

Decoding Optical Responses of Contact-Printed Arrays of Thermotropic Liquid Crystals Using Machine Learning: Detection and Reporting of Aqueous Amphiphiles with Enhanced Sensitivity and Selectivity

Fengrui Wang,^{1,¶} Shiyi Qin,^{2,¶} Claribel Acevedo-Vélez,³ Reid C. Van Lehn,² Victor M. Zavala,^{2,4,*} and David M. Lynn^{1,2,*}

¹*Dept. of Chemistry, Univ. of Wisconsin–Madison, 1101 University Ave., Madison, WI 53706, United States;* ²*Dept. of Chemical and Biological Engineering, Univ. of Wisconsin–Madison, 1415 Engineering Dr., Madison, WI 53706, United States;* ³*Dept. of Chemical Engineering, University of Puerto Rico-Mayagüez, Call Box 9000, Mayagüez, PR 00681-9000, United States;* ⁴*Mathematics and Computer Science Division, Argonne National Laboratory, 9700 S. Cass Ave, Lemont, IL, 60439, United States;* [¶]*These authors contributed equally to this work. Email: (D.M.L.) dlynn@engr.wisc.edu, (V.M.Z.) victor.zavala@wisc.edu*

ABSTRACT: Surfactants and other amphiphilic molecules are used extensively in household products, industrial processes, and biological applications, and are also common environmental contaminants; as such, methods that can detect, sense, or quantify them are of great practical relevance. Aqueous emulsions of thermotropic liquid crystals (LCs) can exhibit distinctive optical responses in the presence of surfactants and have thus emerged as sensitive, rapid, and inexpensive sensors or reporters of environmental amphiphiles. However, many existing LC-in-water emulsions require the use of complicated or expensive instrumentation for quantitative characterization, owing to variations in optical responses among individual LC droplets. In many cases, the responses of LC droplets are also analyzed by human inspection, which can miss subtle color or topological changes encoded in LC birefringence patterns. Here, we report an LC-based surfactant sensing platform that takes a step toward addressing several of these issues and can reliably predict concentrations and types of model surfactants in aqueous solutions. Our approach uses surface-immobilized, microcontact printed arrays of micrometer-scale droplets of thermotropic LCs and hierarchical convolutional neural networks (CNNs) to automatically extract and decode rich information about topological defects and color patterns available in optical micrographs of LC droplets to classify and quantify adsorbed surfactants. In addition, we report computational capabilities to determine relevant optical features extracted by the CNN from LC micrographs, which can provide insights on surfactant adsorption phenomena at LC-water interfaces. Overall, the combination of microcontact-printed LC arrays and machine learning provides a convenient and robust platform that could prove useful for developing high-throughput sensors for on-site testing of environmentally or biologically relevant amphiphiles.

Keywords: liquid crystals, amphiphiles, surfactants, machine learning, neural networks

Introduction

Surfactants are amphiphilic molecules that contain hydrophilic head groups and hydrophobic tail groups, the combination of which allows them to self-assemble both in solution and at interfaces (*e.g.*, at solid surfaces or at interfaces between immiscible liquid phases).¹ Because of their unique properties, surfactants are used in a wide range of industrial and pharmaceutical products, including detergents, food additives, and as components of drug delivery vehicles.²⁻⁴ Amphiphilic molecules also play essential roles in living organisms, and many have been identified as disease biomarkers.^{5, 6} As a result of their widespread use in practical applications, surfactants can also end up as contaminants in consumer or pharmaceutical products or may be discharged into the environment where they can harm surrounding ecosystems.⁷⁻⁹

State-of-the-art methods for detecting and confirming the identities and structures of natural and synthetic amphiphiles include mass spectrometry and high-performance liquid chromatography (HPLC).^{5, 10} While these methods provide excellent sensitivity and selectivity, they often require complex sample preparation procedures, costly laboratory infrastructure, and highly trained personnel. Other lower-tech approaches have been developed, including colorimetric and two-phase titration assays that can monitor different levels of anionic¹¹ and cationic¹² detergents. However, this type of method often suffers from low reproducibility and tends to be time-consuming.¹³ Several other approaches have been developed to detect and quantify specific analytes of interest with great sensitivity and specificity;⁵ nevertheless, these methods can also be expensive, complex, and resource-intensive (for example, bacterial endotoxins can be detected at very low concentrations and with high specificity using assays

with biomolecular components derived from the farming and large-scale milking of blood from horseshoe crabs^{14, 15}).

Owing to these general limitations, there has been an increasing number of studies dedicated to developing new approaches to surfactant detection and quantification, including many optical, electrochemical, and mechanical-based techniques.^{5, 16} Among these new approaches, systems that exploit the unique properties of liquid crystals (LCs) have garnered a growing amount of attention.¹⁷⁻²² Liquid crystals (LCs) are liquid-phase materials that possess long-range directional order. Several decades of past research has shown that the adsorption of amphiphiles to LC-water interfaces can alter surface anchoring energy and translate molecular-level interfacial interactions into the bulk of the material, leading to optical changes, including birefringence patterns, that can be readily observed using cross-polarized light.¹⁸ In recent years, researchers have successfully designed various LC-based sensing platforms to detect amphiphilic lipids, proteins, endotoxins, and nucleic acids at biologically relevant concentrations.^{17, 19, 21, 22}

Overall, LC-based systems address several practical issues associated with existing methods for the detecting and reporting of environmental amphiphiles and, with further development, have the potential to lead to affordable, mobile, and potentially high-throughput sensing platforms. Many past studies have relied on visual inspection of polarized light micrographs for qualitative analysis (e.g., to report on the presence or absence of surfactants).^{17, 22} These methods are useful when changes in LC optical appearance are substantial; however, optical changes in color and defect patterns that are subtle or masked by more salient visual features can be easily overlooked by the use of the naked eye. The work reported here was motivated by recent advances in data-centric modeling techniques and the hypothesis that machine learning (ML) and related methods, including convolutional neural nets (CNNs), can be

used to reveal, extract, and exploit additional rich information contained in polarized light micrographs of LC emulsions in ways that can further enhance sensitivity or selectivity for the detection, reporting, and quantification of aqueous amphiphiles.

CNNs are a powerful ML tool for feature extraction from 1D (*e.g.*, feature array), 2D (*e.g.*, images and videos), and other higher-dimensional signals. These models have recently been implemented in LC-based sensing systems; for example, Cao *et al.*,²³ Xu *et al.*,²⁴ Bao *et al.*,²⁵ and Smith *et al.*²⁶ used CNN architectures to extract features from micrographs of LC thin films (*e.g.*, obtained from videos that recorded optical responses under polarized light) for detection of aqueous and gaseous analytes. In these cases, the CNN improved sensor selectivity and responsiveness. Jiang *et al.*¹⁴ investigated the use of free-floating LC-in-water emulsions for the detection of bacterial endotoxins; in this approach, the authors characterized differences in the scattering of light by LC droplets in emulsions using flow cytometry output as input signals to CNNs to predict endotoxin concentrations. In a more recent study, Frazao *et al.*²⁷ investigated LC emulsions embedded in hydrogels and used images of LC droplets as inputs to CNN-based frameworks to classify volatile organic compounds with an accuracy (F1-score) above 93%.

Overall, these past studies reveal that there exist characteristic features in the optical responses of various planar and colloidal LC-based materials that can be exploited using ML techniques to improve sensitivity and selectivity. However, these studies have either focused on the interactions between LCs and gaseous molecules²⁷ or require additional and expensive characterization equipment, such as flow cytometry,¹⁴ to characterize or detect stimuli adsorption at LC-water interfaces. Amphiphiles have not been fully explored as the stimuli of interest. Furthermore, since the optical properties of LC droplets depend upon their size, variations in the droplet size in LC-emulsion-based systems can lead to high variations in optical responses to

even a single analyte. Such high variations also make the chromatograms of the LC droplets challenging and less explored as the inputs for ML. Systems that use videos as the model input are further limited by the computational resources and video processing time.²⁷ Here, we address these limitations to demonstrate the potential of using LC micrographs as machine learning inputs to detect and quantify amphiphiles in solution.

In this study, we report a strategy for the detection, reporting, and quantification of model aqueous surfactants based on the use of surface-immobilized, microcontact printed LC droplet arrays and a hierarchical CNN for the analysis of LC optical transitions captured using polarized light microscopy. We demonstrate that microcontact printing yields uniform and immobile LC droplets that exhibit optical responses upon contact with aqueous amphiphiles that are diagnostic and, in several ways, similar to those of free-floating LC droplets investigated in past studies. The uniform size of the immobilized LC droplets in these printed arrays significantly decreases variations in optical responses upon the introduction of surfactants and thus decreases the complexity needed to decode optical responses using CNNs to classify both surfactant type and concentration. Our results demonstrate that this approach can extract rich information contained in polarized light micrographs of LC droplets to improve the sensitivity of the system. We also show that the system can be used to distinguish LC droplets treated with 0.05 nM concentrations of the model surfactant sodium dodecyl sulfate (SDS), with a prediction accuracy of greater than 90%, from those treated with water as a control, despite the optical images having no apparent meaningful or discernable differences when viewed and interpreted by trained human eyes. In addition, our results show that this overall strategy and a hierarchical CNN can be used to improve the selectivity of the system and permit accurate distinctions between multiple different synthetic and natural surfactants with differences in both head group and tail group structures

(again, based on images that have no meaningful or reliable differences when viewed or interpreted by human eyes). Generation of saliency maps reveals that the CNN can distinguish prominent defect patterns using grayscale microscopy images, and that color information is necessary for capturing subtle pattern changes in concentric rings at low surfactant concentrations. Overall, this approach leads to improved sensitivity and selectivity and provides a robust sensing workflow that does not require access to expensive instrumentation to provide accurate and useful information about amphiphilic species in aqueous solutions. This approach could therefore serve as a step toward the development of low-cost, convenient, and potentially high-throughput methods for in-line or on-site testing, identification, and reporting of surfactants and other environmentally or biologically relevant amphiphiles.

Experimental and Computational Methods

1. Experimental methods

Materials. The nematic thermotropic LC 4'-pentyl-cyanobiphenyl (5CB) was purchased from HCCH Jiangsu Hecheng Display Technology Co., Ltd. (Jiangsu, China). Sodium dodecyl sulfate (SDS) and dodecyltrimethylammonium bromide (DTAB) were obtained from Sigma-Aldrich (Milwaukee, WI). PDMS (Sylgard 184 Silicon Elastomer) was purchased from Ellsworth Adhesives (Germantown, WI). Glass coverslips were obtained from Fisher Scientific (Pittsburgh, PA). Rhamnolipids, 90% pure (commercially obtained rhamnolipids used in this study were isolated from *P. aeruginosa* as a mixture of congeners with different numbers of rhamnose sugar moieties, mono- and di-, linked to one or two molecules of hydroxy acid that differ in their aliphatic chain length) were purchased from AGAE Technologies (Corvallis, OR). Deionization of distilled water was performed using a Milli-Q system (Millipore, Bedford, MA) to yield water with a resistivity of 18.2 M Ω . All materials were used as purchased without further purification unless noted otherwise.

