Received: 7 July 2017 | Revised: 6 October 2017 | Accepted: 12 October 2017

DOI: 10.1002/cem.2973

SPECIAL ISSUE ARTICLE

WILEY Journal of CHEMOMETRICS

Integrating spatial, morphological, and textural information for improved cell type differentiation using Raman microscopy

Sascha D. Krauß[1] | Hesham K. Yosef[1] | Tatjana Lechtonen[1] | Hendrik Jütte[2] | Andrea Tannapfel[2] | Heiko U. Käfferlein[3] | Thomas Brüning[3] | Florian Roghmann[4] | Joachim Noldus[4] | Samir F. El-Mashtoly[1] | Klaus Gerwert[1] | Axel Mosig[1]

[1]Department of Biophysics, Ruhr-University Bochum, 44780 Bochum, Germany

[2]Institute of Pathology, Bergmannsheil Hospital, Ruhr-University Bochum, 44789 Bochum, Germany

[3]Institute for Prevention and Occupational Medicine of the German Social Accident Insurance, Institute of the Ruhr-University Bochum (IPA), 44789 Bochum, Germany

[4]Department of Urology, Marien Hospital Herne, Ruhr-University Bochum, 44625 Herne, Germany

Correspondence
Axel Mosig, Department of Biophysics, Ruhr-University Bochum, 44780 Bochum, Germany.
Email: axel.mosig@bph.rub.de; axel.mosig@bph.ruhr-uni-bochum.de

Funding information
Protein Research Unit Ruhr within Europe (PURE); Ministry of Innovation, Science and Research (MIWF) of North-Rhine Westphalia, Germany; European Regional Development Fund

Abstract

Raman microscopy is a well-established tool for distinguishing different cell types in cell biological or cytopathological applications, since it can provide maps that show the specific distribution of biochemical components in the cell, with high lateral and spatial resolution. Currently, established data analysis approaches for differentiating cells of different types mostly rely on conventional chemometrics approaches, which tend to not systematically utilise the advantages provided by Raman microscopic data sets. To address this, we propose 2 approaches that explicitly exploit the large number of spectra as well as the morphological and textural information that are available in Raman microscopic data sets. *Spatial bagging* as our first approach is based on a statistical analysis of majority vote over classification results obtained from individual pixel spectra. Based on the Condorcet's Jury Theorem, this approach raises the accuracy of a relatively weak classifier for individual spectra to nearly perfect accuracy at the level of characterising whole cells. Our second approach extracts morphological and textural (*morpho-textural*) features from Raman microscopic images to differentiate cell types. While using few wavenumbers of the Raman spectrum only, our results indicate on a quantitative basis that Raman microscopic images carry more morphological and textural information than haematoxylin and eosin (H&E) stained images as the current gold standard in cytopathology. Our 2 approaches promise improved protocols for the fast acquisition of Raman imaging data, for instance, for the morphological analysis of coherent anti-Stokes Raman spectroscopy microscopic imaging data or for improving the accuracy of fibre optical probe systems by resampling spectra and utilising spatial bagging.

KEYWORDS

cytopathology, morphological classification, Raman microscopy, supervised learning, spatial bagging

1 | INTRODUCTION

Raman microscopy provides a label-free approach to characterise cellular samples at high spatial and lateral resolution, where the pixel spectra of a Raman microscopic image are both highly location specific and representative of the

biochemical sample at the respective pixel position. As such, Raman microscopy has been utilised in a broad bandwidth of cell biological applications, ranging from drug efficacy and distribution studies,[1,2] resolving subcellular components,[3,4] or cytopathological applications.[5] Throughout these applications, the predominant approach to analyse Raman microscopic pixel spectra is to apply methods established in the chemometric analysis of conventional, nonmicroscopic Raman spectra: Some approaches rely on factorisation algorithms, principal component analysis, or clustering techniques, while others utilise supervised learning approaches such as linear discriminant analysis or artificial neural networks.[6,7]

Conventional chemometrics approaches limit the analysis of Raman microscopic images in several ways. First, when dealing with larger collections of images, these approaches do not exploit the situation that the sample is represented by a large number of spectra, rather than one single spectrum. Second, conventional chemometrics approaches typically ignore *morpho-textural* characteristics, ie, features regarding the morphology or texture of the sample. In fact, there is a large body of image analysis approaches that have been established for the analysis of microscopic images obtained from other modalities such as bright-field, phase contrast, fluorescence, or histopathological staining microscopy, which explicitly analyse morpho-textural characteristics. These approaches typically operate on either pure intensity images, as obtained from fluorescence microscopes, or images with only few colour channels such as red, green, and blue (RGB) images obtained from histopathologically stained samples. Somewhat surprisingly, these approaches did not catch significant attention of the Raman microscopic research community, so that a major objective of this contribution is to assess and validate the application of morphological and textural image analysis methods to Raman microscopic image data sets. Specifically, we will study the application of the image feature approach by Boland et al[8] to a Raman microscopy–based cytopathology study.

Another major contribution of this work is to introduce the concept of *spatial bagging* for Raman microscopic image analysis, which relies on a majority vote over the classification results for a large number of pixel spectra. It has been known historically as the Condorcet's Jury Theorem (CJT) that a majority vote over a larger number of decisions of lesser competence leads to decisions with better competence as long as independence between the individual decisions can be assumed.[9] As majority votes also constitute the foundation of well-established *bagging classifiers*,[10] we refer to majority votes over the classification results of pixel spectra as *spatial bagging*.

The essential idea of training a bagging classifier on a single given training data set is to create several random subsamples of the training data and train a classifier on each of the subsamples. Observed data points can then be classified by majority vote of these subsample classifiers, which yields the bagging classifier. As has been demonstrated by Breiman, bagging classifiers may turn classifiers such as decision trees and neural networks that are unstable in a certain sense into stable classifiers.[10,11] The concept of spatial bagging introduced here explicitly exploits a substantial advantage of Raman microscopy. The large number of pixel spectra available for each data set is utilised for a majority-vote classification. In other words, we *resample observations to obtain more classification results* rather than resampling the training data to obtain more classifiers. Elementary statistical considerations promise significant gains in accuracy under certain independence assumptions. In Section 3.1 of this contribution, we investigate how far this gain can be realised on actual data.

