Detecting Eating Episodes by Tracking Jawbone Movements with a Non-Contact Wearable Sensor

KEUM SAN CHUN, University of Texas at Austin, USA
SARNAB BHATTACHARYA, University of Texas at Austin, USA
EDISON THOMAZ, University of Texas at Austin, USA

Eating is one of the most fundamental human activities, and because of the important role it plays in our lives, it has been extensively studied. However, an objective and usable method for dietary intake tracking remains unrealized despite numerous efforts by researchers over the last decade. In this work, we present a new wearable computing approach for detecting eating episodes. Using a novel multimodal sensing strategy combining accelerometer and range sensing, the approach centers on a discreet and lightweight instrumented necklace that captures head and jawbone movements without direct contact with the skin. An evaluation of the system with 32 participants comprised of three phases resulted in eating episodes detected with 95.2% precision and 81.9% recall in controlled studies and 78.2% precision and 72.5% recall in the free-living study. This research add technical contributions to the fields of wearable computing, human activity recognition, and mobile health.

CCS Concepts: • **Human-centered computing** → **Empirical studies in ubiquitous and mobile computing**; • **Applied computing** → *Health informatics*;

Additional Key Words and Phrases: automated dietary monitoring, eating detection, food intake, gesture recognition, commodity sensing

1 INTRODUCTION

Good nutrition is vital for optimal growth, development, and prevention of disease [21, 27]. In particular, poor eating habits are strongly tied to obesity, a major public health problem [28]. The World Health Organization estimates that more than 1.9 billion adults, 18 and older, were overweight in 2014 [1]. Of these, over 600 million were obese [1]. Due to the importance of diet in human life, health researchers have been interested in understanding the science of measuring dietary intake for decades [40], often relying on self-report-based instruments such as food frequency questionnaires (FFQ), 24-hour recalls, and food diaries. Unfortunately, these methods have been widely recognized to have serious shortcomings such as biases, which increase the risk of false characterization of dietary habits.

To circumvent the weaknesses observed in self-report-based methods, and by leveraging advances in mobile and sensing technologies, a large body of research work around automatic dietary monitoring has emerged

Authors' addresses: Keum San Chun, University of Texas at Austin, 2501 Speedway, Austin, Texas, 78712, USA, gmountk@utexas.edu; Sarnab Bhattacharya, University of Texas at Austin, 2501 Speedway, Austin, Texas, 78712, USA, sarnab2008@gmail.com; Edison Thomaz, University of Texas at Austin, 2501 Speedway, Austin, Texas, 78712, USA, ethomaz@utexas.edu.

over the last decade. Automatic eating detection has grown to become a key area of interest since it captures the temporal dynamics of eating behaviors and can trigger processes such as behavioral interventions [41]. Although steady progress has been made in the field, many proposed systems for eating detection have required individuals to wear obtrusive devices. These systems, exemplified by neck collars for swallow detection [3, 12, 26] or microphones inside the ear canal to detect chewing [30] proved to be largely impractical and unpopular for everyday usage. Consequently, few human subject experiments targeting automated dietary monitoring have been conducted in naturalistic settings for an extended period of time. In the last couple of years, techniques based on off-the-shelf consumer devices such as mobile phones, smart watches, and wearables (e.g., Google Glass) have been developed. However, despite their practicality, these methods have not yet demonstrated performance results that meet application requirements. These limitations have prevented eating detection, and thus automatic dietary monitoring, from having significant real-world impact.

In this paper, we present a wearable device in the form factor of an instrumented necklace that can detect eating episodes, a keystone towards implementing a fully autonomous dietary monitoring system. The device operates by capturing jawbone movements through a distance measure obtained with a proximity sensor and using this signal to distinguish eating versus non-eating activities. Even though it is a neck-mounted device, the sensor does not need to be in direct contact with skin, allowing for a lightweight design and form-factor if compared to previous systems. The contributions of this work are:

- The design and implementation of the wearable system alongside a computational approach for analyzing and classifying proximity sensor data as eating or non-eating.
- An evaluation of the system with 32 participants comprised of three phases, a controlled laboratory study, a controlled-field study, and an in-the-wild study. Eating episodes were detected with 95.2% precision and 81.9% recall in the controlled studies and 78.2% precision and 72.5% recall in the free-living study.
- An annotated dataset of proximity sensor data collected in the user studies that we share with the community to encourage other researchers to expand upon our work.

2 THE NEED FOR OBJECTIVE DIETARY MONITORING

Due to the importance of diet in human life, health researchers have been interested in understanding the science of measuring dietary intake for decades [40]. Bingham traced the first attempts to perform this measurement outside of a controlled setting to the 1930s and 1940s [9]. Widdoson et al., for instance, presented an examination of English diets using the weighted food record in 1936 [48, 49]; the process involved recording the weight of each item of food and beverage consumed. Soon thereafter, Wiehl, Turner and Reed pioneered interview-based dietary recall and food frequency methods, with the goal of estimating energy intake[47, 51, 52].

Dietary recalls, food records and food frequency questionnaires (FFQ) remain the primary dietary assessment mechanisms in use today, and are considered to be the gold-standard by nutritional epidemiologists. In dietary recall, an interviewer assists an individual in remembering what was eaten over a period of time, typically 24 hours. Dietary records are different in that participants are asked to write down what is consumed shortly after the eating moment. Jacobs observed that in practice people often wait until the end of the day to record what they ate [23]. In this case, the dietary record becomes a self-administered recall. With food frequency questionnaires (FFQ), which come in many flavors in terms of the number and specificity of questions, the objective is to obtain more general dietary knowledge and habits. For instance, a question in a FFQ might be "How often do you eat pizza, and if so, how often and how many slices do you typically consume?". More detailed questions might be asked, such as "When you drink milk, is it typically fat free, 1%, or whole-milk?" or "Do you prefer white or whole-wheat bread?". Despite the use of these self-report methods, observations have shown that people tend to forget items that were eaten, underestimate large portion sizes, over-estimate small ones and, in general, are susceptible to a large variety of errors and biases [23, 33]. Recently it has become possible to measure the

Table 1. A summary of different dietary monitoring approaches.

