

# Fast and Noninvasive Diagnosis of Cervical Cancer by Coherent Anti-Stokes Raman Scattering

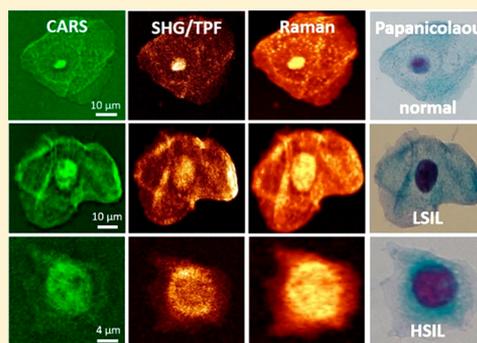
Karim Aljakouch,<sup>†,||</sup> Ziad Hilal,<sup>§,||</sup> Ibrahim Daho,<sup>†,||</sup> Martin Schuler,<sup>†,‡,||</sup> Sascha D. Krauß,<sup>†,||</sup> Hesham K. Yosef,<sup>†</sup> Johann Dierks,<sup>†,‡</sup> Axel Mosig,<sup>†,‡</sup> Klaus Gerwert,<sup>†,‡</sup> and Samir F. El-Mashtoly<sup>\*,†,‡</sup>

<sup>†</sup>Department of Biophysics and <sup>‡</sup>Center for Protein Diagnostics (ProDi), Biospectroscopy, Ruhr University Bochum, 44801 Bochum, Germany

<sup>§</sup>ZyDoLab, 44137 Dortmund, Germany

## Supporting Information

**ABSTRACT:** Cervical cancer is the fourth most common cancer in women worldwide, and early detection of its precancerous lesions can decrease mortality. Cytopathology, HPV testing, and histopathology are the most commonly used tools in clinical practice. However, these methods suffer from many limitations such as subjectivity, cost, and time. Therefore, there is an unmet clinical need to develop new noninvasive methods for the early detection of cervical cancer. Here, a novel noninvasive, fast, and label-free approach with high accuracy is presented using liquid-based cytology Pap smears. CARS and SHG/TPF imaging was performed at one wavenumber on the Pap smears from patients with specimens negative for intraepithelial lesions or malignancy (NILM), and low-grade (LSIL) and high-grade (HSIL) squamous intraepithelial lesions. The normal, LSIL, and HSIL cells were selected on the basis of the ratio of the nucleus to the cytoplasm and cell morphology. Raman spectral imaging of single cells from the same smears was also performed to provide integral biochemical information of cells. Deep convolutional neural networks (DCNNs) were trained independently with CARS, SHG/TPF, and Raman images, taking into account both morphotextural and spectral information. DCNNs based on CARS, SHG/TPF, or Raman images have discriminated between normal and cancerous Pap smears with 100% accuracy. These results demonstrate that CARS/SHG/TPF microscopy has a prospective use as a label-free imaging technique for the fast screening of a large number of cells in cytopathological samples.



Cervical cancer is the fourth most common cancer in women worldwide, with 266 000 deaths in 2012.<sup>1,2</sup> The peak rate of cervical cancer cases is found in middle-aged women between 35 and 44 years of age. Early detection of the precancer stage is necessary to reduce the mortality associated with cervical cancer significantly. Mostly, cervical cancer develops in the basal layer of cells lining the cervix and progresses gradually, revealing several dysplastic changes that can lead to invasive cancer. The Papanicolaou (Pap) test or so-called Pap smear is the most common screening method for identifying an abnormality in the cervix.<sup>3</sup> Abnormal Pap smears are followed by colposcopy, biopsy, and histopathological investigation to confirm the diagnosis.

The Pap test is a noninvasive method, extensively accepted, and its results include the following categories according to the Bethesda system:<sup>4</sup> negatives for intraepithelial lesions or malignancy (NILM), atypical squamous cells of undetermined significance (ASCUS), low-grade squamous intraepithelial lesions (LSIL), and high-grade squamous intraepithelial lesions (HSIL). This cytology method depends on the visual evaluation of individual cell morphology and detects cancer and precancer cells, making it highly subjective with a large variation in the sensitivity (50–96%).<sup>5–7</sup> Persistent infection

with human papillomavirus (HPV) is the major risk factor for the development of cervical cancer.<sup>8</sup> An HPV-DNA test is used to screen for HPV-DNA fragments and determine whether the patient is infected with one of the HPV high-risk types. Although the HPV-DNA test has a higher sensitivity (~95%), it suffers from low specificity (~84%) and is also expensive.<sup>9</sup> Therefore, there is an unmet clinical need to develop a new noninvasive method for cervical cancer screening.

Raman spectroscopic methods including conventional Raman spectroscopy, coherent anti-Stokes Raman scattering (CARS), and stimulated Raman scattering (SRS) are emerging biophotonic tools in the bioanalysis and imaging of biomaterials such as body fluids, cells, and tissues.<sup>10–21</sup> These methods are nondestructive and label-free approaches with a spectroscopic capability to probe different biomolecules and monitor their changes with the progression of cancer or neurodegenerative diseases. Conventional Raman spectroscopy has been applied extensively to cervical cancer specimens, especially tissues and cell lines; however, limited applications

Received: July 26, 2019

Accepted: September 4, 2019

Published: September 4, 2019

on cervix cytology were reported.<sup>22–27</sup> For instance, Raman spectroscopy was used to differentiate between HPV-positive and HPV-negative Pap smears with high accuracy.<sup>22</sup> On the other hand, the classification of normal and cancerous Pap smears was achieved with lower accuracy (80%).<sup>23</sup> This is most likely because Raman spectra were acquired using cell pellets instead of individual cells, leading to sample heterogeneity.

Recent studies have shown that few Raman spectra from the nuclei of single cells can discriminate between normal and abnormal Pap smears with high accuracy using principal component analysis–linear discriminant analysis (PCA–LDA) and partial least-squares discriminant analysis (PLS–DA).<sup>25,26</sup> However, these few spectra neither represent the integral biochemical composition of cells nor allow access to the morphological or textural features of cells because of the lack of a Raman imaging modality.<sup>25,26</sup> In addition, the use of conventional chemometric approaches such as PCA–LDA and PLS–DA does not consider the morphological or textural features of cells.

