ACS SENSORS

pubs.acs.org/acssensors

Article

Detection of Chemical Warfare Agents by Colorimetric Sensor Arrays

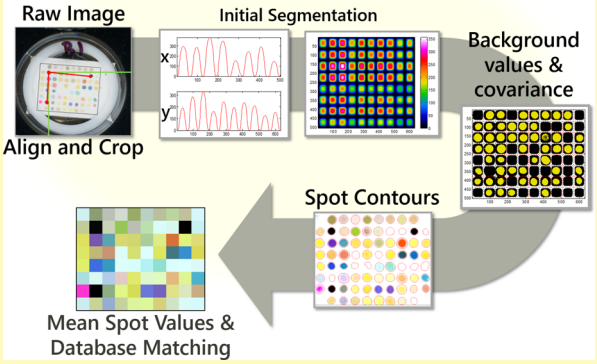
Charles E. Davidson, Melissa M. Dixon, Barry R. Williams, Gary K. Kilper, Sung H. Lim, Raymond A. Martino, Paul Rhodes, Melissa S. Hulet, Ronald W. Miles, Alan C. Samuels, Peter A. Emanuel, and Aleksandr E. Miklos*

Cite This: *ACS Sens.* 2020, 5, 1102−1109

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: We report the successful use of colorimetric arrays to identify chemical warfare agents (CWAs). Methods were developed to interpret and analyze a 73-indicator array with an entirely automated workflow. Using a cross-validated first-nearest-neighbor algorithm for assessing detection and identification performances on 632 exposures, at 30 min postexposure we report, on average, 78% correct chemical identification, 86% correct class-level identification, and 96% correct red light/green light (agent versus non-agent) detection. Of 174 total independent agent test exposures, 164 were correctly identified from a 30 min exposure in the red light/green light context, yielding a 94% correct identification of CWAs. Of 149 independent non-agent exposures, 139 were correctly identified at 30 min in the red light/green light context, yielding a 7% false alarm rate. We find that this is a promising approach for the development of a miniaturized, field-portable analytical equipment suitable for soldiers and first responders.



KEYWORDS: *colorimetric sensor array, chemical warfare agent, chemical identification, field-portable, defense*

Chemical warfare agents (CWAs) are extremely toxic chemicals that are used to kill or incapacitate personnel and to deny areas to an adversary. CWAs are categorized by their mode of action (nerve, blood, blistering, or choking) and rapidly take effect at low levels of exposure.[1] Gear systems that protect against CWAs typically include bulky, durable, and comprehensive barrier garments. Given the rapid onset of symptoms, treatment for CWA exposure (when available) must be rapid as well. Those giving aid must also be protected because once exposed, victims' skin and garments can cause secondary CWA exposure to the responding personnel, for example, when the nerve agent sarin was released on the Tokyo subway in 1995.[2] For medical personnel weighing treatment options in these hazardous environments, agent identification is essential to support rapid selection of effective countermeasures.[3] Development of a robust capability for detecting and identifying CWA threats is a high priority for response personnel; faster notification of potential exposure translates to increased effectiveness of applied countermeasures. The more robust a detector is against a false alarm, the fewer resources that must be expended for protective measures such as donning and doffing protective gear and evacuating personnel.

CWAs comprise a wide variety of chemicals. Some share structural similarities, but considering the broad range of toxic chemicals, fielding a general detection scheme is problematic. Furthermore, many common, less-toxic chemicals with properties similar to those of CWAs will confound detectors.[4,5] Examples include pesticides, industrial chemicals, cleaning agents, personal care products, and workshop chemicals.

Military and domestic first responders currently use multiple methods for rapid CWA detection.[6] Portable devices used for CWA detection and identification rely on a variety of technologies, including Fourier transform infrared spectroscopy (FT-IR),[7] Raman spectrometry,[8] ion mobility spectrometry (IMS),[9] and gas chromatography−mass spectrometry (GC−MS).[10,11] No single system represents an ideal fit for all samples under all conditions, particularly once systems are miniaturized and ruggedized for field use.[12]

Alternatively, colorimetric tests capable of identifying different functional groups of the CWAs provide a simple, cost-effective option for presumptive identification. Although

Received: January 7, 2020
Accepted: March 26, 2020
Published: March 26, 2020

https://dx.doi.org/10.1021/acssensors.0c00042
ACS Sens. 2020, 5, 1102−1109

this approach may be less accurate than spectrometric options, they are cheap and lightweight; therefore, they can be issued widely.[8] The most commonly used CWA colorimetric tests include Dräger tubes, the U.S. military's M256 kit, and M8 and M9 papers.[13,14]

Colorimetric sensor arrays (CSAs) are an attractive sensor technology that combines the favorable cost and robustness of colorimetric assays with greatly enhanced selectivity by increasing dimensionality. The CSAs described herein incorporate a wide variety of cross-reactive indicators to produce characteristic response patterns for each analyte. The function is analogous to human olfaction where a few hundred semispecific receptor types enable discrimination of an enormous number of distinct olfactory stimuli.[15] The CSAs implemented here originated in Prof. Kenneth Suslick's laboratory at the University of Illinois.[16] They are capable of detecting and identifying a wide range of toxic chemicals over a broad concentration range (ppb to ppm) within minutes.[17,18] In addition, these high-dimensional sensor arrays have been demonstrated to differentiate between multiple complex mixtures, including pathogenic bacteria, beer, and coffee.[19] This work was expanded with the use of a CSA technology to detect 20 different toxic industrial chemicals at their immediately dangerous to life or health (IDLH) and permissible exposure limit (PEL) concentrations well within exposure time limits.[18] In addition to this work on industrial hazards, Chulvi et al. used nerve agent simulants and a 16-indicator array to explore the potential application of CSAs to CWA detection,[20] and the use of this approach for CWAs and explosives has been reviewed.[21]