Sensing Modality	Proxy	Sensor	Sensor Location	Performance	
Inertial	Bite	Accelerometer	Wrist	F-scores of 76.1%	[45]
Inertial	Bite	Gyroscope	Wrist	81% Accuracy at 1-sec resolution	[15]
Acoustic	Mastication	Microphone	Neck/Ear	detection accuracy of 83%	[35]
Physiological	Mastication	EMG Sensor	Neck	80% classification accuracy	[3]
Piezoelectric	Swallowing	Piezoelectric Sensor	Neck	F1-score of 85% on solid foods, 86% on liquid	[26]
Proximity	Mastication	Proximity Sensor	Ear	95.3% Accuracy	[8]

accuracy of these gold-standard dietary assessment instruments thanks to the doubly-labeled water technique [29]. Findings confirmed their weaknesses. In light of these limitations, researchers have begun to question the validity of nutritional data collected by self-report methods. Archer et al. focused on the National Health and Nutrition Examination Survey (NHANES), stating that "*methodological limitations compromise the validity of U.S. nutritional surveillance data and the empirical foundation for formulating dietary guidelines and public health policies*" [4]. Dhurandhar et al. believe traditional instruments like dietary recalls and records should not be used at all for energy intake (EI) and physical activity energy expenditure (PAEE) assessment. In their own words, "*...it is time to move from the common view that self-reports of EI and PAEE are imperfect, but nevertheless deserving of use, to a view commensurate with the evidence that self-reports of EI and PAEE are so poor that they are wholly unacceptable for scientific research on EI and PAEE.*" [13].

3 RELATED WORK

Dietary assessment challenges and limitations have fueled interest in automated processes starting in the 1980s. At the time, researchers tried to detect chews and swallows using oral sensors in order to measure the palatability and satiating value of foods [42]. Today, most of the work in dietary assessment involves mobile and wearable devices and the utilization of various forms of sensing modalities such as acoustic, inertial, physiological, visual, capacitive, and piezoelectric sensing.

3.1 Acoustic Sensing

Wearable-based acoustic sensing has been widely explored in dietary monitoring. Sazonov et al. proposed a system for monitoring swallowing and chewing that included a small microphone located over the laryngopharynx [31, 39]. Similarly, Olubanjo and Ghovanloo tracked swallowing events from tracheal acoustics [34]. Passler investigated the problem of intake monitoring using microphones in the outer ear canal [35]. Bodyscope explored how accurately a large number of activities, including eating and drinking, could be recognized with a single acoustic sensor attached to the user's neck [53]. In 2012, Liu et al. developed a food logging application based on the capture of audio and first-person point-of-view images [30]. The system processed incoming sounds through a head-mounted microphone and a classifier identified when chewing was taking place, prompting a wearable camera to capture a video of the eating activity.

3.2 Inertial Sensing

The widespread availability of small wearable accelerometers and gyroscopes has opened up a new avenue for detecting eating activities through on-body inertial sensing. Almost a decade ago, Amft et al. detected eating gestures with a measurement system comprised of five inertial sensors placed on the body (wrists, upper arms and on the upper torso) [2, 3, 25]. In 2013, Dong at al. put forth a method for detecting intake gestures in real-world settings based on a wrist-motion energy heuristic [14, 15]. The authors evaluated the approach by having participants wear a smartphone on the wrist collecting continuous inertial sensor data. As a follow up to

Fig. 1. Several instrumented neckbands aimed at chewing and swallowing detection have been developed over the years but their size, form factor, and aesthetic have limited their mainstream adoption. [3, 12, 17, 26]

this work, Thomaz evaluated an approach for inferring eating moments based on 3-axis accelerometry collected with a popular smartwatch, the Pebble watch [44]. Trained with data collected in a semi-controlled laboratory setting with 20 subjects, the system recognized eating moments in two free-living condition studies with F1 scores of 76.1% and 71.3%. These results were highly encouraging since one of the field experiments was conducted over 30 days, longer than previous ecologically-valid automated dietary monitoring studies.

3.3 Physiological Sensing

Electroglottography (EGG) and electromyography (EMG) have been used to detect chewing and swallowing activities. EGG uses two electrodes placed on the neck to measure the impedance changes induced by the movements of larynx during eating [6]. Farooq et al. evaluated the approach of using EGG on 25 participants (12 male and 13 female) and compared the performance with that of acoustic sensing [17]. The EGG approach resulted in better performance than acoustic sensing approach on both male and female participants. While EMG also uses two electrodes like EGG, the measurement made for EMG is the activation of mandibular muscles as opposed to the impedance change in EGG. Since EMG focuses on detecting action potentials associated with the contraction of mandibular muscles, EMG is more often used for characterizing chewing behavior while EGG is more focused on characterizing swallowing behavior. While both EMG and EGG can accurately monitor eating activities, both EMG and EGG necessitate that the sensors be attached to the skin. In addition, the attachment of electrodes often leads to larger form factors. As all the previous studies involving EGG and EMG took place in laboratory settings, and the wearability and the practical application of physiological sensing in the real world need to be further investigated.

