


Data-driven development of sparse multi-spectral sensors for urological tissue differentiation

Felix Fischer^{1,*} , Karsten Frenner¹, Massimo Granai², Falko Fend², and Alois Herkommer¹

¹Institute of Applied Optics, University of Stuttgart, 70569 Stuttgart, Germany

²Department of Pathology, University Hospital Tübingen, Hoppe-Seyler Strasse 3, 72076 Tübingen, Germany

Received 30 January 2023 / Accepted 17 May 2023

Abstract. Infrared spectroscopy is often used to spot differences between benign and malignant tissue. Due to the proliferation of tumorous cells, the composition of tissue changes drastically. In the consequence shifts occur in its optical properties that are indicated by spectral biomarkers in the so-called fingerprint region. In this work, we propose a new concept for a sparsified multi-spectral measurement of the most important and informative biomarker signals. The results of a data-driven feature selection approach show that a reliable discrimination of the tissue is still possible, even though utilizing only a small fraction of the measured data. A selected arrangement of only a few narrow-band quantum cascade lasers could provide proficient signal-to-noise ratios and can noticeably reduce the data acquisition time. Consequentially, real-time applications will be possible in short-term and in-vivo diagnostics in the long-term. First measurements of silicone phantoms validate the imaging capability of the sensor concept.

Keywords: Biomarker detection, Feature selection, Infrared spectroscopy, Tissue differentiation.

1 Introduction

Most spectral information in analyses of biological structures in human tissue is found in the broad (mid-infrared) MIR region [1]. In order to achieve reasonable signal-to-noise ratios (SNR) throughout the complete spectrum, either the light source has to be utterly powerful, the detector more efficient or the spectrum averaged over multiple measurements [2]. Whereas the former yields an increased energy consumption and naturally more waste heat that has to be dealt with, the latter entails a longer acquisition time. Both of which is not ideal for preferential real-time and intra-operative tissue discrimination with spectral measurement technologies.

Thus, modern MIR spectrometers tend to use quantum cascade lasers (QCLs), that are more powerful light sources with a higher spectral irradiance than conventionally used thermal SiC sources [3–5]. When confined to narrower bands, e.g. in gas analyses of particular media, a real-time analysis of the spectrum is viable [6]. Since most discriminative information in tissue spectra is also found in these distinct peaks (i.e. biomarkers) [7], the logical conclusion is that a single targeted approach to using only the relevant areas yields a drastic decline of the spectral acquisition time. Therefore, we want to concentrate on narrowbanded

distributed feedback (DFB)-QCLs with center wavelengths that provide most discriminative power. A more recent approach in prostate cancer diagnostics is encouraging a similar approach utilizing breath analyses for early stage detection [8].

For a first estimate of the discriminability of human tissue with an extremely sparsified infrared (IR) spectrum, we measured urothelia and urothelial carcinoma of bladder entities. By utilizing several different data-driven feature selection (FS) methods [9–13], we show that only a small fraction of the original MIR spectrum is necessary to reliably separate benign from malignant urothelia.

According to that, we introduce an optical design for an attenuated total reflectance (ATR) spectrometer that enables the spatially resolved real-time discrimination of bladder tissue and show first measurements with a monochromatic light source. Simulations validate these measurements and show limits in the spatial resolution.

2 MIR spectroscopy of urological specimen

In this work, we have used the commercially available Fourier-transform infrared spectroscopy (FT-IR) spectrometer Spectrum Two[®] by PerkinElmer in combination with an ATR accessory that incorporates a diamond crystal. A spectral range from 4000 to 450 cm⁻¹ is available with this

* Corresponding author: felix.fischer@ito.uni-stuttgart.de

setup. The spectrometer additionally provides a plunger with an adjustable pressure p .

An overview of the constituents of intact bladder walls is given in Figure 1a. Since more than 90% of bladder cancer occurs in the transitional cells (urothelium [14]) and ATR spectroscopy is well suited for shallow measurements, we focus on these tightly packed innermost layers. The urothelium usually consists of 3 to 6 cell layers with cell sizes of 10–50 μm [15]. Urothelial cancer can further be categorized into papillary bladder tumor (pTa), non-muscle invasive tumor (carcinoma in situ [CIS] and pT1), and muscle invasive tumors (pT2–pT4 [16]). The investigated tumors are staged between pT0 and pT4a.

All spectra were obtained from eleven different patients, that have been diagnosed with bladder cancer. Every sample has been prepared and labeled by professional pathologists and physicians from the urology department of the University of Tübingen. Immediately after radical cystectomy, the bladder was transported to the laboratory. There, it has been opened completely and the urothelial surfaces of interest were cut into approx. 1 cm^2 areas in order to fit into the spectrometer. For each patient, we have extracted two clearly separable excerpts of healthy urothelium and (if available) urothelial carcinoma. A detailed overview on the tumor staging is given in Table 1. Examples of the measured samples can be seen in Figure 1b. Measurements have been conducted on the inner bladder wall with varying pressure applied. The mean spectra of healthy and tumorous specimen are shown in Figure 1c. Most of the variances occur in the fingerprint region (1800–900 cm^{-1} , [7]).

