Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 347 (2026) 127054

Contents lists available at ScienceDirect

Spectrochimica Acta Part A:
Molecular and Biomolecular Spectroscopy

journal homepage: www.journals.elsevier.com/spectrochimica-acta-part-a-molecular-and-biomolecular-spectroscopy

# Living breast cancer subtype classification by membrane-interfacing 3D surface-enhanced Raman spectroscopy substrates with multivariate analysis

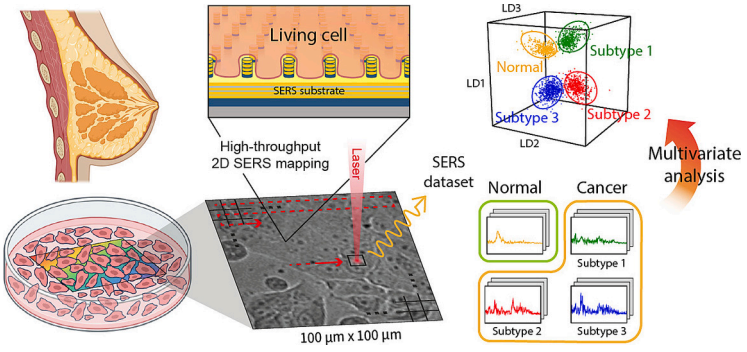
Hyeim Yu [a], Xiang Ren [b], Wonil Nam [a,*]

[a] Department of Electronic Engineering, Pukyong National University, Busan 48513, Republic of Korea
[b] Department of Microelectronics, Tianjin University, Tianjin 300072, China

## HIGHLIGHTS

- 3D nanostructure enables uniform membrane-hotspot interfacing for sensitive detection.
- Wafer-scale platform allows high-throughput, rapid SERS mapping of living cells.
- Large, reproducible spectral datasets enable robust multi-class subtype classification.
- PCA-LDA analysis achieves 92.5 % accuracy in classifying breast cancer subtypes.

## GRAPHICAL ABSTRACT




## ARTICLE INFO

*Keywords:*
Surface-enhanced Raman spectroscopy
Nano-bio interface
Living cancer cells
Multi-class classification
Cancer diagnostics

## ABSTRACT

Surface-enhanced Raman spectroscopy (SERS) is a non-destructive and highly sensitive technique widely used for analyzing complex biological samples. However, conventional SERS approaches for living cell analysis face significant challenges. In particular, nanoparticle-based intracellular probes can induce cytotoxicity, and attempts to use 3D protruding nanostructures to interrogate cells externally have shown limited success so far. Furthermore, existing cell-interfacing SERS substrates typically capture only a small fraction of cellular biomolecular signals, producing sparse data that often limits classification tasks to simple binary outcomes. To address these limitations, we developed a 3D multilayer nanolaminate SERS substrate composed of gold and silica ($Au/SiO_2$) that interfaces directly with the cell membrane. Using this platform, we achieved high-speed, high-throughput, label-free SERS mapping of living breast cancer cells, obtaining large and information-rich spectral datasets. In contrast to prior methods, our approach enabled the classification of four breast cancer subtypes with 92.5 % accuracy using conventional machine learning algorithms. This label-free SERS platform demonstrates potential for more complex cellular analyses via advanced machine learning, including studies of cellular responses to external stimuli such as drug treatments.

* Corresponding author at: Department of Electronic Engineering, Pukyong National University, Busan 48513, Republic of Korea.
*E-mail address:* nam@pknu.ac.kr (W. Nam).

https://doi.org/10.1016/j.saa.2025.127054
Received 25 August 2025; Received in revised form 25 September 2025; Accepted 14 October 2025
Available online 16 October 2025

## 1. Introduction

Surface-enhanced Raman spectroscopy (SERS) has emerged as an ultrasensitive detection technique, significantly amplifying chemical-specific vibrational fingerprints through plasmonic hotspots [1]. The strong electromagnetic (EM) coupling between the surface plasmon-mediated local EM field and the inelastic Raman scattering of molecules in the plasmonic hotspots enables SERS sensitivity to reach single-molecule detection [2]. Such ultrahigh sensitivity allows label-free sensing of the intrinsic chemical signatures of analyte molecules in various samples without complex pre-treatment or pre-processing. Additionally, as an optical sensing technique, SERS is non-destructive, and its narrow peak bandwidth offers powerful multiplexing capability for analyzing complex matrices. Due to these prominent features, SERS has been extensively utilized for diverse biological samples, including DNA, RNA, biomolecules, extracellular vesicles, and biofluids, as well as in vitro, ex vivo, and in vivo models [3–9].

For cellular SERS applications, numerous studies on living cells have been reported to investigate cellular mechanisms involving proteins, lipids, and DNA [10]. SERS bio-analysis of living cells enables molecular profiling of cellular components with minimum disruption to the cell's natural status, providing enhanced biological reliability compared to dead or fixed cells. To probe the intracellular structure and dynamics of living cells, bottom-up-based plasmonic anoparticles utilizing endocytosis have been employed [11–14]. Plasmonic nanoparticles internalized into cells by endocytosis can serve as nanoantennas by directly delivering enhanced Raman spectroscopic signatures of biological molecules and structures of intracellular environments. However, challenges remain as the use of nanoparticles can lead to toxicity, and their eventual localization within lysosomes precludes the investigation of various intracellular mechanisms [10,15]. One interesting approach to mitigate the limited localization of nanoparticles involves label-based techniques with spatially and spectrally selective Raman reporters functionalized onto nanoparticle surfaces [16–20]. However, these methods still require complex chemical processes and target pre-knowledge.

It has been reported that membrane engulfment over 3D protruding nanostructures provides a more consistent nano-bio interface compared to 2D planar substrates with random adhesion points [21–23]. For example, nanoelectrodes with 3D vertical nanotopography showed a significantly enhanced signal-to-noise ratio in electrophysiological applications. The underlying biological mechanism involves topography-induced membrane curvature, which serves as a biochemical signal to initiate actin reorganization within cells through the curvature-sensitive protein FBP17 [24]. Considering the extremely sensitive distance dependence of SERS intensity ($I \propto r^{-12}$) [25,26], the nanotopography-induced tight nano-bio interface can provide a consistent membrane-hotspot interface for cellular SERS studies and few top-down-based 3D vertical nanoantennas have been reported as an alternative to colloidal nanoparticles [27–31]. However, realizing a suitable 3D vertical SERS substrate for label-free living cell analysis has still been challenging due to the poor sensitivity based on low hotspot density and small device footprint that precludes investigation of the cellular network.

Therefore, the amount of biomolecular fingerprint data acquired from previous cell-interfaced SERS substrates has generally been limited to tens of spectra, leading to a binary classification of intrinsically different sample groups, for instance, live versus dead cells, normal versus cancer cells, or cancer cells from different organs [6,32,33]. In order for SERS to play a more significant role in biomedical applications, multiple sample classifications, such as cell subtypes or cells with different external stimuli conditions, need to be achieved. For example, different types of cancer cells from the same organ exhibit significantly different clinical prognoses; therefore, precise and rapid diagnosis of the subtype is critically essential for patients. However, SERS-based multi-class classification for similar cell groups with subtle biological differences remains challenging due to the difficulty in acquiring a large and reliable SERS dataset. To establish a reliable label-free cell-interfaced SERS platform, crucial factors should be ensured, including high-performance SERS substrates with a large device footprint and a consistent interface with cell membranes.

