Leveraging mid-infrared spectroscopic imaging and deep learning for tissue subtype classification in ovarian cancer

Journal:	*Analyst*
Manuscript ID	AN-ART-06-2022-001035.R1
Article Type:	Paper
Date Submitted by the Author:	27-Sep-2022
Complete List of Authors:	Gajjela, Chalapathi ; University of Houston, Electrical and Computer Engineering Brun, Matthew; Rice University Mankar, Rupali; University of Houston, Electrical and Computer Engineering Corvigno, Sara; The Univ. of Texas M. D. Anderson Ctr., Cancer Systems Imaging Kennedy, Noah; Milwaukee School of Engineering Zhong, Yanping ; The Univ. of Texas M. D. Anderson Ctr., Cancer Systems Imaging Liu, Jinsong; The Univ. of Texas M. D. Anderson Ctr., Department of Pathology Sood, Anil; The Univ. of Texas M. D. Anderson Ctr., Cancer Systems Imaging Mayerich, David; University of Houston System, Electrical and Computer Engineering Berisha, Sebastian; Milwaukee School of Engineering Reddy, Rohith; University of Houston, Electrical and Computer Engineering

Journal Name

ARTICLE TYPE

Leveraging mid-infrared spectroscopic imaging and deep learning for tissue subtype classification in ovarian cancer[†]

Chalapathi Charan Gajjela,[a] Matthew Brun,[b] Rupali Mankar,[a] Sara Corvigno,[d] Noah Kennedy,[c] Yanping Zhong,[d] Jinsong Liu,[d] Anil K. Sood,[d] David Mayerich,[a] Sebastian Berisha,[c] and Rohith Reddy[*a]

Mid-infrared spectroscopic imaging (MIRSI) is an emerging class of label-free techniques being leveraged for digital histopathology. Modern histopathologic identification of ovarian cancer involves tissue staining followed by morphological pattern recognition. This process is time-consuming, subjective, and requires extensive expertise. This paper presents the first label-free, quantitative, and automated histological recognition of ovarian tissue subtypes using a new MIRSI technique. This technique, called optical photothermal infrared (O-PTIR) imaging, provides a $10\times$ enhancement in spatial resolution relative to prior instruments. It enables sub-cellular spectroscopic investigation of tissue at biochemically important fingerprint wavelengths. We demonstrate that enhanced resolution of sub-cellular features, combined with spectroscopic information, enables reliable classification of ovarian cell subtypes achieving a classification accuracy of 0.98. Moreover, we present statistically robust validation from 78 patient samples with over 60 million data points. We show that sub-cellular resolution from five wavenumbers is sufficient to outperform state-of-the-art diffraction-limited techniques from up to 235 wavenumbers. We also propose two quantitative biomarkers based on the relative quantities of epithelium and stroma that exhibits efficacy in early cancer diagnosis. This paper demonstrates that combining deep learning with intrinsic biochemical MIRSI measurements enables quantitative evaluation of cancerous tissue, improving the rigor and reproducibility of histopathology.

1 Introduction

Epithelial ovarian cancer is the leading cause of death among gynecological malignancies in the United States. Serous ovarian cancer, its most common subtype, is often diagnosed at a late stage (III or IV), where 5-year survival is 51% and 29%, respectively.[1] Standard treatments involve surgery and at least six courses of chemotherapy.[2] Several novel compounds have been studied and approved over the past 20 years; however, none substantially modify overall survival.[3] The most decisive prognostic factor remains the complete eradication of neoplastic tissue through radical surgery.[4–6] Outcomes are affected by (1) late diagnosis resulting in unresectable disease and (2) unclear identification of neoplastic margins. Therefore, objective and early identification of neoplastic tissue are essential for optimal surgical attempts.

Recent advances reveal the complex organization of the ovarian tumor microenvironment, highlighting inter-cellular pathways[7] as potential treatment targets. New methods quantifying biomolecular characteristics reveal detailed structural and molecular changes that may reveal novel therapeutic targets. The surgical pathological staging systems (International Federation of Gynecology and Obstetrics staging, FIGO staging) continue to be the most crucial tool for determining stage in ovarian cancer.[8] The FIGO staging criteria for cancer are based on the extent of spreading of tumor cells from the ovary. This is determined by using contrast-inducing stains on biopsy sections followed by microscopic examination by a pathologist, which has been the current clinical standard for ovarian cancer diagnosis. Hematoxylin and eosin (H&E) stain is widely used to identify cellular and extracellular components. *Epithelial* carcinoma is the most common histologic type, accounting for about 90 percent of cancers of the ovary, fallopian tube, and peritoneum.[9,10] In high-grade serous carcinoma (HGSC), a pathologist identifies various architectural patterns, including complex papillary, glandular, microcystic, and solid patterns. HGSC infiltrates, destroys, and/or replaces the normal *stroma*. Therefore, histological identification of cellular subtypes is an important step[11] in ovarian cancer diagnosis and

a Address, 4226 Martin Luther King Boulevard, N308 Engineering Building 1, Houston TX, 77584, USA; E-mail: rkreddy@uh.edu

prognosis.

Inter-pathologist variability is a significant challenge,[12] and grading schemes have been proposed to reduce this variability.[13–15] However, these methods have only been successful in resource-rich hospitals with comprehensive training.[16] Automated and semi-automated techniques to reduce inter-pathologist variability, especially in lower-resource settings, are critical for equitable care.[17–19] Automated tissue subtyping is an essential step in this effort.

Automated tissue classification into the epithelium and stroma subtypes is challenging, and several techniques have been proposed. Most use H&E[20–22] and immunohistochemical staining[23] combined with machine learning (ML). Staining quality and variability can confound ML and lead to inconsistent results.[18] Immunofluorescence staining has effectively detected circulating tumor cells that can form micro-metastases in other organs beyond the original tumor site.[24][25] But the specificity of stains for a biomarker/cell type can render it ineffective in detecting a different biomarker/cell type. However, this issue has been addressed recently by employing sequential fluorescent quenching and re-staining[26], effectively allowing subsequent fluorescence staining targeting different markers multiple times.[27] However, this method is limited by the number of antibodies that can be used for a limited set of stains, and the process of staining itself can cause a change in the structure and biochemical composition of the tissue. Our goal is to perform label-free recognition of tissue subtypes without using chemical contrast agents. Moreover, we obtain intrinsic quantitative and repeatable biochemical measurements that are independent of operator tissue processing.