General considerations. All glass slides used for microcontact printing were soaked in ethanol overnight and then rinsed under a stream of ethanol three times prior to use. Bright-field and polarized-light microscopy images were acquired using an Olympus IX71 inverted microscope (Waltham, MA) equipped with cross-polarizers (Olympus analyzer slider IX2-AN and condenser attachment IX-LWPO). Fields of view were recorded using an OPTO-EDU (Beijing, China) eyepiece camera model A59.2211 connected to a computer and controlled through ImageView imaging software version A30.2201.

Microcontact Printing of LC Droplet Arrays. PDMS stamps consisting of an array of pillars (100 μm x 100 μm square, spaced 100 μm apart) were prepared as previously described.²⁸ Briefly, photolithography techniques were used to produce a silicon master with depressions, or negative features, corresponding to final desired pillar structures and patterns. A mixture (10:1 w/w) of PDMS pre-polymer and curing agent was then degassed and poured onto the silicon master. The PDMS layer was then cured at 70 °C in an oven for 1 hr to obtain the desired square-array stamp. Inking of the stamps was achieved by first spreading a 5 μL droplet of 5CB on a clean glass substrate using another glass plate to produce a thin layer of LC. The stamp was gently placed onto the LC-spread surface and then quickly transferred and placed gently on a glass cover slide. Stamps were left in contact with the glass cover slides for 10 s and then removed. The printing process was repeated one more time without reinking and this second print was used for collecting data. Different aqueous surfactant solutions (20 μL) were then carefully introduced to the printed LCs using a micropipette and LC droplet arrays were imaged under cross polarized light. We estimate the height at the centers of droplets after introduction of aqueous surfactant to be ~ 20 μm .

2. Computational Methods

Droplet Detection and Size Calculation. We adapted the canny edge detection method²⁹ to extract only the center region that contains the droplet (removing the dark peripheral area that does not contain any information) with high resolution from raw micrographs of LC droplet arrays (4912 $\times$ 3684); this algorithm smooths a grayscale image using a Gaussian filter, generates the corresponding edge mask (binary) based on gradient computation through a Sobel operator,

and links the broken edges via non-maximum suppression coupled with thresholding. To account for the halo effect, we performed edge dilation that keeps the peripheral area of the detected droplet. After this, we applied contour detection³⁰ to obtain a square bounding box around the droplet, followed by image cropping. The size of each cropped image was used to approximate the diameter of the droplet by a unit conversion from the pixels to microns. Since the droplet detection algorithm includes the background halo, the calculated diameters are expected to be larger than the actual droplet sizes, yet the relative sizes can be studied to analyze the uniformity of the LC droplets prepared by microcontact printing. Additionally, this workflow has the capability of detecting multiple droplets in complex systems with high computational efficiency compared with more common machine learning approaches that are intended for the same purpose, such as YOLOv3³¹ used by Frazao *et al.*²⁷ for optical gas sensing with LC, which requires model training to fit millions of parameters (*e.g.*, Darknet-53 has 53 convolutional layers). Since the droplets have sufficient distinguishing factors from the background, a simple gradient-based method is efficient to accomplish the task.

Image Preparation and Augmentation. In the study reported here 1,371 raw LC micrographs were obtained directly from approximately 130 different experimental trials. For each raw micrograph, we first applied the droplet detection algorithm to extract the droplet region. After droplet detection, the micrographs were resized to 256×256 and augmented 20-fold with combinations of random rotation, shifting, shearing, zooming, and flipping. The augmentation parameters were controlled in a way that simulates experimental variations. For instance, the zooming range was set to 95% to 105%, which is consistent with the standard deviation (~10%) of the droplet size reflected by the dimension of the cropped micrographs. Overall, the resulting

image dataset contained 27,420 pre-processed LC micrographs that were categorized into 13 classes (based on the type and concentration of the corresponding surfactant exposure) for training and validation.

Topological Analysis using Radial Distribution Profiles. We computed a radial distribution (RD) profile of the color intensity I as a function of radius (pixel) r for each micrograph; the RD profiles served as simple topological descriptors to represent the LC droplet data and are defined as shown in Equation 1:

$$I(r) = \frac{1}{\sum_i \sum_j \mathbb{1}_{\{d_{ij}=r\}}} \sum_i \sum_j \mathbb{1}_{\{d_{ij}=r\}} I_{ij} \quad (\text{Equation 1})$$

where i and j are the row and column of a pixel, I_{ij} represents the corresponding color intensity, and $d_{ij} = \sqrt{(i - c_x)^2 + (j - c_y)^2}$ is the Euclidean distance between a pixel and the center point (c_x, c_y) of the droplet. The location of the center point was determined by the weighted average of the coordinates along the row and column of a micrograph; the weights were calculated from the binary mask M_{ij} obtained after image thresholding using Equations 2-4

$$c_x = \sum_i \sum_j i M_{ij} \quad (\text{Equation 2})$$

$$c_y = \sum_i \sum_j j M_{ij} \quad (\text{Equation 3})$$

$$M_{ij} = \begin{cases} 1, & I_{ij} \geq 70^{th} \text{ percentile} \\ 0, & \text{otherwise} \end{cases} \quad (\text{Equation 4})$$

Color Space Transformation of LC Micrographs. Digital images are viewed as grid data objects with single or multiple color channels, each being a 2D matrix with numerical values

describing the color intensity at individual pixels. The LC micrographs are color images with red, green, and blue (RGB) channels. In this study, we performed color space transformations to identify important color channels that contribute most to surfactant quantification. Apart from the RGB color space, we generated grayscale (a linear combination of RGB channels) images and explored the $L^*a^*b^*$ color space (nonlinear conversion from RGB) introduced by the International Commission on Illumination (CIE). The $L^*a^*b^*$ channels represent the lightness, red/green value, and blue/yellow value, respectively; since red/green and blue/yellow are opponent color pairs, the $L^*a^*b^*$ color space has a strong correlation with human visual perception.³² According to Levit *et al.*,³³ the thermochromic behavior of cholesteryl ester liquid crystals can be associated with the cumulative change in the $L^*a^*b^*$ color space. Therefore, we investigated various combinations of the individual color channels from the RGB/ $L^*a^*b^*$ /grayscale color spaces as alternative ways to represent the input data.

CNN Architectures for Different Data Representations. Based on the type of input data representations, we developed three CNN architectures to extract and summarize features from data for classification of surfactant concentrations.

1D CNN. A convolution operation computes the weighted sum across local domains of the input data for feature extraction. For RD profiles that are mathematically represented as 1D vectors with 3 channels (RGB), we applied 1D CNNs. In a 1D CNN, we define an input RD vector v in a discrete domain of radius r and a kernel (or filter) vector k . The 1D convolution operator for each channel c is formulated as shown in Equation 5

$$h^c[r] = v^c[r] * k[r] = \sum_{r'=-\infty}^{\infty} v^c[r'] \cdot k[r - r'] \quad (\text{Equation 5})$$

and the channel-wise hidden features are summarized by Equation 6

$$h[r] = \sum_{c \in \{red, green, blue\}} h^c[r] \quad (\text{Equation 6})$$

2D CNN. Similar to the 1D CNN, 2D CNNs incorporate the same architecture on images, which are 2D objects (matrices with multiple channels). Using the same set of annotations, the 2D convolution operator for each channel c is formulated as shown in Equation 7

$$h^c[i, j] = v^c[i, j] * k[i, j] = \sum_{j'=-\infty}^{\infty} \sum_{i'=-\infty}^{\infty} v^c[i', j'] \cdot k[i - i', j - j'] \quad (\text{Equation 7})$$

where i and j denote the row and column of a pixel, and the channel-wise hidden features are again calculated by the sum of the hidden features of each color channel.

For both 1D and 2D CNNs, we applied zero padding, a technique that adds zero entries to the start and the end of each input, to ensure that the input and output signals have the same dimension. The convolution operation is followed by nonlinear activation σ ; here, we chose Rectified Linear Unit (ReLU), which is a commonly used activation function that avoids vanishing gradients while retaining sensitivity.³⁴ The output from nonlinear activation $\sigma(h)$ then undergoes batch normalization as a regularization method to avoid overfitting as well as to increase convergence speed.³⁵ Next, the normalized data are sent to a pooling layer for size reduction through techniques such as averaging the values within a local domain. Besides size reduction, the pooling layer also allows for invariance to small input translation,³⁶ which is a desired feature in data processing. The abovementioned convolution, batch normalization, and

pooling layers are bundled together and repeated to extract hidden features from the input data at a deeper level.

Hierarchical CNN. The hierarchical CNN proposed in this study operates on images and contains two levels of 2D CNNs that are trained on different data representations. The selection of the data representation for each level relies on the detection focus. The first level is intended to differentiate defect patterns, including bipolar, pre-radial/escaped radial, and radial configurations, and therefore grayscale images, which highlight the contrast between light and shade and contain sufficient information for defect pattern classification. The second level, on the other hand, is forced to learn patterns that are less distinguishable and thus require color information (e.g., RGB). We note here that although this hierarchical CNN architecture selects grayscale images as the input for the first level and RGB images as the input for the second level, alternative data representations were explored, such as the $L^*a^*b^*$ color space and individual color channels. However, these data representations did not improve model accuracy in our early studies and therefore were not selected for further study.

Model Implementation. Images were pre-processed using OpenCV (version 4.4.0) for tasks including resizing, droplet detection, and color space transformation mentioned previously. The CNN models were implemented using PyTorch (version 1.2.0).³⁷ The major hyperparameters we varied include the kernel size (5,7,9), number of convolution layers (1,2,3), pooling function (max-pooling or average-pooling), and learning rate (0.0001,0.0005,0.001). The tested models were trained with the cross-entropy loss, the Adam optimizer, and a learning rate of 0.0001. We performed 5-fold cross-validation (CV) with stratified random sampling for hyperparameter

tuning and model evaluation. Stratification was based on the class labels. All evaluation metrics were computed using the compilation of the validation data in each fold to obtain a realistic estimation of the model performance. More implementation details can be found in Table 1.

