Biosensors and Bioelectronics 145 (2019) 111718

Contents lists available at ScienceDirect

Biosensors and Bioelectronics

journal homepage: http://www.elsevier.com/locate/bios

Label-free surface-enhanced Raman spectroscopy with artificial neural network technique for recognition photoinduced DNA damage

O. Guselnikova [a,b,1], A. Trelin [a,1], A. Skvortsova [a,1], P. Ulbrich [c], P. Postnikov [a,b], A. Pershina [b,d], D. Sykora [e], V. Svorcik [a], O. Lyutakov [a,b,*]

[a] Department of Solid State Engineering, University of Chemistry and Technology, 16628, Prague, Czech Republic
[b] Research School of Chemistry and Applied Biomedical Sciences, Tomsk Polytechnic University, 634049, Tomsk, Russian Federation
[c] Department of Biochemistry and Microbiology, University of Chemistry and Technology, 16628, Prague, Czech Republic
[d] Siberian State Medical University, 2, Moskovsky Trakt, 634050, Tomsk, Russia
[e] Department of Analytical Chemistry, University of Chemistry and Technology, 16628, Prague, Czech Republic

ARTICLE INFO

Keywords:
DNA
Photo-damage
SERS
Detection and recognition
Artificial neural network

ABSTRACT

Taking advantage of surface-enhanced Raman scattering (SERS) methodology with its unique ability to collect abundant intrinsic fingerprint information and noninvasive data acquisition we set up a SERS-based approach for recognition of physically induced DNA damage with further incorporation of artificial neural network (ANN). As a proof-of-concept application, we used the DNA molecules, where the one oligonucleotide (OND) was grafted to the plasmonic surface while complimentary OND was exposed to UV illumination with various exposure doses and further hybridized with the grafted counterpart. All SERS spectra of entrapped DNA were collected by several operators using the portable spectrometer, without any optimization of measurements procedure (e.g., optimization of acquisition time, laser intensity, finding of optimal place on substrate, manual baseline correction, etc.) which usually takes a significant amount of operator's time. The SERS spectra were employed as input data for ANN training, and the performance of the system was verified by predicting the class labels for SERS validation data, using a spectra dataset, which has not been involved in the training process. During that phase, accuracy higher than 98% was achieved with a level of confidence exceeding 95%. It should be noted that utilization of the proposed functional-SERS/ANN approach allows identifying even the minor DNA damage, almost invisible by control measurements, performed with common analytical procedures. Moreover, we introduce the advanced ANN design, which allows not only classifying the samples but also providing the ANN analysis feedback, which associates the spectral changes and chemical transformations of DNA structure.

1. Introduction

Surface-enhanced Raman scattering (SERS) is considered as the most promising analytical tool for (bio)sensing due to the reliable opportunities for the simple, specific and non-destructive detection of various (bio)molecules (nucleic acids, lipids, peptides, and proteins) (Shan et al., 2018; Bruzas et al., 2018; Puebla and Marzán, 2012), as well as for *in vivo* sensing, diagnostics and therapy (Vendrell et al., 2013; Lane et al., 2018; Chao et al., 2016). Despite the potential of SERS for (bio)analytes monitoring, certain critical challenges remain unsolved for wide-world applications: (i) poor reproducibility of SERS and (ii) interference of SERS signal due to the complex structure of biological molecules (Yuan et al., 2017).

The solution of the above-stated challenges was proposed through the surface modification of plasmon-active substrates for the formation of sites for selective analyte entrapping (Szlag et al., 2018). The most common methods for functionalization include grafting of antibodies, enzymes, aptamers, and molecular imprinted polymers (Jazayeri et al., 2016; Ochsenkühn and Campbell, 2010; Cottat et al., 2015; Ashley et al., 2017). This approach can increase the SERS reproducibility and selectivity; however, the problem connected with spectral interference remains unsolved and leads to the complicated interpretation of spectral data. Moreover, the surface grafting by selective "bio-ligands" gives only limited information about target/molecules, and it is unable to register

* Corresponding author. Department of Solid State Engineering, University of Chemistry and Technology, 16628, Prague, Czech Republic.
E-mail address: lyutakoo@vscht.cz (O. Lyutakov).
[1] These authors contributed equally.

https://doi.org/10.1016/j.bios.2019.111718
Received 27 June 2019; Received in revised form 13 September 2019; Accepted 19 September 2019
Available online 20 September 2019

other potentially important changes in (bio)sample composition.

To address the issue connected with complicated spectra analysis and interpretation, mathematical data processing of massive amount of SERS spectra was recently proposed (Santos et al., 2017; Etchegoin et al., 2007; Kahraman et al., 2017). In these regards, different methods can be employed, including the commonly used principal component analysis, 2D correlation, partial least squares, linear discriminant analysis, and support vector machine (Moore et al., 2005; Zhang et al., 2010; de Groot et al., 2001; Shin et al., 2018; Hoonejani et al., 2015; Gromski et al., 2015; Li et al., 2013). However, in contrast to these linear spectral analysis methods, the artificial neural networks (ANN) have proved their superior capability (Naik et al., 2016; Egrioglu et al., 2015). The main advantages of ANN over classical approach are the ability to find nonlinear dependencies, what is more appropriate for complex (bio) samples, which do not obey the linear laws, providing the unique possibility to solve urgent challenges in the field of analytical chemistry (Chan et al., 2016a, 2016b; Lake et al., 2015; Kermany et al., 2018; Segler et al., 2018).

Starting from the utilization of ANN for Raman-based detection of glucose content and the discrimination of sugar additives in honey, the ANN is lately shifted to SERS application (Özbalci et al., 2013; Amato et al., 2013). Recent examples of SERS sensing followed by ANN implementation include detection of small organic molecules in human urine, identification of tumor suppressor genes, classification of aqueous pollen extracts and drug derivatives (Kasera et al., 2015; Shi et al., 2018; Seifert et al., 2016; Alharbi et al., 2015). However, up to date, ANN was used only for differentiation or classification of samples to specific groups and no attempt for obtaining of information about significant and meaningful spectral changes associated with the composition of (bio)analyte (in another word, the spectra interpretation feedback, obtaining using the ANN interpretation of SERS data) has not been reported.

Here, for the first time, we demonstrate that functionalized SERS substrate in combination with ANN is suitable for the collection of information about the molecular structure of analyte and simultaneous interpretation of spectra in the automated regime. As an explicit example, we realized the detection of UV-induced damage of oligonucleotide (OND) using surface plasmon-polariton supported gold gratings grafted by probe OND. Specific spectral changes were highlighted by using a novel feature selection method, binary stochastic filtering (Trelin and Prochazka, 2019), which allows finding of specific spectral changes corresponding to chemical transformations of OND. The proposed approach allows detecting the UV-induced damage in OND structure, which can be hardly measured by alternative methods (Goto et al., 2015; Peccia and Hernandez, 2002).

2. Experimental

2.1. Materials

Acetic acid (reagent grade, ≥99%), diethyl ether, deionized water, 4-aminobenzoic acid (≥99.0%), p-toluenesulfonic acid monohydrate (ACS reagent, ≥98.5%), high-purity water (EMD-Millipore) were purchased from Sigma-Aldrich. The oligonucleotides (ONDs) were obtained from Biosynthesis (Russia). The OND sequences used in this study are presented in Table S1.

2.2. Preparation of samples

Grating preparation. Briefly, deposited by spin-coating Su-8 thin films were patterned by excimer laser (3500 laser pulses, 9 mJ cm^{-2} laser fluencies). The Au thin film was deposited onto a patterned polymer surface by vacuum sputtering (discharge power of 7.5 W, sputtering time 200 s, resulted in thickness 25 nm) (Guselnikova et al., 2017a). The working SERS active area was $1 \times 0.5\ cm^2$.

