

# Surface-Enhanced Raman Scattering-Based Odor Compass: Locating Multiple Chemical Sources and Pathogens

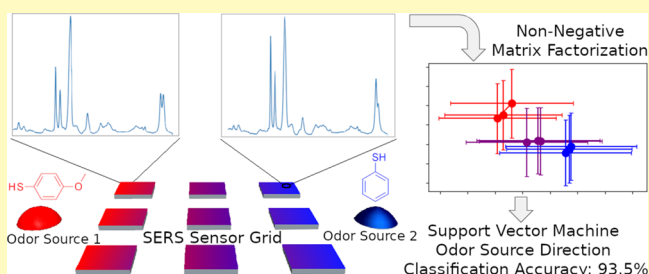
William John Thrift,<sup>†</sup> Antony Cabuslay,<sup>‡</sup> Andrew Benjamin Laird,<sup>§</sup> Saba Ranjbar,<sup>||</sup>  
Allon I. Hochbaum,<sup>†,‡,||,⊥</sup> and Regina Ragan<sup>\*,†,||</sup>

<sup>†</sup>Department of Materials Science and Engineering, <sup>‡</sup>Department of Chemistry, <sup>§</sup>Department of Information and Computer Science, <sup>||</sup>Department of Chemical and Biomolecular Engineering, and <sup>⊥</sup>Department of Molecular Biology and Biochemistry, University of California, Irvine, Irvine, California 92697, United States

## Supporting Information

**ABSTRACT:** Olfaction is important for identifying and avoiding toxic substances in living systems. Many efforts have been made to realize artificial olfaction systems that reflect the capacity of biological systems. A sophisticated example of an artificial olfaction device is the odor compass which uses chemical sensor data to identify odor source direction. Successful odor compass designs often rely on plume-based detection and mobile robots, where active, mechanical motion of the sensor platform is employed. Passive, diffusion-based odor compasses remain elusive as detection of low analyte concentrations and quantification of small concentration gradients from within the sensor platform are necessary. Further, simultaneously identifying multiple odor sources using an odor compass remains an ongoing challenge, especially for similar analytes. Here, we show that surface-enhanced Raman scattering (SERS) sensors overcome these challenges, and we present the first SERS odor compass. Using a grid array of SERS sensors, machine learning analysis enables reliable identification of multiple odor sources arising from diffusion of analytes from one or two localized sources. Specifically, convolutional neural network and support vector machine classifier models achieve over 90% accuracy for a multiple odor source problem. This system is then used to identify the location of an *Escherichia coli* biofilm via its complex signature of volatile organic compounds. Thus, the fabricated SERS chemical sensors have the needed limit of detection and quantification for diffusion-based odor compasses. Solving the multiple odor source problem with a passive platform opens a path toward an Internet of things approach to monitor toxic gases and indoor pathogens.

**KEYWORDS:** surface-enhanced Raman scattering spectroscopy, odor compass, chemical sensing, machine learning, convolutional neural networks, statistical spectral analysis, self-assembly



From October 2015 through February 2016, 100 000 tons of methane and 2.5 tons of benzene were leaked from the Aliso Canyon natural gas storage facility, making it the second worst natural gas leak in US history.<sup>1</sup> Rapid and early detection and identification of hazardous gas leaks are essential to reduce the damage of these disasters, dangerous industrial gas leaks, and contamination of virulent bacteria. Yet, most gas sensors merely alert of the presence of gases and give no indication of the gas source direction. Often, leaks are identified using a mobile odor detector and gas distribution mapping, an incredibly time-consuming process.

To improve upon existing methods, the odor compass was developed in 1991.<sup>2</sup> The canonical odor compass is composed of two semiconductor gas sensors mounted to a rotating stage with a fan mounted to draw analyte containing air toward the gas sensors.<sup>3,4</sup> Heuristics are used to identify odor source direction by comparing the relative signal from the gas sensors at various rotation angles.<sup>5,6</sup> These odor compasses are mounted onto mobile robots so that they can travel through the gas concentration gradient and eventually find the odor

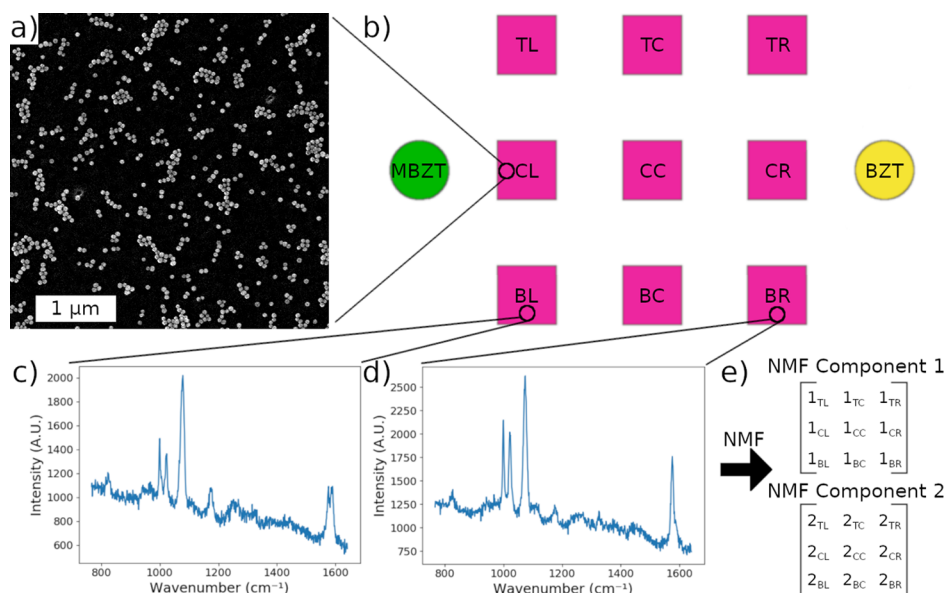
source, a process called chemotaxis in the biological context. Often, simple algorithms inspired by Braitenberg vehicles<sup>7</sup> are used, but sophisticated methods for odor source localization have been developed using genetic algorithms,<sup>8,9</sup> fuzzy logic,<sup>10,11</sup> and infotaxis.<sup>12</sup> Mobile odor compasses have now been used on land, air,<sup>13,14</sup> and sea.<sup>15,16</sup>

While most odor compasses rely on semiconductor gas sensors,<sup>17,18</sup> other gas sensors have been used, such as conductive polymers,<sup>19,20</sup> quartz crystal microbalance gas sensors,<sup>21,22</sup> and even silkworm antennae.<sup>23</sup> Yet (aside from silkworm antennae), these detectors cannot effectively differentiate different types of gases, limiting their usefulness in many applications. More sophisticated sensors are thus necessary for odor identification. Classification of odors during odor localization was first achieved with an electronic nose,<sup>24,25</sup> composed of an array of semiconductor gas sensors with

Received: May 2, 2019

Accepted: August 16, 2019

Published: August 16, 2019



**Figure 1.** (a) SEM image of nanosphere assemblies that comprise the SERS sensors. (b) Schematic of SERS sensor arrays and example of multi-analyte placement with respect to arrays. This schematic depicts the case where MBZT is exposed to the left of the sensor array and BZT is exposed to the right of the sensor array. Representative SERS spectra of a BZT, MBZT mixture acquired from the SERS sensor in the (c) bottom left of the array and (d) bottom right of the array. SERS spectra are acquired with a 785 nm diode laser at 760  $\mu$ W and 0.3 s exposure time. (e) Schematic of the resulting input into the model that is constructed from NMF decomposition of the SERS spectra acquired across the sensor array.