Here, we demonstrate the ability of CSAs to accurately identify CWAs and toxic industrial chemicals in an automated workflow. We evaluated an iSense-proprietary 73-indicator array[22−24] that was previously tested against volatile metabolites from pathogenic microorganisms and urine headspace gases. This universal sensor array covered a diverse portfolio of indicator−analyte interactions, and its sensor responses to CWAs, their precursors and degradation products, fuels, lubricants, cleaning products, solvents, agricultural chemicals, and personal care products were evaluated. To the best of our knowledge, this study is novel in that actual CWAs and an entirely automated data-reduction workflow were used in the experiments. In addition to exploring the potential of this technology for militarily relevant sensing, we also sought to evaluate how much development would be necessary to achieve a "push-button" functionality for easy and correct application by nonspecialist users.

MATERIALS AND METHODS

CSAs. CSAs were manufactured and supplied by iSense, LLC in accordance with previously reported procedures.[25,26] This 73-indicator CSA incorporated a broad spectrum of indicators, including metal-ion-containing dyes to sense Lewis basicity (e.g., amines), pH indicators to sense Brønsted acidity/basicity (e.g., amines and organic acids), solvatochromic dyes with large dipoles to sense molecular polarity (e.g., alcohols), inorganic metal salts and organic redox indicators to monitor redox and chelation properties (e.g., sulfides and arsine), and nucleophilic indicators to detect electrophilic analytes (e.g., aldehydes). Three additional fiducial spots were printed on the array for alignment. Once printed and cured under dry nitrogen for at least 3 days at room temperature, the sensors were individually double-packaged inside polyethylene and aluminized Mylar bags filled with dry nitrogen to protect them until use. Sensors were opened on the day of use or were assembled into test fixtures under nitrogen if intended for later use.

Sensor Cartridges. Headspace analysis was performed by injecting 10 μL of each analyte into a sensor cartridge with the CSA. Liquid analytes were used neat, and nonliquid analytes were dissolved into solutions prior to injection. The cartridge is made by modifying a container (QOSMEDIX no. 29258) as illustrated in Figure 1.

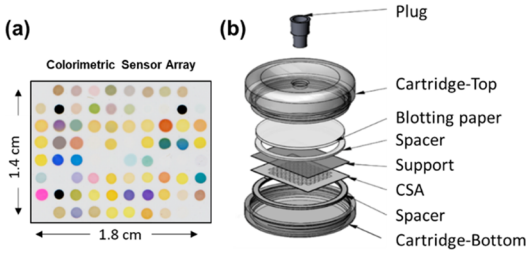


Figure 1. Image of the CSA and an exploded view of the sensor cartridge. Two Teflon split rings suspend the CSA in the housing while a polycarbonate support prevents the CSA from coming in contact with the liquid samples. The blotting paper absorbs the injected liquid sample, and a rubber plug is used to keep the vapor inside.

Experiments were performed by placing multiple test fixtures simultaneously on an Epson V600 flatbed scanner for periodic imaging. A series of "survey" experiments at 5 min scan intervals were performed with multiple analytes on the same scanning bed to maximize throughput. Filling the scanner bed allowed greater throughput but introduced uncertainties in exposure times as spiking all sensor cartridges took several minutes. Repeat exposures with only one analyte tested per run ("standard" exposures) were also performed. This allowed for all test fixtures to be exposed in under 1 min and enabled a shorter 2 min scan interval. In all cases, at least one pre-exposure image was acquired that captured the initial state of the CSAs. Test duration was at least 60 min, and some experiments continued to 120 min for less volatile or reactive analytes.

Chemical Samples. The test matrix included 56 analytes chosen from CWAs, selected CWA precursor and degradation products, solvents commonly used as CWA diluents, and potential confounders, such as household and personal care products, insecticides, and herbicides. Chemicals were selected to (i) provide some coverage across a broad range of threats and (ii) include operational background chemicals that a user might encounter in the field. The complete list of tested analytes can be found in the Supporting Information (Tables S1 and S2). Chemicals were grouped loosely by type or class. Non-agent chemical classes were assigned numeric codes (1−5) to make it easy to distinguish them from the agent classes, which were designated using letters. Commercial products were used as purchased. Cut grass, used at ∼1 g per test, was collected near a sidewalk on Aberdeen Proving Ground, Maryland, shredded finely, and frozen.

A total of 623 CSAs were exposed at ambient laboratory conditions and observed for at least 60 min. Each chemical was tested at least in triplicate.