For validation of both morphological-textural classifiers and spatial bagging, we utilise data from a recently published cytopathology study[12] in the context of bladder cancer diagnostics. Within this study, cells extracted from urine sediment were imaged using Raman microscopy and classified into tumour and normal cells based on the Raman pixel spectra. Bladder cancer is a very common[13] and expensive[14] type of cancer, where non-invasive cytopathological tests are particularly relevant.[15-18] In the present contribution, we reanalyse the data from our recent study[12] using the concepts of morphological and textural classifiers as well as spatial bagging.

Vibrational spectroscopy has been utilised extensively for identifying bladder cancer[19] previously. Yet most of these studies focus on either tissue or cells from cell culture. In contrast, our recent study[12] as well as the data investigated in this contribution use cells from voided urine of actual patients. Our main focus is on urothelial cells, which are the starting point of 95% of all bladder cancer cases,[20] so that distinguishing cancerous urothelial cells from normal urothelial cells is the key step for cytopathological classification.

1.1 | Background

1.1.1 | Spatial bagging

Raman microscopy can acquire a large number of spectra in a given sample. However, capturing Raman spectra is a time-consuming process. Since the acquisition time is a critical parameter in routine medical applications, an essential

question is raised concerning how many spectra are actually required for a reliable classification of a sample. From a statistical point of view, measuring a large number of pixel spectra from a single sample can be viewed as a *resampling* process.

To be more precise, we consider this resampling process in a cell classification setting where cells are to be classified into 2 (or potentially more) classes, typically healthy or disease. We now consider each spectrum being classified using a classifier with a certain accuracy P, and the whole cell is classified according to majority vote of the individual classification results.

To transfer CJT to Raman spectral images, we assume that the classification errors of different pixels are independent, so that the majority vote constitutes a *Bernoulli trial*. The probability of success in the trial P is constituted by the accuracy of the classifier for single spectra. Considering a 2-class setting and assuming a majority vote over classifying n pixel spectra with an accuracy of $P > .5$ for each individual classification, the probability that majority vote will be achieved for the correct class reads as

$$P(X \geq n/2) = 1 - \sum_{1 \leq i < n/2} \binom{n}{i} P^i (1-P)^{n-i}. \tag{1}$$

For $n = 100$ and $P = .55$, we obtain $P(X \geq 50) \sim .87$, in other words, majority vote promises to raise the classification accuracy from 55% to 87% under the assumption of mutually independent misclassification of pixel spectra. While this independence assumption is not realistic in general, it is still realistic to assume independence to a certain degree, as we will investigate this further in Section 3.1 (in particular Figure 3).

While the gold standard in diagnosing bladder cancer is cystoscopy, urine cytology is used as adjunctive to cystoscopy. Cytology is commonly conducted by visual inspection of haematoxylin and eosin (H&E) stained cytopathological samples through a human pathologist. As this is a cognitive performance, it is an obvious approach to assess the potential of morphology-based classifiers to improve the automated identification of cancer cells. We provide a systematic validation of morphology-based classifiers, along with a comparison to the established spectra-based counterparts. As we will discuss, a thorough understanding of the differences between spectral and morphological classifiers promises insights in how to include coherent anti-Stokes Raman spectroscopy or second harmonic generation imaging into automated label-free cytopathology approaches.

1.1.2 | Morphology-based microscopic image analysis

Morphological and textural classification of microscopic image data of cells was systematically addressed in the seminal work by Murphy and others,[8] who gathered features from a broad range across the image-processing literature and validated their application in cell biological applications. Originally, it was demonstrated that these features exhibit characteristic traces of the subcellular location of fluorescently labelled proteins.[8] In subsequent work, it was also found that they can be used for distinguishing histopathological patterns in tissue samples.[21]

The features utilised in this context include Zernike moments[22] and the morphological features proposed by Haralick.[23] Zernike moments are a set of circle polynomials in 2 polar variables. While the moments themselves are complex numbers and sensitive to rotation of the image, their magnitudes are invariant under image rotation and thus used as features. Haralick features, on the other hand, are a collection of 13 different features derived from autocorrelation, optical and digital transforms, textural edgeness, and other features that can be computationally derived from an image. Zernike moments and Haralick features are complemented by a set of 22 further measures that represent features such as convexity or excentricity of images, which are potential characteristics of cellular images. It has been shown previously that morpho-textural features are distinctive for cancer cells.[24-27] In many cases, the features collected in these studies quantify morphological and textural information contained in H&E stained images, raising the question how much morpho-textural information Raman microscopic images carry in comparison to H&E stained images of the same sample. Since suitable data for this comparison are available from our previous spectral cytopathology study,[12] we aim to address this question in our present study.

An essential part of our contributions involves the identification of individual Raman wavenumbers that will be most informative towards distinguishing cancer cells from normal cells. This translates to a feature selection problem, for which numerous approaches have been proposed.[28,29] In our context, it will be particularly relevant to identify a small number of wavenumbers, which, on the one hand, are highly distinctive between normal cells and cancer cells, while on the other hand, the selected wavenumbers should also be uncorrelated. A well-established approach to achieve this is the max-relevance min-redundancy (MRMR)[30] approach, which optimises information theoretic measures to identify features that will be informative to distinguish different classes. A crucial property of the MRMR approach is that it

identifies spectral bands that tend to carry uncorrelated information, so that the morphology for each selected band can be expected to carry different and thus additional information for morphological feature-based classification.

