Sensors & Transducers, Vol. 238, Issue 11, November 2019, pp. 94-98

www.sensorsportal.com

Sensors & Transducers

Published by IFSA Publishing, S. L., 2019
http://www.sensorsportal.com

Rapid Strain Differentiation of *E. coli*-inoculated Urine Using Olfactory-based Smart Sensors

[1] Aminat Adebiyi, [1] Sarath Swaminathan, [2] Nam Than, [1] Mohammed Abdi, [1] Amy N. Bowers, [1] Andrea Fasoli, [1] Alberto Mannari and [1] Luisa Bozano

[1] IBM Research, Almaden, 650 Harry Road, San Jose, CA USA, 95120
[2] Department of Biomedical Engineering, San Jose State University, San Jose, CA USA
Tel.: + 14089271579
E-mail: aminat.adebiyi@ibm.com

Received: 30 August 2019 /Accepted: 15 October 2019 /Published: 30 November 2019

Abstract: Urinary tract infections (UTIs) have a large effect on the public healthcare system. Millions are affected each year, and are likely to have the infection recur with increased resistance to antibiotics. This is partly due to the lack of a targeted healthcare regimen that requires pathogen identification. As traditional cultures can take several days, there is a need for a real-time diagnostic tool for identifying uropathogens. In our study, the use of rapid olfactory-based smart sensors for the strain identification was explored by taking measurements of in-vitro assays involving the inoculation of two strains of *E. coli* (K12 and Uropathogenic *E. coli*) in urine. Using a Support Vector Machine classifier to train on two previous data collections yielded 95.83 % accuracy when tested on the third subsequent collection. These results demonstrate the promise of smart sensing in differentiating close metabolomic signatures of *E. coli* sand are expected to improve with more extensive data collection, assay development and sensor augmentation.

Keywords: Smart sensors, Electronic nose, *E. coli,* Urinary tract infection, Precision healthcare, Internet of things.

1. Introduction

With over 3 million cases a year in the United States alone, Urinary Tract Infections (UTIs) are the most prevalent bacterial infection and poses a great risk of morbidity in vulnerable populations, including children, pregnant women and the elderly [1]. In 2011, the estimated economic burden was estimated at $2.8 billion in the U.S. and the cost is only expected to increase with increasing hospitalization [2].

Accurate diagnosis of UTIs is difficult at the point of care, as it depends on both the presence of symptoms and a positive urine culture, which takes a minimum of 24 hours or up to multiple days for results [3]. A rapid test would lead to the correct empirical course of treatment targeted to the specific bacterial strain - significantly improving treatment outcomes, reducing medical costs and the likelihood of antimicrobial resistance.

Olfactory-based sensors like electronic noses have been widely studied and applied across industries to associate volatile organic compound profiles with the emitting analyte [4]. In healthcare, metabolomic markers can be used for studying disease diagnostics and has motivated the study of volatile organic compounds that correspond to these biomarkers [5].

In this paper, we introduce the Electronic Volatile Analyzer (EVA), a rapid, gas sensing IoT platform

under development at IBM Research, Almaden - designed to deliver classification results under two minutes. A main objective of this study was to demonstrate the utility of rapid gas sensing IoT platforms in differentiating K12 and Uropathogenic (UPEC) strains of *Escherichia coli* (*E. coli*) inoculated into urine at comparable cellular counts.