Table 1. Hyperparameters of CNN architectures.

CNN Architecture	1D CNN	2D CNN	2-Level 2D CNN	
			1 st Level	2 nd Level
Input Data	Radial Distribution (RGB)	Micrograph (RGB)	Micrograph (Gray)	Micrograph (RGB)
Input Data Size	$\mathbb{R}^{180 \times 3}$	$\mathbb{R}^{256 \times 256 \times 3}$	$\mathbb{R}^{256 \times 256 \times 1}$	$\mathbb{R}^{256 \times 256 \times 3}$
# Convolution+Pooling Layers	3	3	2	3
Convolution Kernel Size	7	7×7	7×7	7×7
Pooling Type		Average Pooling		
Batch Normalization		Enabled		
Zero Padding		Enabled		
# Fully-connected Layers		2		
Activation		ReLU		

Saliency Calculations. Saliency maps were generated using integrated gradients (IG)³⁸ as shown in Equation 8

$$IG_i^{approx}(x) = (x_i - x'_i) \times \sum_{k=1}^m \frac{\partial F_{CNN}\left(x' + \frac{k}{m}(x - x')\right)}{\partial x_i} \times \frac{1}{m} \quad (\text{Equation 8})$$

where F_{CNN} denotes the trained CNN model, i denotes individual features (in this case color intensity of a pixel), x denotes input images, x' denotes a baseline image (zero tensor with the same dimension as the input), and m (set to 50) denotes the number of steps in the approximation. IG is one of the state-of-the-art methods for model interpretation, which has proven to be sensitive to important features and insensitive to unimportant features and thus reduce visual diffusion.³⁹ In this study, IG was implemented using Captum⁴⁰ (version 0.4.1).

Results and Discussion

Fabrication of Surface-Immobilized Arrays of LC Droplets by Microcontact Printing

For sensing applications, microscale LC droplets with uniform sizes are desirable because variations in droplet size (*e.g.*, as present in aqueous emulsions prepared using many conventional methods) can lead to variations in the types of optical responses that are observed upon the adsorption of amphiphiles.^{41 42 43} In past studies, researchers have used various methods to create monodispersed LC droplets, including microfluidic devices,⁴⁴ the filling and subsequent removal of monodisperse hollow capsule templates,⁴² capture of LCs on micropatterned self-assembled monolayers (SAMs),⁴⁵ and inkjet printing⁴⁶ to create either free-floating LC droplets or arrays of surface-immobilized droplet arrays. Here, we used a simple PDMS-based microcontact printing method to create immobile and relatively uniform LC droplet arrays. Compared to other methods used in past studies noted above, this microcontact-printing method requires minimal material or preparation, is low cost, and produced arrays of uniformly sized droplets sufficient for all subsequent experiments described below.

We first performed a series of experiments to determine whether the surface-immobilized droplets in these arrays, which are non-spherical and in substantial contact with an underlying glass surface, respond to stimuli in ways that are similar to those often associated with free-floating spherical LC droplets and immobilized LC droplets created using other methods. For these and all other studies described below, we selected the well-studied LC 5CB as a model LC and characterized the behavior of arrays in the absence or presence of an overlayer of water or other aqueous surfactant solutions using bright field and cross-polarized light microscopy.

As shown in Figure 1A, contact-printed LC droplets under air had relatively circular shapes when viewed from above in the bright-field. When observed using cross-polarized light (Figure 1B), the birefringent patterns of the LC droplets resembled those generally characteristic of LC

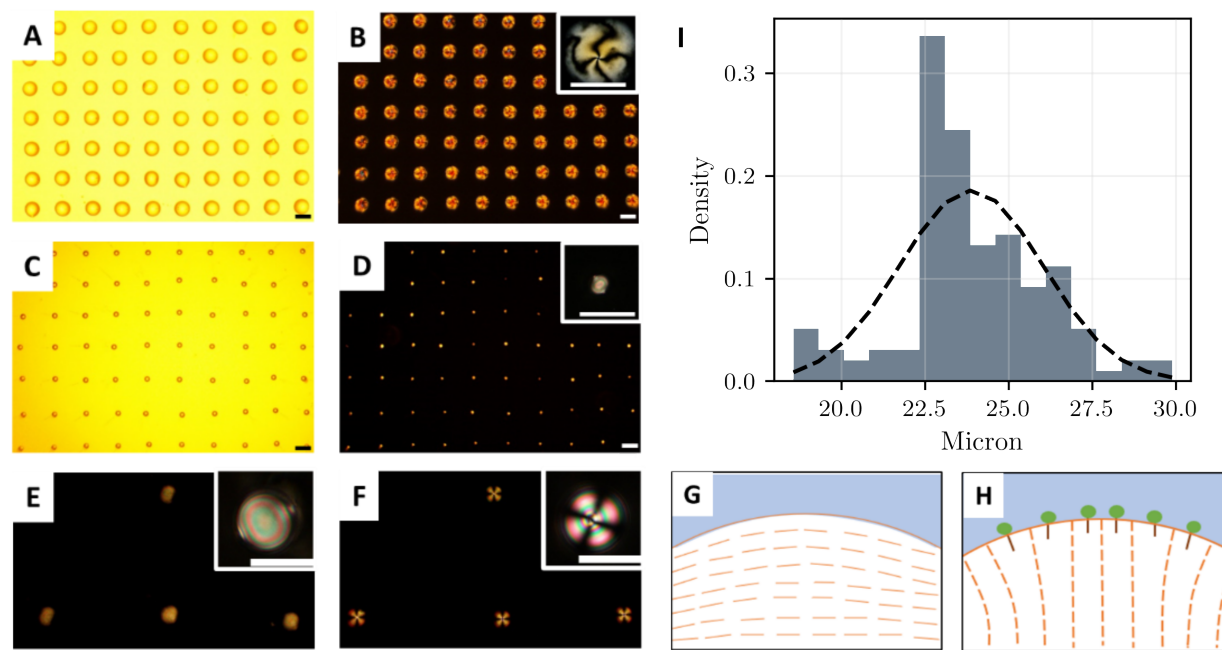


Figure 1. LC droplets prepared by microcontact printing. (A) Printed LC droplets under a bright-field microscope. (B) Printed LC droplets under a cross-polarized microscope. (C) LC droplets with water treatment under a bright-field microscope. (D) LC droplets with water treatment under a cross-polarized microscope. (C) and (D) display a smaller droplet size due to the change in wetting behavior of printed LCs when water is introduced. (E) Magnified view of (D) where a bipolar configuration is observed. (F) Magnified view of the microcontact-printed LC droplets when treated with 1 mM model surfactant SDS. A radial configuration is observed. (G)-(H) Schematic illustrations of the surface anchoring of printed LC molecules with bipolar (planar anchoring) and radial (homeotropic anchoring) configurations, respectively. (I) Size distribution of microcontact-printed LC droplets based on cross-polarized LC micrographs of water-treated LCs. The computationally obtained droplet diameters take into account the background halo and have a mean of 23.1 μm with a standard deviation of 2.1 μm , resulting in a polydispersity index of 0.09.

thin films.⁴⁷ Upon the introduction of an overlayer of water, the printed LCs beaded up to form partially wetted droplets and had substantially smaller footprints (Figure 1C) as compared to the droplets in arrays under air (Figure 1A). When observed under cross-polarized light (Figure 1D), microcontact-printed LC droplets covered by an overlayer of water exhibited optical appearances similar in many ways to those of LC droplets in emulsions created using other methods. As shown in the higher magnification view shown in the inset of Figure 1E, two defect regions can

be observed at opposite poles of the individual droplets. This optical texture is similar to the so-called ‘bipolar’ configuration²¹ observed in free-floating spherical LC droplets, in which the LC adopts a planar anchoring at the oil-water interface (as depicted in Figure 1G for the present case of surface-immobilized droplets; for simplicity, we refer to immobilized, non-spherical droplets having this appearance as being in a ‘bipolar’ configuration from here on).

We then characterized the response of these microcontact-printed LC droplets after the introduction of an overlayer of water containing the model surfactant SDS (1 mM). Figure 1F shows cross-polarized images of printed LC droplets after the introduction of SDS; all droplets show a characteristic cross-like pattern with a single defect near the geometric center of the droplets in this top-down view. This optical appearance is similar in many ways to the so-called ‘radial’ configuration²¹ observed in spherical LC droplets, in which the LC adopts homeotropic anchoring at the oil-water interface (as depicted in Figure 1H for the present case of surface-immobilized droplets; for simplicity, we refer to immobilized, non-spherical droplets having this general optical appearance as being in a ‘radial configuration in the discussion below). The results of these experiments demonstrate that the microcontact-printed LC droplet arrays reported here respond to the presence or absence of water or aqueous surfactant in ways that are similar to those of spherical droplets reported in past studies and sufficiently diagnostic to provide optical patterns that can be used to identify the presence or absence of amphiphilic species in the surrounding aqueous environment. The LC droplet arrays used here were stable when stored under air, and were generally stable (*e.g.*, droplets did not delaminate or decrease substantially in size) when stored for periods of at least several days under an overlayer of water. These features confer potential practical advantages relative to conventional LC-in-water

emulsions used in past studies, which can coalesce, wet surrounding surfaces, or undergo changes in size distribution over a time scale of several hours.⁴⁸

We then conducted experiments to characterize the uniformity of the microcontact-printed LC droplets and the consistency of their response to amphiphilic species. The droplet size distribution of a representative array is shown in Figure 1I (the method used to determine the diameters of these droplets is summarized in the Methods section above). The average diameter of the printed droplets was $23.9 \pm 2.1 \mu\text{m}$ and the polydispersity index based on the distribution was 0.09, which is comparable to droplets created by microfluidic devices⁴⁹ and indicates that the printed LC droplets are relatively monodisperse. To examine the consistency of optical responses to chemical stimuli, we collected multiple micrographs of LC arrays treated with a variety of different model surfactants at various specified concentrations. As shown in each column of Figure 2A, the droplets exhibited comparable and generally similar visual optical appearances in response to each surfactant across multiple trials; additional analysis and discussion of similarities and difference in these images is provided below.

We note that different surfactants can cause different changes in the surface tension and that differences in the charge of surfactant head groups can lead to different interactions at LC/solid interfaces. Because the printed LC droplets are partially wetting the underlying glass surface, both factors could cause LC droplets to have different wetting behaviors and lead to changes in birefringence patterns that differ from those of free-floating LCs.^{20, 50} For example, when high concentrations of positively charged DTAB adsorb to the LC, it can interact with the negatively charged glass surface, resulting in optical appearances that look like a four-lobed structure and resemble the appearance of flattened droplets with homeotropic anchoring.^{46, 51} Overall, despite some differences in droplet wetting behavior in the presence of different

surfactants, the microcontact printed LC droplets do show characteristic optical appearances that are similar in many ways to those of free-floating LC droplets and that, as described in the sections below, can be used as diagnostic inputs for processing using ML methods. The uniformity and immobility of the microcontact-printed droplets also reduce variations in the LC responses and facilitate the rapid imaging of large numbers of droplets, respectively, further facilitating the collection of high-throughput data for ML processing.