Coherent Raman techniques such as CARS and SRS imaging have been used recently in many biomedical applications. These methods are much faster than conventional Raman spectroscopy and can be performed at a speed of up to a video rate, allowing fast diagnosis.<sup>10,11,19,28–30</sup> We have recently used a combination of CARS and second harmonic generation (SHG) imaging as a fast tool for prescreening urothelial cells in urine sediments.<sup>10</sup> Afterward, Raman spectral imaging of the selected cells was performed to differentiate between noncancerous and cancerous urothelial cells using deep convolutional neural networks (DCNNs). The results have shown the advantage of DCNN for Raman data compared to the conventional chemometric methods.<sup>31,32</sup> This is because the DCNN classifications were based on not only the spectral information but also the morphological features of the cell. Thus, a combination of CARS imaging, much faster than Raman imaging, and DCNNs would be a perfect candidate for a fast label-free imaging approach for the diagnosis of cancer using cytopathological samples. Such combination was used for the detection of lung and head and neck carcinoma using only tissue sections that were collected invasively from patients but not using liquid-based cytology.<sup>33,34</sup>

Here, we report for the first time fast CARS, SHG, and two-photon excited autofluorescence (TPF) imaging of liquid-based cytology including normal, LSIL, and HSIL Pap smears that were collected noninvasively from patients. Cells were screened within a very short time. Raman spectral imaging of single cells from the same Pap smears was also acquired, and the results provided not only integral biochemical information of single cells but also morphological features of cells. DCNNs were used to discriminate with very high accuracy among normal, LSIL, and HSIL cells in Pap smears based on morphological features extracted from CARS, SHG/TPF, and Raman microscopic images. Finally, the results demonstrate that CARS and SHG/TPF imaging has the potential to be a fast and noninvasive method for the diagnosis of cervical cancer with high accuracy.

## EXPERIMENTAL SECTION

**Pap Smears.** Pap smears were collected from 10 healthy women, 10 patients diagnosed with LSIL, and 10 patients diagnosed with HSIL by ZYDOLAB (Institute for Clinical

Cytology and Immune Cytochemistry; Dortmund, Germany). Institutional review board approval (IRB 16-5654) and written informed consent from all patients have been obtained. Pap smears were provided using liquid-based cytology. For this method, samples from the cervix uteri were collected using a cervical sampler and deposited into preservative liquid. This technique allows accurate results because of the removal of mucus, blood, and other elements. The preparation of Pap smears for Raman and CARS/SHG/TPF measurements is shown in the [Supporting Information \(SI\)](#).

**CARS and SHG Microscopic Imaging.** CARS and SHG images were acquired using a commercial setup (TCS SP5 II CARS; Leica Microsystems, Heidelberg, Germany) consisting of a picosecond pulsed laser setup (APE picoEmerald, Berlin, Germany). It generates and synchronizes two collinearly aligned beams to a confocal inverted microscope as reported previously.<sup>10,28</sup> The pump and Stokes wavelengths were adjusted to 810.5 and 1064 nm, respectively. Laser beams are focused on the microscope using a water-immersion objective (HCX IRAPO L, 25X/0.95 W, Leica Microsystems). CARS and SHG/TPF images of Pap smears from 30 patients were acquired simultaneously at 2935 cm<sup>-1</sup> with a pixel dwell time of 180  $\mu$ s and a pixel resolution of 250 nm. CARS imaging is displayed in one channel, while SHG and TPF imaging is shown in another channel.

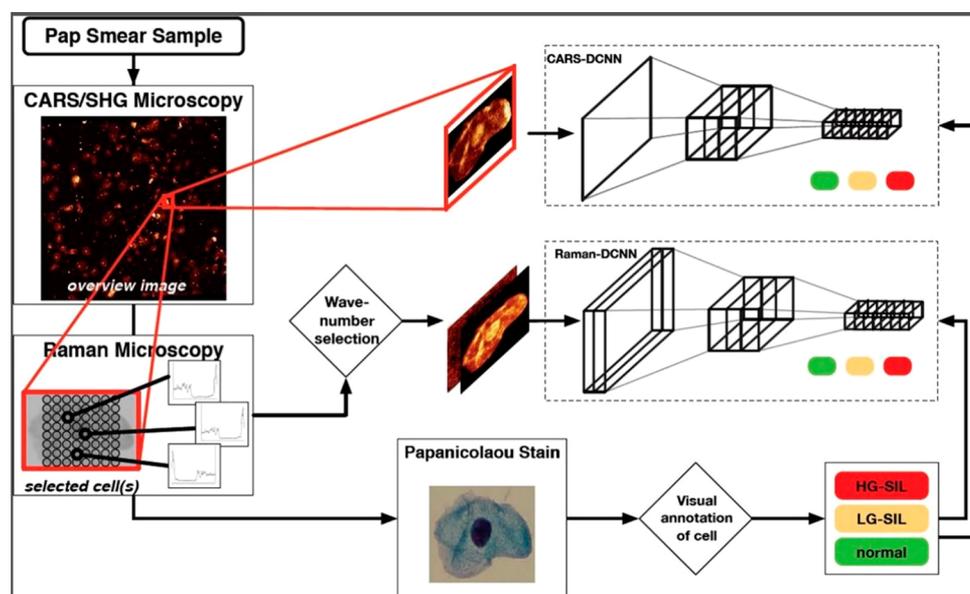
**Raman Spectral Imaging.** Raman spectral imaging of cells in Pap smears was measured using an alpha300 RA confocal Raman microscope (WITec, Ulm, Germany) as described previously.<sup>12–14,35–37</sup> A frequency-doubled Nd:YAG laser operating at 532 nm (Crystal Laser, Reno, NV, USA) is the Raman excitation source. The laser beam is coupled into a microscope using a single-mode optical fiber, and it is collimated and then focused on the sample through a Nikon NIR APO (60x/1.00 NA) water-immersion objective. Raman measurements were performed using a raster scanning laser beam over cervical cells in order to measure full Raman spectra (0.5 s per pixel) with a pixel resolution of 500 nm. The Raman measured 82 normal cells, 32 LSIL cells, and 41 HSIL cells that were selected from 30 patients and produced 590 676, 80 183, and 73 574 Raman spectra, respectively.

