitskiy, A.; Beyer, L.; Kolesnikov, A.; Weissenborn, D.; Zhai, X.; Unterthiner, T.; Dehghani, M.; Minderer, M.; Heigold, G.; Gelly, S.; Uszkoreit, J.; Houlsby, N. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. *CoRR* **2020**, *abs/2010.11929*.

[40] Ando, J.; Palonpon, A. F.; Sodeoka, M.; Fujita, K. High-speed Raman imaging of cellular processes. *Current Opinion in Chemical Biology* **2016**, *33*, 16–24, Chemical genetics and epigenetics * Molecular imaging.

[41] Muddiman, R.; Harkin, S.; Butler, M.; Hennelly, B. Broadband CARS hyperspectral classification of single leukemia cells. **2024**,

[42] Cheng, J.; Xie, X. *Coherent Raman Scattering Microscopy*; Series in Cellular and Clinical Imaging; CRC Press, 2016.

[43] Ma, L.; Luo, K.; Liu, Z.; Ji, M. Stain-Free Histopathology with Stimulated Raman Scattering Microscopy. *Analytical Chemistry* **2024**, *96*, 7907–7925, PMID: 38713830.

[44] Nishiyama, R.; Furuya, K.; Tamura, T.; Nakao, R.; Peterson, W.; Hiramatsu, K.; Ding, T.; Goda, K. Fourier Transform Coherent Anti-Stokes Raman Scattering Spectroscopy: A Comprehensive Review. *Analytical Chemistry* **2024**, *96*, 18322–18336, PMID: 39436740.

[45] Kerr, L.; Hennelly, B. A multivariate statistical investigation of background subtraction algorithms for Raman spectra of cytology samples recorded on glass slides. *Chemometrics and Intelligent Laboratory Systems* **2016**, *158*, 61–68.

[46] Vanna, R.; Masella, A.; Bazzarelli, M.; Ronchi, P.; Lenferink, A.; Tresoldi, C.; Morasso, C.; Bedoni, M.; Cerullo, G.; Polli, D., *et al.* High-resolution Raman imaging of¿ 300 patient-derived cells from nine different leukemia subtypes: a global clustering approach. *Analytical chemistry* **2024**,

[47] Kerr, L. T.; Byrne, H. J.; Hennelly, B. M. Optimal choice of sample substrate and laser wavelength for Raman spectroscopic analysis of biological specimen. *Analytical Methods* **2015**, *7*, 5041–5052.

[48] Mochizuki, K.; Kumamoto, Y.; Maeda, S.; Tanuma, M.; Kasai, A.; Takemura, M.; Harada, Y.; Hashimoto, H.; Tanaka, H.; Smith, N. I., *et al.*

High-throughput line-illumination Raman microscopy with multislit detection. *Biomedical Optics Express* **2023**, *14*, 1015–1026.

[49] Cheng, J.-X.; Xie, X. S. Vibrational spectroscopic imaging of living systems: An emerging platform for biology and medicine. *Science* **2015**, *350*, aaa8870.

[50] De la Cadena, A.; Vernuccio, F.; Ragni, A.; Sciortino, G.; Vanna, R.; Ferrante, C.; Pediconi, N.; Valensise, C.; Genchi, L.; Laptenok, S. P., *et al.* Broadband stimulated Raman imaging based on multi-channel lock-in detection for spectral histopathology. *Apl Photonics* **2022**, *7*.

[51] Camp, C. H., Jr; Lee, Y. J.; Heddleston, J. M.; Hartshorn, C. M.; Hight Walker, A. R.; Rich, J. N.; Lathia, J. D.; Cicerone, M. T. High-Speed Coherent Raman Fingerprint Imaging of Biological Tissues. *Nat Photonics* **2014**, *8*, 627–634.

[52] Wang, Z.; O'Dwyer, K.; Muddiman, R.; Ward, T.; Camp Jr, C. H.; Hennelly, B. M. VECTOR: Very deep convolutional autoencoders for non-resonant background removal in broadband coherent anti-Stokes Raman scattering. *Journal of Raman Spectroscopy* **2022**, *53*, 1081–1093.