Image Capture and Processing. Once the sensor cartridges were arranged on the scanner, the CSAs were scanned before the addition of the test analyte and then at regular intervals after addition. Custom software was used to capture, at specified intervals and durations, raw 16 bit color images of the scanner bed that contained no color correction or sharpness manipulation. Each scanner bed image was then subdivided into individual sensor images. The resulting series of CSA images were analyzed to monitor the sensor response over time. Digital image-processing tools were designed to overcome several challenges in extracting reliable color values from each spot on the CSAs, which included non-uniformity in the spot layouts and sizes. The result of the processing was to record average

https://dx.doi.org/10.1021/acssensors.0c00042
ACS Sens. 2020, 5, 1102−1109

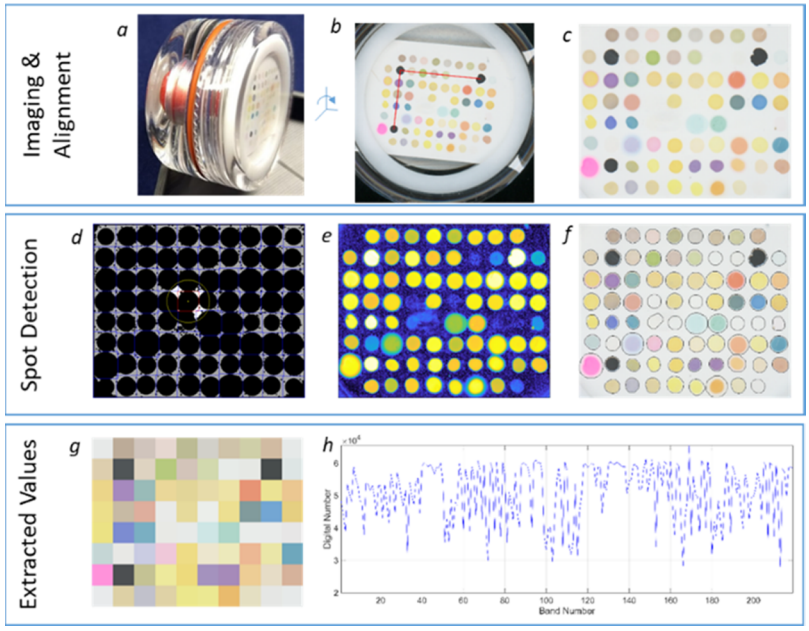


Figure 2. Image capture and processing shown in three major steps.

intensities of red (R), green (G), and blue (B) color channels for each indicator over time. These extracted RGB values formed the data set for chemometric analysis. The procedure (described in more detail in Supporting Information Section S-1) was implemented in MATLAB. As outlined in Figure 2, image processing involves three major steps. First, the sensor cartridge (a) is placed face down on the flatbed scanner, producing a sensor image with arbitrary orientation (b). The image is analyzed to find the three black fiducial spots from which an aligned and cropped image is created (c). In the spot detection phase, the image is broken up into tiles, and regions that are likely blank background areas are identified (gray areas in image (d)). Statistics of the background areas are estimated locally in a region around each tile (shown in white around a central tile in image (d)) to compute an anomaly detection score (e), which is large (bright) for pixels unlikely to be a blank ticket. A threshold is applied to the anomaly detection score to identify spot boundaries (f). The average RGB value for the central portions of each tile is computed (g). The response of the three RGB channels for each of the 73 chromogenic spots characterizes the colors of the ticket at any given time and may be stacked into a single vector in $3 \times 73 = 219$ dimensions or "bands" (h).

Data Analysis. A common method for characterizing CSA technology performance is to visualize the clustering using a dendrogram created from a hierarchical clustering analysis.[27] However, our primary focus is quantitative estimation of performance. Therefore, we must consider partitioning the data into training and test sets.[28,29]

To minimize deviations resulting from production lot variations among CSA batches, each measurement set included CSAs from the same production batch that were collected under identical conditions with the same lot of the analyte. Repeat experiments performed on different days (possibly with different ticket batches) typically showed some systematic differences. This is evidence that within-group variance is typically smaller than between-group variance. CSAs were not available in sufficient quantity to allow performing a complete set of repeat exposures for each chemical. Additionally, CSAs were used as they were available, and no deliberate randomization of manufacturers' batches was performed. Therefore, we evaluated the performance using (i) a truly independent (albeit more limited) test set and (ii) a cross-validated approach using repeated random subsampling where 50% of the data was randomly selected to populate the training set that was used to predict the remaining data. Fifty repeats were used. Selection rules used to populate the independent training and test sets are as follows: Standard exposures (data collected with 2 min intervals between scans) were assigned to the training set, except for any repeat experiments, which were placed in the test set. Survey trials (5 min interval) were placed in the test set unless an exposure was not present in the training set, in which case the first survey experiment was placed into the training set. Any remaining survey exposures were assigned to the test set. These rules generated a training set with 130 agent-exposed tickets and 168 non-agent-exposed tickets. The independent test set consisted of 174 agent-exposed tickets and 160 non-agent-exposed tickets. The ticket distribution among the independent training and test sets is included in Tables S1 and S2. Note that the choices for test-set assignment were conservative and often resulted in more trials being in the test set than in the training set. This conservatism was motivated by a desire to minimize the chances of overestimating the identification performance.

In the Results and Discussion, we present cross-validated prediction performance values (based on repeated, random data segregation into test and training sets) separately from the prediction performance estimates from the (more limited) independent test set. We chose to apply the first-nearest-neighbor (1-NN) rule, which assigns the test sample the identity of the closest training sample. The 1-NN rule is often justified due to its simplicity and because asymptotically (as the number of samples approaches infinity), its performance is no worse than a factor of two from the optimum.[27] Here, we apply the 1-NN rule in an "untimed" fashion: the training data are not limited to only those elapsed times for which the test data point was measured; instead, a closest match may be found at any elapsed time value. This choice was made to minimize sensitivity to expected temporal errors. All distances were computed from the 73 chemically active spots, and the three black control spots (used for alignment) were ignored.

RESULTS AND DISCUSSION

To clarify the appearance of the CSAs and the types of color changes that occur, a small case study is presented for the Windex cleaning fluid and VX nerve agent. Pre-exposure and post-exposure images for two select trials are shown in Figure 3.