2.1 Benefits of ATR spectroscopy

The key advantage of ATR spectroscopy compared to transmission spectroscopy is the obsolete sample preparation. Whereas transmission spectroscopy essentially depends on the sample thickness, ATR measurements always only measure the first few microns in depth of the surface. Figure 2 shows the differences between these two types of spectroscopy.

According to Lambert's law, the absorbance:

$$A = \mu d \cdot \log(e) = \log \frac{I_0}{I}, \quad (1)$$

is proportional to the sample thickness d and the attenuation coefficient μ . It quantifies the attenuation of the intensity I after the incident light intensity I_0 has transmitted through a uniformly distributed medium and is complementary to the transmittance $T = 10^{-A}$. For well-defined transmission measurements, the samples are usually powdered, mixed with non-absorbing powders such as KBr and pressed into pellets of specific thicknesses. The attenuation coefficient:

$$\mu = 4\pi\tilde{\nu}\kappa, \quad (2)$$

depends on scattering effects and absorption by particles or molecules when interacting with absorbing samples, as seen in Figure 2a. Both effects are summarized in the absorption index κ and also depend on the wavenumber

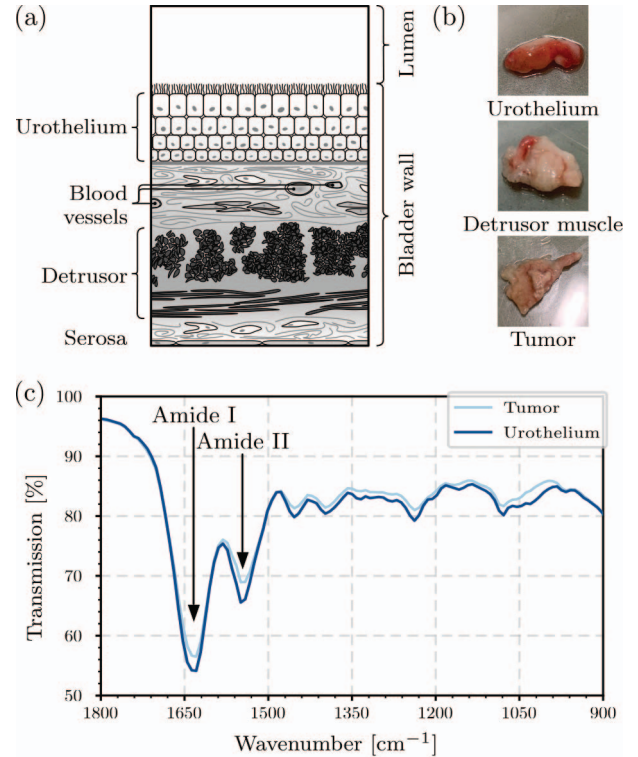


Fig. 1. Overview of measured samples: cross-section of a bladder wall according to [17, 18] (a), freshly resected specimen (b) and their mean spectra in the fingerprint region (c).

$\tilde{\nu} = 1/\lambda$ and thus on the wavelength λ of the incident light.

Another factor that decreases the signal intensity are reflections at each surface. However, spectral application usually use reference measurements without the sample for the calculation of the signal. Therefore, only the attenuation coefficient is recorded and the reflection can be neglected.

In contrast to the transmission spectroscopy, the absorbance in an ATR setup does not depend on the absolute thickness of the sample. As the evanescent wave at the boundary surface of an internal reflective element (IRE) can only penetrate the sample up to a certain depth,

$$d_p = \frac{\lambda}{2\pi\sqrt{n_1^2 \sin^2 \theta - n_2^2}}, \quad (3)$$

preparations of the samples are not required given that $d > d_p$. The refractive indices of IRE and sample are denoted by n_1 and n_2 respectively and θ is the angle of incidence. The depth of penetration is defined as the depth where the strength of the evanescent wave decayed to approx. 37%. Typically, this depth lies in single-digit microns when analyzing biological specimen [19, 20]. Figure 2b shows a setup with a collimated illumination for macroscopic ATR imaging [21]. Single-point spectrometers, such as the Spectrum Two[®], focus the incident light onto the boundary surface using parabolic mirrors to enhance the SNR.

Table 1. Tumor staging of the aquired spectral data using FT-IR spectroscopy. Number of patients and age information is differentiated by gender (female/male). The average age of all 11 patients is 67.0 ± 11.1 .

Subset	Patients	Average age	pT0	pTis	pT2a	pT2b	pT3a	pT3b	pT4a	Sum
Urothelium	5/6	$70.6 \pm 12.0/64.0 \pm 10.3$	160	91	70	46	22	120	11	520
Tumor	3/2	$77.7 \pm 4.5/68.0 \pm 15.6$	0	0	57	45	68	117	34	321
Total	5/6	$70.6 \pm 12.0/64.0 \pm 10.3$	160	91	127	91	90	237	45	841

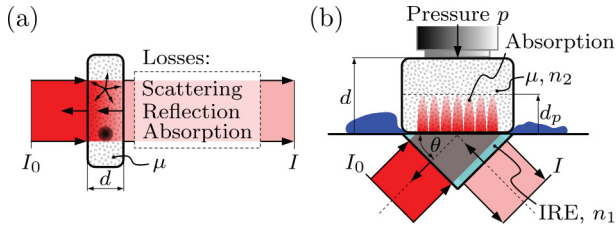


Fig. 2. Differences between transmission (a) and ATR spectroscopy (b).