Spectroscopic techniques are used widely to identify molecules in chemical and biochemical analysis[28] with high sensitivity and specificity. Vibrational spectroscopy is used routinely to identify organic biomolecules by matching measurements to large commercial spectral libraries containing over 260,000 spectra.[29] Prior work on label-free ovarian tissue analysis has utilized spectroscopy. Raman spectroscopy,[30–32] conventional Fourier Transform Infrared (FTIR) spectroscopy,[33–35] and attenuated total reflection (ATR) FTIR spectroscopy[36,37] have been applied to detect and diagnose ovarian cancer. However, prior work lacked spatial specificity and required long data collection times. MALDI imaging[38] has also been used to analyze ovarian histotypes; however, this technique destroys the sample. Second-harmonic generation (SHG) can identify collagen in the stroma,[39–41] and multiphoton microscopy[42] has been used on murine tissue. Raman imaging has been used for ovarian cancer diagnosis and tissue analysis, often with added nanoparticles[43,44] to obtain more robust signals. Taken together, none of these techniques provide classified images in an automated, label-free, quantitative, and non-destructive manner.

Mid-infrared spectroscopic imaging (MIRSI) can extract spectral and spatial information without using contrast agents by utilizing intrinsic biochemical properties of tissue. This technology is non-destructive and therefore compatible with other technologies.[45] Fourier transform infrared (FTIR) spectroscopic imaging, the best known MIRSI technology, can classify cell subtypes in a variety of diseases, including breast,[46] lung,[47] prostate,[48] and colon[49] cancers. We hypothesize that it is also helpful for ovarian cancer tissue subtyping. HGSC is the marked cytologic atypia with prominent mitotic activity in ovarian tissue. The atypical nuclei are hyperchromatic with an over threefold variation in nuclear size. Phosphate spectroscopic peaks (1080, 1201, 1236, 1262 cm^{-1}) corresponding to nucleic acids strongly correlate to mitotic activity, and spectroscopic imaging has shown increased phosphate signals in a variety of cancers.[50–52].

FTIR imaging is limited by the diffraction of mid-infrared light (2.5 μm - 11 μm). Since typical cells are of 5 μm in size, FTIR imaging cannot provide sub-cellular information potentially important for analysis. Optical-photothermal infrared (O-PTIR) imaging overcomes this resolution limitation. This technique combines a visible laser beam and a mid-infrared beam in a pump-probe architecture and estimates the sample's absorbance by measuring the change in intensity of the visible laser caused by the photothermal effect. Therefore the image resolution is determined by the wavelength of visible light (0.5 μm), which is much shorter than the wavelength of IR light incident on the sample, allowing 5x to 22x improvement in spatial resolution.[53,54] The improved spatial resolution is comparable to an optical microscope image, as demonstrated in Figure 1. This technology has been used previously for studying the chemistry of inorganic 2D perovskite and allowed us to understand the edge emission phenomenon in inorganic 2D perovskite.[55] O-PTIR has also been successfully used in chemical imaging of live human ovarian cancer cells,[56] where high resolution is needed to analyze sub-cellular structures in a small area. However, our work is the first large-scale (78 cancer patients) study of clinical ovarian tissue biopsy samples, each with large sample areas (1 mm diameter each).

Machine learning algorithms, including random forest (RF) and Bayesian classifiers, use individual spectra to classify tissue from MIRSI data. These approaches do not leverage spatial information, although MIRSI provides both spatial and spectroscopic information. Convolutional neural networks (CNNs) are deep learning architectures that learn local spatial features. CNNs have been successful in hyperspectral[57] and mid-infrared spectroscopic classification of breast cancer tissues.[58] Traditional CNNs consist of alternating convolution and pooling layers, followed by a fully-connected classifier. In this paper, we use CNNs to determine the impact of improved O-PTIR resolution on classification. Previous work in ovarian cancer spectroscopy relied on point spectra to identify malignant tissue or grade tumors.[32,36] We report the first application of the O-PTIR tissue classification and the first reported application of MIRSI and deep learning to ovarian histology.

1.1 O-PTIR imaging

Photothermal microscopy obtains a measurement of sample absorbance by estimating the thermal expansion caused by the absorption of infrared light using a co-localized visible laser beam. The visible and IR laser are incident on the sample collinearly, as shown in Figure2. The thermal expansion caused in the sample due to IR absorbance causes variation in the refractive index due to the photothermal effect. This change in refractive index is de-
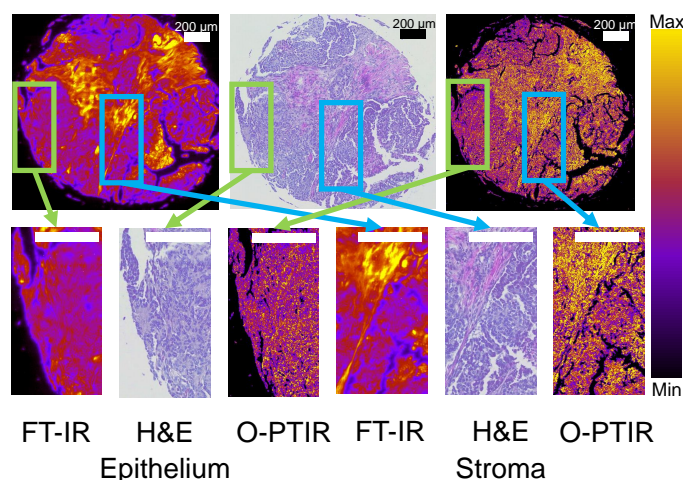
Fig. 1 Spatial differences between different cell types in FT-IR image on the left, H & E image in the center and O-PTIR image on the right. Cropped regions around pixels from the same core in FT-IR and O-PTIR images collected at 1664 cm^{-1}.

tected by measuring the change in intensity of the back-reflected green laser (visible laser) using a point detector.
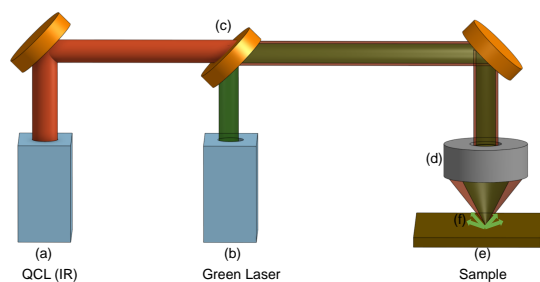

Fig. 2 Schematic of the optical path of the IR and green (532 nm) lasers in our O-PTIR instrument. A pulsed Quantum Cascade Laser (QCL) shown in (a) is the source of mid-IR light which acts as a pump causing photothermal expansion at the sample. A Continuous Wave (CW) green laser shown in (b) is incident collinearly on the sample and acts as a probe beam. A dichroic mirror (c) combines the green and QCL light and focuses them on the sample (e) using a reflective Cassegrain objective (d). The modulation in the intensity of the green light (f) scattered back from the sample enables the measurement of its IR absorbance.

1.1.1 Improved Spatial Resolution

FTIR has been the standard spectroscopic imaging technique for characterizing a material's chemistry. While FTIR provides spectral data across all mid-infrared wavenumbers, its spatial resolution is diffraction-limited by the long wavelengths of light resulting in modest image quality. O-PTIR overcomes this limitation and provides higher spatial resolution ($0.5\,\mu m$) images, with data quality comparable to H&E stained microscopy images. Figure 1 compares the image quality of O-PTIR, FTIR, and H&E on the same cancer tissue. The improved spatial resolution of O-PTIR relative to FTIR is evident. Furthermore, O-PTIR data quality is comparable to microscopy data after H&E staining on an adjacent section. Finer spatial details in the epithelium and stromal tissue regions are also observed in the O-PTIR data presented in Figure

1.