In this work, we present a label-free approach for classifying living breast cancer subtypes enabled by membrane-interfacing 3D vertical SERS substrates. We performed label-free high-throughput SERS mapping for four types of normal and cancerous living breast cells, including MCF-10A, MCF-7, HCC-1806, and MDA-MB-231, and achieved a multi-class classification accuracy of 92.5 %. A vertically stacked metal-insulator-metal (MIM) structure-based SERS substrate with uniform and highly sensitive hotspots, named nanolaminate, exhibits a broad plasmonic resonance that spectrally covers the biological Raman window. The 3D protruding nanotopography of MIM building blocks induces cell engulfment, facilitating a tight membrane-hotspot interface. Such a nano-bio interface allows a highly sensitive hotspot for directly acquiring extracellular signals from living cells.

## 2. Experimental sections

### 2.1. Fabrication of 3D nanostructure SERS substrate

First, we made a composite polydimethylsiloxane (PDMS) stamp of nanowell arrays with a period of 400 nm, a diameter of 100 nm, and a height of 150 nm from a silicon wafer patterned with nanopillar structures by soft lithography. By molding with the PDMS stamp, we used UV-curable polyurethane (PU) (NOA83H, Norland Product Inc.) to fabricate nanopillar arrays on a flexible and optically transparent polyester film. After 10 min UV curing, we performed an additional heat-curing process at 80 °C in a convection oven overnight. Next, alternating layers of gold and $SiO_2$ were deposited by electron-beam deposition (PVD250, Kurt J. Lesker Company). The thickness for four gold layers is 30 nm, and the thicknesses of three $SiO_2$ layers are 6 nm, 8 nm, and 12 nm from bottom to top. We deposited 1 nm of chromium between the polymer nanopillar array and the first layer of gold, and 0.7 nm thick titanium between gold and insulator layers as adhesion layers. Buffered oxide etchant (BOE, 10:1) (Transene Inc.) was then used to etch $SiO_2$ layers for 20 s and expose embedded MIM plasmonic hotspots.

### 2.2. Finite-difference time domain (FDTD)

3D FDTD simulations were performed by commercial software (FDTD solution, Ansys Lumerical Inc., USA). A uniform 2 nm mesh was used for x-, y-, and z-directions. The optical constants of gold were taken from Johnson and Christy. The Bloch boundary condition was used in x- and y-directions with a periodicity of 400 nm and the perfectly matched layer boundary condition was used in the z-direction. The refractive indices of $SiO_2$, PU, and background were set as 1.5, 1.56, and 1.33, respectively.

### 2.3. Cell culture

MDA-MB-231 (American Type Culture Collection, ATCC) was grown in F12:DMEM (Dulbecco's Modified Eagle Medium, Lonza, Basel, Switzerland) with 4 mM glutamine, 10 % fetal bovine serum (FBS), and penicillin−streptomycin (100 units per mL). HCC-1806 (ATCC) was grown in an ATCC-formulated RPMI-1640 medium (Roswell Park Memorial Institute 1640 medium, enriched with L-glutamine, 4-(2-hydroxyethyl)-1 piperazine ethanesulfonic acid (HEPES), and sodium pyruvate, ATCC 30–2001) with 10 % FBS and 1 % PenStrep (100 units/mL penicillium and 100 μg/mL streptomycin). MCF-7 cells (ATCC) were grown in EMEM with 10 % FBS and 2× L glutamine. MCF-10A cells (Lombardi Comprehensive Cancer Center, Georgetown University in Washington, D.C.) were grown in F12:DMEM with penicillin−streptomycin (100 units/mL), 20 ng/mL epidermal growth factor (EGF), 2.5 mM L-glutamine, 10 μg/mL insulin, 0.1 μg/mL cholera

toxin, 0.5 μg/mL hydrocortisone, and 5 % horse serum. All cells were grown in T-25 cm2 culture flasks (Corning, NY) at 37 °C in a 5 % CO2 in air atmosphere. Cells were then trypsinized and seeded on nanolaminate SERS substrates.

### 2.4. SEM for cells

SEM was performed using FEI Helios 600 Nanolab Dual-beam. Cultured cells were rinsed by PBS solution twice, followed by fixation with 2.5 % glutaraldehyde in PBS solution at room temperature for 1 h. Cells were then rinsed by PBS solution twice, followed by post-fixation with 1 % osmium tetroxide and dehydration in graded ethanol series from 15 % to 100 % (each condition was carried out for 15 min). A critical point dryer dried cells in liquid $CO_2$. 5 nm of PtPd was sputtered as a conducting layer to reduce the charging in SEM measurements.

### 2.5. SERS measurement

We used a confocal Raman microscope (Alpha 300 RSA+, WITec, Germany) for SERS measurements under laser excitation at 785 nm (Toptica Photonics, Germany) via a 20× objective (NA = 0.4). For 2D Raman mapping of living cells, a 20× water immersion objective (NA = 0.5) was used with 5 mW laser power and 10 ms integration time per pixel over a 100 μm × 100 μm area. Before the measurement, the instrumental calibration was verified by the silicon peak at 520 $cm^{-1}$. All measurements were conducted in the backscattering geometric configuration at room temperature. A long-pass filter blocks elastically scattered radiation at the wavelength corresponding to the laser line (Rayleigh scattering). Simultaneously, the rest of the collected light was guided through a multimode fiber (100 μm core diameter), acting as the pinhole for a confocal microscope, to a spectrometer (UHTS 300, WITec, Germany). The backscatter photons were dispersed with a 300 groove $mm^{-1}$ (750 nm blaze grating) and detected by a CCD camera which was thermoelectrically cooled and maintained at −60 °C.

### 2.6. Statistical analysis

Cosmic ray removal was conducted by instrument-embedded software (Project v4.1, WITec). Smoothing interpolation and data truncation were carried out with the R package hyperSpec. Principal component analysis (PCA) and peak picking were done with the R packages ChemoSpec and MALDIquant, respectively. Linear discriminant analysis (LDA) and leave-one-out cross-validation (LOOCV) were performed using the R package MASS.

