NANO · MICRO small
www.small-journal.com

Machine Learning-Assisted High-Throughput SERS Classification of Cell Secretomes

Javier Plou, Pablo S. Valera, Isabel García, David Vila-Liarte, Carlos Renero-Lecuna, Jesús Ruiz-Cabello, Arkaitz Carracedo, and Luis M. Liz-Marzán**

Dedicated to Prof. M. Arturo López-Quintela on the occasion of his 70th birthday

During the response to different stress conditions, damaged cells react in multiple ways, including the release of a diverse cocktail of metabolites. Moreover, secretomes from dying cells can contribute to the effectiveness of anticancer therapies and can be exploited as predictive biomarkers. The nature of the stress and the resulting intracellular responses are key determinants of the secretome composition, but monitoring such processes remains technically arduous. Hence, there is growing interest in developing tools for noninvasive secretome screening. In this regard, it has been previously shown that the relative concentrations of relevant metabolites can be traced by surface-enhanced Raman scattering (SERS), thereby allowing label-free biofluid interrogation. However, conventional SERS approaches are insufficient to tackle the requirements imposed by high-throughput modalities, namely fast data acquisition and automatized analysis. Therefore, machine learning methods were implemented to identify cell secretome variations while extracting standard features for cell death classification. To this end, *ad hoc* microfluidic chips were devised, to readily conduct SERS measurements through a prototype relying on capillary pumps made of filter paper, which eventually would function as the SERS substrates. The developed strategy may pave the way toward a faster implementation of SERS into cell secretome classification, which can be extended even to laboratories lacking highly specialized facilities.

1. Introduction

The study of cell death has been traditionally underpinned by the development of imaging and monitoring technologies. The strides made over the past decades have revealed the multiple roles of cell death, which are not only associated with various pathophysiological processes but are also essential for the homeostatic regulation of the whole organism.[1] In this regard, extensive variability in cell death modalities has been reported. Despite such differences among mechanisms, it is now widely accepted that the processes orchestrating cell death can be classified into two general subgroups: regulated cell death (RCD) and accidental cell death (ACD).[2] Whereas programmed cell death pathways are precisely executed in RCD, generating a time-dependent response to stress (as occurring e.g. in caspase-mediated apoptosis), ACD is dominated by an instantaneous cell

J. Plou, P. S. Valera, I. García, D. Vila-Liarte, C. Renero-Lecuna, J. Ruiz-Cabello, L. M. Liz-Marzán
CIC biomaGUNE
Basque Research and Technology Alliance (BRTA)
Donostia-San Sebastián 20014, Spain
E-mail: jplou@cicbiomagune.es; luis.liz.marzan@uvigo.es

J. Plou, P. S. Valera, I. García, D. Vila-Liarte, C. Renero-Lecuna, L. M. Liz-Marzán
Biomedical Research Networking Center in Bioengineering Biomaterials, and Nanomedicine (CIBER-BBN)
Donostia-San Sebastián 20014, Spain

P. S. Valera, A. Carracedo
CIC bioGUNE
Basque Research and Technology Alliance (BRTA)
Derio 48160, Spain

P. S. Valera
Department of Applied Chemistry
University of the Basque Country
Donostia 20018, Spain

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/smll.202207658.

DOI: 10.1002/smll.202207658

J. Ruiz-Cabello, A. Carracedo, L. M. Liz-Marzán
IKERBASQUE
Basque Foundation for Science
Bilbao 48009, Spain

J. Ruiz-Cabello
Biomedical Research Networking Center in Respiratory Diseases (CIBERES)
Madrid 28029, Spain

J. Ruiz-Cabello
Universidad Complutense de Madrid
Madrid 28040, Spain

A. Carracedo
Biomedical Research Networking Center in Cancer (CIBERONC)
Derio 48160, Spain

A. Carracedo
Translational Prostate Cancer Research Lab
CIC bioGUNE-Basurto
Biocruces Bizkaia Health Research Institute
Derio 48160, Spain

L. M. Liz-Marzán
Cinbio
Universidade de Vigo
Vigo 36310, Spain

collapse due to unfavorable conditions (necrosis by temperature, pH variations, or severe oxidative damage).[3–5]

Recent studies have supported the notion that different soluble metabolites are released to the extracellular milieu upon the activation of cell death by a combination of regulated processes and passive diffusion through impaired membranes. This subset of molecules, termed secretome, varies among different types of cell death and can function as paracrine extracellular messages.[6] Hence, released metabolites transmit either danger signals –alerting the organism about a potential threat– or messages to promote cell proliferation and suppress inflammation within the tissue.[7,8] Noteworthily, the secretome of cancer cells plays a pivotal role in the fate of malignant tumors during treatment, transforming immunologically "cold" environments –those that limit immune activity– into "hot" environments, capable of stimulating antitumoral responses.[9] In turn, the accurate classification of cell secretomes could lead to better cancer treatments, which has stimulated the development of sophisticated analytical tools to readily identify such secreted signatures.[10–12]

Conventional methods for the characterization of cell death secretomes involve time-consuming and invasive analysis techniques (e.g., mass spectrometry).[13] Therefore, new strategies are needed that entail a fast, reliable, and straightforward determination of the different cell death types and their associated metabolic profiles. Surface-enhanced Raman scattering (SERS) has the potential to overcome current drawbacks in extracellular metabolomics, owing to its high sensitivity and fast data acquisition, which enables screening of the characteristic molecular fingerprint of complex media.[14,15] Indeed, we have recently demonstrated that multiplex SERS monitoring of metabolites in the extracellular milieu is feasible,[16] even observing specific profiles that could be correlated with dying cells. Notwithstanding, this approach only focused on specific intensity fluctuations at sharply defined peaks (particularly at 727 cm^{-1}, associated with Hypoxanthine), so additional variations in the SERS spectra were not sufficiently investigated. Previous studies have otherwise used SERS to track intracellular changes during cell death, identifying peaks associated with protein denaturation, amino acid residues, changes in the protein conformation, or lipid degradation.[17,18] However, the initial uptake of nanoparticles (NPs) by the cells (required for intracellular monitoring) may affect their behavior, while ignoring the information contained in the extracellular milieu.[19,20] Overall, although SERS promises to speed up and simplify the task of metabolic profiling, the development of efficient protocols in synergy with advanced data processing is still required for robust cell death monitoring.[21,22]

Herein, we employ label-free SERS monitoring of cell supernatants to demonstrate that the secretome in dying cells is influenced by the underlying stresses. In contrast to previous studies, an exhaustive analysis of the recorded spectra allowed us to capture the diversity of the cell death secretome, while extracting common features among similar conditions by machine learning (ML) techniques. Furthermore, with the aim of performing high-throughput cytotoxic studies, we devised a strategy that allows us to accelerate the collection of cell supernatants and acquire their respective SERS spectra, thereby minimizing external operation. This customized strategy was subsequently utilized as a proof of concept to monitor multiple stress conditions on cell cultures. The obtained results foster our methodology as a powerful tool for fast and accurate classification of the secreted signatures.

2. Results and Discussion

2.1. Secretome Monitoring by SERS

To induce the release of a distinct subset of metabolites, we challenged HeLa cells with a panel of different cell death-inducing conditions. We initially selected ultraviolet radiation (UV), heating at 55 °C, and hypotonic media, owing to their well-known properties as strong stressors.[23,24] Specifically, UV exposure leads to the formation of DNA lesions that activate death pathways; heating and hypotonic stress, by contrast, may cause a direct disruption of other cellular structures, such as the cell membrane or protein denaturation. In addition, other cell assays were maintained without disturbances as the control group.