3.4 Camera Based Dietary Monitoring

Camera-based approach provides an effective means of monitoring eating activities using images captured in intervals. Unlike other methods which leverage swallowing or chewing as proxy for eating detection, camera-based approach relies on the visual information from photos or videos that are captured either continuously or at a regular interval [7, 43]. The camera-based approach has been successfully employed in many applications, and proven to be useful for recording the ground truth [5, 7]. In addition, with image processing, the method has demonstrated that the amount of food consumption can be accurately estimated [10, 24]. Despite its strength, the camera-based approach presents a privacy concern as the camera would capture other people's images as well as their daily activities [43].

3.5 Capacitive Sensing

In 2010, Cheng et al. proposed a new sensing approach that leverages capacitance change induced by contraction of muscles or movements of body. [11]. Cheng et al. demonstrated that the conductive textile can be incorporated in a neckband for eating detection. [12] The application of capacitive sensor in neckband resulted in 84.4% accuracy. While Cheng et al. proposed that neckband can be useful for elderly care or cognitive disease monitoring, the neckband presents a significant challenge for mainstream adoption for dietary monitoring in real world situations due to its large form factor. (Shown in the fourth image on Figure 1). [12]

3.6 Piezoelectric Sensing

Piezoelectric sensors generate electric signal in response to mechanical changes.[38] This principle has been leveraged in multiple applications of dietary monitoring, specifically for chewing detection and swallowing detection. [18–20, 26] In 2016 Farooq and Sazonov demonstrated that eyeglasses that leverages piezoelectric sensor and an accelerometer for dietary monitoring resulted in the F1 score of 99.85% [19] for sedantary eatig, and 94.16% for eating while walking. [19]

3.7 Multimodal Sensing

In the last two years, the utilization of *multiple* sensing modalities to infer eating activity has gained momentum. Recently, Rahman et al. and Merck et al. have merged multiple sensing modalities in eating recognition and prediction with promising results. Rahman et al. built a wearable sensing framework consisting of an array of sensors that captured physical activity, location, heart rate, electrodermal activity, skin temperature and caloric expenditure [37]. Merck et al. combined head and wrist motion (Google Glass, smartwatches on each wrist), with audio (custom earbud microphone) totaling 72 hours of data from 6 participants [32].

3.8 Challenges and Opportunities

As these examples of prior work make evident, acoustic and inertial sensing have been extensively used in dietary monitoring with promising results. However, adoption of these approaches has been hampered by shortcomings. A major reason why inertial sensing is appealing is that accelerometers and gyroscopes have been widely incorporated into commodity devices such as smartwatches and activity tracking bands, making it easy to collect data without specialized and custom-built systems. Unfortunately, the rate of food intake gesture misclassification (i.e., false positives) from inertial measurements has remained high in naturalistic settings despite many years of research in the field.

A similar functional tradeoff also affects acoustic sensing. Although audio is a very rich signal reflecting dietary markers such as chewing and swallowing, audio processing is computationally intensive. This is a particular problem for wearable devices powered by small batteries and even more so when real-time activity classification is required. Another issue of audio sensing, which also applies to inertial and muscle activity sensing, is that the sensors must be in direct contact with the body so that the acoustic signatures of eating can be properly captured. Due to this requirement, researchers have developed instrumented neckbands to keep sensors compressed against the skin [3, 12, 26]; some of these neckbands are shown in Figure 1. While effective for data capture, neckbands are largely impractical for everyday use; its design form factor and aesthetic are considered unappealing, limiting mainstream adoption of these devices.

4 APPROACH: MASTICATION AS EATING PROXY

A distinctive behavioral marker of eating is *mastication*, the process by which food is cut and crushed by teeth so that it can be more easily broken down by enzymes. During mastication, the jaw moves up and down by four muscles in the mouth (i.e., masseter, temporalis, medial and lateral pterygoids) in a mostly deterministic pattern

that varies depending on what is being consumed (i.e., food type). The hypothesis behind the proposed work is that it is possible to infer the presence of dietary activity by continuously measuring the distance between the jaw and the base of the neck, as shown in Figure 4. In principle, this distance signal should be indicative of whether an individual is eating, be predictive of what is being eaten, and capable of discriminating dietary behaviors from other activities such as talking. In contrast to previous efforts that measured markers of mastication with devices in contact with the skin (e.g., acoustic, inertial, electromyography), the jawbone-neckbase distance can be captured with a small, inexpensive and non-contact proximity sensor mounted either on a specialized necklace or adapted to be outfitted into existing clothing or jewelry.

In order to use mastication as a proxy for detecting eating episodes, we take a bottom-up hierarchical approach. The first step involves detecting *chewing*. Once chewing has been identified, we group periods of semi-continuous chewing as a *chewing bout*. Finally, multiple chewing bouts comprise an *eating episode*. We utilize the following operational definitions for these terms:

- **Chewing**: We define *chewing* as an non-interrupted sequence of chews lasting 5 seconds or longer. It is considered the smallest detectable unit in the context of this study. The rationale for choosing 5 seconds as the minimum duration of *chewing* stems from the desire to reduce the rate of false positives. In preliminary experiments, we found that a temporal window shorter than 5 seconds resulted in yawning or nodding being incorrectly detected as chewing, for instance. On the other hand, if longer windows are used, the system might miss actual chews. A 5-second chewing window proved to be a good compromise between these edge cases.
- **Chewing Bout**: While chewing is a marker and proxy for eating, individuals do not typically chew continuously throughout an entire meal. Small breaks are common while reaching out for food, during intake gestures, or even if just for breathing. These pauses represent gaps in-between chewing, and as they do not mark the beginning or end of eating episodes, they should be ignored. To abstract these pauses away, we introduce the notion of *chewing bout*, a group of consecutive *chewing* periods that are within 30 seconds of each other. In effect, a chewing bout can be conceptualized as an entity that chains chewing periods together as a single-unit.
- **Eating Episode**: Recognizing chewing bouts is critical but not sufficient to identify eating episodes. This is because an eating episode such as breakfast, lunch, dinner is often characterized by chewing bouts interspersed with other activities and interruptions, such as chatting with friends, assisting children with their meals, and going to the restroom. When eating alone, it is also common to take pauses during eating, especially when multi-tasking, e.g., eating while browsing the web on a mobile device, reading a magazine, or watching TV. Having the flexibility to consider longer breaks as part of eating episodes is important because it accounts for naturalistic behaviors and enhances the practicality of the approach. Within our inferential framework, we define an eating episode as a sequence of one of more *chewing bouts* that are no more than 5 minutes apart from each other. More than 5 minutes without chewing and there is a good chance that the eating episode has ended. While we experimented with different durations, we found that 5 minutes led to a good balance in terms of performance. We also feel that this parameter can be fine-tuned depending on individual characteristics.

5 SYSTEM DESIGN

To realize our approach, we designed and implemented a wearable necklace system consisting of a proximity sensor, a 3D printed sensor mount, a microcontroller, and a bluetooth module. The components are compactly packaged as shown in Figure 3. The system was powered by a 400 mAh LiPo battery, and we were able to reliably use the device for more than 18 hours. For data collection, annotation and visualization, an Android phone was also used.

	Eating Episode		Chewing Bout		Chewing		Ground Truth

Fig. 2. The figure shows a hypothetical data set for demonstrating the hierarchical classification levels employed in our detection algorithm. False positives are indicated in red color. In the first and the second talk sessions, there are some false positives detected at chewing level. However, they fail to be grouped at chewing bout level. During the restroom visit, there was one false positive eating episode. However, it is discarded in our algorithm since the duration of detected eating episode (red diagonal block) is too short to be considered an eating episode.

Fig. 3. The three sensor mounts at different angles (0, 30, and 45) were prepared to account for the differences in neck postures among individuals. The second and third image show the front and the side view of the sensor on a subject. In our application, the distance *d* between the proximity sensor and jaw is measured as shown in the third image.

5.1 Hardware

5.1.1 Proximity Sensor. The keystone of the proposed wearable platform is the VL6180X sensor by STMicroelectronics. It is a proximity sensor that measures absolute distance between itself and a target. The distance is calculated by measuring the time emitted light takes to travel to the nearest object and reflect back to the sensor (i.e., time-of-flight). In the system, the sensor sits on a board attached to a necklace and the measurement corresponds to the distance between the sensor and the wearer's jawbone. As the jawbone moves up and down due to mastication, these rhythmic patterns are captured as a distance measure (see Figure 4). In principle, eating patterns should have a distinctive pattern signature if compared to non-eating patterns. The proximity sensor draws 1.7 on average, making this component suitable for low-power applications. The maximum distance the sensor can measure is 25. The distance between the sensor and the jaw was maintained below 20 across the entire study. The sensor was placed at the center of the necklace such that it was possible to measure the vertical movement of the jaw.

5.1.2 3D Printed Sensor Mount. The sensor mount was designed to provide a consistent positioning of the sensor. To account for the variations of neck shapes and posture, three sensor mount with different angles (0, 30, and 45) were designed (see Figure 3). The mount contains two larger holes through which a rubber band can pass and two smaller holes for fixing the sensor. The sensor mount is used to hold a proximity sensor in place as well as to connect it to the rubber band that goes around the neck. On one side of the rubber band goes the wires that carry signal collected from the sensor to the micro-controller. The necklace is shown in the right-most image of the Figure 3.
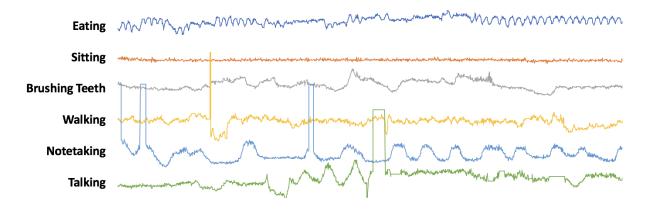
Fig. 4. Six plots featuring different activities a participant was involved in during lab study are shown. The eating signal features the participant chewing celery given as an appetizer. The sitting signal was captured while the participant was watching a movie clip. For note-taking task, the participant copied a short paragraph written on the board. During the note-taking, the participants had to look up on the board periodically, and this is reflected in the periodic rise in amplitude in the second plot from the bottom. In the eating signal, visually distinctive periodic changes in amplitude are observed whereas for the other non-eating activities, no such characteristic is noticeable.

5.1.3 Bluetooth LE Module. In wearable computing, size, weight and heat are critical design factors. To minimize the hardware footprint of the device, only the minimum components necessary for data collection were instrumented in the necklace itself. The sensor data was collected and transmitted to a smartphone in real-time using the Bluetooth LE (i.e., Low Energy) communication protocol. This module relies on the nRF51822 system-on-a-chip (SoC) by Nordic Semiconductor. It is built around the 32-bit ARM Cortex M0 CPU and supports the Bluetooth LE protocol stacks. Unlike other Bluetooth modules which consume approximately 50 during data transmission, BLE consumes approximately 8.5.

5.1.4 Microcontroller. The Arduino Pro Mini was chosen to control the proximity sensor and the Bluetooth module. To reduce the amount of energy consumed by the microcontroller, the 3.3V/8MHz version was used. With its default settings, the Arduino alone consumes approximately 13 when in regular operation. To reduce battery consumption (we used a 400h battery), the Arduino was programmed to remain in sleep-mode for 15 prior to sampling data. Through this modification, we were able to reduce the microcontroller's current draw down to approximately 200.

