Supporting Information

Label-Free Spectroscopic SARS-CoV-2 Detection on Versatile Nanoimprinted Substrates

Debadrita Paria[1,#], Kam Sang Kwok[2,#], Piyush Raj[1], Peng Zheng[1], David H. Gracias [2-7*], Ishan Barman[1,7,8*]

[1]Department of Mechanical Engineering, Johns Hopkins University, Baltimore, MD, USA.

[2]Department of Chemical & Biomolecular Engineering, Johns Hopkins University, Baltimore, MD, USA.

[3]Department of Materials Science and Engineering, Johns Hopkins University, Baltimore, MD, USA.

[4]Department of Chemistry, Johns Hopkins University, Baltimore, MD, USA.

[5] Laboratory for Computational Sensing and Robotics (LCSR). Johns Hopkins University, Baltimore, MD, USA.

[6] Sidney Kimmel Comprehensive Cancer Center (SKCCC), Johns Hopkins University School of Medicine, Baltimore, MD, USA

[7]Department of Oncology, Johns Hopkins University School of Medicine, Baltimore, MD, USA.

[8]Department of Radiology & Radiological Science, Johns Hopkins University School of Medicine, Baltimore, MD, USA

[#]These authors contributed equally to this work.

Keywords: COVID sensors, Surface Enhanced Raman Spectroscopy, Machine Learning, Nano-imprint Lithography, Flexible sensors

***Correspondence:**

David H. Gracias

Johns Hopkins University

Whiting School of Engineering

Department of Chemical & Biomolecular Engineering

Maryland Hall 125

Baltimore, MD 21218, USA

E-mail: dgracias@jhu.edu

Ishan Barman

Johns Hopkins University

Whiting School of Engineering

Department of Mechanical Engineering

Latrobe Hall 103

Baltimore, MD 21218, USA

E-mail: ibarman@jhu.edu

Details of the Finite Element Electromagnetic Simulation Model
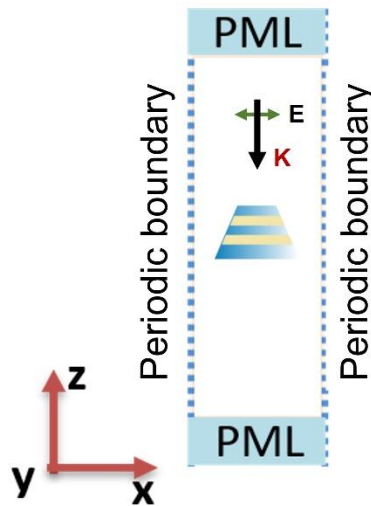
Figure S1: Conceptual scheme of simulation. The direction of polarization and the propagation of the incident field is shown by the green and black arrow, respectively.