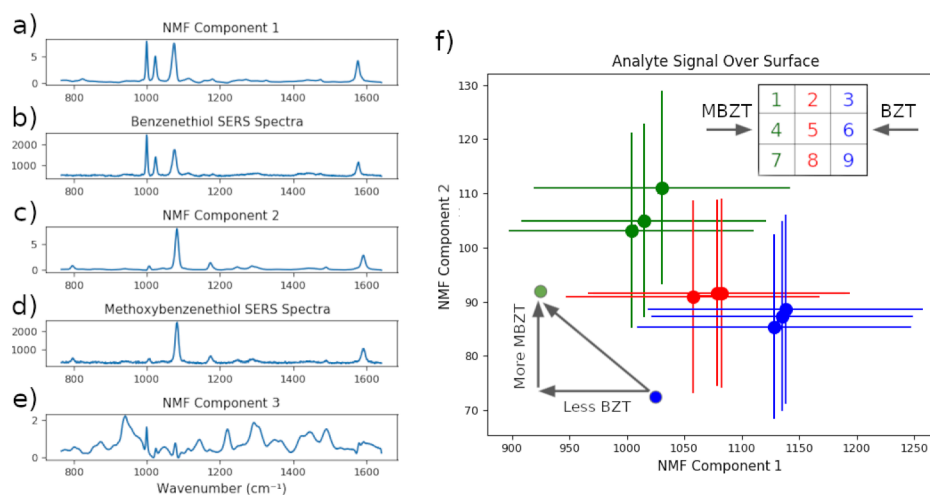
differing composition. Electronic noses have been shown to be very effective at identifying odors in a range of applications, particularly food<sup>26</sup> and health applications.<sup>27</sup> Nevertheless, electronic noses suffer from two limitations: (1) the number of unique sensors is often as little as four,<sup>24</sup> making classification of similar odors challenging and (2) a limit of detection (LOD) of analytes in the parts per million range is typical for electronic noses that are used as an odor compass.<sup>28,29</sup> In comparison, many biological systems have detection limits reaching the parts per trillion range.<sup>30</sup> The achievable LODs of typical artificial odor compasses have limited their applications to plume detection. This necessitates the use of fans to draw in odor plumes, mechanical rotation<sup>31</sup> of the odor compass apparatus, and a robotic system to travel through large areas in order to identify the true odorant direction, greatly slowing the odor localization process.<sup>5,6</sup>

We present the use of surface-enhanced Raman scattering (SERS) sensors to obviate many of the challenges associated with odor source direction identification. SERS brings two main benefits: (1) the use of vibrational spectroscopy enables superior differentiation of even extremely similar analytes<sup>32</sup> and (2) SERS has extremely low LODs of molecules, with typical operation in the parts per billion, and can reach LODs as low as 1 ppt with the use of statistical analysis and appropriate nanoarchitectures that can achieve single molecule SERS.<sup>33</sup> In this work, we demonstrate that a passive (i.e., without fans or motors) SERS sensor array can identify and differentiate the direction of multiple odor sources with similar vibrational spectra relying only on diffusion of analytes to reach the array. We use a machine learning approach to analyze the SERS spectra, which has been shown to be useful in this context.<sup>34–37</sup> Raman spectra are first analyzed with non-negative matrix factorization (NMF) to differentiate the analytes, benzenethiol (BZT) and 3-methoxybenzenethiol (MBZT), via characteristics in vibrational spectra. Then, various machine learning classifiers are compared in terms of

their ability to correctly identify the odor source direction via evaluation of SERS data from the sensor array. Machine learning interpretation of signals from sensor arrays realizes a marked improvement of well over 20% classification accuracy compared to heuristics typically used for odor compasses. The best in class models, support vector machine (SVM) classifier and convolutional neural network (CNN), correctly identify the direction of one or two odor sources over 90% of the time versus 4.2% chance. Models are also evaluated with  $k$ -fold cross-validation, and categorical accuracy achieves a cross-validation standard deviation of less than 0.2%. Transfer learning is implemented on the single odor source problem. Specifically, a CNN model and a fully connected artificial neural network (ANN) model are trained using a BZT dataset, and the weights of these pretrained models are fine-tuned on an MBZT dataset. This strategy significantly reduces the amount of data necessary to achieve good odorant localization, with one-shot accuracies of 86.1 and 84.5% for the CNN and ANN models, respectively. Finally, we demonstrate that this approach can be used to detect the presence of *Escherichia coli* biofilms as well as identify its source direction by tracking volatile organic compounds (VOCs) from the bacteria.

## RESULTS AND DISCUSSION

**SERS Sensor Array Fabrication and Validation.** Passive odor compasses relying on diffusion of analytes to sensor surfaces require significantly better sensitivity, LOD, and precision compared to typical plume-based detection because of both lower gas concentrations and smaller differences in analyte concentration from point to point. Uniformity of signal response across a passive sensor is critical to ensure variations are related to analyte concentration rather than variable sensor responses. In order to fabricate nanosensors with spatially uniform response, we use two-dimensional (2D), physically activated chemical (2PAC) self-assembly to produce SERS sensors with billion-fold and uniform Raman scattering



**Figure 2.** NMF components determined from analyte training datasets. NMF components (a) 1, (c) 2, and (e) 3 most strongly correspond to (b) neat BZT SERS spectrum, (d) neat MBZT SERS spectrum, and (e) background spectrum, respectively. (f) NMF scores of the BZT and MBZT components in the sensor array. Upper right inset: schematic of the SERS sensor array and location of odor source for the plotted data. Circles correspond to the average value of the 819 test spectra acquired in each SERS sensor in the array. The lines represent one standard deviation of NMF values within a sensor element. Lower left inset: schematic of the trajectory of analyte signal across the sensor array.

enhancements over large areas.<sup>38</sup> 2PAC fabrication involves seeded growth of close-packed gold nanosphere assemblies, described in detail elsewhere.<sup>34,38</sup> Briefly, nanosphere seeds are driven to deposit on select regions of a copolymer-coated Si electrode by electrophoresis. After seeding, the interaction between the electrical double layers of the electrode and the nanospheres drives electrohydrodynamic flow. This lateral, attractive flow entrains nearby particles, forming close-packed nanosphere assemblies with carbodiimide-mediated cross-linking between carboxylic acid groups on adjacent nanoparticles. 2PAC results in a uniform gap spacing of approximately 0.9 nm, which leads to reproducible SERS enhancements and thereby signals. A scanning electron microscopy (SEM) image of a 2PAC fabricated SERS surface is depicted in Figure 1a. One may observe discrete, close-packed nanosphere assemblies. Discrete assemblies have greater field enhancements than 2D close-packed films, while still having a relatively dense collection of hotspots for analyte molecules to reside. The reproducibility of the gap spacings due to chemical cross-linking and the rotational invariance of the close-packed oligomers enable the surfaces to be optically uniform within the laser spot size across large sample surfaces.<sup>39</sup> We have previously used 2PAC fabricated sensors for the detection of biofilm formation via sensing bacterial metabolites; the demonstrated LOD is 100 parts per trillion for pyocyanin, a metabolite of *Pseudomonas aeruginosa*.<sup>34</sup>

In this work, we focus on using SERS sensors to not only detect but also identify the direction of analyte sources. We show that 2PAC fabricated SERS sensors in the vicinity of liquid phase analyte sources are sensitive to gradients arising from diffusion-based gas transport of the vapor. Analytes from the gas phase will chemisorb onto nanosphere assemblies, and the spatial concentration gradient measured across sensors is used to locate the source of the analyte. In order to ensure that diffusive gas transport of analytes dominates the measured concentration gradient (as opposed to fan-driven plume transport typically used in odor compasses), we place SERS sensors in a  $3 \times 3$  array into a sealed desiccator at a pressure of 0.5 atm. The arrangement of SERS sensors and analyte source is depicted schematically in Figure 1b. A 5  $\mu$ L drop of neat

analyte is placed on one side of the array. Analytes used in this work, BZT and MBZT, are chosen because of the similarity of their SERS spectra and affinity to chemisorb on Au. It is also worth noting that BZT is an important molecule in the agrochemical and pharmaceutical industries.<sup>40</sup> Both analytes are in the liquid phase and have vapor pressures of 0.13 and 1.33 kPa at room temperature for BZT and MBZT, respectively. Analytes are exposed to SERS sensors for 15 min before characterizing the sensor response ex situ. For SERS data acquisition, nine Raman maps (one from each sensor in the array) are acquired in  $128 \mu\text{m} \times 128 \mu\text{m}$  regions across each sensor surface. The point-to-point reproducibility across a sensor allows for acquisition of a higher quality dataset than repeated illumination of a single spot. It is known that repeated, prolonged laser illumination of a diffraction limited spot on a SERS surface can lead to photodegradation of nanosphere assemblies because of the high temperature in the hotspots.<sup>41</sup> Also, when using sensors having hotspots with high signal enhancements, it is essential to sample many regions to properly train the model to account for small signal variations, which may arise because of variations in analyte chemisorption in hotspots.