## 3. Results and discussion

Fig. 1 depicts an overall flowchart illustrating the label-free cell-interfaced SERS platform for multi-class classification of four living breast cell subtypes. We fabricated nanolaminate SERS substrates, consisting of multilayered gold and silicon dioxide, that support uniform and dense 3D vertical hotspots with a high SERS EF value of $1.6 \times 10^8$ [34]. The high SERS sensitivity is crucial as ultralow concentrations of biomolecules may significantly contribute to subtype classification. The 3D protruding nanotopography of the nanolaminate structure can induce cell engulfment, allowing a tight membrane-hotspot interface and reliable label-free biomolecular sensing from living cells. As shown in Fig. 1 (middle), we confirmed the distinct morphology of living cancer cells cultured on the SERS substrate, reflecting their healthy status, and the scanning electron microscopy (SEM) image shows the membrane-hotspot interface. By leveraging the high SERS performance of a uniform membrane-hotspot interface and a large device footprint (~16 $cm^2$), we performed high-throughput 2D SERS mapping for living cells over a large area (100 μm × 100 μm). Four different types of normal and cancerous cells were studied, and a total of 10,000 data points for each cell type were acquired and filtered, resulting in a high-quality large volume of SERS dataset (total $n > 2000$). To achieve a multi-class classification of living breast cancer subtypes, we employed a conventional multivariate analysis, principal component analysis (PCA) combined with linear discriminant analysis (LDA). PCA was used as a dimension reduction tool, and the processed principal components (PCs)

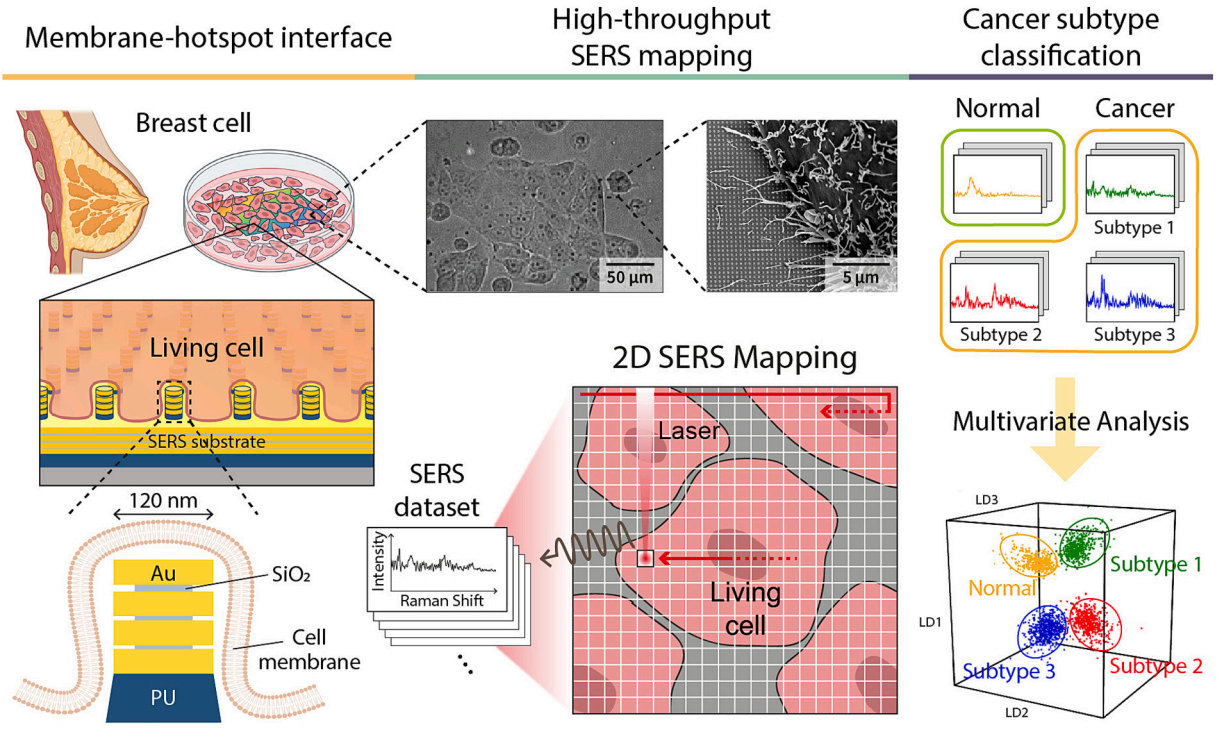



**Fig. 1.** A schematic illustration for classifying four different types of living breast cells (normal and cancerous) using multivariate analysis. High-throughput SERS mapping using nanolaminate SERS substrates for a large-volume living cell dataset acquisition.

were used as input into the LDA model for the classification of four different breast cell lines, including normal breast cell (MCF-10A), moderately malignant breast cancer cell (MCF-7), and two highly malignant breast cancer cells (HCC-1806 and MDA-MB-231).

### 3.1. Fabrication of nanolaminate structures and optical properties

Fig. 2A shows an SEM image of the nanolaminate SERS substrate with a uniform MIM array. The diameter of the MIM building block is 120 nm with a periodicity of 400 nm. To fabricate SERS substrates with MIM building block arrays, we employed soft lithography to create a nanowell composite hard-polydimethylsiloxane (PDMS) stamp [35]. These three dielectric layer thicknesses are designed based on plasmon resonance positions of the magnetic dipole modes to achieve broad resonance from the visible to near-infrared (NIR) range. The fabricated SERS substrate has a large device footprint of 16 $cm^2$, providing a sufficient area for direct cell culture.

To investigate the optical properties of nanolaminate SERS substrate interfacing with living cells, we measured its reflectance spectrum with a water background. Additionally, to gain a deeper insight into the optical properties of the nanolaminate SERS substrate, we employed the finite-difference time-domain (FDTD) method to investigate far-field (reflectance) and near-field responses. The simulation background refractive index was set to 1.33 (water) to mimic the biological condition. Fig. 2B shows the measured and calculated reflectance spectra from 400 nm to 1300 nm. In the calculated spectrum, multiple resonant dips were observed at 590 nm, 728 nm, 752 nm, 850 nm, and 934 nm, each originating from different physical origins. These physical origins include delocalized plasmonic modes from the bottom MIM nanohole arrays, localized gap modes from the vertically stacked nanodisks, and their plasmonic hybridizations [33]. The near-field electric field profiles at these resonant dips are shown in **Fig. S1**. While the measured spectrum shows a good agreement with the calculated reflectance dip positions, it exhibits a broader dip profile. Such a broad plasmonic resonant feature arises from inhomogeneous broadening caused by geometric variations between individual nanostructures and homogeneous broadening resulting from increased metal losses associated with the interface roughness between metal and dielectric layers. The broadband resonant property extending from vis to NIR range is advantageous for label-free living cell SERS measurement under NIR excitation (785 nm), as it encompasses both the excitation laser and the biological window of the Stokes Raman scattering region.

The near-field optical properties of the MIM nanoantenna were analyzed at an excitation wavelength of 785 nm. Fig. 2C displays the electric field intensity map and the $SiO_2$ nanogaps (GAP1, GAP2, and GAP3) at different z-positions. Strongly enhanced electric fields are observed at the nanogap region and sidewall (gap entrance), directly interacting with living cell membranes. In Fig. 2D, we calculated the electric field intensity of three nanogaps along the x-axis and confirmed that the field enhancements are mainly observed in the nanogap and gap entrance regions. At each nanogap GAP1, GAP2, and GAP3, the maximum values of the $|E|^2$ are observed near the gap entrance, with values of $3.2 \times 10^3$, $2.4 \times 10^3$, and $1.1 \times 10^3$, respectively. Based on $|E|^4$ approximation [36], their corresponding estimated SERS EFs are $1.0 \times 10^7$, $5.6 \times 10^6$, and $1.2 \times 10^6$, respectively.