## Absorption characteristics of the FEMIA

An approximate location of the plasmon resonance for the FEMIA with different number of layers can be identified from the absorption coefficient.
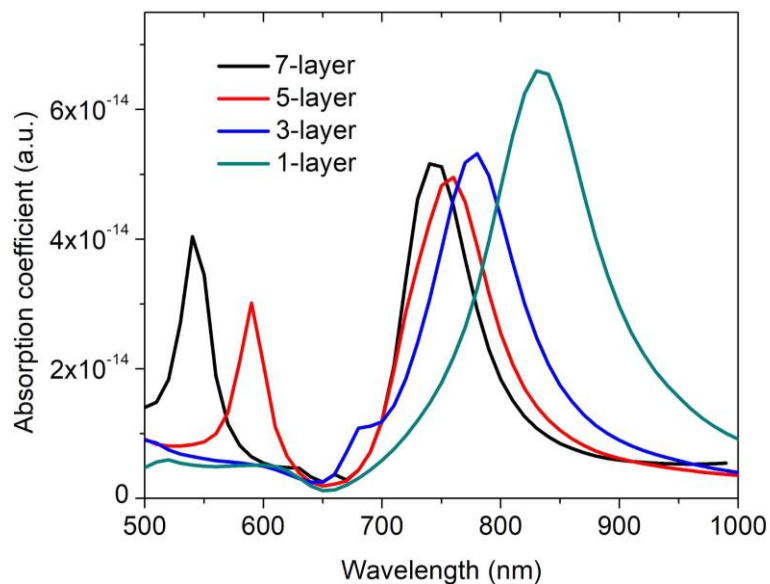
Figure S2: Absorption of a 2-dimensional array of FEMIA nanostructures with a single layer of silver (1-layer). 1 layer of silica sandwiched between 2 layers of silver (3-layer); 2 layers of silica sandwiched
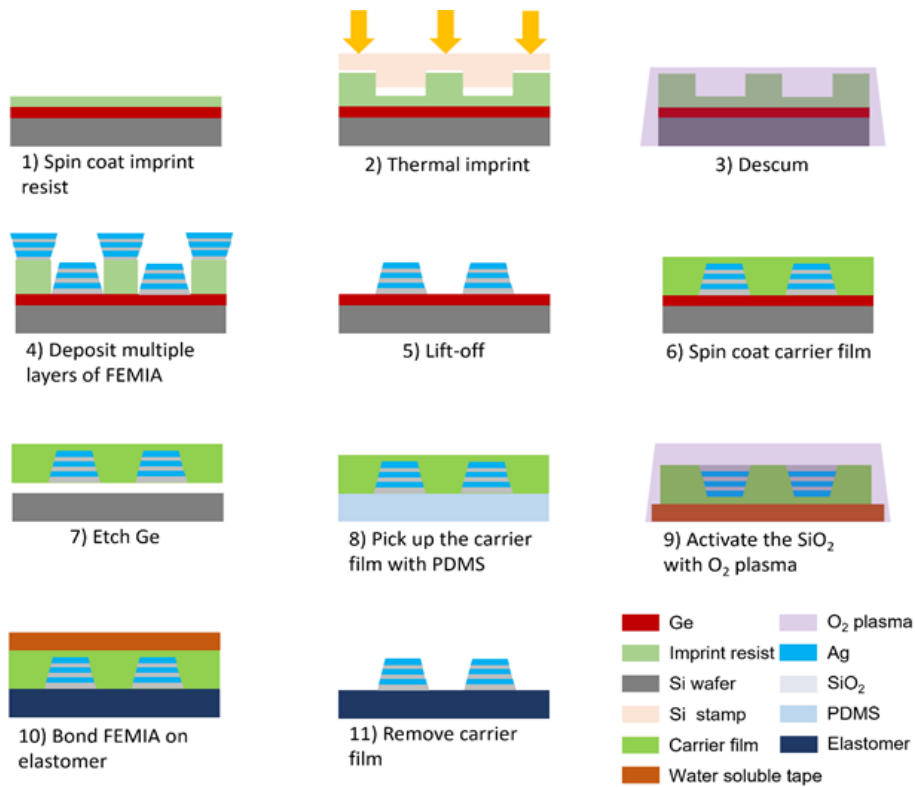
between 3 layers of silver with layers alternating (5-layer); 3 layers of silica sandwiched between 4 layers of silver with layers alternating (7-layer).

FEMIA fabrication on rigid (Si) substrate
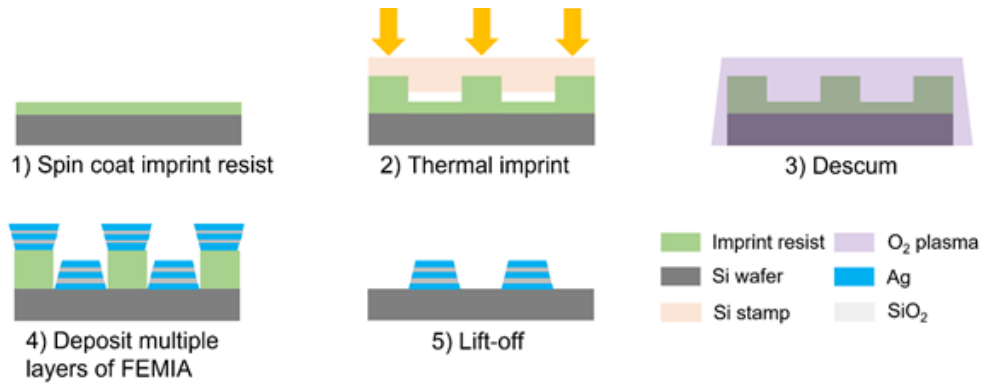
**Figure S3:** Schematic of the detailed steps for fabrication of FEMIA nanostructures on Si wafer.