Au grating surface modification. The obtained gratings were spontaneously modified by soaking in 1 mM freshly prepared aqueous solution of 4-carboxybenzenediazonium tosylate (ADT-COOH), previously prepared according to the procedure described in (Guselnikova et al., 2019a) for 20 min. After modification Au substrates were rinsed under sonication sequentially with water, methanol, and acetone for 10 min and dried in a desiccator.

Probe OND grafting. 0.5 nmol of OND 1 (5-$NH_2(CH)_6$–5′-cg CCAA-TAC GACCAAATCCG-3′) was grafted to the surface through EDC/sulfo-NHS coupling: 4 mL of 0.5 mM solution of EDC (MES buffer pH 4.5) and 1.25 mM Sulfo–NHS–SO_3Na (MES buffer pH 4.5) were mixed and modified gold grating was immersed in EDC/sulfo-NHS solution for 20 min. Afterward, the sample was washed by MES buffer and ultrapure water. For the covalent grafting between probe OND and sulfo-NHS ester on surface, 100 μL of probe OND (0.5 nmol) was dissolved in 5 mL of PBS buffer (pH = 7.2) and activated Au grating was immersed in this solution for 2 h, after coupling modified grating was washed by PBS buffer, ultrapure water and dried.

UV irradiation and hybridization. Solutions (3 mL) of complementary OND at concentration 5×10^{-9} M (annealing buffer) were placed in a cuvette ($S = 1\ cm^2$) and irradiated by LED-lamp light at wavelength 375 nm for 1.5, 2, 3, 7, 10, 12, 20, 24 h. After this solution was diluted by annealing buffer to the concentration 1×10^{-9} M for the following hybridization. For the hybridization experiments annealing buffer (10 mM Tris, pH 7.5, 50 mM NaCl, 1 mM EDTA) was used. Briefly, the Au grating grafted by probe OND in the irradiated solution was heated to 95 °C and gradually cooled to 25 °C. Afterward, the sample was washed by annealing buffer, ultra-pure water and dried at ambient conditions. The samples were further grouped according to the expected UV-induced OND damage into four groups: non-damaged (labeled as 0), slightly damaged (labeled as 1), moderately damaged (labeled as 2), and significantly damaged (labeled as 3). Detailed samples referring, grouping, and description is presented in Table 1.

Table 1
Samples referring and description.

Samples designation	Time of UV illumination, h	Description	Samples label, according the expected UV-induced OND damage
p-OND	0	Probe OND, grafted to gold grating surface	-
OND-0	0	Fully complementaryOND	0 (not damaged)
OND-1.5	1.5	UV damaged complementary further coupled with p-OND	1 (slightly damaged)
OND-2	2	UV damaged complementary further coupled with p-OND	
OND-3	3	UV damaged complementary further coupled with p-OND	
OND-7	7	UV damaged complementary further coupled with p-OND	2 (moderately damaged)
OND-10	10	UV damaged complementary further coupled with p-OND	
OND-12	12	UV damaged complementary further coupled with p-OND	
OND-20	20	UV damaged complementary further coupled with p-OND	3 (significantly damaged)
OND-24	24	UV damaged complementary further coupled with p-OND	

2.3. Measurement techniques

The X-ray photoelectron spectroscopy (XPS) was performed using an Omicron Nanotechnology ESCAProbeP spectrometer fitted with monochromated Al K Alpha X-ray source working at 1486.6 eV (pass energy - 30 eV, analyzed area -2x3 mm^2). Concentrations of elements were calculated in at. % using the manufacturer's sensitivity factors. For characterization of the sample surface and nanomechanical mapping before and after the surface modification, the peak force AFM technique with the Icon (Brucker) microscope was used. UV/Vis absorption spectra were measured by using a Lambda 25 UV/Vis/NIR Spectrometer (PerkinElmer, USA) at a scanning rate of 480 nm min^{-1}.

Raman scattering was measured on a portable ProRaman-L spectrometer (Laser power 35 mW) Raman spectrometer with 785 nm excitation wavelengths. Each spectrum was averaged from 30 measurements, each of them with 3 s accumulation time. The spectra were measured by four operators (Ph. D. students with different level of experience with Raman spectrometer). In general, 1800 spectra were collected from 60 SERS substrates. The measurements were carried out on the randomly chosen places on the active area of the sample $1 \times 0.5\ cm^2$.

C***ontrol measurements of UV-induced OND damage***. Solutions of Probe OND and illuminated were analyzed by FTIR-ATR, Matrix-assisted laser desorption/ionization, agarose electrophoresis, and qPCR for high resolution melting analysis. For detailed experimental procedures and related discussion, see Supporting Information (SI).

2.4. SERS data analysis

Data preprocessing. Before feeding in the neural network, the collected spectral data was initially processed, using the developed automatic algorithm, which also eliminates the potential learning/validation uncertainty due to spectral background disturbance. The data preprocessing includes following subsequent steps:(i) - background subtraction (according to (Eilers and Boelens, 2005)); (ii) - noise suppression (using the mathematical procedures, described in (Savitzky, 1951; Meurer et al., 2017a)); (iii) - MinMax normalization; (iv) – labeling. A detailed description of data preprocessing is given in SI.

ANN architecture. Obtained spectra, labeled according to Table 1, were processed with deep convolutional neural network (CNN), a special kind of ANN. The classification problem was formulated, having given spectrum as an input and the corresponding degree of damage as an output (Meurer et al., 2017b). The created network consists of the feature extraction module, composed from a few convolutional layers and multilayer perceptron made from fully-connected layers. It was found experimentally that convolutional layers improve the model accuracy in comparison to classical multilayer perceptrons. Proposed CNN model consists of 4 convolutional layers with an increasing number of kernels (8, 16, 32, 64) and Leaky ReLU activation with negative slope coefficient $\alpha = 0.3$, alternated with max pooling layers which increase information density by performing the two-fold reduction of the previous layers outputs dimensionality. Each pooling layer is followed with batch normalization layer, which stabilizes the training process (Ioffe and Szegedy, 2015). Those set of layers form feature extraction module, the output of which is flattened and then passed to the fully-connected classifier. This module is built from 5 fully-connected layers with ReLU activation and output layer, having probabilistic softmax activation. Consequently, categorical cross-entropy was used as a loss function, which is the standard loss in classification problem. Dropout (rate = 0.5) was added to the fully connected layers to improve generalization ability (Srivastava et al., 2014). The model structure is visualized in Fig. S1. CNN was implemented in Python programming language with Keras framework (Abadi et al., 2016) using TensorFlow library (Chollet, 2017) as back-end.

ANN Training and validation. The model was trained by Adam optimizer with the following parameters, recommended in the original article (Kingma and Ba, 2014) - learning rate equal to 5×10^{-4} and zero learning rate decay. The model was trained in total for 500 epochs, which finally led to training loss equal to 3.53×10^{-4} and validation loss equal to 0.0217. Validation was performed by comparing CNN predicted labels for validation dataset with actual ones after each training epoch. Such an on-line approach allows estimation of CNN generalization ability during weights evolution, and initial convergence was achieved in less than 200 epochs (Fig. S2).