To determine the metabolic fingerprint in the extracellular environment, we collected the supernatant of HeLa cells after 4 h of treatment, which ensured that the stress response would be triggered, and spiked the SERS substrate prior to the measurements (Figure S1a, Supporting Information). We opted for a plasmonic substrate created by drop-casting Ag colloidal nanoparticles (NPs) (Figure S1b, Supporting Information) on aluminum film supports, on account of their simple preparation and homogeneity (Figure S1c, Supporting Information). For each experiment, we collected 400 spectra at different locations of the same SERS substrate and carried out three independent biological replications per condition. The average of the registered spectra is displayed in **Figure 1**a, showing distinct vibrational fingerprints for each condition. The corresponding stacks of SERS spectra are plotted in Figure 1b, so that the distribution of samples is shown in detail. The acquired SERS profiles exhibit inevitable variability among biological repetitions, as well as stochastic noise within the same label. To better visualize and analyze this variability, a linear principal component analysis (PCA) was conducted, which examines the full spectrum for each condition. The results of this analysis are presented in Figure 1c. We can observe the formation of defined clusters for each class, which was particularly clear for the spectra acquired as control. Still, we noticed that some conditions present minor overlap. This is a consequence of the non-linearity consistently present in complex SERS spectra; for example, we must consider that one metabolite might exhibit multiple peak shifts, according to its orientation toward the plasmonic surface.[25]

Apart from the variability among conditions, we examined whether SERS could monitor the changes in secreted metabolites over time, i.e., the evolution of the system upon stress. In previous experiments, all measurements were recorded from supernatants after 4 h. However, the frequency of death-specific molecules is highly dependent on the selected time point. As shown in Figure 1d, we carried out a multivariate spectral analysis by means of other unsupervised methods typically used to promote clustering (in this case, nonlinear t-distributed

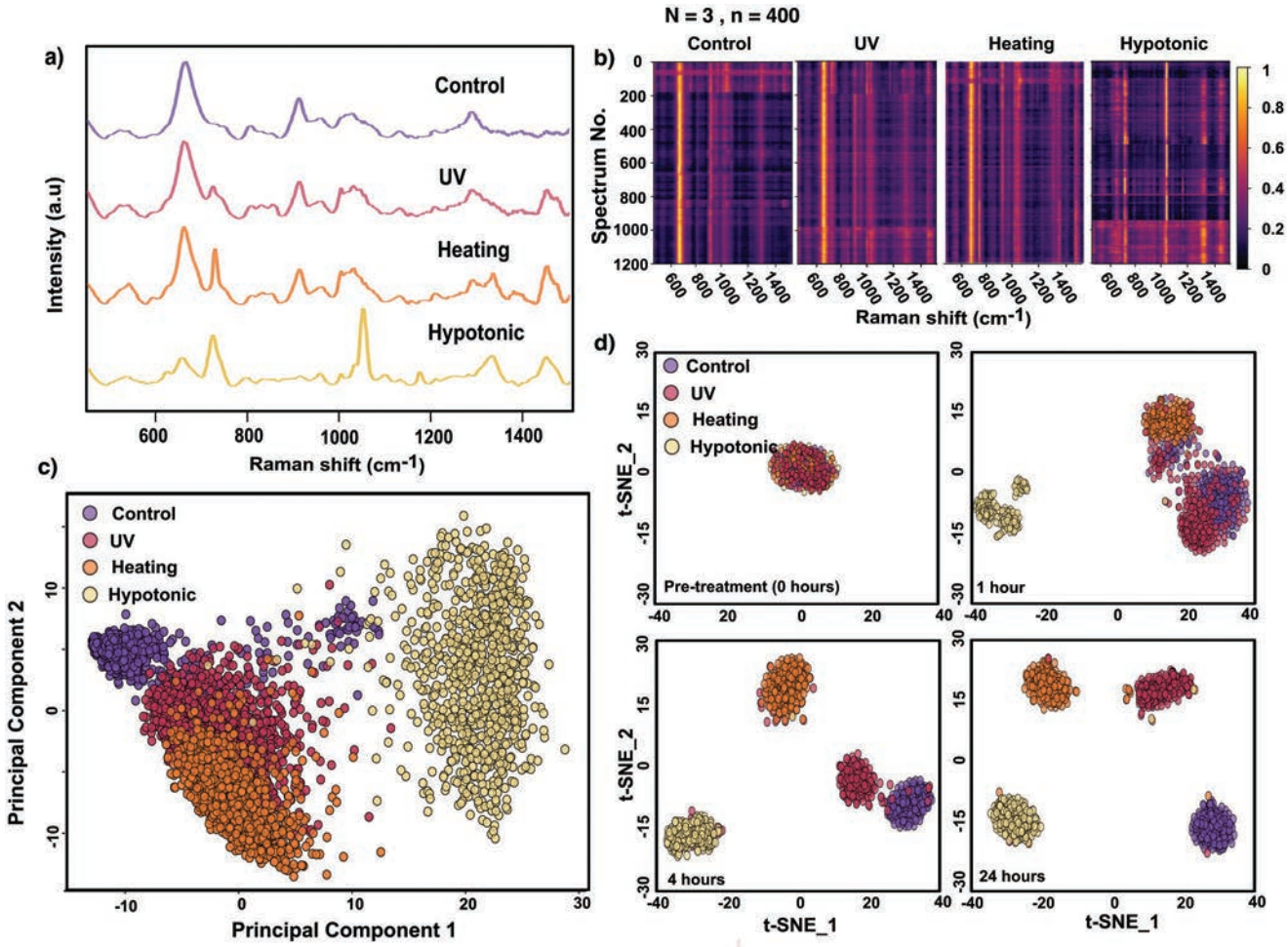

Figure 1. a) Average SERS spectra recorded from supernatants of HeLa cells challenged under different stress conditions, named as Control, UV, Heating, and Hypotonic (see main text for description). b) Heatmaps displaying 1200 spectra ($N = 3$, $n = 400$) collected for each of the 4 monitored conditions. The color map indicates SERS intensity, normalized from 0 to 1. c) Principal component analysis (PCA) score plot of the first two principal components (PC1 – 80.5%, PC2 – 19.4%), showing the relative secretome separation after 4 h of stress stimulus. d) t-distributed stochastic neighbor embedding (t-SNE) analysis of the spectra acquired at successive times: pre-treatment (0 h), 1 h, 4 h, and 24 h of experiment, as labeled. Cluster formation can be visualized over time.

stochastic neighbor embedding, or t-SNE, which better preserves the internal structure of the data than PCA).[25] As expected, at initial stages the spectra were very similar among groups, albeit samples from the hypotonic class could be quickly distinguished from the others. The hypotonic solution presents a lower concentration of salts and metabolites compared with standard media, which resulted – not surprisingly- in a distinct SERS profile. It should also be noted that the plotted spectra offer meaningful information on the timing of cell death events. For example, a clear separation was observed after 1 h of heating – high temperatures cause immediate damage to exposed cells (ACD). In contrast, UV illumination activates mechanisms that trigger slower dying processes (RCD), so the registered spectra barely differed from the control group at the initial stages (see Figure 1d,h). The data retrieved after 24 h resulted in 4 well-separated clusters, supporting the notion of specific SERS signatures for each condition.

2.2. Classification of SERS-Recorded Secretomes by Supported Vector Machine Analysis

Following recent developments in machine-learning (ML) techniques for SERS monitoring, we investigated whether our SERS platform could classify different cell death mechanisms according to the recorded secretomes. As depicted in **Figure 2**a, we initially deployed a supported vector machine (SVM) architecture to identify unique patterns in the secretome of dying cells. In general, the main idea of the supervised SVM method is to disclose the linear decision surface that better separates the different classes of spectra.[26] Thereby, we sought to predict the cytotoxic impact (live/dead) of different conditions by SERS. To create a robust and stable SVM model capable of dealing with complex SERS spectra, we utilized the data collected in the previous section (5500 spectra from 4 different conditions, acquired at 4 different time points). Importantly, in such a supervised analysis, we had to initially add informative