Not all spots were visible before exposure; for example, spots on rows 2, 4, 5, and 6 showed very low contrast against the surrounding background regions (note that the corner

https://dx.doi.org/10.1021/acssensors.0c00042
ACS Sens. 2020, 5, 1102−1109

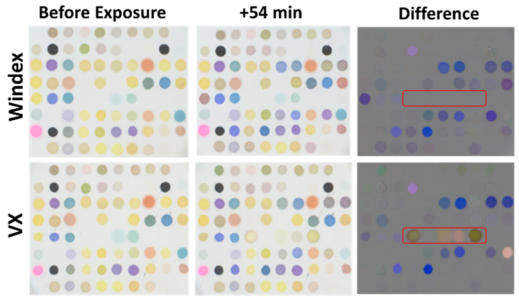


Figure 3. Ticket appearance for two select exposures: Windex and VX nerve agent. The first column shows the CSAs prior to exposure. The middle shows the CSAs after 54 min of exposure. Color difference maps are shown on the right column with the key difference between two analytes in a red box. Gray corresponds to no change.

locations are blank and do not have printed spots). Upon visual inspection, colors of the same spot on the CSA were similar, although there were minor size, shape, and spacing differences. At 54 min post-exposure, the most obvious change was the two "colorless" spots on row 5 that become visible upon exposure to VX but not Windex. The difference images (last column in the figure) show that many spots (20 or more) changed for each chemical, and quite a few color changes (at least 10 spots) were in common between Windex and VX. For these particular trials, the color-change patterns were sufficiently different to support identification, although the overall performance will be determined by the repeatability and consistency of the color-change patterns, which we investigated next.

Figure 4a,b shows the stacked color-change vectors for multiple Windex and VX trials at 54 min without statistical whitening applied. As was observed in Figure 3, both similarities and differences were evident when Windex and VX were compared. Figure 4c shows the raw color change over time with a 3D principle component (PC) plot built from the same Windex and VX trials. Every set of color changes for all spots across the CSA is represented by a point in the PC space, and the origin of the PC space was translated to correspond to no color change. This 3D PC plot accounts for 88% of the variation in the data. Each trial traces out a path or track as the spots on the CSA change color. Some differences in the tracks among replicates of the same chemical are evident, which indicates differences in the CSA response. Most notably, for the atypical set of VX trials that stayed closest to the Windex trials (and did not curve to the right with the other VX trials), the spots on row 5 of the CSA that became visible with exposure to VX (shown in Figure 3) did not change their color. For earlier times, Windex and VX samples are close to one another in the plot but deviate as time progresses. From the spacing between data points in the plot, we can infer that Windex caused quick initial color changes, whereas VX caused the CSA to develop color changes more slowly. Note that the color-change paths in the PC plot are not linear.

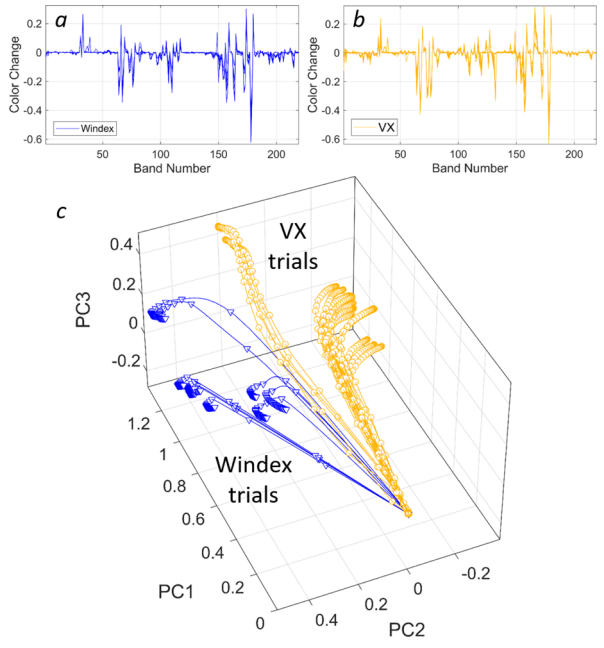


Figure 4. Color change vector response for multiple Windex and VX exposures at 54 min elapsed (panels (a) and (b), respectively), and a principal component representation of the color changes (c) for all time points up to 54 min elapsed. Windex and VX color changes share many similarities but also exhibit differences. For VX, those trials that do not curve to the right in panel (c) did not show the strong color changes on row 5 of the CSA that were pointed out in Figure 3.

Figures 3 and 4 show that two particular exposures, Windex and VX, shared some similarities and exhibited some variation in color change upon repeat exposure. We then investigated relationships between all exposures. Statistically whitened color changes for all trials were pooled at an elapsed time of 54 min, and a dendrogram was computed using the 1-NN and Ward's linkage method.[30] In lieu of a full dendrogram, a more concise image of the relationships among the data is obtained by applying a distance threshold to the dendrogram to cluster the data into groups and then selectively merging the groups to form 11 clusters. Figure 5 is a 3D PC plot (describing 81% of the variation) of the data at 54 min elapsed showing the clusters.

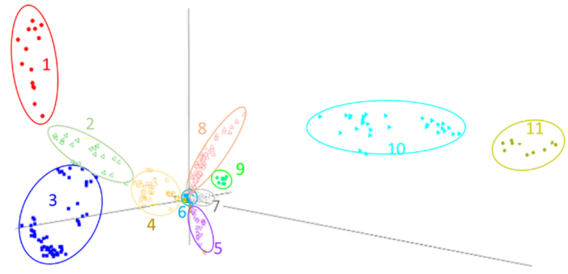


Figure 5. Principle component plot of the color change for all trials at 54 min elapsed. The origin at (0,0,0) represents no color change. The marker color and shape are determined by cluster number. Cluster membership is given in Table 1.