**Deep Learning/Minimal Net (MNi).** We devised a minimal topology with  $i = 1$  spectral bands in the input layer and one with  $i = 6$  spectral bands, both including six hidden layers. The topology of MNi was described previously.<sup>31</sup> Varying the number  $i$  of spectral bands allowed us to investigate the impact of including different amounts of spectral information for classification, where  $i = 6$  was previously shown to be the best selection for Raman images.<sup>31</sup>

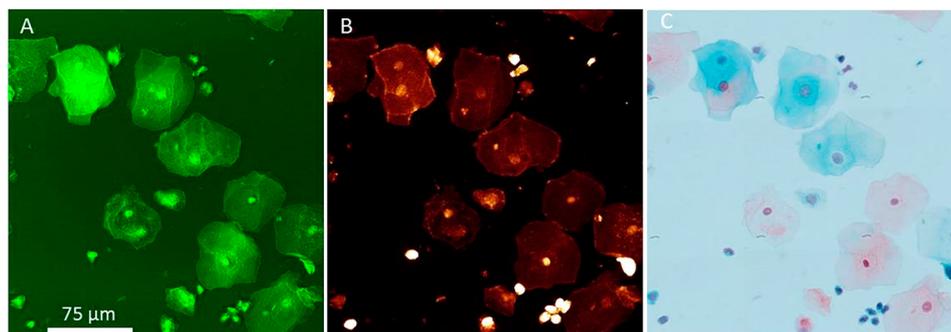
**Implementation.** The MNi networks were implemented in Python 3.6 using the Keras 2.0.9 and TensorFlow 1.4.0 deep learning frameworks. Preprocessing of Raman spectra was based on implementations, and the training of the networks is described in the [SI](#). Experiments were run on a server with 20 CPUs and 1 Nvidia GTX Titan X graphics card running Ubuntu 16.04.

## RESULTS AND DISCUSSION

**Workflow for the Identification of Cervical Cancer Cells in Pap Smears.** A workflow including Raman spectral imaging, CARS/SHG/TPF imaging, cluster analysis, and DCNNs was established to identify cervical cancer cells in Pap smears of patients diagnosed with negative for intra-epithelial lesion or malignancy (NILM), low-grade squamous



**Figure 1.** Workflow for the detection of cervical cancer cells in Pap smears using deep learning.



**Figure 2.** Unstained CARS (A) and SHG/TPF (B) images and a Papanicolaou-stained image (C) of the same HSIL Pap smear.

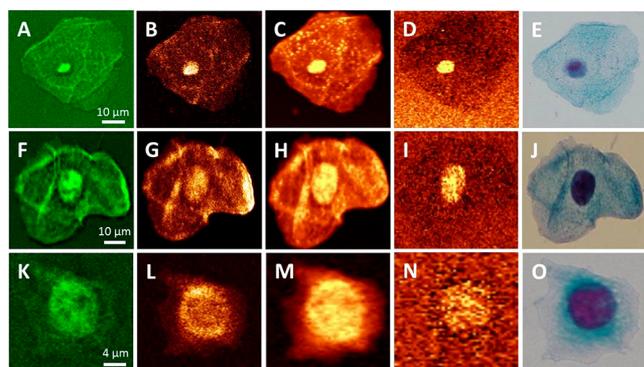
intraepithelial lesions (LSIL), and high-grade squamous intraepithelial lesions (HSIL). As shown in Figure 1, fast CARS and SHG/TPF imaging techniques were first used to acquire images of Pap smears in order to identify cervical cancer cells. Subsequently, the identified cells were measured by Raman spectral imaging. These images or spectral images were used to train DCNNs.

CARS is a nonlinear method that depends on high-intensity excitation pulses to produce an efficient signal. Consequently, additional nonlinear effects such as SHG and TPF are simultaneously generated.<sup>11,29</sup> CARS and SHG/TPF images of Pap smears were measured simultaneously within a short time covering a large area as shown in Figure S1 and were used for the fast screening of cervical cancer cells. A similar approach was applied recently to the screening of urothelial cancer cells in urine sediments.<sup>10</sup> The nucleus to cytoplasm ratio is the main feature used to distinguish normal, LSIL, and HSIL cervical cancer cells in addition to the color of the cytoplasm and cell shape in Papanicolaou-stained images. Slater et al. suggested calculating this ratio by utilizing the mean diameter of the cytoplasm rather than the increasing size of the nucleus.<sup>38,39</sup> Accordingly, a cell with a nucleus-to-cytoplasm ratio of less than 25% is considered to be a normal intermediate cell, a cell with <50% as LSIL and >50% as HSIL. A similar approach was applied here to identify normal, LSIL, and HSIL cells using label-free CARS/SHG/TPF images

because the cell nuclei were visible in these images as shown in Figure 2. Cancer cells were detected in all cancer samples measured by CARS/SHG/TPF imaging. After CARS/SHG/TPF and Raman measurements, Papanicolaou staining was performed and the annotation of these cells was confirmed by a cytologist (Z.H.) using the stained images of the same cells.

It is noted that Pap smears contain intermediate and superficial cells and these cells cannot be differentiated in unstained slides (Figure 2A,B). In the stained image (Figure 2C), the intermediate and superficial cells are blue and orange to pink, respectively, in color. It was reported that a mixed population of superficial and intermediate cell types can be used for the classification based on Raman spectra of different Pap smears.<sup>26</sup> In the present study, a mixed population of cell types was used, but most of the measured cells are intermediate cells.