According to [22], the depth of penetration cannot be inserted in equation (1) directly, since there is no consideration of the attenuation by the sample itself. Eventhough the effective thickness,

$$d_{\text{eff}} = \frac{n_1}{n_2 \cdot \cos \theta} \cdot |t_{12}|^2 \cdot \frac{d_p}{2}, \quad (4)$$

linearly depends on d_p , it also takes the complex transmission coefficient t_{12} for supercritical reflectance into account. This depth is different for parallel and perpendicular polarizations. Since the penetration depth is rather small, one problem is the influence of other materials in between sample and prism.

2.2 Influences of the applied pressure on the tissue spectra

As already mentioned, we are also able to apply a defined pressure on top of the sample using a plunger. The maximum possible pressure is $p_{\text{max}} = 5.3$ MPa utilizing a plunger with a diameter of 6 mm. This is needed to restrain as much water from the sample as possible in order to reduce the interfering influences due to its high absorption coefficient in the MIR region. Figure 1b clearly shows films of water transpiring from the samples onto the sample area.

The maximum pressure corresponds to an applied force of 150 N. In the dataset, each sample has been measured using forces in the range between 0 and 150 N with an increment of 10 per measurement.

The increasing pressing force has direct influence on the signal quality of the transmission spectrum, see Figure 3. When measuring without any applied pressure, the spectrum is clearly influenced by the remaining water within cells and bound to molecules. For comparative reasons, we also recorded a spectrum of pure water (dashed line). Since water has a nearly constant level of absorption between 1450 and 1000 cm^{-1} , the influence can be considered as a downscaling dampening effect on the tumor

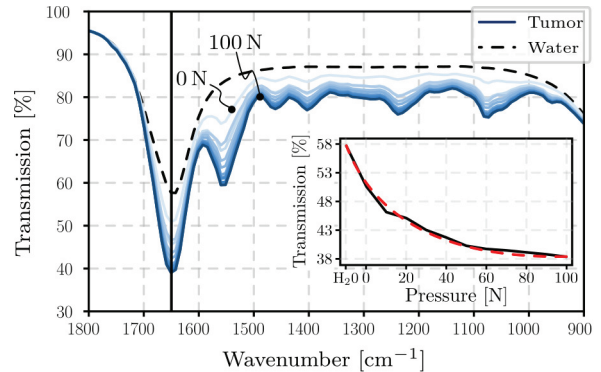


Fig. 3. Spectra of an urothelial carcinoma under increasing pressing force. Lighter colors indicate lower pressure. The transmission is gradually reduced with increasing pressure (see peak transmission of amide I band displayed in the inset). The red line indicates a polynomial fit.

spectrum. Looking at the changes of peak intensities within each increment, an exponential approximation towards the signal of 100 N is noticeable.

As the signal is heavily influenced by the pressing force, we also have to investigate the discriminability of samples with varying water content in the next section.

3 Feature selection to optimize spectral scans

The subsequent selection of relevant wavelengths for the spectral discrimination between healthy and diseased urological specimen is based on the measurements described in the previous section. This is the crucial part in engineering these types of sparse spectrometers.

In an earlier study, we have already shown that a differentiation of urothelia and detrusor muscles is possible with an extremely sparsified signal consisting of only two wavenumbers [23]. Since there are naturally different properties between epithelial tissue and deeper muscular tissue, which is responsible for the bladder contraction, we want to investigate a more difficult discrimination task in this study.

In order to select the q most relevant wavenumbers for an optimal differentiation, we have chosen different data-driven FS methods. When speaking of features in this context, wavenumbers are meant. The results of classical multivariate statistical approaches using principal component analysis (PCA) [9, 11, 12] and linear discriminant analysis (LDA) [13] are compared to those of a recently

developed FS algorithm called FeaSel-Net [10], that is based on recursive elimination of the input signal in neural networks. Whereas both former analyses are based on linear transformations and covariance evaluations, the latter is inherently non-linear and promises more complex solutions for the selection.

In total, the acquired dataset \mathcal{D} consists of $n_C = 520$ control spectra (healthy urothelium, C) and $n_T = 321$ spectra of tumorous tissue (T). One single measurement with a spectral resolution of $\Delta\tilde{\nu} = 8 \text{ cm}^{-1}$ consists of $p = 113$ features in the fingerprint region (the region of interest). Each spectrum is normalized dividing it by its respective maximum, such that it becomes 1.0. The difference in the numbers of spectra is due to the pre-treatment of some bladders, where macroscopically differentiable tumor has already been resected transurethrally and only control tissue could be secured in a follow-up surgery. For validation purposes, this dataset is split into two datasets using a train-test-split ratio of 0.2. A more detailed overview on the sets is given in Table 2. For a fair comparison of the FS methods, the subsets are perpetually allocated.