Characteristic SERS spectra are plotted in Figure 1c,d. Preprocessing of SERS spectra before machine learning analysis is essential to aid in distinguishing analyte signals from other molecular signals that may emerge from the sample environment and will limit the generalizability of machine learning models for environmental sensing. The processed spectra are then divided into training and validation sets in an 80/20 split. Next, the training dataset is used to train NMF dimensional reduction (when closure, unimodality, and local rank constraints are used with alternating least squares optimization, NMF is called multivariate curve resolution, which is also used in spectroscopy).<sup>42,43</sup> The dimension of spectra is reduced from 1011 wavenumber features to 3 NMF scores. Finally, a spectrum from each sample in the nine sample grid are bundled into a  $3 \times 3$  matrix for each NMF score; the value of each spectra bundle is scaled to a mean of 0 and variance of 1 to improve model convergence.<sup>44</sup> These NMF score matrices are depicted schematically in Figure 1e.

**Visualization of Odorant Chemisorption on SERS Surfaces.** An intrinsic advantage of SERS over comparable methods, such as an electronic nose, is that the signal is composed of the vibrational spectra of analyte molecules enabling visual matching of signal contributions from individual analytes to the observed SERS spectra. In a complex mixture, SERS signals are additive; thus, NMF is well suited to isolate molecular signals in these complex spectra. The non-negativity constraint results in a decomposition of the spectra into a parts-based representation with contributions from different analytes primarily emerging as the different parts (components).<sup>45</sup> Figure 2a,c,e depicts plots of the three NMF components extracted from the total training dataset. The training dataset contains spectra of both odor sources, spectra of BZT only, and spectra of MBZT only. Comparing raw SERS spectra from sensors only exposed to BZT or only to MBZT, plotted in Figure 2b,d, respectively, one can easily identify NMF component 1 as mainly corresponding to the signal caused by BZT and NMF component 2 mainly corresponding to the signal caused by MBZT. NMF component 3 is associated with background signals emerging from the sensor surface chemistry (ligands on the gold nanospheres, polymer template, etc.) and molecules in the ambient environment because of the ultralow detection capacity of the sensor.

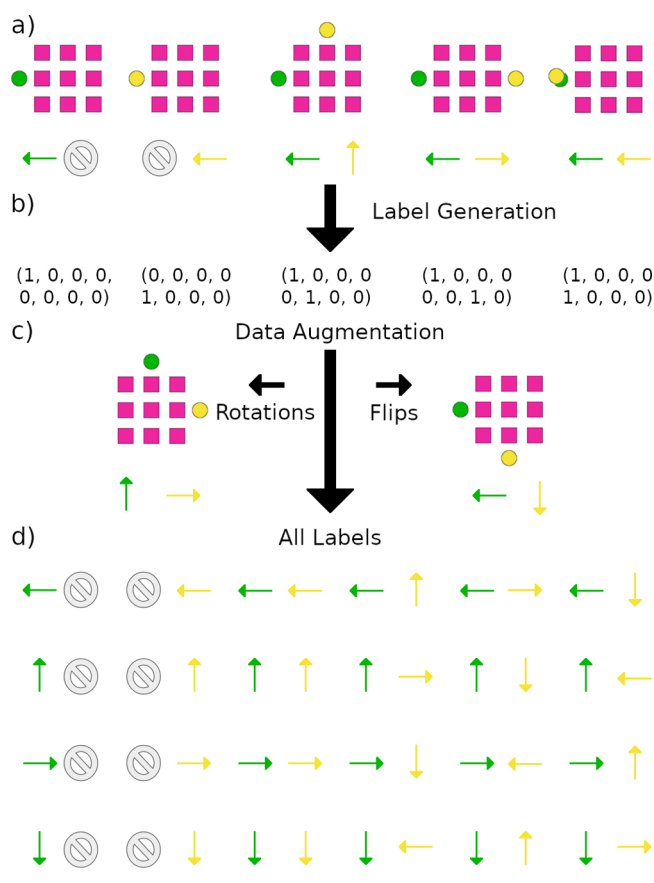
Figure 2f depicts the values of NMF scores extracted from sensor elements in the array exposed to MBZT and BZT sources emerging from opposite directions. The average values of NMF scores acquired from a sensor element in the same column in the grid, which are expected to have similar scores because of the symmetry of the deposition, are within 10% of one another. In order to evaluate the uniformity of the sensor surface, an overnight soak in 0.1 mMol BZT is performed. This results in a self-assembled monolayer of analytes on gold nanoparticles across a sensor surface. The data in Figure S1 show that the spectra have an 11.1% relative standard deviation (RSD) in their 1572  $\text{cm}^{-1}$  vibrational mode, nearly half of the 20% RSD reported for state-of-the-art top-down fabricated SERS substrates.<sup>46</sup> More variance is observed within a sensor element's Raman map in the two-source odor compass configuration, with the largest RSD approaching 20% for BZT, depicted in Figure 2f. We attribute this variance to two effects: (1) chemisorption of analytes on gold nanosphere assemblies is submonolayer and will vary with proximity to source and (2) leakage of the NMF components into one another. Leakage may be observed by comparison of neat analyte spectra with their associated NMF components.

From Figure 2f, one observes that a clear trajectory across the SERS sensors in the array emerges in the NMF scores, clearly reflecting the location of analytes. The NMF score increases for sensor elements in the array closer to their corresponding analyte. Although promising for the general multiple source problem, the variance in NMF score complicates source direction assignment and will be further complicated when the odor sources are separated by just 90°. Thus, instead of simply comparing average sensor element contributions from a row or column on the sensor array to different NMF components, a more holistic, supervised machine learning method must be used to achieve high prediction accuracy for analyte location.

**Model Performance on Multisource Task.** While supervised machine learning methods have been applied to the problem of tracking an odor source using a mobile odor compass, comparatively little effort has been put into

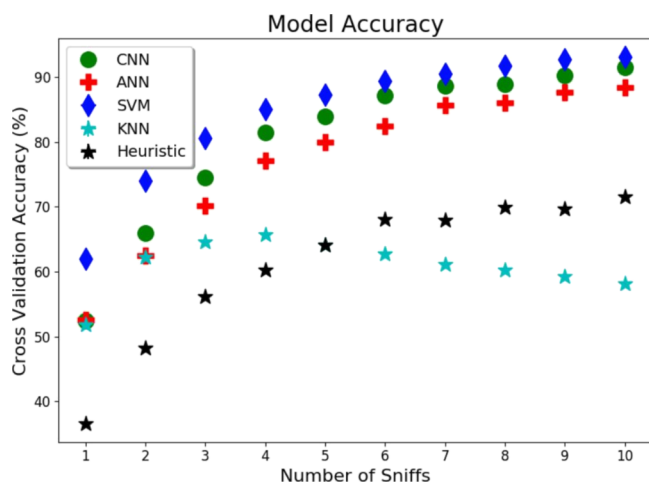
supervised methods of identifying odor source direction from a stationary measurement. We evaluate four machine learning models and one heuristic to identify the model best fit for this task, a 2D CNN, a fully connected ANN, a SVM classifier, and a  $k$ -nearest neighbors (KNN) classifier. The label generation for odor location in our supervised approach is depicted in Scheme 1. Specifically, we consider the one or two odor source