5.2 Software

5.2.1 Android Application. One of the components of the wearable system is an Android smartphone running an application with Bluetooth LE support. This application serves three important roles. Firstly, it receives and stores the sensor data collected by the wearable necklace. Secondly, its interface is designed to assist with the annotation of the sensor data. And thirdly, the application leverages the smartphone's accelerometer and a simple heuristic to determine whether the individual carrying the phone is walking. This measure is used to help discriminate walking from eating. The Android application receives the acceleration with respect to x, y, and z axis in real time. The magnitude of the acceleration vector is calculated using the euclidean norm, and if the magnitude is greater than 1 m/s^2, it is assumed that the individual is in motion. In that case, the distance measure obtained with the proximity sensor is discarded (i.e., no eating is taking place).

5.3 Data Analysis and Modeling

At the heart of the method is a proposed *four-phase* pipeline for analyzing the data collected by the VL6180X sensor. To reiterate, the goal of the wearable device is to identify episodes of eating from a distance signal measured between the jawbone and the base of the neck. In the first phase of the pipeline, the signal is pre-processed and conditioned. In the second and third phases, chewing and chewing bouts are identified. Lastly, in the fourth phase, eating episodes are inferred from chewing bouts using a clustering algorithm.

5.3.1 Preprocessing. : The VL6180X sensor produces a time series signal with distance measures at a rate of 20Hz. The signal is first pre-processed using a median filter for smoothing. Next, frames are extracted using a 5-second, point-by-point sliding window across the time series (See Figure 5). The window size was determined empirically with the goal of optimizing the balance between inference resolution and rate of false positives. In preliminary experiments, we found that a temporal window shorter than 5 seconds resulted in yawning or nodding being incorrectly detected as chewing, for instance. On the other hand, if longer windows were used, the system could miss actual chews. A 5-second chewing window proved to be a good compromise between these edge cases. Each extracted frame was baseline shifted to eliminate patterns and trends that are not intrinsic to the data. The baseline shift was achieved by subtracting the root-mean-square (RMS) of each frame. Next, each frame was bandpass-filtered to remove artifacts that do not represent dietary activity. The majority of chewing happens in frequencies between 0.94 Hz and 2.17 Hz [36]; therefore a pass band around this frequency range was used (See Figure 5).

5.3.2 Chewing Detection. : Chewing is a distinctive mechanical movement that distinguishes eating from other activities such as talking, yawning, singing or drinking. When captured with a sensor such as the one we propose, this distinction becomes clear in terms of the regularity, frequency and amplitude of the signal. In the chewing detection step, non-eating signals are filtered out using a method called *level-crossing*. With this technique, we define an amplitude and count the number of crossing over the threshold, as shown in figure 5. If the number of amplitude thresholds crossings are aligned with the chewing frequency reported by Po et al., the frame is labeled as chewing [36]. Otherwise, it is labeled as non-chewing. Based on our analysis, we found that the best results are obtained with amplitude thresholds between 1 and 2, as shown in figure 8.

5.3.3 Chewing Bouts. : To eliminate small pauses in-between chewing, we group chewing periods over time into clusters using the DBSCAN clustering algorithm [16], and call these clusters *chewing bouts*. In prior work, DBSCAN has been identified as appropriate for this task since it does not require the specification of a pre-defined number of clusters prior to analysis; this is not possible with other clustering techniques such as K-Means). Furthermore, this algorithm allows for the parametrization of important settings such as the minimum number of chewing periods per cluster [46]. Parametrization also makes DBSCAN amenable to fine-tuning for detecting snacking behavior, when shorter chewing windows might be needed.

5.3.4 Eating Episode Detection. : To identify eating episodes, the beginning and end of each chewing bout are obtained. Then, the duration between each consecutive pair of chewing bouts is calculated. If this duration does not exceed a certain amount (we experimented with duration values between 1 and 10 minutes), the chewing bouts are combined to represent an eating episode.

5.3.5 Evaluation Metric. : To evaluate the eating episode detector, we perform a comparison between the output of the detector and our acquired ground truth on a segment-by-segment basis, where a segment corresponds to a specified time period (e.g., 5 minutes). We assign true positives, true negatives, false positives and false negatives accordingly. Given a particular segment, it is common for a partial match between inference and ground truth; for example, the detector might output that an eating episode took place for only half of a 5-minute segment (i.e., for 50% of the segment), whereas study participants ate during the entire 5-minute period. To avoid missing
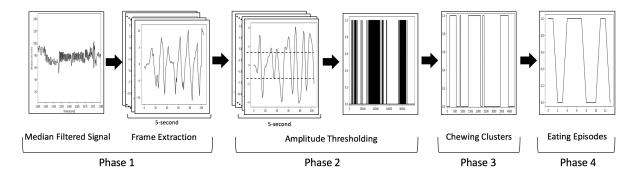
Fig. 5. The first plot shows the median filtered signal. In the first plot, y-axis shows the distance in millimeter and x-axis shows time in second. The second plot shows signals in each frame with baseline adjusted. And the third plot shows the signal after band pass filter (BPF). In the second and third plot, x-axis represents data points and y-axis represents amplitude of filtered signal. Then, the number of crossing about the amplitude thresholds indicated by dotted lines is used to classify if the given frame represents eating or not (fourth plot). DBSCAN clusters the classification output to form clusters as shown in the fifth plot. Finally, eating episode is detected by segmenting the DBSCAN result into non-overlapping time frames and evaluating if each frame contains more than the minimum required proportion of eating (shown in the last plot).

eating episodes (but at a risk of more false positives), we marked segments as "eating" under these partial match cases and tested for different amounts of overlap (i.e., 10%, 20% and 30%). We found that the system performs best with overlaps between 10% and 20%.