Decoding LC Micrographs using ML to Detect and Quantify SDS

Quantifying SDS Concentrations Using Radial Distribution Profiles: We performed a series of computational experiments in which we used ML methods to extract quantitative and qualitative information from micrographs of LC droplet arrays and determine whether this information could be used to reliably detect and quantify aqueous surfactant concentrations. We first selected SDS as a model surfactant and attempted to quantify its concentration using direct feature extraction from the LC micrographs to establish a baseline measurement of the predictive power of our proposed approach. We first generated radial distribution (RD) profiles for each micrograph as described in the Methods, where the color intensity of image pixels was plotted as a function of the droplet radius. RD profiles provide a useful summary that helps visualize differences in the topological features of LC droplets. Figure 2B shows the RD profiles of two representative LC micrographs, one with a bipolar configuration (treated with water) and the other with a radial configuration (treated with 1 mM SDS). For the LC droplet with the bipolar configuration, the RD profile shows an overall decreasing trend in the color intensity from the center to the edge of the droplet; for the LC droplet with the radial configuration, the RD profile reveals low intensity near the center of the droplet. Overall, these two different RD profiles

reflect large and characteristic differences in the defect patterns of these two LC droplet optical configurations that can also be readily discerned by human visual inspection.

We used statistical analysis to determine whether images of LC droplets were representative across different arrays and trials using water and analyte solutions at seven

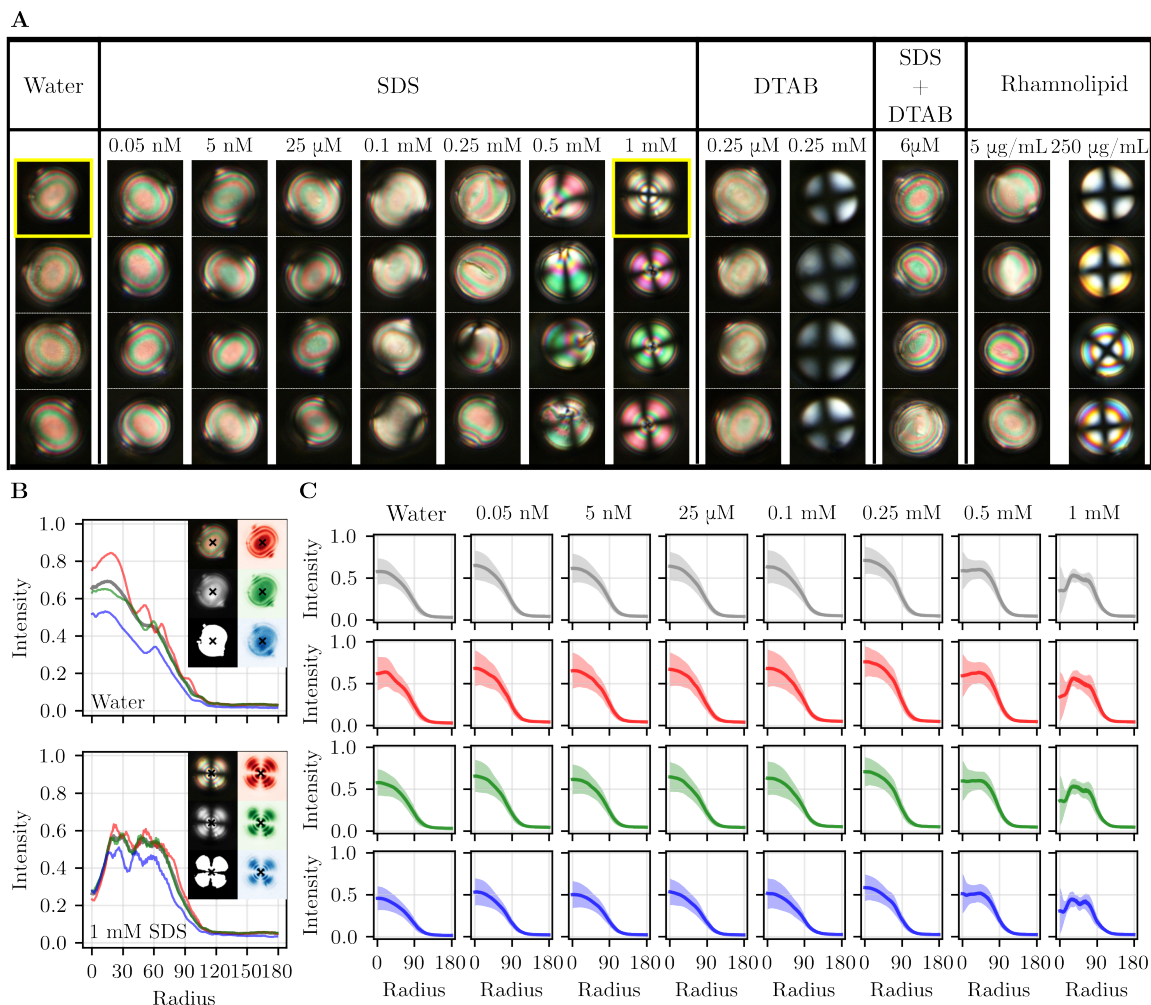


Figure 2. Examples of LC micrographs and radial distribution (RD) profiles. (A) Examples of LC micrographs after droplet extraction and resizing. The micrographs are labeled by surfactant type and concentration, and they are categorized into 13 classes. (B) The RD profiles of the LC micrograph examples boxed in yellow in (A). Each LC micrograph is first converted to a grayscale image, which is used to generate a binary mask for droplet center detection. With the center point coordinate, the LC micrograph is then broken down into its composing RGB channels, and the corresponding radial distribution is computed using the color intensities of individual pixels. The radius is units of pixels. (C) Statistical summary of the RD profiles for water (0 mM SDS) and seven SDS concentration classes. For each SDS quantification class, the mean and standard deviation are calculated using all the LC micrographs from different experimental trials for that class.

different SDS concentrations. The SDS concentrations used were selected to encompass a range of concentrations that lead to a wide range of different optical appearances of LC droplets,

including bipolar and radial configurations and several different intermediate optical textures. Figure 2C shows the mean and standard deviation of the RD profiles summarized over individual SDS classes. The summary plots smooth out the local peaks that appear in each individual plot due to variation in the relative positions of the concentric rings observed in different LC micrographs, but the general shape of the RD profile is retained. At low concentrations of SDS (<0.5 mM), all droplets exhibit bipolar-like configurations and hence the color intensities in the RD profiles drop as the radius increases. At 0.5 mM and 1 mM SDS, the RD profiles reveal unique patterns that were different from those at other concentrations. At 0.5 mM, the immobilized LC droplets exhibited optical appearances similar in many ways to so-called ‘pre-radial’ or ‘escaped radial’ configurations⁴² reported previously for spherical droplets, for which the defect regions start to appear closer to the center of the droplets. Therefore, in the RD profiles of these droplets, the intensities at the center of the droplet are smaller and show a plateau region. At 1 mM SDS, the LC droplets exhibited a radial configuration and RD profiles having a bimodal distribution with declining peak heights (indicative of maximum color intensities that move away from the centers of the droplets). In general, the appearance and shift of the peaks in these RD profiles consistently and successfully captured the most salient visual differences in the defect patterns of LC droplets treated with different concentrations of SDS in this range.

Because changes in LC defect regions reflect different levels of ordering transitions triggered by surfactant adsorption, we hypothesized that the RD profiles generated above could be used to quantify surfactant concentrations. To test this hypothesis, as shown in Figure 3A, we used RD profiles as the input for a simple 1D CNN architecture to predict eight different concentrations of SDS (including pure water, which corresponds to zero SDS concentration). We

chose this 1D CNN architecture due to its feature extraction power by inspecting the correlations between signals at local domains. For 8-class multiclassification of SDS concentration, the 1D CNN architecture achieved an overall accuracy of 59% (after 5-fold cross-validation) with an