### 3.2. High-throughput SERS data acquisition

Fig. 3A-D show bright-field images of normal breast cells and three different types of breast cancer cells directly cultured on SERS substrates. Fig. 3A shows the normal MCF-10A cells, Fig. 3B shows the moderately malignant MCF-7 breast cancer cells, and Fig. 3C and D show the two triple-negative breast cancer cells, HCC-1806 and MDA-MB-231. Triple-negative refers to a subtype of breast cancer in which the cancer cells lack expression of the estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2). Triple-negative breast cancer (TNBC) cells have been reported not to respond to hormonal therapies or treatments targeting HER2,

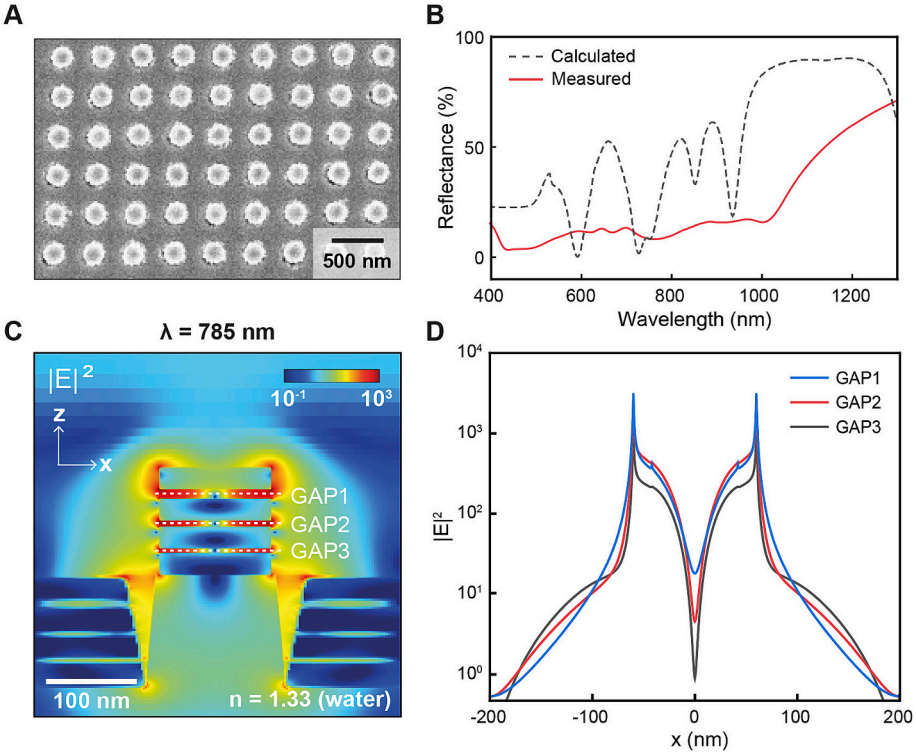



**Fig. 2.** SERS substrate characterization. (A) Top-view SEM image of the nanolaminate SERS substrate. (B) Measured and FDTD-calculated reflectance spectrum of the nanolaminate substrate. (C) FDTD-calculated x-z distribution map of $|E|^2$ at 785 nm under a background refractive index of 1.33. (D) FDTD-calculated $|E|^2$ as a function of x for the three MIM plasmonic nanogaps (GAP1, GAP2, and GAP3).

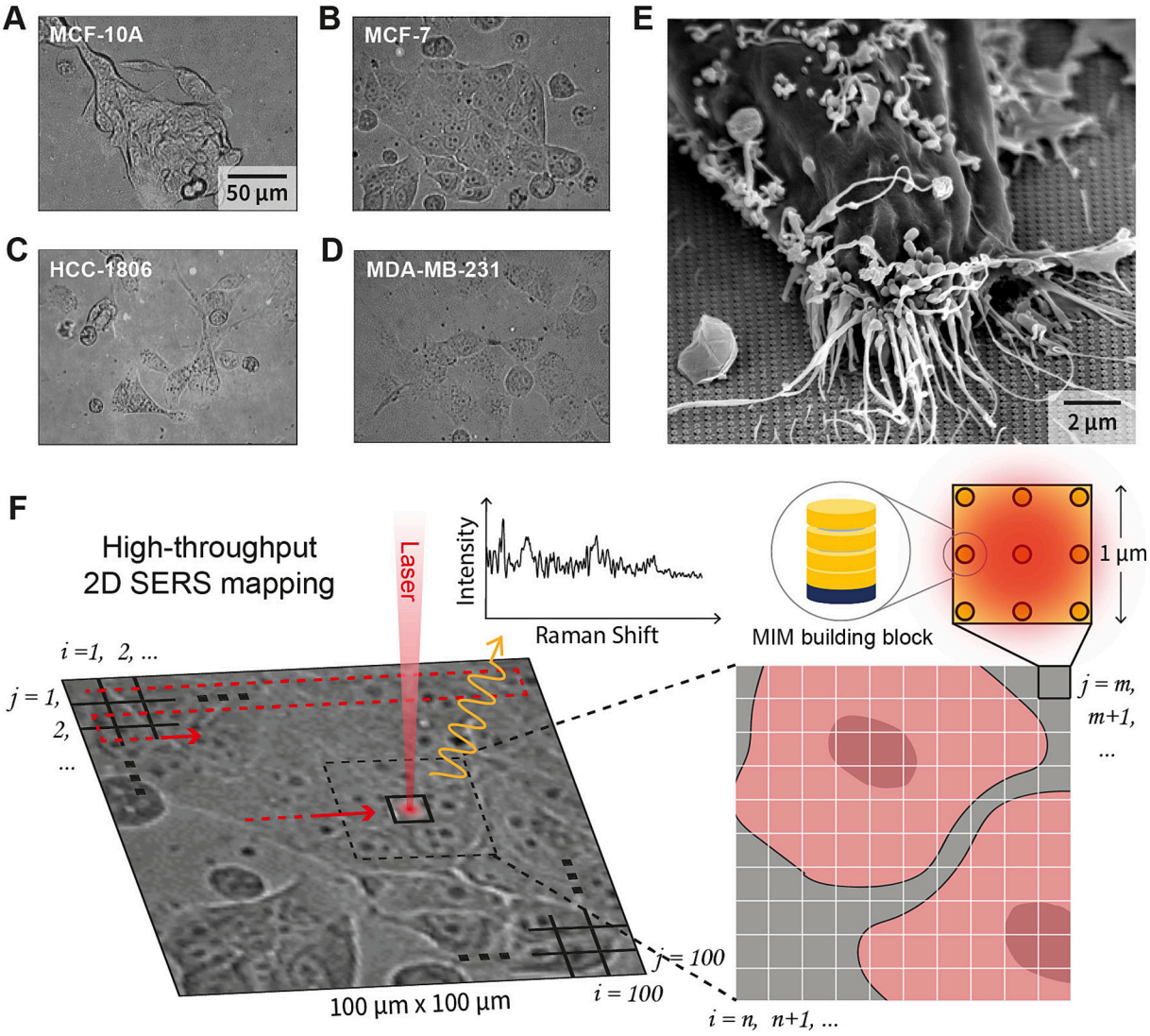



**Fig. 3.** High-throughput label-free SERS mapping for living cells. Bright-field images of cells: (A) MCF-10A, (B) MCF-7, (C) HCC-1806, and (D) MDA-MB-231. (E) SEM image of an MDA-MB-231 cell on the nanolaminate SERS substrate. (F) High-throughput 2D SERS mapping over a 100 μm × 100 μm region, with each pixel size of 1 μm × 1 μm, covering approximately five MIM buildings per laser spot.

making them more challenging to treat and considered a highly malignant cancer type.