FEMIA transfer on flexible substrate

Figure S4: Schematic of the detailed steps for transfer of FEMIA nanostructures on flexible substrates

Optical Performance of FEMIA

Our normalized experimental reflectance measurement (details in Materials and Methods) closely agrees with the simulation and demonstrates two localized surface plasmon resonance (LSPR) peaks for the 5-layer and 7-layer samples (Figure **S5**). From the resonance positions in the simulation, we can conclude that the peaks in the reflectance spectra represent the localized surface plasmon resonance (*1*, *2*).
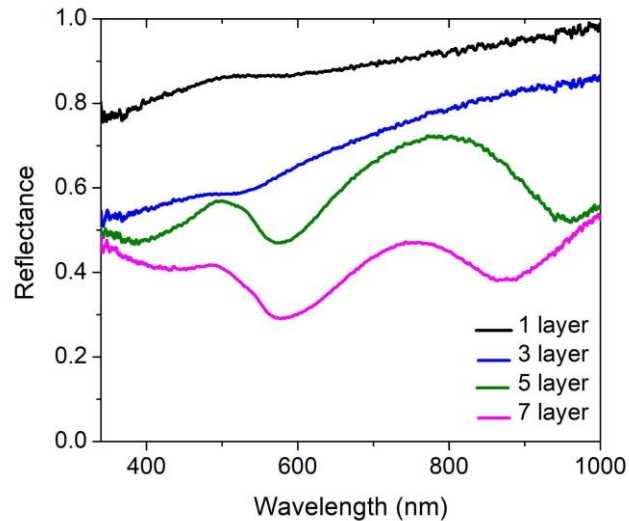
Figure S5: Experimental reflectance spectra of the FEMIA with different layers on silicon.

**Dependence of plasmon resonance on array periodicity**

An important point to note in the optical characterization is the array periodicity. When the nanostructures are placed further apart (numerically), the dipolar resonance becomes weaker. Thus, considering a realistic dimension from the fabrication point of view we chose a periodicity of 400 nm in our experiments.
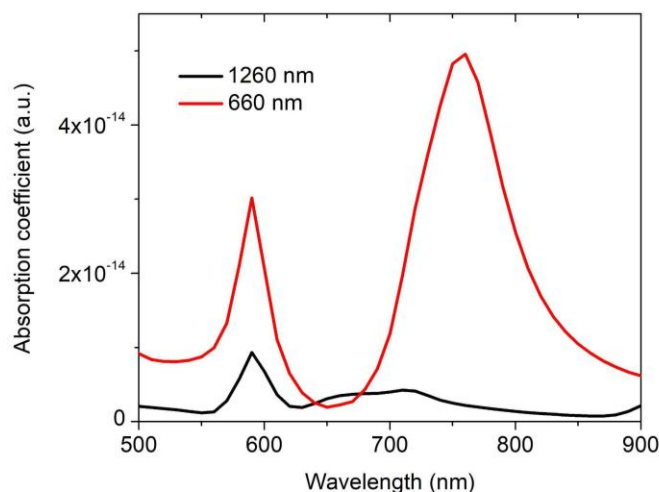
Figure S6: Absorption of an infinite 2-dimensional array of 5-layer FEMIA for different array periodicities, i.e., center to center distance of 1260 nm and 660 nm.

Enhancement Test with 4-NTP

The SERS signal enhancing capabilities of the different samples are assessed using 4-nitrothiophenol (4-NTP), a standard Raman reporter. An aqueous solution of 500 μM 4-NTP is dropped on FEMIA substrates with different layers. Raman spectra acquired from the dried drop using a 785 nm laser demonstrate strong signatures of the 4-NTP molecule (*3*). The 5-layer FEMIA, which provides the maximum enhancement (Figure **3B)** at 785 nm laser excitation as compared to the other samples, is considered for our study. We chose a 785 nm laser for the rest of the study, considering it lies in the tissue transparent region, causing minimal damage to biological samples, along with offering a low fluorescence background while maintaining a sufficient signal-to-noise ratio (*4*).
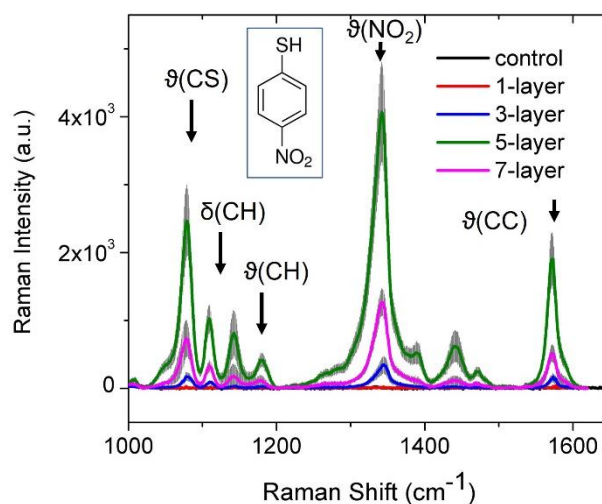

Figure S7: Raman spectra of 500 μM 4-NTP acquired on different FEMIA substrates and a plain silicon wafer (control). Some of the prominent Raman peaks corresponding to 4-NTP are marked with an arrow. ϑ indicates stretching, and δ indicates bending. The inset shows the chemical structure of the molecule.