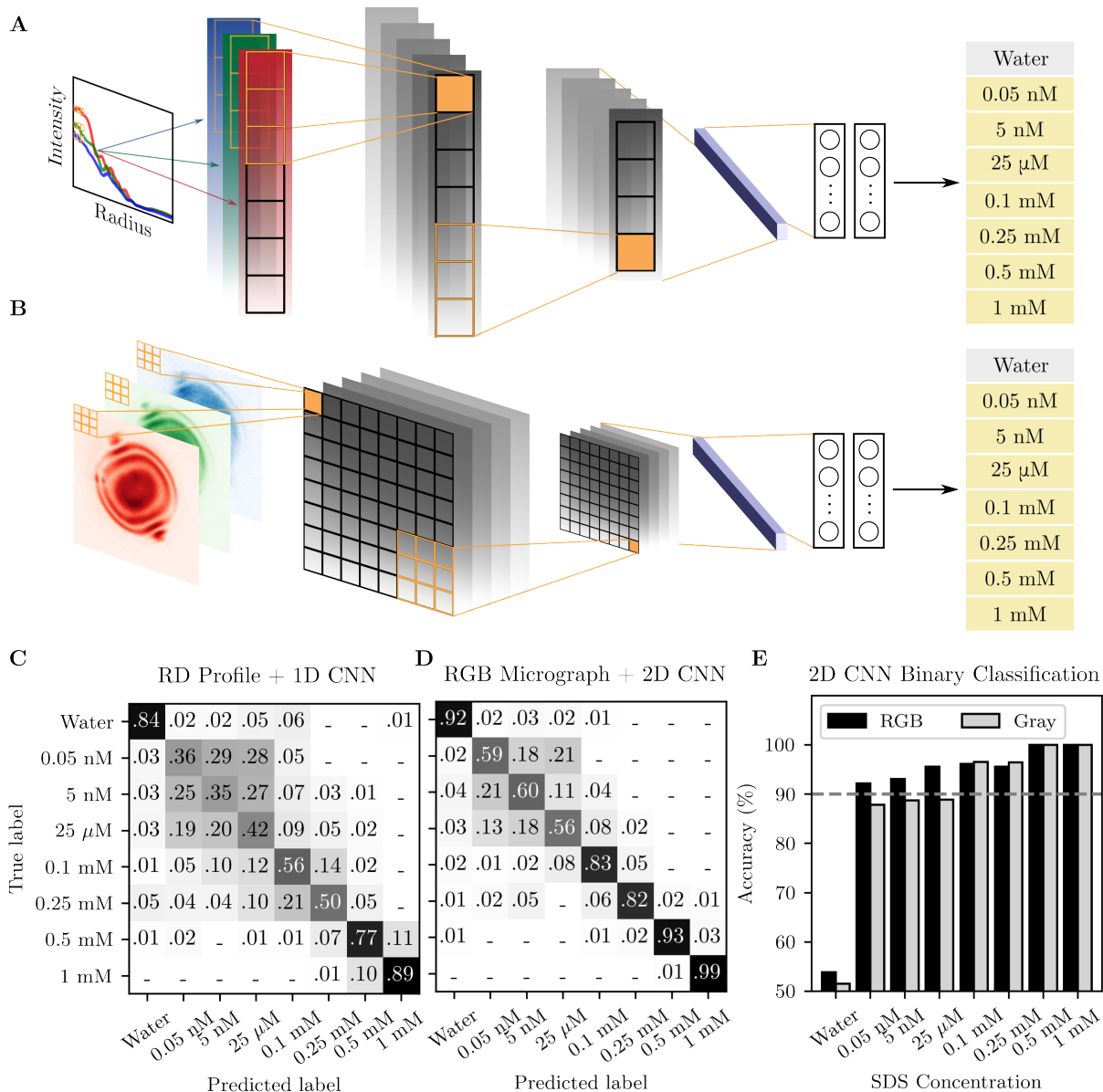


Figure 3. SDS classification with 1D and 2D CNN architectures. (A) Schematic of 1D CNN using RD profiles as the inputs. (B) Schematic of 2D CNN using LC micrographs as the inputs. Both input data representations in (A) and (B) keep the three RGB color channels, and each CNN is trained for 8-class classification using cross entropy loss. The categorical-wise accuracies are summarized in (C) and (D), respectively. Numerical values are colored in grayscale with different shades of gray corresponding to the numerically indicated prediction accuracy with a scale varying from white (for values equal to zero) to black (for values equal to 1). (E) Bar plot of the binary classification using RGB and grayscale LC micrographs as the inputs. The details of the architecture are summarized in Table 1. The dashed line shown at 90% accuracy is included to guide the eye.

average F1-score of 0.59, which is much higher than a random guess (12.5% for 8-class classification). The categorical-wise accuracy is summarized in the confusion matrix shown in Figure 3C. In general, highlighted concentrations, at which droplets adopt configurations with topological defects that are readily distinguishable by visual inspection, such as pre-radial, escaped radial, and radial, have a relatively high prediction accuracy. These results are consistent with the visual observations of the LC micrographs. We note that the “water” category (no SDS) also exhibited a relatively high accuracy, which was not initially anticipated, given that bipolar configurations at very low SDS concentrations were almost identical to the trained naked eye. This observation is also supported by Figure 2C; the differentiating RD features, although subtle, show a unique pattern in the red color channel. Because even trace amounts of surfactant adsorption can cause changes in color/light intensity, it is possible for the CNN to distinguish droplets with a perfect bipolar configuration (where the defects are always at the edge of the droplet) from those with bipolar-like configurations (where the defects are in slightly different positions).

To summarize, RD profiles encode sufficient information for LC pattern classification between bipolar, pre-radial/escaped radial, and radial configurations using a 1D CNN, but this approach fails to make accurate predictions for low SDS concentrations. Because RD profiles are only a high-level summary of the topological features of the LC droplets, it is understandable that they could miss subtle pattern differences among bipolar-like configurations at low SDS concentrations that are not sufficiently high to trigger substantial topological defect changes. Nevertheless, RD profiles have direct physical relevance to the appearance of the LC after surfactant adsorption and can be interpreted easily. Moreover, as a simple data representation

method that retains only the high-level topological features of a micrograph, RD profiles have better overall prediction accuracy than a random guess.

Quantifying SDS Concentrations Using LC Micrographs: With the baseline established by the RD profiles in hand, we further investigated whether it was possible to improve classification accuracy using LC micrographs directly as inputs to a 2D CNN (Figure 3B). Because the micrographs contain information about the LC droplets that is much richer than that captured in the corresponding RD profiles, we reasoned that they would serve as a more effective data representation for surfactant quantification with improved sensitivity. The resulting confusion matrix (Figure 3D) reveals an increased accuracy across all classes of SDS concentrations compared to the baseline, and the overall cross-validation accuracy increased from 59% to 78%, with the F1-score being 0.78. The prediction accuracy for water, 0.5 mM, and 1 mM SDS exceeded 90%, which again confirms the capability of 2D CNNs to discriminate between LC droplets with stable topological defects. Another significant improvement lies in the successful classification ($> 80\%$) of the LC droplets that are treated with lower concentrations of 0.1 and 0.25 mM SDS. For droplets treated with 0.25 mM SDS, their configurations fall, by visual inspection, into the transition region intermediate to bipolar and pre-radial/escaped radial configurations. In this transition region, the defect structures begin to move around the surface of the droplets, leading to complex optical responses that do not have any apparent distinctive pattern, which adds difficulty to pattern recognition. However, these complex optical responses that sometimes resemble both bipolar and pre-radial configurations are still being classified successfully by the strong feature extraction power of 2D CNNs. For LC droplets treated with even lower 0.1 mM SDS, all the droplets are in bipolar-like configurations, and yet the CNN is still able to distinguish these droplets from those treated with water or other lower concentrations

of SDS. It is also worth noting that, for LC droplets treated with micro- or nanomolar concentrations of SDS, the class-wise accuracy rises by approximately 20% as compared to the analysis of the RD profiles, demonstrating that the LC micrographs themselves contain more information that is useful for surfactant quantification.

In summary, our results of these initial studies demonstrate the potential utility of combining micrographs of microcontact-printed LCs with ML methods to quantify aqueous surfactant concentrations in surrounding media. With the rich information in the LC micrographs decoded by a 2D CNN, the classification accuracy was greatly improved. The tuned 2D CNN architecture can successfully quantify concentrations of SDS ranging from 0.1 mM to 1 mM, and even at concentrations of SDS that are not sufficiently high to lead to any observable pattern changes that can be reliably interpreted as distinct by the trained human eye. Such results show the power of 2D CNN to extract hidden features and to use such information to differentiate between similar bipolar droplets to some extent.

Detecting Trace Amounts of SDS through Binary Classification: The results above demonstrate that 2D CNNs can distinguish micrographs of the water-treated LC droplets (no SDS) from the remainder of the surfactant concentration categories. Motivated by this observation, we further explored whether hidden information in the optical responses could be used to detect trace amounts of SDS with high accuracy (thus testing the detection limit of the sensing system using this approach). To investigate this, we conducted a binary classification between the control group (“water”) and any SDS-treated group. The bar plot in Figure 3E shows the results for the binary classification. The cross-validation accuracy exceeded 90% in each task, even for SDS concentrations in the nanomolar regime. As SDS concentration decreased, overall, the accuracy

decreased. We tested an extreme condition where the labels were randomly assigned to the control group (“random water”) to mimic negligible SDS treatment, and the resulting prediction accuracy was approximately 50%. This result confirms that there indeed exists a detection limit and illustrates that the 2D CNN is not relying on random features but, rather, capturing hidden features that are invisible to or difficult to interpret by the trained naked eye.

Because some concentrations of SDS tested here were extremely low, we speculate that, in addition to defect patterns, the algorithm is observing other optical features, such as the unique color patterns that originate from the long-range directional ordering of the LCs. To provide further insight, we performed binary classification using only grayscale images (weighted averages of the RGB color channels). Since these grayscale images are only summaries of the detailed information contained in the individual color channels, we expected to see a slight decrease in prediction accuracy. This was confirmed, as illustrated in Figure 3E, where the prediction accuracies for low SDS concentrations using grayscale images dropped by 5% to 10%, as compared to RGB images. This result indicates that the inclusion of individual color channels is necessary for detecting low SDS concentrations.

Overall, with the addition of the 2D CNN, it is possible to detect SDS concentrations in the nanomolar regime with high accuracy. However, the binary classification model has a limited number of use cases, as it can only identify the existence of SDS rather than distinguishing between various SDS concentrations (which was achieved earlier). Here, we investigated binary classification to push the predictive power to the extreme for potential applications of low SDS detection. This part of the study was also used to show the existence of a detection limit and confirm that the information encoded in individual color channels is useful for surfactant quantification. It is worth noting that we picked SDS as a proof-of-concept model surfactant to

determine whether CNNs can help improve the sensitivity of this LC-based system. In the sections below, we show that this general approach can also be applied to detect and quantify trace amounts of other aqueous amphiphiles.

Maximizing Selectivity Through Hierarchical CNNs

Differentiating various surfactants using LC micrographs: The results above demonstrate that CNNs can extract and decode hidden information from LC micrographs to quantify concentrations of aqueous SDS with improved sensitivity. We also investigated whether CNNs can uncover information that can be used to increase selectivity to this LC sensing system. Past studies report that as surfactants adsorb to LC-water interfaces, their hydrophobic regions penetrate into the LC surface and that variations in tail branching or tail number can lead to differences in optical transitions.⁵² Such structural differences can lead to different concentrations at which bipolar-to-radial transitions are triggered. However, to the human eye, micrographs of bipolar or radial LC droplets treated with single-tail surfactants generally look identical (or very similar) to those treated with two-tailed surfactants. Past studies have determined the hydrophilic head groups of surfactants to play less of a direct role in influencing LC alignment.⁵³ However, the structure of surfactant head groups can still affect the assembly of surfactants at LC-water interfaces through charge interactions and hydrogen bonding, which could lead to changes in the configurations of LC droplets that are less visible or more difficult to interpret by the naked eye.