6 EXPERIMENTS

Our system and approach were evaluated in three experiments. The first experiment was run in controlled laboratory settings; the second experiment was conducted in controlled *field* settings, a hybrid experimental format that combines the benefits of a free living study with the controlled characteristics of a lab study. Finally, the third experiment was an in-the-wild study in completely naturalistic settings. For each study, precision and recall measures were calculated at the eating episode level to assess the performance of the system.

The food types offered in the studies were chosen to reflect the diverse selection of foods available in the real world (see Figure 7). In particular, food groups were picked to include foods with crunchy textures (e.g. nuts, tacos), soft textures (e.g. ice cream, yogurt), and wet crisp textures (e.g. celery, fruits). A wide array of food types allowed us to examine the performance of the system under different conditions (e.g., soft vs. crunchy foods) and eating styles (e.g., eating with utensils vs. grabbing and holding food with hands).

At the start of each study, it was made sure that the sensor was pointing at the jaw of the participant. In laboratory study, the sensor position was adjusted based on visual examination by the lab staff. In controlled field study and wild study, the distance measure was visualized on Android phone in real-time. Thus, participants were asked to move their jaws for a few seconds as if they would when chewing food, and if the signal captured the vertical movement of the jaw, the study began.

6.1 Laboratory Study

The goal of the lab study was to validate the feasibility of the proposed approach by obtaining examples of dietary and non-dietary behaviors in a controlled setting where participants could be closely monitored. Twenty participants were asked to come to our laboratory and were offered lunch. After check-in, the study staff helped participants put on the necklace device, and adjust the sensor such that it was pointing in the direction of the

Table 2. The study design of lab study, controlled field study and wild study. The average time for each activity is indicated below the corresponding activity.

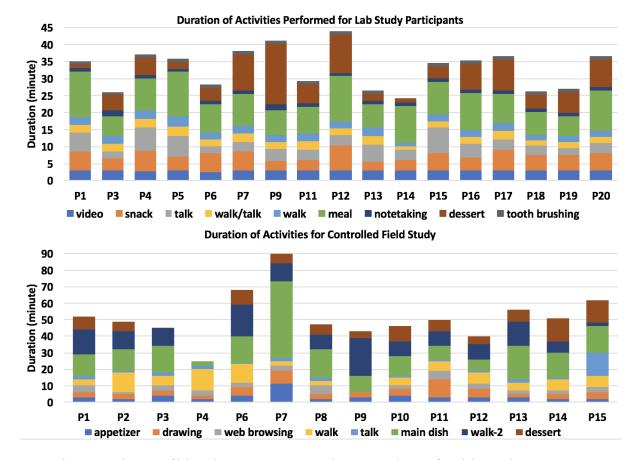| 2*Lab	Check-in	Video	Video/Snack	Talk	Walk/Talk	Walk	Meal	Note-taking	Dessert	Brushing Teeth	Total Time
	5 min	3 min	4.6 min	3.8 min	2 min	2 min	9.4 min	1 min	6 min	1 min	37.8 min
2Controlled Field	Check-in	Appetizer	Drawing	Browsing	Walk	Talk	Meal	Walk		Dessert	Total Time
	5 min	3.4 min	4.2 min	2.9 min	7 min	2.5 min	15.1 min	11.5 min		7.8 min	59.6 min
2*Wild	Various activities with at least one eating episode										Average Time
	2 - 8 hours										4.6 hours

participant's jawbone and chin. Over the course of each session, which lasted for approximately 40 minutes on average, participants were offered a variety of foods with different characteristics, such as celery, nuts, soup, lasagna, fried rice, burrito, yogurt and chocolate stick. Each participant was provided with a three-course meal that consisted of an appetizer, a main dish and a dessert (see Figure 7). For each food served, participants were provided with proper utensils (e.g. fork and knife for lasagna, spoon for soup and fried rice). While they were encouraged to use the provided utensils, they were not required to use them. A common challenge with current eating detection systems is their high rate of false positives. Therefore, participants were also asked to perform common, non-dietary everyday tasks, such as walking, note-taking, talking and brushing teeth. Eating and non-eating activities were inter-weaved (see Table 2). To obtain ground truth labels, participants were observed by study staff and captured by video cameras throughout the study. The study sessions were then annotated for chewing, eating episodes, and non-eating activities.

6.2 Controlled Field Study

While useful, controlled laboratory studies rarely replicate truly naturalistic experiences and settings. Consequently, results obtained in the lab often lack external validity, and thus do not generalize well to real-world conditions. On the other hand, free living studies place participants in their natural everyday environments, but collecting data, and ground truth data in particular, is a significant hurdle. Other than experience sampling techniques using paper diaries or mobile phones, an increasingly popular method for recording ground truth in-the-wild include collecting photos or videos using wearable cameras. However, annotating thousands of photos and hours of videos is a time-consuming and burdensome process. Moreover, continuously capturing media in real-world settings results in a wide range of privacy challenges that cannot be easily overcome.

As an attempt to benefit from the advantages of both laboratory and in-the-wild studies while sidestepping their shortcomings, we explored a hybrid study design that we call *controlled field study*. In our controlled field study, participants met the experimenter in a public place or commercial establishment instead of in the laboratory. All the experimental tasks took place in real-world environments such as cafes, ice-cream shops and restaurants. The experimenter followed and monitored each and every participant throughout the study, and participants were allowed to interact with the experimenter as if he or she were a friend. The advantage of this approach is that all activities take place in the wild, and since the experimenter is together with the participants, ground truth can be reliably recorded.