MCF-10A cells exhibit a round or oval shape, characteristic of normal breast epithelial cells, and tend to form a monolayer due to cell-cell contact [37]. MCF-7 cells, in contrast, have a relatively polygonal shape with stronger cell-cell interactions compared to MCF-10A [38]. HCC-1806 cells display a distinct spindle shape and scattered distribution, reflecting the characteristics of basal-like TNBC [39]. MDA-MB-231 cells, known for their more aggressive cancer behavior, have less cell-cell contact and exhibit a higher degree of atypical morphology [40]. The morphology of cells directly cultured on SERS substrates matches the reported morphological features well, indicating the healthy status of cells and the biocompatibility of gold-based nanolaminate SERS substrates.

To further investigate the membrane-hotspot interface, Fig. 3E presents an SEM image of the MDA-MB-231 cell cultured on the SERS substrate. Cancerous cells differ from normal cells in terms of interaction with the extracellular matrix (i.e., SERS substrate), with the structural feature of a brush-like layer being a key consideration. The brush-like layer interacting with the substrate exhibits morphology and length consistent with previously reported findings [41]. Additionally, the SEM image reveals how cells adhere to the nanolaminate SERS substrate, with the brush-like layer appearing to grasp or wrap the 3D protruding nanotopography. Such adhesion of cells induced by the 3D nanotopography enables a tight membrane–hotspot interface, facilitating the direct acquisition of biomolecular fingerprints from the cellular membrane environment.

Fig. 3F demonstrates the high-throughput SERS measurement for living cells. We performed 2D mapping over a 100 μm × 100 μm region containing several living cells with a pixel size of 1 μm × 1 μm and an integration time of 10 ms. The high sensitivity of nanolaminate SERS substrates allows a short integration time and low laser power during measurement, minimizing sample damage from high laser power-induced and/or plasmon-mediated local heating. By rapidly scanning the entire area within 3–5 min, we can maximize the signal reliability of living cancer cells with minimum perturbations from dynamic cellular processes. For example, we experimentally observed active cell division after tens of minutes during bright-field imaging. Rapid mapping across nanolaminate SERS substrates ensures improved reliability in label-free living cell SERS measurements.

### 3.3. Biomolecular peak assignment

For 10,000 data points acquired from 2D SERS mapping, areas

without cells and spectra with no peaks were discarded to focus on cell-specific SERS signals. We obtained 422, 538, 561, and 560 SERS spectra for MCF-10A, MCF-7, HCC-1806, and MDA-MB-231, respectively, and performed spectral smoothing to reduce noise. Fig. 4A-D show the average SERS spectra for each cell type, with data distribution ranging from the 5th to 95th percentile. For all cases, the absence of broad carbon-based D (1350 $cm^{-1}$) and G (1580 $cm^{-1}$) bands, which can possibly mask the observation of low-intensity SERS signals, confirms that the laser excitation did not generate photoinduced graphitization of organic components of cells. By comparing SERS spectra among normal and the other three types of cancer cells, we can find that the SERS signal of MCF-10A cells shows noticeably different peak distribution from those of cancer cells. Specifically, the normal cells show a prominent peak only in the 600–800 $cm^{-1}$ region, whereas the cancer cells show peaks scattered over a broader range. Among the three cancerous cell groups, the SERS spectra of highly malignant HCC-1806 and MDA-MB-231 cells show significant intensity variations with a much wider signal distribution compared to moderately malignant MCF-7 cells. This difference reflects the inherent cancerous heterogeneities associated with the degree of malignancy, including distinct molecular signatures and variations in the molecular composition of cancer cells [42].

As shown in Fig. 4, all cancer cells (4B-4D) reveal more peaks in the lipid-relevant region (1050–1500 $cm^{-1}$) compared to normal cells. Notably, highly malignant cancer cells show higher SERS intensities in the lipid-relevant region, reflecting enhanced lipid content due to increased synthesis of fatty acids and phospholipids [43]. Furthermore, the characteristic peaks of each cell type can indicate specific and comprehensive biochemical information. A few characteristic peaks observed in the spectra reveal that the measured SERS signals originate from viable living cells. The presence of the C-C stretching skeletal of the acyl backbone in the lipid peak (1129 $cm^{-1}$), which is directly related to components of cell membranes, reflects a healthy state of the cells [27,44]. Moreover, the appearance of the cholesterol peak (416 $cm^{-1}$) supports cellular membrane fluidity [30,44]. The good viability is further confirmed by the absence of benzene ring stretching (1000 $cm^{-1}$) and N−H out-of-plane bending (1585 $cm^{-1}$) modes associated with the cellular death dynamics [45]. Additionally, the observation of characteristic peaks can reveal the biological differences between normal and cancer cells. For example, breast cancer cells show stronger phenylalanine (645 $cm^{-1}$) and tyrosine (1178 $cm^{-1}$) peaks compared to normal cells [44]. These differences may indicate overexpression of aromatic amino acids and increased protein content in cancer cells, reflecting metabolic abnormalities and enhancement of protein disorder within the cells [46]. While the averaged spectra effectively reveal biological characteristics, the averaging process within each cell group may mask critical information present at low concentrations of biomolecules. Furthermore, manual interpretation of data is not only unreliable due to its subject peak assignment and lack of consistency, but it is also highly inefficient. Therefore, multivariate analysis can play a significant role in extracting distinctive characteristics from a large volume of complex SERS data containing various biochemical information of each cell group.

### 3.4. Multi-class classification by PCA-LDA

We performed multi-class classification using PCA-LDA to efficiently analyze rich bio-fingerprint information from living cells and distinguish cell types. PCA, a dimensionality reduction technique, simplifies a large dataset into a smaller one while preserving key patterns and trends [47]. Fig. 5A shows the plot of PC scores for four different cells. The PC score plot reveals that HCC-1806 (red) and MDA-MB-231 (blue) show broader dispersion compared to MCF-10A (green) and MCF-7 (yellow). These data distributions for each group of cells are consistent with their biological property, reflecting inherent cancerous heterogeneities associated with the degree of malignancy. PC scatter plot allows for a visual interpretation of the dataset variability that may help determine similarities and differences between cell types. However, since PCA does not consider interclass variability, it was used as a tool for dimensionality reduction to extract essential characteristic information as input for subsequent LDA.

To conduct multi-class classification, LDA, a supervised technique used to find a linear combination of features that best separates different classes, was performed. Since using an unnecessarily high number of PCs may lead to over-training, using the entire PCA dataset for LDA is not feasible. Therefore, to preserve the essential characteristics of the data while reducing dimensions and improving computational efficiency, only the PCA data within the range of 95 % cumulative variance was used as input for the LDA [48,49]. Fig. 5B-D show PCA-LDA score plots based on the orthogonal LD1, LD2, and LD3 axes that maximize the distinction between classes, with the axes for LD1-LD2, LD1-LD3, and LD2-LD3. The results visualized in the 2D plot of PCA-LDA reveal that intra-class cohesion is high, and the distinction between different classes is more pronounced. The 3D representation of PCA-LDA for four classes in Fig. 5E clearly shows the distinction between the cell types through the visible clustering of each group. While each cell type was successfully distinguished, slight overlapping scatter distributions were observed between the two MCF groups and the two TNBC groups. These overlap may reflect their similar cellular network expression, such as lipid contents and surface proteins.