Members of each cluster are shown in Table 1. The origin, which corresponds to no color change, is contained within cluster 6. Color changes are dominated by two major "arms" in the PC plot: the first is largely aligned with the first PC (clusters 1−4) and includes nitrogen-bearing compounds such as amines and ammonia-based cleansers, the second (clusters 8−11) includes most G agents, bleach, and lewisite. Cluster 6 contains exposures with very weak color changes and includes all blank exposures. Any compound not listed as a member of a different cluster in Table 1 is a member of cluster 6. In many cases, similar chemical exposures clustered closely, and for the most part, each chemical was only part of a single cluster. The

https://dx.doi.org/10.1021/acssensors.0c00042
ACS Sens. 2020, 5, 1102−1109

Table 1. Cluster Members

cluster	chemical
1	DEA and TEA
2	Hoppes 9 and Windex
3	DIAET, QL, VR, and VX
4	DiCDi, GB, HN1, and WBG-purple
5	GA and malathion
6	various; includes HN3, DIMP, DMMP, and EMPA blanks
7	CEES, HCl, HD, HT, and T
8	GB, GD, GF, L3, permethrin, and Pine-Sol
9	bleach
10	DCP, GB, GD, GF, and L2
11	L-lewisite

major exceptions were GB, which is in clusters 4, 8, and 10; and HCl, which is in clusters 6 and 7. In this study, most G agents were obtained from at least two different sources; in particular, GB was obtained with varying purity. HCl was also exposed at three different concentrations; one experiment in particular showed extreme time variability where the color changes progressed at very different rates.

To characterize the identification performance, two methods were used (discussed in Data Analysis). We first present the resampling strategy, cross-validation. In this method, half of the data is repeatedly segmented into training and test sets. The identities of the test set members are predicted based on the closest sample in the training set using the Euclidean distance in the statistically whitened space (1-NN rule). The cross-validated strategy allows for computing of the average identification performance, but it may overestimate the identification performance due to insufficient variability in uncontrollable environmental variables. Identification results as a function of time are shown in Figure 6. Three levels of stringency were applied to the identification: chemical identification, class, and alarm. Correct chemical identification meant that the compound abbreviation of the prediction matched the truth. A correct class meant that the chemical or agent "type" code associated with the prediction matched the type of the test sample (i.e., a G agent came back as a G agent, a household cleanser was predicted to be a household cleanser, etc.). Finally, a correct alarm meant that chemical weapons (of any type) were categorized as some type of chemical weapon, and all other chemicals were categorized as a non-agent. In general, the same initial protective actions are warranted regardless of which type of CWA may be present; any CWA detection should result in an alarm. Likewise, misidentifying a nonweapon chemical as another nonweapon chemical does not incur much risk. However, interclass confusion between agent-related and non-agent chemical classes will lead to incorrect decisions. False alarms (identifying non-agent as an agent, also known as a type I error) cause inconvenience, lost time, and unnecessary cost. The most serious errors, termed "missed detections", involve declaring an agent as a non-agent (also known as a false negative or type II error). Missed detections can lead to exposure and death.

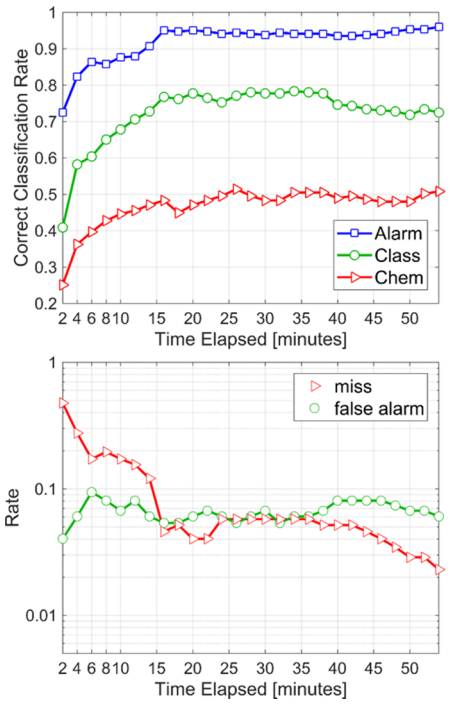


Figure 6. Classification performance over time for the independent test set. Top: average classification rate at three levels of stringency. Bottom: average false alarm and missed detection rates.

Cross-validation performance estimates give average correct chemical identification rates of about 64% at 6 min, 78% at 30 min, and 81% at just under 1 h. Rates of correct identification by chemical class were 79% at 6 min, 86% at 30 min, and 87% at just under 1 h. Correct identification at the alarm level was over 90% at 6 min or longer, and there was little improvement after 15 min. The false alarm rate (declaring a nonthreat chemical a threat) started at about 9% and decreased slowly, hovering at around 5% from about 22 min on. The miss rate (declaring a threat a nonthreat) as a function of time started at 21% and decreased steadily, reaching 1% after 44 min elapsed. Time-resolved curves of the identification rates may be viewed in Figure S1.

At 30 min elapsed, the most egregious analyte-level misidentifications as estimated by the cross-validation procedure included identifying 48% of EMPA trials as blank (a false negative) and 80% of DEG exposures as HD (a false alarm). A graphical representation of the confusion matrix for all analytes at 30 min is shown in Figure S2. The average class confusion matrix is shown in Table S3.