3. Results and discussion

3.1. Experimental concept and SERS grating creation

In this work, we propose a simple experimental route for label-free detection of DNA damage in an extremely fast and reliable way. Our approach is based on the combination of unique advantages provided by SERS and ANN combination. It is well known that the SERS method is favorable for detection of a molecular fingerprint with very low detection limit (Guselnikova et al., 2017b; Chen et al., 2018), but it is more suitable for recognition of low molecular weight analytes. The complications caused by the presence of interfering Raman signal from different compounds in a sample are a major drawback for the application of SERS for real biomedical samples, containing biomolecules with high molecular weight. At this moment utilization of ANN seems to be a promising alternative, allowing finding even exiguous spectral features in highly noised Raman signal. So, a combination of the high sensitivity of SERS and ability of ANN to discriminate minimal key spectral features despite the highly intensive background can become an upcoming alternative for DNA analysis.

Schematic description of the proposed experimental concept is given in Fig. 1. As plasmonic support, we utilized the periodic Au grating surface, created by excimer laser patterning. The single probe-OND (p-OND) (Fig. 1) was grafted to the Au grating surface using the two steps approach, including surface covalent grafting by 4-carboxyphenylene moieties and their further coupling with amino-terminated p-OND. Simultaneously complementary OND (hereinafter referred as OND-x, where the x is a time of UV illumination) was exposed to UV-A light, which causes the damage of OND-x structure and presents the main endanger for human health (Sinha and Häder, 2002). The spectral range of used UV light is presented in Fig. S3 and the degree of OND damage were regulated using the time of illumination. In the next step, we applied the hybridization of OND-x with grafted p-OND and measured the raw SERS data. The measured SERS spectra were then used as a data set for ANN learning and testing.

According to the previous report (Guselnikova et al., 2017a, 2019a), we utilized the highly regular Au grating surface (Fig. 2A), which provide the homogeneous distribution of plasmon intensity along the SERS substrate and minimize the uncertainty, related to the random distribution of hot spots, typical for many SERS measurements (Guselnikova et al., 2017b, 2019a, 2019b; Eilers and Boelens, 2005). The grating parameters were tuned to obtain the SPP related plasmon absorption band near the 785 nm, in consistence with further used SERS excitation wavelength (Fig. 2B). Utilization of SPP-based SERS approach (instead of more known LSP – localized surface plasmon resonance) allows the more smoother decay of plasmon evanescent wave above the grating surface, making the analysis of relatively large (bio)molecules possible. Consequently, the SERS enhancement factor (EF) of the created plasmonic substrate reaches 10^6 (Fig. S4, details of the EF calculation are given in SI). The numerical simulation of the distribution of plasmon intensity under the illumination with 785 nm is presented in Fig. 2C. The uniformity of the substrate was evaluated experimentally by analyzing SERS spectra obtained from different randomly selected points (five spectra were taken from each point) on the grating surface with deposited R6G as a model SERS analyte (Fig. S5). Quantitatively, the measured relative standard deviation (RSD) value of SERS intensity between the different points on several substrates was found to be as low

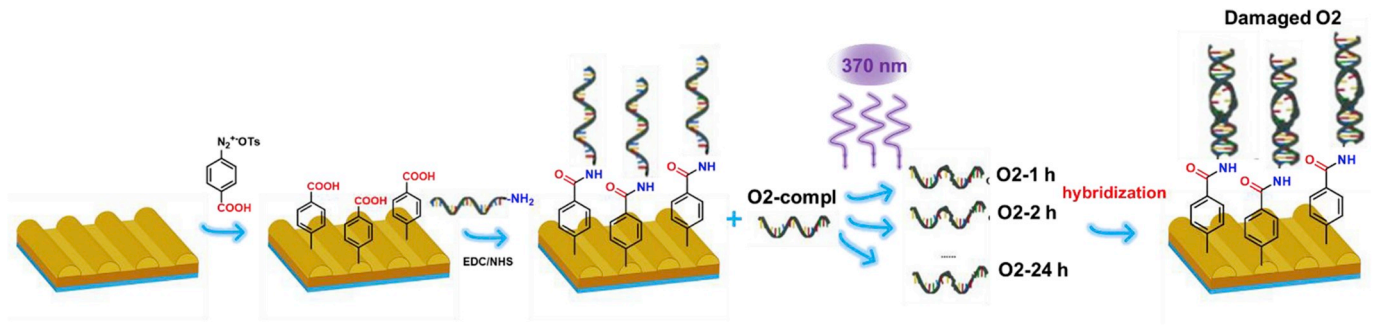


Fig. 1. Schematic of proposed experimental concept.

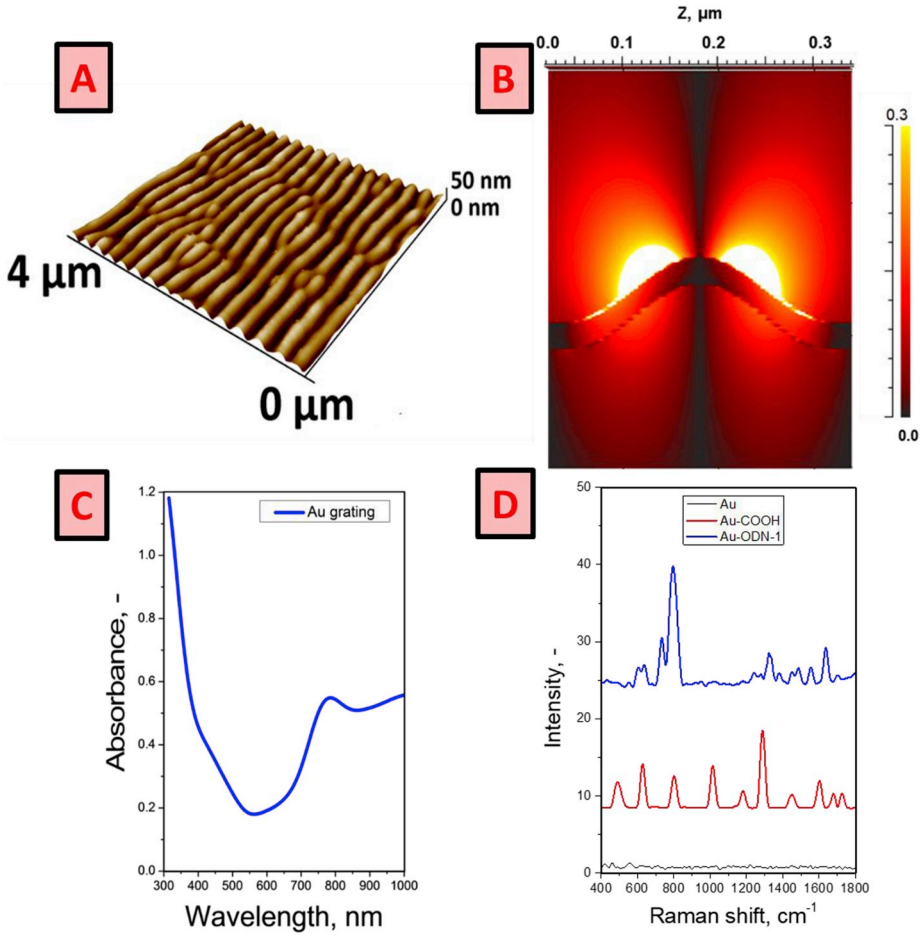


Fig. 2. Characterization of gold grating properties and functionality: (A) - AFM measured surface morphology of Au grating; (B) - results of numerical simulation of electric field distribution on Au grating surface under the illumination with 785 nm; (C) – UV–Vis spectra of Au grating surface measured with light polarization, perpendicular to grating orientation; (D) - Raman spectra, measured at different stages of probe OND grafting. . (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

as 5.5%.