**Scheme 1.** (a) Schematic of the Deposition Conditions Used in This Work and (b) Their Corresponding Labels; (c) Schematic of the Data Augmentation Used in This Work, Where Small Random Numbers Are Also Added to the NMF Scores of the Training Data; and (d) Schematic of 24 Labels Possible in This Work



problem with distinct odors and a location without multiple sources of the same odor and a resolution of 90°, resulting in 24 possibilities. The CNN model uses the full 3 × 3 pixel spatial relationships of the input data, while the inputs are flattened for the other classifiers. The architectures of neural networks are described in the Methods section. This results in a multiclass, multilabel problem as one or two labels may be true; thus, both neural networks are trained with a binary cross-entropy loss function. We simplify the problem for the SVM and KNN models, evaluating them with labels transformed from a binary representation of the four possible directions per odor to 24 labels total, one for each possible outcome. The heuristic first evaluates the variance of each input matrix to determine whether or not an odor exists based on a user-defined threshold, and then the average value from the rows or columns is compared with its opposite and the largest difference is used to identify the odorant direction for each matrix. Each model performance and the heuristic are

evaluated using the categorical accuracy and plotted in Figure 3a.



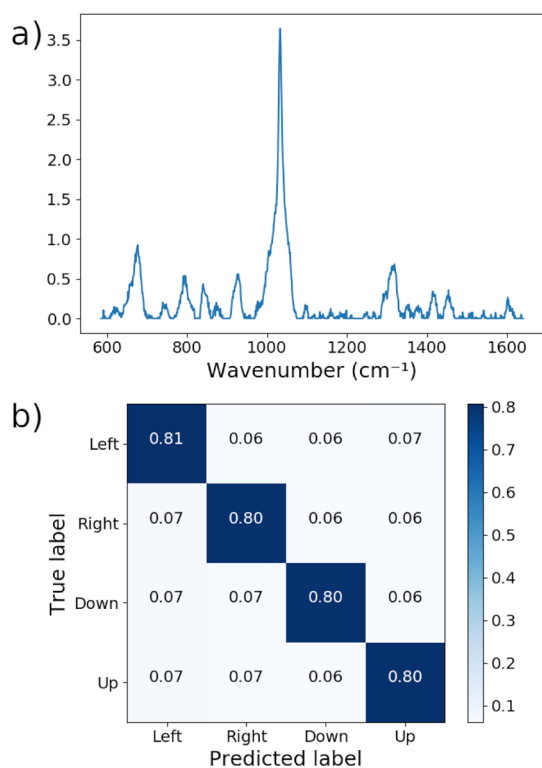
**Figure 3.** Cross-validation accuracy of the models used. Error bars are obstructed by the markers but are less than 0.45% in all cases. For a prediction to be considered correct, the direction of both odors must be correctly identified by the model.

The prediction accuracy can be greatly improved by simultaneously considering multiple input data matrices. Averaging sensor inputs has been an important method for improving sensor precision and is ubiquitous in natural olfaction in the form of sniffing, where it greatly improves the accuracy of predictions made by mammals<sup>47</sup> and spike timing-dependent plasticity models of a moth olfactory bulb.<sup>48</sup> We implement an averaging procedure, with each extra input referred to as a sniff, and plot the categorical accuracy over sniffs in Figure 3a. Explicitly, each input, or sniff, is an additional SERS spectral measurement that is NMF-transformed to  $3 \times 3$  BZT and MBZT score matrices. Each additional sniff comes from the same sensor grid array. Extra sniffs are added to the inputs as time series in long short-term memory (LSTM) layers for the CNN model and are added to lengthen the input vectors for the other models. The inputs are averaged element-wise for the heuristic.

All models and the heuristic greatly benefit from the averaging procedure, with the 7+ sniff SVM and 9+ sniff CNN achieving >90% categorical accuracy. The CNN model receives a disproportionate boost in model performance from the extra sniffs, especially the 2–4 sniff models because of the integration into a LSTM as a time series, whereas the KNN model actually declines in accuracy beyond 4 sniffs because of the increased sparsity of the input space. Although the sniffing procedure increases the acquisition time, acquiring one SERS matrix only consumes 0.9 s of exposure time, so even a 10 sniff input requires only 9 s of exposure time, which is significantly less than is necessary for many plume-based odor compasses;<sup>6</sup> for the latter, the typical measurement time for one of several iterations needed for odor location is on the order of 20 s, for example, in Ishida et al.<sup>4</sup> These results demonstrate that the passive, SERS-based approach to odor source direction identification significantly reduces the identification time compared to fan-driven approaches typically used. Highly confident predictions can be made in less than half the time using the SERS-based approach.

For this two odor source problem, we have used two molecules which chemisorb onto gold nanoparticles on the sensor surface. Chemisorption is beneficial to ensure that molecular concentration does not change over the time necessary to acquire the large training dataset, that is, 20 480  $3 \times 3$  NMF score matrices per analyte in total, used. Large datasets are necessary to train complex neural network models with large numbers of parameters to avoid overfitting. These models have a key advantage over traditional, less data-intensive models, that is, transfer learning. When using transfer learning, a neural network model is first fully trained with a large dataset, and then network weights are fine-tuned with a new, smaller dataset, such that high performance can be achieved even with few training examples. In order to demonstrate this ability, a BZT-only dataset (6553 examples per class) is used to train single odor source (4 label) ANN and CNN models having an architecture identical to those used in the above-described two odor source problem. These CNN and ANN models are applied to a MBZT-only dataset, and the weights are fine-tuned. This process is described in more detail in the Methods section. Figure S3 depicts the accuracy of the transfer-learned MBZT model as a function of training examples, where 10 sniff examples are used. The one-shot, 10 sniff learning accuracy achieved via transfer learning is determined to be  $86.1 \pm 4.6$  and  $84.5 \pm 3.5\%$  for the transfer-learned CNN and ANN models, respectively. For reference, the fully trained CNN and ANN models were 94.6 and 91.5% for the CNN and ANN models, respectively, similar to the results observed for 10 sniff accuracy of the full dataset in Figure 3. Thus, using transfer learning, one is able to implement trained models to evaluate significantly smaller datasets, even one-shot datasets. Training models with the large well-behaved two odor source dataset provides a basis for models using transfer learning to evaluate SERS-based odor compass data in more challenging experimental conditions where noise and confounding factors often contribute to signals. For example, in large rooms, factors such as lower analyte concentration (decreased signal to noise) and air currents (reduced contrast) make analysis more challenging. Transfer learning would enable better model convergence in these noisy, reduced contrast environments compared to standard Xavier weight initialization.<sup>49</sup>

**Biofilm Localization.** SERS's rich spectral information and low LOD for small molecules have made it an important technique in the detection and discrimination of bacteria by sensing the VOCs that they produce.<sup>50–52</sup> Here, we use *E. coli* as an odor source to test the efficacy of our odor compass with complex mixtures of VOCs. SERS arrays are exposed to *E. coli* VOCs in a desiccator at ambient pressure and for 6 h. The *E. coli* source is placed adjacent to the SERS grid array similarly as illustrated in Figure 1b. After exposure, the sensor elements are characterized ex situ. The resulting dataset is decomposed into NMF components which were compared to SERS spectra of *E. coli* acquired at 785 nm wavelength excitation.<sup>53</sup> The NMF component in Figure 4a is in good agreement with the previous work reporting SERS analysis of *E. coli*.<sup>53</sup> While we will discuss model performance in identifying the location of the biofilm, it is also important to note here that this NMF component provides the highest accuracy of predictions for the location of the *E. coli* biofilm. Inspection of the *E. coli* NMF component shows an important vibrational mode at  $1023 \text{ cm}^{-1}$ , an energy usually associated with ring breathing modes. The observation that the SERS spectra of volatile metabolites



**Figure 4.** (a) NMF component acquired from bacterial VOC training dataset used in bacterial odor source localization. (b) Normalized confusion matrix produced by a 1 sniff SVM model applied to the bacterial VOC test dataset.

from *E. coli* exhibit high signal intensity near this energy region is unsurprising for a few reasons. First, ring breathing modes are typically associated with larger Raman scattering cross sections than other molecular vibrations because of their delocalized electrons, so they tend to dominate SERS signal in complex mixtures.<sup>54</sup> Second, volatile molecules tend to be hydrophobic as this drives them to evaporate out of solution, so hydrophobic phenyl groups and other functional groups with ring breathing modes will be overrepresented among volatile metabolites. Third, the *E. coli* metabolome contains many metabolites with functional groups that have strong ring breathing mode vibrations.