[53] Muddiman, R.; O'Dwyer, K.; Camp, C. H.; Hennelly, B. Removing non-resonant background from broadband CARS using a physics-informed neural network. *Analytical Methods* **2023**, *15*, 4032–4043.

[54] Tamamitsu, M.; Sakaki, Y.; Nakamura, T.; Podagatlapalli, G. K.; Ideguchi, T.; Goda, K. Ultrafast broadband Fourier-transform CARS spectroscopy at 50,000 spectra/s enabled by a scanning Fourier-domain delay line. *Vibrational Spectroscopy* **2017**, *91*, 163–169.

[55] Francis, A.; Berry, K.; Chen, Y.; Figueroa, B.; Fu, D. Label-free pathology by spectrally sliced femtosecond stimulated Raman scattering (SRS) microscopy. *PLoS One* **2017**, *12*, e0178750.

[56] Sarri, B.; Canonge, R.; Audier, X.; Simon, E.; Wojak, J.; Caillol, F.; Cador, C.; Marguet, D.; Poizat, F.; Giovannini, M., *et al.* Fast stimulated Raman and second harmonic generation imaging for intraoperative gastro-intestinal cancer detection. *Scientific reports* **2019**, *9*, 10052.

[57] Li, Y.; Xiao, X.; Yang, C. Spectral-compressed fiber laser source for broadband-CARS microspectroscopy. 2017 16th International Conference on Optical Communications and Networks (ICOCN). 2017; pp 1–3.

[58] Liu, D.; Hennelly, B. M. Wavenumber calibration protocol for raman spectrometers using physical modelling and a fast search algorithm. *Applied Spectroscopy* **2024**, 00037028241254847.

[59] Barton, S. J.; Hennelly, B. M. An algorithm for the removal of cosmic ray artifacts in spectral data sets. *Applied spectroscopy* **2019**, *73*, 893–901.

[60] Hollon, T. C.; Pandian, B.; Adapa, A. R.; Urias, E.; Save, A. V.; Khalsa, S. S. S.; Eichberg, D. G.; D'Amico, R. S.; Farooq, Z. U.; Lewis, S., *et al.* Near real-time intraoperative brain tumor diagnosis using stimulated Raman histology and deep neural networks. *Nature medicine* **2020**, *26*, 52–58.

[61] Freudiger, C. W.; Min, W.; Saar, B. G.; Lu, S.; Holtom, G. R.; He, C.; Tsai, J. C.; Kang, J. X.; Xie, X. S. Label-free biomedical imaging with high sensitivity by stimulated Raman scattering microscopy. *Science* **2008**, *322*, 1857–1861.

[62] Jiang, C.; Bhattacharya, A.; Linzey, J. R.; Joshi, R. S.; Cha, S. J.; Srinivasan, S.; Alber, D.; Kondepudi, A.; Urias, E.; Pandian, B., *et al.* Rapid automated analysis of skull base tumor specimens using intraoperative optical imaging and artificial intelligence. *Neurosurgery* **2022**, *90*, 758–767.

[63] Orringer, D. A.; Pandian, B.; Niknafs, Y. S.; Hollon, T. C.; Boyle, J.; Lewis, S.; Garrard, M.; Hervey-Jumper, S. L.; Garton, H. J.; Maher, C. O., *et al.* Rapid intraoperative histology of unprocessed surgical specimens via fibre-laser-based

stimulated Raman scattering microscopy. *Nature biomedical engineering* **2017**, *1*, 0027.

[64] Sarri, B.; Appay, R.; Heuke, S.; Poizat, F.; Franchi, F.; Boissonneau, S.; Caillol, F.; Dufour, H.; Figarella-Branger, D.; Giovaninni, M., *et al.* Observation of the compatibility of stimulated Raman histology with pathology workflow and genome sequencing. *Translational Biophotonics* **2021**, *3*, e202000020.