Leave-one-out cross-validation (LOOCV) was then performed to validate the predictive ability of the PCA-LDA classifier. This approach maximizes the use of data by ensuring that each single sample serves as a test point while the model is trained on the remaining data points, offering a thorough evaluation of the model. Fig. 5F displays the multi-class classification results as a histogram, showing accuracies of 86.7 % for MCF-10A, 95.9 % for MCF-7, 93.4 % for HCC-1806, and 92.5 % for MDA-MB-231, resulting in an average accuracy of 92.5 %. The corresponding confusion matrix is shown in **Fig. S2**. False predictions in the

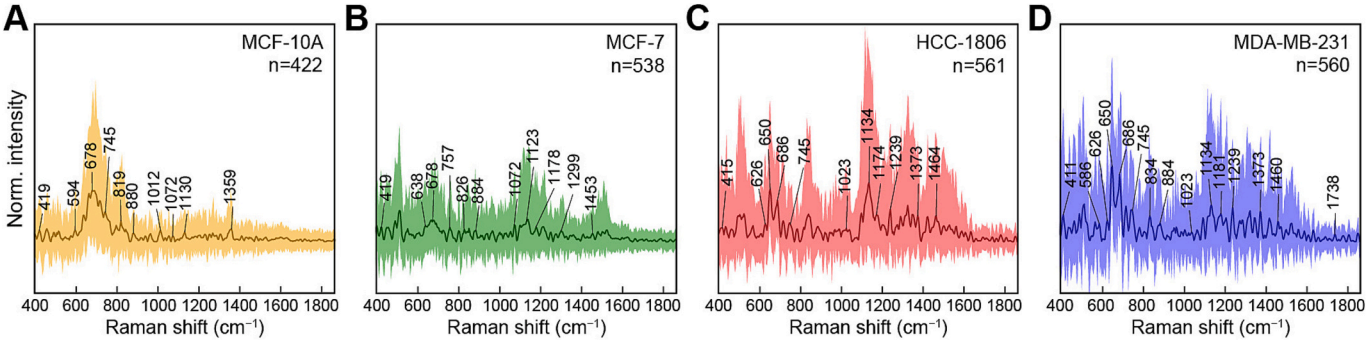



**Fig. 4.** Representative average SERS spectra of living cells with the 5th and 95th percentiles and biological characteristic peaks (A) MCF-10A, (B) MCF-7, (C) HCC-1806, and (D) MDA-MB-231. The numbers of spectra for each cell type are 422 for MCF-10A, 538 for MCF-7, 561 for HCC-1806, and 560 for MDA-MB-231.

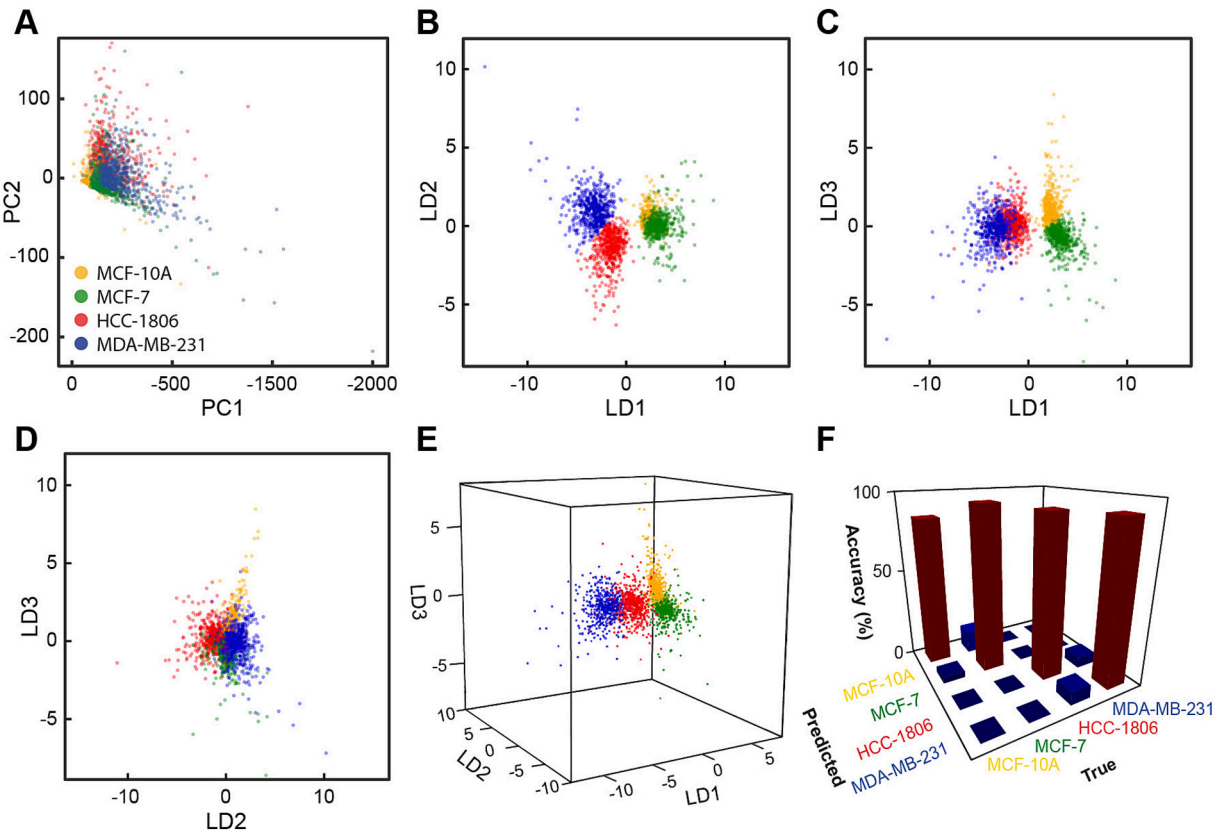



**Fig. 5.** Multivariate analysis of living cells using four different types of breast cancer cells. (A) PC score plot. (B-E) PCA-LDA score plots: (B) LD1-LD2, (C) LD1-LD3, (D) LD2-LD3, and (E) 3D plot along axis of LD1, LD2, and LD3. (F) Histograms of confusion matrix for PCA-LDA with LOOCV.

confusion matrix between the two MCF groups and the two TNBC groups further support their scatter overlap observed in Fig. 5E, likely due to their similar cellular expression. A high classification accuracy confirms that the 3D vertical nanostructure substrate successfully enabled a reliable label-free SERS for living cells and captured a large volume of rich biological information. This result supports the potential for multi-class label-free bioanalysis of living systems, as it demonstrates the ability to classify four different subtypes from the same organ, which is more challenging than classifying cells from intrinsically different organs.