The other method for estimating performance used training and test sets that were selected to be as independent as possible. When using the cross-validated strategy, it is unlikely through random sampling that all exposures from a particular experiment (collected simultaneously on the same scanner) would be eliminated from the training set. Here, the training and test sets were carefully chosen so that exposures from the same experiment could not be assigned to both: all exposures from a given experiment were assigned to either the training or the test set, exclusively. This helped ensure that systematic differences in the response that occurred because of differences in the chemical batch or environmental factors were treated fairly. In an actual use case, there may not be any library exposures performed under these exact test conditions. If these conditions affect the CSA response, there will be a potential performance penalty. Therefore, prediction estimates produced from this carefully chosen train and test set combination are more likely to be accurate. The disadvantages, however, are that the test set is more sparsely populated, and fewer chemicals are represented. Figure 4, which compares Windex and VX exposures, provides a good example of this concept.

https://dx.doi.org/10.1021/acssensors.0c00042
ACS Sens. 2020, 5, 1102−1109

Five atypical VX trials were collected during one experiment. Selected randomly, it is likely that a few trials of each VX group would be represented in both the training and test sets; thus, we would observe good performance. However, if all atypical trials were placed in the test set, it would be less likely to correctly identify those trials as VX. The estimated performance would be less optimistic and more likely to represent the actual performance observed when identifying an as-yet unmeasured VX experiment.

Figure 6 shows the identification rates as a function of time for the independent test set. Correct identification down to the individual chemical level (Figure 6, top: red triangles) was indeed worse than predicted on average by the cross-validation strategy (compared to Figure S1), remaining near 50% correct from about 15 min. However, class-level identification was less severely affected (staying in the range of 72−78% by 14 min elapsed), and alarm-level performance was on par with the cross-validated estimate, exceeding 90% beyond 15 min. False-alarm and miss rates (Figure 6, bottom) were higher than those found in cross-validation. We noted that the slight bump in the miss rate between 6 and 14 min elapsed was in fact caused by VX exposures being identified as Windex (see discussion regarding Figure 4). Analyte-level confusion for the independent test set at 30 min elapsed is summarized graphically in Figure S3, and the class confusion matrix is given in Table S4.

The largest difference in estimated performance between the cross-validated strategy and the independent strategy involved pesticides (class 2) in which all five permethrin test exposures were identified as chemical agent L3. This again points to a systematic difference between the repeat permethrin experiment that populated the test set and the training set replicates. The two experiments were from the same permethrin source but were executed a year apart. Further analysis of this specific case is shown in Figure S4.

The next largest difference was a decrease in performance of the G agent class where the GD and GF exposures that were members of cluster 10 (see Figure 5 and Table 1) were misidentified as DCP (a class 1, non-agent compound). These two particular G exposures were obtained from the same source and were more similar to one another than to any other exposure, but both appeared in the test set. They also resembled the color changes of GB samples that were members of cluster 10 (which was known to be of low purity), although the GB changed more quickly and intensely. Based on our assignment rules, all three of these G agents were placed in the test set. For GF and GD, the closest training set member was DCP; and for GB, the closest training set member was L-lewisite. Due to the similarity of the GF and GD responses to the "impure" GB exposure, and given that GF and GD were obtained from the same source, we suspect that they may also have been of lower purity. A 3D PCA plot of cluster 10 exposures is shown in Figure S5.

These examples highlight the sensitivity of the CSA technology to differences that occur between experiments. For permethrin, we hypothesize that CSA batch differences or degradation due to sample aging contributed to the slight differences in color response. For G agents in cluster 10, there are suspected composition differences as compared to other G exposures and a correspondingly different color response of the CSA ticket. This observation, that replicates within an experiment more closely resembled one another than replicates from other experiments, was consistently observed across the database. This sensitivity is both good and bad: while it demonstrates response specificity and the potential to identify closely related compounds, it may also lead to misidentifications of compounds not captured by the training database. In the case of the cluster 10 G agents, they are identifiable exposures (separable from non-agent exposures); however, if they have no representation in the training database, they will be missed.

We also note that feature selection (eliminating unreliable or unhelpful spots or color channels) could potentially improve identification performance by reducing variation irrelevant or deleterious to the classification task, but this was outside the scope of our initial work described here.

CONCLUSIONS

First-response personnel and warfighters have long sought a low-cost, low-power, low-weight chemical detection system that has high fidelity. The size and weight are even more important for already-burdened warfighters; they must consider the value of a device like this, which, if left behind, could permit them to carry additional ammunition. In this study, we report the successful use of colorimetric arrays to identify CWAs using a cross-validated first-nearest-neighbor (1-NN) algorithm for assessing detection and identification performances on 632 exposures. At 30 min post-exposure, the CSA technology provided 78% correct chemical identification, 86% correct class-level identification, and 96% correct red light/green light (agent versus non-agent) detection. Of 174 total independent agent testing, 164 were correctly identified in the red light/green light context, yielding a 94% correct identification of CWAs. Of 149 independent non-agent exposures, 139 were correctly identified in the red light/green light context, yielding a 7% false alarm rate. This was encouraging for an initial exploratory effort, especially given that the colorimetric array used in this testing was not in any way optimized for CWAs.