[65] Sarri, B.; Poizat, F.; Heuke, S.; Wojak, J.; Franchi, F.; Caillol, F.; Giovannini, M.; Rigneault, H. Stimulated Raman histology: one to one comparison with standard hematoxylin and eosin staining. *Biomedical optics express* **2019**, *10*, 5378–5384.

[66] Sarri, B.; Chevrier, V.; Poizat, F.; Heuke, S.; Franchi, F.; De Franqueville, L.; Traversari, E.; Ratone, J.-P.; Caillol, F.; Dahel, Y., *et al.* In vivo organoid growth monitoring by stimulated Raman histology. *npj Imaging* **2024**, *2*, 18.

[67] Pekmezci, M.; Morshed, R. A.; Chunduru, P.; Pandian, B.; Young, J.; Villanueva-Meyer, J. E.; Tihan, T.; Sloan, E. A.; Aghi, M. K.; Molinaro, A. M., *et al.* Detection of glioma infiltration at the tumor margin using quantitative stimulated Raman scattering histology. *Scientific reports* **2021**, *11*, 12162.

[68] Orringer, D. A.; Pandian, B.; Niknafs, Y. S.; Hollon, T. C.; Boyle, J.; Lewis, S.; Garrard, M.; Hervey-Jumper, S. L.; Garton, H. J.; Maher, C. O., *et al.* Rapid intraoperative histology of unprocessed surgical specimens via fibre-laser-based stimulated Raman scattering microscopy. *Nature biomedical engineering* **2017**, *1*, 0027.

[69] Eichberg, D. G.; Shah, A. H.; Di, L.; Semonche, A. M.; Jimsheleishvili, G.; Luther, E. M.; Sarkiss, C. A.; Levi, A. D.; Gultekin, S. H.; Komotar, R. J., *et al.* Stimulated Raman histology for rapid and accurate intraoperative diagnosis of CNS tumors: prospective blinded study. *Journal of neurosurgery* **2019**, *134*, 137–143.

[70] Hollon, T. C.; Lewis, S.; Pandian, B.; Niknafs, Y. S.; Garrard, M. R.; Garton, H.; Maher, C. O.; McFadden, K.; Snuderl, M.; Lieberman, A. P., *et al.* Rapid intraoperative diagnosis of pediatric brain tumors using stimulated Raman histology. *Cancer research* **2018**, *78*, 278–289.

[71] Einstein, E. H.; Ablyazova, F.; Rosenberg, A.; Harshan, M.; Wahl, S.; Har-El, G.; Constantino, P. D.; Ellis, J. A.; Boockvar, J. A.; Langer, D. J., *et al.* Stimulated Raman histology facilitates accurate diagnosis in neurosurgical patients: A one-to-one noninferiority study. *Journal of Neuro-Oncology* **2022**, *159*, 369–375.

[72] Williams, R. M.; Zipfel, W. R.; Webb, W. W. Interpreting second-harmonic generation images of collagen I fibrils. *Biophysical journal* **2005**, *88*, 1377–1386.

[73] Jing, Y.; Yang, Y.; Feng, Z.; Ye, J.; Yu, Y.; Song, M. Neural style transfer: A review. *IEEE transactions on visualization and computer graphics* **2019**, *26*, 3365–3385.

[74] Gatys, L. A. A Neural Algorithm of Artistic Style. *arXiv preprint ArXiv:1508.06576* **2015**,

[75] Zhu, J.-Y.; Park, T.; Isola, P.; Efros, A. A. Unpaired image-to-image translation using cycle-consistent adversarial networks. Proceedings of the IEEE international conference on computer vision. 2017; pp 2223–2232.

[76] Pradhan, P.; Meyer, T.; Vieth, M.; Stallmach, A.; Waldner, M.; Schmitt, M.; Popp, J.; Bocklitz, T. Computational tissue staining of non-linear multimodal imaging using supervised and unsupervised deep learning. *Biomedical Optics Express* **2021**, *12*, 2280–2298.