We turn now to evaluate the SERS-based odor compass performance for bacteria localization. With the NMF component selected as described above, the spectra acquired for this experiment are then reduced to the score of this component, resulting in one  $3 \times 3$  NMF score matrix. As the SVM model had the best model accuracy for the two odor source problem, we only consider this model here, and we evaluate the dataset with a 4 class, 1 sniff SVM model. The classes are chosen to represent the four possible outcomes of the single odor source problem: bacteria left, right, up, or down. A categorical accuracy of 82.95% is achieved on the test dataset versus 25% chance. A confusion matrix of the test dataset is plotted in Figure 4b, demonstrating well-behaved predictions that are invariant with rotation. These high accuracies for bacteria localization demonstrate the advantages for a SERS-based odor compass. First, the large amount of spectral information in the “fingerprint region” of the vibrational spectrum enables us to identify the complex mixture of molecules associated with the volatile metabolites of *E. coli*. Second, it shows that relatively low concentrations of

the various molecules produced by *E. coli* can be detected, which are typically in the parts per million and parts per billion concentration range.<sup>53</sup> Third, it shows that localized populations of *E. coli* cells can be detected, as these experiments were performed with *E. coli* biofilms rather than cell cultures in suspension. This demonstrates that SERS-based odor compasses can be used in practical applications such as locating the presence of a bacterial biofilm.

## CONCLUSIONS

In this work, we have demonstrated the efficacy of a diffusion-based odor compass for the multiple odor source problem using SERS sensors. Specifically, we have implemented a sensor grid array to identify small spatial variations in analyte chemisorption and used various machine learning models to identify multiple analyte source directions. We have shown that SVM classifier and CNN models can achieve greater than 90% categorical accuracy for this multiple source problem with 90° resolution. Transfer learning was demonstrated to attain good one-shot learning accuracy for the single odor source problem, with a BZT pretrained CNN model achieving 86.1% accuracy on an MBZT dataset. We have further shown that the sensor grid array can be used to locate complex odor sources such as *E. coli*. This work uses SERS as the gas sensor in an odor compass and paves the way for its use in passive Internet of things devices and mobile odor compasses. We envision integrating this strategy with newly emerging waveguide excitation of SERS sensors for device miniaturization and continuous monitoring.<sup>55,56</sup> This approach will greatly benefit from the transfer learning methodology introduced in this work, which enables accurate models to be trained quickly. These devices promise to make a material impact on people’s health by identifying and locating toxic gases and VOCs produced by indoor pathogens.

## METHODS

**Materials.** Random copolymer poly(styrene-*co*-methyl methacrylate)- $\alpha$ -hydroxyl- $\omega$ -tempo moiety (PS-*r*-PMMA) ( $M_n = 7400$ , 59.6% PS) and diblock copolymer poly(styrene-*b*-methyl methacrylate) (PS-*b*-PMMA) ( $M_n = 170$ - $144 \text{ kg mol}^{-1}$ ) were purchased from Polymer Source, Inc. (Dorval, Canada). Forty nanometer diameter gold nanospheres were purchased from nanoComposix (San Diego, CA, USA). Si(001) wafers with a resistivity of  $0.004 \Omega \text{ cm}$  were purchased from Virginia Semiconductor (Fredericksburg, VA, USA). Hydrofluoric acid (HF) was purchased from Fisher Scientific (Pittsburgh, PA, USA). 2-(*N*-Morpholino)ethanesulfonic acid (MES) 0.1 M buffer, 1-ethyl-3-[3-(dimethylamino)propyl]carbodiimide hydrochloride (EDC), and *N*-hydroxy sulfosuccinimide (*s*-NHS), dimethyl sulfoxide (DMSO), ethylenediamine, BZT, toluene, ethanol, isopropyl alcohol (IPA), potassium carbonate, and 52-mesh Pt gauze foil were purchased from Sigma-Aldrich (St. Louis, MO, USA). Nanopure deionized water ( $18.2 \text{ M}\Omega \text{ cm}^{-1}$ ) was obtained from a Milli-Q Millipore system.

**SERS Sensor Fabrication.** Sensor fabrication has been described in depth in previous work.<sup>34,38</sup> First, block copolymer templates for nanoparticle attachment are prepared. Random PS-*b*-PMMA block copolymer is spin-coated onto a HF-cleaned (the potential of HF to cause severe injury mandates extreme caution during usage),  $0.004 \Omega \text{ cm}$  Si wafer and annealed for 3 days, followed by a toluene rinse and spin coating of lamella forming a PS-*b*-PMMA block copolymer and further 3 days of annealing, as described elsewhere.<sup>57–59</sup> Then, the PMMA regions are selectively functionalized with amine end groups by immersing the entire substrate in DMSO and then in ethylenediamine/DMSO solution (5% v/v), both for 5 min without rinsing

between steps. The functionalized template is then washed with IPA for 1 min, dried under nitrogen, and used immediately.

Au nanosphere assemblies are deposited on the surface as follows: a Au nanosphere solution (0.1 mg/mL, 3 mL) is added to a 10 mL glass beaker. Then, s-NHS (20 mM) in a MES (0.1 M) buffer (35  $\mu$ L) is added to the beaker and swirled. Afterward, EDC (8 mM) in a MES (0.1 M) buffer (35  $\mu$ L) is added to the beaker and swirled. This beaker is placed on a hot plate and brought to 60 °C. Next, a 1 cm  $\times$  1 cm functionalized template-coated Si substrate is placed into the solution vertically and held in place with alligator clips. One millimeter away from the substrate, a 1 cm  $\times$  1 cm Pt mesh is placed into the solution vertically. A dc power supply is used to apply a voltage of 1.2 V across the Pt mesh and substrate for 10 min. The substrate, Pt mesh, and beaker are rinsed with IPA for 1 min and dried under nitrogen. This process is repeated with the same substrate and fresh nanosphere solution as described above but with 25  $\mu$ L of EDC and s-NHS solution.

**Analyte Deposition.** SERS sensors are exposed to analytes with nine sensors arranged in a grid. Depositions are performed in a sealed desiccator with a 10 cm diameter. A 5  $\mu$ L drop of neat analyte is placed on a glass slide at the edge of the desiccator; when a second analyte is used, it is similarly placed on a glass slide at the edge of the desiccator either 90° offset, 180° offset, or adjacent to the first analyte. The sensor grid array is organized such that each sensor is placed 3 cm away from other sensors to form a square grid, with an analyte–sensor spacing of 2 cm for the closest sensor. The desiccator was then sealed and brought to 0.5 atm for 5 min, followed by 15 min of static deposition.