CSAs represent a viable and novel technology around which field-portable chemical identifiers and detectors could be constructed. We have shown that the approach is robust against a wide range of analytes, and we have demonstrated that interpretation of the arrays may be entirely automated. Given that microprocessors and high-resolution cameras are available at very low price points and tiny physical sizes, a robust, pocket-sized reader paired with a CSA cartridge has the potential to meet the needs of many first responders entering situations involving unknown chemical threats. Beyond device development, future work to optimize the performance of CSAs for the CWA detection tasks could focus both on indicator chemistries specific to high-hazard agents and on algorithm improvements such as indicator and feature selection strategies.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acssensors.0c00042.

(Table S1) Agent compounds, (Table S2) non-agent compounds, (Section S-1), RGB data extraction, (Figure S1) cross-validation performance, (Figure S2) graphical representation of cross-validation confusion, (Figure S3) graphical representation of independent test confusion, (Table S3) average class confusion matrix at 30 min

https://dx.doi.org/10.1021/acssensors.0c00042
ACS Sens. 2020, 5, 1102−1109

estimated by cross-validation, (Table S4) class confusion matrix at 30 min for the independent test set, and (Figures S4 and S5) color change PCA visualizations (PDF)

AUTHOR INFORMATION

Corresponding Author

Aleksandr E. Miklos − *U.S. Army CCDC Chemical Biological Center, Aberdeen Proving Ground, Aberdeen, Maryland 21010, United States;* orcid.org/0000-0001-6375-2304; Phone: (410)436-5975; Email: aleksandr.e.miklos.civ@mail.mil

Authors

Charles E. Davidson − *Science and Technology Corporation, Belcamp, Maryland 21017, United States*

Melissa M. Dixon − *U.S. Army CCDC Chemical Biological Center, Aberdeen Proving Ground, Aberdeen, Maryland 21010, United States*

Barry R. Williams − *Leidos, Abingdon, Maryland 21009, United States*

Gary K. Kilper − *Excet, Springfield, Virginia 22150, United States*

Sung H. Lim − *iSense, Mountain View, California 94043, United States*

Raymond A. Martino − *iSense, Mountain View, California 94043, United States*

Paul Rhodes − *iSense, Mountain View, California 94043, United States*

Melissa S. Hulet − *Leidos, Abingdon, Maryland 21009, United States*

Ronald W. Miles − *U.S. Army CCDC Chemical Biological Center, Aberdeen Proving Ground, Aberdeen, Maryland 21010, United States*

Alan C. Samuels − *U.S. Army CCDC Chemical Biological Center, Aberdeen Proving Ground, Aberdeen, Maryland 21010, United States*

Peter A. Emanuel − *U.S. Army CCDC Chemical Biological Center, Aberdeen Proving Ground, Aberdeen, Maryland 21010, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acssensors.0c00042

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Funding

This work was funded by the Defense Threat Reduction Agency/Joint Science and Technology Office for Chemical and Biological Defense.

Notes

The authors declare the following competing financial interest(s): Drs. Lim and Rhodes and Mr. Martino have interests in iSense LLC which engineers and supplies the CSAs (colorimetric sensor arrays) described in this manuscript.

ACKNOWLEDGMENTS

The authors wish to acknowledge the early funding for this work from the Edgewood Chemical Biological Center (ECBC) Director's Office and The Joint Program Executive Office Nuclear Biological and Chemical Contamination Avoidance division (JPEO NBCCA). The authors also wish to thank Lisa Clark Hersey for considerable assistance in editing this manuscript and Drs. Brandi Vann and Anthony Esposito for their guidance and support.

ABBREVIATIONS

CWA, chemical warfare agent; CSA, colorimetric sensor array. Please see Tables S1 and S2 for chemical names.

REFERENCES

(1) Hilmas, C. J.; Smart, J. K.; Hill, B. A. *Medical aspects of chemical warfare*. US Army Office of the Surgeon General, 2008.

(2) Okumura, S.; Okumura, T.; Ishimatsu, S.; Miura, K.; Maekawa, H.; Naito, T. Clinical review: Tokyo−protecting the health care worker during a chemical mass casualty event: an important issue of continuing relevance. *Crit. Care* **2005**, *9*, 397−400.

(3) Sidell, F. R.; Newmark, J.; McDonough, J. H. Nerve agents. In *Medical aspects of chemical warfare*; US Army Office of the Surgeon General: Rockville 2008.

(4) Rosen, P.; Abood, L.; Benjamin, G.; Bowler, R.; Daniels, J.; DeAtley, C.; Goldfrank, L.; Hauer, J.; Larson, K.; Lederberg, J. *Chemical and biological terrorism: research and development to improve civilian medical response*. Institute of Medicine, National Research Council; National Academy Press: 1999.

(5) Sferopoulos, R. *A review of chemical warfare agent (CWA) detector technologies and commercial-off-the-shelf items*; Defence Science And Technology Organization (Australia): 2009.

(6) Emanuel, P.; Caples, M. *Global CBRN Detector Market Survey*; Edgewood Chemical Biological Center Aberdeen Proving Ground United States: 2017.

(7) US DHS, D. P. *Portable Infrared Spectroscopy Chemical Detectors Assessment Report*; https://www.dhs.gov/sites/default/files/publications/Portable-Infrared-Spectroscopy-Chemical-Detectors-ASR_0616-508.pdf, 2016.

(8) Kondo, T.; Hashimoto, R.; Ohrui, Y.; Sekioka, R.; Nogami, T.; Muta, F.; Seto, Y. Analysis of chemical warfare agents by portable Raman spectrometer with both 785nm and 1064nm excitation. *Forensic Sci. Int.* **2018**, *291*, 23−38.

(9) Mäkinen, M. A.; Anttalainen, O. A.; Sillanpää, M. E. T. Ion mobility spectrometry and its applications in detection of chemical warfare agents. *Anal. Chem.* **2010**, *82*, 9594−9600.