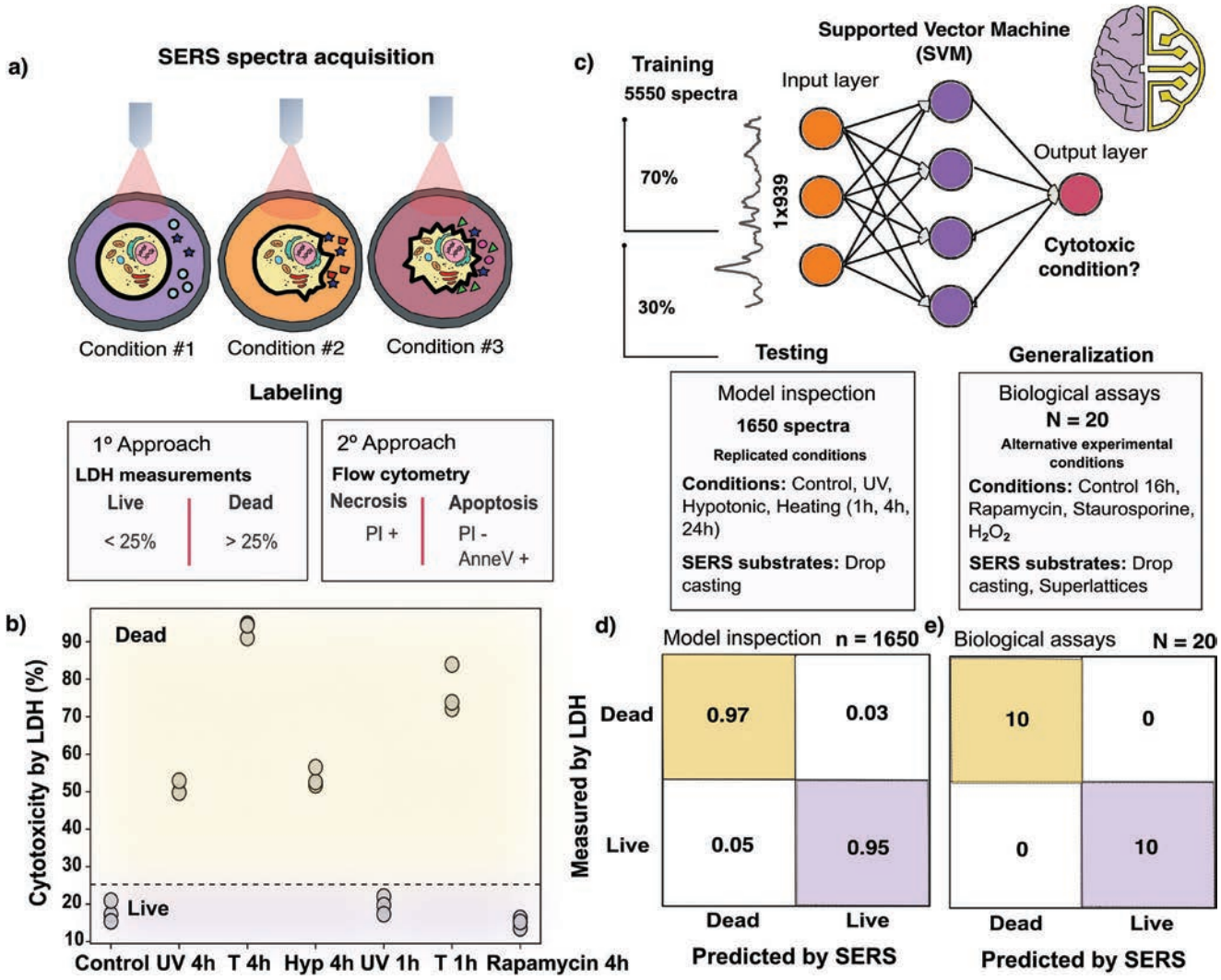

Figure 2. a) Scheme depicting the workflow of our machine learning (SVM) classification method strategy. Different cell conditions are initially labeled by different viability assays, as live or dead. Besides, among dying cells, we differentiate between necrosis/apoptosis cell death. b) LDH assay results for representative cell conditions. On the basis of LDH readouts for control cells, the threshold that determines data labeling was established at 25%, with confidence of $p = 0.99$ (Hyp: Hypotonic). c) The SVM is trained with 70% of the data, previously analyzed in Figure 1 ($n = 5550$), and then validated with the remaining 30% to inspect the model. Finally, the generalization of the model was assessed with data from additional stress conditions. d) Confusion matrix for the live/dead classification based on the data that were not used in the training step (model inspection dataset, $n = 1650$). The figures in each box are normalized by the number of elements in each class. e) Confusion matrix showing the generalization of the trained network to spectra collected from different biological assays ($N = 20$).

labels to the raw data so that our ML model could learn therefrom. For this specific approach, data labeling was a function of the outcomes in the Lactate dehydrogenase (LDH) assay, measuring the percentage of induced cell death at each condition. Considering the readouts in Figure 2b (and Figure S2, Supporting Information), we labeled as viable (live) those conditions that showed LDH values <25% while noting as cytotoxic (dead) those above the same threshold (see the Experimental section for details on threshold determination). In this manner, SERS spectra from conditions retrieving an LDH value >25%, e.g., UV illumination after 4 h, were labeled as dead.

We randomly split the 5500 spectra into the training and testing (70% – 30%) data subsets (see Figure S3, Supporting Information, for information on datasets). The resulting 70% subset was fed into the model during the so-called training phase, to learn and connect the specific SERS signatures with their defined output variables (live and dead, see the sketch in Figure 2c). The created machine learning-based model was next tested with the remaining 30% of the data – inspection dataset, not previously used in the training phase, and displaying the retrieved predictions in the confusion matrix of Figure 2d. This type of table is routinely used to visualize and summarize the performance of classification models. Specifically, the rows present the actual classes (measured by LDH), whereas the columns dictate the predictions by the model (outputs of the SVM). The number of correctly classified spectra is given in the diagonal boxes (colored boxes), and thus the off-diagonal elements are those that resulted mislabeled (all the data in this

matrix were normalized by the number of elements per class). Overall, the network could recognize the SERS spectra satisfactorily from either living or dying cells, with a sensitivity of 97%, a specificity of 95%, and a total validation accuracy of 96%.

It should be noted that, the model must also establish connections between spectra and labels, to the extent that it enables a generalization to other stress conditions –not seen during the training step. The created model was thus evaluated with unused spectra from cells exposed to alternative biological assays, as depicted in Figure S3, Supporting Information. At this generalization step, we challenged HeLa cells with different stresses and drugs, including Rapamycin ($N = 3$), Staurosporine ($N = 5$), and H_2O_2 (N = 5). In addition, we measured SERS spectra at an additional time point, 16 h, for a further control ($N = 2$), which along with control replicates at 4 h ($N = 5$), raised the total number in the biological assay dataset to 20 independent experiments. Besides, some of the spectra were also recorded on different SERS substrates –so-called plasmonic superlattices (details in the Experimental Section and Figure S4a, Supporting Information), to thoroughly assess the general application of the method. Upon collecting the corresponding SERS measurements, we fed the model with one averaged spectrum per experiment (representative average spectra shown in Figure S4b, Supporting Information). From the resulting confusion matrix in Figure 2e, we could confirm that the model also achieves a high accuracy in classifying new conditions, i.e., regardless of the imposed stress the model could differentiate secretomes from living and dead cells. It should be specially noted that, for the experiments with Rapamycin, a drug affecting proliferation but not viability (Figure S4c,d, Supporting Information, viability studies and labeling of new conditions), the model correctly classified their secretome as still pertaining to living cells, meaning that it can distinguish the effects of cell death from cell growth inhibition.

Once the general discrimination between live/dead conditions was probed, we addressed the classification of different cell death types, particularly necrosis and apoptosis. To do so, we relabeled the data according to the outcomes of flow cytometry (**Figure 3**a). As also observed in Figure S5, Supporting Information, necrotic cell death exhibited a high percentage of Propidium Iodide positive cells (PI+), whereas apoptotic cells could be identified by labeling external phosphatidyl serine (PS) with Annexin V. On this account, the percentage of cells only marked with Annexin V (PI-/+Annexin V) was used to establish the cell death mechanism for each condition (threshold definition explained in the Experimental section).[27] The model was further trained with 70% of the dataset and their newly registered labels (apoptosis/necrosis). Such a training run enabled the identification of the most obvious parameters for each class until the model reached a level of accuracy of 95%, as detailed in Figure 3b. We then plotted the coefficient scores in Figure 3c, which indicated the most important variables (wavenumbers) for the separation between heating (necrosis) and UV (apoptosis) classes in the created SVM model, and therefore it was expected to offer valuable information on the secretome variability among them. It should be noticed that the principal vibrations of the adenosine triphosphate (ATP) metabolite, unlike other common metabolites in cell secretomes (Figure 3d), closely matched the wavenumbers with the highest scores for heating in Figure 3c – 730 cm^{-1} and 1334 cm^{-1}. Hence, ATP molecules and related metabolites are key components of the resulting secretome upon heating, in contrast with other measured conditions.[28] Furthermore, the commercial bioluminescent assay of Figure 3e revealed a peak release of ATP in the Heating class, even after only 15 min, whereas its concentration in the Control and UV groups remained stable over time.