[77] Liu, Z.; Chen, L.; Cheng, H.; Ao, J.; Xiong, J.; Liu, X.; Chen, Y.; Mao, Y.; Ji, M. Virtual formalin-fixed and paraffin-embedded staining of fresh brain

tissue via stimulated Raman CycleGAN model. *Science Advances* **2024**, *10*, eadn3426.

[78] Orringer, D. A.; Pandian, B.; Niknafs, Y. S.; Hollon, T. C.; Boyle, J.; Lewis, S.; Garrard, M.; Hervey-Jumper, S. L.; Garton, H. J.; Maher, C. O., *et al.* Rapid intraoperative histology of unprocessed surgical specimens via fibre-laser-based stimulated Raman scattering microscopy. *Nature biomedical engineering* **2017**, *1*, 0027.

[79] Hollon, T. C.; Lewis, S.; Pandian, B.; Niknafs, Y. S.; Garrard, M. R.; Garton, H.; Maher, C. O.; McFadden, K.; Snuderl, M.; Lieberman, A. P., *et al.* Rapid intraoperative diagnosis of pediatric brain tumors using stimulated Raman histology. *Cancer research* **2018**, *78*, 278–289.

[80] Liu, Z.; Su, W.; Ao, J.; Wang, M.; Jiang, Q.; He, J.; Gao, H.; Lei, S.; Nie, J.; Yan, X., *et al.* Instant diagnosis of gastroscopic biopsy via deep-learned single-shot femtosecond stimulated Raman histology. *Nature communications* **2022**, *13*, 4050.

[81] Tipping, W. J.; Wilson, L. T.; An, C.; Leventi, A. A.; Wark, A. W.; Wetherill, C.; Tomkinson, N. C.; Faulds, K.; Graham, D. Stimulated Raman scattering microscopy with spectral phasor analysis: applications in assessing drug–cell interactions. *Chemical science* **2022**, *13*, 3468–3476.

[82] Ellis, D. I.; Goodacre, R. Metabolic fingerprinting in disease diagnosis: biomedical applications of infrared and Raman spectroscopy. *Analyst* **2006**, *131*, 875–885.

[83] Ellis, D. I.; Cowcher, D. P.; Ashton, L.; O'Hagan, S.; Goodacre, R. Illuminating disease and enlightening biomedicine: Raman spectroscopy as a diagnostic tool. *Analyst* **2013**, *138*, 3871–3884.

[84] Stone, N.; Matousek, P. Advanced transmission Raman spectroscopy: a

promising tool for breast disease diagnosis. *Cancer Research* **2008**, *68*, 4424–4430.

[85] Kerr, L. T.; Domijan, K.; Cullen, I.; Hennelly, B. M. Applications of Raman spectroscopy to the urinary bladder for cancer diagnostics. *Photonics & Lasers in Medicine* **2014**, *3*, 193–224.

[86] Jermyn, M.; Mok, K.; Mercier, J.; Desroches, J.; Pichette, J.; Saint-Arnaud, K.; Bernstein, L.; Guiot, M.-C.; Petrecca, K.; Leblond, F. Intraoperative brain cancer detection with Raman spectroscopy in humans. *Science translational medicine* **2015**, *7*, 274ra19–274ra19.

[87] McKinnon, K. M. Flow cytometry: an overview. *Current protocols in immunology* **2018**, *120*, 5–1.

[88] Toseland, C. P. Fluorescent labeling and modification of proteins. *Journal of chemical biology* **2013**, *6*, 85–95.

[89] O'Dwyer, K.; Domijan, K.; Dignam, A.; Butler, M.; Hennelly, B. M. Automated Raman micro-spectroscopy of epithelial cell nuclei for high-throughput classification. *Cancers* **2021**, *13*, 4767.

[90] Traynor, D.; Duraipandian, S.; Bhatia, R.; Cuschieri, K.; Tewari, P.; Kearney, P.; D'Arcy, T.; O'Leary, J. J.; Martin, C. M.; Lyng, F. M. Development and validation of a Raman spectroscopic classification model for Cervical Intraepithelial Neoplasia (CIN). *Cancers* **2022**, *14*, 1836.