Table S1: Peak assignment of the prominent peaks in the Raman spectra of HA and S protein

Raman shift (cm⁻¹)	Assignment
1048, 1138, 1328, 1586, 1621	Tryptophan
1104, 1615	Phenylalanine
1155	C-N stretching
1189, 1258	Amide III
1441, 1589	Glycine
1453	C-H stretching of glycoprotein
1567	Valine
1525	Amide II
1655, 1667	Amide I

PCA on SERS spectra of HA and S protein
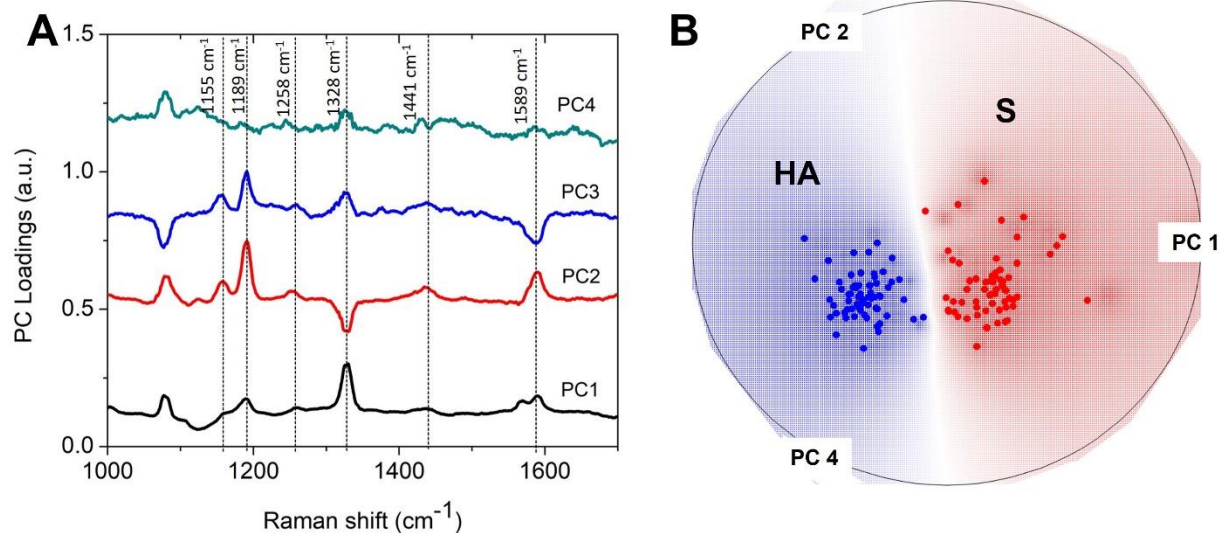
Figure S8: (A) The first four principal components obtained from the entire date set of HA and S protein. The prominent features are represented by the dotted lines. (B) Radial visualization plot of the most relevant PC scores demonstrating the clustering of the two types of proteins.

A dried drop of viral lysate on FEMIA
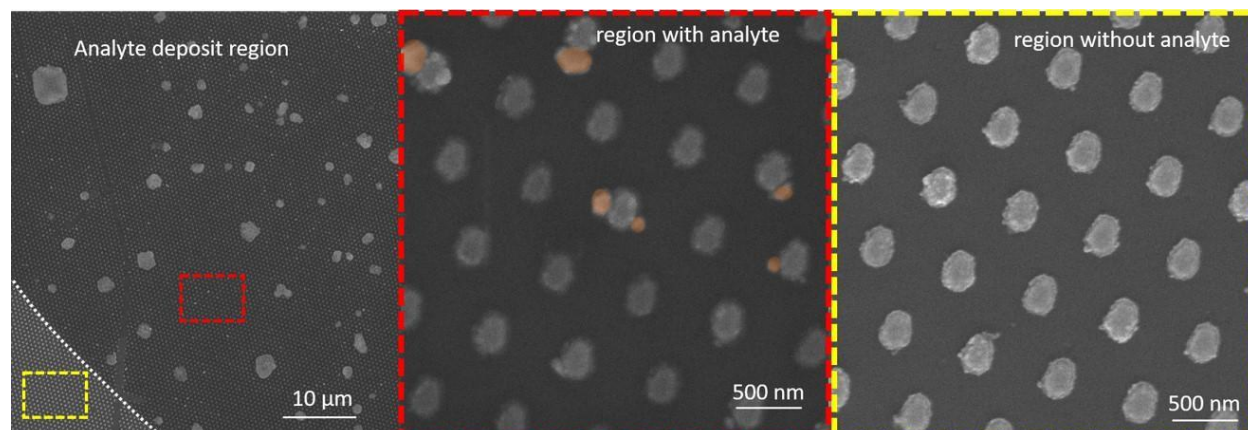
Figure S9: SEM image showing a dried drop of cell lysate infected with SARS-CoV-2 (left panel) on the FEMIA. Zoomed-in image of the region with the analyte (middle panel) and without the analyte (right panel). The virus particles are false-colored orange for better visualization.