Fifteen participants were recruited in our controlled field study, which began at a university library. After putting on and having the necklace device adjusted for comfort and fit, participants were asked to perform an initial baseline of activities that included drawing on a tablet device and performing a simple web search on a computer. Immediately afterwards, they were offered a small snack of either nuts or fruits and were asked where they wanted to go for lunch. Lunch options were determined by the availability of restaurants around our university campus and included Chipotle, Panda Express, Shake Shack, Wendy's, Chick-Fil-A and Taco Bell. After lunch at one of these establishments, participants were offered to walk to an ice-cream or frozen yogurt

Fig. 6. The average duration of lab study was 37.8 minutes, and 20 minutes (52.9% of total duration) were spent on eating. For the lab study, the average time spent on appetizer, main dish, and dessert were 4.6 minutes (12.2%), 9.4 minutes (24.9%), and 6 minutes (15.9%), respectively. In controlled field study, the duration information for P5 was failed to be saved, and for P4 the duration information for walk-2 and dessert was missing. P3 refused to get dessert. For the controlled field study, the average duration of total activities was 59.5 minutes, and 26.4 minutes (44.5% of total duration) were spent on eating. The average time spent on appetizer, main dish, and dessert were 3.4 minutes (5.8%), 15.1 minutes (25.4%), and 7.9 minutes (13.3%), respectively.

place and have dessert. Throughout the study, the experimenter accompanied participants and made annotations of all activities in real-time.

The controlled field study was a pseudo-wild study in the sense that every activity took place in the wild. However, due to the natural context provided in the controlled field study, some participants (e.g. P7) initiated and engaged in a long conversation with the researcher after having only a few bites of food (see Figure 6). Whenever participants attempted to get involved in a long conversation longer than 5 minutes without eating, they were reminded of the study and asked to focus on eating.
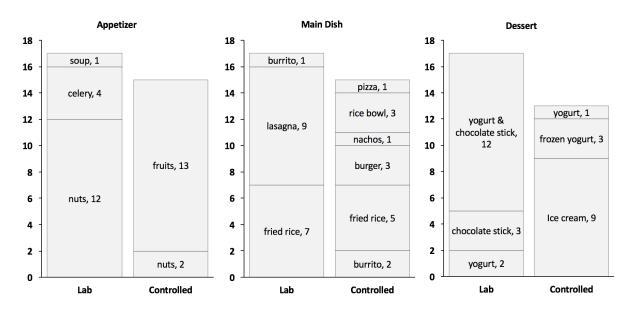
Fig. 7. The composition of foods served for lab study and controlled field study is shown in the figure. The number of servings is shown in the y-axis of the bar chart. The served food is written in each segment of bar along with the number of servings for each food. The bar height for dessert from controlled field study is shorter since two participants (P3 and P4) refused to have desserts. In controlled field study, the menus for appetizer and main dish were prepared such that eating would require significant amount of chewing while the menus for desserts were prepared with those that require minimal chewing.

Table 3. A variety of foods were consumed in the wild. P1 and P11 failed to have a meal session in the wild, and thus, were excluded from the table.

Participant	P2	P3	P4	P5	P6	P7	P8	P9	P10	P12	P13	P14	P15
2**Food**	Pasta	Burger Cookies	Sandwich	Bread	Fried Rice Stir Fried Tofu	Pasta Cookies	Chicken Salad	Salad Apple	Rice Curry	Rice Bowl	Rice Soup	Pasta	Salad Apple

6.3 In-the-Wild Study

To evaluate our system and approach in real-world settings, we conducted a 1-day in-the-wild study with 19 participants. While wearing the eating detection device, which was given to them in the morning, participants were free to perform their normal everyday activities for the rest of the day. Each participant was given an Android phone and asked to annotate the beginning and end of any eating activities. Additionally, they were instructed to make a note of what they ate. Two participants (P1 and P11) reported that they failed to eat a meal during the study and thus their data was excluded from the analysis. The remaining seventeen participants reported that they had at least one meal. The average duration of the in-wild study was 4 hours and 38 minutes per participant; the shortest duration was 2 hours and the longest was 8 hours and 48 minutes. The types of food consumed in this study can be seen in Table 3.
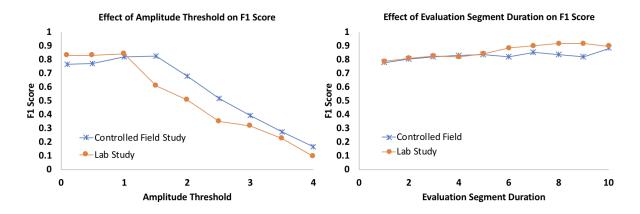
Fig. 8. The difference in plots for the lab study and the controlled field study is due to the difference in necklace design, which resulted the necklace for the lab study being placed consistently closer to the chin. The figure on the right shows the average F1 score as evaluation segment duration varied from 1 to 10 minutes. The amplitude threshold used for this analysis was 1.5 for controlled field and wild study, and 1.0 for lab study. For each evaluation segment duration, the minimum proportion of eating required within each evaluation segment duration for eating detection was 10%.

7 RESULTS

7.1 Laboratory Study Results

Of the twenty participants recruited in the lab study, three participants were excluded from the analysis due to sensor misalignment. A 10-minute evaluation segment with 10% overlap and an amplitude threshold of 1.0 resulted in precision of 91.2% and recall of 92.6%.