## 4. Conclusion

Previously, the difficulty of reliably acquiring a large volume of SERS data from living cells posed a significant challenge for the multi-class classification. In this work, we have employed a 3D protruding nanostructure-based high-performance nanolaminate SERS substrate and obtained high-quality and large-volume SERS datasets via label-free living cell SERS mapping. The key enabling factor is the uniform arrangement of MIM building blocks over a large area with 3D vertical nanotopography, which induces a tight membrane-hotspot interface. We investigated the optical properties of nanolaminate SERS substrates, including broad plasmon resonances and gap-based high-field enhancements, which are beneficial for high-quality data acquisition. We conducted label-free high-throughput 2D SERS mapping for four different living breast cell types, including normal (MCF-10A) and cancerous (MCF-7, HCC-1806, and MDA-MB-231) cells. We acquired a large volume of datasets ($n > 2000$) containing characteristic biological signatures and employed PCA-LDA to extract key features and perform multi-class classification. We demonstrate an accuracy of 92.5 % with LOOCV, indicating that the conventional multivariate analysis method alone can effectively classify cancer subtypes as long as reliable and high-quality SERS datasets are provided. This result implies that the label-free living cell SERS platform can potentially be used for precise and rapid diagnosis of cancer subtypes, which may facilitate timely treatment and improved prognosis. Furthermore, future integration with advanced machine learning methods may enable more complicated cancer studies with external stimuli such as chemotherapy and resistance development.

## 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.

## Acknowledgements

This work was supported by the Global Joint Research Program funded by the Pukyong National University (202412300001).

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.saa.2025.127054.

## Data availability

Data will be made available on request.

## References

[1] X. Wang, S.-C. Huang, S. Hu, S. Yan, B. Ren, Fundamental understanding and applications of plasmon-enhanced Raman spectroscopy, Nat. Rev. Phys. 2 (2020) 253–271.