The grafting of probe OND to the Au grating surface was confirmed using the SERS and XPS techniques. XPS results (Table S2) shows the apparent increase of carbon and oxygen surface concentration as well as attenuation of Au 4f signal and appearance of N1s peaks after the surface decoration with probe OND. In turn, characteristic spectral features (Fig. 2D) after the grafting of ADT-COOH and p-OND confirm the successful grating surface functionalization with p-OND (detailed peak affiliation is given in Table S3). So, both methods (SERS and XPS) convincingly confirm the surface functionalization with probe OND, which further serves as reactive moieties for entrapping of OND-x from solution. In addition, XPS based calculation allows to approximately estimate the p-OND grafting density which was founded to be $1.08*10^{-14}$ mol/cm^2 (detailed description is given in Supplementary Information).

3.2. UV-induced OND damage

In the next step, the UV illumination of OND was used to induce the chemical transformations resulting in the damage of DNA damage. The control measurements were performed using the FTIR, MALDI, electrophoresis, and temperature resolved DNA dehybridization. Results of IR (Fig. S6 and related discussion in SI) indicate the gradual changes of IR spectra, which becomes more apparent with increasing of UV illumination time. In turn, MALDI results show negligible changes of OND-x mass under the slight or moderate illumination time. Only in the case of prolonged UV triggering, the changes of MALDI spectra were observed on the border of distinguishability of this method (Fig. S7).

Further control experiments include electrophoresis test of hybridized p-OND and OND-x, but in this case, no significant signal of DNA damage was observed (compare to control, uncomplimentary OND-

Fig. S8A and related discussion). Finally, advanced control experiments of p-OND and OND-x hybridization were performed using the temperature-resolved luminescence measurements (Fig. S8B and related discussion). After the statistical evaluation of obtained results, the apparent effect of UV illumination became evident, but mainly in the case of samples with a long UV-illumination time. So, the results of control measurements can be summarized as follow: (i) UV illumination leads to chemical changes in the OND structures (FTIR), (ii) such chemical changes are not accompanied by significant changing of molecular mass (MALDI), (iii) observed changes affect the DNA hybridization (i.e., potential DNA mutation), which can be detected only in the case of prolonged UV-illumination time and with utilization of time-consuming techniques.

3.3. SERS measurements and ANN

After the control experiments, we performed the hybridization of complementary or UV damaged ONDs with grafted probe OND and the SERS measurements. It should be noted that the grating morphology cannot affect the orientation of grafted p-OND or hybridized DNA, due to mismatch in the surface curvature (amplitude – 27 nm, periodicity – 350 nm) and DNA sizes (approximately 6.5 nm). We tried to maximally complicate the initial ANN task with the aim to maximally highlight its capability in the final. In particular, we avoided the procedure for optimization of SERS measurements procedure and did not apply any baseline correction. Moreover, the spectra were independently collected by several (4) operators to introduce additionally the „human factor error“ (Brougham et al., 2011). The raw SERS data for different irradiation time is presented in Fig. 3 for different sets, obtained after the hybridization of OND-x with surface grafted p-OND. As can be expected, the measured at such unpretentious conditions SERS spectra represents sophisticated interference pattern. It is almost impossible to recognize separated, DNA-base related vibration bands as well as recognize the difference between the sets of SERS spectra (i.e., to determine the time of UV illumination).

In the next step, we applied the raw datasets of SERS data for ANN learning and validation. Among the different ANN design, the so-called Convolutional Neural Network (CNN) has been used for data processing. This architecture is proven to work efficiently in the background presence or even to unwind the signal-baseline interference, (the convolution operation may be thought of as a form of digital filtering (Liu et al., 2017)). In addition, the present ANN design is featured by analysis feedback, which associates the spectral changes and related biochemical changes in DNA structure. It is well known that identification of ANN response to changes in specific input features (in our case – determination of spectral region, responsible for recognition of the DNA damage) is a great challenge. ANN operates with data of extreme dimensionalities, and it is usually considered to be a “black box”, which is able of efficient solving of specific kind of problems, but the process how the solution is done cannot be analyzed. To overcome this problem, and introduce the new, attractive possibility in the ANN-based SERS spectra interpretation, so-called Binary Stochastic Filtering (BSF) was introduced, which allows highlighting of SERS spectra features responsible for DNA damage identification (Trelin and Prochazka, 2019). BSF estimates the importance of every input feature (value of each pixel in the spectrum) thus recognizes regions of interest in the original spectra

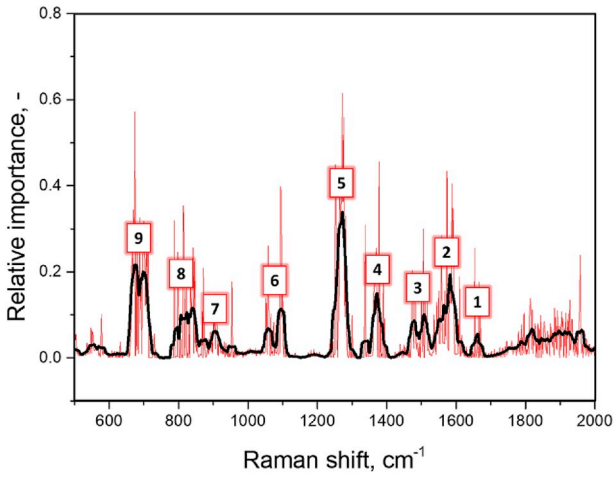


Fig. 4. Raman areas of interest, determined by ANN during the SERS spectra evaluation, responsible for discrimination of UV damages degree.

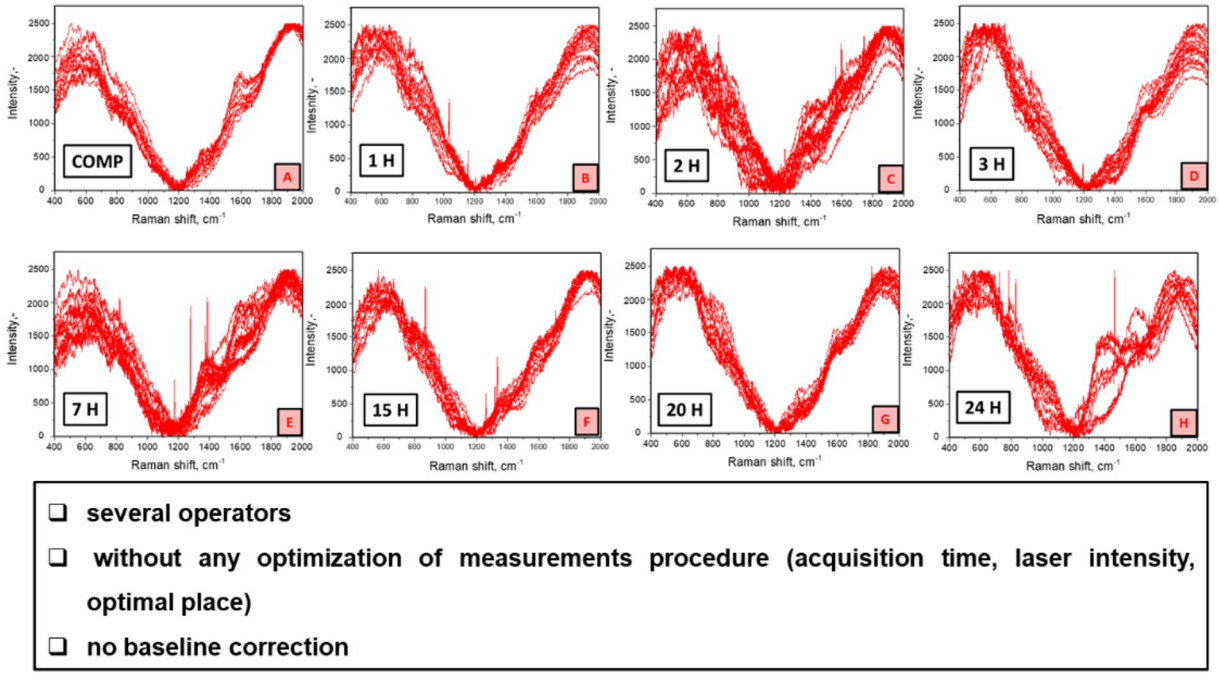


Fig. 3. Set of raw SERS spectra, measured after the interaction of grafted probe OND with pristine UV illuminated for different time intervals complementary OND.