Following the identification of potential distinctive metabolites in secretomes from apoptosis and necrosis, we tested the model with the same dataset of Figure 2c, labeling H_2O_2 ($N = 5$) as necrosis and Staurosporine ($N = 5$) as apoptosis. Although we see one misclassification event, the created model unveiled the state of evaluated cell conditions (confusion matrix in Figure 3f), thereby confirming the automatic recognition of cytotoxic mechanisms from SERS spectra.

2.3. Design of SERS Detection on Paper Microfluidic Biosensors

A major shortcoming in the implemented protocol was the time consumed prior to SERS measurements, which impaired high-throughput assays: on one hand, preparing the SERS substrates by drop casting takes time (>1 h); on the other hand, the supernatants to be spiked onto the SERS substrates had to be collected by hand. Additional time was again required for cell media to evaporate, prior to acquisition of the spectra. Even though this protocol was relatively straightforward, it still required several hours and was greatly limited by the fabrication of SERS substrates in advance (see Figure S6, Supporting Information). Considering all of the above, we aimed at refining the technology for rapid SERS analysis compatible with high throughput operation. We thus developed an approach based on polydimethylsiloxane (PDMS) microfluidic devices (for small-volume control) powered by capillary paper pumps, thereby achieving a significant reduction in waiting times (from hours to minutes). In the configuration illustrated in **Figure 4**a, growing cells are attached at the bottom of one microfluidic inlet, filling the flowing circuit with the media required for cell growth. At the selected measurement time, the NPs required for SERS interrogation are incorporated from a second inlet. These devices require the use of a paper pump,[29] which plays a dual role: i) it drives the flow of components in both chambers by capillarity, working as a wick that starts the flow and facilitates the mixing of components along the central channel (purposely designed for mixing, see Figure S6b, Supporting Information); ii) the filter paper is eventually leveraged to function (upon NP accumulation) as the plasmonic substrate from which SERS signals are recorded. Hence, such paper-based substrates, whose utility and simplicity have been described in previous studies,[29–31] were generated by the progressive accumulation of flowing NPs on the paper fibers (see Figure S7 and Movie S1, Supporting Information). It should be stressed that, striking differences in the arrangement of nanoparticles and resulting signal enhancement were observed between substrates generated by the above method and those created by conventional direct micropipetting, as illustrated in Figure S8, Supporting Information.

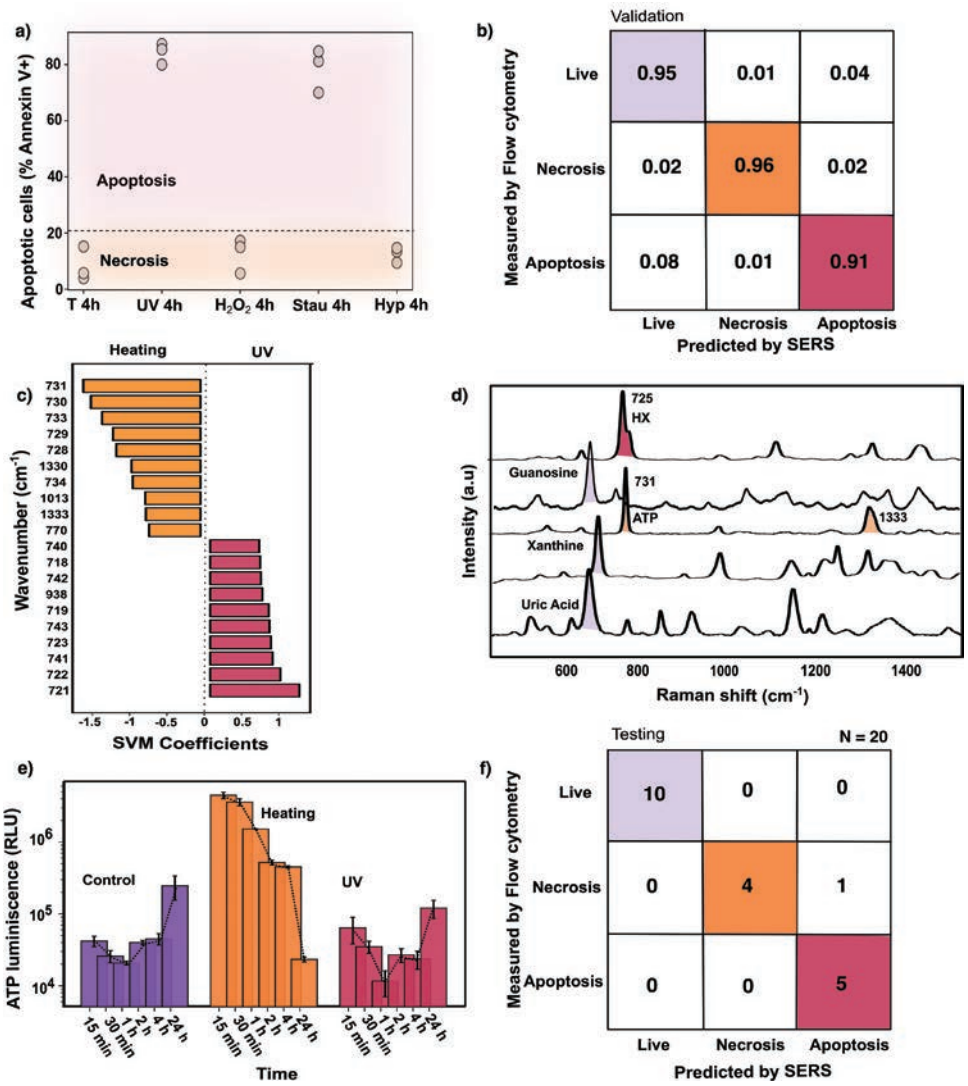

Figure 3. a) Flow cytometry results, showing the percentage of Annexin V-positive populations, for representative cell conditions. The presence of cells displaying Annexin V-positive (Annexin V +) and PI-negative cells (PI -) staining was employed for the identification of cell death mechanisms as apoptotic or necrotic profiles (Stau – Staurosporine, Hyp – Hypotonic). b) Confusion matrix for the live/necrosis/apoptosis SVM classification with 30% of the dataset, normalized to the number of elements per class. c) Coefficients of the 10 most important features (wavenumbers, cm^{-1}) for each class (UV and Heating) in the deployed SVM model. d) SERS spectra corresponding to metabolites with potential roles in cell death. e) Quantification of extracellular ATP by luminescence over time ($N = 3$). f) Confusion matrix for the multiplexing classification with the testing dataset ($N = 20$).

In addition, we found that the selection of paper type entails a critical step to grant homogeneous substrate formation toward reliable SERS data acquisition. A detailed study of substrate formation with different types of commercial papers and NPs is summarized in Figure 4b. We selected paper types on the basis of their chemical composition and pore size differences, which in combination determined the capillary diffusion rate. The employed NPs varied in size and composition (Au or Ag NPs) to identify the most convenient features. Three different patterns could be distinguished, which are marked with frames of different colors, namely "dispersed" (orange), "optimally filled" (violet), and "overpacked" (red). These parameters of capillary flow strongly shaped the substrate general appearance, which also affects resulting localized surface plasmon resonance (LSPR). For instance, a flow rate of 0.34 mm s^{-1} rendered spots with a diameter of 0.1 cm that was optimally filled with nanoparticles and matched the area of the microfluidic outlet in contact with the paper, as shown in Figure 4c. We additionally noted that the use of bigger NP sizes (60 nm diameter) resulted in more homogenous substrates for both AuNPs and AgNPs, except for the lowest flow rates, in which a state of over-agglomeration was reached (red frames). Therefore, we selected the combination of 60 nm Ag nanoparticles and a capillary speed of 0.34 mm s^{-1} as the optimized conditions to fabricate our SERS-on-paper sensors. Malachite Green (MG) was used as a standard probe molecule to evaluate the performance of these substrates in a sensing application. As expected, the SERS spectra in Figure 4c exhibit a clear correspondence between acquired average intensity and type of substrate.