3.1 Feature selection with PCA and LDA

As a first estimate on the separability of both classes C and T, the multivariate analyses PCA and LDA are used.

The linear transformations underlying these methods aim for a lower-dimensional representation of the original spectral dataset \mathbf{X} . This dataset is standardized along the feature axis, making their means $\mu_f = 0$ and standard deviations $\sigma_f = 1$. Multiplying the so-called loadings matrix \mathbf{A} with the input yields the transformed dataset,

$$\mathbf{Y} = \mathbf{A}^T \mathbf{X}. \quad (5)$$

We obtain this loadings matrix by solving the eigenvalue problem for the covariance matrix or scatter matrices respectively. What differs both types of transformations is that the LDA is not only using the covariances of the features but also profits from the pre-labeled class information that is encoded in the scatter matrices. However, both analyses try to explain as much data as possible while using as little latent variables q as possible (i.e. $p \gg q$). Since these variables are usually hard to interpret and even harder to measure directly, the weights of the transformation (i.e. the loadings) are used as a measure for the feature importance,

$$\mathcal{I}_f = \sum_{i=1}^q |\mathbf{a}_{fi}|, \quad (6)$$

where the absolutes of all weights belonging to a particular feature f are summed up.

Figure 4 qualitatively shows score plots of the transformed data applying PCA and LDA on the measurements of the two different tissue types. Each point in the scatter plot corresponds to a sample spectrum and is colored according to their class (C: dark and T: light blue). We used the training set \mathcal{D}_t , that contains all aforementioned pressure variations, to calculate the weights for the transformations and plotted this set as well as the validation \mathcal{D}_v into the score plots. Afterwards, equation (6) is applied to each

Table 2. Overview on the data used for the FS.

Dataset	Urothelium	Tumor	Sum
Training	413	259	672
Validation	107	62	169
Complete	520	321	841

method resulting in the three wavenumbers that are highlighted by red rectangles. Their values can also be found in Table 3 together with the findings from Section 3.2. Subsequently, we apply a second transformation using only the information that is given by these three wavenumbers. The resulting score plot can be retrieved from the right-hand side. Since the first two components describe at least more than 96% in each method, the focus lies on a $q = 2$ dimensional representation and all higher components can be neglected.

Unfortunately, the results of the PCA do not look promising for a good discriminability. The confidence intervals of both classes overlap almost entirely and their means are located at almost the exact same spot. However, the high explained variance of the first principal component (PC) with 95% show that the signals are highly correlated. This correlation is pronounced, since the differences in spectra measured with low applied pressures are always superimposed by the water signal. Reducing the input information to $p = 3$ wavenumber results in a small benefit in terms of the overlap of means and confidence intervals, but the classes are still not distinguishable.

As expected, the latent space of the LDA is noticeably more structured. By drawing a line parallel to the second LD, we are able to correctly classify the samples using the correct value of LD_1 as a threshold (dotted line). Except from one measurement of the urothelium class, all samples can be clearly distinguished from another. Even though the correlation between both mean spectra in Figure 1c is nearly $\rho = 100\%$, it is still possible to transform the data into distinct clusters using LDA.

3.2 Feature selection with neural networks

Another evaluation of the wavenumbers is done using FeaSel-Net [10]. This is a non-linear recursive FS method, that has recently been developed and already been used in several analyses of Raman and IR spectra [9, 10, 23]. It can be embedded in any 1D neural network.

During the training process of a classifier model, the feature importance,

$$\mathcal{I}_f(\mathcal{D}_{t,f}) = -\frac{1}{n_t} \sum_{i=0}^{n_t} \sum_{\mathbf{x} \in \mathcal{D}_{t,f}} \hat{\mathbf{y}}(\mathbf{x}) \cdot \log \mathbf{y}(\mathbf{x}), \quad (7)$$

is evaluated as soon as the model performs well and the least informative features are pruned before the training continues. The evaluation is different to linear transformations in equation (6). Here, the training dataset \mathcal{D}_t with $p = 113$ features and $n_t = 672$ samples is masked at each feature f . Subsequently, the impact on the classification performance is analyzed and features are discarded,