(i.e. spectral regions responsible for the estimation of DNA damage). These regions are presented in Fig. 4 (and Table 2) and further interpreted from the spectroscopic point of view. Generally, observed spectral pattern perfectly corresponds with typical DNA spectrum and include cytosine (C), guanine (G), adenine (A), or thymine (T) Raman bands. The major part of "peaks of interest" (the detailed assignation of the area of interest is given in Table 2) can be attributed to the nitrogenous ring stretching and bending, and vibration of amino groups, bringing us to the conclusion that the UV illumination results in chemical transformations of heterocyclic fragments of DNA. With a high probability, these transformations result in formation of cyclobutane pyrimidine dimer (CPD) and (6–4) photoproducts, which are typical products of UV induced damage for double and single-stranded oligonucleotides (Fornace et al., 1976; Doddridge et al., 1998). Also, we found that the peaks assigned to symmetrical stretching of PO^{2-} take part in the classification of SERS spectra. We suppose that the contribution of PO^{2-} peaks can be associated with changes in the chemical environment of phosphate groups, proved by the effect of C═O stretching, due to the CPD formation or sugar part oxidation (Doddridge et al., 1998). The perfect coincidence of ANN estimated spectral features with typical DNA SERS fingerprints convincingly proofs that the created network can highlight the most informative spectral regions that have real physical meaning. Thus, the developed method allows finding specific spectral and even chemical patterns, which accompany DNA damage.

3.4. SERS/ANN based DNA damage classification

The predictive power of ANN (in terms of neither, slight, moderate, or strong UV-induced OND damage – see Table 1) was studied with validation dataset and obtained results of ANN training are presented in Fig. 5. For data evaluation, the training part of the spectral dataset (i.e. the data with defined UV-illumination time, but previously unknown for ANN) was used as the input to the ANN classifier. Results of ANN-based spectra evaluation provide their class, e.g. the recognition of a sample was performed as it belonging to the specific class (for example, to the class 2, which corresponds to [0, 0, 1, 0] output). Obtained outputs were compared with known times of illumination, corresponded to different OND damage and stacked forming a 2D array, which was visualized as an image (Fig. 5), where each row corresponds to class probabilities for each sample. So, in Fig. 5 the rows of the image correspond to the probabilities (encoded with colors) that the input spectra belong to the given class. The real classes are depicted with curly braces on the left classification map; the corresponding braces on the right are omitted for obvious reasons. It is visible that only single datum (highlighted with a red arrow – Fig. 5A) was incorrectly classified, this fact indicating maximum probability of correct classification in the other classes. Black arrows highlight SERS data that were classified correctly (in terms of DNA damage), but with a lower degree of confidence (Fig. 5A). So, classification and correspondence maps demonstrate predictive accuracy of the present model higher than 98%, and the fact that from the 58 validation spectra, only one was predicted incorrectly. The probable reason for this inconsistency is spectral collection errors or random source of noise, such as cosmic spikes. Overall, this experiment demonstrates that the model is able of highly accurate classification of SERS spectral data and precise estimation of DNA damage (Fig. 5B and C). Especial and strong attention also deserves the fact that the proposed methods give the opportunity to recognize even negligible DNA damage, which can hardly be detected by common analytical route (see control experiments section). This ability is convincingly proved by perfect classification of previously "unknown" spectra labeled as 1 (low degree of OND damage, undetectable by MALDI or electrophoresis) or 2 (moderate degree of DNA damage).

Table 2
Affiliation of SERS peak from the areas of interest for the discrimination of UV damages oligonucleotides.

Number of peaks	Frequency, cm^{-1}	Affiliation
1	1660	Str C═O, sciss NH_2
2	1585	str C═C, str N–C
3	1480, 1510	str N–C
4	1378	bend C–H, str C–N–C
5	1270	str N–C
6	1056, 1098	Symm str PO_2^-
7	904	str C–C
8	802–843	Symmstr O–P–O, bend N–H, C–H
9	674–703	Ring breath, bend C–C, N–C

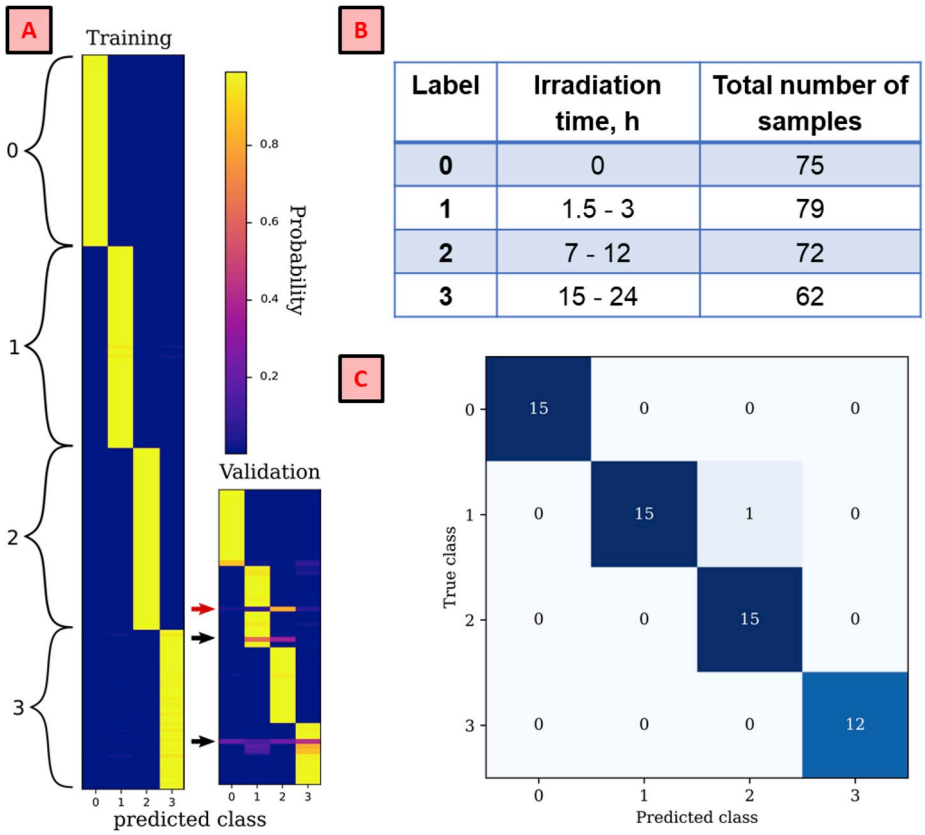
Considering the applicability of SERS biosensor in real conditions, the one of the ultimate goals is the preparation of stable functional SERS substrate. For the test of proposed substrates stability, the p-OND grafted Au gratings were kept under ambient conditions (room temperature and laboratory lightning, in an air atmosphere) for several months – Fig. S9A (or continually measured during the first month of storage – Fig. S9B). From the Fig. SXXB it can clearly be seen that the samples provide the highly stable SERS signal for at least 30 day, while Fig. S9A indicates the slight decrease (by ca 10%) of measured SERS signal after the 3 months of previous samples storage. So, proposed functional SERS substrate can considered as relatively stable (their stability can be further increased by improving of storage conditions – inert atmosphere, dark, etc.).