2 Materials and Methods

An ovarian biopsy tissue microarray (TMA) was obtained from Biomax US, Rockville, MD (TMA ID: BC11115c) and imaged using O-PTIR. The TMA consists of paraffin-embedded cores mounted on a 1 mm thickness CaF2 substrate. These cores are from separate patients with cases of normal, hyperplastic, dysplastic, and malignant tumors. The patient cohort was composed of women aged 29 to 69; ovarian tumor stages varied between stage I to stage IIIC; histological subtypes include clear cell carcinoma, high-grade serous carcinoma, and Mucinous adenocarcinoma. The deparaffiniztion was done following the protocol along the lines as described in [50] before undergoing O-PTIR imaging. The paraffin-embedded samples were deparaffinized by washing the sample in 100 percent xylene twice for 5 minutes each and then with 100 percent ethanol thrice. The corresponding adjacent histological section was stained with H&E and examined by an expert pathologist. Cell subtypes were identified across disease stages. We trained a random forest (RF) classifier, support vector machine (SVM), k-nearest neighbor (KNN), and a CNN model by using the cores on the left half of TMA for training and testing on the other cores on the right half of TMA, ensuring that training pixels and testing pixels come from mutually exclusive cores on TMA.

2.1 Data acquisition

FTIR images were acquired using an Agilent Stingray imaging system equipped with a 680-IR spectrometer connected to a 620-IR imaging microscope with a numerical aperture of 0.62. Each core was imaged with 16 co-additions in transmission mode at a spectral resolution of $8\,cm^{-1}$ truncated from $902\,cm^{-1}$ to $3892\,cm^{-1}$, and a pixel size of $1.1\,\mu m$.We collected the background scan at 128 co-additions and ratioed to the single beam data to remove spectral contributions from the substrate, atmosphere, and globar source.

The adjacent H&E stained TMA was imaged with a Nikon inverted optical microscope with a 10X, 0.4NA objective in the brightfield mode, and has diffraction-limited spatial resolution in the visible range ($0.4\,\mu m$ - $0.7\,\mu m$).

The O-PTIR dataset was acquired using a Photothermal mIRage microscope with a silicon photodiode, a pixel size of $0.5\,\mu m \times 0.5\,\mu m$ and a 0.65 numerical aperture. A Quantum Cascade Laser (QCL) source sweeps through the range of $902\,cm^{-1}$ to $1898\,cm^{-1}$. Each core was imaged at five selected wavenumbers ($1162\,cm^{-1}$, $1234\,cm^{-1}$, $1396\,cm^{-1}$, $1540\,cm^{-1}$, and $1661\,cm^{-1}$). An image of the entire TMA acquired at the Amide I band is shown in Figure 3. Background spectra are collected with 10 co-additions and used to normalize the raw data to calculate the IR absorbance at each band. Band images are then normalized with Amide I ($1664\,cm^{-1}$) to bring the data range between 0 and 1. Note that some tissue biopsies are missing because the collected data was corrupted. We used an Aperio Scanscope system to acquire the Light microscope images of the whole slide's H&E chemically stained sections.
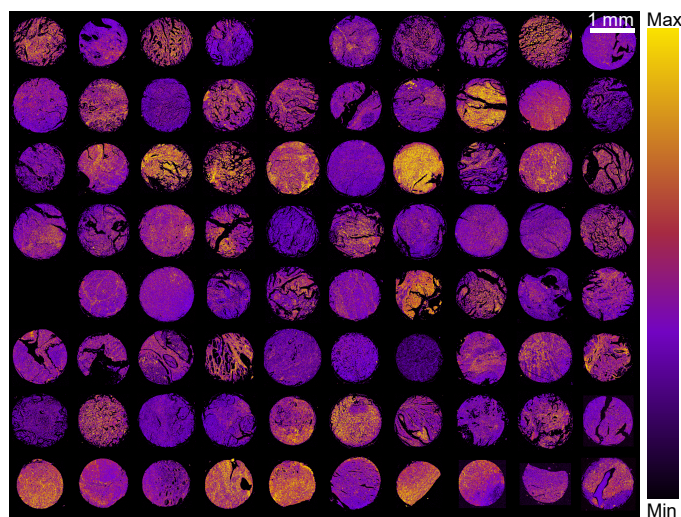
Fig. 3 O-PTIR microarray (8×10) shown at band $1664\,\text{cm}^{-1}$. Data from 78 ovarian cancer patients is shown. The biochemical variations in tissue are evident from the differences in color in the Figure. Machine learning algorithms combine biochemical data from multiple bands enabling tissue subtype identification and disease diagnosis.

2.2 Feature Selection

Since the O-PTIR signal is detected using a point detector, the time taken to collect an image of a single core varies between 90 to 100 minutes per wavenumber.[59] Collecting a hyperspectral data cube of a core at all wavenumbers ($900\,\text{cm}^{-1}$ to $1900\,\text{cm}^{-1}$ at $2\,\text{cm}^{-1}$ spacing) would take approximately 35 day. We, therefore, collect fewer bands, focusing on acquiring important biochemical information. These wavenumbers are determined by analyzing FTIR spectra of ovarian tissue to determine absorbance peaks corresponding to functional groups relevant to ovarian tissue analysis based on prior work. We acquired O-PTIR data at wavenumbers $1162\,\text{cm}^{-1}$, $1234\,\text{cm}^{-1}$, $1396\,\text{cm}^{-1}$, $1540\,\text{cm}^{-1}$, and $1661\,\text{cm}^{-1}$ (Figure 4), which correspond to glycogen, amide III, nucleic acids and lipids, amide II, and amide I, respectively.[50,60]

2.3 Data annotation

Two pathologists labeled areas in tissue cores as *stroma* or *epithelium* using H&E stained microscopy data. H&E with IR images were aligned manually, and the labels were transferred to O-PTIR images to create annotated data for machine learning. The TMA was divided into two halves. The right half was used for training and the left for testing, with an equal number of cores in each cohort.

2.4 Classification Models and Hyperparameters

The SVM classifier was trained on $10,000$ randomly selected pixels per class from the training dataset (Table 1). An equal number of data points are drawn from each class to balance the training data and optimize classifier performance. The RF classifier was trained with 100 trees using $10,000$ samples per class. Classifier inputs consisted of five-element vectors containing the IR absorption values at each pixel and the corresponding pixel label.