A representative example of different imaging results of a cell from one patient's Pap smear of each type is shown in Figure 3. CARS (A, F, K) and SHG/TPF (B, G, L) images of normal and LSIL and HSIL cells are displayed, respectively. CARS microscopic imaging at  $2935\text{ cm}^{-1}$  monitors mainly the distribution of proteins in cells.<sup>10,11,40</sup> On the other hand, SHG visualizes the collagen and myosin structures in cells.<sup>41</sup> Myosin is present in the cell nuclei and is involved in several nuclear functions.<sup>42</sup> TPF shows the distribution of nicotinamide adenine dinucleotide (NADH) and flavins/flavoproteins.<sup>43</sup>

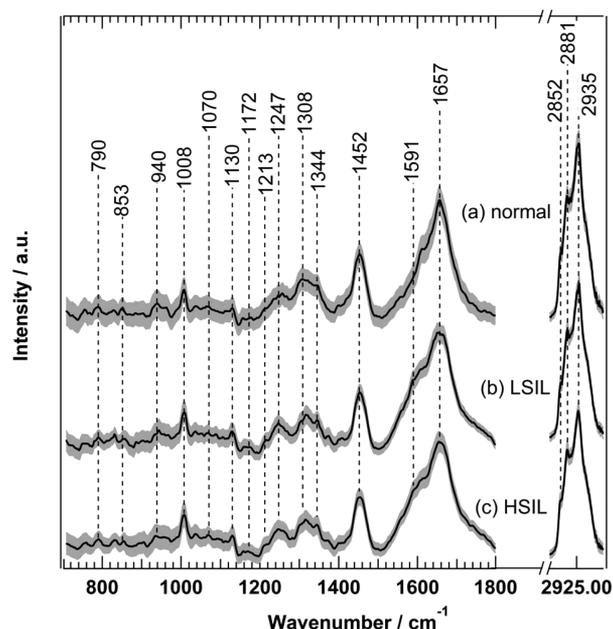


**Figure 3.** Multimodal imaging techniques of normal (A–E), LSIL (F–J), and HSIL (K–O) cells in Pap smears. CARS images (A, F, K), SHG/TPF images (B, G, L). Integrated Raman intensities of cells in the 2850–3050  $\text{cm}^{-1}$  region (C, H, M) and in the 785–805  $\text{cm}^{-1}$  region (D, I, N). Papanicolaou-stained images (E, J, O).

Nuclei are clearly visible in both CARS and SHG/TPF images, and not only the size of the nucleus is increasing with the progression of cancer but also the nucleus-to-cytoplasm ratio as mentioned before.

Raman microspectroscopy of these cells was acquired afterward, and the integrated intensity images in the C–H stretching region (2850–3050  $\text{cm}^{-1}$ ) of the normal, LSIL, and HSIL cells are depicted in panels C, H, and M, respectively. Raman spectra were collected from each cell with a pixel resolution of 500 nm providing thousands of spectra per cell. These spectra enable the label-free imaging of the cell based on its biochemical composition. This approach is different than that used in the previous Raman studies of Pap smears in which few spectra from cell nuclei were measured from single cells, lacking the imaging modality.<sup>25,26</sup> The integrated Raman intensity images near 790  $\text{cm}^{-1}$ , a DNA marker band,<sup>10,13</sup> show the nuclei of cells as indicated in panels D, I, and N. After Raman measurements, cells were stained with Papanicolaou staining and a cytologist (Z.H.) annotated these stained cells shown in panels E, J, and O as normal, LSIL, and HSIL cells, respectively. The cell nuclei displayed by label-free CARS (A, F, K), SHG/TPF (B, G, L), and Raman (D, I, N) images agree well with Papanicolaou-stained images (E, J, O). Therefore, label-free Raman, CARS, and SHG/TPF imaging can be used for the screening of cancerous cervical cells in Pap smears based on morphological features of the cell and the ratio of the nucleus to the cytoplasm. Because CARS/SHG/TPF imaging at one wavenumber is much faster than Raman spectral imaging, it allows fast screening of a large number of cells within a short time.

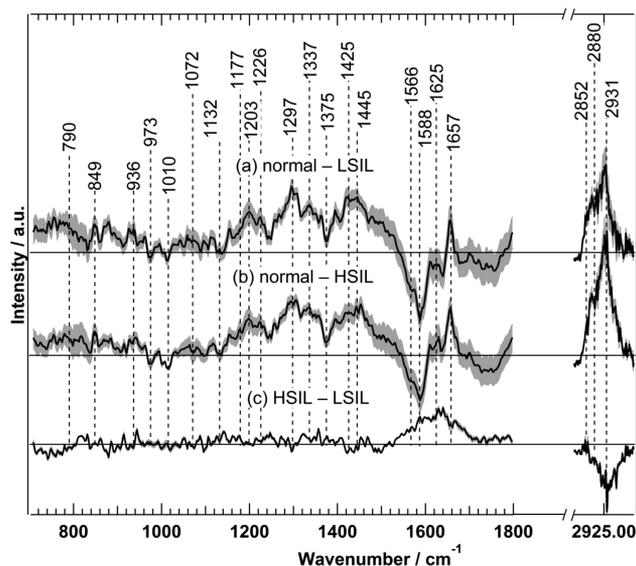
**Molecular Differences among Normal, LSIL, and HSIL Smears.** The average Raman spectra for normal (a), LSIL (b), and HSIL (c) cells from several Pap smears are shown in Figure 4. These spectra represent an integration of the biochemical composition of the cells in each smear type. Raman bands are observed at 790 (DNA: O–P–O; pyrimidine ring breathing mode), 853 (proteins: tyrosine and proline), 940 (proteins: proline and valine), 1008 (proteins: phenylalanine), 1130 (lipids), 1172 (nucleic acids), 1213 (proteins: tryptophan and phenylalanine), 1247 (proteins: amide III), 1308 (collagen and lipids), 1344 (collagen, nucleic acids, and tryptophan), 1452 ( $\text{CH}_2\text{CH}_3$  bending mode in proteins and lipids), 1591 (nucleic acids and phenylalanine), and 1657 (proteins: amide I) and in the C–H stretching



**Figure 4.** Average Raman spectra for normal (a), LSIL (b), and HSIL (c) cells from several Pap smears used to train DCNN. The shading represents the standard deviation. The spectra are offset for clarity.