3.5. Advantages of proposed route and future perspective

Finally, it should also be noted that the proposed route for identification of UV induced DNA damage is significantly favorable for its potential portability, fast processing times (after the ANN training, the time is limited only by spectra collection, which usually takes less than 30 min in the case of manual collection and could be improved by process automation), simplicity and low cost. The developed approach does not require the more complicated and time-consuming chromatographic separation and mass spectrometry techniques. The developed method can be automated with ease for high throughput analyses and could have a tremendous impact on a broad range of applications. Additionally, the proposed route allows the correct evaluation of the degree of DNA damage despite demonstrated operator-induced sources of variation. It should be also noted that proposed algorithm of DNA analysis is not restricted to the damage induced by UV, but also can be potentially applied for the discrimination of mismatched DNA, or for screening responsible for development of relevant mutation (for example, detection of antimicrobial resistance genes (Kuo et al., 2015)). Our further work will be focused on the analysis of long DNA molecules. In this case the two main questions will be further solved: (i) the maximal (or optimal) lengths of DNA fragments (or OND), which can be discriminated by SERS/ANN combination, (ii) the ability to perform DNA fragmentation, before analysis and applicability of this approach for the analysis of relevant DNA molecules. In particular, the future works will be focused on expanding of our approach for the analysis of longer DNA structures (like New Delhi Metallo-β-Lactamase (NDM)-coding gene and TEM -Lactamase (blaTEM) Genes) with the utilization of specific restriction enzyme, which have recognition signs and separate the sights of our interest.

4. Conclusion

In this study, a SERS spectroscopy in combination with ANN is proposed for the identification and quantification of light-induced DNA damage. The experimental set-up includes the plasmon-active Au grating surface, which provides the high SERS enhancement factor and homogeneous distribution of SERS intensity. The grating surface was grafted with single OND for entrapping of complementary or UV-damaged OND from solution. SERS spectra collection was performed without any optimization of measurements procedure and obtained



Label	Irradiation time, h	Total number of samples
0	0	75
1	1.5 - 3	79
2	7 - 12	72
3	15 - 24	62

Fig. 5. (A) - classification maps, produced with trained model for training SERS spectra dataset (left) and testing SERS spectra dataset (right); (B) – table of labels corresponding to the "zero", "low", "moderate", and "high" irradiation time; (C) – relation between true and predicted classes for tested SERS spectra of UV-light damaged OND.

spectra were processed by ANN. It was successfully demonstrated that proposed experimental concept is able to recognize even invisible by common analytical route DNA damage. Additionally, the developed method allows finding specific spectral and even chemical patterns, which accompany DNA damage. Proposed route for identification of DNA damage is favorable for its sensitivity, portability, processing times, simplicity, low cost and possibility to be automated with ease for high throughput analyses.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

O. Guselnikova: Conceptualization, Investigation, Methodology, Visualization, Writing - review & editing. **A. Trelin:** Software, Investigation. **A. Skvortsova:** Investigation, Methodology. **P. Ulbrich:** Methodology, Investigation. **P. Postnikov:** Data curation, Validation, Formal analysis, Funding acquisition. **A. Pershina:** Data curation, Validation, Writing - review & editing, Writing - original draft. **D. Sykora:** Methodology, Investigation. **V. Svorcik:** Conceptualization, Writing - review & editing. **O. Lyutakov:** Supervision, Conceptualization, Writing - original draft.

Acknowledgment

This work was supported by the GACR under the project 18-26170S and Tomsk Polytechnic University (project VIU-RSCABS-68/2019).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bios.2019.111718.

References

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D.G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., Wicke, M., Yu, Y., Zheng, X., 2016. TensorFlow: A System for Large-Scale Machine Learning, 12th Symposium on Operating Systems Design and Implementation, pp. 265–283. https://www.usenix.org/conference/osdi16/technical-sessions/presentation/abadi.

Alharbi, O., Xu, Y., Goodacre, R., 2015. Simultaneous multiplexed quantification of caffeine and its major metabolites theobromine and paraxanthine using surface-enhanced Raman scattering. Anal. Bioanal. Chem. 407, 8253–8261. https://doi.org/10.1007/s00216-015-9004-8.

Amato, F., López, A., Peña-Méndez, E.M., Vanhara, P., Hampl, A., Havel, J., 2013. Artificial neural networks in medical diagnosis. J. Appl. Biomed. 11, 47–58. https://doi.org/10.2478/V10136-012-0031-X.

Ashley, J., Shahbazi, M.-A., Kant, K., Chidambara, V.A., Wolff, A., Bang, D.D., Sun, Y., 2017. Molecularly imprinted polymers for sample preparation and biosensing in food analysis: progress and perspectives. Biosens. Bioelectron. 91, 606–615. https://doi.org/10.1016/J.BIOS.2017.01.018.

Brougham, D.F., Ivanova, G., Gottschalk, M., Collins, D.M., Eustace, A.J., O'Connor, R., Havel, J., 2011. Artificial neural networks for classification in metabolomic studies of whole cells using 1H nuclear magnetic resonance. J. Biomed. Biotechnol. 158094. https://doi.org/10.1155/2011/158094.

Bruzas, I., Lum, W., Gorunmez, Z., Sagle, L., 2018. Advances in surface-enhanced Raman spectroscopy (SERS) substrates for lipid and protein characterization: sensing and beyond. Analyst 143, 3990–4008. https://doi.org/10.1039/C8AN00606G.

Chan, W., Jaitly, N., Le, Q., Vinyals, O., 2016. Listen, attend and spell: a neural network for large vocabulary conversational speech recognition. In: 2016 IEEE Int. Conf. Acoust. Speech Signal Process. IEEE, pp. 4960–4964. https://doi.org/10.1109/ICASSP.2016.7472621.

Chan, W., Jaitly, N., Le, Q., Vinyals, O., 2016. Listen, attend and spell: a neural network for large vocabulary conversational speech recognition. In: 2016 IEEE Int. Conf. Acoust. Speech Signal Process. IEEE, pp. 4960–4964. https://doi.org/10.1109/ICASSP.2016.7472621.

Chao, J., Cao, W., Su, S., Weng, L., Song, S., Fan, C., Wang, L., 2016. Nanostructure-based surface-enhanced Raman scattering biosensors for nucleic acids and proteins. J. Mater. Chem. B. 4, 1757–1769. https://doi.org/10.1039/C5TB02135A.

Chen, C., Li, Y., Kerman, S., Neutens, P., Willems, K., Cornelissen, S., Lagae, L., Stakenborg, T., Van Dorpe, P., 2018. High spatial resolution nanoslit SERS for single-molecule nucleobase sensing. Nat. Comput. 9 https://doi.org/10.1038/s41467-018-04118-7.

Chollet, F., 2017. Xception: deep learning With depthwise separable convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1251–1258. http://openaccess.thecvf.com/content_cvpr_2017/html/Chollet_Xception_Deep_Learning_CVPR_2017_paper.html (accessed June 7, 2019).