**Characterization.** Au nanoparticle assemblies are imaged with a Magellan XHR SEM (FEI). All Raman spectroscopy measurements are conducted using a confocal Renishaw inVia micro Raman system with a 785 nm diode laser, a laser power of 760  $\mu$ W, an exposure time of 0.3 s, and a 50 $\times$  air objective with a 0.75 NA. Raman maps are collected with a spacing of 2  $\mu$ m between points. For the sensor grid array, one 64  $\times$  64 pixel Raman map is acquired per sensor. As depicted in Scheme 1, five analyte deposition configurations are necessary to provide sufficient data for the two odor source problem. This results in a total of 184 320 spectra acquired for the two odor source problem and 36 864 spectra for the single bacteria source problem discussed in Figure 4.

**Spectra Preprocessing.** Raman spectra preprocessing was performed using Python 3.3 programming language. Savitzky–Golay smoothing was implemented with Scikit-learn using an 11 pixel window and polynomial order 3. Asymmetric least squares background subtraction was implemented in NumPy with  $\lambda = 10\,000$ ,  $p = 0.001$ . NMF was implemented with Scikit-learn, three components, and default settings, trained only on the dataset segregated for use in training the models. This trained NMF transformation was used to reduce the dimension of all data to 2. NMF-reduced maps from each point on the sensor grids were then rearranged into a 3  $\times$  3 matrix, with two NMF scores per map per data point. Each data point was then scaled to a variance of 1 and mean of 0. This results in 20 480 NMF score matrices for the two odor source problem and 4096 NMF score matrices for the single bacteria source problem.

Data augmentation was performed by rotating the rank 2 tensors that compose each data point 90° and similarly rotating the labels. A Gaussian random number ( $\mu = 0$ ,  $\sigma =$  standard deviation of that component extracted from each deposition condition  $\times 0.1$ ) is added to each component for the training data. This process augments the data to 163 840 NMF score matrices for the two odor source problem and 32 768 NMF score matrices for the single bacteria source problem. These data are then stacked along a third dimension, called “sniffs” producing a rank 4 tensor with shape (sniffs, sensor grid axis 1, sensor grid axis 2, number of analyte). For the SVM, KNN, and ANN models, these inputs are flattened into a vector before being fed into the model. For the heuristic, the components are averaged along the sniff axis and then flattened, and each analyte is considered separately.

**Compass Models.** The SVM classifier and KNN classifier were both implemented in Scikit-learn using default settings. The ANN and CNN models were implemented in Keras. The ANN model is

composed of an input layer, three fully connected layers with succeeding dropout layers, and one fully connected output layer with eight sigmoid nodes. The CNN model is composed of an input layer, 4 time-distributed 2D convolutional layers performed over the sensor grid axes with succeeding batch normalization layers, and then a time-distributed 2D maximum pooling layer with a pool size of 2  $\times$  2 and strides of 1  $\times$  1. The pooling layer is followed by 2 more time-distributed 2D convolutional layers and a time-distributed flattening layer. Next, a LSTM layer is used across “sniff” axis. Following that layer are three fully connected layers with succeeding batch normalization layers and dropout layers, outputting ultimately to a fully connected output layer with eight sigmoid nodes. Each output node on the ANN and CNN models is associated with a direction and analyte, so binary cross-entropy is used as the loss function. Early stopping and reduction of learning rate on a performance plateau are implemented using test loss as the metric. The heuristic identifies odorant direction in two steps: first identifying the odorant and then the direction. First, the variance of each input matrix is used to determine whether or not an odor exists based on a user-defined threshold. Then, the average value from the rows or columns is compared with its opposite, and the largest difference is used to identify the odorant direction for each matrix.

The generalizability of each model is performed using  $k$ -fold cross-validation. First, 20% of the data are removed as the validation dataset. The remaining 80% of the data are used to train the NMF transformation, which is also applied to the validation dataset. Then,  $k$ -fold cross validation is implemented on the remaining 80% of the data using Scikit-learn, with the number of folds equal to 5. Thus, the training dataset for each fold included 104 857 NMF score matrices, the test dataset for each fold included 26 214 NMF score matrices, and the validation dataset included 32 768 NMF score matrices. The training fold is used to train each model. Each test fold is used to train the early stopping for the ANN and CNN models but are discarded for the SVM and KNN models. Finally, each of the trained models is evaluated on the validation dataset; the average and standard deviation across the 5 folds are plotted in Figure 3.

The code used in this work will be made available on GitHub.

**Transfer Learning.** Transfer learning is achieved using the same ANN and CNN architectures described above in the Compass Models section. First, the BZT only dataset is split into 80% training data and 20% testing data, and an NMF transformation is trained using the training dataset. The spectra are grouped into 3  $\times$  3 matrices corresponding to the different elements in the sensor grid array and are reduced to the first NMF component that corresponds to BZT. Spectra are preprocessed, and data are augmented as described in the earlier Methods subsections. The models are trained using the 3  $\times$  3 BZT NMF scores set aside for training, and early stopping is implemented with the remaining testing dataset. The weights of the neural network models are then saved. Next, the same procedure (without training models) is repeated for the MBZT-only dataset. Then, 10 3  $\times$  3 MBZT NMF score matrices per class (4 in this case) are randomly selected for each example used (which ranges between 1 and 10 in Figure S3). These training examples are then used to train ANN and CNN models that have their weights set from the already trained BZT models. The accuracy is evaluated based on the remaining examples. This process is repeated 10 times per model per different numbers of examples to identify an average accuracy and the standard deviation of the accuracies. When only one example is used to train a model, this is referred to as one-shot learning. These results are plotted in Figure S3.

**Static Biofilm Preparation and Characterization.** All materials for bacterial culture were purchased from VWR International unless otherwise noted. *E. coli* MC4100 (*E. coli* K12 derivative *araD*  $\Delta$ (*argF-lac*)U169 *rpsL* *relA* *flbB* *deoC* *ptsF* *rbsR*)<sup>60</sup> frozen stock was streaked onto lysogeny broth (LB) (Lennox) agar (1.5%) plates and grown overnight at 37 °C. One colony was harvested, inoculated into 2 mL of liquid LB, and grown overnight at 37 °C on an orbital shaker at 200 rpm. The overnight culture was diluted with 10 g/L tryptone (BD Bacto Tryptone) to an optical density of 0.02. Two milliliters of the diluted culture was added to each well of the sterile six-well plate. One

18 × 18 mm coverslip was submerged in each well and was incubated at room temperature (22 °C) for 24 h. Three of the coverslips and their inoculum solution were transferred into a sterile Petri dish that was placed into a covered 10 cm diameter desiccator. A sensor grid was then placed in the desiccator as described above, and the system was kept isolated for 4 h. These sensors were then removed, and SERS measurements were performed as described above.

## ■ ASSOCIATED CONTENT

### 📄 Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: [10.1021/acssens-9b00809](https://doi.org/10.1021/acssens.9b00809).

SERS spectra from an overnight BZT treatment depicting signal enhancement uniformity and SERS spectra from across a sensor grid array used as training data; transfer learning model accuracy as a function of training examples; and multilinear regression of NMF scores to determine the LOD of aqueous 2-aminoacetophenone (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Author

\*E-mail: [rragan@uci.edu](mailto:rragan@uci.edu).

### ORCID

Allon I. Hochbaum: [0000-0002-5377-8065](https://orcid.org/0000-0002-5377-8065)

Regina Ragan: [0000-0002-8694-5683](https://orcid.org/0000-0002-8694-5683)

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

The authors acknowledge the National Science Foundation EECS-1449397 for funding this work. W.J.T. acknowledges support from the National Science Foundation BEST IGERT Program (NSF DGE-1144901). S.R. and A.I.H. acknowledge support from the Air Force Office of Scientific Research grant FA9550-17-1-0193 and NSF award CHE-1808332. The authors also acknowledge the use of the facilities within the Laser Spectroscopy Facility and the UC Irvine Materials Research Institute (IMRI) at the University of California, Irvine.