2 | MATERIALS AND METHODS

2.1 | Wet lab

In this study we reanalysed data from a recently published study.[12] For the sake of completeness, we briefly describe the sample preparation and image acquisition already previously described.[12]

2.1.1 | Urine sampling

This study was conducted within the framework of the Protein research Unit Ruhr within Europe (PURE) at the Ruhr-University Bochum (RUB), Germany. Prior to human urine sampling, institutional review board approval (IRB 3674-10) has been acquired, and written informed consent from all patients has been obtained. Urine samples were collected from 10 patients diagnosed with high-grade urothelial bladder cancer and from 10 patients with pathologically confirmed urocystitis but without cancer at Marienhospital Herne, Germany. Collected urine samples were spun at 3700 rpm (10 min, 10°C). The supernatant urine was discarded and followed by suspension and fixation of the precipitated pellets of urine cells using 1 mL of 4% formaldehyde solution (Roti-Histofix, Carl Roth GmbH, Karlsruhe, Germany) and stored at 4°C. In case of haematuria, urine cell solution was filtered using a Millipore nylon-net filter of 11-μm pore size (Merck Chemicals GmbH, Darmstadt, Germany) and washed with phosphate buffered saline (Life Technologies, Darmstadt, Germany), to remove blood cells, bacteria, yeast, and all other colouring contaminants.

Calcium fluoride (CaF_2) slides (Korth Kristalle, Kiel, Germany) were used as cell substrate for spectroscopic measurements. CaF_2 slides were coated with 0.01% (w/v) poly-L-Lysine (Sigma-Aldrich GmbH, Munich, Germany). An amount of 100 μL of urine cell solution was added to a cytofunnel assembly (Shandon, ThermoFisher GmbH, Dreieich, Germany), which contains a CaF_2 slide. The cytofunnel assembly was spun using cytospin centrifuge (Cytospin 4, ThermoFisher GmbH, Dreieich, Germany) at 1500 rpm for 20 minutes. Next, the CaF_2 slide carrying a spot of adhered urine cells was removed from the assembly and subsequently immersed in phosphate buffered saline buffer to proceed with Raman measurements.

2.1.2 | Spectroscopy

A confocal Raman microscope (alpha300 AR, WITec, Ulm, Germany) was implemented for Raman measurements of urine cells, as described previously.[1,2] Raman excitation source is a frequency-doubled Nd:YAG laser 532 nm (Crystal laser, Reno, USA) with the output power of $\sim$ 40 mW. The excitation laser beam is directed into a Zeiss microscope by a wavelength-specific single-mode optical fibre, which is followed by collimation of the laser beam and focused on the sample by a Nikon NIR APO (60×/1.00 NA, Nikon, Düsseldorf, Germany) water immersion objective. The urine cell slide is fixed on a piezoelectrically driven microscope scanning stage. The collected Raman-scattered light was directed to a back-illuminated deep-depletion charge-coupled device camera that is operated at −60°C, which can detect the Raman signal. In this study, Raman imaging is conducted using a raster scanning laser beam over cells, to acquire the full Raman spectra at speed of 0.5 second per pixel and a pixel resolution of 500 nm. From the 20 patients, 60 high-grade cancer urothelial cells, and 61 non-cancerous cells were selected, which resulted in 375,203 and 161,937 Raman spectra from cancerous and non-cancerous cells, respectively.

2.1.3 | Staining

The H&E staining of cells was conducted after Raman measurements. The cells were fixed via spraying of a commercially fixative solution Merkofix (Merck KGaA, Darmstadt, Germany) that was left for 10 minutes to dry. Next, the cell slide was immersed in Haris haematoxylin solution (Merck KGaA, Darmstadt, Germany) for 1 minute and subsequently washed in water stream for 1 minute. Afterwards, cells were immersed in eosin Y (0.5% alcoholic, Merck KGaA, Darmstadt, Germany) for 1 minute and then washed by a water stream for 1 min. The cell slide was immersed subsequently in multiple washing solvents (15 s each): ethanol (96%), ethanol (100%), and Xylol (100%) (Sigma-Aldrich GmbH, Munich, Germany). Lastly, a liquid cover glass (Merkoglass, Merck KGaA, Darmstadt, Germany) was added on the cell slide and

then covered with a glass coverslip (Servoprax GmbH, Wesel, Germany) and left overnight to dry. For imaging of cells stained with H&E, a Nikon upright microscope (Eclipse Ni-U, Düsseldorf, Germany) was used. The microscopic imaging was conducted using a Nikon Plan APO (60×/1.4 NA, Nikon, Düsseldorf, Germany) oil immersion objective. All collected images were sent for cytopathology annotation.

2.2 | Data analysis

A complete sample consists of thousands of cells on a microscopy slide. From each sample, a small number of cells relevant for classification are visually selected. These selected cells are first measured by Raman microscopy and subsequently stained with H&E (see Section 2.1.3). These stained cells were annotated by a pathologist as cancer or normal urothelial cells.

2.2.1 | Spectral preprocessing

Raman hyperspectral results were exported to Matlab (The MathWorks, Natick, Massachusetts), and in-house built scripts were implemented for data preprocessing. All Raman spectra without a C–H band at 2850 to 3000 cm^{-1} were treated as background and deleted. To remove cosmic spikes, an impulse noise filter was applied and the Raman spectra were interpolated to a reference wavenumber scale. Furthermore, all spectra were vector normalised. We performed classification on both uncorrected spectra and baseline-corrected spectra. Baseline correction had only a very minor effect on classification results (see Supporting Information), so that all results are presented for spectra without baseline correction.

2.2.2 | Training spectral classifiers

To train classifiers for individual spectra, we use a modified version of a previously proposed colocalization scheme[4] on the preselected cells. While this previously proposed colocalization scheme extracts training spectra through colocalization analysis between a spectral image and a fluorescence image, it is adapted here to colocalization analysis between a spectral image and an H&E image.