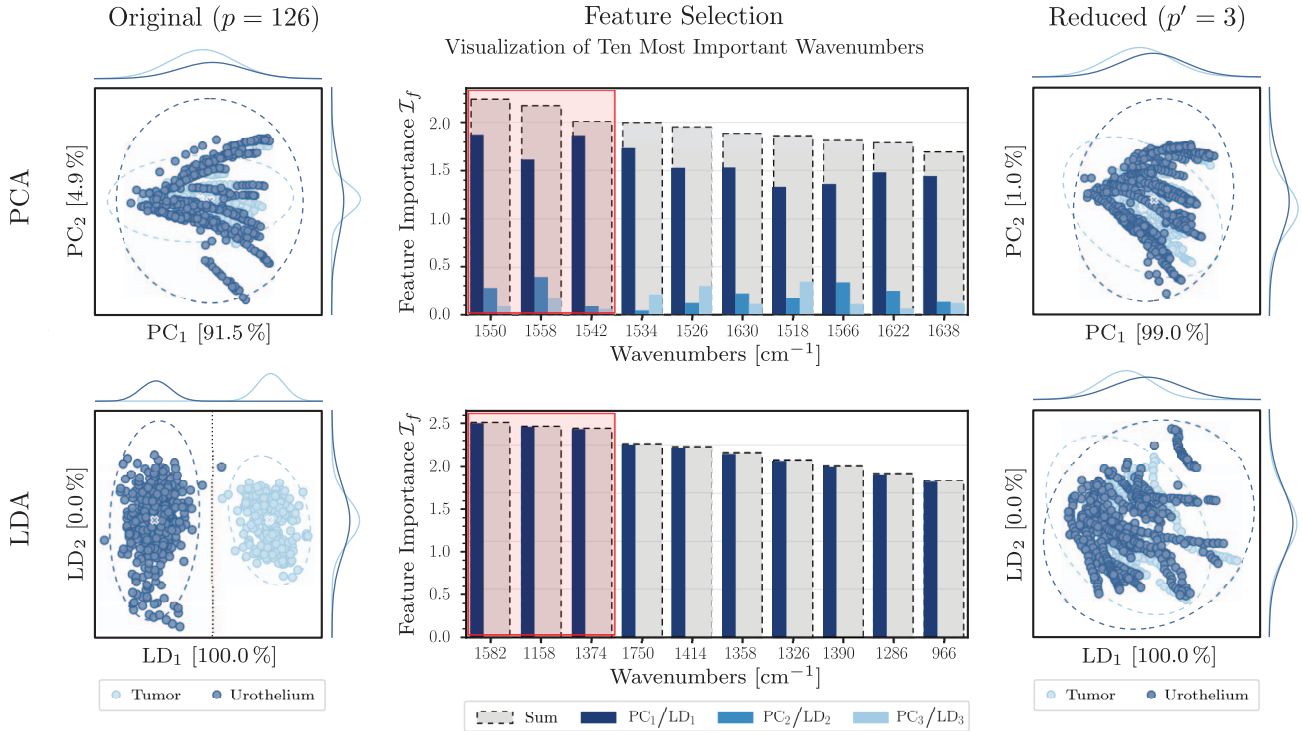


Fig. 4. Comparison of the feature selection with PCA and LDA. The left side shows score plots of the complete spectra, where each point represents one measurement. Confidence intervals with 3σ are depicted as dashed ellipses and projected onto each component. The score plots on the right are calculated using only the $p' = 3$ most important features (red rectangle). Note that the discriminative power of the LDA collapses.

Table 3. Wavenumbers of relevant biomarkers [cm^{-1}] for different FS methods.

FS method	\hat{v}_1	\hat{v}_2	\hat{v}_3
PCA	1542	1550	1558
LDA	1158	1374	1582
FeaSel (best run)	1022	1479	1655
FeaSel (Jaccard)	1014	1022	1495

if the impact of masking is negligible. Therefore, we calculate the cross-entropy (CE) for each prediction \mathbf{y} and the target vector $\hat{\mathbf{y}}$ for each subset $\mathcal{D}_{t,f}$ and average over all samples. Low results indicate a smaller impact of the mask and thus unnecessary features and vice versa.

The binary classifier model is optimized using the Adam [24] with a learning rate $\eta = 5 \times 10^{-4}$ and the applied loss function again is the CE. It consists of three fully-connected hidden layers with a decreasing number of nodes per layer (113–113–29–8–2). Nodes are activated using the non-linear ReLU function.

This classifier is trained until the accuracy threshold $\tau = 95\%$ is surpassed for $\Delta e = 20$ training epochs in a row. We also implemented a decay of this threshold with 5×10^{-4} per epoch that starts as soon as the algorithm is theoretically allowed to prune again. When these criteria are met, the training pauses. At training halt, the feature selection process is executed with a pruning rate of

$\pi = 0.2$ for every recursive pruning step. A convergence is obtained as soon as $p' = 3$ distilled wavenumbers are remaining. Inherent randomness in neural networks makes it indispensable to statistically evaluate multiple selections. Hence, we apply FeaSel-Net 25 times in a row and create a weighted Jaccard matrix $\mathbf{J}_w(x, y)$ that shows the relationship amongst chosen wavenumbers, see Figure 5. The diagonal of this matrix corresponds to a histogram of the selected data and show that wavenumbers 1495 cm^{-1} , 1014 cm^{-1} , 1022 cm^{-1} are all chosen seven times, which is approximately every fourth run and more than ten times as often as random guessing. The wavenumbers that are chosen most often are given in Table 3.

FeaSel-Net converged in every application with the best run's training history depicted in Figure 6. This run resulted in an accuracy with approx. 95% for the validation and training data set, with the validation accuracy being slightly higher. Since unnecessary wavenumbers are discarded during the training, an overfitting effect is less likely to be observed.