In this part of the study, we performed experiments to determine whether ML can further extract hidden features that are relevant to surfactant chemical structure and thereby introduce new degrees of selectivity to this LC sensing platform. To evaluate both the influence of hydrophobic tail groups and hydrophilic head groups, we acquired micrographs of LC droplets

treated with aqueous solutions containing SDS, DTAB, and rhamnolipid (RL), an amphiphilic bacterial toxin that is also used as a commercial biosurfactant. The chemical structures of these

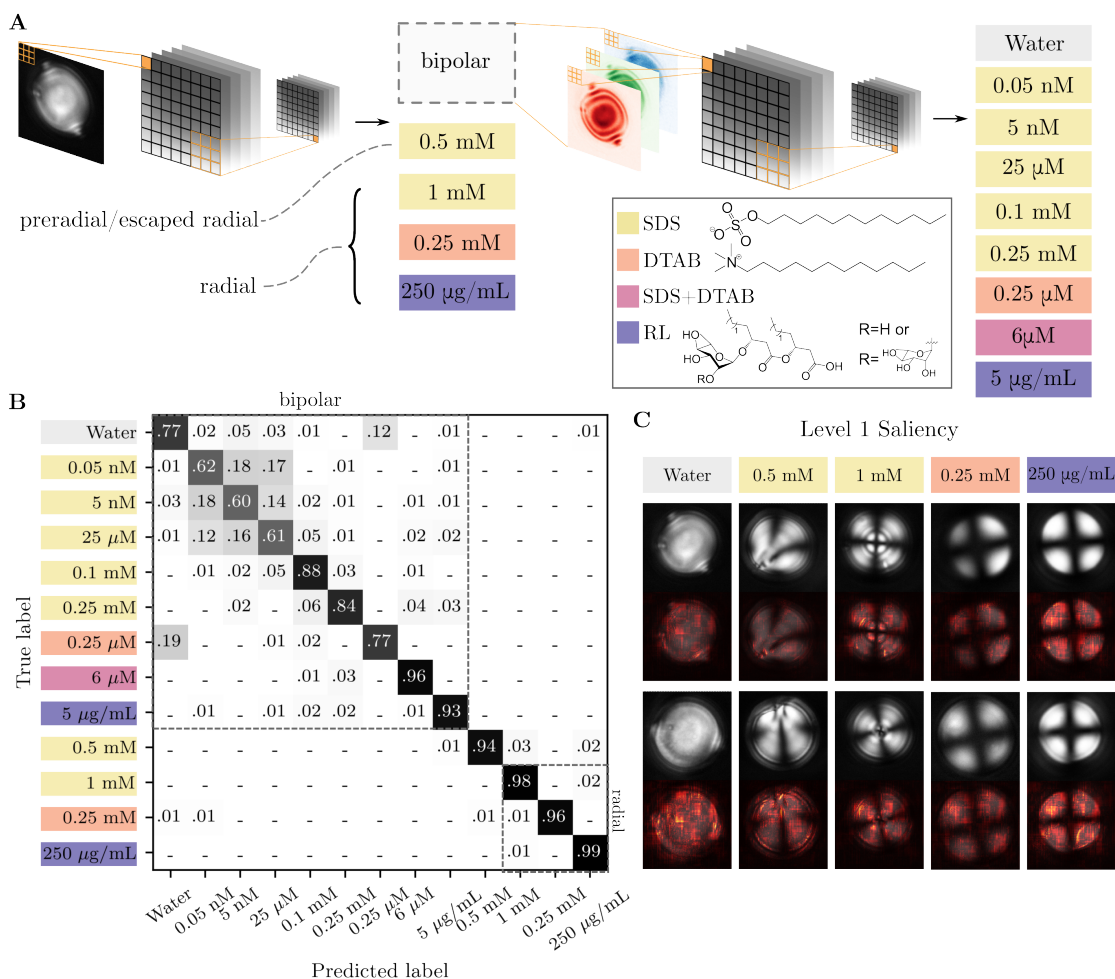


Figure 4. Classification of multiple surfactants using hierarchical CNN. (A) Schematic of proposed hierarchical CNN architecture. The first level focuses on classifying droplet patterns that have unique visual or RD patterns using grayscale micrographs while the second level concentrates on further distinguishing the subtle differences between bipolar-like LC droplets and labeling their specific surfactant type and concentration. The details of the CNN architecture are summarized in Table 1. (B) The confusion matrix from cross-validation. The surfactant types are color-coded, and the surfactant concentrations are listed. (C) Examples of the saliency maps overlaid on top of the original LC micrograph for each class in level 1.

amphiphiles are shown in Figure 4. For both SDS (anionic) and DTAB (cationic), the hydrophobic region contains a single 12-carbon chain. Rhamnolipid contains a mixture of molecules with either one or two fatty acid tails. As shown in Figure 2A, at low concentrations for each surfactant system, droplets remain in bipolar configurations. At high concentrations, all droplets show a cross pattern representing radial configurations. For the SDS-treated droplets,

the patterns remain colorful, while the DTAB-treated droplets appear to be monochromatic, which is more commonly observed for flattened radial droplets. For rhamnolipid, both monochromatic and colorful radial configurations were observed. Additionally, a surfactant mixture system containing an equimolar ratio of SDS and DTAB was prepared as a representation of a more complex surfactant system. The concentration of the surfactant mixture selected here (6 μM) is low enough not to trigger bipolar-to-radial transitions, and leads to LC configurations similar to those of droplets treated with 0.25 mM SDS.

Other results discussed above demonstrate that color-rich images are important for differentiating micrographs that have similar optical responses, whereas grayscale images are sufficient to classify distinguishable defect patterns. Hence, to increase model interpretability and potentially lower computational burden without sacrificing accuracy, we used a hierarchical CNN approach. As illustrated in Figure 4A, grayscale images were submitted to the first level of the CNN to classify the images into five major droplet configurations, all of which have unique visual or RD patterns. At this level, all of the bipolar-like droplets were grouped into a single class. After that, a second-level CNN was trained using the colored LC micrographs to label each bipolar-like droplet with a specific surfactant type and concentration. In general, the LC micrographs were classified into 13 classes (Figure 2A), addressing both the sensitivity and selectivity of surfactant detection.

The results of the hierarchical CNN approach are summarized in the confusion matrix shown in Figure 4B. The CNN in the first hierarchical level achieved a cross-validation accuracy of around 98.9%. This is not surprising based on the RD profiles of the LC micrographs. Despite the presence of some overlapping optical features across radial droplets treated with different surfactants, the RD profiles have already shown some distinctive peaks. We also created saliency

maps, a gradient-based calculation to highlight the key regions in an image used to make a classification, to decipher what information the CNN was capturing. Figure 4C shows a few examples of the original grayscale images with their corresponding saliency maps overlaid on top. The highlighted regions in the saliency map resemble the overall optical patterns of the original images, indicating that the first level of the CNN is classifying the images based on what a trained human eye would consider to be obvious optical features. For the second hierarchical level, the overall classification accuracy was around 77.4%. The categorical-wise prediction accuracy for 0.1 mM SDS and 0.25 μ M DTAB were 88% and 77%, respectively. Despite showing no apparent reliable visible differences in the LC micrographs all at bipolar configurations, the prediction accuracy was relatively high. A closer examination of the confusion matrix revealed that low DTAB concentrations were more likely to be misclassified as water (19%) than low SDS concentrations (3%). This indicates that the trained CNN might be picking out hidden features that are not visually apparent, but are unique to DTAB adsorption. As mentioned above, for droplets treated with 0.25 mM SDS, an equimolar mixture of SDS and DTAB, or 5 μ g/mL rhamnolipid, micrographs share a substantial number of common optical features. These droplets are within the transition spectrum between bipolar and pre-radial configurations and thus cover many complex patterns that either look alike or are difficult to distinguish reliably using visual inspection. Despite that, they were all classified here with relatively high accuracies; for 0.25 mM SDS, the SDS+DTAB mixture, and 5 μ g/mL rhamnolipid, the F1 scores were 84%, 96%, and 93%, respectively.

Understanding Hierarchical CNN Predictions Using Saliency Maps: To further understand the results discussed above, we statistically analyzed the saliency maps using RD profiles of the saliency maps themselves (rather than the LC micrographs, which was performed

to generate the results in Figure 2) as topological descriptors. Figure 5 shows the summarized RD profiles of the saliency maps for each class in the second level of the hierarchical CNN

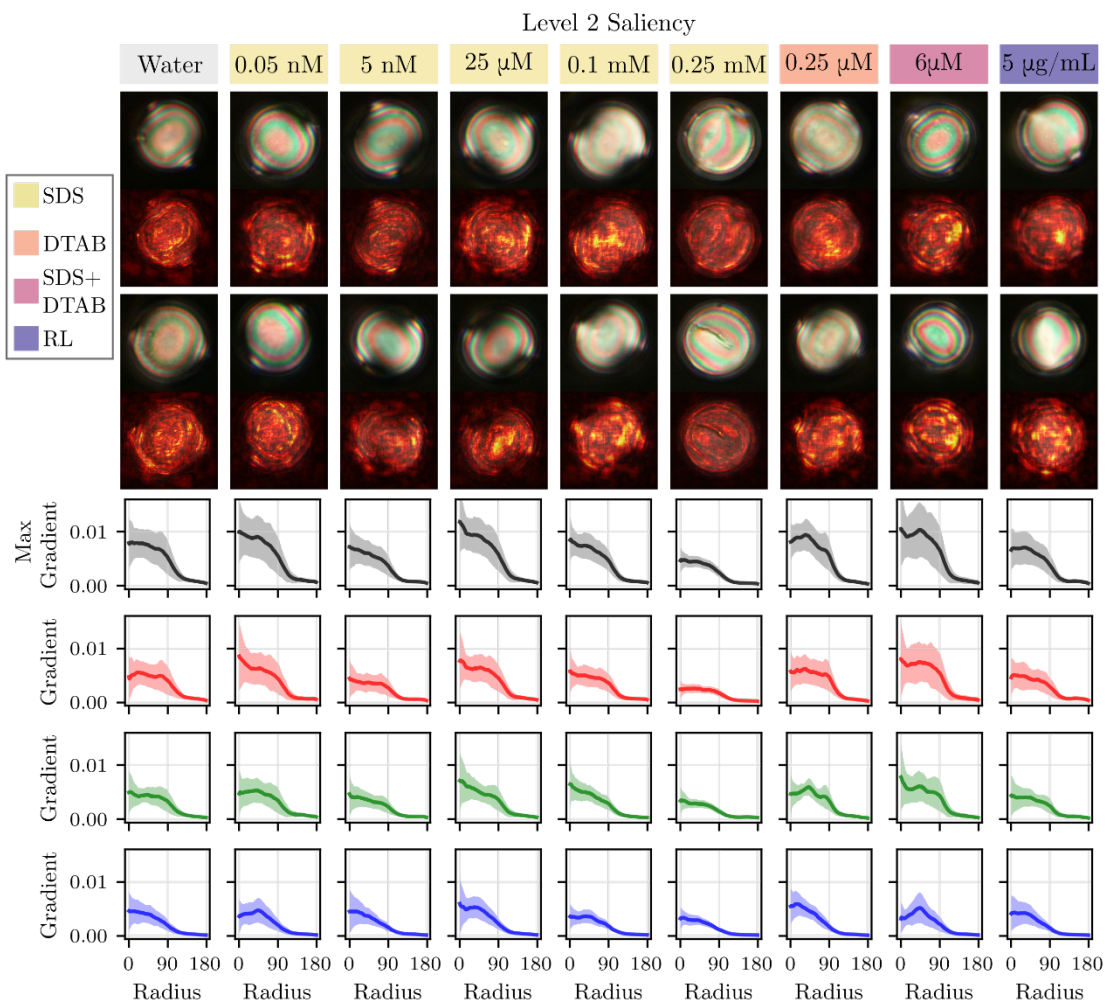


Figure 5. Examples of the saliency maps overlayed on top of the original LC micrograph for each class in level 2. We performed the same RD analysis on the saliency map (instead of the original micrographs) and visualized the summary statistics for all the samples in each class. Most of SDS treated droplets show an overall decreasing trend with the highest intensity point located around the origin of the RD profiles, suggesting that the model is mostly relying on information around the center of the droplets for classification. Most interestingly, we see a unique RD pattern in the green channel for cases where DTAB is present. Unlike SDS or RL, DTAB contains a positively charged head group, which might be the reason for this peak although the micrographs are visually similar.