## ■ REFERENCES

- (1) Conley, S.; Franco, G.; Faloona, I.; Blake, D. R.; Peischl, J.; Ryerson, T. B. Methane Emissions from the 2015 Aliso Canyon Blowout in Los Angeles, CA. *Science* **2016**, *351*, 1317.
- (2) Rozas, R.; Morales, J.; Vega, D. Artificial Smell Detection for Robotic Navigation. *Fifth International Conference on Advanced Robotics Robots in Unstructured Environments*, 1991; Vol. 2, pp 1730–1733.
- (3) Nakamoto, T.; Ishida, H.; Moriizumi, T. An Odor Compass for Localizing an Odor Source. *Sens. Actuators, B* **1996**, *35*, 32–36.
- (4) Ishida, H.; Kobayashi, A.; Nakamoto, T.; Moriizumi, T. Three-Dimensional Odor Compass. *IEEE Trans. Robot. Autom.* **1999**, *15*, 251–257.
- (5) Ishida, H.; Wada, Y.; Matsukura, H. Chemical Sensing in Robotic Applications: A Review. *IEEE Sens. J.* **2012**, *12*, 3163–3173.
- (6) Kowadlo, G.; Russell, R. A. Robot Odor Localization: A Taxonomy and Survey. *Int. J. Robot. Res.* **2008**, *27*, 869–894.
- (7) Braitenberg, V. *Vehicles: Experiments in Synthetic Psychology*; MIT Press, 1986.
- (8) de Croon, G. C. H. E.; O'Connor, L. M.; Nicol, C.; Izzo, D. Evolutionary Robotics Approach to Odor Source Localization. *Neurocomputing* **2013**, *121*, 481–497.
- (9) Villarreal, B. L.; Olague, G.; Gordillo, J. L. Synthesis of Odor Tracking Algorithms with Genetic Programming. *Neurocomputing* **2016**, *175*, 1019–1032.
- (10) Jiang, P.; Wang, Y.; Ge, A. Multivariable Fuzzy Control Based Mobile Robot Odor Source Localization via Semitensor Product. <https://www.hindawi.com/journals/mpe/2015/736720/abs/> (accessed Oct 24, 2018).
- (11) Husni, N. L.; Handayani, A. S.; Nurmaini, S.; Yani, I. Odor Localization Using Gas Sensor for Mobile Robot. *2017 4th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, 2017; pp 1–6.
- (12) Vergassola, M.; Villermaux, E.; Shraiman, B. I. 'Infotaxis' as a Strategy for Searching without Gradients. *Nature* **2007**, *445*, 406–409.
- (13) Luo, B.; Meng, Q.-H.; Wang, J.-Y.; Zeng, M. A Flying Odor Compass to Autonomously Locate the Gas Source. *IEEE Trans. Instrum. Meas.* **2018**, *67*, 137–149.
- (14) Neumann, P. P.; Hernandez Bennets, V.; Lilienthal, A. J.; Bartholmai, M.; Schiller, J. H. Gas Source Localization with a Micro-Drone Using Bio-Inspired and Particle Filter-Based Algorithms. *Adv. Robot.* **2013**, *27*, 725–738.
- (15) Grasso, F. W.; Consi, T. R.; Mountain, D. C.; Atema, J. Biomimetic Robot Lobster Performs Chemo-Oriented in Turbulence Using a Pair of Spatially Separated Sensors: Progress and Challenges. *Robot. Auton. Syst.* **2000**, *30*, 115–131.
- (16) Consi, T. R.; Atema, J.; Goudey, C. A.; Cho, J.; Chryssostomidis, C. AUV Guidance with Chemical Signals. *Proceedings of IEEE Symposium on Autonomous Underwater Vehicle Technology (AUV'94)*, 1994; pp 450–455.
- (17) Miyatani, I.; Ishida, H. Active Stereo Nose: Using Air Curtain to Enhance the Directivity. *2010 IEEE SENSORS*, 2010; pp 1522–1525.
- (18) Duckett, T.; Axelsson, M.; Saffiotti, A. Learning to Locate an Odour Source with a Mobile Robot. *Proceedings 2001 ICRA. IEEE International Conference on Robotics and Automation (Cat. No. 01CH37164)*, 2001; Vol. 4, pp 4017–4022.
- (19) Kazadi, S.; Goodman, R.; Tsikata, D.; Green, D.; Lin, H. An Autonomous Water Vapor Plume Tracking Robot Using Passive Resistive Polymer Sensors. *Aut. Robots* **2000**, *9*, 175–188.
- (20) Hayes, A. T.; Martinoli, A.; Goodman, R. M. Distributed Odor Source Localization. *IEEE Sens. J.* **2002**, *2*, 260–271.
- (21) Ishida, H.; Tokuhito, T.; Nakamoto, T.; Moriizumi, T. Improvement of Olfactory Video Camera: Gas/Odor Flow Visualization System. *Sens. Actuators, B* **2002**, *83*, 256–261.
- (22) Deveza, R.; Thiel, D.; Russell, A.; Mackay-Sim, A. Odor Sensing for Robot Guidance. *Int. J. Robot. Res.* **1994**, *13*, 232–239.
- (23) Kuwana, Y.; Nagasawa, S.; Shimoyama, I.; Kanzaki, R. Synthesis of the Pheromone-Oriented Behaviour of Silkworm Moths by a Mobile Robot with Moth Antennae as Pheromone Sensors. This Paper Was Presented at the Fifth World Congress on Biosensors, Berlin, Germany, 3–5 June 1998. *Biosens. Bioelectron.* **1999**, *14*, 195–202.
- (24) Trincavelli, M.; Coradeschi, S.; Loutfi, A. Classification of Odours with Mobile Robots Based on Transient Response. *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2008; pp 4110–4115.
- (25) Trincavelli, M.; Coradeschi, S.; Loutfi, A. Odour Classification System for Continuous Monitoring Applications. *Sens. Actuators, B* **2009**, *139*, 265–273.
- (26) Baietto, M.; Wilson, A.; Baietto, M.; Wilson, A. D. Electronic-Nose Applications for Fruit Identification, Ripeness and Quality Grading. *Sensors* **2015**, *15*, 899–931.
- (27) de Meij, T. G.; Larbi, I. B.; van der Schee, M. P.; Lentferink, Y. E.; Paff, T.; sive Droste, J. S. T.; Mulder, C. J.; van Bodegraven, A. A.; de Boer, N. K. Electronic Nose Can Discriminate Colorectal Carcinoma and Advanced Adenomas by Fecal Volatile Biomarker Analysis: Proof of Principle Study. *Int. J. Cancer* **2014**, *134*, 1132–1138.