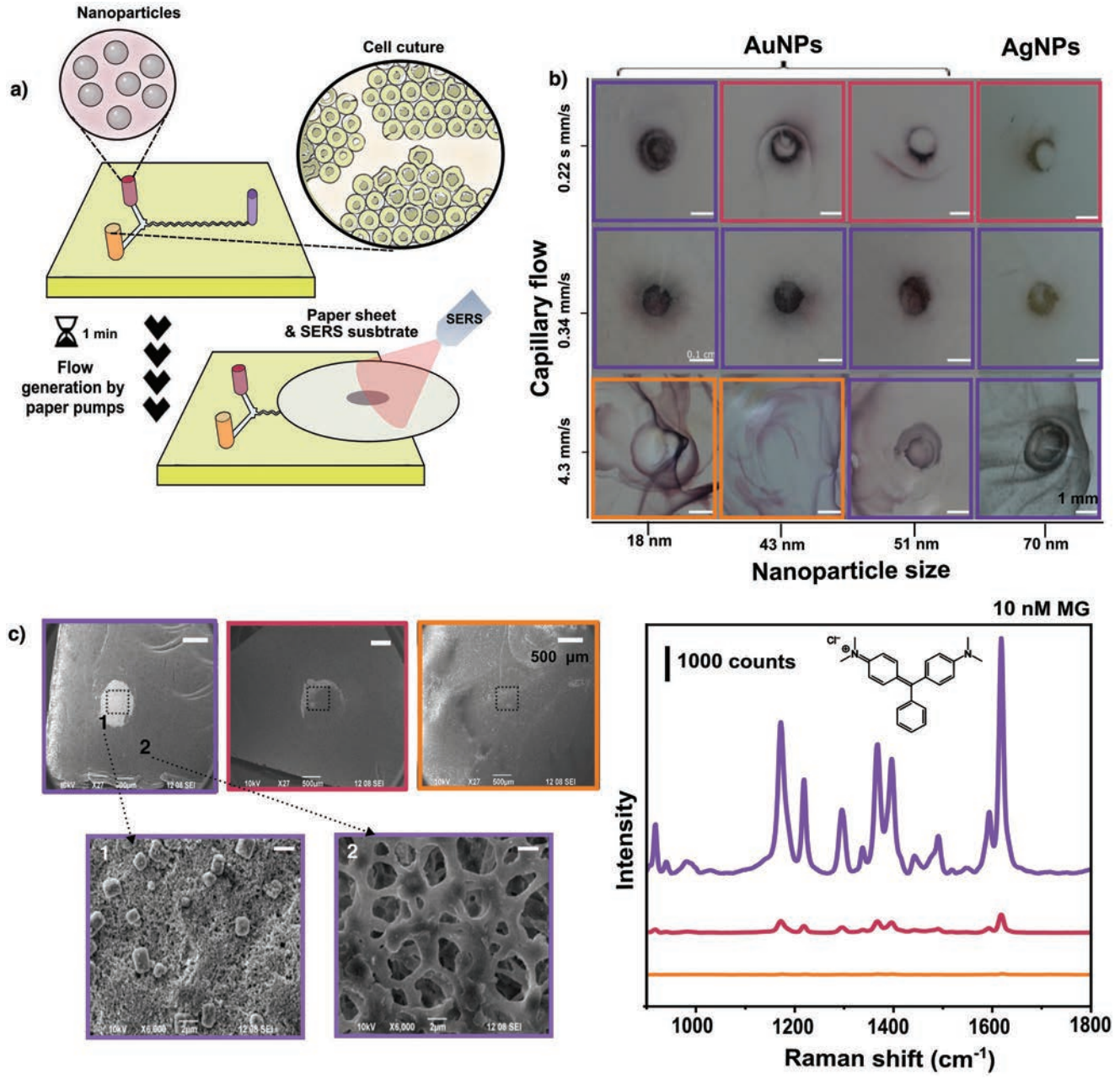

Figure 4. a) Scheme of the custom-made device with one inlet for cell culture and another for addition of NPs. The double role of the paper sheet (as the pump and support for nanoparticles), speeds up sample collection and SERS substrate preparation, thus reducing waiting time down to ca. 1 min. b) Influence of paper capillary action and NP diameter on SERS substrate formation. In general, three different substrate qualities can be distinguished. Higher speed yielded disparate substrates (orange boxes), whereas lower velocities led to more homogenous SERS substrates, created around the area in contact with the microfluidic outlet (violet boxes). Nanoparticles with larger sizes present restricted diffusion through the paper, resulting in a dense accumulation of NPs over heterogeneous small areas (red boxes). c) Representative SEM images of substrates with different nanoparticle distributions and their corresponding SERS spectra averaged over the indicated squared area, upon incubation with 10 nM Malachite Green (MG). Higher magnification images show general features of paper with (1) or without (2) NP accumulation.

2.4. High-Throughput SERS Screening Assays

We finally interrogated our SVM models with spectra obtained from cells cultured in these microfluidic chips. We started by culturing HeLa cells in microfluidic devices with dimensions specified in **Figure 5**a, and we subsequently exposed such cells to either control or cytotoxic conditions (detergent-based lysis buffer, L-chip). The results of LDH viability tests for cells growing in the customized chips (C_Chip) were found to be in the same range as those for cells cultured in conventional 96 well-plates (C), as can be seen in Figure 5b. In the next step (4 h later), Ag NPs were flown into the chips, so as to create

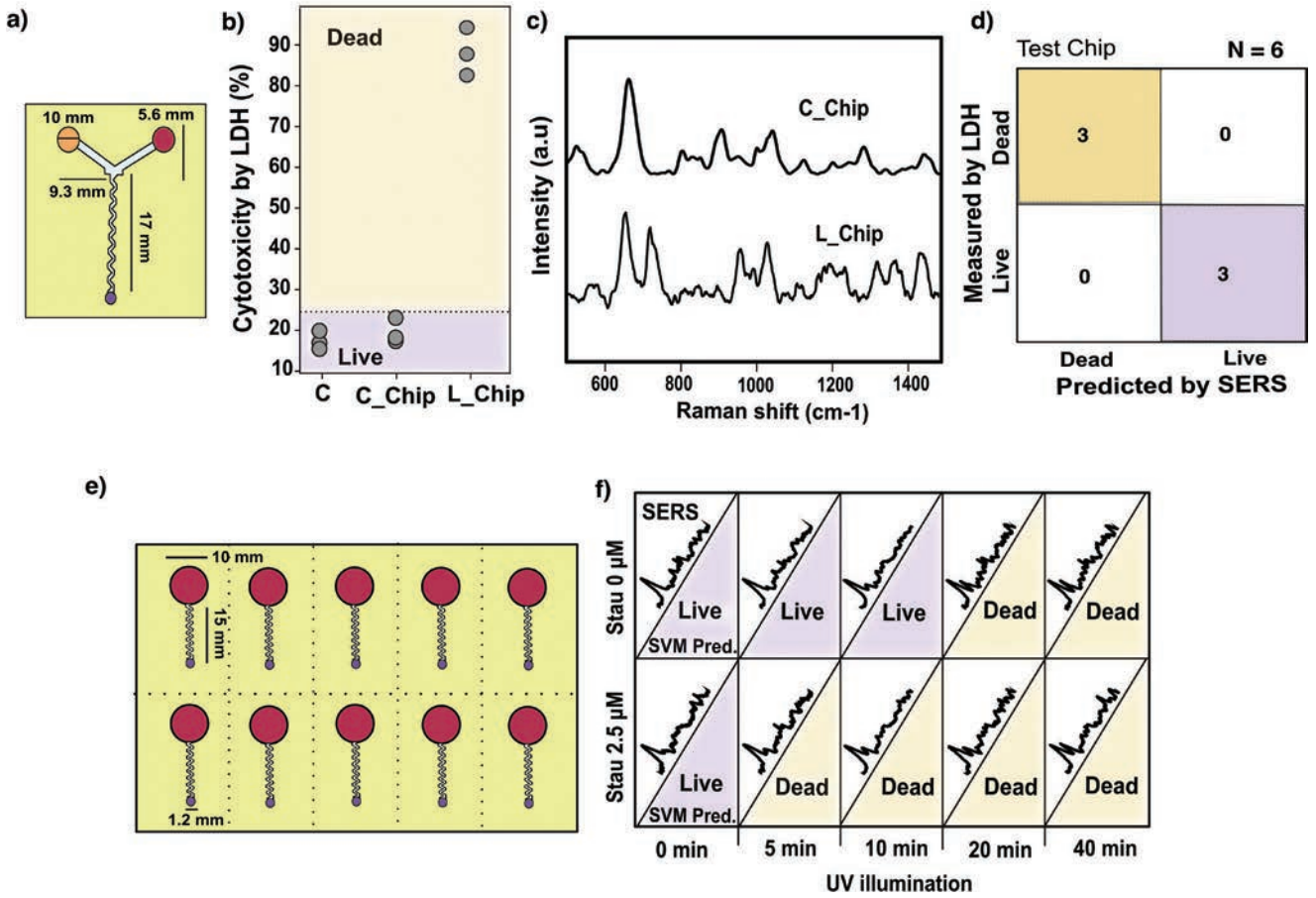

Figure 5. a) Scheme of the two-inlet device, depicting its dimensions. b) LDH results of viability assays for cells growing on chips, termed C_Chip, in comparison with cells cultured on 96 well plates, C, and with an injection of a commercial Lysis buffer into the Chips, L_Chip. c) Representative average SERS spectra registered upon 633 nm illumination on paper sensors. d) Confusion matrix displaying the predictions of the SVM model when spectra collected from C_chips and L-chips were loaded in the ML network. e) Multiwell device and corresponding dimensions. The larger inlet in red is leveraged for cell culture and AgNP administration, whereas the violet outlet served to draw the mixture of metabolites and AgNPs, due to the action of capillary paper pumps. f) Matrix representation of the SERS spectra acquired for different stress conditions (UV + Staurosporine) and the prediction provided for each case by the trained ML model (SVM Pred.).