[91] Palonpon, A. F.; Sodeoka, M.; Fujita, K. Molecular imaging of live cells by Raman microscopy. *Current opinion in chemical biology* **2013**, *17*, 708–715.

[92] Potma, E.; Srinivasan, G. Tissue imaging with coherent anti-Stokes Raman scattering microscopy. 2010.

[93] Camp Jr, C. H.; Lee, Y. J.; Heddleston, J. M.; Hartshorn, C. M.; Walker, A.

R. H.; Rich, J. N.; Lathia, J. D.; Cicerone, M. T. High-speed coherent Raman fingerprint imaging of biological tissues. *Nature photonics* **2014**, *8*, 627–634.

[94] Camp Jr, C. H.; Cicerone, M. T. Chemically sensitive bioimaging with coherent Raman scattering. *Nature photonics* **2015**, *9*, 295–305.

[95] Muddiman, R.; Hennelly, B. Broadband CARS high-throughput single-cell imaging. EPJ Web of Conferences. 2023; p 03018.

[96] O'Dwyer, K.; Domijan, K.; Dignam, A.; Butler, M.; Hennelly, B. M. Automated Raman micro-spectroscopy of epithelial cell nuclei for high-throughput classification. *Cancers* **2021**, *13*, 4767.

[97] Hutsebaut, D.; Vandenabeele, P.; Moens, L. Evaluation of an accurate calibration and spectral standardization procedure for Raman spectroscopy. *Analyst* **2005**, *130*, 1204–1214.

[98] Liu, D.; Hennelly, B. M. Improved wavelength calibration by modeling the spectrometer. *Applied Spectroscopy* **2022**, *76*, 1283–1299.

[99] Liu, D.; Hennelly, B. M. Wavenumber Calibration Protocol for Raman Spectrometers Using Physical Modelling and a Fast Search Algorithm. *Applied Spectroscopy* **Accepted for publication March 2024**,

[100] Zimmermann, B.; Kohler, A. Optimizing Savitzky–Golay parameters for improving spectral resolution and quantification in infrared spectroscopy. *Applied spectroscopy* **2013**, *67*, 892–902.

[101] Khmaladze, A.; Jasensky, J.; Price, E.; Zhang, C.; Boughton, A.; Han, X.; Seeley, E.; Liu, X.; Holl, M. M. B.; Chen, Z. Hyperspectral imaging and characterization of live cells by broadband coherent anti-Stokes Raman scattering (CARS) microscopy with singular value decomposition (SVD) analysis. *Applied spectroscopy* **2014**, *68*, 1116–1122.

[102] Liu, Y.; Lee, Y. J.; Cicerone, M. T. Broadband CARS spectral phase retrieval using a time-domain Kramers–Kronig transform. *Optics letters* **2009**, *34*, 1363–1365.

[103] Camp Jr, C.; Lee, Y.; Cicerone, M. Quantitative, Comparable Coherent Anti-Stokes Raman Scattering (CARS) Spectroscopy: Correcting Errors in Phase Retrieval. *Journal of Raman Spectroscopy* **2016**, *47*, 408–415.

[104] Abdullah Chandra Sekhar Talari, S. R., Zanyar Movasaghi; ur Rehman, I. Raman Spectroscopy of Biological Tissues. *Applied Spectroscopy Reviews* **2015**, *50*, 46–111.

[105] Valensise, C. M.; Giuseppi, A.; Vernuccio, F.; De la Cadena, A.; Cerullo, G.; Polli, D. Removing non-resonant background from CARS spectra via deep learning. *APL Photonics* **2020**, *5*.

[106] Muddiman, R.; Hennelly, B. Deep-learning autoencoders for unsupervised BCARS chemical imaging. Optics in Health Care and Biomedical Optics XIII. 2023; pp 375–389.

[107] Neuville, D. R.; De Ligny, D.; Henderson, G. S. Advances in Raman spectroscopy applied to earth and material sciences. *Reviews in mineralogy and geochemistry* **2014**, *78*, 509–541.