(10) Alcaraz, A. *Gas Chromatography/Mass Spectrometry in Analysis of Chemicals Relevant to the Chemical Weapons Convention*. John Wiley and Sons, Ltd.: Chichester, UK, 2012.

(11) Murray, G. M. *Detection and Screening of Chemicals Related to the Chemical Weapons Convention*. John Wiley and Sons, Ltd.: Chichester, UK, 2013.

(12) Emanuel, P.; Caples, M. *WMD Detector Selector Web Tool*. https://www.wmddetectorselector.army.mil.

(13) Reich, N.; Wagner, P.; Geelhaar, J. *Report of test for Canadian paper, chemical agent detector, 3-way liquid*; 1963.

(14) Thoraval, D.; Bets, R. W.; Bovenkamp, J. W.; Lacroix, B. V. *Development of Paper, Chemical Agent Detector, 3-Way Liquid Containing Non-Mutagenic Dyes. 2. Replacement of the Blue Indicator Dye Ethyl-bis-(2, 4-Dinitrophenyl Acetate (EDA))*; Defence Research Establishment Ottowa (Canada): 1988.

(15) Bushdid, C.; Magnasco, M. O.; Vosshall, L. B.; Keller, A. Humans can discriminate more than 1 trillion olfactory stimuli. *Science* **2014**, *343*, 1370−1372.

(16) Rakow, N. A.; Suslick, K. S. A colorimetric sensor array for odour visualization. *Nature* **2000**, *406*, 710−713.

(17) Feng, L.; Musto, C. J.; Suslick, K. S. A simple and highly sensitive colorimetric detection method for gaseous formaldehyde. *J. Am. Chem. Soc.* **2010**, *132*, 4046−4047.

(18) Feng, L.; Musto, C. J.; Kemling, J. W.; Lim, S. H.; Suslick, K. S. A colorimetric sensor array for identification of toxic gases below permissable exposure limits. *Chem. Commun.* **2010**, *46*, 2037−2039.

(19) Li, Z.; Askim, J. R.; Suslick, K. S. The Optoelectronic Nose: Colorimetric and Fluorometric Sensor Arrays. *Chem. Rev.* **2019**, *119*, 231−292.
(20) Chulvi, K.; Gaviña, P.; Costero, A. M.; Gil, S.; Parra, M.; Gotor, R.; Royo, S.; Martínez-Mañez, R.; Sancenón, F.; Vivancos, J.-L. Discrimination of nerve gases mimics and other organophosphorous derivatives in gas phase using a colorimetric probe array. *Chem. Commun.* **2012**, *48*, 10105−10107.
(21) Kangas, M. J.; Burks, R. M.; Atwater, J.; Lukowicz, R. M.; Williams, P.; Holmes, A. E. Colorimetric Sensor Arrays for the Detection and Identification of Chemical Weapons and Explosives. *Crit. Rev. Anal. Chem.* **2017**, *47*, 138−153.
(22) Lim, S. H.; Martino, R.; Anikst, V.; Xu, Z.; Mix, S.; Benjamin, R.; Schub, H.; Eiden, M.; Rhodes, P. A.; Banaei, N. Rapid Diagnosis of Tuberculosis from Analysis of Urine Volatile Organic Compounds. *ACS Sens.* **2016**, *1*, 852−856.
(23) Lim, S. H.; Mix, S.; Xu, Z.; Taba, B.; Budvytiene, I.; Berliner, A. N.; Queralto, N.; Churi, Y. S.; Huang, R. S.; Eiden, M.; Martino, R. A.; Rhodes, P.; Banaei, N. Colorimetric Sensor Array Allows Fast Detection and Simultaneous Identification of Sepsis-Causing Bacteria in Spiked Blood Culture. *J. Clin. Microbiol.* **2014**, *52*, 592−598.
(24) Shrestha, N. K.; Lim, S. H.; Wilson, D. A.; SalasVargas, A. V.; Churi, Y. S.; Rhodes, P. A.; Mazzone, P. J.; Procop, G. W. The combined rapid detection and species-level identification of yeasts in simulated blood culture using a colorimetric sensor array. *PLoS One* **2017**, *12*, e0173130.
(25) Lim, S. H.; Feng, L.; Kemling, J. W.; Musto, C. J.; Suslick, K. S. An optoelectronic nose for the detection of toxic gases. *Nat. Chem.* **2009**, *1*, 562−567.
(26) Lonsdale, C. L.; Taba, B.; Queralto, N.; Lukaszewski, R. A.; Martino, R. A.; Rhodes, P. A.; Lim, S. H. The use of colorimetric sensor arrays to discriminate between pathogenic bacteria. *PLoS One* **2013**, *8*, No. e62726.
(27) Duda, R. O.; Hart, P. E.; Stork, D. G. *Pattern Classification*. 2nd ed.; Wiley-Interscience: New York, 2001.
(28) Hastie, T.; Tibshirani, R.; Friedman, J. *The elements of statistical learning: data mining, inference, and prediction*. 2nd ed.; Springer: New York, 2013.
(29) Mohri, M.; Rostamizadeh, A.; Talwalkar, A. *Foundations of machine learning*. MIT press: 2018.
(30) Ward, J. H., Jr. Hierarchical Grouping to Optimize an Objective Function. *J. Am. Stat. Assoc.* **1963**, *58*, 236−244.

https://dx.doi.org/10.1021/acssensors.0c00042
ACS Sens. 2020, 5, 1102−1109