The H&E stained image is computationally transformed so that a thresholding algorithm can be applied to segment the cell into nucleus and cytoplasm (the latter including all non-nuclear organelles). For the corresponding spectral image, a dendrogram is computed using hierarchical clustering of the pixel spectra. Then colocalization analysis[4] uses the dendrogram to identify the best matching cluster for nucleus and cytoplasm, respectively. The set of training spectra can now be derived from the overlap between the best matching cluster and the thresholded area belonging to the nucleus or cytoplasm. This process yields a training data set that consists of 4 classes, namely, cytoplasm and nucleus for normal urothelial cells and cancer urothelial cells, respectively. Using the training spectra obtained from colocalization analysis, random forests[11] were trained for classifying individual spectra applying default parameters (100 trees and drawing 27 features per node, where 27 is roughly the square root of the number of features in each spectrum).

The resulting classifier (see classifier **S** in Figure 1) can be used for majority vote–based spatial bagging as described in Section 2.2.3. After training the random forest classifier, new and unknown cell samples are provided for validation, which we conduct at different levels: First, we measure accuracy at the level of single spectra, where the accuracy constitutes the probability P in the Bernoulli trial behind spatial bagging. Second, we assess accuracy at the level of whole cells after the majority voting process. Finally, since in general, more than one cell will be available from each patient, we also consider accuracy at patient level. Here, a patient will be considered as affected by cancer if most cells are classified as cancer cells. The validity of these results was assessed based on the accuracies obtained from *leave-one-patient-out* cross-validation. For each round of cross-validation, all cells and samples associated with one patient were withheld for validation, and training was performed on the remaining patients' cells, as proposed and discussed previously.[4] This means that the division into training and validation sets was performed at the highest possible hierarchical level (being the patient) to correctly evaluate the classification models.[31]

2.2.3 | Spatial bagging

Spatial bagging classification (classifier **SB** in Figure 1) is performed as follows: We assume each pixel spectrum is measured at D many wavenumbers, so that an individual spectrum x can be considered as a D-dimensional vector $x \in \mathbb{R}^D$. We assume a given classifier $c : \mathbb{R}^D \to \{0, 1\}$ that classifies a spectrum in 2 classes, in cytopathology typically cancer vs normal. Furthermore, a cell is measured as a spectral image within a coordinate system Γ as a finite set of pixel positions,

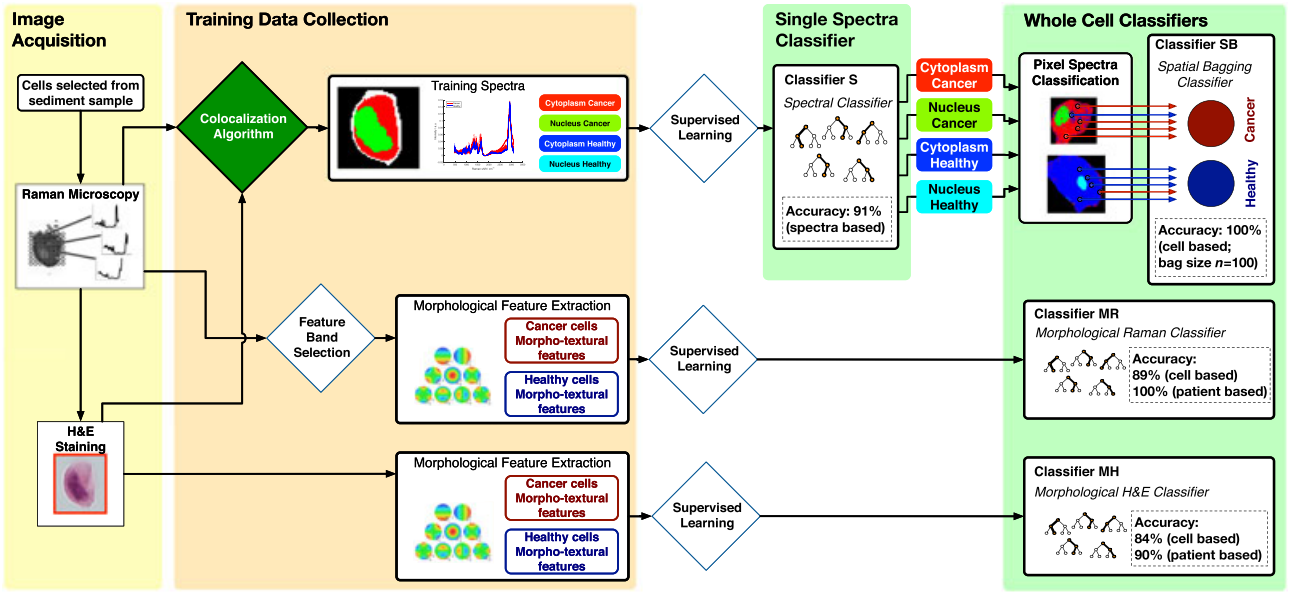

FIGURE 1 Overview of the workflow and results for training and validating the spectral and morphological cell classifiers. Cells from urine sediment are imaged using Raman microscopy and then stained with H&E for conventional light microscopic imaging. Training data for the *spectral* (**S**) classifier are obtained from colocalization analysis between the 2 microscopic image modalities. The training data comprise 4 classes: cytoplasm and nucleus, each in a cancerous and non-cancerous version. The accuracy of 90% on single spectra can be enhanced even further to 100% via *spatial bagging* as described in Section 2.2.3. For the *morphological Raman* (**MR**) classifier, 3 wavenumbers are selected using MRMR feature selection as detailed in Section 2.2.4 and Figure 2. Morphological features are extracted from these images as described in Section 2.2.4. The random forest trained on these features results in an accuracy of 89% correctly classified cells and 100% based on patients. The second classifier (**MH**) is trained on morphological features extracted from the H&E images rather than the Raman images, which results in lower accuracies (84% and 90%)

so that the spectral image is available as a mapping

$$I : \Gamma \to \mathbb{R}^D.$$

To classify a complete image, we obtain a bagging classifier B_c for I through a majority vote

$$B_c(I) = \begin{cases} 1 \text{ if } |\{x \in \Gamma | c(I(x)) = 1\}| > |\Gamma|/2 \\ 0 \text{ otherwise,} \end{cases}$$

where $|\Gamma|$ reads as the number of pixels in the image. Following the assumption that the classification errors of different pixel spectra are (at least to a certain degree) mutually independent, the bagging classifier B_c conducts a Bernoulli trial and promises a higher accuracy than the original classifier c according to Equation 1.