In the following, we make use of the wavenumbers retrieved from the best run. To compare the linear and non-linear methods, we need to apply the same reduction of input signal for the training of another neural network and train the model anew. Whereas the discriminative power of the LDA gets lost using only $p' = 3$ and has never been available in PCA, the neural network approach seems to achieve significantly better results. Figure 7 shows the prediction of the complete dataset by an exemplary trained

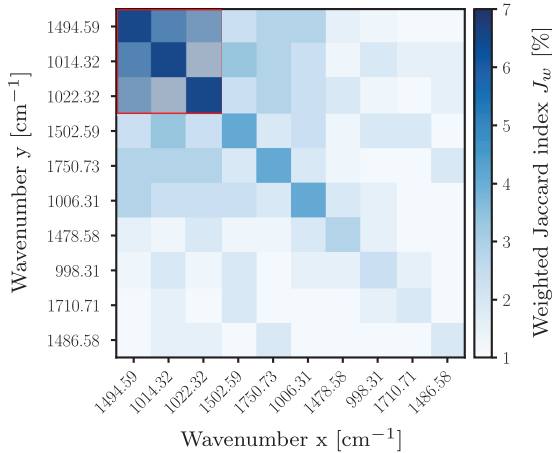


Fig. 5. Weighted Jaccard matrix for 25 FeaSel-Net applications. Features that are often chosen together appear darker than less chosen ones. The chosen features are indicated by the red square.

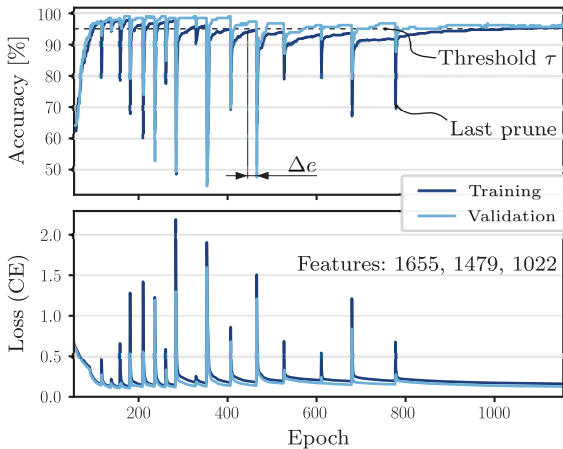


Fig. 6. Training history of one FS run with FeaSel-Net. The typical saw tooth pattern can be observed.

model. The class probability is shown on the left side. The average probabilities are depicted with a dashed line in their respective colors (C: dark and T: light blue). A shaded background indicates the standard deviations. The means of both classes are clearly separated and together with the confusion matrix on the right, the results can be interpreted as successful predictions.

Whilst the first model has trained using approx. 16.5k parameters, the second model only needs 368. Using only two percent of the original number of parameters is an enormous saving of computational costs. Despite all these reductions in complexity the training accuracy is satisfyingly good with $92.7 \pm 2.9\%$ and the validation accuracy is even better with $95.4 \pm 2.9\%$ on average over ten optimizations with 250 training epochs. The losses amount to 0.21 ± 0.04 for the training and 0.16 ± 0.05 for the validation.

An analysis of the impact on the classification using smaller pressures has additionally been conducted.

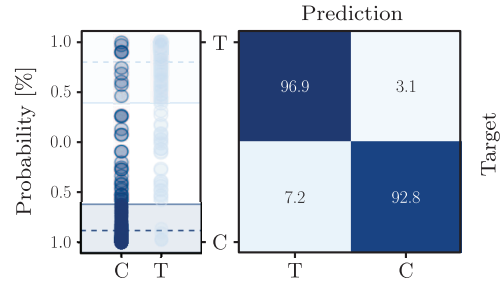


Fig. 7. Discriminability of neural networks restricted to an ultra-sparse spectral input ($p' = 3$).

When using small forces (0–10 N), an overall classification accuracy (OCA) of 81% is achieved. This instantly gets better when using forces greater or equal to 10 N where it is always more than 85% OCA.

4 Multi-spectral sensor concept

With all the savings of computational costs and the findings that a complete MIR spectrum is not necessary for tissue discrimination, we are allowed to confine ourselves on the spectral information from ultra-sparse illumination. Therefore, we propose a multi-spectral approach for the measurement of tissue samples using a spatially resolved ATR measurement system.

As depicted in [Figure 8](#), it is possible to use an echellette grating for example to combine all DFB-QCL onto the optical axis. The QCL are ideally chosen such that they correspond to our previous findings. With well-timed shutters, the sample is illuminated sequentially. In this configuration, the lasers are imaged on the focal plane array (FPA) bolometer as a reference at first. As soon as there is close contact with the tissue, a fraction of the laser light is absorbed depending on the tissue's absorption coefficient at the interface according to equations (1)–(4). The angle of incidence is 45° .