SERS substrates on the filter paper wetted by the solution drawn from the microfluidic circuits –containing a mixture of AgNPs and metabolites from the cell supernatant. The spectra collected under 633 nm laser illumination (Figure 5c) were eventually introduced in our previously trained SVM model, which could predict the cell state from the examined secretomes. As illustrated in Figure 5d, the developed strategy yielded similar accuracies to those obtained in the above sections, i.e., it was able to identify dead and living cells, while helping streamline the data acquisition process.

In addition, thanks to the implementation of a microfabrication protocol that integrates 3D stereolithography (SLA) printing and PDMS soft lithography (see Experimental Section),[32] we could generate a wide variety of microfluidic chips, tailored to our specific requirements of cell culture and SERS substrate formation. A high-throughput device was devised (Figure 5e) with 10 circular wells that enabled cell seeding and adequate flow, induced by external capillary pumps (see Figure S9, Supporting Information, for the 3D printed mold). It should be noted otherwise that, in this approach, AgNPs are directly injected into the same well where the cells are growing (central chamber, 1 cm diameter). Although this design may lead to a higher variability of the SERS signal, considering that the mixing of both components (AgNPs and supernatants) is less controlled than in the two inlets strategy of Figure 5a, it was also less prone to form bubbles along the channels, thereby facilitating a successful implementation for multiassay experiments. To exploit the applicability of such multiwell devices in high-throughput screening assays, we evaluated different stress conditions and their impact on cell viability after 4 h (Figure 5f). Hence, we challenged the cells growing on each well with two different apoptotic triggers: UV illumination (varying exposition times) and Staurosporine 2.5 µM. Once the secretomes were measured on the paper-based substrates, the recorded SERS spectra were analyzed by our SVM framework, and proved to accurately predict cell viability; the retrieved outputs along with the corresponding input spectra are shown in Figure 5f. We could observe how the combination of both treatments elicited higher cytotoxicity, to an extent that the spectra from multimodal treatments of UV illumination for 5 and 10 min,

and Staurosporine, were classified as dying cells, in contrast to the separate therapies (see Figure S10, Supporting Information, for retrieved full spectra). We additionally explored the general application of this ML-assisted high-throughput method to analyze cell death in other cancer cell lines with diverse tissue origin. For this screening trial, we challenged prostate cancer (PC3) and human fibrosarcoma (HT-1080) cell lines with UV (20 min, 40 min) and compared the retrieved spectra with those obtained under control conditions (Figure S11a, Supporting Information). The results provided in the matrix of Figure S11b, Supporting Information, clearly confirmed the capability of the trained SVM model to predict the state of different cell types, not being restricted to standard HeLa cells, thereby promising a broad generalization of the method. Overall, despite representing a simple embodiment of potential high-throughput assays, our results demonstrate that the combination of paper-based sensors (created through microfluidic volume control) and ML automatized analysis considerably improves the accessibility and readability of SERS screening assays.

3. Conclusions

In connection with the need for approaches that can trace information contained in extracellular secretomes while avoiding time-consuming protocols, we have introduced the combination of label-free SERS monitoring and microfluidics together with ML analysis, to unveil specific biochemical signatures in tumor secretomes, even deciphering time-dependent variations. A key feature of our observations was that a trained machine learning network (SVM) was able to recognize shared signatures among different conditions, so that accurate classification could be performed according to similar cell states, e.g., identifying living and dying cells. It should be stressed that the high accuracy achieved in the model (96%) demonstrates the efficacy of SERS for assessing changes at the extracellular metabolic level. Moreover, we purposely devised PDMS microfluid chips, based on 3D SLA printing and PDMS soft lithography, to accelerate the collection of cell supernatants and subsequent acquisition of SERS spectra. The integration of filter paper that drives liquid flow along microfluidic channels, while sustaining the accumulation of NPs, resulted in a cornerstone to simplify protocols, on the way toward high-throughput implementation.

4. Experimental Section

Materials: Aluminum foil (ALUGRAM RP-18W), $HAuCl_4{\cdot}3\ H_2O$ (≥99.9%, trace metal basis) was purchased from Alfa Aesar. Glassware from Menzel-Gläser 24×24 #1. Trisodium citrate and silver nitrate ($AgNO_3$) were purchased by Sigma–Aldrich. Nitrocellulose membranes working as paper pumps were purchased from Sartorius (NC180; NC95, capillary flow 0.22 and 0.34 m s^{-1} respectively), whereas PES membrane discs were obtained from PALL (capillary flow; 4.3 m s^{-1}). Dead Cell Apoptosis Kits with Annexin V for Flow Cytometry of Thermo Fisher and LDH Cytotoxicity Assay Kit were from Thermo Scientific. Standard Photopolymer resin by Elegoo, SYLGARD 184 Silicone Elastomer Kit of Dow, and enamel acrylic spray was purchased from Pintyplus Evolution. Hydrogen peroxide (H_2O_2, 28%), Staurosporine, and Rapamycin were purchased from Sigma-Aldrich. All solutions, except $HAuCl_4$ and $AgNO_3$, were prepared immediately before use. Purified Milli-Q water was used in all experiments (Millipore, 18.2 MΩ cm). Glassware was from Menzel-Gläser. Commercial metabolites and probe molecules: Malachite Green chloride (MG). Staurosporine (Stau), Rapamycin, Adenosine Triphosphate (ATP), Hypoxanthine, Guanosine, Xanthine, and Uric acid were purchased from Sigma-Aldrich

Synthesis of Citrate-Stabilized Nanoparticles: AuNPs with varying sizes (18 nm, 43 nm, 51 nm) were synthesized according to a previously reported seeded growth method.[33] Representative transmission electron microscopy (TEM) images and UV-Vis spectra can be found in Figure S7, Supporting Information. AgNPs were synthesized following the well-established Lee-Meisel method,[34] comprising the reduction of $AgNO_3$ with sodium citrate in an aqueous solution. Again, representative TEM images and UV-Vis spectra are included in the supplementary information.

Fabrication of Plasmonic SERS Substrates: *Drop-casting*: To produce substrates with a random distribution of AgNPs, a 1.4 mM dispersion was employed. Subsequently, 5 µL of the Ag concentrated colloid (centrifugation –1520g, 10 min) was spiked onto aluminum sheets. Once dried under room temperature (≈1 h), 2 µL of selected supernatants were deposited onto freshly made SERS substrates, and let dried until complete evaporation of the solvent, prior to SERS measurements.

Plasmonic superlattices: we followed the protocol established by Matricardi et al.[35] Briefly, a 2 µL droplet of nanoparticle dispersion (PEG-AuNPs)[36] was placed on a nanostructured PDMS stamp. After 40 s, a glass slide was laid down on top of the droplet. After 2 h, the glass slide was carefully removed from the PDMS template, transferring the plasmonic superlattice onto the glass substrate with the inverse structure of the PDMS mold.

Sensors on paper: As a first step, metal nanoparticles (AgNPs or AuNPs, 1.4 mM) were added at the desired moment into the microfluidics inlets, specifically devised for that role. Afterward, nitrocellulose paper pumps (NC95 baked type) were placed over the chip, controlling carefully the contact of the paper with the outlet channel of the chip. As a result, the paper pumps drawn the liquid contained in the circuit by capillary action, carrying nanoparticles and metabolites along the zigzag mixing channel. In turn, the accumulation of nanoparticles occurred on the surface of the paper pumps. Once the paper is completely wet, and therefore the flow inside the chip stops, the paper is lifted from the device and transported to the SERS instrument for collection of the spectra (see Movie S1, Supporting Information).