Beside majority voting, classifying the mean spectrum through what we will refer to as the *mean classifier*

$$M_c(I) = c\left(\frac{1}{|\Gamma|}\sum\nolimits_{x\in\Gamma} I(x)\right), \tag{2}$$

which constitutes another straightforward and commonly used approach to classify a complete spectral image based on the single-spectrum classifier c.[2,32-37]

2.2.4 | Morphological classification

For morphological classification, the features proposed by Boland et al[8] were computed for each spectral image. As these features are computed from grey scale intensity images, we proceeded as visualised in Figure 2. We collected one set of spectra comprising all tumour cell spectra, and a second set comprising all spectra contained in normal cells of our data set. On these 2 classes of spectra, we performed feature selection using the MRMR approach,[30] to select 3

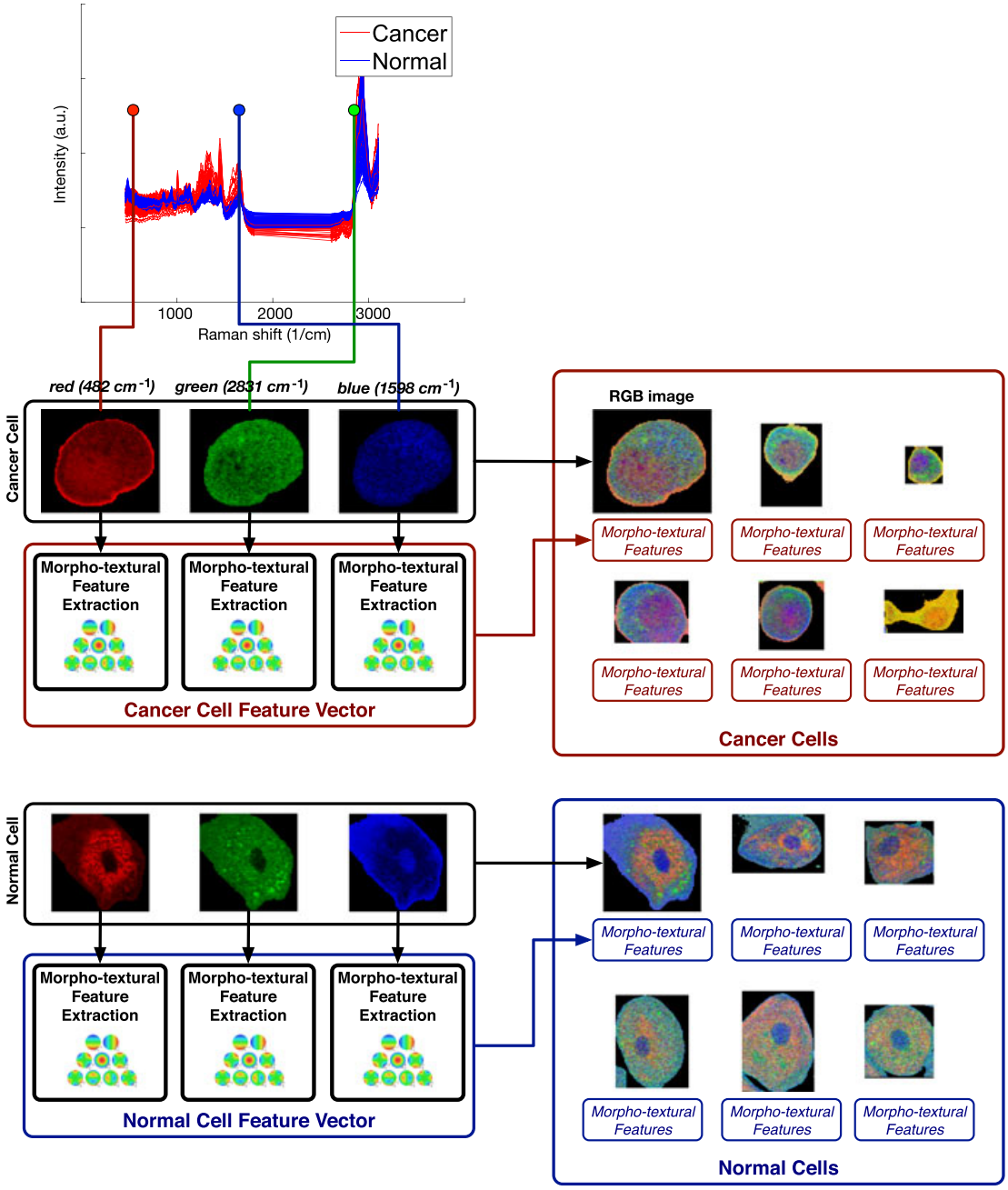

FIGURE 2 Wavenumber selection and feature extraction for morpho-textural classification. MRMR feature selection on a cancer vs normal spectral data set yields 482, 2831, and 1598cm^{-1} as 3 discriminative wavenumbers. Morpho-textural features are computed for the intensity images obtained from each of these wavenumbers, which are concatenated to obtain a feature vector representing each cell. Visual inspection of the RGB images resulting from the 3 selected wavenumbers indicate distinct morphology and textures between cancer and normal cells, for instance, nuclei sharply separated from cytoplasm in normal cells, but not in cancer cells

wavenumbers. Each of these wavenumbers yields a grey scale intensity image, for which morpho-textural features were computed and concatenated into one morpho-textural feature vector representing the complete image. A random forest was trained on the resulting image feature vectors, yielding the classifier depicted as classifier **MR** in Figure 1. In a similar manner, we obtained classifier **MH** by computing the morphological feature vector for each of the H&E stained RGB images.