A magnification of the IRE surface is made possible using an $4f$ beam expander. The macroscopic imaging technique from [Figure 2](#) is extended and magnified with the imaging scale,

$$\beta' = \frac{f_{\text{asphere}}}{f_{\text{PM}}}, \quad (8)$$

to obtain a mesoscopic image. For the reduction of chromatic aberrations, a parabolic mirror (PM) with $f_{\text{PM}} = 15$ mm has been selected as the first optical element and an aspheric lens (Thorlabs: AL72550-E3) with $f_{\text{asphere}} = 50$ mm as the second resulting in a magnification of $\beta' = 3.3\times$. The lens as well as the right-angled prism are made of ZnSe, since it provides transmittance in the complete MIR region. The bolometer used for the light detection is a Lepton 3.5 by FLIR with 160×120 pixels, a framerate of 8 Hz and a pixel size of $=12 \mu\text{m}$. Within this work, we focus on using only one continuous wave laser for the illumination. The DFB-QCL light source has a

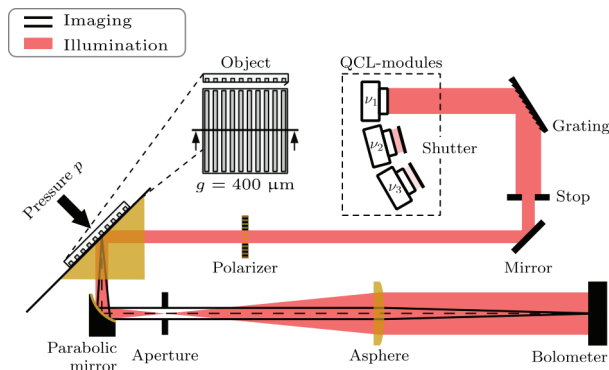


Fig. 8. Schematic drawing of the proposed multi-spectral sensor for tissue discrimination.

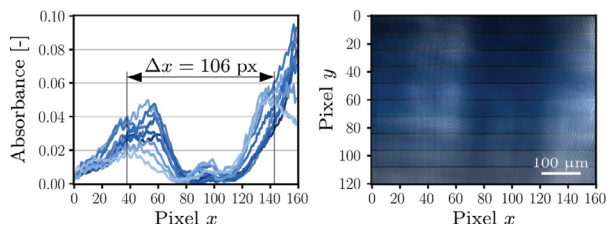


Fig. 9. Measurements with the proposed multi-spectral sensor. The absorbance plot (left) shows the absorbance averaged over 12 px each. The colors indicate the area of averaging.

center wavelength of $\lambda = 9286$ nm and a maximum optical power of 79 mW.

In an experiment with silicone phantoms we validate the image magnification and want to assess the imaging quality, see Figure 9. The simulated image-sided airy disk has a diameter of $102.4 \mu\text{m}$ using an aperture of the size $D = 2.1$ mm. In order to have a pattern that can easily be analyzed and validated, we have chosen a grating structure with a grating constant of $g = 400 \mu\text{m}$. We measured the peak-to-peak distance in the absorption plot with $\Delta x = 106$ px, which corresponds to a measured grating constant of $g' = 385 \mu\text{m}$. The difference of approximately 4% could be explained by deformations due to an applied pressure during the measurement or by fabrication errors. Since the fabrication has been done using a 3D-printed mold, there could be additional errors due to the printing process parameters such as lateral motion resolution or nozzle size.

However, the results clearly show, that it is possible to image the absorption of silicone at an IRE's boundary surface using a mesoscopic magnification setup. The slits with absent material are visible and the simulated magnification could be validated.

5 Conclusion

In this work, we have shown that it is possible to distinguish tumorous from healthy tissue using sparse mid-infrared spectroscopy in combination with IRE crystals and ultra-sparse illumination. The selection of the most important

wavelengths for the engineering of the light source is done using different linear and non-linear feature selection algorithms.

Linear transformations (PCA and LDA) are not able to cope with the difficulties induced by the superimposing effects of water that tend to equalize the net spectra of both classes. Especially in low pressure applications of ATR spectroscopy, where the water content within the tissue is at high level, the results are sobering. Inherently non-linear transforming neural networks are a better alternative to discriminate the tissues and find suitable illuminating wavelengths. The influences of superimposition in the spectra due to water signals on the discriminative power is only a problem in measurements when absolutely no pressure is applied.

Furthermore, we have proposed a concept that enables real-time spatially resolved measurements, that could potentially even be used for in-vivo applications when integrated in endoscopes for example. One of the biggest challenges will be an economical production of such endoscopes and the miniaturization of IR optics.

Availability of data and material

Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

Review board statement

The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the local Ethics Committee (UKT IRB# 804/2020/B02).

Competing interests

The authors declare that they have no competing interests.

Acknowledgments. The work described in this paper was conducted in the framework of the Graduate School 2543/1 "Intraoperative Multisensory Tissue Differentiation in Oncology" (project A2) funded by the German Research Foundation (DFG – Deutsche Forschungsgemeinschaft).

We would like to thank the urological department of the University Hospital Tübingen for providing us with clinical samples. Special thanks are addressed to Dr. med. Simon Walz and Jörg Hennenlotter.

References

- 1 Kalasinsky V.F. (1996) Biomedical applications of infrared and Raman microscopy, *Appl. Spectrosc. Rev.* **31**, 3, 193–249.
- 2 Ewing G.W. (1971) Signal-to-noise enhancement in infrared absorption spectrophotometry, *Rev. Sci. Instrum* **42**, 1, 169–170.
- 3 Schwaighofer A., Brandstetter M., Lendl B. (2017) Quantum cascade lasers (QCLs) in biomedical spectroscopy, *Chem. Soc. Rev.* **46**, 19, 5903–5924.