2. Methods

2.1. The IBM Electronic Volatile Analyzer™

The IBM Electronic Volatile Analyzer™ (EVA, Fig. 1) is an electronic nose under development at IBM Research Almaden. The design of the platform is modular: each gas sensor is mounted on its own sensor module, a Printed Circuit Board (PCB) of common design carrying an integrated microcontroller and the circuitry needed to operate the sensor. The sensor modules communicate via I2C protocol with a central hub (BeagleBone Black), a single-board computer that orchestrates the sensor modules and processes the multisensorial output. A set of six commercial Metal Oxide (MOX) gas sensors was used to collect the measurements described in the following sections. The selection of the sensor portfolio was based on the indications of their respective manufacturers regarding the nominal target gases of each sensor. Consequently, the as-selected sensor array was expected to respond to a variety of volatile molecules, including alcohols, hydrocarbons, ammonia, methane, as a well as a wide range of volatile organic compounds (VOC), with a lower limit of detection at parts per million (ppm) level. Each MOX sensor was operated using an individualized multi-step, periodic voltage profile applied to the heating element, which resulted in step-wise modulation of the device temperature (Figs. 1b(i) and (ii)). Modulation of the temperature of MOX sensing elements is a well-known technique that can be used to enhance and tune the dependence of the sensing element resistance to the surrounding environment (Fig. 2b) [7]. The heater voltage profiles applied to the EVA sensors were optimized to maximize the sensitivity of each sensor with respect to a subset of target VOC, as well as to improve response orthogonality. Although the duration and amplitude of the individual waveforms was adjusted independently for each sensor, all waveforms were synchronized to a period of 80 s to simplify the handling and processing of the sensor array outputs. The resistance of each MOX sensing element was monitored at fixed voltage at a rate of 10 Hz. During operation, a constant flow of 150 ± 10 sccm was established by means of a small pump, drawing air from the environment in a continuous fashion while also collecting the headspace of the sample of interest. The vapor-carrying flow was directed towards an enclosed chamber containing the six MOX gas sensors, for the exposure and detection to take place.

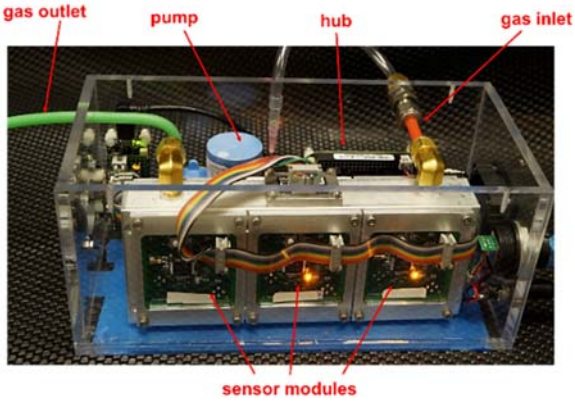

Fig. 1a. Electronic Volatile Analyzer (EVA) Sensor Array prototype with parts labeled.

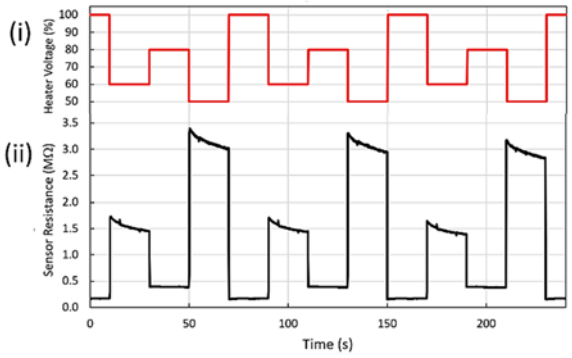

Fig. 1b. Example of a temperature profile modulation for MOX sensors for IBM EVA™: (i) Periodic waveform of heater voltage, expressed as a percentage of the maximum operating voltage recommended by the sensor manufacturer; (ii) Corresponding variations in MOX sensor resistance under constant environment.

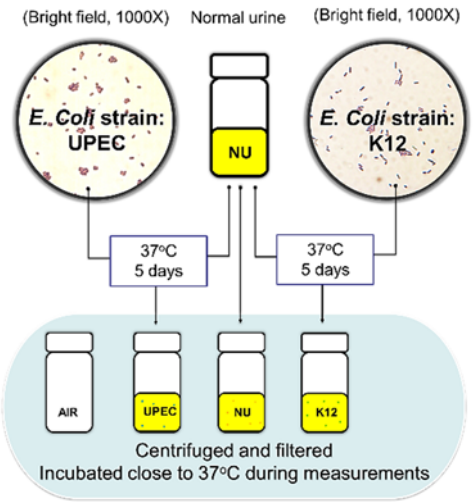

Fig. 2. Diagrammatic representation of assay development and conditions of data collection. Each strain of *E. coli* (K12 and UPEC) was inoculated separately into Normal Urine (NU) and incubated over five days, where it was filtered to trap volatiles. Each data collection involved four sample vials with Lab Air, NU, K12 and UPEC. Samples were incubated close to 37°C during measurements.

2.2. Assay Development and Sample Preparation

Two liquid UTI assays were developed by culturing *Escherichia coli* (Migula) Castellani and Chalmers (ATCC® 29425™) and *Escherichia coli* (Migula) Castellani and Chalmers (ATCC® 700928™) in Normal Human Urine (purchased from UTAK Laboratories). The Normal Human Urine (NU) was centrifuged and filtered with 0.2 µm filters prior to inoculation, to remove sediments and to aid in the homogeneity of samples. Each of the *E. coli* strains were then separately inoculated into NU at a fixed concentration, and incubated at 37 ºC over five days to guarantee a terminal concentration of 10^9 CFU/ml. After samples were plated, they were further filtered to remove cellular bodies of the bacteria, leaving only their emitted metabolomic volatiles behind [6].

Sample volumes of 5ml were dispensed in 20ml Wheaton Septa Top vials for measurement. The list of samples measured were filtered Uropathogenic *E. Coli* (UPEC) in NU, K12 *E. coli* in NU, Normal Urine, and Lab Air as reference.

2.3. Data Collection and Processing

Collection of data for the training and testing datasets were made at different times, with a minimum of eight hours gap in between. The training dataset was comprised of two experiments, and the test dataset was the third sequential acquisition. Headspace measurements of each of the four sample vials were made using the IBM EVA™. The sequence of collection from the sample vials was randomized over the course of an eight hour collection period, with an air flush between vials. Samples were incubated close to 37ºC during measurements, to aid in the evaporation of trapped volatiles. The response of the sensors was saved using Data Acquisition System (DAS) before feature extraction and labeling of the sensorial data were performed. Various feature extraction approaches were explored on sensor raw resistance curves to capture sensitivity and amplitude differences in the samples. Machine learning was performed mainly on the extracted feature set to identify the odor classes from the testing dataset (Testing Batch) by training on the two prior collections of the training set (Training Batches A and B), which is the main analytical model in view of the intended application. Given the limited number of total examples, traditional machine learning classifiers were selected, including Support Vector Machine (SVM), Logistic regression, Random Forest and Multilayer Perceptron. Implementation of these classifiers was from the Scikit-learn 0.19.1 package on Python [8]. The entire dataset was cross validated using three-fold validation with each batch serving as a fold – also referred to here as leave-one-batch-out. The parameters for cross validation are shown in Table 1.

Table 1. Cross validation parameters for selected traditional machine learning classifiers

Classifier Type	Tuning Parameters
Logistic Regression (LR)	'l1' and 'l2' loss; C values from 0.001 to 1.0; optimization algorithms – 'newton-cg', 'libfgs', 'liblinear', 'sag', 'saga'; multiclass classifications using 'one-vs-rest' and 'multinomial'
Support Vector Machine (SVM)	'linear' and 'rbf' kernels; C values from 0.001 to 1.0; 'one-vs-one' and 'one-vs-rest' for multiclass classification
Random Forest (RF)	estimators from 5 to 20; maximum depth of tree from 2 to 10 or None; max_features - 'auto', 'sqrt', 'log2' and None; and criterion – 'gini' or 'entropy'
Multi-layer Perceptron (MLP)	one to two hidden layers with size ranging from 10 to 25; sigmoid and relu activation functions; stochastic gradient descent and adam solvers; l2 penalty from 0.0001 to 0.01

3. Data Analysis

3.1. Data Visualization

Resultant data from the projected input set by sample label is shown in Fig. 3. The two main principle components of the dataset represented 98.7% of variation in the data, showing the dataset is separable and easily defined. Training Batch A showed more variation across labels compared to Training Batch B and Testing Batch. Based on the sequence of collection, measurement drift was experienced in Training Batch A but stabilized between Training Batch B and Testing Batch. This phenomenon could also represent the correlation of concentration of volatiles in samples between the batches.

3.2. Features

The raw resistance sensor output values are high dimensional and noisy and are not suitable to be directly used as input for classifiers. To transform this raw values into usable features and thereby to represent the data better for machine learning algorithms, we explored the following feature extraction approaches:

Max: Maximum response values for sensors corresponds to their degree of reaction to the gas. For each sensor, the resistance curve is equally divided into ten sections and extracted the mean resistance value for each sections.

Fractional Difference: Fractional difference is an another kind of maximum value feature, which is defined to correct additive errors in the raw resistance values. For each sensor, the maximum resistance value (Rmax) and minimum resistance value (Rmin) are taken. The fractional difference value for a sensor is defined as (Rmax – Rmin) / Rmin.

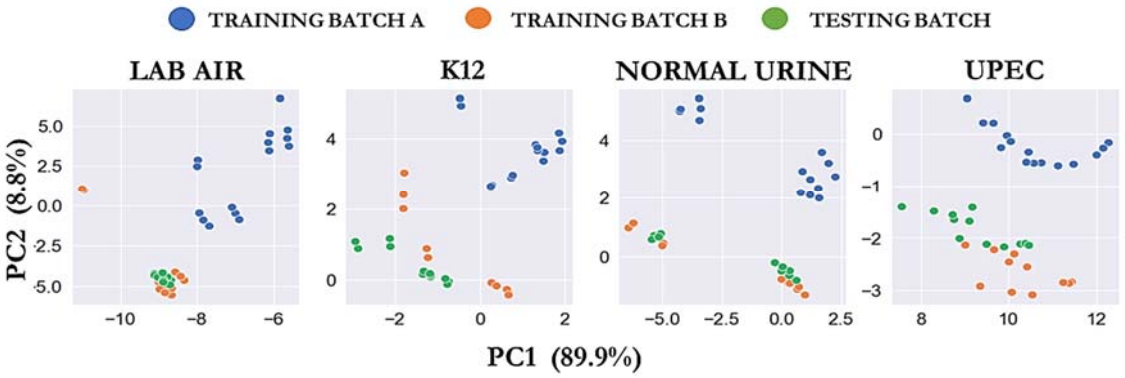

Fig. 3. Principal Component Analysis (PCA) visualization of each sample label (Lab Air, Normal Human Urine with K12 E.coli strain, Normal Human Urine and Normal Human Urine with UPEC *E. coli* strain) by data collection batch. The two axes represent Princpal Components (PCs) 1 and 2 which describes 98.7% percent of the variation in the data. Training Batches 1(blue) and 2 (orange) were collected in sequence before the terminal collection of the Testing Batch (green). Both Training Batches were used to develop the classifier for which the Testing Batch was tested on. Training Batch B and Testing Batch overlap closely, while Training Batch A is more distant with respect to the other batches.

First Derivative: To smoothen the resistance curve for each sensor, the curve is divided into ten equal sections and the mean is taken. The first derivative is then taken over this "smoothened" mean representation.

Mean: For each sensor, the resistance curve is equally divided into ten sections and the mean resistance value for each section is taken. This ten-dimensional mean representation of the raw data, smoothens the initial raw resistance curve of sensors.

PCA: This common statistical technique was used to reduce dimensionality of mean values to obtain a six-dimensional representation of the sensor data.

Effectiveness of these features was measured by finding the average accuracy of the models with leave-one-batch-out classification. Initially the best parameter for the classifiers were tuned for each of these features. Cross validation based on the parameters list in Table 1 was used for parameter selection. Leave-one-batch-out or three-fold validation was done by training on two batches and testing on the third batch. This was repeated where each of the three batches served as the Testing Batch at least once. Average test accuracy for all three batches of data is reported in Table 2.

Based on the accuracies from Table 2, mean and max features performed equally well while fractional difference fell short. And one interesting thing to note is when combined mean and max features accuracy reduced significantly. PCA features did slightly better than max and mean features. First derivative feature gave surprisingly great result and when combined with max features it improved further. As we can see from Table 2. *Max & First derivative* feature gave perfect result of accuracy 1.0 for SVM, Random Forest and Multi-layer Perceptron. And Random Forest gave slightly better result for other features compared to other classifiers.

Based on the tuning parameters the best parameters for each classifier is outlined in Table 3. Given the linearity of the data shown in Fig. 3, the cross validation parameters for classification show very high separability based on the high accuracy achieved for all feature types and models attempted, and is confirmed through the PCA visualization.

Table 2. Average leave-one-batch-out accuracy scores for selected classifiers and features

Features	LR	SVM	RF	MLP
Mean	0.91	0.96	0.91	0.95
Max	0.94	0.99	0.93	0.95
First Derivative	1.00	1.00	1.00	0.96
Fractional Difference	0.84	0.87	0.92	0.87
PCA	0.91	0.98	1.00	0.93
Mean & Max	0.85	0.85	0.90	0.92
Max & First Derivative	0.98	1.00	1.00	1.00
Average	**0.92**	**0.95**	**0.95**	**0.94**