For classifiers **MR** and **MH**, we determined most discriminative feature vectors through the gains in Gini importance provided for each feature by the random forest implementation.

3 | RESULTS AND DISCUSSION

3.1 | Spatial bagging

Figure 3 demonstrates the gain in accuracy achieved by spatial bagging with increasing bag size. While the Bernoulli process promises a theoretically perfect classification with a bag size of $n < 20$ in the case of mutually independent classification errors, this perfect classification is achieved in practice with a bag size of $n \geq 100$. In addition to spatial bagging, we also assessed the performance of a mean classifier, whose accuracy failed to converge to the accuracy obtained with majority decisions. We also compared the classification accuracies obtained from nucleus, cytoplasm, and full cell spectra. As it turned out, the cytoplasm data set required twice the number of spectra compared to the full cell to achieve 100% accuracy. Classification based on nucleus spectra did not exceed a value of 91%. These results clearly indicate that a separation of organelles is not relevant for the identification of urothelial bladder cancer.

Default parameters were used for training the random forest classifiers (100 trees, 27 features drawn for each node). Using less than 100 trees and drawing fewer features per node have no observable effect on accuracy while saving computation time (data not shown). The accuracy of 100% could be reproduced consistently with no exception for spatial bag sizes > 100 when rerunning the pipeline with different random reinitialisation.

3.2 | Morphological classification

The morphological classifier on Raman intensity images (**MR** in Figure 1) yielded an accuracy of 89%. Performing a majority decision over all cells associated with one patient, we achieve a perfect classification with 100% accuracy. Limiting feature extraction to the intensity image of the 482-cm^{-1} band, the cell-based and the patient-based accuracies dropped slightly to 88% and 98%, respectively. The **MH** classifier trained on features extracted from the H&E image resulted in an accuracy of 84% (cell-based) and 90% (patient-based).

Additionally, we extracted the 2 most discriminating features from urine cell images based on the intensities of the 3 wavenumbers determined by the MRMR algorithm, which can be seen in Figure 4 together with 2 components from principal component analysis. The most discriminating features are determined via the mean drop in Gini index during the training of the random forest classifier—the higher the drop in this measure of inequality by permuting the value of a feature, the more important it is for discrimination of classes. Interestingly, the 2 selected features consist of 2 Haralick features from 2 different wavenumbers stressing the importance of textural features in this discrimination task.

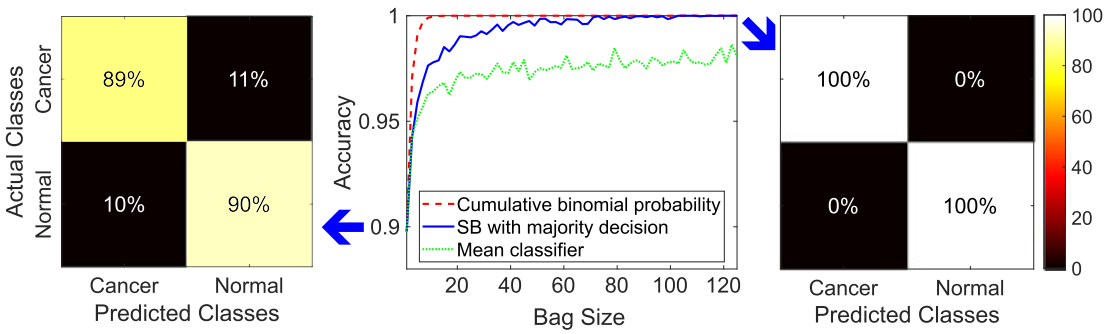

FIGURE 3 Spatial bagging increases accuracy and identifies optimal number of spectra to measure per patient. The left panel shows the confusion matrix of leave-one-patient out cross validation on a per-spectra basis. With a sensitivity of 89% and a specificity of 90% correctly classified healthy cell spectra, the classifier achieves an accuracy of 89.5%. The central panel shows the gain in accuracy of the spatial bagging procedure as the number of spectra involved in majority voting is being increased. Besides the actually observed accuracies during spatial bagging with majority decision (solid-blue line), the theoretical optimum obtained from the cumulative binomial probability is indicated by the dashed red line and results of the mean classifier by the dotted green one. Spatially bagging $n > 100$ spectra results in an accuracy, sensitivity, and specificity of 100% (as additionally shown in the right panel), while the mean classifier scores lower with 98% maximum accuracy