Cottat, M., D'Andrea, C., Yasukuni, R., Malashikhina, N., Grinyte, R., Lidgi-Guigui, N., Fazio, B., Sutton, A., Oudar, O., Charnaux, N., Pavlov, V., Toma, A., Di Fabrizio, E., Gucciardi, P.G., Lamy de la Chapelle, M., 2015. High sensitivity, high selectivity SERS detection of MnSOD using optical nanoantennas functionalized with aptamers. J. Phys. Chem. C 119, 15532–15540. https://doi.org/10.1021/acs.jpcc.5b03681.

de Groot, P., Postma, G., Melssen, W., Buydens, L.M., Deckert, V., Zenobi, R., 2001. Application of principal component analysis to detect outliers and spectral deviations in near-field surface-enhanced Raman spectra. Anal. Chim. Acta 446, 71–83. https://doi.org/10.1016/S0003-2670(01)01267-3.

Doddridge, Z.A., Warner, J.L., Cullis, P.M., Jones, G.D.D., 1998. UV-induced strand break damage in single stranded bromodeoxyuridine-containing DNA oligonucleotides. Chem. Commun. 0, 1997–1998. https://doi.org/10.1039/a804416c.

Egrioglu, E., Yolcu, U., Aladag, C.H., Bas, E., 2015. Recurrent multiplicative neuron model artificial neural network for non-linear time series forecasting. Neural Process. Lett. 41, 249–258. https://doi.org/10.1007/s11063-014-9342-0.

Eilers, P.H.C., Boelens, H.F.M., 2005. Baseline Correction with Asymmetric Least Squares Smoothing, vol 1. Leiden University Medical Centre Report, p. 5, 1. https://zanran_storage.s3.amazonaws.com/www.science.uva.nl/ContentPages/443199618.pdf.

Etchegoin, P.G., Meyer, M., Blackie, E., Le Ru, E.C., 2007. Statistics of single-molecule surface enhanced Raman scattering signals: fluctuation analysis with multiple analyte techniques. Anal. Chem. 79, 8411–8415. https://doi.org/10.1021/AC071231S.

Fornace, A.J., Kohn, K.W., Kann, H.E., 1976. DNA single-strand breaks during repair of UV damage in human fibroblasts and abnormalities of repair in xeroderma pigmentosum. Proc. Natl. Acad. Sci. U.S.A. 73, 39–43. https://doi.org/10.1073/pnas.73.1.39.

Goto, N., Bazar, G., Kovacs, Z., Kunisada, M., Morita, H., Kizaki, S., Sugiyama, H., Tsenkova, R., Nishigori, C., 2015. Detection of UV-induced cyclobutane pyrimidine dimers by near-infrared spectroscopy and aquaphotomics. Sci. Rep. 5, 11808. https://doi.org/10.1038/srep11808.

Gromski, P.S., Muhamadali, H., Ellis, D.I., Xu, Y., Correa, E., Turner, M.L., Goodacre, R., 2015. A tutorial review: metabolomics and partial least squares-discriminant analysis – a marriage of convenience or a shotgun wedding. Anal. Chim. Acta 879, 10–23. https://doi.org/10.1016/J.ACA.2015.02.012.

Guselnikova, O., Postnikov, P., Erzina, M., Kalachyova, Y., Švorčík, V., Lyutakov, O., 2017. Pretreatment-free selective and reproducible SERS-based detection of heavy metal ions on DTPA functionalized plasmonic platform. Sens. Actuators B Chem. 253, 830–838. https://doi.org/10.1016/J.SNB.2017.07.018.

Guselnikova, O., Postnikov, P., Kalachyova, Y., Kolska, Z., Libansky, M., Zima, J., Svorcik, V., Lyutakov, O., 2017. Large-scale, ultrasensitive, highly reproducible and reusable smart SERS platform based on PNIPAm-grafted gold grating. ChemNanoMat 3, 135–144. https://doi.org/10.1002/cnma.201600284.

Guselnikova, O., Postnikov, P., Pershina, A., Svorcik, V., Lyutakov, O., 2019. Express and portable label-free DNA detection and recognition with SERS platform based on functional Au grating. Appl. Surf. Sci. 470, 219–227. https://doi.org/10.1016/J.APSUSC.2018.11.092.

Guselnikova, O., Postnikov, P., Trelin, A., Švorčík, V., Lyutakov, O., 2019. Dual mode chip enantioselective express discrimination of chiral amines via wettability-based mobile application and portable surface-enhanced Raman spectroscopy measurements. ACS Sens. 4, 1032–1039. https://doi.org/10.1021/acssensors.9b00225.

Hoonejani, M.R., Pallaoro, A., Braun, G.B., Moskovits, M., Meinhart, C.D., 2015. Quantitative multiplexed simulated-cell identification by SERS in microfluidic devices. Nanoscale 7, 16834–16840. https://doi.org/10.1039/C5NR04147C.

Ioffe, S., Szegedy, C., 2015. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. http://arxiv.org/abs/1502.03167.

M. H. Jazayeri, H. Amani, A. A. Pourfatollah, H. Pazoki-Toroudi, B. Sedighimoghaddam, Various methods of gold nanoparticles (GNPs) conjugation to antibodies, Sens. Bio Sens. Res. 9 2016 17–22. doi:10.1016/J.SBSR.2016.04.002.

Kahraman, M., Mullen, E.R., Korkmaz, A., Wachsmann-Hogiu, S., 2017. Fundamentals and applications of SERS-based bioanalytical sensing. Nanophotonics 6, 831–852. https://doi.org/10.1515/nanoph-2016-0174.

Kasera, S., Herrmann, L.O., del Barrio, J., Baumberg, J.J., Scherman, O.A., 2015. Quantitative multiplexing with nano-self-assemblies in SERS. Sci. Rep. 4, 6785. https://doi.org/10.1038/srep06785.

Kermany, D.S., Goldbaum, M., Cai, W., Valentim, C.C.S., Liang, H., Baxter, S.L., McKeown, A., Yang, G., Wu, X., Yan, F., Dong, J., Prasadha, M.K., Pei, J., Ting, M.Y. L., Zhu, J., Li, C., Hewett, S., Dong, J., Ziyar, I., Shi, A., Zhang, R., Zheng, L., Hou, R., Shi, W., Fu, X., Duan, Y., Huu, V.A.N., Wen, C., Zhang, E.D., Zhang, C.L., Li, O., Wang, X., Singer, M.A., Sun, X., Xu, J., Tafreshi, A., Lewis, M.A., Xia, H., Zhang, K., 2018. Identifying medical diagnoses and treatable diseases by image-based deep learning. Cell 172, 1122–1131. https://doi.org/10.1016/J.CELL.2018.02.010 e9.

Kingma, D.P., Ba, J., 2014. Adam: a method for stochastic optimization. arXiv preprint arXiv, 1412. http://arxiv.org/abs/1412.6980 (accessed June 7, 2019), 6980.

Kuo, T.M., Shen, M.Y., Huang, S.Y., Li, Y.K., Chuang, M.C., 2015. Facile fabrication of a sensor with a bifunctional interface for logic analysis of the New Delhi metallo-β-lactamase (NDM)-Coding gene. ACS Sens. 1, 124–130. https://doi.org/10.1021/acssensors.5b00080.

Lake, B.M., Salakhutdinov, R., Tenenbaum, J.B., 2015. Human-level concept learning through probabilistic program induction. Science 350, 1332–1338. https://doi.org/10.1126/science.aab3050.

Lane, L.A., Xue, R., Nie, S., 2018. Emergence of two near-infrared windows for in vivo and intraoperative SERS. Curr. Opin. Chem. Biol. 45, 95–103. https://doi.org/10.1016/J.CBPA.2018.03.015.