- (28) Burgués, J.; Jiménez-Soto, J. M.; Marco, S. Estimation of the Limit of Detection in Semiconductor Gas Sensors through Linearized Calibration Models. *Anal. Chim. Acta* **2018**, *1013*, 13–25.
- (29) Szulczyński, B.; Rybarczyk, P.; Gębicki, J. Monitoring of N-Butanol Vapors Biofiltration Process Using an Electronic Nose Combined with Calibration Models. *Monatsh. Chem.* **2018**, *149*, 1693–1699.
- (30) Walker, D. B.; Walker, J. C.; Cavnar, P. J.; Taylor, J. L.; Pickel, D. H.; Hall, S. B.; Suarez, J. C. Naturalistic Quantification of Canine Olfactory Sensitivity. *Appl. Anim. Behav. Sci.* **2006**, *97*, 241–254.
- (31) Zheng, J. B.; Yang, L.; Chen, J. B.; Wang, Y. M. Study on Odor Source Localization Method Based on Bionic Olfaction. *Appl. Mech. Mater.* **2013**, *448–453*, 391–395.
- (32) Sánchez-Purrà, M.; Roig-Solvas, B.; Rodríguez-Quijada, C.; Leonardo, B. M.; Hamad-Schifferli, K. Reporter Selection for Nanotags in Multiplexed Surface Enhanced Raman Spectroscopy Assays. *ACS Omega* **2018**, *3*, 10733–10742.
- (33) de Albuquerque, C. D. L.; Sobral-Filho, R. G.; Poppi, R. J.; Brolo, A. G. Digital Protocol for Chemical Analysis at Ultralow Concentrations by Surface-Enhanced Raman Scattering. *Anal. Chem.* **2018**, *90*, 1248–1254.
- (34) Nguyen, C. Q.; Thrift, W. J.; Bhattacharjee, A.; Ranjbar, S.; Gallagher, T.; Darvishzadeh-Varcheie, M.; Sanderson, R. N.; Capolino, F.; Whiteson, K.; Baldi, P.; et al. Longitudinal Monitoring of Biofilm Formation via Robust Surface-Enhanced Raman Scattering Quantification of *Pseudomonas Aeruginosa*-Produced Metabolites. *ACS Appl. Mater. Interfaces* **2018**, *10*, 12364–12373.
- (35) Lussier, F.; Missirlis, D.; Spatz, J. P.; Masson, J.-F. Machine-Learning-Driven Surface-Enhanced Raman Scattering Optophysiology Reveals Multiplexed Metabolite Gradients Near Cells. *ACS Nano* **2019**, *13*, 1403–1411.
- (36) Dong, R.; Weng, S.; Yang, L.; Liu, J. Detection and Direct Readout of Drugs in Human Urine Using Dynamic Surface-Enhanced Raman Spectroscopy and Support Vector Machines. *Anal. Chem.* **2015**, *87*, 2937–2944.
- (37) Alharbi, O.; Xu, Y.; Goodacre, R. Simultaneous Multiplexed Quantification of Nicotine and Its Metabolites Using Surface Enhanced Raman Scattering. *Analyst* **2014**, *139*, 4820–4827.
- (38) Thrift, W. J.; Nguyen, C. Q.; Darvishzadeh-Varcheie, M.; Zare, S.; Sharac, N.; Sanderson, R. N.; Dupper, T. J.; Hochbaum, A. I.; Capolino, F.; Abdolhosseini Qomi, M. J.; et al. Driving Chemical Reactions in Plasmonic Nanogaps with Electrohydrodynamic Flow. *ACS Nano* **2017**, *11*, 11317–11329.
- (39) Rahmani, M.; Lei, D. Y.; Giannini, V.; Lukiyanchuk, B.; Ranjbar, M.; Liew, T. Y. F.; Hong, M.; Maier, S. A. Subgroup Decomposition of Plasmonic Resonances in Hybrid Oligomers: Modeling the Resonance Lineshape. *Nano Lett.* **2012**, *12*, 2101–2106.
- (40) Sun, Q.; Yang, S.-H.; Wu, L.; Yang, W.-C.; Yang, G.-F. A Highly Sensitive and Selective Fluorescent Probe for Thiophenol Designed via a Twist-Blockage Strategy. *Anal. Chem.* **2016**, *88*, 2266–2272.
- (41) Jones, S.; Sinha, S. S.; Pramanik, A.; Ray, P. C. Three-Dimensional (3D) Plasmonic Hot Spots for Label-Free Sensing and Effective Photothermal Killing of Multiple Drug Resistant Superbugs. *Nanoscale* **2016**, *8*, 18301–18308.
- (42) Felten, J.; Hall, H.; Jaumot, J.; Tauler, R.; de Juan, A.; Gorzsás, A. Vibrational Spectroscopic Image Analysis of Biological Material Using Multivariate Curve Resolution–Alternating Least Squares (MCR-ALS). *Nat. Protoc.* **2015**, *10*, 217–240.
- (43) Hong, K. Y.; de Albuquerque, C. D. L.; Poppi, R. J.; Brolo, A. G. Determination of Aqueous Antibiotic Solutions Using SERS Nanogratings. *Anal. Chim. Acta* **2017**, *982*, 148–155.
- (44) Ioffe, S.; Szegedy, C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. **2015**, arXiv:1502.03167 [cs].
- (45) Lee, D. D.; Seung, H. S. Learning the Parts of Objects by Non-Negative Matrix Factorization. *Nature* **1999**, *401*, 788–791.
- (46) Kanipe, K. N.; Chidester, P. P. F.; Stucky, G. D.; Moskovits, M. Large Format Surface-Enhanced Raman Spectroscopy Substrate Optimized for Enhancement and Uniformity. *ACS Nano* **2016**, *10*, 7566–7571.
- (47) Laing, D. G. Natural Sniffing Gives Optimum Odour Perception for Humans. *Perception* **1983**, *12*, 99–117.
- (48) Delahunt, C. B.; Kutz, J. N. Putting a Bug in ML: The Moth Olfactory Network Learns to Read MNIST. **2018**, arXiv:1802.05405 [cs, q-bio].
- (49) Pan, S. J.; Yang, Q. A Survey on Transfer Learning. *IEEE Trans. Knowl. Data Eng.* **2010**, *22*, 1345–1359.
- (50) Galvan, D. D.; Yu, Q. Surface-Enhanced Raman Scattering for Rapid Detection and Characterization of Antibiotic-Resistant Bacteria. *Adv. Healthcare Mater.* **2018**, *7*, 1701335.
- (51) Liu, Y.; Zhou, H.; Hu, Z.; Yu, G.; Yang, D.; Zhao, J. Label and Label-Free Based Surface-Enhanced Raman Scattering for Pathogen Bacteria Detection: A Review. *Biosens. Bioelectron.* **2017**, *94*, 131–140.
- (52) Bodelón, G.; Montes-García, V.; Costas, C.; Pérez-Juste, I.; Pérez-Juste, J.; Pastoriza-Santos, I.; Liz-Marzán, L. M. Imaging Bacterial Interspecies Chemical Interactions by Surface-Enhanced Raman Scattering. *ACS Nano* **2017**, *11*, 4631–4640.
- (53) Premasiri, W. R.; Moir, D. T.; Klempner, M. S.; Krieger, N.; Jones, G.; Ziegler, L. D. Characterization of the Surface Enhanced Raman Scattering (SERS) of Bacteria. *J. Phys. Chem. B* **2005**, *109*, 312–320.
- (54) Larkin, P. *Infrared and Raman Spectroscopy; Principles and Spectral Interpretation*; Elsevier, 2011.
- (55) Darvishzadeh-Varcheie, M.; Guclu, C.; Ragan, R.; Boyraz, O.; Capolino, F. Electric Field Enhancement with Plasmonic Colloidal Nanoantennas Excited by a Silicon Nitride Waveguide. *Opt. Express* **2016**, *24*, 28337–28352.
- (56) Zhao, Q.; Guclu, C.; Huang, Y.; Capolino, F.; Ragan, R.; Boyraz, O. Plasmon Optical Trapping Using Silicon Nitride Trench Waveguides. *J. Opt. Soc. Am. B* **2016**, *33*, 1182–1189.
- (57) Choi, J. H.; Adams, S. M.; Ragan, R. Design of a Versatile Chemical Assembly Method for Patterning Colloidal Nanoparticles. *Nanotechnology* **2009**, *20*, 065301.
- (58) Adams, S. M.; Campione, S.; Caldwell, J. D.; Bezares, F. J.; Culbertson, J. C.; Capolino, F.; Ragan, R. Non-Lithographic SERS Substrates: Tailoring Surface Chemistry for Au Nanoparticle Cluster Assembly. *Small* **2012**, *8*, 2239–2249.
- (59) Adams, S. M.; Campione, S.; Capolino, F.; Ragan, R. Directing Cluster Formation of Au Nanoparticles from Colloidal Solution. *Langmuir* **2013**, *29*, 4242–4251.
- (60) Casadaban, M. J. Transposition and Fusion of the Lac Genes to Selected Promoters in *Escherichia coli* Using Bacteriophage Lambda and Mu. *J. Mol. Biol.* **1976**, *104*, 541–555.