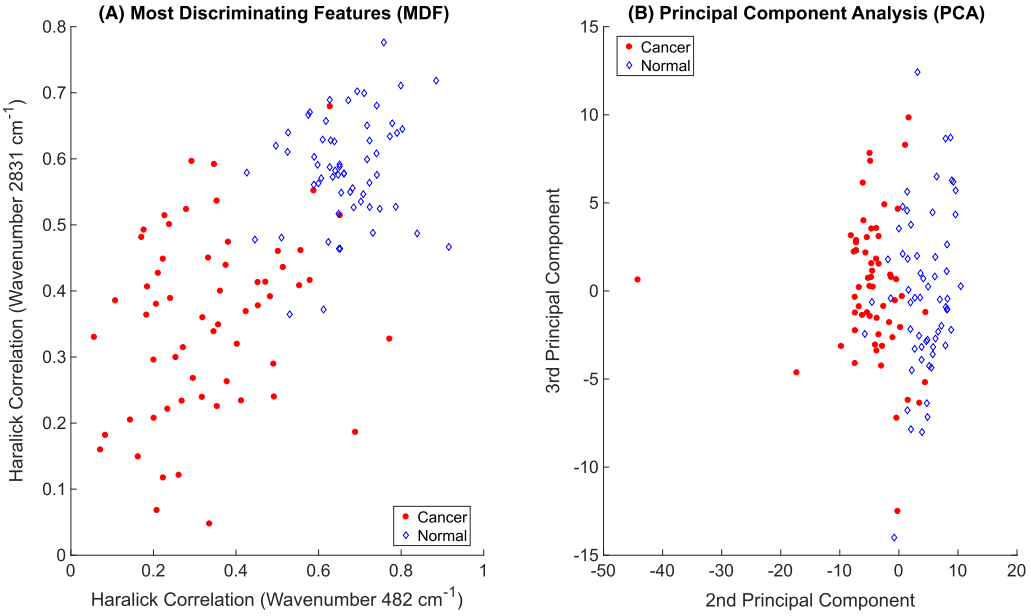

FIGURE 4 Two-dimensional projections of (A) the 2 most discriminative features and (B) principal components of the extracted morphological features of urine cell images based on the intensities of the 3 wavenumbers determined by the MRMR algorithm. The cancer dataset (red circles) appears distinct from the healthy cell samples (blue diamonds) in both of these projections

4 | CONCLUSIONS

We introduced and assessed 2 approaches for spectral image classification. First, we introduced spatial bagging as an application of CJT to spectral imaging data. Second, we proposed morphological classifiers using a set of well-established morphological and textural features that can be extracted from spectral imaging data.

The majority vote behind spatial bagging is a straightforward idea supported by the simple statistics of a Bernoulli trial and achieves a high gain in accuracy even for relatively small bag sizes around $n = 100$. It performs particularly well in comparison to mean classifiers. For novel Raman spectra acquisition protocols, this suggests that it will be preferable to resample a larger number of relatively noisy spectra with short acquisition times from a sample, rather than collecting a single spectrum with high signal-to-noise ratio and correspondingly high acquisition time. At the same time, spatial bagging provides a framework to minimise the number of spectra to be acquired for cell identification applications of Raman microscopy.

It is remarkable that morpho-textural classification of cells with an accuracy of 89% almost matches the accuracy of a single-spectrum Raman classifier (90%), since only a small part—3 out of several hundred wavenumbers—of the Raman spectrum is involved in this classifier. This potential of morphology-based classification is of high relevance for coherent anti-Stokes Raman spectroscopy and second harmonic generation microscopy, where a few wavenumbers of the spectrum can be acquired at high speed.

Finally, morpho-textural classifiers based on Raman spectral images clearly exceed the accuracy of morpho-textural classifiers based on H&E images, which achieve only 84% per cell accuracy. This provides a strong quantitative indication that Raman spectral images are morphologically and texturally more informative than H&E images, which are the basis of currently established urine cytology protocols. Again, this is accomplished using only a small fraction of the information provided by the Raman spectral image.

The 2 approaches introduced here promise to be useful beyond cytopathology applications of Raman microscopy. In fact, Raman microscopy has been utilised in numerous other settings to identify different cell types in a label-free manner, where the concept of spatial bagging will be useful immediately. In a broader context, spatial bagging exploits what is possibly the biggest advantage in the analysis of both Raman and infrared microscopic imaging data: Each data set is not just represented by one single feature vector but actually by thousands or, in some cases, even millions[38] of feature vectors. It appears that the potential advantages resulting from this have not been fully exploited to date, and spatial bagging constitutes one step in this direction. For example, spatial bagging may be of relevance in the context of fibre optical Raman probes,[39] where spatial bagging could be applicable by collecting several spectra around one investigated spot.

ACKNOWLEDGEMENTS

This research was supported by the Protein Research Unit Ruhr within Europe (PURE), Ministry of Innovation, Science and Research (MIWF) of North-Rhine Westphalia, Germany, and the European Regional Development Fund, European Union and North-Rhine Westphalia, Germany.

ORCID

Sascha D. Krauß http://orcid.org/0000-0001-6271-9940
Samir F. El-Mashtoly http://orcid.org/0000-0001-6087-8817