region from 2850 to 2935  $\text{cm}^{-1}$  (lipids and proteins).<sup>16,26,44–46</sup> These are the main Raman bands accompanied by their tentative assignments. It is noted that a little blood contamination of the smears may contribute to the Raman band at  $\sim 1591$   $\text{cm}^{-1}$ , which originates from the heme of the hemoglobin.<sup>10,47</sup>

To detect the changes in the cellular biochemical composition as a result of cervical cancer progression, Raman difference spectra between normal and cancerous cervical cells were calculated and are shown in Figure 5. The Raman difference spectrum of normal–LSIL (a) reveals large spectral changes near 790, 849, 936, 973, 1010, 1177, 1203, 1226, 1297, 1337, 1375, 1425, 1445, 1566, 1588, 1625, 1657,



**Figure 5.** Comparison among the Raman difference spectra of normal–LSIL (a), normal–HSIL (b), and HSIL–LSIL (c). The shading represents the standard deviation.

2852, 2880, and 2931  $\text{cm}^{-1}$ . These bands originate from nucleic acids, polysaccharides, proteins, and lipids. Similar results were observed for the difference spectrum of normal–HSIL (b). The positive and negative bands in the spectra (a, b) indicate higher contributions from normal and cancerous (LSIL and HSIL) cells, respectively.

The positive features in the difference spectra (a, b) can be attributed to decreases in the levels of lipids, proteins, polysaccharides, and nucleic acids in the cancerous cervical cells.<sup>48,49</sup> These results reflect the underlying metabolic changes in cells upon the progression of cervical cancer. The difference spectrum of HSIL–LSIL (c) also shows positive bands near 1566 (nucleic acids), 1588 (nucleic acids and phenylalanine and/or blood), 1625 (proteins), and 1657  $\text{cm}^{-1}$  (proteins) as well as a negative band near 2850–2880  $\text{cm}^{-1}$  (mainly lipids).<sup>25,26,48</sup> These results suggest that the protein levels are higher in HSIL cells, whereas those of lipids are lower in comparison to LSIL cells.

**Deep Convolutional Neural Networks of CARS/SHG/TPF and Raman Images.** We have recently shown that DCNNs are a more powerful method for Raman microscopy than conventional chemometric approaches which typically ignore the morphology or texture of the sample.<sup>31</sup> The DCNNs of Raman spectral images discriminated between urocytitis and high-grade cancerous urothelial cells with high accuracy for the diagnosis of urothelial carcinoma using urine sediments.<sup>31</sup> In the present study, DCNNs were trained from scratch not only on the Raman spectral images of different smear types but also on both CARS and SHG/TPF images. To obtain Raman spectral images representative of a cell, the identification of a few Raman wavenumbers that are most informative for differentiating among normal, LSIL, and HSIL cells was determined using the MRMR approach as explained in the SI.<sup>50</sup> The selected wavenumbers based on MRMR should also be uncorrelated so that each selected wavenumber is expected to carry different/additional information for morphological feature-based classification. This approach was successfully applied to the Raman spectral images of urothelial cells.<sup>32,51</sup> For different smear types, six wavenumbers were selected: 1011, 1375, 1657, 2845, 3042, and 3049  $\text{cm}^{-1}$ . DCNNs were trained on the resulting image of normal, LSIL, and HSIL cells to identify these cells.

Tables 1 and S1 show the sensitivity, specificity, and accuracy for the classification based on per-patient cross-

**Table 1. Classification of Different Pap Smears by Applying DCNNs to Different Imaging Modalities**

classification (accuracy)	normal/LSIL (%)	normal/HSIL (%)	LSIL/HSIL (%)	normal/LSIL/HSIL (%)
Raman	100	100	100	94
CARS	100	100	85	90
SHG/TPF	95	100	90	83
CARS/SHG/TPF	100	100	85	90

validation with deep learning of the Raman spectral images. The differentiation between normal and cancerous cells was achieved with high accuracy. For instance, the binary classification results indicate that LSIL or HSIL cells can be differentiated from the normal cells with 100% accuracy. The differentiation between LSIL and HSIL is also achieved with 100% accuracy. These results are better than those obtained for the discrimination between normal and HSIL smears with

~95% accuracy using PLS-DA based on a few Raman spectra measured from single cell nuclei.<sup>26</sup>

Furthermore, DCNNs were also trained from scratch on both CARS and SHG/TPF images at one wavenumber (2935  $\text{cm}^{-1}$ ) and on their combination, and the classification results are shown in Tables 1 and S1. Accuracies of 100 and 95% were achieved to differentiate between the normal and cancerous cells in the case of CARS and SHG/TPF, respectively, similar to those obtained with Raman spectral images using six wavenumbers (Table 1). In addition, the accuracy for the differentiation among normal, LSIL, and HSIL in the case of CARS/SHG/TPF (90%) images is close to that observed for the Raman results (94%). Interestingly, using CARS or SHG/TPF images at one wavenumber (2935  $\text{cm}^{-1}$ ) produces accuracy similar to that observed for Raman images produced using six wavenumbers. This is most likely attributed to the higher spatial resolution of CARS compared to Raman imaging.<sup>11,29</sup> In addition, the pixel resolution in CARS and SHG/TPF (250 nm) images is higher than that in Raman images (500 nm), providing images with more morphotextural features.

**Potential of Raman and CARS/SHG/TPF Microscopy in Cervical Cancer Diagnosis.** Raman spectroscopy has been applied in the diagnosis of cervical cancer for almost two decades.<sup>52</sup> For instance, it has been used for the in vivo detection of cervical precancer with an accuracy of ~85% using a fiber optic probe.<sup>53–55</sup> Several studies have also shown that Raman spectroscopy can differentiate between normal and malignant cervical ex vivo tissues with high accuracy.<sup>52,56,57</sup> However, these biopsies were taken from patients invasively during colposcopy.

The previous and present studies have also revealed the potential of Raman spectroscopy as a noninvasive diagnostic tool using easily accessible Pap smears taken from patients, in contrast to the aforementioned invasive methods. For example, normal, LSIL, and HSIL smears were classified with ~95% accuracy using conventional chemometric methods based on a few Raman spectra from single cell nuclei.<sup>25,26</sup> In the present study, Raman spectral imaging of cells, representing the integral biochemical composition of cells, was used in combination with DCNNs to discriminate among different smear types with 100% accuracy.

One disadvantage of the conventional Raman spectroscopy is that the Raman intensity is weak. This leads to a longer accumulation time to improve the signal-to-noise ratio of the Raman signals. To overcome this issue, coherent Raman techniques such as CARS or SRS imaging at single wavenumbers that can be performed at a speed of up to the video rate have to be used.<sup>10,11,19,28–30</sup> In addition, fast multiplex/broadband CARS or SRS microspectroscopy, in which the acquired vibrational spectra would allow the detection of a larger number of biochemical molecules, can be used.<sup>20,58–60</sup> In the present study, we have established a new noninvasive diagnostic method based on Pap smears using a combination of CARS/SHG/TPF imaging and DCNNs. With this, cancerous cervical cells were discriminated from noncancerous cells with 100% accuracy.

Papanicolaou and HPV-DNA tests are the most commonly used noninvasive tests in clinical practice using Pap smears for the early detection of cervical cancer.<sup>3,8,9</sup> However, the HPV-DNA test is expensive and has lower specificity (~84%).<sup>9</sup> In addition, the Papanicolaou test suffers from a large variation in the sensitivity (50–96%).<sup>6,7</sup> There is also a lack of

cytopathologists in the labor market because training in cytopathology is a deep specialization in public health, where we have a general lack of skilled employees. Therefore, the present label-free CARS approach that automatically detects cervical cancer cells appears to be superior to the other conventional tests and it is also much faster than the Raman spectroscopic approach.

## CONCLUSIONS

CARS, SHG/TPF, Raman spectral imaging, and DCNNs were used for the label-free detection of cancerous cervical cells in Pap smears. The discrimination between normal and cancerous cells was achieved automatically with 100% accuracy. This classification is based on morphotextural information, which is obtained from CARS/SHG/TPF images at one wavenumber using leave-one-patient-out cross-validation. Similar accuracy was obtained when the classifications were performed on the basis of both morphotextural and spectral information obtained from Raman spectral images of the same Pap smears at six wavenumbers. These results demonstrate that high accuracy for discrimination among different Pap smears is still achieved even if the number of spectral images is reduced to one, as in the case of CARS/SHG/TPF imaging ( $2935\text{ cm}^{-1}$ ). Thus, CARS/SHG/TPF microscopy has the potential to be a fast, label-free imaging tool for screening large numbers of cells in Pap smears. Finally, the present study reports on a cohort at a single tertiary academic center. In the future, a multi-institution approach will be required to further validate the findings of this report in a prospective setting in LSIL and HSIL as well as healthy controls using fast CARS, SHG, and TPF imaging tools.

## ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: [10.1021/acs.analchem.9b03395](https://doi.org/10.1021/acs.analchem.9b03395).

Preparation of samples for spectroscopic measurements; Papanicolaou staining; data processing and multivariate analysis; spectral band selection; validation and testing; references; classification of different Pap smears applying DCNNs to different imaging modalities; and CARS and SHG/TPF images of a large area of a Pap smear (PDF)

## AUTHOR INFORMATION

### Corresponding Author

\*E-mail: [samir.el-mashtoly@rub.de](mailto:samir.el-mashtoly@rub.de).

### ORCID

Sascha D. Krauß: [0000-0001-6271-9940](https://orcid.org/0000-0001-6271-9940)

Hesham K. Yosef: [0000-0002-2385-0046](https://orcid.org/0000-0002-2385-0046)

Samir F. El-Mashtoly: [0000-0001-6087-8817](https://orcid.org/0000-0001-6087-8817)

### Author Contributions

<sup>||</sup>These authors contributed equally to this work.

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

We thank Tatjana Frick, Irina Chenska, and Raphael Roy for discussions. This research was supported by the Protein Research Unit Ruhr within Europe (PURE) funded by the Ministry of Innovation, Science and Research (MIWF) of

North-Rhine Westphalia, Germany (grant number 233-1.08.03.03-031-68079).

## REFERENCES

- (1) Ferlay, J.; Soerjomataram, I.; Dikshit, R.; Eser, S.; Mathers, C.; Rebelo, M.; Parkin, D. M.; Forman, D.; Bray, F. *Int. J. Cancer* **2015**, *136* (5), E359–E386.
- (2) Torre, L. A.; Bray, F.; Siegel, R. L.; Ferlay, J.; Lortet-Tieulent, J.; Jemal, A. *Ca-Cancer J. Clin.* **2015**, *65* (2), 87–108.
- (3) *Koss' Diagnostic Cytology and Its Histopathologic Bases*, 5th ed.; Koss, L. G., Melamed, M. R., Koss, L. G., Eds.; Lippincott Williams & Wilkins: Philadelphia, 2006.
- (4) Kurman, R. J.; Solomon, D. *The Bethesda System for Reporting Cervical/Vaginal Cytologic Diagnoses: Definitions, Criteria, and Explanatory Notes for Terminology and Specimen Adequacy*; Springer: New York, 1994.
- (5) Ikenberg, H.; Bergeron, C.; Schmidt, D.; Griesser, H.; Alameda, F.; Angeloni, C.; Bogers, J.; Dachez, R.; Denton, K.; Hariri, J.; et al. *JNCI J. Natl. Cancer Inst.* **2013**, *105* (20), 1550–1557.
- (6) Kitchener, H. C.; Blanks, R.; Dunn, G.; Gunn, L.; Desai, M.; Albrow, R.; Mather, J.; Rana, D. N.; Cubie, H.; Moore, C.; et al. *Lancet Oncol.* **2011**, *12* (1), 56–64.
- (7) McCrory, D. C.; Matchar, D. B.; Bastian, L.; Datta, S.; Hasselblad, V.; Hickey, J.; Myers, E.; Nanda, K. *Evid. Rep. Technol. Assess. (Summ.)* **1999**, No. 5, 1–6.
- (8) Walboomers, J. M.; Jacobs, M. V.; Manos, M. M.; Bosch, F. X.; Kummer, J. A.; Shah, K. V.; Snijders, P. J.; Peto, J.; Meijer, C. J.; Muñoz, N. *J. Pathol.* **1999**, *189* (1), 12–19.
- (9) Ronco, G.; Giorgi-Rossi, P.; Carozzi, F.; Confortini, M.; Dalla Palma, P.; Del Mistro, A.; Ghiringhello, B.; Girlando, S.; Gillio-Tos, A.; De Marco, L.; et al. *Lancet Oncol.* **2010**, *11* (3), 249–257.
- (10) Yosef, H. K.; Krauß, S. D.; Lehtonen, T.; Jütte, H.; Tannapfel, A.; Käfferlein, H. U.; Brüning, T.; Roghmann, F.; Noldus, J.; Mosig, A.; et al. *Anal. Chem.* **2017**, *89* (12), 6893–6899.
- (11) Krafft, C.; Schmitt, M.; Schie, I. W.; Cialla-May, D.; Matthäus, C.; Bocklitz, T.; Popp, J. *Angew. Chem., Int. Ed.* **2017**, *56* (16), 4392–4430.
- (12) Aljakouch, K.; Lehtonen, T.; Yosef, H. K.; Hammoud, M. K.; Alsaidi, W.; Kötting, C.; Mügge, C.; Kourist, R.; El-Mashtoly, S. F.; Gerwert, K. *Angew. Chem., Int. Ed.* **2018**, *57* (24), 7250–7254.
- (13) El-Mashtoly, S. F.; Yosef, H. K.; Petersen, D.; Mavarani, L.; Maghnouj, A.; Hahn, S.; Kötting, C.; Gerwert, K. *Anal. Chem.* **2015**, *87* (14), 7297–7304.
- (14) El-Mashtoly, S. F.; Petersen, D.; Yosef, H. K.; Mosig, A.; Reinacher-Schick, A.; Kötting, C.; Gerwert, K. *Analyst* **2014**, *139* (5), 1155–1161.
- (15) Kong, K.; Rowlands, C. J.; Varma, S.; Perkins, W.; Leach, I. H.; Koloydenko, A. A.; Williams, H. C.; Nottingher, I. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (38), 15189–15194.
- (16) Stone, N.; Kendall, C.; Smith, J.; Crow, P.; Barr, H. *Faraday Discuss.* **2004**, *126*, 141.
- (17) Byrne, H. J.; Baranska, M.; Puppels, G. J.; Stone, N.; Wood, B.; Gough, K. M.; Lasch, P.; Heraud, P.; Sulé-Suso, J.; Sockalingum, G. D. *Analyst* **2015**, *140* (7), 2066–2073.
- (18) Tipping, W. J.; Lee, M.; Serrels, A.; Brunton, V. G.; Hulme, A. N. *Chem. Soc. Rev.* **2016**, *45* (8), 2075–2089.
- (19) Petersen, D.; Mavarani, L.; Niedieker, D.; Freier, E.; Tannapfel, A.; Kötting, C.; Gerwert, K.; El-Mashtoly, S. F. *Analyst* **2017**, *142* (8), 1207–1215.
- (20) Ji, M.; Arbel, M.; Zhang, L.; Freudiger, C. W.; Hou, S. S.; Lin, D.; Yang, X.; Bacskai, B. J.; Xie, X. S. *Sci. Adv.* **2018**, *4* (11), eaat7715.
- (21) Michael, R.; Lenferink, A.; Vrensen, G. F. J. M.; Gelpi, E.; Barraquer, R. I.; Otto, C. *Sci. Rep.* **2017**, *7* (1), 15603.
- (22) Vargis, E.; Tang, Y.-W.; Khabele, D.; Mahadevan-Jansen, A. *Transl. Oncol.* **2012**, *5* (3), 172–179.
- (23) Rubina, S.; Amita, M.; Kedar, K. D.; Bharat, R.; Krishna, C. M. *Vib. Spectrosc.* **2013**, *68*, 115–121.