Cell Culture and Stress Conditions: HeLa (Cervical cancer), PC3 (Prostate cancer), HT-1080 (Fibrosarcoma) cells were routinely cultured in Dulbecco's modified Eagle medium (DMEM) supplemented with 10% fetal Bovine serum (FBS); the cells were detached for passages or experiments upon reaching 80% confluence. In most experiments, HeLa cells were harvested either in 12-well plates at 6×10^4 cell mL^{-1} in 1 mL DMEM (without FBS), or in our custom-made PDMS chips at 6×10^5 cell mL^{-1} in 0.1 mL. Of note, the cells were exposed to different stresses after 24 h of initial cell seeding, challenging HeLa cells as follows: Staurosporine and Rapamycin were administrated at a final concentration of 10 µM and 20 nM, respectively, whereas H_2O_2 was administrated to reach 50 µM in the cell solution. For the hypotonic stress, the cells were exposed to a 1:5 dilution of Hank's balanced Salt solution (HBSS) in milli-Q Water. Heating damage was provoked by placing the cell device on a heating plate at 55 °C for 5 min. Illumination with UV light (365 nm) for 15 min was utilized to induce UV-stress.

Flow Cytometry: For flow cytometry studies, HeLa cells were detached from the well plates upon treatment, using 1X Trypsin. Those cells were concentrated afterward by centrifugation (5 min, 1500 rpm), leaving a final volume of 100 ml that contains Annexin buffer (1:5, in PBS) + 2% FBS. Annexin V-FITC (Annexin V) and Propidium Iodide (PI) were added at the concentrations detailed in the commercial kit. After 15 min incubation, the cells were resuspended again in diluted Annexin Buffer (1:5, PBS), reaching a final volume of 400 µL. Finally, stained cells with PI and Annexin V were measured by flow cytometry. A minimum of 5000 cells were analyzed in every experiment collecting emission for Annexin V-FITC at 530 nm and for PI-PE at 585 nm using the blue laser (488 nm excitation) in a FACSCantoII Cytometer. Data were finally analyzed by the BD FACSDiva software.

Lactate Dehydrogenase (LDH) Cytotoxic Assay: LDH Cytotoxicity assay was selected to assess the cell viability. The enzyme LDH, present in all cell types, is typically added into the cell culture medium upon damage of the plasma membrane. These assays were performed following the commercial protocol, therefore 50 µL of cell media from evaluated conditions were plated in a 96-well plate. Next, the absorbance of such solutions was measured at 490 and 680 nm in a Thermo Scientific Varioskan Flash plate reader. The percentage of cell viability was calculated as follows:

$$\text{Cytotoxicity}(\%): \frac{A490[\text{Experiment}] - A490[\text{Low Control}]}{A490[\text{High Control}] - A490[\text{Low Control}]} \quad (1)$$

Screening of Extracellular ATP Release: To measure the ATP that has been released into the cultured cell environment as a result of cell death, the RealTime-Glo Extracellular ATP assay was employed. This assay generated a nearly simultaneous luminescence response proportional to current extracellular ATP levels in cell culture. Consequently, a 4X ATP assay reagent was mixed with the culture medium and dispensed on the growing cells. Then, upon stress, the resulting luminescence was recorded over time (15 min, 30 min, 1 h, 2 h, 4 h, and 24 h) by an Orion II Microplate Luminometer. Mean ± standard deviation (SD) values of luminescence were calculated from the three independent cell assays over indicated times.

SERS Measurements: Cell supernatant derived from different biological assays was sampled and 2 µL of the liquid was deposited on a plasmonic substrate, measuring the SERS signal after solvent evaporation (in substrates fabricated by drop casting or plasmonic superlattices). Importantly, all conditions were performed in at least three independent cell assays, and spiked on three different plasmonic substrates. Finally, SERS spectra were collected by an InVia Reflex Raman microscope (Renishaw plc). Renishaw equipment comprises an optical microscope (Leica) with a XYZ scanning stage coupled to a high-throughput Raman spectrometer equipped with a 1024×512 front-illuminated CCD detector and a grating of 1200 grooves mm^{-1} for 785 and 1800 grooves mm^{-1} for 633 nm. A line-shaped 785 nm laser excitation source (maximum 190 mW) was used, recording in static mode at standard confocality through a 50× objective in long distance (numerical aperture NA = 0.5) with an integration time of 1s at 15.15 mW (10% of the maximum power). In contrast, the signals from paper-based substrates were recorded under a 633 nm laser line (maximum 18 mW), again with the L50x objective and an integration time of 1s. In this case, to avoid burning of the nitrocellulose membrane, the power of the laser was 0.54 mW (5% of the maximum power). To calculate the averaged spectrum to be finally sent to the SVM network, 10 points from different substrate areas were measured in every experiment.

Microfluidic Chip Fabrication: 3D Stereolithography (SLA) printing was used to develop 3D patterns with the desired geometry on light-curable resin, presenting a µm-sized resolution. Here, we combined conventional PDMS soft lithography with SLA 3D printed molds, mimicking the procedure used for micropatterned silicon masters. Initially, 3D designs of the microfluidic chips were created using the Tinkercad online tool. Such 3D models were then printed by the Anycubic Photon Mono printer via a Standard Photopolymer resin. Subsequently, 3D printed resin-based masters were cleaned using 2-propanol and cured by a 365 nm UV light. After that, the resin was covered with a thin layer of enamel to improve PDMS curation process, as reported in Hagemann et al.[32] Polydimethylsiloxane was mixed at a 10:1 weight ratio of base to curing agent. The mixed solution was poured into the resin master and then degassed to remove air bubbles. Once the solution was cured at 70 °C, the replica-molded layer was trimmed, perforated, and cleaned with ethanol 70%. In the final step, the PDMS device were plasma-bonded on glass coverslips with varying dimensions.

SERS Data Analysis, Statistical and Machine Learning Methods: The obtained SERS spectra were initially processed to remove cosmic rays by using the Renishaw software. For the following steps, we employed different libraries for machine learning in Python Colab. First, a baseline correction data pre-processing step based on BaselineRemoval library was applied, in particular, Modpoly fitting (following a polynomial curve of 7 order), to remove interferences from the background arising from the substrate and biological fluorescence. The spectra were subsequently normalized to the most intense signal between 0 and 1.

The processed data sets (5500 spectra) were then analyzed with principal component analysis (PCA) and t-Distributed Stochastic Neighboring Entities (t-SNE, perplexity = 40, number of iterations = 3000) and supported vector machines (SVM, kernel = linear) using standard algorithms of Scikit Learn library. Both t-SNE and PCA, are unsupervised algorithms for exploring the data without previous training and require a preliminary step of data standardization (mean = 0, variance = 1). For data labeling in the supervised SVM classification, threshold estimations were made according to the results obtained in control conditions (for the LDH and flow cytometry). Upper value of the 99% confidential interval was calculated, considering normal distribution, by this formula:

$$\text{Confidence interval} = \overline{x} + z \times \frac{\sigma}{\sqrt{n}} \quad (2)$$

The general framework of the SVM presented an input layer –a SERS spectrum with 1×939 dimensions– and a hidden layer devoted to extracting its principal vibrational features. Once the model was trained and tested with the model inspection dataset (n = 5500, divided in 70/30 training-testing respectively), classification of different biological assays was done based on the averaged spectra of 10 points localized at random spots of the SERS substrate (N = 20 biological assays, see scheme in Figure S3, Supporting Information). In the experiments with the 633 nm, the retrieved spectra displayed lower numbers of pixels: 918 wavenumber points at 633 nm versus 939 wavenumber points in 785 nm. Therefore, we calculated the missing data points in the spectra by linear interpolation with the neighboring wavenumber before passing the spectra through the SVM model. Both the code and the data are available in GitHub. https://github.com/JaviPlou/Machine-learning-high-throughput-SERS-classification-of-cell-secretomes

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

J.P. and P.S.V. contributed equally to this work. L.M.L.-M. acknowledges funding from the European Research Council (ERC AdG 787510, 4DbioSERS). A.C. was funded by MICINN (PID2019-108787RB-I00 (FEDER/EU)) and the European Research Council (ERC Consolidator Grant 819242). Funding for open access charge: Universidade de Vigo/CISUG.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available in the supplementary material of this article.