REFERENCES

1. El-Mashtoly SF, Petersen D, Yosef HK, et al. Label-free imaging of drug distribution and metabolism in colon cancer cells by Raman microscopy. *Analyst*. 2014;139(5):1155-1161.
2. El-Mashtoly SF, Yosef HK, Petersen D, et al. Label-free Raman spectroscopic imaging monitors the integral physiologically relevant drug responses in cancer cells. *Anal Chem*. 2015;87(14):7297-7304.
3. El-Mashtoly SF, Niedieker D, Petersen D, et al. Automated identification of subcellular organelles by Coherent anti-Stokes Raman scattering. *Biophys J*. 2014;106(9):1910-1920.
4. Krauß SD, Petersen D, Niedieker D, et al. Colocalization of fluorescence and Raman microscopic images for the identification of subcellular compartments: a validation study. *Analyst*. 2015;140(7):2360-2368.
5. Diem M, Mazur A, Lenau K, et al. Molecular pathology via IR and Raman spectral imaging. *J Biophotonics*. 2013;6(11-12):855-886.
6. Krafft C, Steiner G, Beleites C, Salzer R. Disease recognition by infrared and Raman spectroscopy. *J Biophotonics*. 2009;2(1-2):13-28.
7. Hedegaard M, Matthäus C, Hassing S, Krafft C, Diem M, Popp J. Spectral unmixing and clustering algorithms for assessment of single cells by raman microscopic imaging. *Theor Chem Acc*. 2011;130(4-6):1249-1260.
8. Boland MV, Murphy RF. A neural network classifier capable of recognizing the patterns of all major subcellular structures in fluorescence microscope images of HeLa cells. *Bioinformatics*. 2001;17(12):1213-1223.
9. Boland PJ. Majority systems and the Condorcet jury theorem. *Statistician*. 1989;38(3):181-189.
10. Breiman L. Bagging predictors. *Mach Learn*. 1996;24(2):123-140.
11. Breiman L. Random forests. *Mach Learn*. 2001;45(1):5-32.
12. Yosef HK, Krauß SD, Lechtonen T, et al. Non-invasive diagnosis of high-grade urothelial carcinoma in urine by Raman spectral imaging. *Anal Chem*. 2017;89(12):6893-6899.
13. Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer*. 2015;136(5):E359-E386.
14. Botteman MF, Pashos CL, Redaelli A, Laskin B, Hauser R. The health economics of bladder cancer. *Pharmacoeconomics*. 2003;21(18):1315-1330.
15. Bonberg N, Taeger D, Gawrych K, et al. Chromosomal instability and bladder cancer: the UroVysion™test in the UroScreen study. *BJU Int*. 2013;112(4):E372-E382.
16. Huber S, Schwentner C, Taeger D, et al. Nuclear matrix protein-22: a prospective evaluation in a population at risk for bladder cancer. Results from the UroScreen study. *BJU Int*. 2012;110(5):699-708.
17. Johnen G, Gawrych K, Bontrup H, et al. Performance of survivin mRNA as a biomarker for bladder cancer in the prospective study UroScreen. *PlOS ONE*. 2012;7(4):e35363 pages 1-10.
18. Pesch B, Taeger D, Johnen G, et al. Screening for bladder cancer with urinary tumor markers in chemical workers with exposure to aromatic amines. *Int Arch Occup Environ Health*. 2014;87(7):715-724.
19. Kerr LT, Domijan K, Cullen I, Hennelly B. M. Applications of Raman spectroscopy to the urinary bladder for cancer diagnostics. *Photonics Lasers Med*. 2014;3(3):193-224.
20. Dahm P, Gschwend JE. Malignant non-urothelial neoplasms of the urinary bladder: a review. *Eur Urol*. 2003;44(6):672-681.
21. Newberg J, Murphy RF. A framework for the automated analysis of subcellular patterns in human protein atlas images. *J Proteome Res*. 2008;7(6):2300-2308.
22. Zernike F. Beugungstheorie des schneidenverfahrens und seiner verbesserten form, der phasenkontrastmethode. *Physica*. 1934;1(7-12):689-704.
23. Haralick RM. Statistical and structural approaches to texture. *Proc IEEE*. 1979;67(5):786-804.
24. Doyle S, Hwang M, Shah K, Madabhushi A, Feldman M, Tomaszeweski J. Automated grading of prostate cancer using architectural and textural image features. In: Biomedical Imaging: from Nano to Macro, 2007. ISBI 2007. 4th IEEE International Symposium on IEEE. Washington DC, USA; 2007:1284-1287.
25. Tahmasbi A, Saki F, Shokouhi SB. Classification of benign and malignant masses based on Zernike moments. *Comput Biol Med*. 2011;41(8):726-735.

26. Moura DC, López MAG. An evaluation of image descriptors combined with clinical data for breast cancer diagnosis. *IJCARS*. 2013;8(4):561-574.
27. Yuan Y, Failmezger H, Rueda OM, et al. Quantitative image analysis of cellular heterogeneity in breast tumors complements genomic profiling. *Sci Transl Med*. 2012;4(157):157ra143-157ra143.
28. Guyon I, Elisseeff A. An introduction to variable and feature selection. *J Mach Learn Res*. 2003;3(Mar):1157-1182.
29. Saeys Y, Inza I, Larrañaga P. A review of feature selection techniques in bioinformatics. *Bioinformatics*. 2007;23(19):2507-2517.
30. Peng H, Long F, Ding C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans Pattern Anal Mach Intell*. 2005;27(8):1226-1238.
31. Guo S, Bocklitz T, Neugebauer U, Popp J. Common mistakes in cross-validating classification models. *Anal Methods*. 2017;9(30):4410-4417.
32. Tolstik T, Marquardt C, Matthäus C, et al. Discrimination and classification of liver cancer cells and proliferation states by Raman spectroscopic imaging. *Analyst*. 2014;139(22):6036-6043.
33. Nawaz H, Garcia A, Meade AD, Lyng FM, Byrne HJ. Raman micro spectroscopy study of the interaction of vincristine with a549 cells supported by expression analysis of bcl-2 protein. *Analyst*. 2013;138(20):6177-6184.
34. Yosef HK, Mavarani L, Maghnouj A, Hahn S, El-Mashtoly SF, Gerwert K. In vitro prediction of the efficacy of molecularly targeted cancer therapy by Raman spectral imaging. *Anal Bioanal Chem*. 2015;407(27):8321-8331.
35. Isabelle M, Dorney J, Lewis A, et al. Multi-centre Raman spectral mapping of oesophageal cancer tissues: a study to assess system transferability. *Faraday Discuss*. 2016;187:87-103.
36. Santos IP, Caspers PJ, Bakker Schut TC, et al. Raman spectroscopic characterization of melanoma and benign melanocytic lesions suspected of melanoma using high-wavenumber Raman spectroscopy. *Anal Chem*. 2016;88(15):7683-7688.
37. Schie IW, Kiselev R, Krafft C, Popp J. Rapid acquisition of mean Raman spectra of eukaryotic cells for a robust single cell classification. *Analyst*. 2016;141(23):6387-6395.
38. Kallenbach-Thieltges A, Großerüschkamp F, Mosig A, Diem M, Tannapfel A, Gerwert K. Immunohistochemistry, histopathology and infrared spectral histopathology of colon cancer tissue sections. *J Biophotonics*. 2013;6(1):88-100.
39. Petersen D, Naveed P, Ragheb A, et al. Raman fiber-optical method for colon cancer detection: cross-validation and outlier identification approach. *Spectrochim Acta Mol Biomol Spectrosc*. 2017;181:270-275.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Krauß SD, Yosef HK, Lechtonen T, et al. Integrating spatial, morphological, and textural information for improved cell type differentiation using Raman microscopy. *Journal of Chemometrics*. 2017;e2973. https://doi.org/10.1002/cem.2973