Table S2: Peak assignment of the prominent peaks in the Raman spectra of the cell lysates containing the virus

Raman shift (cm^{-1})	Assignment
1004	Phenylalanine
1076	Lipids
1123	C-N (protein)
1138, 1622	Tryptophan
1156	C-N (protein)
1246, 1280	Amide III
1347	Guanine
1367	Lipid
1465	Adenine
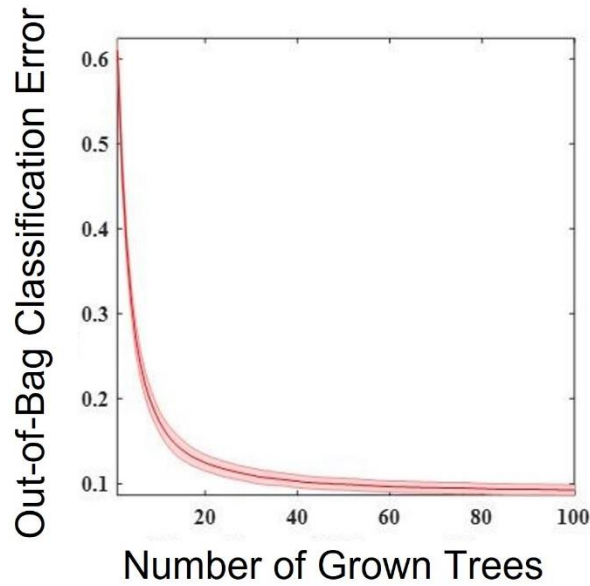
Figure S10: Evaluation of classification error vs. the number of decision trees shows that the error falls below 10% with the inclusion of 100 trees.

Figure S11: Classification accuracy for binary random forest classification with SARS-CoV-2 sample in one class and the other viruses (H1N1 A, Zika, and Marburg) in the other class.
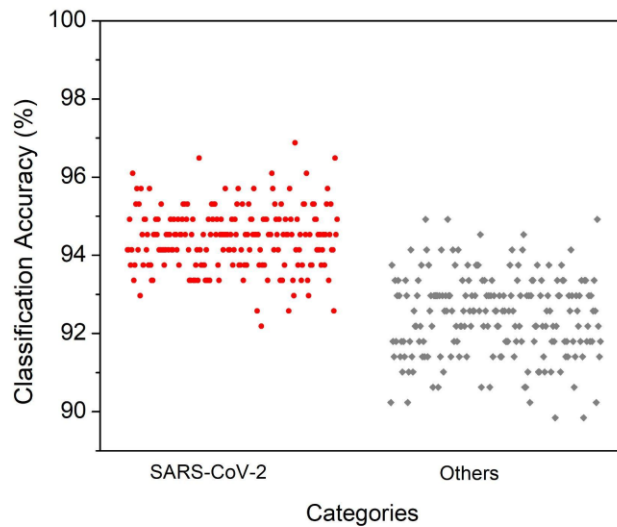
LOD measurement for SARS-CoV-2