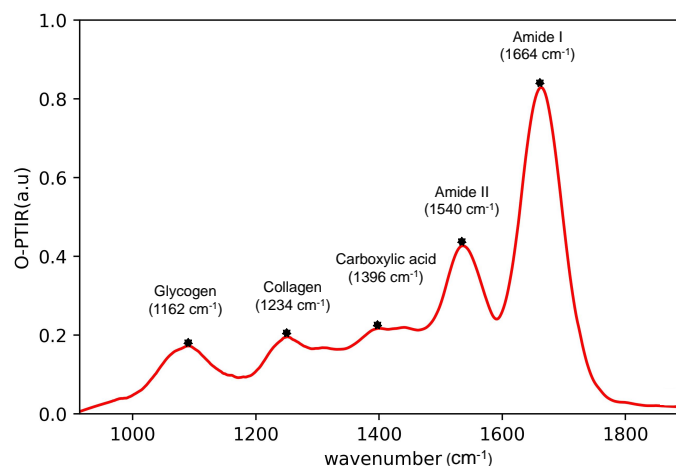The CNN model uses the same general structure as our previous

Fig. 4 The Figure shows the absorption spectrum of ovarian tissue collected using O-PTIR. The spectrum shows the IR absorption values (Y-axis) for the wavenumbers (X-axis) in the fingerprint region. The marked peaks on the spectrum correspond to biochemically relevant functional groups[60] of glycogen at $1162\,\text{cm}^{-1}$, amide III at $1234\,\text{cm}^{-1}$, nucleic acids and lipids at $1396\,\text{cm}^{-1}$, amide II at $1540\,\text{cm}^{-1}$, and amide I at $1661\,\text{cm}^{-1}$.

Table 1 The total number of O-PTIR pixels in the training and testing datasets separated by class is presented. The TMA is split in half to create the training and testing cohorts. A small, random data subset is chosen during the first training step, and the classifier is optimized. Equal numbers of pixels are selected from each class to prevent class bias in training. $10,000$ O-PTIR pixels per class are used in the SVM and RF classifiers and $400,000$ pixels per class for CNNs.

Class	Training	Testing
Epithelium	$22,766,257$	$19,249,625$
Stroma	$11,001,575$	$8,719,719$
Total	$33,767,832$	$27,969,344$

breast classifier[58]. We optimized the network for O-PTIR data classification using the following parameters. Inputs are cropped into 32×32 regions around the center pixel to leverage the local spatial information. The network consisted of: (1) a convolution layer with 32-feature maps, (2) a 2×2 max-pooling layer, (3) a convolution layer with 32-feature maps, (4) a convolution layer with 64-feature maps, (5) and finally a fully connected layer with 64 nodes and a softmax output for class probabilities. The O-PTIR network contains an additional 32-feature convolution layer and 2×2 max pooling layer before the 64-feature convolution layers to adjust to the larger input size, ensuring that the feature size in the final layers is equivalent in both models (Figure 5). The same architecture is optimized for the FTIR data to make an apt comparison of the classification results.

The following hyperparameters are used for our CNN models:

1. **Optimization:** The Adam optimizer was used.[61]

2. **Dropout:** The networks had a dropout layer before the fully connected layer with a keep probability of 0.5. This layer aids regularization and prevents overfitting by randomly disabling nodes in the first and last layers in training. The dropout layer was included before the softmax layer and before the first 64-feature convolution layer.
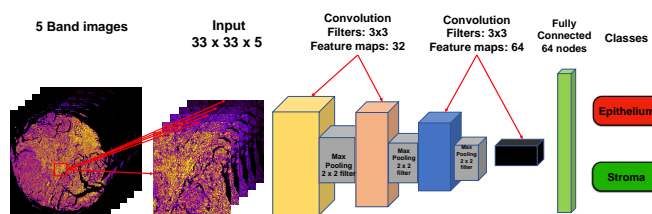
Fig. 5 Schematic of the CNN architecture used for classification of O-PTIR data. A spatial region of size 33×33 is cropped around each pixel. Data cubes of size $33 \times 33 \times 5$ are fed into the first convolution layer. Each input is convolved with filters of size 3×3 outputting 32 feature maps. The following layer is a max pooling layer, which reduces the spatial dimensions by half. Feature extraction continues with three more convolution layers consisting of one 32 and two 64 feature maps consecutively. The extracted features are then flattened and fed to a fully connected layer with 128 units. The last layer, softmax, consisting of 2 units (number of classes) outputs a vector of class probabilities. In the end, the maximum probability is used to map each input pixel to its corresponding class labels.

3. **Non-linearity:** A rectified linear unit (ReLU) activation function is used for each layer.

4. **Weight initialization:** The initial weights of keras layers are initialized using the *randomnormal* class from the built-in initializer, which generates tensors with a normal distribution with mean at 0.0 and standard deviation of 0.05.

5. **Batch Size:** The networks are trained on batches of 128 images each of size $33 \times 33 \times 5$.

6. **Epochs:** The networks were trained for 8 epochs, with data randomly shuffled between epochs.

2.5 Implementation

All data pre-processing was performed using our open-source SIproc, software[62] implemented in C++ and CUDA. Training and testing were performed in Python using open-source software packages. The CNNs were implemented in Python with the Keras library,[63] built on TensorFlow.[64] RF and SVM classifiers and accuracy scores were computed using the Scikit-learn library.[65] The CNN classifier's performance was calculated by testing the classifiers on ten different sets of randomly selected training pixels and averaging the overall accuracy run on an NVIDIA Tesla K40m GPU. The computation time for training the model with the aforementioned parameters is 10 minutes and for classification on the testing dataset is 30 minutes.

2.6 Cancer detection metrics

We define a pair of metrics that utilize tissue subtype classification results to aid early ovarian cancer detection. The "stromal ratio" (SR) is calculated by dividing the number of pixels classified as stroma by the total number of pixels in each core. Similarly, the "epithelial ratio" (ER) is the number of epithelial pixels divided by the total number of pixels in the core.

3 Results

We evaluated classifier performance using the overall accuracy (OA) and receiver operating characteristic (ROC) curves. OA is beneficial for binary and multi-class classification, representing the percentage of pixels mapped correctly to the appropriate class. The ROC curves delineate the correlation between specificity and sensitivity to ascertain acceptable false positive and true positive indicators.

We performed tissue segmentation using multiple machine learning algorithms, including those based on spectra alone, such as RFs and SVMs, and those that utilize both spatial and spectral features, such as CNNs. The class-wise and overall accuracy obtained after the classification of the testing dataset are summarized in Table 2. CNNs outperform RF and SVM classifiers in both class-wise and overall accuracy (by 40-50%). The low OA scores for RF (53.21%) and SVM (45.57%) can be attributed to using spectral information from only 5 wavenumbers instead of all the wavenumbers in the 900-1900 cm^{-1} range. Meanwhile, the high overall accuracy achieved by CNN (94.61%) is due to the utilization of both spectral and spatial features.