[108] Altangerel, N.; Ariunbold, G. O.; Gorman, C.; Alkahtani, M. H.; Borrego, E. J.; Bohlmeyer, D.; Hemmer, P.; Kolomiets, M. V.; Yuan, J. S.; Scully, M. O. In vivo diagnostics of early abiotic plant stress response via Raman spectroscopy. *Proceedings of the National Academy of Sciences* **2017**, *114*, 3393–3396.

[109] Dodo, K.; Fujita, K.; Sodeoka, M. Raman spectroscopy for chemical biology research. *Journal of the American Chemical Society* **2022**, *144*, 19651–19667.

[110] Liu, Y.; Lee, Y. J.; Cicerone, M. T. Broadband CARS spectral phase retrieval using a time-domain Kramers–Kronig transform. *Optics letters* **2009**, *34*, 1363–1365.

[111] Muddiman, R.; O'Dwyer, K.; Tang, Z.; Hennelly, B. Field guide for building a broadband CARS system for biomedical applications. Optics in Health Care and Biomedical Optics XI. 2021; pp 242–258.

[112] Tipping, W. J.; Wilson, L. T.; An, C.; Leventi, A. A.; Wark, A. W.; Wetherill, C.; Tomkinson, N. C.; Faulds, K.; Graham, D. Stimulated Raman scattering microscopy with spectral phasor analysis: applications in assessing drug–cell interactions. *Chemical science* **2022**, *13*, 3468–3476.

[113] Xie, S.; Iglesia, E.; Bell, A. T. Effects of temperature on the Raman spectra and dispersed oxides. *The Journal of Physical Chemistry B* **2001**, *105*, 5144–5152.

[114] Takeuchi, M.; Kajimoto, S.; Nakabayashi, T. Experimental evaluation of the density of water in a cell by Raman microscopy. *The journal of physical chemistry letters* **2017**, *8*, 5241–5245.

[115] Xu, Z.; He, Z.; Song, Y.; Fu, X.; Rommel, M.; Luo, X.; Hartmaier, A.; Zhang, J.; Fang, F. Topic review: application of Raman spectroscopy characterization in micro/nano-machining. *Micromachines* **2018**, *9*, 361.

[116] Vogt, F. P.; Bonaccini Calia, D.; Hackenberg, W.; Opitom, C.; Comin, M.; Schmidtobreik, L.; Smoker, J.; Blanchard, I.; Espinoza Contreras, M.; Aranda, I., *et al.* Detection and Implications of Laser-Induced Raman Scattering at Astronomical Observatories. *Physical Review X* **2017**, *7*, 021044.

[117] Talari, A. C. S.; Movasaghi, Z.; Rehman, S.; Rehman, I. U. Raman spectroscopy of biological tissues. *Applied spectroscopy reviews* **2015**, *50*, 46–111.

[118] Han, K.; Boziki, A.; Tkatchenko, A.; Berryman, J. T. TIHI Toolkit: A Peak

Finder and Analyzer for Spectroscopic Data. *ACS omega* **2024**, *9*, 49397–49410.

[119] Kensert, A.; Bosten, E.; Collaerts, G.; Efthymiadis, K.; Van Broeck, P.; Desmet, G.; Cabooter, D. Convolutional neural network for automated peak detection in reversed-phase liquid chromatography. *Journal of Chromatography A* **2022**, *1672*, 463005.

[120] Chen, M.; Shi, X.; Zhang, Y.; Wu, D.; Guizani, M. Deep feature learning for medical image analysis with convolutional autoencoder neural network. *IEEE transactions on big data* **2017**, *7*, 750–758.

[121] Jaiswal, G.; Rani, R.; Mangotra, H.; Sharma, A. Integration of hyperspectral imaging and autoencoders: Benefits, applications, hyperparameter tunning and challenges. *Computer Science Review* **2023**, *50*, 100584.

[122] Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18. 2015; pp 234–241.

[123] Passos, D.; Mishra, P. A tutorial on automatic hyperparameter tuning of deep spectral modelling for regression and classification tasks. *Chemometrics and Intelligent Laboratory Systems* **2022**, *223*, 104520.