For the LOD measurement, the cell lysate containing the virus is diluted with pristine cell lysates to make solutions of virus concentration ranging from 10^4 copies/ml to 10^9 copies/ml. The same technique as before is followed for measurement and data acquisition. Partial Least Square regression with leave-one-spectra-out cross-validation is applied on the acquired data set. Eight spectra are averaged into one in the data set

along with the application of outlier rejection (*5*). From the figure below saturation is observed below 10^5 copies/ml, while the model performs better for higher concentrations.
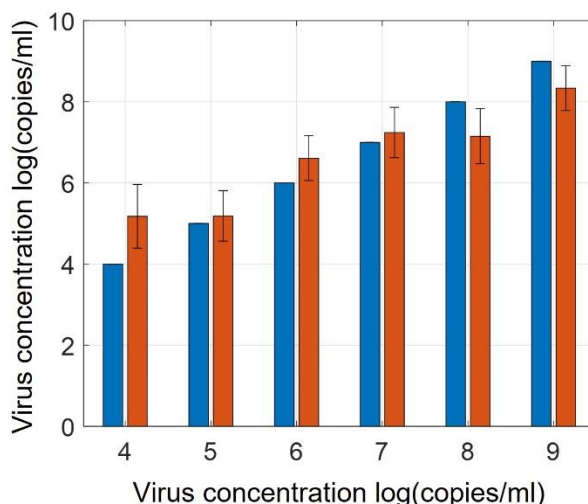
Figure S12: PLS prediction of SARS-CoV-2 concentrations. Blue represents actual concentration, and orange represents predicted concentration

Materials and Methods

Fabrication of FEMIA arrays on a rigid (silicon wafer) substrate

We fabricated the array of metal insulator metal (MIM) nanostructures by nanoimprint lithography (NIL). Briefly, we spin-coated a layer of the NIL resist (mr-I 7030, micro resist technology) with a thickness of 350 nm on top of a silicon wafer. We imprinted the nanopatterns using a commercial low-cost Si master stamp (LightSmyth grating) and a Nanonex Advanced Nanoimprint Tool NX-B200 with a pressure of 350 psi at 130 °C. After imprinting, we stripped away the residue of the resist using an oxygen (O_2) plasma at 60 W for 2 minutes. Our FEMIA nanostructures have multiple, i.e., 1, 3, 5, or 7 MIM layers. Each layer has silver (Ag, 20 nm), silicon dioxide (SiO_2, 10 nm), and silver (Ag, 20nm) which were deposited by e-beam evaporation. We also deposited 2 nm of chromium (Cr) between the silicon wafer and Ag and in between each layer of Ag and SiO_2 and SiO_2 and Ag as well as to improve adhesion. After deposition, we sonicated the sample in acetone to dissolve the mr-I 7030 and generate a large area array (8 mm by 8 mm) of FEMIA nanostructures on top of the Si wafer.

Fabrication of FEMIA arrays on a flexible (elastomer) substrate

We thermally evaporated a thin film of germanium (Ge) with a thickness of 100 nm on a Si wafer as the sacrificial layer. We created arrays of multiple FEMIA nanostructures on top of the Ge coated wafer using the same approach described in the previous section.

We spin coated a layer of polymethylglutarimide (PMGI SF6, Kayaku Advanced Materials) as a carrying film to transfer the arrays of FEMIA nanostructures. We dissolved the Ge sacrificial layer in water and picked up the film using a mm-scale thick piece of PDMS. After the film was completely dried, we peeled off the film from the PDMS using water-soluble tape (3M 5414). We activated the surfaces of the bottom of the FEMIA nanostructures (SiO_2) and elastomer (Dragon Skin™) by O_2 plasma. We bonded the FEMIA and elastomer by pressing both surfaces together and removing the tape and PMGI in water and alkaline developer (MF 26A, Kayaku Advanced Materials).

Previous studies of transfer printing with irreversible bonding require exposing the nanostructures to oxygen (O_2) plasma or molecular crosslinking (*6*, *7*), hence not suitable for SERS since they tend to oxidize surfaces of Ag, thereby reducing the electrical field enhancement (*8*). Here, for the first time, we transfer-printed a large area of Ag-based FEMIA nanostructures on an elastomer without compromising e-field enhancement or Raman activity. A key innovation lies in the use of a 100 nm thick germanium (Ge) layer deposited on the silicon wafer, which acts as a sacrificial layer and creates a large area of FEMIA nanostructures on the Ge coated wafer by NIL and multi-layered e-beam evaporation. After lift-off, we utilized PMGI as a carrying film to release the FEMIA nanostructures from the wafer and protect the nanostructures from oxidation during the plasma process. The advantage of PMGI over other polymer films (e.g., polymethyl methacrylate) is that we can remove PMGI in an alkaline developer, whereas organic solvents easily swell any elastomer (*9*).