- 4 Klocke J.L., Mangold M., Allmendinger P., Hugi A., Geiser M., Jouy P., Faist J., Kottke T. (2018) Single-shot sub-microsecond mid-infrared spectroscopy on protein reactions with quantum cascade laser frequency combs, *Anal. Chem.* **90**, 17, 10494–10500.
- 5 Weida M.J., Yee B. (2011) Quantum cascade laser-based replacement for FTIR microscopy, *Proc. SPIE* **7902**, 280–286.
- 6 Ghorbani R., Schmidt F.M. (2017) Real-time breath gas analysis of CO and CO₂ using an EC-QCL, *Appl. Phys. B* **123**, 5, 144.
- 7 Ellis D.I., Dunn W.B., Griffin J.L., Allwood J.W., Goodacre R. (2007) Metabolic fingerprinting as a diagnostic tool, *Pharmacogenomics J.* **8**, 9, 1243–1266.
- 8 Maiti K.S., Fill E., Strittmatter F., Volz Y., Sroka R., Apolonski A. (2021) Towards reliable diagnostics of prostate cancer via breath, *Sci. Rep.* **11**, 1, 18381.
- 9 Becker L., Fischer F., Fleck J.L., Harland N., Herkommer A., Stenzl A., Aicher W.K., Schenke-Layland K., Marzi J. (2022) Data-driven identification of biomarkers for in situ monitoring of drug treatment in bladder cancer organoids, *Int. J. Mol. Sci.* **23**, 13, 6956.
- 10 Fischer F., Birk A., Somers P., Frenner K., Tarín C., Herkommer A. (2022) FeaSel-Net: A recursive feature selection callback in neural networks, *Mach. Learn. Knowl. Extr.* **4**, 4, 968–993.
- 11 Han H. (2010) Nonnegative principal component analysis for mass spectral serum profiles and biomarker discovery, *BMC Bioinform.* **11**, Suppl 1, S1.
- 12 Lovergne L., Ghosh D., Schuck R., Polyzos A.A., Chen A.D., Martin M.C., Barnard E.S., Brown J.B., McMurray C.T. (2021) An infrared spectral biomarker accurately predicts neurodegenerative disease class in the absence of overt symptoms, *Sci. Rep.* **11**, 1, 15598.
- 13 Ollesch J., Drees S.L., Heise H.M., Behrens T., Bruning T., Gerwert K. (2013) FTIR spectroscopy of biofluids revisited: An automated approach to spectral biomarker identification, *Analyst* **138**, 14, 4092–4102.
- 14 Kaufman D.S., Shipley W.U., Feldman A.S. (2009) Bladder cancer, *Lancet* **374**, 9685, 239–249.
- 15 Jost S.P., Gosling J.A., Dixon J.S. (1989) The morphology of normal human bladder urothelium, *J. Anat.* **167**, 103–115.
- 16 John B.A., Said N. (2017) Insights from animal models of bladder cancer: Recent advances, challenges, and opportunities, *Oncotarget* **8**, 34, 57766–57781.
- 17 Wang S., Jin S., Shu Q., Wu S. (2021) Strategies to get drugs across bladder penetrating barriers for improving bladder cancer therapy, *Pharmaceutics* **13**, 2, 166.
- 18 Zacche M., Srikrishna S., Cardozo L. (2015) Novel targeted bladder drug-delivery systems: A review, *Res. Rep. Urol.* **7**, 169.
- 19 Andrew Chan K.L., Kazarian S.G. (2016) Attenuated total reflection Fourier-transform infrared (ATR-FTIR) imaging of tissues and live cells, *Chem. Soc. Rev.* **45**, 7, 1850–1864.
- 20 Vongsvivut J., Pérez-Guaita D., Wood B.R., Heraud P., Khambatta K., Hartnell D., Hackett M.J., Tobin M.J. (2019) Synchrotron macro ATR-FTIR microspectroscopy for high-resolution chemical mapping of single cells, *Analyst* **144**, 10, 3226–3238.
- 21 Kazarian S.G., Chan K.L.A. (2010) Micro- and macro-attenuated total reflection Fourier transform infrared spectroscopic imaging. Plenary Lecture at the 5th International Conference on Advanced Vibrational Spectroscopy, 2009, Melbourne, Australia, *Appl. Spectrosc.* **64**, 5, 135A–152A.
- 22 Milosevic M. (2012) *Internal reflection and ATR spectroscopy*, Vol. **176** of Chemical Analysis: A Series of Monographs on Analytical Chemistry and Its Applications, John Wiley and Sons. ISBN 978-0-470-27832-1.
- 23 Fischer F., Frenner K., Herkommer A.M. (2022) Sparse mid-infrared spectra enable real-time and in-vivo applications in tissue discrimination, *EPJ Web Conf.* **266**, 02004.
- 24 Kingma D.P., Ba J. (2014) Adam: A method for stochastic optimization. arXiv preprint.