Table 2 Accuracy scores for epithelium and stroma classification using (a) SVM, (b) RF, and (c) CNNs averaged across 80 repetitions are presented below. CNNs utilize a combination of spatial and spectral features and outperform SVMs and RFs that employ spectral features alone.

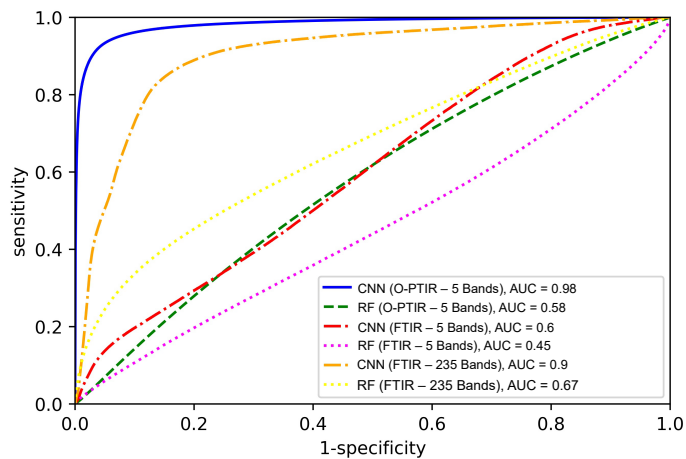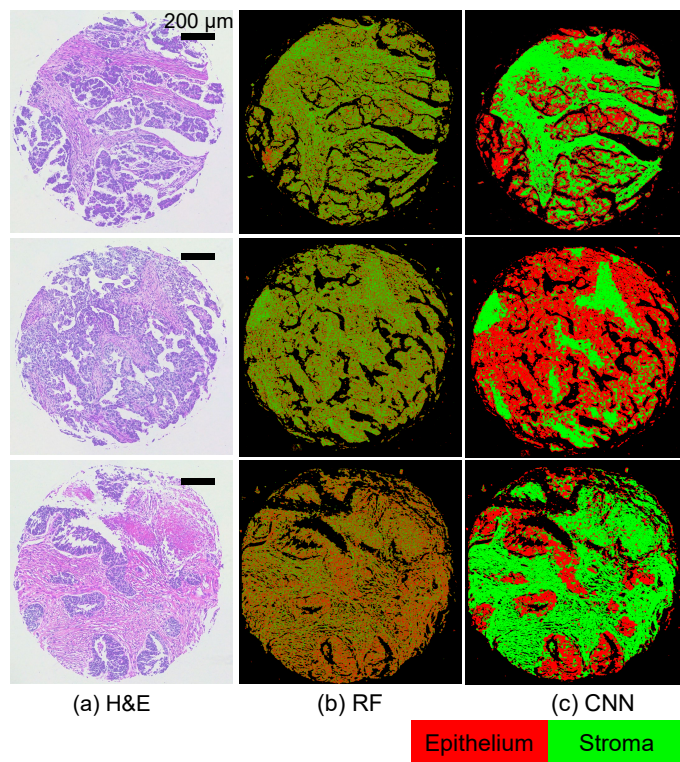| Class	SVM	RF	CNN
Epithelium	80.31 ± 0.18	60.18 ± 0.29	95.33 ± 1.52
Stroma	29.84 ± 0.26	44.27 ± 0.21	93.00 ± 1.97
Total	45.57 ± 0.3	53.21 ± 0.05	94.61 ± 0.82

Fig. 6 ROC curves and associated AUC values for binary classification of tissue type, separated by classifier type and datasets used. Due to the use of two-class models, each tissue class curve is a reflection of the curve from the other class, and thus the AUC values are equal across tissue class. CNN classifiers exhibit superior results to the RF classifiers, indicating that spatial information is essential in distinguishing tissue types.

Results that characterize the performance of all classifiers using the Area Under the Curve (AUC) in a Receiver Operating Characteristic (ROC) plot are presented in Figure 6. Our CNN classi-

fier on O-PTIR data from 5 bands outperforms all others with an AUC of 0.98. An RF classifier on the same O-PTIR data shows an AUC of 0.58. Since CNNs use spatial features and RFs don't, these results highlight the importance of combining spatial and spectroscopic information for improved tissue classification. RF classification on FTIR data with 5 bands provides a poor AUC of 0.45. This AUC increases to 0.67 by incorporating the 235 fingerprint bands. Classification of FTIR data from 5 bands using a CNN yields an AUC of 0.6, which increases to 0.9 when we include the 235 fingerprint bands. Comparing CNN performance for FTIR with fingerprint spectrum (AUC=0.9) and O-PTIR with 5 bands (AUC=0.98) implies that spatial details obtained from O-PTIR compensate for the loss in spectroscopic information due to the reduction in the number of bands.

Figure 7 presents results when the RF and CNN classification models are used to segment ovarian tissue cores, including regions outside annotated areas. O-PTIR data is consistent with H&E stained microscopy data, whereas RF results show poor concordance. These results show that our CNN classification results extend beyond annotated data indicating effective tissue segmentation into epithelium and stroma.

Fig. 8 Stromal ratio (SR) and Epithelial ratio (ER) are plotted as a function of the pathologist-assigned FIGO stage of ovarian cancer. These ratios are calculated by dividing the number of pixels classified as stroma or epithelium by the total number of pixels in each biopsy core. The trend line in plots (a) and (c) show a nonlinear fit for SR and ER as a function of the cancer stage. These trend lines display a substantial reduction in SR and an increase in ER from normal to grade II, but do not change appreciably from grades II to IV. There is a statistically significant ($P < 0.01$) decrease in SR and a significant ($P < 0.01$) increase in ER during the early stage of cancer from normal tissue until stage II as presented in (b) and (d). The graphs show the mean SR and ER vs. early cancer grade. The error bands correspond to one standard deviation (SD).

(a) H&E   (b) RF   (c) CNN

Epithelium   Stroma

Fig. 7 (a) H&E images of tissue cores are compared to O-PTIR classification results from (b) RFs and (c) CNNs. There is a good correspondence between O-PTIR class images in (c) and the corresponding H&E images in (a) indicating that our classification results generalize beyond annotated tissue regions. The correspondence between RF and H&E is poor as expected from the AUC values.