(referred to as level 2 and corresponding to visually indistinguishable droplets with low surfactant concentrations). As shown in Figure 5, most SDS-treated droplets showed an overall decreasing trend in the sum (grey) channel, with the highest intensity point located around the origin of the RD profiles. This suggests that the algorithm is mostly relying on information

around the center of the droplets to classify them. There are no characteristic peaks that can set different SDS concentrations apart, which is consistent with the relatively low accuracy for low concentrations of SDS. However, closer inspection of the individual color channels revealed that there are some differences in these three channels that facilitate SDS classification. For both 0.1 mM and 0.25 mM SDS-treated droplets, we see a relatively smaller standard deviation, suggesting that the observed highlighted regions are more consistent across all the images as compared to other classes. The relatively low variations in information extracted from the original images might be the reason for their higher prediction accuracy.

For droplets treated with only water (no surfactant), both the sum, the red, and the green channels have a much longer plateau as compared to other treatments. This indicates that the algorithm is using information collected from not only the center but also the near edge of the droplets to classify water. Since defect regions of the bipolar configurations are present around the near edge of the droplets, the presence or positions of the defects are potentially the distinctive feature the algorithm is relying on for classifying water.

We also observed a unique peak in the green channel for the 0.25 μ M DTAB condition. This peak was also present in the surfactant mixture condition, and so might be caused by some hidden but distinctive optical response resulting from adsorption of DTAB. Since both rhamnolipid and SDS are anionic surfactants and their RD profiles lack this peak, the positively charged head group of DTAB might be the reason for the difference in adsorption. We also observed a plateau in the red channel of the 0.25 μ M DTAB condition; this potentially explains the tendency for this condition to be misclassified as water. Finally, we note that the blue channel of the surfactant mixture system also has a unique peak. This peak can also be found in some of the low concentrations of SDS, but it becomes much more prominent in the mixture. This could

be diagnostic of some unique optical features triggered by synergetic effects of the mixed surfactants. We did not see many distinctive patterns in the RD profiles for low concentrations of rhamnolipid, however we note that the highest intensity point in the red and the sum channel is no longer at the center, suggesting that the distinctive features for rhamnolipid are not at the center of the droplet. Overall, on the basis of the results above, we conclude that ML algorithms can be used in combination with this LC droplet array platform to distinguish the presence of different types of aqueous surfactants.

Summary and Conclusions

We have developed a surfactant sensing method using arrays of LC droplets prepared by microcontact printing; the surfactant-induced optical responses of LC droplets were captured by cross-polarized microscopy and analyzed by feature extraction with CNN-based architectures. This experimental system allows for consistent optical responses from uniformly sized LC droplets, and the two-level hierarchical CNN (where the first level classifies defect patterns using grayscale images and the second level classifies color patterns using RGB images) enables the classification of surfactant type and concentration with enhanced selectivity and sensitivity. Delegating different focuses of feature extraction into separate levels also helps model interpretation to obtain physical insights via statistical analysis of the saliency maps, such as RD profiles, which serve as a physically intuitive descriptor for LC micrographs. Overall, these findings demonstrate the potential of the proposed method for on-site and high-throughput sensing of surfactants.

The experimental, materials, and computational components of this study could be expanded in a number of ways to develop new array-based platforms for the rapid detection

and/or identification of many different classes of natural and synthetic amphiphiles with increased accuracy and sensitivity. This array-based approach could also lead to new insights useful for the design and iterative improvement of sensing platforms based on LC-in-water emulsions reported in past studies. The computational framework and analysis methods could also be extended to other image-based models. For instance, separating training tasks into multiple CNN levels and analyzing saliency maps with RD descriptors may be applied to similar LC systems to obtain meaningful insights. Moreover, the droplet detection algorithm could be directly applied to micrographs of LC-in-water emulsions to extract micrographs of individual LC droplets and bypass the need for complex object segmentation algorithms such as YOLO.²⁷ Future studies will also explore the use of different data representations, such as Euler characteristics⁵⁴ (a topological descriptor as an alternative to RD) to decode optical responses of LC micrographs.

Acknowledgments. Financial support for this work was provided, in part, by the NSF through grants provided to the Wisconsin-Puerto Rico Partnership for Research and Education in Materials [Wi(PR)₂EM] at the University of Puerto Rico—Mayaguez and the University of Wisconsin—Madison (PREM; DMR-1827894), the BIGDATA grant IIE-1837812 to V. Z. and R. C. V., and the UW-Madison Materials Research Science and Engineering Center (MRSEC; DMR-1720415). The authors acknowledge the use of instrumentation supported by the NSF through the UW MRSEC (DMR-1720415). The authors gratefully acknowledge Oscar H. Piñeres-Quñones (UPR-Mayagüez) for helpful discussions.

Supporting Information. All data and scripts needed to reproduce the results can be found at https://github.com/zavalab/ML/tree/master/LC_CNN_Micrograph.

ORCID

Fengrui Wang: 0000-0002-2054-2160
Shiyi Qin: 0000-0002-2045-0995
Claribel Acevedo-Vélez: 0000-0001-6297-7747
Reid C. Van Lehn: 0000-0003-4885-6599
Victor M. Zavala: 0000-0002-5744-7378
David M. Lynn: 0000-0002-3140-8637

References

- (1) Rosen, M. J.; Kunjappu, J. T. *Surfactants and Interfacial Phenomena: Fourth Edition*; John Wiley and Sons, 2012.
- (2) Myers, D. *Surfactant Science and Technology: Third Edition*; John Wiley and Sons, 2005.
- (3) Gallou, F.; Isley, N. A.; Ganic, A.; Onken, U.; Parmentier, M. Surfactant Technology Applied Toward an Active Pharmaceutical Ingredient: More than a Simple Green Chemistry Advance. *Green Chemistry* **2015**, *18*, 14-19.
- (4) Torchilin, V. P. Structure and Design of Polymeric Surfactant-Based Drug Delivery Systems. *J. Control. Release* **2001**, *73*, 137-172.
- (5) Kubicek-Sutherland, J. Z.; Vu, D. M.; Mendez, H. M.; Jakhar, S.; Mukundan, H. Detection of Lipid and Amphiphilic Biomarkers for Disease Diagnostics. *Biosensors* **2017**, *7* (3), 25.
- (6) Occhipinti, A.; Eyassu, F.; Rahman, T. J.; Rahman, P. K. S. M.; Angione, C. *In silico* Engineering of *Pseudomonas* Metabolism Reveals New Biomarkers for Increased Biosurfactant Production. *PeerJ* **2018**, *6*, e6046.
- (7) Chiu, C.-F.; Tsai, H.-P.; Chen, Y.-C.; He, Y.-X.; Lin, K.-Y. A.; Yang, H. Self-Assembled Curved Macroporous Photonic Crystal-Based Surfactant Detectors. *ACS Applied Materials & Interfaces* **2017**, *9* (31), 26333-26340.
- (8) Smith, B. C.; Curran, C. A.; Brown, K. W.; Cabarrus, J. L.; Gown, J. B.; McIntyre, J. K.; Moreland, E. E.; Wong, V. L.; Grassley, J. M.; Grue, C. E. Toxicity of Four Surfactants to Juvenile Rainbow Trout: Implications for Use Over Water. *Bull Environ Contam Toxicol* **2004**, *72* (3), 647-654.

- (9) Arora, J.; Ranjan, A.; Chauhan, A.; Biswas, R.; Rajput, V. D.; Sushkova, S.; Mandzhieva, S.; Minkina, T.; Jindal, T. Surfactant Pollution, an Emerging Threat to Ecosystem: Approaches for Effective Bacterial Degradation. *J Appl Microbiol* **2022**, *133* (3), 1229-1244.
- (10) Behrens, B.; Helmer, P. O.; Tiso, T.; Blank, L. M.; Hayen, H. Rhamnolipid Biosurfactant Analysis Using Online Turbulent Flow Chromatography-Liquid Chromatography-Tandem Mass Spectrometry. *J Chromatogr A* **2016**, *1465*, 90-97.
- (11) Cohen, L.; Moreno, A.; Berna, J. L. Two Phase Titration of Anionic Surfactants - A New Approach. *Tenside Surface Det* **1997**, *34* (3), 183-185.
- (12) Gasiorowska, M.; Wroblewska, E. K. Two-Phase Titration Method for Cationic Surface Active Agents Determination with Use of 5-(3-Bromo-4-hydroxy-5-methoxyphenyl)-7,7-dimethyl-7H-indolo[1,2-a]quinolinium Perchlorate Dye. *Tenside Surface Det* **2012**, *49* (1), 23-25.
- (13) Sakač, N.; Jozanović, M.; Karnaš, M.; Sak-Bosnar, M. A New Sensor for Determination of Anionic Surfactants in Detergent Products with Carbon Nanotubes as Solid Contact. *Journal of Surfactants and Detergents* **2017**, *20* (4), 881-889.
- (14) Jiang, S.; Noh, J.; Park, C.; Smith, A. D.; Abbott, N. L.; Zavala, V. M. Using Machine Learning and Liquid Crystal Droplets to Identify and Quantify Endotoxins from Different Bacterial Species. *The Analyst* **2021**, *146* (4), 1224-1233.
- (15) Su, W.; Ding, X. Methods of Endotoxin Detection. *J Lab Autom* **2015**, *20* (4), 354-364.
- (16) Chen, Y.-T.; Yang, J.-T. Detection of an Amphiphilic Biosample in a Paper Microchannel Based on Length. *Biomedical Microdevices* **2015**, *17* (3), 52.
- (17) Lin, I. H.; Miller, D. S.; Bertics, P. J.; Murphy, C. J.; de Pablo, J. J.; Abbott, N. L. Endotoxin-Induced Structural Transformations in Liquid Crystalline Droplets. *Science* **2011**, *332* (6035), 1297-1300.
- (18) Carlton, R. J.; Hunter, J. T.; Miller, D. S.; Abbasi, R.; Mushenheim, P. C.; Tan, L. N.; Abbott, N. L. Chemical and Biological Sensing Using Liquid Crystals. *Liq Cryst Rev* **2013**, *1* (1), 29-51.
- (19) Agarwal, H.; Nyffeler, K. E.; Manna, U.; Blackwell, H. E.; Lynn, D. M. Liquid Crystal-Infused Porous Polymer Surfaces: A "Slippery" Soft Material Platform for the Naked-Eye Detection and Discrimination of Amphiphilic Species. *ACS Appl Mater Interfaces* **2021**, *13* (28), 33652-33663.
- (20) Guo, X.; Manna, U.; Abbott, N. L.; Lynn, D. M. Covalent Immobilization of Caged Liquid Crystal Microdroplets on Surfaces. *ACS Appl Mater Interfaces* **2015**, *7* (48), 26892-26903.