Simulation details

For the numerical simulation, we used finite element analysis (COMSOL Inc., Burlington, MA, USA). We used a periodic boundary condition to simulate infinitely repeating FEMIA nanostructures in 2 dimensions. We considered a constant dielectric constant value of 2.09 for the silica layer and a frequency-dependent dielectric constant for the silver layer (*10*). We considered dynamic meshing where the maximum and minimum size for the FEMIA nanostructure is 50 nm and 0.5 nm, respectively and $\lambda/6$ (λ - wavelength of incident light) for the rest. The boundary conditions at the input and output ports were assigned as a perfectly matched layer. The dimensions for the FEMIA were considered as measured by SEM, i.e, the thickness of silver and silica were 20 nm and 10 nm, respectively. We considered the shape of the nanostructure as conical (with rounded edges), having a bottom radius of 130 nm, and the center-to-center distance between adjacent nanostructures was 660 nm. In an actual fabrication, it is difficult to achieve sharp edges of the FEIMA. Thus, the simulation model is considered to have rounded edges to prevent the inclusion of additional enhancement in our numerical calculation due to the "lightning-rod effect" at sharp corners (*11*).

Reflectance measurement

Reflection spectra were acquired by a USB 4000 spectrometer (Ocean Optics) and a tungsten halogen lamp (LS-1) (Ocean Optics). The spectra were averaged over 3 measurements and normalized.

Raman measurement

We used an XploRA PLUS Raman microscope (HORIBA Instruments Inc., Edison, NJ, USA) for the Raman measurements. For the 4-NTP study, the Raman scattered signal was acquired through a 100x Olympus objective for 1 s and averaged 3 times, 10% power. For the purified protein study, a 100x Olympus objective was used for 0.5 s and averaged 20 times, 25% power. The laser power was about 1 μW. For the 4-NTP study, signals were collected from 64 spatially distinct spots from a single substrate. For the other analytes (protein, cell lysates and saliva samples), signals were collected from 64 spatially distinct spots on two different substrates.

For the viral samples, the acquisition parameters were: 100x Olympus objective, 0.5 s acquisition time, 30 accumulations, 25% power, 128 spectra for each sample.

For the saliva samples: 50x Olympus objective, 0.5 s acquisition time, 30 accumulations, 25% power, 128 spectra for each sample.

For all the samples 2.5 μl of the solution was used. A thermoelectrically cooled charge-coupled device camera (1024 × 256OE Syncerity, HORIBA Instruments Inc., Edison, NJ, USA) collects the scattered spectra. We have calibrated the Raman shift using a silicon substrate.

Samples details

4-NTP was purchased from Sigma-Aldrich, St. Louis, MO.

HA protein was purchased from HA protein from Sino Biological A/WSN/1933 (Cat# 11692-V08H), S protein BEI Resources

Following virus samples and control cell lysates were purchased from BEI Resources

Influenza A Virus, A/New York/18/2009 (H1N1) (10^7 copies/ml)

SARS-Related Coronavirus 2 (*47*), Isolate USA-WA1/2020 (10^9 copies/ml)

Zika virus, PRVABC59 (10^{10} copies/ml)

Marburg Marburgvirus, H. sapiens-tc/AGO/2005/Angola-200501379 (10^{11} copies/ml)

All samples have cell lysate from *Cercopithecus aethiops* kidney epithelial cells (Vero E6) as the base media.

Control - Vero 6 cell lysate from BEI Resources

Human pooled saliva - Innovative Research

All measurements were performed after drying the analyte drop on the substrate in the ambient, which permits physical adsorption of the analytes to the nanostructured substrate. No chemical functionalization was performed for specific binding of the analyte.

The virus samples were handled in BSL-2 laboratory conditions. After the measurement, the samples were discarded in a 10% bleach solution.

Random Forest Analysis

For random forest analysis, the number of spectra was made equal for all classes. This equalization was done by first getting a minimum number of spectra from each class and then by randomly picking a minimum number of spectra from other classes. This method ensured that the random forest results were not biased to any class. Thereafter, the TreeBagger built-in MATLAB® function was used to perform classification-based random forest with sample replacement and curvature-based predictor selection. After that, the out-of-bag prediction was used to predict classes. Since the initial spectra selection from different classes were at random, we repeated the analysis 200 times to ensure robustness in our prediction, and the final classification results included all the iterations.

Data Analysis

The Raman spectra collected were subjected to a background subtraction via a polynomial fitting algorithm (*12*). We have applied a 5th Savitzky Golay filter with frame length 21 on the resultant spectra. This was followed by peak normalization with the maximum intensity value. We used MATLAB 2019b for the spectral pre-processing.

TP - true positive, TN - true negative, FP - false positive, FN - false negative.

Precision: TP/(TP+FP); Recall/Sensitivity: TP/(TP+FN); Specificity: TN/(TN+FP).

For our multiclass classification, accuracy is defined as the percentage of correctly classified spectra in each class for each iteration.

PCA

The pre-processed data were subjected to PCA using Orange data mining software (*13*). Operating without any *a priori* knowledge of the sample composition, PCA projects the

spectral data onto a set of linearly uncorrelated (orthogonal) directions (that are called principal components), such that the variance in the original data is captured using only a few PCs. The radial visualization plot maps the PC scores onto a two-dimensional space to permit analysis of inter- vs. intra-class variance of the different analytes. We employed the VizRank algorithm in the Orange data mining software to select the PCs based on their ability to obtain informative projections of the class-labeled data.

1. G. Si, Y. Zhao, J. Lv, M. Lu, F. Wang, H. Liu, N. Xiang, T. J. Huang, A. J. Danner, J. Teng, Y. J. Liu, Reflective plasmonic color filters based on lithographically patterned silver nanorod arrays. *Nanoscale.* **5**, 6243–6248 (2013).

2. Y. Shen, J. Zhou, T. Liu, Y. Tao, R. Jiang, M. Liu, G. Xiao, J. Zhu, Z.-K. Zhou, X. Wang, C. Jin, J. Wang, Plasmonic gold mushroom arrays with refractive index sensing figures of merit approaching the theoretical limit. *Nature Communications.* **4**, 1-9 (2013).

3. M. Abdelsalam, Surface enhanced raman scattering of aromatic thiols adsorbed on nanostructured gold surfaces. *Open Chemistry.* **7**, 446-453 (2009).

4. L. T. Kerr, H. J. Byrne, B. M. Hennelly, Optimal choice of sample substrate and laser wavelength for Raman spectroscopic analysis of biological specimen. *Analytical Methods.* **7**, 5041–5052 (2015).

5. M. Hubert, P. J. Rousseeuw, K. Vanden Branden, ROBPCA: A New Approach to Robust Principal Component Analysis. *Technometrics.* **47,** 64-79 (2005).

6. L. Gao, Y. Zhang, H. Zhang, S. Doshay, X. Xie, H. Luo, D. Shah, Y. Shi, S. Xu, H. Fang, J. A. Fan, P. Nordlander, Y. Huang, J. A. Rogers, Optics and Nonlinear Buckling Mechanics in Large-Area, Highly Stretchable Arrays of Plasmonic Nanostructures. *ACS Nano.* **9**, 5968–5975 (2015).

7. S. H. Hwang, S. Jeon, M. J. Kim, D.-G. Choi, J.-H. Choi, J.-Y. Jung, K.-S. Kim, J. Lee, J. H. Jeong, J. R. Youn, Covalent bonding-assisted nanotransfer lithography for the fabrication of plasmonic nano-optical elements. *Nanoscale.* **9**, 14335–14346 (2017).

8. C. Ma, M. J. Trujillo, J. P. Camden, Nanoporous Silver Film Fabricated by Oxygen Plasma: A Facile Approach for SERS Substrates. *ACS Applied Materials & Interfaces.* **8**, 23978–23984 (2016).

9. K. S. Kwok, P. K. Luthra, D. H. Gracias, in *2020 IEEE 20th International Conference on Nanotechnology (IEEE-NANO).* 207-208 (2020).

10. E. D. Palik, *Handbook of Optical Constants of Solids* (Academic Press, 2012).

11. Y. Tian, Q. Cui, H. Ma, A. Jiao, C. Wang, M. Zhang, L. Zheng, X. Wang, S. Li, G. Li, M. Chen, Extra electric field-enhanced lightning rod effect in pine needle-like Au microarrays for boosting direct plasmon-driven photoelectrochemical hydrogenation reactions via in-situ SERS monitoring. *Applied Surface Science.* **578**, 152100 (2022).

12. C. A. Lieber, A. Mahadevan-Jansen, Automated method for subtraction of fluorescence

from biological Raman spectra. *Applied spectroscopy.* **57**, 1363–1367 (2003).

13. J. Demšar, B. Zupan, G. Leban, T. Curk, Orange: From Experimental Machine Learning to Interactive Data Mining. *Lecture Notes in Computer Science.* 537–539 (2004).