We can utilize classification results to facilitate early ovarian cancer detection. Figure 8 presents plots of stromal ratio (SR) and epithelial ratios (ER) as a function of Federation Internationale de Gynecolgie et d'Obstetrique (FIGO) cancer stage. SR and ER
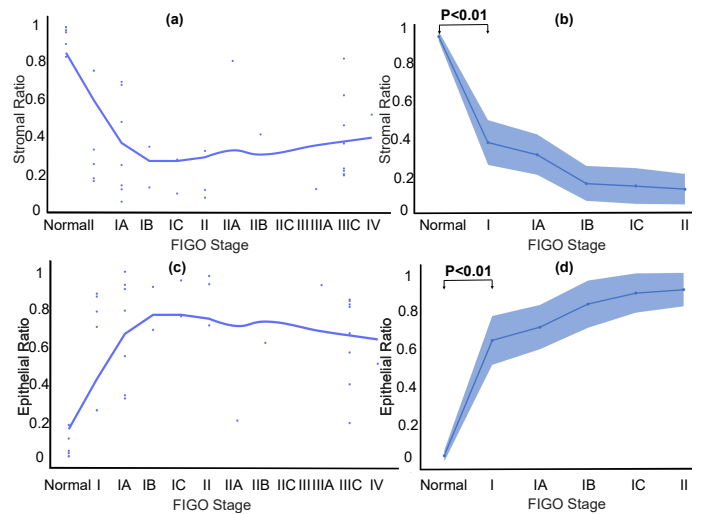
showed a nonlinear relationship with the FIGO stage, and the curve fitting function was calculated using nonlinear least-squares on the ratios. We observe a sharp reduction in SR with increasing cancer grade until grade II and a subsequent flattening of the curve. Figure 8(b) emphasizes the SR trend in early cancer stages: normal, grade I, IA, IB, IC, and II. There is a statistically significant ($P < 0.01$) reduction in SR from normal to grade I, illustrating its effectiveness as an early detection biomarker. A complimentary trend is observed in the ER metric in (c). A statistically significant ($P < 0.01$) increase in ER is observed with progressively worsening early stage cancer in (d). Pathologists have *qualitatively* observed changes in the relative quantities of stroma and epithelium in early cancer stages. Our technique *quantifies* the number of epithelial and stromal pixels, thereby enabling measurement of *quantitative* trends in the aforementioned metrics in the early stages of ovarian cancer which is consistent with the earlier studies made on tumor-stroma ratio in ovarian cancer.[8]

4 Discussion

FTIR imaging measures tissue absorbance across all mid-IR wavenumbers and constructs a hyperspectral data cube. The technology does not allow the measurement of individual wavenumbers. O-PTIR uses a tunable QCL to measure tissue absorbance at discrete wavenumbers. We can measure only a subset of mid-IR wavenumbers relevant to a specific application, thereby reducing data collection time. On the other hand, O-PTIR uses a pump-probe architecture to obtain super-resolution images. This im-

proved resolution results in an increase in the quantity of data and a corresponding increase in data collection time for each image. The effect of the improvement in resolution on tissue classification and the tradeoff between spatial and spectral resolution have not been studied until this paper. Our results demonstrate that we can maintain excellent tissue classification accuracy by reducing the number of bands and increasing the spatial resolution. This work presents a framework for making spatial-spectral tradeoffs in spectroscopic imaging while retaining good tissue segmentation accuracy.

Deep learning is used routinely in image classification.[66] However, it's application to hyperspectral data is limited.[58] Furthermore, it has never been applied to super-resolution hyperspectral data, and our paper is the first to demonstrate efficacy. Hyperspectral data being three-dimensional (3D) requires a large memory bandwidth. Super-resolution images have finer spatial details and require larger convolutional kernels to identify the same area as FTIR, increasing computational costs and making classification more challenging. We have optimized our novel deep-learning architecture to achieve an excellent tissue segmentation AUC of 0.98 despite these challenges.

Our results are obtained on 78 independent cancer patient cores and are statistically robust. Prior work often utilizes pixels within the same set of tissue cores during classification.[67,68] This can lead to misleading results since the machine learning algorithm may learn features that correspond to specific patient traits that are challenging to generalize beyond the current dataset. We perform training and validation on mutually exclusive patient cores, achieving robust, generalizable results that enhance scientific rigor and reproducibility.

Results in Figure 6 show a significant improvement in efficacy between RF (AUC=0.58) and CNNs (AUC=0.98) on the same O-PTIR data. Since CNNs not only utilize spectroscopic information, but also extracts spatial data, these results highlight the advantages of combining spatial and spectral features. This work builds on prior spatial-spectral FTIR classification work[51] and affirms the validity of this research approach.

The increased spatial resolution of O-PTIR leads to larger within-class spectral variation. Spectra in FTIR imaging are averaged over $\approx 5\mu m$ pixel, which is approximately the size of one cell. On the other hand, spectra in O-PTIR correspond to more localized $(0.5\mu m)$ sub-cellular features such as cell nucleus or Golgi apparatus, which have disparate biochemical constituents. This leads to a larger spectra variation in O-PTIR even within the same tissue class. A large within-class variance can be a potential disadvantage in tissue segmentation and analysis. However, our data analysis approach that combines spatial-spectral features turns this variation into an advantage.

The performance of machine learning algorithms depends critically on the quality and quantity of annotated data. Since annotations are performed using images of stained adjacent-sections that are several microns away from the MIRSI section, this imposes limitations on labeling accuracy. We mitigate annotation errors by limiting our labeling to unambiguous tissue areas and avoiding class boundaries. Furthermore, the alignment of images from adjacent sections is challenging,[69] and we obtained the best

fit through manual adjustment. The five wavenumbers that we chose for O-PTIR imaging offer good classification performance, but optimizing the set of wavenumbers could lead to improved performance. We will explore this optimization and the effects of improved spatial resolution on identifying other tissue subtypes in future work involving multi-class segmentation.

To our knowledge, this is the first large scale analysis of ovarian cancer tissue using mid-IR spectroscopic imaging. This analysis affords quantitative insights into ovarian cancer. Pathologists utilize the extent of epithelial infiltration into stroma and the relative proportion of stroma or epithelium to the rest of the tissue to subjectively assess cancer grade. Since our approach can precisely quantify the number of pixels of these subtypes, we can *quantify* these assessments and observe trends in a reliable manner. Furthermore, we analyze 78 cancer patients, enabling statistically robust analysis. Figure 8 (a) and (b) show that there is a statistically significant ($P < 0.01$) reduction in the stromal ratio (SR) between normal tissue (SR ≈ 0.9) and early stage (grade I - SR < 0.4) ovarian cancer. SR Furthermore, SR reduces from grade I to II and then shows no appreciable change from grade II to IV. The grades are obtained directly from Biomax. A complimentary trend is observed in the epithelium ratio (ER) in (c) and (d). ER for normal tissue is ≈ 0.1 and that for grade I cancer is > 0.6. These results illustrate the utility of quantitative tissue classification in cancer diagnosis. SR and ER are quantitative biomarkers for early stage ovarian cancer diagnosis and will be explored in greater detail in future work.