Keywords

artificial Intelligence, biosensors, cell devices, drug screening, metabolic profiles

Received: December 7, 2022
Revised: January 25, 2023
Published online:

[1] J. C. Ameisen, *Cell Death Differ.* **2002**, *9*, 367.
[2] L. Galluzzi, I. Vitale, S. A. Aaronson, J. M. Abrams, D. Adam, P. Agostinis, E. S. Alnemri, L. Altucci, I. Amelio, D. W. Andrews, M. Annicchiarico-Petruzzelli, A. V Antonov, E. Arama, E. H. Baehrecke, N. A. Barlev, N. G. Bazan, F. Bernassola, M. J. M. Bertrand, K. Bianchi, M. V Blagosklonny, K. Blomgren, C. Borner, P. Boya, C. Brenner, M. Campanella, E. Candi, D. Carmona-Gutierrez, F. Cecconi, F. K.-M. Chan, N. S. Chandel, *Cell Death Differ.* **2018**, *25*, 486.
[3] M. D. Jacobson, M. Weil, M. C. Raff, *Cell* **1997**, *88*, 347.
[4] G. Yan, M. Elbadawi, T. Efferth, *World Acad. Sci. J.* **2020**, *2*, 39.
[5] L. Ouyang, Z. Shi, S. Zhao, F.-T. Wang, T.-T. Zhou, B. Liu, J.-K. Bao, *Cell Prolif.* **2012**, *45*, 487.
[6] C. B. Medina, P. Mehrotra, S. Arandjelovic, J. S. A. Perry, Y. Guo, S. Morioka, B. Barron, S. F. Walk, B. Ghesquière, A. S. Krupnick, U. Lorenz, K. S. Ravichandran, *Nature* **2020**, *580*, 130.
[7] V. Labi, M. Erlacher, *Cell Death Dis.* **2015**, *6*, e1675.
[8] C. D. Gregory, J. D. Pound, *J. Pathol.* **2011**, *223*, 177.
[9] T. A. van Schaik, K.-S. Chen, K. Shah, *Front Oncol.* **2021**, *11*, 678562.
[10] E. C. Madden, A. M. Gorman, S. E. Logue, A. Samali, *Trends Cancer* **2020**, *6*, 489.
[11] J. Tüshaus, S. A. Müller, E. S. Kataka, J. Zaucha, L. Sebastian Monasor, M. Su, G. Güner, G. Jocher, S. Tahirovic, D. Frishman, M. Simons, S. F. Lichtenthaler, *EMBO J.* **2020**, *39*, 105693.
[12] P. Mukherjee, S. Mani, *Biochim. Biophys. Acta* **2013**, *1834*, 2226.
[13] S. Han, H. Shin, J.-K. Lee, Z. Liu, R. Rabadan, J. Lee, J. Shin, C. Lee, H. Yang, D. Kim, S. H. Kim, J. Kim, J.-W. Oh, D.-S. Kong, J.-I. Lee, H. J. Seol, J. W. Choi, H. J. Kang, D.-H. Nam, *Exp. Mol. Med.* **2019**, *51*, 147.
[14] N. Kim, M. R. Thomas, M. S. Bergholt, I. J. Pence, H. Seong, P. Charchar, N. Todorova, A. Nagelkerke, A. Belessiotis-Richards, D. J. Payne, A. Gelmi, I. Yarovsky, M. M. Stevens, *Nat. Commun.* **2020**, *11*, 207.
[15] J. Plou, B. Molina-Martínez, C. García-Astrain, J. Langer, I. García, A. Ercilla, G. Perumal, A. Carracedo, L. M. Liz-Marzán, *Nano Lett.* **2021**, *21*, 8785.
[16] J. Plou, I. García, M. Charconnet, I. Astobiza, C. García-Astrain, C. Matricardi, A. Mihi, A. Carracedo, L. M. Liz-Marzán, *Adv. Funct. Mater.* **2020**, *30*, 1910335.
[17] K. Koike, K. Bando, J. Ando, H. Yamakoshi, N. Terayama, K. Dodo, N. I. Smith, M. Sodeoka, K. Fujita, *ACS Nano* **2020**, *14*, 15032.
[18] E. Genova, M. Pelin, G. Decorti, G. Stocco, V. Sergo, A. Ventura, A. Bonifacio, *Anal. Chim. Acta* **2018**, *1005*, 93.
[19] B. Kang, L. A. Austin, M. A. El-sayed, *ACS Nano* **2014**, *8*, 4883.
[20] M. Aioub, M. A. El-Sayed, *J. Am. Chem. Soc.* **2016**, *138*, 1258.
[21] F. Lussier, D. Missirlis, J. P. Spatz, J.-F. Masson, *ACS Nano* **2019**, *13*, 1403.
[22] W. J. Thrift, S. Ronaghi, M. Samad, H. Wei, D. G. Nguyen, A. S. Cabuslay, C. E. Groome, P. J. Santiago, P. Baldi, A. I. Hochbaum, R. Ragan, *ACS Nano* **2020**, *14*, 15336.
[23] B. Leber, U. Mayrhauser, B. Leopold, S. Koestenbauer, K. Tscheliessnigg, V. Stadlbauer, P. Stiegler, *Anticancer Res.* **2012**, *32*, 915.
[24] A.-S. Gary, P. J. Rochette, *Sci. Rep.* **2020**, *10*, 18946.
[25] A. R. L. Marshall, J. Stokes, F. N. Viscomi, J. E. Proctor, J. Gierschner, J.-S. G. Bouillard, A. M. Adawi, *Nanoscale* **2017**, *9*, 17415.
[26] S.-X. Li, Q.-Y. Zeng, L.-F. Li, Y.-J. Zhang, M.-M. Wan, Z.-M. Liu, H.-L. Xiong, Z.-Y. Guo, S.-H. Liu, *J. Biomed. Opt.* **2013**, *18*, 027008.
[27] M. van Engeland, L. J. W. Nieland, F. C. S. Ramaekers, B. Schutte, C. P. M. Reutelingsperger, *Cytometry* **1998**, *31*, 1.
[28] M. Michaud, I. Martins, A. Q. Sukkurwala, S. Adjemian, Y. Ma, P. Pellegatti, S. Shen, O. Kepp, M. Scoazec, G. Mignot, S. Rello-Varona, M. Tailler, L. Menger, E. Vacchelli, L. Galluzzi, F. Ghiringhelli, F. Di Virgilio, L. Zitvogel, G. Kroemer, *Science* **2011**, *334*, 1573.
[29] M. Yafia, O. Ymbern, A. O. Olanrewaju, A. Parandakh, A. Sohrabi Kashani, J. Renault, Z. Jin, G. Kim, A. Ng, D. Juncker, *Nature* **2022**, *605*, 464.
[30] J. E. L. Villa, N. R. Quiñones, F. Fantinatti-Garboggini, R. J. Poppi, *Anal. Bioanal. Chem.* **2019**, *411*, 705.
[31] L. Polavarapu, A. La Porta, S. M. Novikov, M. Coronado-Puchau, L. M. Liz-Marzán, *Small* **2014**, *10*, 3065.
[32] C. Hagemann, M. C. D. Bailey, V. M. Lionello, N. Khokhar, P. Suklai, C. Moreno-Gonzalez, K. O'Toole, G. Konstantinou, S. Joshi, E. Giagnorio, M. Bergholt, A. Imbert, F. S. Tedesco, A. Serio, https://doi.org/10.1101/2022.02.22.481424.
[33] N. G. Bastús, J. Comenge, V. Puntes, *Langmuir* **2011**, *27*, 11098.
[34] P. C. Lee, D. Meisel, *J. Phys. Chem.* **1982**, *86*, 3391.
[35] C. Matricardi, C. Hanske, J. L. Garcia-Pomar, J. Langer, A. Mihi, L. M. Liz-Marzán, *ACS Nano* **2018**, *12*, 8531.
[36] C. Hanske, G. González-Rubio, C. Hamon, P. Formentín, E. Modin, A. Chuvilin, A. Guerrero-Martínez, L. F. Marsal, L. M. Liz-Marzán, *J. Phys. Chem. C* **2017**, *121*, 10899.