Li, S.-X., Zeng, Q.-Y., Li, L.-F., Zhang, Y.-J., Wan, M.-M., Liu, Z.-M., Xiong, H.-L., Guo, Z.-Y., Liu, S.-H., 2013. Study of support vector machine and serum surface-enhanced Raman spectroscopy for noninvasiveesophageal cancer detection. J. Biomed. Opt. 18, 027008 https://doi.org/10.1117/1.JBO.18.2.027008.

Liu, J., Osadchy, M., Ashton, L., Foster, M., Solomon, C.J., Gibson, S.J., 2017. Deep convolutional neural networks for Raman spectrum recognition: a unified solution. Analyst 142, 4067–4074. https://doi.org/10.1039/c7an01371j.

Meurer, A., Smith, C.P., Paprocki, M., Čertík, O., Kirpichev, S.B., Rocklin, M., Kumar, Am, Ivanov, S., Moore, J.K., Singh, S., Rathnayake, T., Vig, S., Granger, B.E., Muller, R.P., Bonazzi, F., Gupta, H., Vats, S., Johansson, F., Pedregosa, F., Curry, M. J., Terrel, A.R., Roučka, Š., Saboo, A., Fernando, I., Kulal, S., Cimrman, R., Scopatz, A., 2017. SymPy: symbolic computing in Python. PeerJComput. Sci. 3, e103. https://doi.org/10.7717/peerj-cs.103.

Meurer, A., Smith, C.P., Paprocki, M., Čertík, O., Kirpichev, S.B., Rocklin, M., Kumar, Am, Ivanov, S., Moore, J.K., Singh, S., Rathnayake, T., Vig, S., Granger, B.E., Muller, R.P., Bonazzi, F., Gupta, H., Vats, S., Johansson, F., Pedregosa, F., Curry, M. J., Terrel, A.R., Roučka, Š., Saboo, A., Fernando, I., Kulal, S., Cimrman, R., Scopatz, A., 2017. SymPy: symbolic computing in Python. PeerJComput. Sci. 3, e103. https://doi.org/10.7717/peerj-cs.103.

Moore, A.A., Jacobson, M.L., Belabas, N., Rowlen, K.L., Jonas, D.M., 2005. 2D correlation analysis of the continuum in single molecule surface enhanced Raman spectroscopy. J. Am. Chem. Soc. 127, 7292–7293. https://doi.org/10.1021/JA043651U.

Naik, B., Nayak, J., Behera, H.S., Abraham, A., 2016. A self adaptive harmony search based functional link higher order ANN for non-linear data classification. Neurocomputing 179, 69–87. https://doi.org/10.1016/J.NEUCOM.2015.11.051.

Ochsenkühn, M.A., Campbell, C.J., 2010. Probing biomolecular interactions using surface enhanced Raman spectroscopy: label-free protein detection using a G-quadruplex DNA aptamer. Chem. Commun. 46, 2799. https://doi.org/10.1039/b920941g.

Özbalci, B., Boyaci, İ.H., Topcu, A., Kadılar, C., Tamer, U., 2013. Rapid analysis of sugars in honey by processing Raman spectrum using chemometric methods and artificial neural networks. Food Chem. 136, 1444–1452. https://doi.org/10.1016/J.FOODCHEM.2012.09.064.

Peccia, J., Hernandez, M., 2002. Rapid immunoassays for detection of UV-induced cyclobutane pyrimidine dimers in whole bacterial cells. Appl. Environ. Microbiol. 68, 2542–2549. https://doi.org/10.1128/AEM.68.5.2542-2549.2002.

Puebla, R.A., Marzán, L.M., 2012. SERS detection of small inorganic molecules and ions. Angew. Chem. Int. Ed. 51, 11214–11223. https://doi.org/10.1002/anie.201204438.

Santos, M.C.D., Morais, C.L.M., Nascimento, Y.M., Araujo, J.M.G., Lima, K.M.G., 2017. Spectroscopy with computational analysis in virological studies: a decade (2006–2016). TrAC Trends Anal. Chem. 97, 244–256. https://doi.org/10.1016/J.TRAC.2017.09.015.

Savitzky, A., M.J.E., 1951. Smoothing and differentiation of data by simplified least squares procedures. Z. Physiol. Chem. 40, 1832. https://doi.org/10.1021/ac60214a047.

Segler, M.H.S., Kogej, T., Tyrchan, C., Waller, M.P., 2018. Generating focused molecule libraries for drug discovery with recurrent neural networks. ACS Cent. Sci. 4, 120–131. https://doi.org/10.1021/acscentsci.7b00512.

Seifert, S., Merk, V., Kneipp, J., 2016. Identification of aqueous pollen extracts using surface enhanced Raman scattering (SERS) and pattern recognition methods. J. Biophot. 9, 181–189. https://doi.org/10.1002/jbio.201500176.

Shan, B., Pu, Y., Chen, Y., Liao, M., Li, M., 2018. Novel SERS labels: rational design, functional integration and biomedical applications. Coord. Chem. Rev. 371, 11–37. https://doi.org/10.1016/J.CCR.2018.05.007.

Shi, H., Wang, H., Meng, X., Chen, R., Zhang, Y., Su, Y., He, Y., 2018. Setting up a surface-enhanced Raman scattering database for artificial-intelligence-based label-free discrimination of tumor suppressor genes. Anal. Chem. 90, 14216–14221. https://doi.org/10.1021/acs.analchem.8b03080.

Shin, H., Jeong, H., Park, J., Hong, S., Choi, Y., 2018. Correlation between cancerous exosomes and protein markers based on surface-enhanced Raman spectroscopy (SERS) and principal component analysis (PCA). ACS Sens. 3, 2637–2643. https://doi.org/10.1021/acssensors.8b01047.

Sinha, R.P., Häder, D.-P., 2002. UV-induced DNA damage and repair: a review. Photochem. Photobiol. Sci. 1, 225–236. https://doi.org/10.1039/b201230h.

Srivastava, N., Hinton, G., Krizhevsky, A., Salakhutdinov, R., 2014. Dropout: a simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. 15, 1929–1958.

Szlag, V.M., Rodriguez, R.S., He, J., Hudson-Smith, N., Kang, H., Le, N., Reineke, T.M., Haynes, C.L., 2018. Molecular affinity agents for intrinsic surface-enhanced Raman scattering (SERS) sensors. ACS Appl. Mater. Interfaces 10, 31825–31844. https://doi.org/10.1021/acsami.8b10303.

Trelin, A., Prochazka, A., 2019. Binary Stochastic Filtering: a Solution for Supervised Feature Selection and Neural Network Shape Optimization, p. 04510 arXiv preprint arXiv 1902. http://arxiv.org/abs/1902.04510.

Vendrell, M., Maiti, K.K., Dhaliwal, K., Chang, Y.-T., 2013. Surface-enhanced Raman scattering in cancer detection and imaging. Trends Biotechnol. 31, 249–257. https://doi.org/10.1016/J.TIBTECH.2013.01.013.

Yuan, Y., Panwar, N., Yap, S.H.K., Wu, Q., Zeng, S., Xu, J., Tjin, S.C., Song, J., Qu, J., Yong, K.-T., 2017. SERS-based ultrasensitive sensing platform: an insight into design and practical applications. Coord. Chem. Rev. 337, 1–33. https://doi.org/10.1016/J.CCR.2017.02.006.

Zhang, L., Li, Q., Tao, W., Yu, B., Du, Y., 2010. Quantitative analysis of thymine with surface-enhanced Raman spectroscopy and partial least squares (PLS) regression. Anal. Bioanal. Chem. 398, 1827–1832. https://doi.org/10.1007/s00216-010-4074-0.