5 Conclusion

MIRSI is an emerging technology that has the ability to revolutionize digital histopathology. Significant progress has been made in overcoming the technological challenges impeding its clinical adoption. O-PTIR solves the spatial resolution challenge of prior FTIR imaging technology, enabling label-free sub-cellular tissue investigation. In this work, we present the first label-free, automated histological classification of ovarian tissue subtypes using MIRSI. We show that the improved spatial resolution allows us to make fewer spectral band measurements and still achieve reliable tissue segmentation with an AUC of 0.98. These results are enabled using a novel deep-learning architecture optimized for MIRSI data. The results are statistically robust with validation over 78 cancer patients and 60 million data points. We utilize tissue classification and propose new quantitative biomarkers for early ovarian cancer diagnosis. The combination of deep learning and quantitative biochemical measurements using MIRSI enables numerically precise evaluation of previously subjective assessments, improving the rigor and reproducibility of histopathology. O-PTIR also performs measurements in back-reflection geometry, making the instrument easy to use on diverse tissue samples and facilitates future clinical translation.

Author Contributions

CCG: Data curation, Visualization, Formal analysis, Investigation, Methodology, Software. **MB**: Data curation, Visualization, Investigation, Writing- Original draft preparation. **RM**: Methodology,

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

Notes and references

1 L. A. Torre, B. Trabert, C. E. DeSantis, K. D. Miller, G. Samimi, C. D. Runowicz, M. M. Gaudet, A. Jemal and R. L. Siegel, *CA: a cancer journal for clinicians*, 2018, **68**, 284.

2 S. Lheureux, C. Gourley, I. Vergote and A. Oza, *Lancet (London, England)*, 2019, **393**, 1240.

3 J.-M. Lee, L. Minasian and E. C. Kohn, *Cancer*, 2019, **125**, 4623.

4 A. du Bois, A. Reuss, E. Pujade-Lauraine, P. Harter, I. Ray-Coquard and J. Pfisterer, *Cancer*, 2009, **115**, 1234–1244.

5 S. Kehoe, J. Hook, M. Nankivell, G. C. Jayson, H. Kitchener, T. Lopes, D. Luesley, T. Perren, S. Bannoo, M. Mascarenhas *et al.*, *Lancet (London, England)*, 2015, **386**, 249–257.

6 I. Vergote, C. G. Tropé, F. Amant, G. B. Kristensen, T. Ehlen, N. Johnson, R. H. Verheijen, M. E. van der Burg, A. J. Lacave, P. B. Panici *et al.*, *The New England journal of medicine*, 2010, **363**, 943–953.

7 Z. Luo, Q. Wang, W. Lau, B. Lau, L. Xu, L. Zhao, H. Yang, M. Feng, Y. Xuan, Y. Yang *et al.*, *Cancer letters*, 2016, **377**, 174.

8 Y. Chen, L. Zhang, W. Liu and X. Liu, *BioMed research international*, 2015, **2015**,.

9 E. Banks, *Methods in Molecular Medicine*, 2001, **39**, 3–11.

10 A. P. Heintz, F. Odicino, P. Maisonneuve, U. Beller, J. L. Benedet, W. T. Creasman, H. Y. Ngan, M. Sideri and S. Pecorelli, *Journal of Epidemiology and Biostatistics*, 2001, **6**, 107–138.

11 A. Malpica, *International journal of gynecological pathology: official journal of the International Society of Gynecological Pathologists*, 2008, **27**, 175–181.

12 E. Hernandez, B. Bhagavan, T. Parmley and N. Rosenshein, *Gynecologic oncology*, 1984, **17**, 117–123.

13 A. Malpica, M. T. Deavers, K. Lu, D. C. Bodurka, E. N. Atkinson, D. M. Gershenson and E. G. Silva, *The American journal of surgical pathology*, 2004, **28**, 496–504.

14 R. Taylor, J. Zeller, R. Lieberman and D. O'Connor, *Gynecologic oncology*, 1999, **74**, 3–6.

15 F. Zeppernick and I. Meinhold-Heerlein, *Archives of gynecology and obstetrics*, 2014, **290**, 839–842.

16 A. Malpica, M. T. Deavers, C. Tornos, R. J. Kurman, R. Soslow, J. D. Seidman, M. F. Munsell, E. Gaertner, D. Frishberg and E. G. Silva, *The American journal of surgical pathology*, 2007, **31**, 1168–1174.

17 K. Bera, K. A. Schalper, D. L. Rimm, V. Velcheti and A. Madabhushi, *Nature reviews. Clinical oncology*, 2019, **16**, 703–715.

18 M. Wu, C. Yan, H. Liu and Q. Liu, *Bioscience Reports*, 2018, **38**,.

19 D. Komura and S. Ishikawa, *Computational and Structural Biotechnology Journal*, 2018, **16**, 34.

20 Y. Du, R. Zhang, A. Zargari, T. C. Thai, C. C. Gunderson, K. M. Moxley, H. Liu, B. Zheng and Y. Qiu, *Annals of biomedical engineering*, 2018, **46**, 1988–1999.

21 X. Chen, R. Zhang, K.-M. Fung, H. Liu, B. Zheng and Y. Qiu, Biophotonics and Immune Responses XV, 2020, p. 112410F.

22 J. Xu, C. Zhou, B. Lang and Q. Liu, *Deep Learning and Convolutional Neural Networks for Medical Image Computing*, Springer, 2017, pp. 73–95.

23 C. Fiore, D. Bailey, N. Conlon, X. Wu, N. Martin, M. Fiorentino, S. Finn, K. Fall, S.-O. Andersson, O. Andren *et al.*, *Journal of clinical pathology*, 2012, **65**, 496–502.

24 K. Pantel and R. H. Brakenhoff, *Nature reviews cancer*, 2004, **4**, 448–456.

25 K. Pantel and C. Alix-Panabières, *Nature Reviews Clinical Oncology*, 2019, **16**, 409–424.

26 D. L. Adams, R. K. Alpaugh, S. Tsai, C.-M. Tang and S. Stefansson, *Scientific reports*, 2016, **6**, 1–9.

27 D.-B. Asante, M. Morici, G. R. Mohan, E. Acheampong, I. Spencer, W. Lin, P. van Miert, S. Gibson, A. B. Beasley, M. Ziman *et al.*, *Cancers*, 2021, **13**, 6225.

28 I. ur Rehman, Z. Movasaghi and S. Rehman, *Vibrational spectroscopy for tissue analysis*, CRC press, 2012.

29 M. J. D'Souza and F. Koyoshi, *Pharmaceutical reviews*, 2009, **7**, http–www.

30 C. L. Morais, P. L. Martin-Hirsch and F. L. Martin, *Analyst*, 2019, **144**, 2312–2319.

31 M. Paraskevaidi, K. M. Ashton, H. F. Stringfellow, N. J. Wood, P. J. Keating, A. W. Rowbottom, P. L. Martin-Hirsch and F. L. Martin, *Talanta*, 2018, **189**, 281–288.

32 K. Maheedhar, R. A. Bhat, R. Malini, N. Prathima, P. Keerthi, P. Kushtagi and C. M. Krishna, *Photomedicine and Laser Surgery*, 2008, **26**, 83–90.

33 C. M. Krishna, G. Sockalingum, R. A. Bhat, L. Venteo, P. Kushtagi, M. Pluot and M. Manfait, *Analytical and bioanalytical chemistry*, 2007, **387**, 1649–1656.

34 K. R. Flower, I. Khalifa, P. Bassan, D. Démoulin, E. Jackson, N. P. Lockyer, A. T. McGown, P. Miles, L. Vaccari and P. Gardner, *Analyst*, 2011, **136**, 498–507.

35 K. Gajjar, J. Trevisan, G. Owens, P. J. Keating, N. J. Wood, H. F. Stringfellow, P. L. Martin-Hirsch and F. L. Martin, *Analyst*, 2013, **138**, 3917–3926.

36 G. Theophilou, K. M. Lima, P. L. Martin-Hirsch, H. F. Stringfellow and F. L. Martin, *Analyst*, 2016, **141**, 585–594.

37 K. M. Lima, K. B. Gajjar, P. L. Martin-Hirsch and F. L. Martin, *Biotechnology progress*, 2015, **31**, 832–839.

38 O. Klein, F. Kanter, H. Kulbe, P. Jank, C. Denkert, G. Nebrich, W. D. Schmitt, Z. Wu, C. A. Kunze, J. Sehouli *et al.*, *PROTEOMICS–Clinical Applications*, 2019, **13**, 1700181.

39 A. A. Zeitoune, J. S. Luna, K. Sanchez Salas, L. Erbes, C. L. Cesar, L. A. Andrade, H. F. Carvahlo, F. Bottcher-Luiz, V. H. Casco and J. Adur, *Cancer informatics*, 2017, **16**, 1176935117690162.

40 D. Pouli, E. M. Genega, T. B. Sullivan, K. M. Rieger-Christ, V. Wright, I. Georgakoudi and T. Schnelldorfer, *Biomedical Optics Express*, 2019, **10**, 4479.

41 K. Tilbury and P. J. Campagnola, *Perspectives in medicinal chemistry*, 2015, **7**, PMC–S13214.

42 M. J. Huttunen, A. Hassan, C. W. McCloskey, S. Fasih, J. Upham, B. C. Vanderhyden, R. W. Boyd and S. Murugkar, *Journal of biomedical optics*, 2018, **23**, 066002.

43 J. V. Jokerst, A. J. Cole, D. Van de Sompel and S. S. Gambhir, *ACS nano*, 2012, **6**, 10366–10377.

44 A. Oseledchyk, C. Andreou, M. A. Wall and M. F. Kircher, *ACS nano*, 2017, **11**, 1488–1497.

45 C. Petibois and G. Déléris, *TRENDS in Biotechnology*, 2006, **24**, 455–462.

46 A. Benard, C. Desmedt, M. Smolina, P. Szternfeld, M. Verdonck, G. Rouas, N. Kheddoumi, F. Rothé, D. Larsimont, C. Sotiriou *et al.*, *Analyst*, 2014, **139**, 1044–1056.

47 F. Großerueschkamp, A. Kallenbach-Thieltges, T. Behrens, T. Brüning, M. Altmayer, G. Stamatis, D. Theegarten and K. Gerwert, *Analyst*, 2015, **140**, 2114–2120.

48 M. J. Baker, E. Gazi, M. D. Brown, J. H. Shanks, N. W. Clarke and P. Gardner, *Journal of biophotonics*, 2009, **2**, 104–113.

49 C. Krafft, D. Codrich, G. Pelizzo and V. Sergo, *Journal of biophotonics*, 2008, **1**, 154–169.

50 M. J. Baker, J. Trevisan, P. Bassan, R. Bhargava, H. J. Butler, K. M. Dorling, P. R. Fielden, S. W. Fogarty, N. J. Fullwood, K. A. Heys *et al.*, *Nature protocols*, 2014, **9**, 1771.

51 F. N. Pounder, R. K. Reddy and R. Bhargava, *Faraday Discussions*, 2016, **187**, 43–68.

52 S. Pahlow, K. Weber, J. Popp, B. R. Wood, K. Kochan, A. Rüther, D. Perez-Guaita, P. Heraud, N. Stone, A. Dudgeon, B. Gardner, R. Reddy, D. Mayerich and R. Bhargava, *Applied spectroscopy*, 2018, **72**, 52–84.

53 D. Zhang, C. Li, C. Zhang, M. N. Slipchenko, G. Eakins and J.-X. Cheng, *Science advances*, 2016, **2**, e1600521.

54 R. Mankar, C. C. Gajjela, C. E. Bueso-Ramos, C. C. Yin, D. Mayerich and R. K. Reddy, *Applied Spectroscopy*, 2022, **76**, 508–518.

55 Z. Qin, S. Dai, C. C. Gajjela, C. Wang, V. G. Hadjiev, G. Yang, J. Li, X. Zhong, Z. Tang, Y. Yao *et al.*, *Chemistry of Materials*, 2020.

56 Y. Bai, D. Zhang, L. Lan, Y. Huang, K. Maize, A. Shakouri and J.-X. Cheng, *Science advances*, 2019, **5**, eaav7127.

57 W. Hu, Y. Huang, L. Wei, F. Zhang and H. Li, *Journal of Sensors*, 2015, **2015**,.

58 S. Berisha, M. Lotfollahi, J. Jahanipour, I. Gurcan, M. Walsh, R. Bhargava, H. Van Nguyen and D. Mayerich, *Analyst*, 2019, **144**, 1642–1653.

59 M. Lotfollahi, N. Tran, S. Berisha, C. Gajjela, Z. Han, D. Mayerich and R. Reddy, *arXiv preprint arXiv:2008.00566*, 2020.

60 V. Zohdi, D. R. Whelan, B. R. Wood, J. T. Pearson, K. R. Bambery and M. J. Black, *PLoS One*, 2015, **10**, e0116491.

61 D. P. Kingma and J. A. Ba, *arXiv preprint arXiv:1412.6980*, 2019, **434**,.

62 S. Berisha, S. Chang, S. Saki, D. Daeinejad, Z. He, R. Mankar and D. Mayerich, *Analyst*, 2017, **142**, 1350–1357.

63 F. Chollet *et al.*, *Keras*, https://keras.io, 2015.

64 M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu and X. Zheng, *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*, 2015, `http://tensorflow.org/`, Software available from tensorflow.org.

65 F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot and E. Duchesnay, *Journal of Machine Learning Research*, 2011, **12**, 2825–2830.

66 Y. Jiang, M. Yang, S. Wang, X. Li and Y. Sun, *Cancer communications*, 2020, **40**, 154–166.

67 S. Kumar, C. Desmedt, D. Larsimont, C. Sotiriou and E. Goormaghtigh, *Analyst*, 2013, **138**, 4058–4065.

68 R. Wang, A. Naidu and Y. Wang, *Diagnostics*, 2021, **11**, 2133.

69 R. Mankar, C. C. Gajjela, F. F. Shahraki, S. Prasad, D. Mayerich and R. Reddy, *Analyst*, 2021, **146**, 4822–4834.