- (21) Manna, U.; Zayas-Gonzalez, Y. M.; Carlton, R. J.; Caruso, F.; Abbott, N. L.; Lynn, D. M. Liquid Crystal Chemical Sensors That Cells Can Wear. *Angew Chem Int Ed Engl* **2013**, *52* (52), 14011-14015.
- (22) Ortiz, B. J.; Boursier, M. E.; Barrett, K. L.; Manson, D. E.; Amador-Noguez, D.; Abbott, N. L.; Blackwell, H. E.; Lynn, D. M. Liquid Crystal Emulsions That Intercept and Report on Bacterial Quorum Sensing. *ACS Appl Mater Interfaces* **2020**, *12* (26), 29056-29065.
- (23) Cao, Y.; Yu, H.; Abbott, N. L.; Zavala, V. M. Machine Learning Algorithms for Liquid Crystal-Based Sensors. *ACS Sensors* **2018**, *3* (11), 2237-2245.
- (24) Xu, Y.; Rather, A. M.; Song, S.; Fang, J. C.; Dupont, R. L.; Kara, U. I.; Chang, Y.; Paulson, J. A.; Qin, R. J.; Bao, X. P.; Wang, X. G. Ultrasensitive and Selective Detection of SARS-CoV-2 Using Thermotropic Liquid Crystals and Image-Based Machine Learning. *Cell Rep Phys Sci* **2020**, *1* (12), 100276.
- (25) Bao, N.; Jiang, S.; Smith, A.; Schauer, J. J.; Mavrikakis, M.; Van Lehn, R. C.; Zavala, V. M.; Abbott, N. L. Sensing Gas Mixtures by Analyzing the Spatiotemporal Optical Responses of Liquid Crystals Using 3D Convolutional Neural Networks. *ACS Sens* **2022**, *7* (9), 2545-2555.
- (26) Smith, A. D.; Abbott, N.; Zavala, V. M. Convolutional Network Analysis of Optical Micrographs for Liquid Crystal Sensors. *The Journal of Physical Chemistry C* **2020**, *124* (28), 15152-15161.
- (27) Frazão, J.; Palma, S. I. C. J.; Costa, H. M. A.; Alves, C.; Roque, A. C. A.; Silveira, M. Optical Gas Sensing with Liquid Crystal Droplets and Convolutional Neural Networks. *Sensors* **2021**, *21* (8), 2854.
- (28) Broderick, A. H.; Azarin, S. M.; Buck, M. E.; Palecek, S. P.; Lynn, D. M. Fabrication and Selective Functionalization of Amine-Reactive Polymer Multilayers on Topographically Patterned Microwell Cell Culture Arrays. *Biomacromolecules* **2011**, *12* (6), 1998-2007.
- (29) Canny, J. A Computational Approach to Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **1986**, *PAMI-8* (6), 679-698.
- (30) Suzuki, S. Topological Structural Analysis of Digitized Binary Images by Border Following. *Computer vision, graphics, and image processing* **1985**, *30* (1), 32-46.
- (31) Redmon, J.; Farhadi, A. YOLOv3: An Incremental Improvement. **2018**.
- (32) Tkalcic, M.; Tasic, J. F. Colour Spaces: Perceptual, Historical and Applicational Background. In *The IEEE Region 8 EUROCON 2003. Computer as a Tool.*, IEEE. DOI: 10.1109/eurcon.2003.1248032.
- (33) Levit, S. L.; Nguyen, J.; Hattrup, N. P.; Rabatin, B. E.; Stwodah, R.; Vasey, C. L.; Zeevi, M. P.; Gillard, M.; D'Angelo, P. A.; Swana, K. W.; Tang, C. Color Space Transformation-

Based Algorithm for Evaluation of Thermochromic Behavior of Cholesteric Liquid Crystals Using Polarized Light Microscopy. *ACS Omega* **2020**, 5 (13), 7149-7157.

- (34) Nair, V.; Hinton, G. E. Rectified Linear Units Improve Restricted Boltzmann Machines. In *Icml*, 2010.
- (35) Santurkar, S.; Tsipras, D.; Ilyas, A.; Madry, A. How Does Batch Normalization Help Optimization? In *Advances in Neural Information Processing Systems 31 (NeurIPS 2018)*, **2018**; <https://arxiv.org/abs/1805.11604>.
- (36) Goodfellow, I.; Bengio, Y.; Courville, A. *Deep learning*; MIT Press, Cambridge, MA, 2016; pp. 339-344.
- (37) Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L. PyTorch: An Imperative Style, High-Performance Deep Learning Library. *Advances in neural information processing systems* **2019**, 32, 8024-8035.
- (38) Sundararajan, M.; Taly, A.; Yan, Q. Axiomatic Attribution for Deep Networks. In *International conference on machine learning*, 2017; PMLR: pp 3319-3328.
- (39) Adebayo, J.; Gilmer, J.; Muelly, M.; Goodfellow, I.; Hardt, M.; Kim, B. Sanity Checks for Saliency Maps. *Advances in Neural Information Processing Systems* **2018**, 2018-December, 9505-9515.
- (40) Kokhlikyan, N.; Miglani, V.; Martin, M.; Wang, E.; Alsallakh, B.; Reynolds, J.; Melnikov, A.; Kliushkina, N.; Araya, C.; Yan, S. Captum: A Unified and Generic Model Interpretability Library for PyTorch. *arXiv preprint arXiv:2009.07896* **2020**.
- (41) Gupta, J. K.; Sivakumar, S.; Caruso, F.; Abbott, N. L. Size-Dependent Ordering of Liquid Crystals Observed in Polymeric Capsules with Micrometer and Smaller Diameters. *Angew Chem Int Ed Engl* **2009**, 48 (9), 1652-1655.
- (42) Gupta, J. K.; Zimmerman, J. S.; de Pablo, J. J.; Caruso, F.; Abbott, N. L. Characterization of Adsorbate-Induced Ordering Transitions of Liquid Crystals Within Monodisperse Droplets. *Langmuir* **2009**, 25 (16), 9016-9024.
- (43) Miller, D. S.; Abbott, N. L. Influence of Droplet Size, pH and Ionic Strength on Endotoxin-Triggered Ordering Transitions in Liquid Crystalline Droplets. *Soft Matter* **2013**, 9 (2), 374-382.
- (44) Kim, J.; Khan, M.; Park, S. Y. Glucose Sensor Using Liquid-Crystal Droplets Made by Microfluidics. *ACS Appl Mater Interfaces* **2013**, 5 (24), 13135-13139.
- (45) Kolacz, J.; Wei, Q. H. Self-Localized Liquid Crystal Micro-Droplet Arrays on Chemically Patterned Surfaces. *Crystals* **2022**, 12 (1), 13.

- (46) Alino, V. J.; Tay, K. X.; Khan, S. A.; Yang, K. L. Inkjet Printing and Release of Monodisperse Liquid Crystal Droplets From Solid Surfaces. *Langmuir* **2012**, *28* (41), 14540-14546.
- (47) Lockwood, N. A.; Gupta, J. K.; Abbott, N. L. Self-Assembly of Amphiphiles, Polymers and Proteins at Interfaces Between Thermotropic Liquid Crystals and Aqueous Phases. *Surface Science Reports* **2008**, *63* (6), 255-293.
- (48) Pineres-Quinones, O. H.; Lynn, D. M.; Acevedo-Velez, C. Environmentally Responsive Emulsions of Thermotropic Liquid Crystals with Exceptional Long-Term Stability and Enhanced Sensitivity to Aqueous Amphiphiles. *Langmuir* **2022**, *38* (3), 957-967.
- (49) Ramou, E.; Rebordão, G.; Palma, S. I. C. J.; Roque, A. C. A. Stable and Oriented Liquid Crystal Droplets Stabilized by Imidazolium Ionic Liquids. *Molecules* **2021**, *26* (19), 6044.
- (50) Carlton, R. J.; Zayas-Gonzalez, Y. M.; Manna, U.; Lynn, D. M.; Abbott, N. L. Surfactant-Induced Ordering and Wetting Transitions of Droplets of Thermotropic Liquid Crystals "Caged" Inside Partially Filled Polymeric Capsules. *Langmuir* **2014**, *30* (49), 14944-14953.
- (51) Alino, V. J.; Sim, P. H.; Choy, W. T.; Fraser, A.; Yang, K. L. Detecting Proteins in Microfluidic Channels Decorated with Liquid Crystal Sensing Dots. *Langmuir* **2012**, *28* (50), 17571-17577.
- (52) Lockwood, N. A.; de Pablo, J. J.; Abbott, N. L. Influence of Surfactant Tail Branching and Organization on the Orientation of Liquid Crystals at Aqueous–Liquid Crystal Interfaces. *Langmuir* **2005**, *21* (15), 6805-6814.
- (53) Brake, J. M.; Mezera, A. D.; Abbott, N. L. Effect of Surfactant Structure on the Orientation of Liquid Crystals at Aqueous–Liquid Crystal Interfaces. *Langmuir* **2003**, *19* (16), 6436-6442.
- (54) Smith, A.; Zavala, V. M. The Euler Characteristic: A General Topological Descriptor for Complex Data. *Computers & Chemical Engineering* **2021**, *154*, 107463.

For Table of Contents Use Only:

