NOSE: A Novel Odor Sensing Engine for Ambient Monitoring of the Frying Cooking Method in Kitchen Environments

POOYA KHALOO, University of Massachusetts Amherst, USA
BRANDON OUBRE, University of Massachusetts Amherst, USA
JEREMY YANG, University of Massachusetts Amherst, USA
TAUHIDUR RAHMAN, University of Massachusetts Amherst, USA
SUNGHOON IVAN LEE, University of Massachusetts Amherst, USA

How we cook and prepare our food has an enormous impact on our health and well-being. Specific cooking methods, like deep-frying, are linked to obesity and the degradation of food nutrients, which contribute to various diseases and health issues. We present *NOSE*, a Novel Odor Sensing Engine, that passively and continuously monitors gas emissions in the kitchen area using an array of six metal oxide semiconductor (MOS) gas sensors and detects the occurrence of deep-frying. To evaluate *NOSE*, we collected sensor data from five foods (chicken, fish, beef, potato, and onion) cooked with three methods (deep-frying, grilling, and boiling) and three common frying oils (canola, corn, and soybean) in three different kitchens in a controlled manner. We demonstrate that *NOSE* can classify cooking by deep-frying with an average F_1-score of 0.89. Based on the in-laboratory findings, we deployed *NOSE* in two different real-world households throughout a three-week period and successfully detected the occurrence of frying cooking with an average F_1-score of 0.72, which is a promising result considering the relatively small number of data samples collected. Our results show the potential of using *NOSE* as an assistive dietary monitoring tool that periodically reports to users about their cooking habits.

CCS Concepts: • **Human-centered computing → Ubiquitous and mobile computing systems and tools**; • **Applied computing → Health care information systems**; • **Networks**;

Additional Key Words and Phrases: electronic nose, gas sensors, frying detection

ACM Reference Format:
Pooya Khaloo, Brandon Oubre, Jeremy Yang, Tauhidur Rahman, and Sunghoon Ivan Lee. 2019. NOSE: A Novel Odor Sensing Engine for Ambient Monitoring of the Frying Cooking Method in Kitchen Environments. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 3, 2, Article 49 (June 2019), 25 pages. https://doi.org/10.1145/3328920

1 INTRODUCTION

Obesity is a major health problem in developed countries that increases the risk of diabetes, high blood pressure, and dyslipidemia [56]. In 2016, more than 1.9 billion adults aged 18 years and older were overweight and, of them,

Authors' addresses: Pooya Khaloo, College of Information and Computer Sciences, University of Massachusetts Amherst, 140 Governors Dr. Amherst, MA, 01003, USA, pkhaloo@cs.umass.edu; Brandon Oubre, College of Information and Computer Sciences, University of Massachusetts Amherst, 140 Governors Dr. Amherst, MA, 01003, USA, boubre@cs.umass.edu; Jeremy Yang, College of Information and Computer Sciences, University of Massachusetts Amherst, 140 Governors Dr. Amherst, MA, 01003, USA, jeremyyang@umass.edu; Tauhidur Rahman, College of Information and Computer Sciences, University of Massachusetts Amherst, 140 Governors Dr. Amherst, MA, 01003, USA, trahman@cs.umass.edu; Sunghoon Ivan Lee, College of Information and Computer Sciences, University of Massachusetts Amherst, 140 Governors Dr. Amherst, MA, 01003, USA, silee@cs.umass.edu.

2474-9567/2019/6-ART49 $15.00
https://doi.org/10.1145/3328920

over 650 million adults were obese [6]. Fried foods (mostly fast foods) alongside sugary drinks, snacks, portion sizes, and low activity levels are the major causes of obesity [75]. Hence, quantifying diet (i.e., energy intake) and physical activity level (i.e., energy consumption) are both necessary to assist individuals wanting to track progress toward their goal of obesity management and healthier living [43]. Various systems for monitoring energy consumption (by assessing the level of physical activity) have been successful in the market – such as Fitbit [3] and the Apple Watch [1]. However, evaluating and quantifying diet using ubiquitous and wearable sensors remains a challenge that has not yet been comprehensively addressed [46]. Various approaches have been proposed to allow users to track their energy intake and/or eating behavior including manual self-reporting, wearable and hand-held device-based semi-manual reporting, and ambient sensing platforms (see Section 2 for details). Despite these advances in eating activity and food recognition, the characterization of the energy intake (and of cooking methods in particular) remains largely underexplored.

Previous studies have shown that people generally believe home-cooked foods are healthier and may slow down or even prevent obesity [18, 71]. However, not all cooking methods used in the home are necessarily healthy. The frequent consumption of fried foods, even in home settings, can increase caloric intake through fat absorption and lead to serious obesity problems [32, 68]. Furthermore, frying with oil at high temperatures for a long period of time causes the formation and release of food toxins (e.g., aldehydes), which are known to increase the risk of cancer [37]. Therefore, a ubiquitous technology that can longitudinally monitor cooking behaviors (especially the frequent use of deep-frying) and inform us about the effects of these methods on our health could dramatically improve our dietary behavior. Previous studies have attempted to detect cooking activities through a variety of different sensing modalities – including cameras, accelerometers, microphones, and instrumented dining tables – but either require substantial user engagement, provide limited information regarding food processing or nutritional factors, or have notable privacy concerns. (Section 2 provides a more complete overview.)

In this work, we explore an entirely different sensing modality: the odors released from cooking foods. We present a Novel Odor Sensing Engine (*NOSE*) for recognizing cooking methods, especially cooking by deep-frying, in the kitchen environment. The *NOSE* system was implemented with six metal oxide semiconductor (MOS) gas sensors within a 3D-printed enclosure. These gas sensors detect the concentration of specific chemical compounds – called analytes – in the air. *NOSE* is designed to continuously and passively sample data in the kitchen area to detect these analytes – including toluene, alcohol, butane, hydrogen, and aldehydes – which are then used to identify the unique gaseous fingerprints (i.e., the "smell" or "odor") of frying and other cooking methods (e.g., grilling and boiling) that are common in different cuisines.

We evaluated *NOSE* based on 1) controlled, in-laboratory experiments in three different kitchen environments and 2) uncontrolled, in-the-wild experiments in two different real-world households. For the in-laboratory experiments, we collected sensor data from five foods (chicken, fish, beef, potato, and onion) cooked with three common methods (frying, grilling, and boiling), and three common frying oils (canola, corn, and soybean) in three kitchen environments. Each cooking event was controlled to be completely isolated from one another. The collected data were used to investigate the opportunities and limitations of this novel sensing modality and to construct a classification model for distinguishing deep-frying from other cooking methods. Afterward, *NOSE* was deployed in two different real-world households to monitor the occurrences of cooking by frying over a three-week period in an uncontrolled manner. We demonstrate that *NOSE* can accurately detect different cooking methods, and can detect deep-frying in particular. Furthermore, the classification of cooking method can be made independently of the type of foods, the type of cooking oils, and the use of ventilation fans. The main contributions of this paper include:

(1) We design and develop a novel odor sensing engine, *NOSE*, composed of an array of six MOS gas sensors that can continuously and passively monitor common cooking methods, particularly deep-frying. This

Table 1. Summary of related literature proposing different dietary monitoring approaches, including where *NOSE* fits into the bigger picture.

Approach	Sensor	Study	Information
Manual	-	Caroll *et al.*[19], Fitbit [3], MyFitnessPal [5]	-
Wearable & Hand-held	Camera	Arab *et al.*[10], Reddy *et al.*[70], Gemming *et al.*[35] Liu *et a.*[55], O'Loughlin *et al.*[62]	Photos of Food
	Acoustic	Thomaz *et al.*[80], Bi *et al.*[15], Olubanjo and Ghovanloo[63] Rahman *et al.*[69], Nishimura and Kuroda[59], Amft[7]	Chew and Swallow
	Motion	Farooq *et al.*[30], Bedri *et al.*[12, 13], Kalantarian *et al.*[44, 45] Dong *et al.*[27, 28], Amft and Tröster[8]	Wrist, Throat, Jaw and Ear Canal Motion
Ambient Sensing	Camera	Gao *et al.*[33], Cadavid *et al.*[17], Wu amd Yang[83]	Photos of Food, Eating Action
	Scaling Table	Kissileff *et al.*[47], Change *et al.*[20]	Food Weight
	Scaling Table + Camera	Chi *et al.*[21]	Photos of Food, Food Weight
	Scaling Table + Force Sensor	Zhou *et al.*[85]	Eating Action, Food Weight
	Gas Sensor	*NOSE*	Food Aroma

low-maintenance and low-cost platform can instantly turn any kitchen into smart and connected cooking environment with minimal instrumentation.

(2) We discuss the fundamental technical limitations and opportunities associated with an odor-based sensing modality for monitoring cooking activities, and introduce our contributions to address some of these limitations.

(3) We introduce a data analytics pipeline for data pre-processing, feature engineering, and machine learning modeling, which was used to detect the occurrence of cooking by deep-frying. We also investigate the use of an incremental retraining approach to adapt a previously trained machine learning model to new kitchen environments in order to enable location-agnostic deep-frying classification.

(4) We evaluate and benchmark *NOSE* in a set of controlled experiments involving five foods (chicken, fish, beef, potato, and onion) with three cooking methods (frying, grilling, and boiling) and three common frying oils (canola, corn, and soybean) in three kitchens. The resulting data set consists of a total of 207 cooking sessions (approximately 35 hours) of sensor data.

(5) We also deploy and evaluate *NOSE* in two different real-world households over three-week periods. We demonstrate that our system can reliably monitor and detect deep-frying in in-the-wild settings.

2 RELATED WORK

Table 1 provides a comparative summary of related works in the areas of mobile and ubiquitous sensing technologies for eating behavior and food monitoring.

2.1 Eating Behavior Monitoring

Traditionally, food logging has been done manually with either paper- or mobile-device-based diaries [19] or spoken records of diet [51, 54]. Currently, a large number of wearable fitness trackers, such as Fitbit [3] and smartphone applications (e.g. MyFitnessPal [5], LoseIt [4]), provide digital platforms for users to log consumed foods. However, these approaches demand a relatively high level of engagement and effort from users and suffer from low compliance rates for long-term monitoring [11, 23].

2.1.1 Wearable & Hand-held Systems. There is a body of work available in the literature focusing on automating or semi-automating food logging and eating tracking based on various types of on/off-body sensors [46, 67]. Among the wearable and hand-held devices, cameras are one of the most widely studied sensors for dietary monitoring. Users were asked to take photos of the food they consumed using wearable cameras or a smartphone's built-in cameras. These photos were then annotated based on either user recall [10, 35, 62] or crowdsourcing frameworks [60, 79] to obtain labels for food names and caloric information. Some studies attempted to automate this process by replacing human resources with machine-learned image classifiers [49, 86]. These methods could provide reasonably accurate information regarding participants' eating behaviors (e.g., what food they consumed or related caloric information), however, like diary-based approaches, they require a substantial level of user engagement for long-term monitoring.

Wearable microphones are another widely studied wearable sensor for dietary behavior monitoring. These sensors are often placed in the ear or attached to the neck to record the sound of chewing or swallowing [7, 14, 15, 59, 63, 69, 80]. These sounds are processed to recognize different eating events (e.g., drinking or eating) and the texture of consumed foods (e.g., soft vs. hard food). Motion sensors, such as accelerometers and gyroscopes, have also been attached to different body locations – including the jaw [30], neck [8, 44, 45], and ear canal [12, 13] – to recognize different eating events (e.g., chewing and swallowing). Motion sensors on the wrist have also been extensively studied to detect eating motions [27, 28, 65]. The aforementioned wearable sensors can capture important information regarding one's eating behavior, such as the number of food-to-mouth movements, eating frequency, eating duration, and an estimation of food texture. However, these sensors provide relatively limited information regarding how the foods were processed and the associated health-related issues (e.g., whether users consume healthy or non-healthy foods).

2.1.2 Ambient Systems. Sensors embedded in the environment aim to seamlessly monitor one's dietary behavior without requiring much effort or engagement from users. Kissileff *et al.* introduced a table embedded with a weight scale to quantify the amount and rate of food consumed [47]. A number of studies extended the idea of the scale table by fusing it with other types of sensors. Chang *et al.* combined a scale table with Radio Frequency Identification (RFID) tags attached to food containers in order to recognize the placement of specific foods on the table and monitor the amount of food consumption [20]. Although interesting, this system requires users to manually import the names of foods in each container, which may be cumbersome for long-term use. Zhou *et al.* leveraged resistive force distribution sensors to detect different eating actions (e.g., stirring, scooping, cutting, and poking) and used the information to infer the types of food that were consumed [85]. Unfortunately, this approach provides limited information regarding how the foods were cooked and their nutritional factors. Surveillance cameras embedded in the dining area have also been used to recognize various eating activities (e.g., the start and end of a meal) [17, 33] or the foods that were consumed during mealtime [83]. However, cameras in the ambient setting introduce inevitable privacy issues.

2.2 Gas Sensors for Food Monitoring

The use of gas sensors (also referred to as an electronic nose or e-nose), which are sensitive to specific target analytes and attempt to replicate the human olfactory system by detecting various types of odors, was first introduced in the early 1980s [66]. Since then, gas sensors have been applied to a variety of scientific research fields including agricultural, bio-medicine, environmental, and food sciences [82]. Specifically in studies related to food, MOS gas sensors have been used to classify food quality (e.g., whether food is rotten or moldy) for meat [36], dairy products [9], bread [29], and coffee [16]. The reader is directed to [34, 57, 72, 76, 77] for in-depth reviews.

Although MOS gas sensors have been applied to investigate different characteristics of foods, related studies in the field were conducted in highly controlled environments (e.g., placing food and gas sensors together in a

small chamber or controlling gases, such as oxygen, to accurately calibrate the sensors). In the area of ubiquitous computing, there have been some studies that explored the use of gas sensors in more naturalistic environments. Kobayashi *et al.* developed a wearable system with gas sensors to detect various activities (e.g having a meal) [48]. The same research group extended the work to employ a camera to capture pictures of the consumed foods. Hirano *et al.* investigated the use of gas sensors to detect fourteen different types of beverages (e.g hot water, milk, soda, orange juice, coffee, champagne, etc.) in an open space [42]. The same research group also investigated classifying the degree of cooking doneness (i.e., undercooked, cooked, overcooked) in two different foods (i.e., waffles and popcorn) [41]. Although MOS gas sensor-based systems have not yet been used to detect different cooking methods, our work is inspired by these previous studies by Hirano *et al.* that employed signal processing and machine learning techniques to extract meaningful information from the time-series data and investigate different real-world applications.

Our work adds technical contributions to the existing literature by 1) discussing technical limitations and opportunities for the real-world deployment of MOS gas sensor-based systems, 2) discussing the design of machine learning algorithms to address the existing limitations of MOS gas sensors that affect classification accuracy when the sensors are deployed in new kitchen environments; measurements of the target substance concentration are affected by location-dependent factors including the physical size of the kitchen, the existence and capacity of the ventilation fan, and the ambient gas profile in the environment, and 3) the application of gas sensing to detect different cooking methods (specifically, deep-frying) in uncontrolled, free-living settings.

3 DESIGN CONSIDERATIONS

This work hypothesizes that *NOSE* can capture an odor signature that is uniquely associated with frying. In this section, we elaborate on the specific design principles that were considered during the development of *NOSE*.

Low-Cost, Ambient Sensing: The major principle of *NOSE* is that it must minimize the efforts and involvement of users during the cooking monitoring process. As discussed in Section 2, systems that demand consistent user engagement during long-term monitoring suffer from poor compliance. Thus, we focused on implementing a system that can unobtrusively and seamlessly monitor the way that foods are cooked in the kitchen with minimal user involvement. Furthermore, we envisioned a sensing system that is low cost, small in size, and lightweight, such that it can be conveniently installed in conventional kitchen environments.

User Acceptability for Potential Privacy Infringement : Previous studies have investigated the user acceptance of various types of smart-home technologies and reported that users expressed specific concerns regarding the use of cameras compared to other technologies due to their high-dimensional data and the negative impacts of potential privacy violations [24, 25]. *NOSE* captures relatively low-dimensional (six dimensional) non-vision (i.e., odor) information, which we believe poses fewer hurdles for user acceptability regarding potential privacy threats.

Responsive and Reliable Sensing Performance: Sensors need to quickly respond to the presence of or changes in the target analyte concentration with a minimum response time of approximately $5 - 10$ minutes, such that the detection of different cooking methods can be made in a timely manner considering that the average cooking time per day in American families varies between 20 and 30 minutes [39, 61, 81]. Moreover, the sensors needed to provide measurable responses of the target analytes at relatively low concentrations, as we envisioned that the system would be installed in a relatively large space (i.e., the kitchen), possibly equipped with a ventilation system. In this work, we used highly sensitive, fast responding MOS gas sensors to capture the odor signatures of different cooking methods [31, 82].

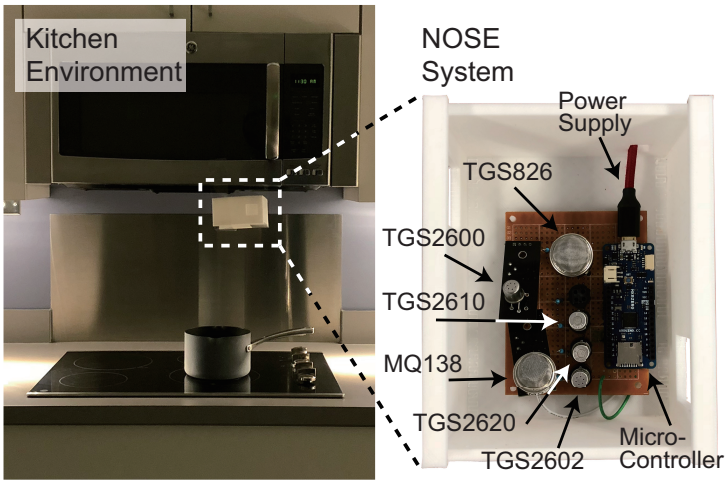

Fig. 1. Envisioned instrumentation of the proposed *NOSE* system in conventional kitchen environments. *NOSE* contains a microprocessor and an array of six gas sensors that are responsive to the different gaseous fingerprints (i.e., odors) associated with deep-frying as well as other common cooking methods such as grilling and boiling.

Table 2. The Metal Oxide Semiconductor (MOS) gas sensors incorporated in *NOSE* and their primary and secondary analytes.

Sensor	**Primary Analyte**	**Secondary Analytes**	**Manufacturer**
TGS2602	Toluene	Ammonia, Alcohol	Figaro sensor
TGS2620	Alcohol	Hydrogen, Iso-Butane	Figaro sensor
TGS2610	Butane	Alcohol, Hydrogen	Figaro sensor
TGS2600	Hydrogen	Alcohol, Carbon Monoxide	Figaro sensor
TGS826	Ammonia	Alcohol, Hydrogen	Figaro sensor
MQ138	Aldehydes	Acetone, Alcohol	Hanwei electronics

4 NOSE SYSTEM DESIGN AND IMPLEMENTATION

With the aforementioned inspiration and design considerations in mind, we developed *NOSE* based on an array of MOS gas sensors that are low-cost, lightweight, reliable, and responsive. Figure 1 shows how *NOSE* can be installed in conventional kitchen areas (left) and the prototype design of the system (right). Various gaseous substances are released into the air when foods are deep-fried, such as hydrocarbons, aldehydes, alcohols, 2-alkenals, ketones, acids, and toluene [22, 64, 73]. The proposed *NOSE* system incorporated MOS gas sensors that can provide insights regarding the presence of the aforementioned substances, which are summarized in Table 2. The fundamental hypothesis of this work is that we can detect when different types of food are deep-fried based on the gaseous fingerprints characterized by these sensors. It is noteworthy that the sensors in Table 2 capture not only the associated primary analyte but also some secondary analytes that are cross-covered by other sensors. Furthermore, all sensors were responsive to temperature and humidity, and one of the sensors (MQ138) was responsive to the water condensation around the sensing area. Hence, we did not include a dedicated temperature and humidity sensor in *NOSE* (see Section 7 for detailed discussions).

The resistances of the MOS gas sensors change depending on the concentrations of the target analytes in the air. Each resistive gas sensor was wired to a voltage divider, such that the output voltage can vary between 0

and 3.3 V. The obtained voltage readings can, in theory, be converted into a measure of concentration (i.e., parts per million (ppm)) using a manufacturer-defined polynomial function if each sensor is calibrated to a known concentration of the target analytes in a chamber (e.g., 50 ppm of ammonia for the TGS826 sensor). However, the measure of the concentration level may significantly vary in different kitchen environments based on a large number of environmental factors, such as the physical volume of the kitchen, the existence or capacity of the ventilation system, and the ambient gas profile of the surrounding environment. Thus, we argue (and demonstrate) in this work that data analytics models to detect a specific cooking method (e.g., deep-frying) need to be trained for each kitchen environment, which can be performed using the raw sensor readings and eliminates the need for the conversion of voltage to the concentration level. Therefore, the proposed *NOSE* system utilizes sensor responses that are directly read by the embedded system without conversion to ppm (see Section 5.2 for details). The sensor readings (i.e., the output voltage from the voltage dividers) were sampled using a low-power embedded system (Arduino MKR ZERO [2]) at 0.77 Hz, the maximum sampling rate that the system could support. The sampled analog signal was digitized by an embedded 10-bit Analog-To-Digital (ADC) converter. The captured data were stored in a local memory (i.e., SD card) for off-line analysis; we envision that future iterations of the system will be equipped with wireless communication modules to periodically or opportunistically push the captured data to a cloud server.

5 CONTROLLED, IN-LABORATORY EXPERIMENTS

The primary objective of *NOSE* is to detect and longitudinally monitor the occurrences of deep-frying during food preparation. We conducted a series of experiments in a controlled manner to both develop the cooking method classification algorithm and systematically understand the technical limitations and opportunities of our odor-based sensing modality. For example, can we identify unique odor signatures of different cooking methods when different foods are cooked (i.e., in a food-agnostic manner)? Are the odor signatures of different cooking methods location-dependent or location-agnostic? Can we robustly classify frying in the presence of different amounts or types of cooking oils? Does the presence of a ventilation fan affect the ability of our system to classify cooking methods?

5.1 Data Collection

The *in-laboratory* experiments were designed to collect the odor sensor data of three common cooking methods (i.e., deep-frying, grilling, and boiling) for five foods (i.e., chicken, fish, beef, potato, and onion) in three kitchen environments. In this work, we consider the use of at least 150 mL of oil to be deep-frying. The five foods considered in this work were foods that are commonly fried in the Western diet [84]. The three cooking methods were determined based on a study [40], which analyzed 1420 recipes from the Internet and identified that frying, grilling, and boiling are the most common cooking methods. More specifically, we considered grilling as a means to validate our hypothesis that odors associated with a specific food, which may be best evaporated by applying heat on a pan, do not affect the classification of whether the food was cooked by frying or not (i.e., food-agnostic classification of frying).

Before each cooking trial, the cooking apparatus (e.g., a pan and spatula) was cleaned and the sensor was exposed to the baseline air for 15 minutes to stabilize the readings to the room's ambient gas profile. Approximately 200 g of foods were prepared for each trial. Foods were cooked using one of the three cooking methods for approximately 650 seconds (≈11 minutes) with the kitchen ventilation turned off; we later validate for the effects of the ventilation fan. All three kitchens were equipped with electric stoves. When deep-frying the foods, 150 mL of canola oil was poured on the pan and heated for two minutes. When grilling the foods, the pan was heated for two minutes without oil. For boiling, foods were placed in a pot containing 500 mL of boiling water. Each trial started the moment when the food was placed on the pan or in the pot. The cooking activity was repeated

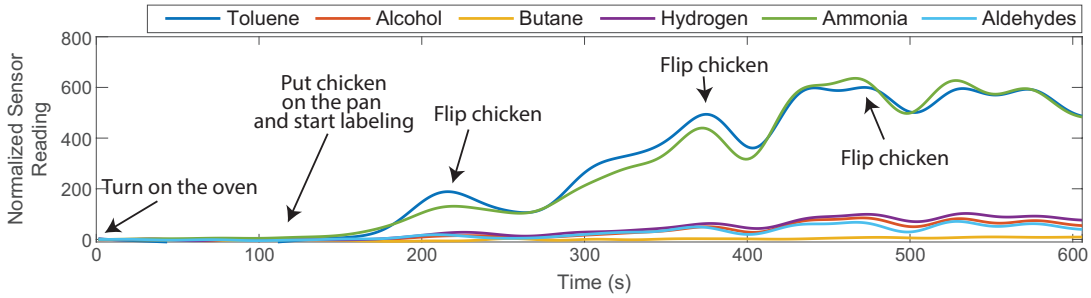

Fig. 2. Illustration of a data sample obtained from grilling a piece of chicken at Location 1. The annotations indicate different actions that occurred during the cooking process.

three times for each food and each cooking method, yielding a total of 45 trials per location: 5 foods × 3 cooking methods × 3 repetitions. We repeated the entire process at three different kitchens: a simulated apartment setting at the University of Massachusetts Amherst (Location 1), a single-family house (Location 2), and a unit of a multi-family apartment complex (Location 3). A total of 135 data samples (45 samples/location × 3 locations) consisting of 23 hours of sensor data were collected from the experiments.

We also conducted three additional experiments designed to evaluate the effects of 1) different cooking oils, 2) different frying methods (i.e., deep-frying vs. stir-frying or sautéing) that use a smaller amount of oil, and 3) the use of the kitchen ventilation fan. These additional experiments were conducted at Location 1, used only three foods (i.e., potato, beef, and chicken), and otherwise followed the same experimental procedure. In the first experiment, we deep-fried the three foods using approximately 150 *mL* of one of the two additional cooking oils: corn and soybean oils. The cooking activities were again repeated three times, generating 18 additional data samples: 2 oils × 3 foods × 3 repetitions. In the second experiment, we stir-fried the three foods using approximately 50 *mL* of canola oil and produced 9 data samples: 3 foods × 3 repetitions. In the third experiment, we cooked the three foods using the three cooking methods while turning on the kitchen ventilation fans. Each cooking event was repeated five times, generating 45 data samples: 3 cooking methods × 3 foods × 5 repetitions. The three additional experiments therefore produced 72 (18 + 9 + 45) data samples.

Overall, we collected a total of 207 (135 + 72) data samples of cooking activities, consisting of approximately 35 hours of cooking in controlled environments. A total of 117 samples were collected from Location 1, 45 samples from Location 2, and 45 samples from Location 3. Figure 2 shows a data sample obtained from grilling a piece of chicken. The time-series data in Figure 2 are annotated with different actions (e.g., flipping the chicken) that occurred during the cooking process, in which the MOS gas sensor data show clear patterns. Although interesting, the detection of different cooking-related actions was not investigated as it is out of the scope of this work.

5.2 Data Analysis

As highlighted earlier, *NOSE* employs MOS gas sensors that respond to the concentration of target analytes in the air. Figure 3 illustrates the data analytics pipeline that processed the sensor data from both the in-laboratory and in-the-wild scenarios to detect when foods are deep-fried during cooking. The proposed pipeline employed a data-driven approach to study the unique gaseous fingerprints of different cooking methods using the in-laboratory data and to validate the feasibility of classifying deep-frying. The following subsection discusses each building block of Figure 3 in detail.

5.2.1 Sensor Signal Pre-processing. This step ensures that the sensor signal remains highly sensitive to the gaseous fingerprints associated with different cooking activities while rejecting information about a kitchen's

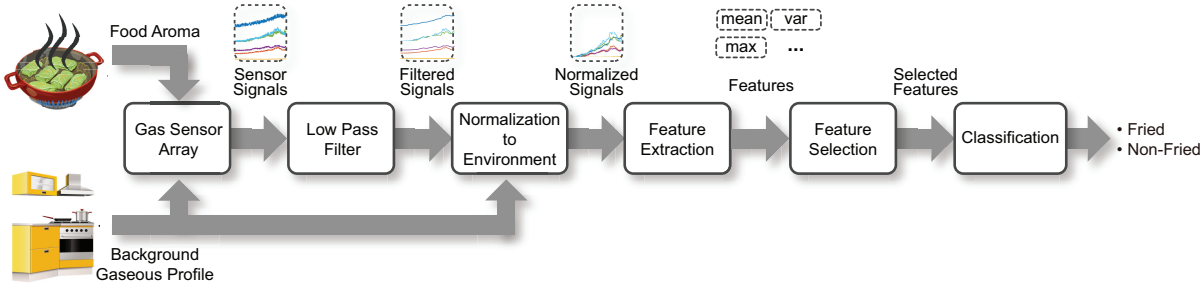

Fig. 3. The data-driven analytic pipeline of *NOSE* that can detect cooking methods (i.e., deep-frying, grilling, and boiling) based on gaseous substances released during cooking.

ambient gas profile. The ambient gas profile was estimated by sampling all sensor data for approximately thirteen seconds (i.e., ten data points) immediately prior to the start of cooking activities. The mean values of the ambient readings were subtracted from each sensor's cooking data in order to capture the *changes in the concentration* of the target analytes during cooking activities. We validate the importance of this normalization process, especially when classification models trained in one location are deployed to another, in Section 5.3.4. Figure 4 shows some sample data collected from trials where a piece of chicken was boiled (1st row), grilled (2nd row), and deep-fried (3rd row) in three different locations. The figures in the 4th row show the sensor data when normalized to each location's ambient gas profile for the deep-fried samples.

5.2.2 Feature Extraction. We employed a set of statistical functions to extract a rich pool of features that capture the temporal changes in the concentration level of the target analytes. Most features were extracted from two different time-series: the normalized sensor readings and the filtered normalized sensor readings that were obtained by applying a digital Butterworth low-pass filter with an empirically-chosen cut-off frequency of $0.02\ Hz$. The low-pass filtered readings were computed to remove any undesired fluctuations in the raw sensor readings caused by different cooking actions, such as flipping the foods (as shown in Figure 2). The statistical features extracted from each trial included the mean, median, harmonic mean, minimum, maximum, range (difference between the minimum and maximum), interquartile range, root mean square value, mean and median absolute deviation, signal entropy, skewness, and kurtosis. These features were similarly extracted from the rate of change of each times-series (both filtered and unfiltered) to capture the diffusion characteristics of gaseous substances. Since our MOS gas sensors cross-cover primary and secondary analytes, we also computed features that compare the similarities and differences between each pair of gas sensor time-series. Features were extracted from the differences of each pair as described above. The cross-correlations between each pair were computed to measure the similarity. This similarity can provide insights regarding the presence of a specific gas substance. For instance, a high correlation between the sensor data of TGS826 (whose primary analyte is ammonia) and TGS2602 (whose secondary analytes include ammonia) can reveal information specific to ammonia.

A few amplitude-related features – such as the mean, median, and harmonic mean of the sensor readings of the ambient gas profile (i.e., the baseline sensor readings) – were extracted from the unnormalized, raw time-series to represent the fraction of gas molecules present in a particular kitchen environment. These features from the raw time-series are important when sufficient data samples are made available from the same location such that a location-specific classification model can be constructed for the accurate classification of deep-frying (see Section 5.3.5 for results).

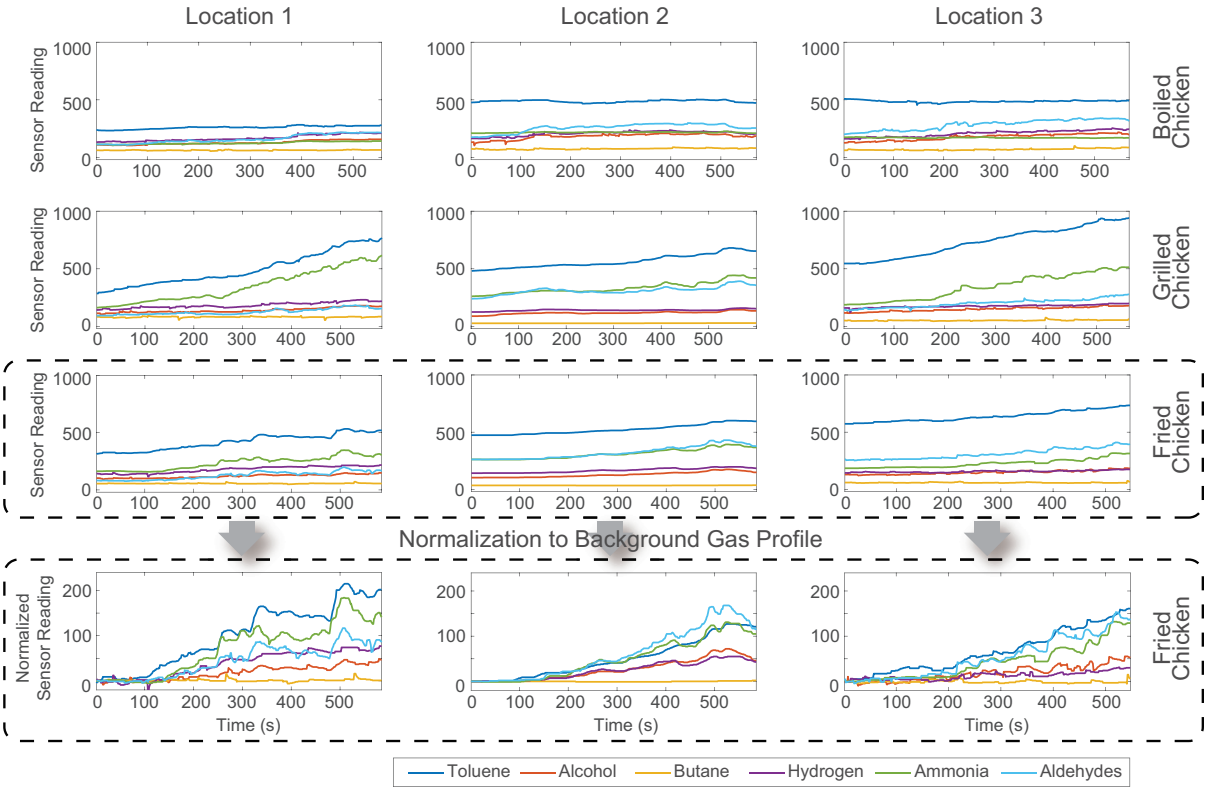

Fig. 4. Low-pass filtered sensor data when chicken was prepared by boiling (1st row), grilling (2nd row), and deep-frying (3rd row) in three different locations. The figures in the 4th row show sensor data for deep-fried samples when they were normalized to the ambient gas profile.

5.2.3 Classification of Cooking Methods for In-Laboratory Data. In this work, we consider two different classification tasks: multi-class classification of the three cooking methods (deep-frying vs. grilling vs. boiling) and binary classification for classifying the occurrences of deep-frying (frying vs. non-frying). The former task is to investigate the unique gaseous fingerprints of different cooking methods and to examine the possibility of accurate classification. The latter task is to construct a data-driven model specifically to classify cooking by deep-frying. The binary classification results were derived from the results of the multi-class classification by combining the boiling and grilling labels into a single non-fried label. We also investigated training an independent binary classification model for fried vs. non-fried, but the proposed approach provided better performance, as will be discussed in Section 5.3.1.

We employed leave-one-food-out as well as leave-one-location-out cross validations to evaluate the classification performance of each model. The former ensures that data from the same food, cooking method, and oil type are not included in both the training and testing sets in order to avoid overfitting and reporting of optimistic classification results. The latter ensures that data from the same location are not included in both the training and testing sets. These cross-validation methods generalize the classification model to a new food and cooking combination, and a new location, respectively. Specifically, leave-one-location-out cross validation examines the transferability of a classification model trained in one kitchen environment to recognizing cooking methods in

another kitchen. However, due to the inherent technical limitations of the MOS sensors in the proposed *NOSE* system – i.e., the detected amounts of gaseous substances in the air may substantially vary depending on the environmental factors (see Section 4 and Section 5.2.4) – we anticipated that the classification performance for the leave-one-location-out cross validation would be inferior compared to leave-one-food-out cross validation. The F_1 score and average classification accuracy were used as the primary metrics of model performance.

A Correlation-based Feature Selection (CFS) algorithm [38] was employed to select a subset of important features that are relevant to the three cooking methods. The CFS algorithm identifies important features based on maximizing individual features' correlation to the label (predictability) and minimizing the correlations among the selected features (redundancy) [38]. Our work employed the best-first search strategy to construct the feature search space. Feature selection was performed independently within each cross-validation fold, involving only the training set. We then constructed models using six different machine learning algorithms capable of multi-class classification: Random Forest, C4.5 (J48) Decision Tree, Support Vector Machine (SVM) with a Pearson Universal Kernel (PUK), SVM with a 3rd order Polynomial Kernel, Multi-Layer Perceptron (Neural Network) , and Logistic Regression. The multi-class SVM used in this work employed the one vs. one ensemble approach.

5.2.4 Location-Dependent Classification of Frying. The concentration level, a measure of gas volume in the air, depends not only on the gaseous analytes released from cooking activities but also on environmental factors. Therefore, it is extremely challenging to develop a machine learning-based algorithm that can classify a specific cooking method (e.g., deep-frying) in a location-agnostic way (we verify this in Section 5.3.5). We argue that machine learning models that classify a specific cooking method need to be tailored to each kitchen environment. In order to address this technical problem, we drew inspiration from a branch of machine learning, namely online (incremental) learning, that can adopt or update a model constructed on data obtained from different kitchen environments for use on newly distributed data. However, our study involves a relatively small number of data and thus, 1) meaningful adaptation of classification models based on location-specific data is not feasible and 2) complete retraining of classification models can be performed fairly easily. Hence, we instead resembled the idea of online learning by initially constructing a model from the two other kitchens and incrementally re-training the model while gradually adding sensor data from the testing kitchen. The data from the testing location was divided into learning (adaptation) set and the testing set in a leave-one-food-out cross validation manner. We gradually increased the size of the learning set from 0% to 95% of the testing location's data in increments of 5%. The learning sets were chosen in a random fashion. The classification model was retrained based on the training and learning sets and evaluated on the testing set. The evaluation of the classification performance for each size of the learning set was repeated 50 times to minimize the effects of the randomization of the learning set. The classification performance of the incremental learning approach was compared to the leave-one-location-out cross validation approach (i.e., equivalent to a 0% learning set). This approach helps us to understand the expected performance of *NOSE* when the classification model is gradually tuned towards a specific kitchen, which can be especially important for the deployment of the system in in-the-wild settings.

5.3 Evaluation

5.3.1 Classification of Cooking Methods. Table 3 summarizes the classification results for recognizing cooking by deep-frying based on the data obtained from the in-laboratory experiments, evaluated using the leave-one-food-out cross validation technique. The results show that the Polynomial Kernel SVM provides the most accurate performance with an average classification accuracy of 0.89 and an average F_1 score of 0.89. The results were further divided into three locations. The F_1 scores for Locations 1, 2, and 3 were 0.92, 0.85, and 0.72, respectively. We believe that the classification performance for Location 1 was substantially greater than other two locations due to the larger data size (i.e., 117 samples from Location 1 vs. 45 samples from Locations 2 and 3). This indirectly demonstrates that we may be able to further improve the F_1 performance if sufficient data are made available. We

Table 3. The classification results (accuracy and F_1 scores) for classifying cooking by frying based on six different algorithms, evaluated using the leave-one-food-out cross validation technique.

Classifier	**All Data**		**Location 1**		**Location 2**		**Location 3**	
	Acc	F_1	Acc	F_1	Acc	F_1	Acc	F_1
SVM with Polynomial Kernel	0.89	0.89	0.95	0.92	0.87	0.85	0.80	0.72
SVM with PUK kernel	0.87	0.87	0.95	0.95	0.84	0.84	0.80	0.78
Neural Network	0.87	0.87	0.92	0.87	0.80	0.73	0.80	0.72
Random Forest	0.85	0.84	0.91	0.91	0.78	0.77	0.76	0.74
Logistic Regression	0.82	0.82	0.86	0.86	0.80	0.79	0.73	0.70
J48 Decision Tree	0.73	0.72	0.78	0.78	0.67	0.65	0.67	0.62

believe the classification performance for Location 3 was lower than Location 2, despite the same number of data samples, due to higher ambient gaseous noise level at Location 3, which was a unit of a multi-family apartment complex. However, more rigorous investigation is necessary to confirm this.

Figure 5a shows the confusion matrix when a Polynomial Kernel SVM was used. Note that, as discussed in Section 5.2.3, the binary classification results were obtained from the multi-class classifier results by combing the grilled and boiled labels into a single non-fried label. The corresponding multi-class classification results are summarized in Figure 5b, where the average classification accuracy was 0.86 and the average F_1 score was 0.81. Collapsing the ternary classification results into binary labels provided better performance compared to training an independent binary classifier, where we obtained an average F_1 score of 0.83. We believe that converting the ternary labels to binary labels performs better because the multi-class training forces to define the boundary between fried and grilled foods more precisely, which is the main source of misclassifications.

Figure 6 shows the results of Principal Component Analysis on the top twelve features shown in Table 4. We can clearly observe the three data clusters of different cooking methods, as well as the difference between fried and non-fried foods. More specifically, Figure 6 shows that the fried and grilled data partially overlap, which agrees with the large number of misclassifications between the two data sets, as shown in Figure 5b. The grilled beef and salmon samples dominated the false positives, with four grilled beef and two grilled salmon samples classified as fried out of 10 false positives (i.e., 60%). We believe this is because beef and salmon contain a large amount of lipids. The fat in the meat breaks down when heat is applied – a process known as lipid degradation – which releases volatile organic compounds, such as aldehydes, into the air [58]. These compounds are also evaporated from cooking oils during frying [22], generating similar gaseous fingerprints and thus confusing the classification model. Three out of the five false negative fried samples that were misclassified as grilled were stir-fried samples. A detailed discussion regarding stir-fried vs. deep-fried samples is provided in the following paragraph. In sum, our results demonstrate the validity of the proposed *NOSE* system for identifying the gaseous fingerprints of common cooking methods, and more specifically for classifying occurrences of deep-frying.

Figure 7 describes the binary classification results (frying vs. non-frying) in terms of different foods, types of oil, amounts of oil (equivalently, stir- vs. deep-frying), and weather or not the ventilation fan was used during the cooking process. Figure 7a summarizes the classification performance, categorized into the five foods used for the in-laboratory experiments. We achieved the F_1 scores of 0.94, 0.88, 0.90, 0.92, and 0.85 for chicken, salmon, beef, potato, and onion, respectively. Classification for salmon, beef, potato, and onion performed similarly to the overall classification performance. Results for chicken were better than the overall performance because chicken contains a relatively small amount of fat – meaning less lipid degradation occurs – as well as a small amount of water for vaporization during frying, both of which may affect the misclassificaiton rate. Figure 7b compares

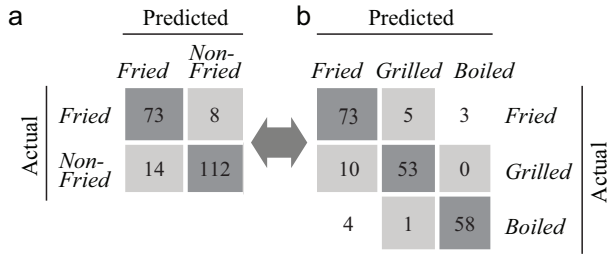

Fig. 5. The confusion matrix for a) classifying the occurrences of cooking by frying (frying vs. non-frying) and b) multi-class classification model for three different cooking methods (deep-frying vs. grilling vs. boiling). An SVM with a Polynomial Kernel was used to derive this result.

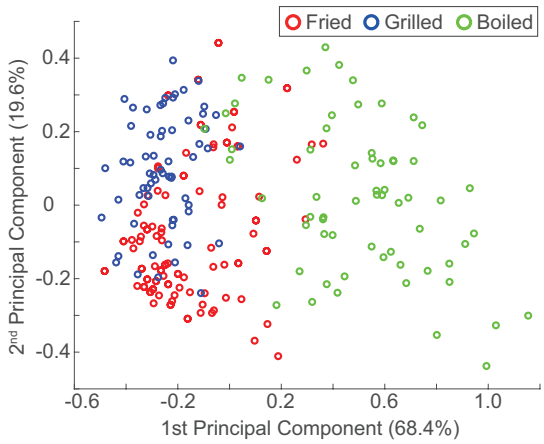

Fig. 6. A scatter plot of the cooking activity data obtained by Principal Component Analysis. The number in the axis label represents the percentage of explained variance for the associated principal component. The figure clearly shows clusters for different cooking methods.

the classification performance between using or not using the ventilation fan. The F_1 score was 0.92 with the ventilation fan and 0.96 without the fan. Although the impact on the classification performance is minimal, we believe that the use of the ventilation fan increased the rate at which the target analytes were eliminated from the kitchen area resulting in slightly inferior classification performance. Figure 7c summarizes the true positive rate for classifying frying when different cooking oils (canola vs. corn vs. soybean oils) were used. Results show that the classification of deep-frying cooking was made consistently with true positive rates of 0.94, 1.0, and 1.0 for canola, corn, and soybean oils, respectively. Note that only chicken, beef, and potato were included for the canola oil data to make the results comparable to the corn and soybean oil data (see Section 5). This result is not surprising since vegetable cooking oils generate similar volatile organic compounds, such as aldehydes, hydrocarbons, and alcohol [53]. Note that the number of data samples from canola oils (i.e., 81 samples) was substantially larger than corn and soybean data (i.e., 9 samples each). Therefore, the true positive rate of 0.94 for canola oil should better represent the generalized classification rate for deep-frying. Figure 7d summarizes the true positive rate for classifying two frying methods that use different amounts of cooking oils: deep-frying with 150 mL of oil vs. stir-frying with 50 mL of oil. The true positive rates were 1.0 and 0.67 for deep-frying and

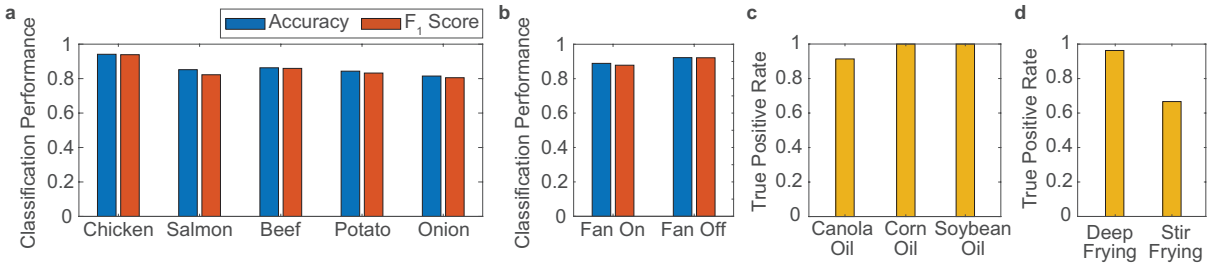

Fig. 7. Comparison of the fried food classification performance for a) different foods, b) the use of kitchen ventilation fan, b) different cooking oils, and b) different amounts of oils.

Table 4. List of the twelve most important features that were selected in every cross validation fold.

Number	Description of Features	Time-Series Type	Sensors Involved
#1	Arithmetic mean of (TGS2600 - TGS826)	Raw Time-Series	TGS2600, TGS826
#2	Maximum of the difference between TGS2600 and TGS826	Normalized	TGS2600, TGS826
#3	Arithmetic mean of (TGS2620 - TGS826)	Raw Time-Series	TGS2620, TGS826
#4	Std. Dev. of (TGS2620 - TGS826)	Normalized	TGS2620, TGS826
#5	Arithmetic mean of (TGS2620 - TGS2600)	Raw Time-Series	TGS2620, TGS2600
#6	Median absolute difference of (TGS2620 - TGS2600)	Trend	TGS2620, TGS2600
#7	Value range of (MQ138 - TGS2602)	Normalized	MQ138, TGS2602
#8	Baseline Amplitude of MQ138	Raw Time-Series	MQ138
#9	Correlation between TGS2602 and TGS2620	Normalized	TGS2602, TGS2620
#10	Correlation between TGS2600 and TGS826	Normalized	TGS2600, TGS826
#11	Correlation between TGS2620 and TGS826	Normalized	TGS2620, TGS826
#12	Mean rate of change of TGS826	Normalized	TGS826

stir-frying, respectively, showing that the amount of oil used in the cooking process significantly affects the classification performance. All of the false negatives of stir-frying were misclassified as grilling, which is not surprising considering that the relatively small amount of oil used in the cooking process makes the gaseous fingerprints similar to grilling. However, it is noteworthy that stir-frying or sautéing has less of an impact on health than deep-frying and thus, these misclassifications do not affect our ultimate goal of monitoring the frequent use of deep-frying in real-world kitchen environments.

5.3.2 Feature Selection Validation. Since each iteration of the leave-one-food-out cross validation independently performs feature selection on different training datasets, the analysis of the important features has been done in a retrospective way. When the entire cross validation process was completed, we analyzed the rate (in percentage) at which features were selected across all of the cross validation folds. Table 4 lists the top twelve features, which were selected in every cross validation fold.

The most-selected features were derived from the sensors that we initially hypothesized would capture the major volatile organic compounds from different cooking activities (e.g., aldehydes, hydrocarbons, and alcohol for deep-frying and aldehydes and ammonia for grilling meats). For example, the mean and the maximum values of the difference between TGS2600 and TGS826 (Feature #1 and #2, respectively) capture the responses of carbon monoxide and ammonia. Similarly, the mean and standard deviation of the difference between TGS2620 and TGS826 (#3 and #4), as well as the mean and the median absolute difference of the difference between TGS2620 and TGS2600 (#5 and #6), focus on capturing the responses of ammonia. These compounds are actively presented in the

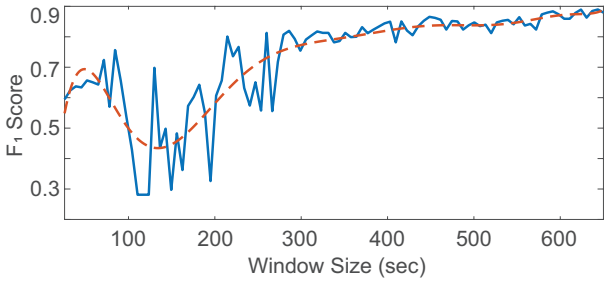

Fig. 8. Effect of the length of the time-series (i.e., window) on the classification performance measured by the F_1 score. The trend line was obtained using a polynomial fitting.

volatiles of heated meats [50, 58], which is highly relevant to grilling. The value range of the difference between MQ138 and TGS2602 (#7) and the baseline sensor amplitude of MQ138 (#8) together capture the concentration of the two major substances associated with frying (i.e., toluene and aldehydes [64, 73]). The correlation coefficients between TGS2602 and TGS2620 (#9), TGS2600 and TGS826 (#10), and TGS2620 and TGS826 (#11) all examine the responses of alcohol, which is cross-covered among all of the sensors and is relevant to deep-frying. Finally, the mean rate of change of TGS826 (#12) captures how fast the concentration level of ammonia and alcohol changes over time, which is also relevant to deep-frying. It is also noteworthy that all sensors were responsive to temperature and humidity, and thus, we believe that the aforementioned features also captured information related to boiling.

5.3.3 Optimal Length of the Window for Time-Series Analysis. We investigated the effects of the length of the time-series sensor data (from the moment when the cooking has started) on classification performance. Considering that the deployment of *NOSE* in in-the-wild settings will produce a continuous stream of sensor time-series, we expected to extract features from a fixed-length sliding window, perform classification on each window, and aggregate the decisions from each window to classify the occurrence of deep-frying. The length of the window determines the amount of data to be processed for feature extraction and thus influences the computational load and time required to support the classification of cooking activities, which is important for continuous and longitudinal monitoring in in-the-wild settings. The length of the window was adjusted from 20 to 650 seconds (incremented by approximately 10 seconds) from the beginning of the cooking activity. The F_1 score was recorded to evaluate the deep-frying classification performance in a leave-one-food-out cross validation manner. Figure 8 displays the effects of the window length on classification performance. The results show that the F_1 score stabilizes after approximately 300 seconds and then an increasing trend as the length increases. This is not surprising since the feature extraction process can capture richer information about sensor behavior when more data are made available for analysis. The maximum classification performance of the average F_1 of 0.89 reported in this work was obtained at the length of 640 seconds (i.e., 10 minutes and 40 seconds). That is, for in-laboratory experiments the window length consisted of almost the entire trial. Given that the average cooking duration per day in the United States is apprixmately 20 - 30 minutes (and even longer in other OECD countries) [61, 81], the selected window length is reasonable. Our data obtained from the in-the-wild experiments also demonstrates that this length is adequate (see Section 6 for details).

5.3.4 Effects of the Normalization Process. As discussed in Section 5.2.1, the sensor readings were normalized with respect to the ambient readings in order to minimize the effect of the ambient gas profile on the classification performance, which we hypothesized to be important when a model learned in one location is transferred (deployed) to data obtained from another location. Although we argue in this work that a classification model

Table 5. Comparison of classification results with/without the normalization process for the leave-one-food-out and leave-one-location-out cross validations.

	With Normalization		**Without Normalization**	
	Acc	F_1	Acc	F_1
Leave-One Food-Out	0.89	0.89	0.89	0.89
Leave-One Location-Out	0.63*	0.59*	0.59*	0.56*

* Statistically significant difference with a 99% confidence.

constructed from data obtained from the same kitchen environment is necessary due to location-dependent environmental factors, the normalization is necessary to maximize the transferrability of the learned knowledge from one location to another. Table 5 summarizes the effects of the normalization process on the classification performance. Classification was again performed using the Polynomial Kernel SVM. When using leave-one-food-out cross validation, the normalization process did not make a statistical difference in the accuracy at a 99% confidence level. This is because there already exist data from the same cooking method and location in the training data and thus the normalization process did not introduce any merits to the process. The statistical significance of the observed difference between two different classification models was performed based on the technique introduced in [78]. On the other hand, when the classification was performed in a leave-one-location-out cross validation manner (i.e., training a machine learning model on data obtained from two locations and testing on the left-out location), the difference was significant at a 99% confidence level with a classification accuracy improvement of 0.036 when the calibration was applied. A detailed discussion regarding the location-independent classification results is provided in the subsequent section. The results reported in this section show that the normalization process plays an important role when the learned knowledge is transferred from one location to another.

5.3.5 Location-Independent Classification. Figure 9 shows the classification results when the analysis was performed in a leave-one-location-out cross validation manner. The F_1 scores for Locations 1, 2, and 3 were 0.55, 0.73, and 0.66, respectively. The overall F_1 score was 0.62, as discussed in the previous section. Given that the number of data collected in Location 1 was sufficiently larger than Locations 2 and 3 (i.e., 117 vs. 45 samples), the classification model applied to Location 1 was trained on a substantially smaller amount of data samples, which resulted in poor classification performance. As hypothesized, the classification performance was inferior compared to the leave-one-food-out cross validation, mainly due to location-dependent environmental factors. This result supports our argument that we need classification models that are tailored towards each kitchen environment.

Figure 10 shows the changes in the classification performance when subsets of the data from the same location (i.e., an adaptation set) were combined with the data from other two locations to train the classification model in an attempt to replicate the concept of online learning, in which the adaptation data are made available to the model incrementally over time. The x-axis represents the number of the adaptation data samples from the sample location, and the y-axis represents the average F_1 score. The dotted lines represent the trend line obtained using a linear fit. Note that the number of data samples collected from Location 1 (i.e., 117) was larger than Locations 2 and 3 (i.e., 45 in each location), which is reflected in the length of the plot for Location 1. This disparity also

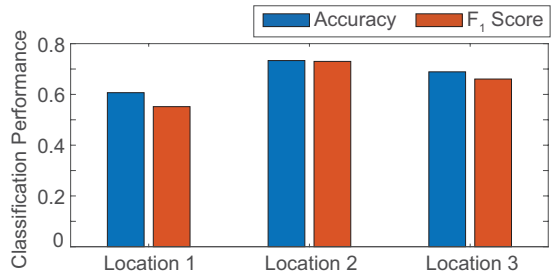

Fig. 9. The classification performance when the analysis was performed in a leave-one-location-out cross validation manner. The average F_1 score was 0.55, 0.73, and 0.66 for Locations 1, 2, and 3, respectively.

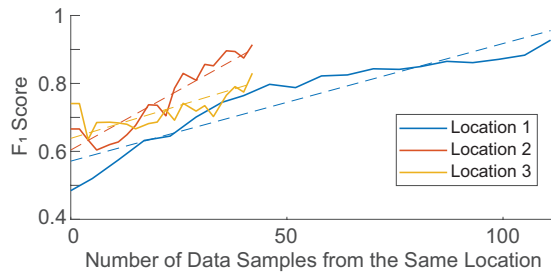

Fig. 10. Classification performance when different subsets of the data from the same location (i.e., adaptation set) were combined with the data from other two locations to train the classification model in an attempt to replicate the concept of online learning. The dotted lines represent the linear fit trend lines.

resulted in larger y-axis intercept values for Locations 2 and 3 – which are equivalent to the F_1 scores of the leave-one-location-out cross validation technique (i.e., empty adaptation set) – because a larger training dataset from Location 1 was used to train the initial model for Locations 2 and 3. However, the slopes of the trend lines of each location were similar to one another, which implies that the rate of improvement in the classification performance was consistent throughout the locations. In summary, these results support our argument regarding the need for classification models that are fine-tuned towards each kitchen environment, with more cooking data from the specific location being made available over time. This result supports a use case scenario in which the proposed *NOSE* system queries users to provide labels for recently cooked meals until an appropriate classification performance is guaranteed. Note that unlike the diary-based approaches discussed in Section 2 , users only need to provide feedback until an acceptable level of performance is reached and not indefinitely, which we believe will result in improved user compliance.

6 UNCONTROLLED, IN-THE-WILD EXPERIMENTS

Based on our in-depth understanding of the technical capacity of *NOSE* from the in-laboratory study and the development of effective machine learning algorithms for detecting the deep-frying cooking method, we deployed the system in two different real-world households and monitored the residents' cooking behaviors over three-week periods. This section focuses on the evaluation of the performance of *NOSE* for detecting deep-frying in in-the-wild settings.

Table 6. A summary of the data samples collected from the in-the-wild experiments.

Household	Number of Fried Samples	Number of Grilled Samples	Number of Boiled Samples	Total
#1	10	6	14	30
#2	9	0	11	20

6.1 Data Collection

The two new locations did not overlap with the kitchens used in the in-laboratory experiments in order to validate the generalizability of the proposed system. Household #1 was a single-family house and Household #2 was one side of a two-family duplex. Participants (i.e., residents of these two households) were instructed to attach *NOSE* to the range hood of the stove. The kitchens of both households were equipped with ventilation fans, but we did not provide any instructions for the use of the fans. The participants were asked to provide a log of all of their cooking activities including 1) the cooking method, 2) the foods cooked, 3) a rough estimate of the amount of oil used, and 4) the start and end times of the cooking activities. No further instruction was given that may alter their day-to-day cooking activities.

Table 6 summarizes the data collected from these free-living experiments. We were able to label all of the cooking activities as one of the three classes of the cooking methods (i.e., frying, grilling, and boiling). For example, preparing a soup or boiling water for instant noodles or pasta was classified as boiling. Any cooking activities that involved cooking oils were labeled as frying. A total of 30 cooking samples were collected from Household #1 (approximately 11.8 hours) and 20 samples from Household #2 (approximately 8.6 hours). The average cooking time per meal for Household #1 was 23.6 minutes with a minimum of 6 minutes and a maximum of 50 minutes. The average cooking time for Household #2 was 25.7 minutes with a minimum and maximum of 13 and 49 minutes, respectively.

6.2 Data Analysis

Unlike the in-laboratory experiments, where the start and end of a cooking activity were clearly defined, our system needs to detect the occurrence of cooking from a continuous stream of data for in-the-wild settings. To do this, we employed a two-layer classification approach. A high-level classifier with a one-minute window was trained based on the in-laboratory data (i.e., non-cooking, ambient monitoring data before the start of the in-laboratory trials vs. cooking data) to determine if a given data sample contained a cooking activity or not. Forty-five samples of non-cooking data were randomly chosen from each of the in-laboratory locations for a total of 135 non-cooking minutes. The identical feature engineering, feature selection, and classification procedures with Polynomial Kernel SVM that we discussed in Section 5.2 were used to detect activity in in-the-wild settings. The only difference was that the statistical features were extracted only from the low-pass filtered time-series to focus on the slowly-varying trend of the sensor responses. Nearby samples were grouped into a single cooking activity. The decisions were further low-pass filtered to eliminate any cooking activities with relatively short duration (e.g., cooking events lasted less than approximately three minutes).

For low-level cooking method classification, the data obtained from the in-laboratory experiments were used to train a base classification model. Again, the same analytic pipeline discussed in Section 5.2 was used with a Polynomial Kernel SVM. A window length of 640 seconds, which provided the best classification performance as we reported in Section 5.3.3, was used to extract features from the detected activity. For activities of length shorter than 640 seconds, we extracted features from the entire detected activity. The sliding window was incremented by approximately 10 seconds throughout the cooking period, producing an array of decisions made by the classification model. For example, if the cooking activity was 740 seconds, then ten classification decisions were

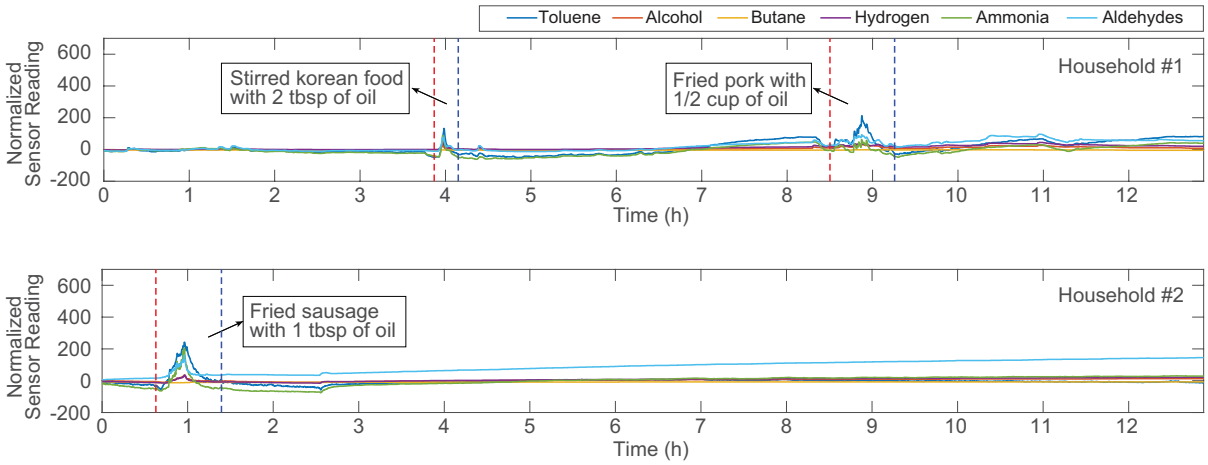

Fig. 11. Illustration of sampled data from the in-the-wild experiments. The dashed lines indicate the Start Time (St) and End Time (Et) of cooking activity. The labels show the corresponding activity as logged by the participant.

made using this approach. The decisions were aggregated per cooking activity by taking the label that most frequently appeared. Note that this sliding window mechanism generated a total of 4,248 windows (data samples) for Household #1, which was converted (aggregated) to cooking method decisions of 30 cooking sessions, and 3,084 windows for Household #2 that was converted to 20 cooking method decisions. Then the sensor data, along with the user-provided label of the actual cooking method, was added to the training set and the classification model was retrained for the subsequent data sample (as discussed in Section 5.3.5). This process was repeated for all samples collected in in-the-wild settings. We again used the average classification accuracy and F_1 score to evaluate the performance of our system.

6.3 Evaluation

The high-level classification algorithm for detecting the occurrences of cooking activities produced 100% true positives, but also produced some false positives – which were all associated with the use of an over-the-range microwave oven. This instance marks an interesting case of our system that we did not consider in our design phase. However, by adding more data related to these cooking activities, we believe we could reduce this type of false positive. In this section, we excluded such false positives from the presentation of the classification results since our model was not trained to detect the use of the microwave oven and classify cooking methods for those data samples.

Figure 11 shows some sample data collected from the in-the-wild experiments. The red and blue vertical dotted lines represent the start and end of the cooking activities detected by the high-level classifier. Figure 12 summarizes the confusion matrices of the obtained classification results from the two locations. We achieved an average accuracy of 0.77 and an average F_1 score of 0.74 in Household #1, and an average accuracy of 0.70 and an average F_1 score of 0.70 in Household #2. Considering that the data collected from these two locations (i.e., 30 samples from Household #1 and 20 samples from Household #2) are substantially smaller than the size of the data collected in the in-laboratory experiments, the achieved classification results are promising. These results, although validated in a relatively small number of kitchen environments with a small data size, demonstrate the potential of *NOSE* to detect cooking by frying in in-the-wild settings in a longitudinal manner.

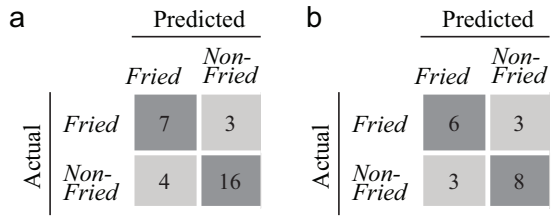

Fig. 12. The confusion matrix for detecting the occurrences of cooking by frying (frying vs. non-frying) for a) Household #1 and b) Household #2.

7 LIMITATIONS, FUTURE WORKS, AND CONCLUSION

In this work, we have designed, developed, and evaluated a novel odor sensing engine, *NOSE*, for the ambient monitoring of different cooking activities in kitchen environments. The *NOSE* hardware platform is implemented with six commercially available metal oxide semiconductor (MOS) sensors that are responsive to the gas emissions of common cooking activities. We have systematically implemented feature engineering and data analytics algorithms to process raw gas profiles and detect deep-frying in kitchen environments. We demonstrated the validity of our *NOSE* system for identifying the gaseous fingerprints of different cooking methods and classifying deep-frying on a total of 207 cooking events collected in a laboratory setting with an average F_1 score of 0.89. We also showed that *NOSE* can reliably classify frying cooking with different ingredients, cooking oils, and with/without the use of kitchen ventilation fan. However, we observed that the performance of *NOSE* was sensitive towards the amount of oil that the system was able to detect as sautéing with an accuracy of 0.67. We also showed that the classification of frying cannot be done accurately in a location-agnostic manner because of location-dependent factors – such as the physical size of the kitchen, the existence and capacity of the ventilation fan, and the ambient gas profile in the environment – and that classification models need to be tailored towards a new kitchen by gradually adding newly available data to the existing model. Finally, *NOSE* was deployed in two real-world households over a three-week period to demonstrate that our system can detect the presence of frying with an average F_1 score of 0.72. Considering the relatively small number of data samples collected from these in-the-wild experiments, the reported results are promising and demonstrate the potential for improvement with a larger data set.

The proposed *NOSE* system does not include humidity and temperature sensors, which seem like obvious choices for recognizing cooking by boiling. Since all sensors are responsive to the temperature and humidity levels, we believe that dedicated humidity and temperature sensors are not necessarily required for our application. We conducted a separate experiment evaluating the contribution of a dedicated humidity and temperature sensor to deep-frying detection performance. Chicken, beef, and potato were cooked with three different methods (deep-frying, grilling, and boiling) at Location 1 of the in-laboratory experiment, each of which was repeated three times. The use of the temperature humidity sensor had a minimal impact on the classification performance as we hypothesized, resulting in an F_1 score of 0.94 with the humidity sensor and 0.93 without the sensor. There was no statistical significance in the difference at a 99% confidence level.

Several limitations and future work associated with this study deserve further discussion. First, the proposed sensing modality has the difficulty in transferring the learned knowledge about cooking activities to new kitchen environments. The proposed system requires new data and retraining for each deployment to a new location. Our model demands some degree of user participation, and the system's performance may depend on the quality of users' labeling of the data. However, we showed that the system may need as few as 40-50 data samples to support reasonable classification performance when a sufficient amount of baseline training data are available

for the classification model (e.g., Locations 2 and 3 in Figure 10). Alternatively, if we deploy *NOSE* on a larger scale, in terms of both locations and duration, we would have a large enough data set to potentially combat this limitation with other methodologies. For example, unsupervised clustering algorithms can be applied to identify kitchens in the training set with similar sensor behaviors, and we can train a classification model based on data obtained from those kitchens in order to optimize the classification performance. Such an approach may also enable generalization without requiring additional user inputs as well as detection of ingredients used. However, unfortunately, a larger data set also implies that retraining machine learning models on newly available data samples from a new kitchen environment might not be technically feasible due to the high computational complexity. In such a case, efficient online learning techniques, such as incremental learning algorithms [26, 74] that optimize the regularization and kernel parameters instead of complete retraining, can be incorporated. These remain as exciting future research questions.

Second, as we reported in Section 6.3, we witnessed that *NOSE* was not only responsive to cooking activities that are performed on the main stove unit, but also to activities using a microwave oven that was placed near the sensor. Constructing classification models to detect more fine-grained activities within the kitchen area, including cooking with the microwave oven (e.g., detection of the ingredients and amount of oil contained in the food), remains as future work. Third, we only experimented with placing *NOSE* directly above the range. Though we expect similar performance, additional experiments need to be conducted to determine if *NOSE* is sensitive to placement and thus whether it is applicable to a broader range of kitchen layouts. Fourth, all kitchens involved in this study were equipped with electric stoves. We believe the reported results herein may not apply to kitchens with gas stoves as gaseous substances from the fuel source may affect the sensor readings. The validation of the proposed system on different types of stoves remains as important future work. Finally, since our in-the-wild experiments only lasted for three weeks, we were not able to study the possibility of long-term variations in the ambient gas profiles of kitchen environments (e.g. seasonal changes). Future work on the longitudinal deployment of *NOSE* will need to investigate the existence or possible effects of long-term variations on the system's performance. We believe such variations, if they exist, can be compensated for by periodic retraining or by unsupervised learning given a large enough data set (e.g. finding a similar environment).

We believe that this proof-of-concept study to understand cooking activities based on odor sensing has great potential to open up new research and development opportunities in the area of automatic dietary monitoring. The results of our in-laboratory and in-the-wild studies are promising and motivate us to further validate our observations with a large-scale deployment of *NOSE*, which will allow us to investigate more advanced data analysis tools (e.g., true online learning, unsupervised learning for finding similar kitchen environments, or deep-learning). We also plan to conduct a longitudinal study to track the cooking practices of different households and correlate to their practices to healthy dietary behavior (e.g., the Healthy Eating Index [52]). In addition to refining the sensing engine, we plan to explore the potential of this single-point *NOSE* platform for enabling longer-term behavior changes towards healthy cooking practices.

ACKNOWLEDGMENTS

This study was supported by the College of Information and Computer Sciences and the Institute for Applied Life Sciences at the University of Massachusetts Amherst.

REFERENCES

[1] [n. d.]. *Apple Watch.* https://www.apple.com/watch
[2] [n. d.]. *Arduino.* https://www.arduino.cc/
[3] [n. d.]. *Fitbit.* https://www.fitbit.com/home
[4] [n. d.]. *Lose it.* https://www.loseit.com/
[5] [n. d.]. *MyFitnessPal.* https://www.myfitnesspal.com/

[6] 2018. *World Health Organization.* http://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight

[7] Oliver Amft. 2010. A wearable earpad sensor for chewing monitoring. In *Sensors, 2010 IEEE*. IEEE, 222–227.

[8] Oliver Amft and Gerhard Tröster. 2008. Recognition of dietary activity events using on-body sensors. *Artificial intelligence in medicine* 42, 2 (2008), 121–136.

[9] S Ampuero and JO Bosset. 2003. The electronic nose applied to dairy products: a review. *Sensors and Actuators B: Chemical* 94, 1 (2003), 1–12.

[10] Lenore Arab, Deborah Estrin, Donnie H Kim, Jeff Burke, and Jeff Goldman. 2011. Feasibility testing of an automated image-capture method to aid dietary recall. *European journal of clinical nutrition* 65, 10 (2011), 1156.

[11] Elizabeth Barrett-Connor. 1991. Nutrition epidemiology: how do we know what they ate? *The American journal of clinical nutrition* 54, 1 (1991), 182S–187S.

[12] Abdelkareem Bedri, Apoorva Verlekar, Edison Thomaz, Valerie Avva, and Thad Starner. 2015. Detecting mastication: A wearable approach. In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*. ACM, 247–250.

[13] Abdelkareem Bedri, Apoorva Verlekar, Edison Thomaz, Valerie Avva, and Thad Starner. 2015. A wearable system for detecting eating activities with proximity sensors in the outer ear. In *Proceedings of the 2015 ACM International Symposium on Wearable Computers*. ACM, 91–92.

[14] Shengjie Bi, Tao Wang, Nicole Tobias, Josephine Nordrum, Shang Wang, George Halvorsen, Sougata Sen, Ronald Peterson, Kofi Odame, Kelly Caine, Ryan Halter, Jacob Sorber, and David Kotz. 2018. Auracle: Detecting Eating Episodes with an Ear-mounted Sensor. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 3, Article 92 (Sept. 2018), 27 pages. https://doi.org/10.1145/3264902

[15] Yin Bi, Mingsong Lv, Chen Song, Wenyao Xu, Nan Guan, and Wang Yi. 2016. AutoDietary: A wearable acoustic sensor system for food intake recognition in daily life. *IEEE Sensors Journal* 16, 3 (2016), 806–816.

[16] Susanna Buratti, Simona Benedetti, and Gabriella Giovanelli. 2017. Application of electronic senses to characterize espresso coffees brewed with different thermal profiles. *European Food Research and Technology* 243, 3 (2017), 511–520.

[17] Steven Cadavid, Mohamed Abdel-Mottaleb, and Abdelsalam Helal. 2012. Exploiting visual quasi-periodicity for real-time chewing event detection using active appearance models and support vector machines. *Personal and Ubiquitous Computing* 16, 6 (2012), 729–739.

[18] Karen J Campbell, David A Crawford, Jo Salmon, Alison Carver, Sarah P Garnett, and Louise A Baur. 2007. Associations between the home food environment and obesity-promoting eating behaviors in adolescence. *Obesity* 15, 3 (2007), 719–730.

[19] Erin A Carroll, Mary Czerwinski, Asta Roseway, Ashish Kapoor, Paul Johns, Kael Rowan, and MC Schraefel. 2013. Food and mood: Just-in-time support for emotional eating. In *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, 252–257.

[20] Keng-hao Chang, Shih-yen Liu, Hao-hua Chu, Jane Yung-jen Hsu, Cheryl Chen, Tung-yun Lin, Chieh-yu Chen, and Polly Huang. 2006. The diet-aware dining table: Observing dietary behaviors over a tabletop surface. In *International Conference on Pervasive Computing*. Springer, 366–382.

[21] Pei-Yu Peggy Chi, Jen-Hao Chen, Hao-Hua Chu, and Jin-Ling Lo. 2008. Enabling calorie-aware cooking in a smart kitchen. In *International Conference on Persuasive Technology*. Springer, 116–127.

[22] E Choe and DB Min. 2007. Chemistry of deep-fat frying oils. *Journal of food science* 72, 5 (2007).

[23] Mary Ruth Craig, Alan R Kristal, Carrie L Cheney, and Ann L Shattuck. 2000. The prevalence and impact of âĂŸatypicalâĂŹdays in 4-day food records. *Journal of the American Dietetic Association* 100, 4 (2000), 421–427.

[24] George Demiris, Brian K Hensel, Marjorie Skubic, and Marilyn Rantz. 2008. Senior residents' perceived need of and preferences for 'smart home' sensor technologies. *International journal of technology assessment in health care* 24, 1 (2008), 120–124.

[25] George Demiris, Marilyn J Rantz, Myra A Aud, Karen D Marek, Harry W Tyrer, Marjorie Skubic, and Ali A Hussam. 2004. Older adults' attitudes towards and perceptions of 'smart home' technologies: a pilot study. *Medical informatics and the Internet in medicine* 29, 2 (2004), 87–94.

[26] Christopher P Diehl and Gert Cauwenberghs. 2003. SVM incremental learning, adaptation and optimization. In *Neural Networks, 2003. Proceedings of the International Joint Conference on*, Vol. 4. IEEE, 2685–2690.

[27] Yujie Dong, Adam Hoover, Jenna Scisco, and Eric Muth. 2012. A new method for measuring meal intake in humans via automated wrist motion tracking. *Applied psychophysiology and biofeedback* 37, 3 (2012), 205–215.

[28] Yujie Dong, Jenna Scisco, Mike Wilson, Eric Muth, and Adam Hoover. 2014. Detecting periods of eating during free-living by tracking wrist motion. *IEEE journal of biomedical and health informatics* 18, 4 (2014), 1253–1260.

[29] Hossein Rezaei Estakhroueiyeh and Esmat Rashedi. 2015. Detecting moldy bread using an e-nose and the KNN classifier. In *Computer and Knowledge Engineering (ICCKE), 2015 5th International Conference on*. IEEE, 251–255.

[30] Muhammad Farooq, Paula C Chandler-Laney, Maria Hernandez-Reif, and Edward Sazonov. 2015. Monitoring of infant feeding behavior using a jaw motion sensor. *Journal of healthcare engineering* 6, 1 (2015), 23–40.

[31] George F Fine, Leon M Cavanagh, Ayo Afonja, and Russell Binions. 2010. Metal oxide semi-conductor gas sensors in environmental monitoring. *Sensors* 10, 6 (2010), 5469–5502.

[32] Taraka V Gadiraju, Yash Patel, J Michael Gaziano, and Luc Djoussé. 2015. Fried food consumption and cardiovascular health: A review of current evidence. *Nutrients* 7, 10 (2015), 8424–8430.
[33] Jiang Gao, Alexander G Hauptmann, Ashok Bharucha, and Howard D Wactlar. 2004. Dining activity analysis using a hidden markov model. In *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, Vol. 2. IEEE, 915–918.
[34] Julian W Gardner and Philip N Bartlett. 1994. A brief history of electronic noses. *Sensors and Actuators B: Chemical* 18, 1-3 (1994), 210–211.
[35] L Gemming, A Doherty, P Kelly, J Utter, and C Ni Mhurchu. 2013. Feasibility of a SenseCam-assisted 24-h recall to reduce under-reporting of energy intake. *European Journal Of Clinical Nutrition* 67 (04 09 2013), 1095 EP –.
[36] Mahdi Ghasemi-Varnamkhasti, Seyed Saeid Mohtasebi, Maryam Siadat, and Sundar Balasubramanian. 2009. Meat quality assessment by electronic nose (machine olfaction technology). *Sensors* 9, 8 (2009), 6058–6083.
[37] Maria D Guillén and Patricia S Uriarte. 2012. Aldehydes contained in edible oils of a very different nature after prolonged heating at frying temperature: Presence of toxic oxygenated α, β unsaturated aldehydes. *Food Chemistry* 131, 3 (2012), 915–926.
[38] Mark Andrew Hall. 1999. Correlation-based feature selection for machine learning. (1999).
[39] Karen S Hamrick and Ket McClelland. 2016. *Americans' Eating Patterns and Time Spent on Food: The 2014 Eating & Health Module Data*. United States Department of Agriculture, Economic Research Service.
[40] Atsushi Hashimoto, Naoyuki Mori, Takuya Funatomi, Yoko Yamakata, Koh Kakusho, and Michihiko Minoh. 2008. Smart kitchen: A user centric cooking support system. In *Proceedings of IPMU*, Vol. 8. 848–854.
[41] Sen H Hirano, Jed R Brubaker, Donald J Patterson, and Gillian R Hayes. 2013. Detecting cooking state with gas sensors during dry cooking. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*. ACM, 411–414.
[42] Sen H Hirano, Gillian R Hayes, and Khai N Truong. 2015. uSmell: exploring the potential for gas sensors to classify odors in ubicomp applications relative to airflow and distance. *Personal and Ubiquitous Computing* 19, 1 (2015), 189–202.
[43] Mandy Ho, Sarah P Garnett, Louise Baur, Tracy Burrows, Laura Stewart, Melinda Neve, and Clare Collins. 2012. Effectiveness of lifestyle interventions in child obesity: systematic review with meta-analysis. *Pediatrics* (2012), peds–2012.
[44] Haik Kalantarian, Nabil Alshurafa, Tuan Le, and Majid Sarrafzadeh. 2015. Monitoring eating habits using a piezoelectric sensor-based necklace. *Computers in biology and medicine* 58 (2015), 46–55.
[45] Haik Kalantarian, Nabil Alshurafa, and Majid Sarrafzadeh. 2014. A wearable nutrition monitoring system. In *Wearable and Implantable Body Sensor Networks (BSN), 2014 11th International Conference on*. IEEE, 75–80.
[46] Haik Kalantarian, Nabil Alshurafa, and Majid Sarrafzadeh. 2017. A survey of diet monitoring technology. *IEEE Pervasive Computing* 16, 1 (2017), 57–65.
[47] HARRY R Kissileff, GARY Klingsberg, and THEODORE B Van Itallie. 1980. Universal eating monitor for continuous recording of solid or liquid consumption in man. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* 238, 1 (1980), R14–R22.
[48] Yasuki Kobayashi, Tsutomu Terada, and Masahiko Tsukamoto. 2011. A context aware system based on scent. In *Wearable Computers (ISWC), 2011 15th Annual International Symposium on*. IEEE, 47–50.
[49] Fanyu Kong and Jindong Tan. 2011. Dietcam: Regular shape food recognition with a camera phone. In *Body Sensor Networks (BSN), 2011 International Conference on*. IEEE, 127–132.
[50] R Koppmann, K von Czapiewski, and JS Reid. 2005. A review of biomass burning emissions, part I: gaseous emissions of carbon monoxide, methane, volatile organic compounds, and nitrogen containing compounds. *Atmospheric chemistry and physics discussions* 5 (2005), 10455–10516.
[51] Mandy Korpusik, Nicole Schmidt, Jennifer Drexler, Scott Cyphers, and James Glass. 2014. Data collection and language understanding of food descriptions. In *Spoken Language Technology Workshop (SLT), 2014 IEEE*. IEEE, 560–565.
[52] Susan M Krebs-Smith, TusaRebecca E Pannucci, Amy F Subar, Sharon I Kirkpatrick, Jennifer L Lerman, Janet A Tooze, Magdalena M Wilson, and Jill Reedy. 2018. Update of the healthy eating index: HEI-2015. *Journal of the Academy of Nutrition and Dietetics* 118, 9 (2018), 1591–1602.
[53] EV Krisilova, AM Levina, and VA Makarenko. 2014. Determination of the volatile compounds of vegetable oils using an ion-mobility spectrometer. *Journal of analytical chemistry* 69, 4 (2014), 371–376.
[54] Ronilda Lacson and William Long. 2006. Natural language processing of spoken diet records (SDRs). In *AMIA Annual Symposium Proceedings*, Vol. 2006. American Medical Informatics Association, 454.
[55] Jindong Liu, Edward Johns, Louis Atallah, Claire Pettitt, Benny Lo, Gary Frost, and Guang-Zhong Yang. 2012. An intelligent food-intake monitoring system using wearable sensors. In *Wearable and Implantable Body Sensor Networks (BSN), 2012 Ninth International Conference on*. IEEE, 154–160.
[56] Tim Lobstein, Louise Baur, and Ricardo Uauy. 2004. Obesity in children and young people: a crisis in public health. *Obesity reviews* 5, s1 (2004), 4–85.
[57] Amy Loutfi, Silvia Coradeschi, Ganesh Kumar Mani, Prabakaran Shankar, and John Bosco Balaguru Rayappan. 2015. Electronic noses for food quality: A review. *Journal of Food Engineering* 144 (2015), 103–111.
[58] Donald S Mottram. 1998. Flavour formation in meat and meat products: a review. *Food chemistry* 62, 4 (1998), 415–424.

[59] Jun Nishimura and Tadahiro Kuroda. 2008. Eating habits monitoring using wireless wearable in-ear microphone. In *Wireless Pervasive Computing, 2008. ISWPC 2008. 3rd International Symposium on*. IEEE, 130–132.
[60] Jon Noronha, Eric Hysen, Haoqi Zhang, and Krzysztof Z Gajos. 2011. Platemate: crowdsourcing nutritional analysis from food photographs. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*. ACM, 1–12.
[61] OECD. 2016. Society at a Glance 2016: OECD Social Indicators. *OECD Publishing* (2016).
[62] Gillian O'Loughlin, Sarah Jane Cullen, Adrian McGoldrick, Siobhan O'Connor, Richard Blain, Shane O'Malley, and Giles D. Warrington. 2013. Using a Wearable Camera to Increase the Accuracy of Dietary Analysis. *American Journal of Preventive Medicine* 44, 3 (2013), 297 – 301. https://doi.org/10.1016/j.amepre.2012.11.007
[63] Temiloluwa Olubanjo and Maysam Ghovanloo. 2014. Real-time swallowing detection based on tracheal acoustics. In *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*. IEEE, 4384–4388.
[64] Santford V Overton and John J Manura. 1995. Analysis of volatile organics in cooking oils by thermal desorption gas chromatography-mass spectrometry. *Journal of agricultural and food chemistry* 43, 5 (1995), 1314–1320.
[65] Abhinav Parate, Meng-Chieh Chiu, Chaniel Chadowitz, Deepak Ganesan, and Evangelos Kalogerakis. 2014. Risq: Recognizing smoking gestures with inertial sensors on a wristband. In *Proceedings of the 12th annual international conference on Mobile systems, applications, and services*. ACM, 149–161.
[66] Krishna Persaud and George Dodd. 1982. Analysis of discrimination mechanisms in the mammalian olfactory system using a model nose. *Nature* 299, 5881 (1982), 352.
[67] Temiloluwa Prioleau, Elliot Moore II, and Maysam Ghovanloo. 2017. Unobtrusive and wearable systems for automatic dietary monitoring. *IEEE Transactions on Biomedical Engineering* 64, 9 (2017), 2075–2089.
[68] Qibin Qi, Audrey Y Chu, Jae H Kang, Jinyan Huang, Lynda M Rose, Majken K Jensen, Liming Liang, Gary C Curhan, Louis R Pasquale, Janey L Wiggs, Immaculata De Vivo, Andrew T Chan, Hyon K Choi, Rulla M Tamimi, Paul M Ridker, David J Hunter, Walter C Willett, Eric B Rimm, Daniel I Chasman, Frank B Hu, and Lu Qi. 2014. Fried food consumption, genetic risk, and body mass index: gene-diet interaction analysis in three US cohort studies. *BMJ* 348 (2014). https://doi.org/10.1136/bmj.g1610 arXiv:https://www.bmj.com/content/348/bmj.g1610.full.pdf
[69] Tauhidur Rahman, Alexander T. Adams, Mi Zhang, Erin Cherry, Bobby Zhou, Huaishu Peng, and Tanzeem Choudhury. 2014. BodyBeat: A Mobile System for Sensing Non-speech Body Sounds. In *Proceedings of the 12th Annual International Conference on Mobile Systems, Applications, and Services (MobiSys '14)*. ACM, New York, NY, USA, 2–13. https://doi.org/10.1145/2594368.2594386
[70] Sasank Reddy, Andrew Parker, Josh Hyman, Jeff Burke, Deborah Estrin, and Mark Hansen. 2007. Image browsing, processing, and clustering for participatory sensing: lessons from a DietSense prototype. In *Proceedings of the 4th workshop on Embedded networked sensors*. ACM, 13–17.
[71] Marla Reicks, Amanda C Trofholz, Jamie S Stang, and Melissa N Laska. 2014. Impact of cooking and home food preparation interventions among adults: outcomes and implications for future programs. *Journal of nutrition education and behavior* 46, 4 (2014), 259–276.
[72] Frank Röck, Nicolae Barsan, and Udo Weimar. 2008. Electronic nose: current status and future trends. *Chemical reviews* 108, 2 (2008), 705–725.
[73] Raffaele Romano, Anella Giordano, Laura Le Grottaglie, Nadia Manzo, Antonello Paduano, Raffaele Sacchi, and Antonello Santini. 2013. Volatile compounds in intermittent frying by gas chromatography and nuclear magnetic resonance. *European Journal of Lipid Science and Technology* 115, 7 (2013), 764–773.
[74] Stefan Ruping. 2001. Incremental learning with support vector machines. In *Data Mining, 2001. ICDM 2001, Proceedings IEEE International Conference on*. IEEE, 641–642.
[75] Krushnapriya Sahoo, Bishnupriya Sahoo, Ashok Kumar Choudhury, Nighat Yasin Sofi, Raman Kumar, and Ajeet Singh Bhadoria. 2015. Childhood obesity: causes and consequences. *Journal of family medicine and primary care* 4, 2 (2015), 187.
[76] Emmanuelle Schaller, Jacques O Bosset, and Felix Escher. 1998. âĂŸElectronic nosesâĂŹ and their application to food. *LWT-Food Science and Technology* 31, 4 (1998), 305–316.
[77] Simon M Scott, David James, and Zulfiqur Ali. 2006. Data analysis for electronic nose systems. *Microchimica Acta* 156, 3-4 (2006), 183–207.
[78] Pang-Ning Tan et al. 2007. *Introduction to data mining*. Pearson Education India.
[79] Edison Thomaz, Aman Parnami, Irfan Essa, and Gregory D Abowd. 2013. Feasibility of identifying eating moments from first-person images leveraging human computation. In *Proceedings of the 4th International SenseCam & Pervasive Imaging Conference*. ACM, 26–33.
[80] Edison Thomaz, Cheng Zhang, Irfan Essa, and Gregory D Abowd. 2015. Inferring meal eating activities in real world settings from ambient sounds: A feasibility study. In *Proceedings of the 20th International Conference on Intelligent User Interfaces*. ACM, 427–431.
[81] United States Department of Agriculture. 2019. Americans Spend an Average of 37 Minutes a Day Preparing and Serving Food and Cleaning Up. (2019). https://www.ers.usda.gov/amber-waves/2016/november/americans-spend-an-average-of-37-minutes-a-day-preparing-and-serving-food-and-cleaning-up/
[82] Alphus D Wilson and Manuela Baietto. 2009. Applications and advances in electronic-nose technologies. *Sensors* 9, 7 (2009), 5099–5148.

[83] Wen Wu and Jie Yang. 2009. Fast food recognition from videos of eating for calorie estimation. In *Multimedia and Expo, 2009. ICME 2009. IEEE International Conference on*. IEEE, 1210–1213.
[84] Alam Zeb. 2019. *Food Frying: Chemistry, Biochemistry, and Safety*. John Wiley & Sons.
[85] Bo Zhou, Jingyuan Cheng, Mathias Sundholm, Attila Reiss, Wuhuang Huang, Oliver Amft, and Paul Lukowicz. 2015. Smart table surface: A novel approach to pervasive dining monitoring. In *Pervasive Computing and Communications (PerCom), 2015 IEEE International Conference on*. IEEE, 155–162.
[86] Fengqing Zhu, Marc Bosch, Carol J Boushey, and Edward J Delp. 2010. An image analysis system for dietary assessment and evaluation. In *Image Processing (ICIP), 2010 17th IEEE International Conference on*. IEEE, 1853–1856.

Received August 2018; revised February 2019; accepted April 2019