[2] S.-Y. Ding, J. Yi, J.-F. Li, B. Ren, D.-Y. Wu, R. Panneerselvam, Z.-Q. Tian, Nanostructure-based plasmon-enhanced Raman spectroscopy for surface analysis of materials, Nat. Rev. Mater. 1 (2016) 1–16.
[3] J. Kneipp, H. Kneipp, K. Kneipp, SERS—a single-molecule and nanoscale tool for bioanalytics, Chem. Soc. Rev. 37 (2008) 1052–1060.
[4] Y.C. Cao, R. Jin, C.A. Mirkin, Nanoparticles with Raman spectroscopic fingerprints for DNA and RNA detection, Science 297 (2002) 1536–1540.
[5] X. Pan, L. Li, H. Lin, J. Tan, H. Wang, M. Liao, C. Chen, B. Shan, Y. Chen, M. Li, A graphene oxide-gold nanostar hybrid based-paper biosensor for label-free SERS detection of serum bilirubin for diagnosis of jaundice, Biosens. Bioelectron. 145 (2019) 111713.
[6] H. Shin, B.H. Choi, O. Shim, J. Kim, Y. Park, S.K. Cho, H.K. Kim, Y. Choi, Single test-based diagnosis of multiple cancer types using exosome-SERS-AI for early stage cancers, Nat. Commun. 14 (2023) 1644.
[7] W. Premasiri, J. Lee, L. Ziegler, Surface-enhanced Raman scattering of whole human blood, blood plasma, and red blood cells: cellular processes and bioanalytical sensing, J. Phys. Chem. B 116 (2012) 9376–9386.
[8] S. Laing, L.E. Jamieson, K. Faulds, D. Graham, Surface-enhanced Raman spectroscopy for in vivo biosensing, Nat. Rev. Chem. 1 (2017) 0060.
[9] W. Nam, H. Chen, X. Ren, M. Agah, I. Kim, W. Zhou, Nanolaminate plasmonic substrates for high-throughput living cell SERS measurements and artificial neural network classification of cellular drug responses, ACS Appl. Nano Mater. 5 (2022) 10358–10368.
[10] C. Zong, M. Xu, L.-J. Xu, T. Wei, X. Ma, X.-S. Zheng, R. Hu, B. Ren, Surface-enhanced Raman spectroscopy for bioanalysis: reliability and challenges, Chem. Rev. 118 (2018) 4946–4980.
[11] A.F. Palonpon, J. Ando, H. Yamakoshi, K. Dodo, M. Sodeoka, S. Kawata, K. Fujita, Raman and SERS microscopy for molecular imaging of live cells, Nat. Protoc. 8 (2013) 677–692.
[12] J.W. Kang, P.T.C. So, R.R. Dasari, D.-K. Lim, High resolution live cell Raman imaging using subcellular organelle-targeting SERS-sensitive gold nanoparticles with highly narrow intra-nanogap, Nano Lett. 15 (2015) 1766–1772.
[13] E. Lenzi, M. Henriksen-Lacey, B. Molina, J. Langer, C.D. de Albuquerque, D. Jimenez de Aberasturi, L.M. Liz-Marzán, Combination of live cell surface-enhanced Raman scattering imaging with chemometrics to study intracellular nanoparticle dynamics, ACS sens. 7 (2022) 1747–1756.
[14] A. Kapara, V. Brunton, D. Graham, K. Faulds, Investigation of cellular uptake mechanism of functionalised gold nanoparticles into breast cancer using SERS, Chem. Sci. 11 (2020) 5819–5829.
[15] J. Taylor, A. Huefner, L. Li, J. Wingfield, S. Mahajan, Nanoparticles and intracellular applications of surface-enhanced Raman spectroscopy, Analyst 141 (2016) 5037–5055.
[16] A. Michałowska, A. Kudelski, Applications of surface enhanced Raman scattering (SERS) spectroscopy for detection of nucleic acids, Nanophotonics 13 (2024) 4577–4603.
[17] N. Guarrotxena, G.C. Bazan, Antitags: SERS-encoded nanoparticle assemblies that enable single-spot multiplex protein detection, Adv. Mater. 26 (2014) 1941–1946.
[18] J. Lv, S. Chang, X. Wang, Z. Zhou, B. Chen, R. Qian, D. Li, Live-cell profiling of membrane sialic acids by fluorescence imaging combined with SERS labelling, Sensors Actuators B Chem. 351 (2022) 130877.
[19] N. Choi, H. Dang, A. Das, M.S. Sim, I.Y. Chung, J. Choo, SERS biosensors for ultrasensitive detection of multiple biomarkers expressed in cancer cells, Biosens. Bioelectron. 164 (2020) 112326.
[20] S. Lee, H. Chon, J. Lee, J. Ko, B.H. Chung, D.W. Lim, J. Choo, Rapid and sensitive phenotypic marker detection on breast cancer cells using surface-enhanced Raman scattering (SERS) imaging, Biosens. Bioelectron. 51 (2014) 238–243.
[21] L. Hanson, W. Zhao, H.-Y. Lou, Z.C. Lin, S.W. Lee, P. Chowdary, Y. Cui, B. Cui, Vertical nanopillars for in situ probing of nuclear mechanics in adherent cells, Nat. Nanotechnol. 10 (2015) 554–562.
[22] M. Dipalo, H. Amin, L. Lovato, F. Moia, V. Caprettini, G.C. Messina, F. Tantussi, L. Berdondini, F. De Angelis, Intracellular and extracellular recording of spontaneous action potentials in mammalian neurons and cardiac cells with 3D plasmonic nanoelectrodes, Nano Lett. 17 (2017) 3932–3939.
[23] M. Dipalo, A.F. McGuire, H.-Y. Lou, V. Caprettini, G. Melle, G. Bruno, C. Lubrano, L. Matino, X. Li, F. De Angelis, Cells adhering to 3D vertical nanostructures: cell membrane reshaping without stable internalization, Nano Lett. 18 (2018) 6100–6105.
[24] H.-Y. Lou, W. Zhao, X. Li, L. Duan, A. Powers, M. Akamatsu, F. Santoro, A. F. McGuire, Y. Cui, D.G. Drubin, Membrane curvature underlies actin reorganization in response to nanoscale surface topography, Proc. Natl. Acad. Sci. 116 (2019) 23143–23151.
[25] P.L. Stiles, J.A. Dieringer, N.C. Shah, R.P. Van Duyne, Surface-enhanced Raman spectroscopy, Annu. Rev. Anal. Chem. 1 (2008) 601–626.
[26] S. Schlücker, Surface-enhanced raman spectroscopy: concepts and chemical applications, Angew. Chem. Int. Ed. 53 (2014) 4756–4795.
[27] O. Liang, P. Wang, M. Xia, C. Augello, F. Yang, G. Niu, H. Liu, Y.-H. Xie, Label-free distinction between p53+/+ and p53−/−colon cancer cells using a graphene based SERS platform, Biosens. Bioelectron. 118 (2018) 108–114.
[28] C. Zhang, S. Siddhanta, D. Paria, Y. Li, C. Zheng, I. Barman, Spectroscopy-assisted label-free molecular analysis of live cell surface with vertically aligned plasmonic nanopillars, Small 17 (2021) 2100161.
[29] J.-H. Choi, T.-H. Kim, W.A. El-Said, J.-H. Lee, L. Yang, B. Conley, J.-W. Choi, K.-B. Lee, In situ detection of neurotransmitters from stem cell-derived neural interface at the single-cell level via graphene-hybrid SERS nanobiosensing, Nano Lett. 20 (2020) 7670–7679.
[30] R. La Rocca, G.C. Messina, M. Dipalo, V. Shalabaeva, F. De Angelis, Out-of-plane Plasmonic antennas for Raman analysis in living cells, Small 11 (2015) 4632–4637.
[31] W. Nam, X. Ren, I. Kim, J. Strobl, M. Agah, W. Zhou, Plasmonically calibrated label-free surface-enhanced Raman spectroscopy for improved multivariate analysis of living cells in cancer subtyping and drug testing, Anal. Chem. 93 (2021) 4601–4610.
[32] J. Plou, P.S. Valera, I. García, D. Vila-Liarte, C. Renero-Lecuna, J. Ruiz-Cabello, A. Carracedo, L.M. Liz-Marzán, Machine learning-assisted high-throughput SERS classification of cell secretomes, Small 19 (2023) 2207658.
[33] W. Nam, X. Ren, S.A.S. Tali, P. Ghassemi, I. Kim, M. Agah, W. Zhou, Refractive-index-insensitive nanolaminated SERS substrates for label-free Raman profiling and classification of living cancer cells, Nano Lett. 19 (2019) 7273–7281.
[34] J. Song, W. Nam, W. Zhou, Scalable high-performance nanolaminated SERS substrates based on multistack vertically oriented plasmonic nanogaps, Advanced Materials Technologies 4 (2019) 1800689.
[35] W. Nam, Y. Zhao, J. Song, S. Ali Safiabadi Tali, S. Kang, W. Zhu, H.J. Lezec, A. Agrawal, P.J. Vikesland, W. Zhou, Plasmonic electronic raman scattering as internal standard for spatial and temporal calibration in quantitative surface-enhanced raman spectroscopy, The Journal of Physical Chemistry Letters 11 (2020) 9543–9551.
[36] E.C. Le Ru, E. Blackie, M. Meyer, P.G. Etchegoin, Surface enhanced Raman scattering enhancement factors: a comprehensive study, J. Phys. Chem. C 111 (2007) 13794–13803.
[37] J. Debnath, S.K. Muthuswamy, J.S. Brugge, Morphogenesis and oncogenesis of MCF-10A mammary epithelial acini grown in three-dimensional basement membrane cultures, Methods 30 (2003) 256–268.
[38] O. Crespo, C. Fernandez-Viadero, R. Verduga, V. Ovejero, S. Cos, Interaction between melatonin and estradiol on morphological and morphometric features of MCF-7 human breast cancer cells, J. Pineal Res. 16 (1994) 215–222.
[39] Y. Lombardo, A. Filipović, G. Molyneux, M. Periyasamy, G. Giamas, Y. Hu, P. S. Trivedi, J. Wang, E. Yagüe, L. Michel, Nicastrin regulates breast cancer stem cell properties and tumor growth in vitro and in vivo, Proc. Natl. Acad. Sci. 109 (2012) 16558–16563.
[40] R.M. Neve, K. Chin, J. Fridlyand, J. Yeh, F.L. Baehner, T. Fevr, L. Clark, N. Bayani, J.-P. Coppe, F. Tong, A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes, Cancer Cell 10 (2006) 515–527.
[41] S. Iyer, R.M. Gaikwad, V. Subba-Rao, C.D. Woodworth, I. Sokolov, Atomic force microscopy detects differences in the surface brush of normal and cancerous cells, Nat. Nanotechnol. 4 (2009) 389–393.
[42] I. Dagogo-Jack, A.T. Shaw, Tumour heterogeneity and resistance to cancer therapies, nature reviews, Clin. Oncol. 15 (2018) 81–94.
[43] K. Czamara, K. Majzner, M.Z. Pacia, K. Kochan, A. Kaczor, M. Baranska, Raman spectroscopy of lipids: a review, J. Raman Spectrosc. 46 (2015) 4–20.
[44] Z. Movasaghi, S. Rehman, I.U. Rehman, Raman spectroscopy of biological tissues, Appl. Spectrosc. Rev. 42 (2007) 493–541.
[45] L.A. Austin, B. Kang, M.A. El-Sayed, A new nanotechnology technique for determining drug efficacy using targeted plasmonically enhanced single cell imaging spectroscopy, J. Am. Chem. Soc. 135 (2013) 4688–4691.
[46] S. Contorno, R.E. Darienzo, R. Tannenbaum, Evaluation of aromatic amino acids as potential biomarkers in breast cancer by Raman spectroscopy analysis, Sci. Rep. 11 (2021) 1698.
[47] M. Greenacre, P.J.F. Groenen, T. Hastie, A.I. d'Enza, A. Markos, E. Tuzhilina, Principal component analysis, Nat. Rev. Meth. Prim. 2 (2022) 100.
[48] C.L.M. Morais, K.M.G. Lima, M. Singh, F.L. Martin, Tutorial: multivariate classification for vibrational spectroscopy in biological samples, Nat. Protoc. 15 (2020) 2143–2162.
[49] P. Pořízka, J. Klus, E. Képeš, D. Prochazka, D.W. Hahn, J. Kaiser, On the utilization of principal component analysis in laser-induced breakdown spectroscopy data analysis, a review, Spectrochim. Acta B At. Spectrosc. 148 (2018) 65–82.