Table 3. Best parameters for classifier type for training data

Classifier Type	Best Parameters
Logistic Regression	C=0.1, multi_class='multinomial', penalty='l2', solver='lbfgs'
Support Vector Machine	C=1.0, decision_function_shape='ovr', kernel='linear'
Random Forest	criterion='gini', max_depth=5, max_features=n_features, n_estimators= 10
Mult-layer Perceptron	activation='relu', alpha=0.01, hidden layer 1 = 20, hidden layer 2 = 10, solver= 'adam'

The detailed results of the performance of the classifiers on the Max and First derivative feature set for the three-fold validation with each batch serving as a fold is shown in Table 4. Table 4a shows the accuracy of using the indicated batch as a hold out while training on the remaining batches. Table 4b shows the F1 scores, which is the harmonic mean of the precision and recall, measuring the accuracy of the classifier on the dataset. In addition to the main sequential model for our application, the results for the remainder of the batches as hold-outs are shown.

Table 4. Normalized accuracy and F1 scores of leave-one-out batch by classifier type when trained on two other batches and tested on indicated sample set with Max and First Derivative as feature type.

(a) Normalized Accuracy Scores

Batch	LR	SVM	RF	MLP
Training Batch A	0.94	1.00	1.00	1.00
Training Batch B	1.00	1.00	1.00	1.00
Testing Batch	1.00	1.00	1.00	1.00
Average	**0.98**	**1.00**	**1.00**	**1.00**

(b) F1 scores

Batch	LR	SVM	RF	MLP
Training Batch A	0.93	1.00	1.00	1.00
Training Batch B	1.00	1.00	1.00	1.00
Testing Batch	1.00	1.00	1.00	1.00
Average	**0.98**	**1.00**	**1.00**	**1.00**

For the main sequential model the application is based on, which involves training on Training Batches A and B and testing on the last collection, all classifiers performed perfectly with an accuracy and f1 score of 1.0. While all classifiers performed well on all combination of leave-one-batch-out testing, Logistic regression did slightly worse for predicting Training Batch A.

All the models performed very well on average for all the feature types, with over 90 % accuracy. SVM and Random Forest, performed the best with a tied 95 % average accuracy. The feature combination of Max and First Derivative proved to be only slightly better that the First Derivative on its own, which performed the worst using Logistic Regression as a classifier, while the Max/First Derivative combination performed worst using Multi-Layer Perceptrons. Otherwise, both feature types achieved a 100 % classification accuracy with SVM and Random Forest models.

4. Conclusions

The results of the pilot experiments presented here highlight the promise of olfactory-based smart sensors in pathogen identification for bacteria of the same genus. Further studies, including extensive data collection, assay development and sensor augmentation are warranted to validate this device as a precision antibiotic matching tool.

References

[1]. A. Flores-Mireles et al., Urinary tract infections: epidemiology, mechanisms of infection and treatment options, *Nature Reviews Microbiology,* 13, 5, 2015, pp. 269-284.

[2]. J. E. Simmering et al., The increase in hospitalizations for urinary tract infections and the associated costs in the United States, 1998–2011, *Open Forum Infectious Diseases,* Vol. 4, No. 1, 2017.

[3]. B. Foxman, Epidemiology of urinary tract infections: incidence, morbidity, and economic costs, *The American journal of Medicine,* 113, 1, 2002, pp. 5-13.

[4]. A. Wilson and B. Manuela, Applications and advances in electronic-nose technologies, *Sensors,* 9, 7, 2009, pp. 5099-5148.

[5]. A. Amann and S. David eds., Volatile biomarkers: non-invasive diagnosis in physiology and medicine, *Elsevier Science Ltd*, 2013.

[6]. R. B. Caldwell and T. L. Claus, Simple and reliable method to precipitate proteins from bacterial culture supernatant, *Appl. Environ. Microbiol.,* 70. 1, 2004, pp. 610-612.

[7]. A. Schutze, T. Baur, M. Leidinger, W. Reimringer, R. Jung, T. Conrad, T. Sauerwald, Highly sensitive and selective VOC sensor systems based on semiconductor gas sensors: how to?, *Environments,* 4, 20, 2017.

[8]. F. Pedregosa, G. Varoquaux, et al., Scikit-learn: Machine learning in Python, *Journal of Machine Learning Research*, 2011, 12, Oct, pp. 2825-2830.

Published by International Frequency Sensor Association (IFSA) Publishing, S. L., 2019 (http://www.sensorsportal.com).