- (24) Bonnier, F.; Traynor, D.; Kearney, P.; Clarke, C.; Knief, P.; Martin, C.; O'Leary, J. J.; Byrne, H. J.; Lyng, F. *Anal. Methods* **2014**, *6* (19), 7831–7841.
- (25) Ramos, I. R.; Meade, A. D.; Ibrahim, O.; Byrne, H. J.; McMenamin, M.; McKenna, M.; Malkin, A.; Lyng, F. M. *Faraday Discuss.* **2016**, *187*, 187–198.
- (26) Duraipandian, S.; Traynor, D.; Kearney, P.; Martin, C.; O'Leary, J. J.; Lyng, F. M. Raman Spectroscopic Detection of High-Grade Cervical Cytology: Using Morphologically Normal Appearing Cells. *Sci. Rep.* **2018**, *8*, 15048.
- (27) Traynor, D.; Duraipandian, S.; Bhatia, R.; Cuschieri, K.; Martin, C. M.; O'Leary, J. J.; Lyng, F. M. *J. Biophotonics* **2019**, No. e201800377.
- (28) El-Mashtoly, S. F.; Niedieker, D.; Petersen, D.; Krauss, S. D.; Freier, E.; Maghnouj, A.; Mosig, A.; Hahn, S.; Kötting, C.; Gerwert, K. *Biophys. J.* **2014**, *106* (9), 1910–1920.
- (29) Cheng, J.-X.; Xie, X. S. *J. Phys. Chem. B* **2004**, *108* (3), 827–840.
- (30) Saar, B. G.; Freudiger, C. W.; Reichman, J.; Stanley, C. M.; Holtom, G. R.; Xie, X. S. *Science* **2010**, *330* (6009), 1368–1370.
- (31) Krauß, S. D.; Roy, R.; Yosef, H. K.; Lechtonen, T.; El-Mashtoly, S. F.; Gerwert, K.; Mosig, A. *J. Biophotonics* **2018**, *11* (10), No. e201800022.
- (32) Krauß, S. D.; Yosef, H. K.; Lechtonen, T.; Jütte, H.; Tannapfel, A.; Käßlerlein, H. U.; Brüning, T.; Roghmann, F.; Noldus, J.; El-Mashtoly, S. F.; et al. *J. Chemom.* **2018**, *32*, e2973.
- (33) Weng, S.; Xu, X.; Li, J.; Wong, S. T. C. *J. Biomed. Opt.* **2017**, *22* (10), 1.
- (34) Rodner, E.; Bocklitz, T.; von Eggeling, F.; Ernst, G.; Chernavskaia, O.; Popp, J.; Denzler, J.; Guntinas-Lichius, O. *Head Neck* **2018**, hed.25489.
- (35) Yosef, H. K.; Mavarani, L.; Maghnouj, A.; Hahn, S.; El-Mashtoly, S. F.; Gerwert, K. *Anal. Bioanal. Chem.* **2015**, *407* (27), 8321–8331.
- (36) Hammoud, M. K.; Yosef, H. K.; Lechtonen, T.; Aljakouch, K.; Schuler, M.; Alsaidi, W.; Daho, I.; Maghnouj, A.; Hahn, S.; El-Mashtoly, S. F.; et al. *Sci. Rep.* **2018**, *8* (1), 15278.
- (37) Yosef, H. K.; Frick, T.; Hammoud, M. K.; Maghnouj, A.; Hahn, S.; Gerwert, K.; El-Mashtoly, S. F. *Analyst* **2018**, *143* (24), 6069–6078.
- (38) Slater, D. N.; Rice, S.; Stewart, R.; Melling, S. E.; Hewer, E. M.; Smith, J. H. F. *Cytopathol. Off. J. Br. Soc. Clin. Cytol.* **2005**, *16* (4), 179–192.
- (39) Herbert, A. *Cytopathol. Off. J. Br. Soc. Clin. Cytol.* **2005**, *16* (4), 165–166.
- (40) Chung, C.-Y.; Potma, E. O. *Annu. Rev. Phys. Chem.* **2013**, *64* (1), 77–99.
- (41) Campagnola, P. J.; Dong, C.-Y. *Laser Photonics Rev.* **2011**, *5* (1), 13–26.
- (42) de Lanerolle, P.; Serebryanny, L. *Nat. Cell Biol.* **2011**, *13* (11), 1282–1288.
- (43) Huang, S.; Heikal, A. A.; Webb, W. W. *Biophys. J.* **2002**, *82* (5), 2811–2825.
- (44) Salzer, R.; Siesler, H. W. *Infrared and Raman Spectroscopic Imaging*; Wiley-VCH: Weinheim, 2009.
- (45) Rygula, A.; Majzner, K.; Marzec, K. M.; Kaczor, A.; Pilarczyk, M.; Baranska, M. *J. Raman Spectrosc.* **2013**, *44* (8), 1061–1076.
- (46) Konorov, S. O.; Schulze, H. G.; Atkins, C. G.; Piret, J. M.; Aparicio, S. A.; Turner, R. F. B.; Blades, M. W. *Anal. Chem.* **2011**, *83* (16), 6254–6258.
- (47) Diem, M. *Modern Vibrational Spectroscopy and Micro-Spectroscopy: Theory, Instrumentation, and Biomedical Applications*; John Wiley & Sons, Inc: Chichester, West Sussex, 2015.
- (48) Rehman, I. ur; Movasaghi, Z.; Rehman, S. *Vibrational Spectroscopy for Tissue Analysis*; Series in Medical Physics and Biomedical Engineering; CRC Press: Boca Raton, 2013.
- (49) Diem, M. G. *Vibrational Spectroscopy for Medical Diagnosis*; Diem, M. G., Ed.; Wiley & Sons, 2008.
- (50) Peng, H.; Long, F.; Ding, C. *IEEE Trans. Pattern Anal. Mach. Intell.* **2005**, *27* (8), 1226–1238.
- (51) Krauß, S. D.; Roy, R.; Yosef, H. K.; Lechtonen, T.; El-Mashtoly, S. F.; Gerwert, K.; Mosig, A. *J. Biophotonics* **2018**, *11* (10), No. e201800022.
- (52) Lyng, F. M.; Traynor, D.; Ramos, I. R. M.; Bonnier, F.; Byrne, H. J. *Anal. Bioanal. Chem.* **2015**, *407* (27), 8279–8289.
- (53) Mo, J.; Zheng, W.; Low, J. J. H.; Ng, J.; Ilancheran, A.; Huang, Z. *Anal. Chem.* **2009**, *81* (21), 8908–8915.
- (54) Duraipandian, S.; Zheng, W.; Ng, J.; Low, J. J. H.; Ilancheran, A.; Huang, Z. *Anal. Chem.* **2012**, *84* (14), 5913–5919.
- (55) Mahadevan-Jansen, A.; Mitchell, M. F.; Ramanujam, N.; Utzinger, U.; Richards-Kortum, R. *Photochem. Photobiol.* **1998**, *68* (3), 427–431.
- (56) Krishna, C. M.; Prathima, N. B.; Malini, R.; Vadhiraja, B. M.; Bhatt, R. A.; Fernandes, D. J.; Kushtagi, P.; Vidyasagar, M. S.; Kartha, V. B. *Vib. Spectrosc.* **2006**, *41* (1), 136–141.
- (57) Lyng, F. M.; Faoláin, E. O.; Conroy, J.; Meade, A. D.; Knief, P.; Duffy, B.; Hunter, M. B.; Byrne, J. M.; Kelehan, P.; Byrne, H. J. *Exp. Mol. Pathol.* **2007**, *82* (2), 121–129.
- (58) Shiozawa, M.; Shirai, M.; Izumisawa, J.; Tanabe, M.; Watanabe, K. *J. Biomed. Opt.* **2016**, *21* (7), No. 076004.
- (59) Rinia, H. A.; Bonn, M.; Müller, M. *J. Phys. Chem. B* **2006**, *110* (9), 4472–4479.
- (60) Kee, T. W.; Cicerone, M. T. *Opt. Lett.* **2004**, *29* (23), 2701–2703.