7.2 Controlled Field Study Results

For the controlled field study, fifteen participants were recruited. The best result was obtained with precision of 95.2% and recall of 81.9% when the evaluation segment was 10 minutes long with overlap set to 10% and amplitude threshold set to 1.5.
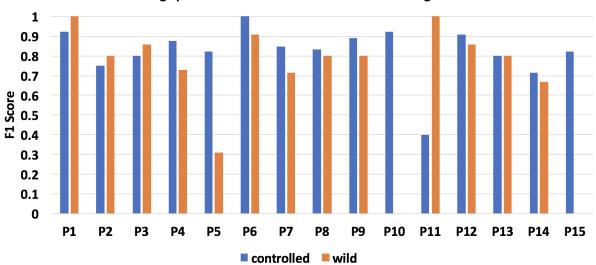
7.3 In-the-Wild Study Results

The fifteen participants from the controlled field study participated in the in-the-wild study. In this study, eating episode detection was achieved with a precision of 78.2% and a recall of 72.5% with the evaluation segment set to 10 minutes. The amplitude threshold of 1.5 and 10% overlap were used for the analysis.

8 DISCUSSION

As expected, the results demonstrated that eating episode detection performance drops as the duration of the evaluation segment decreases. This is because there are fewer instances of chewing behavior in shorter segments, which reduces the ability of the classifier to discriminate eating from non-eating activities. Figure 8 illustrates the effect of evaluation segment on eating episode detection performance and also the *amplitude threshold* setting.

For the lab study, the average F1 score was consistently greater than 0.8 for amplitude thresholds 0.1, 0.5 and 1.0 and started to decrease after 1.0. We set the amplitude threshold of 1.0 for the lab study analysis. The amplitude threshold was varied from 0.1 to 4.0 for the fifteen participants from the controlled field study and wild study, and for the seventeen participants from the lab study. For the controlled field study, the average F1 score increased

Eating Episode Detection at 5-Minute Evaluation Segment

Fig. 9. F1 scores for all participants from the controlled field study and the wild study. The evaluation segment was 5 minutes, the amplitude threshold was 1.5, and the minimum required proportion of eating in each evaluation segment duration was 20%. P1 and P11 failed to have meal in the wild and their wild F1 score was calculated as 1. For P10 and P15, the system failed to capture the eating event in the wild. The average F1 score of controlled field study was 0.83 (0.88 precision and 0.79 recall) and the average F1 score of wild study was 0.76 (0.83 precision and 0.70 recall). The average F1 score of wild study without P1 and P11 was 0.72 (0.80 precision and 0.65 recall).

as the amplitude threshold was increased from 0.1 to 1.5. However, as the amplitude threshold was increased beyond 1.5, the average F1 score began to drop. Based on this result, we set the amplitude threshold at 1.5 for the analysis of data for the controlled field study and the wild study. This difference in optimal amplitude thresholds between lab study and controlled field study is due to the different mechanical designs of the necklaces employed in the studies; after the lab study, sensor mount and rubber band were introduced to the design for ease of use and consistent positioning of the necklace. In the earlier version of the necklace where the proximity sensor was solely suspended by the tension of electrical wires without any support of rubber band and sensor mount, the proximity sensor was consistently placed closer to the chin than that of the second version of the necklace. This resulted in smaller amplitudes in signals, which led to smaller optimal amplitude threshold value in laboratory study than that of the controlled field study.

8.1 Food Type and Drinking Episodes

Our approach to eating episode detection hinges on the identification of chewing, first and foremost. However, some legitimate food types such as ice cream, soup, and yogurt are perceived to require very little chewing. To understand how our system performed with these so-called soft-textured foods, we compared the recognition performance of eating episodes of the controlled field study where participants ate soft-textured foods against the one that involved all the other foods.

For the analysis, the evaluation segment was set to 1 minute to account for the short duration of dessert eating episodes, whose average length was 7.9 minutes. The same setting was applied for the non-soft textured food

episodes. Additionally, an amplitude threshold of 1.5 and 70% overlap in eating segment were applied. To restate, this overlap meant that for every non-overlapping sequence of 1-minute evaluation segments, if each segment contained more than 42 seconds of eating (i.e., 70% of 1 minute), the system inferred that eating occurred in the segment.

The average F1 score for soft textured food was 67% (sd=34%), and the average F1 score for non-soft textured food was 73% (sd=28%). As expected, the lower F1 score for soft textured food suggests that it is indeed harder to identify eating when less chewing occurs. But the difference of only 6% in F1 score between these two conditions is encouraging. Another measure of interest is the standard deviation of the F1 scores across participants. Its high values shine light on the amount of intra-class variability in eating episode detection when data from multiple individuals is considered.
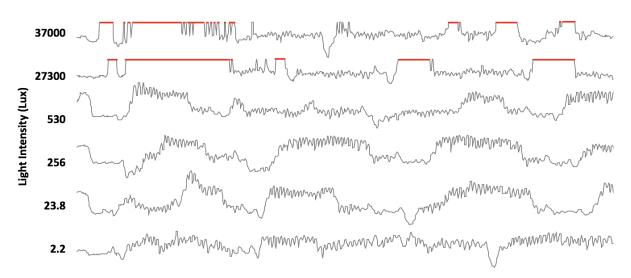
When it comes to understanding the extent to which our wearable system performs across a variety of dietary intake scenarios, it is also natural to question whether our approach is effective for drinking detection. This is important because there has been growing interest in the health research community over the last few years in passively monitoring hydration patterns. While not a focus of this study, we were able to observe characteristic signals that may be used to identify drinking moments. Drinking gave rise to a significant increase in the sensor distance reading as the head is tilted backward and the jawbone is moved further away from the sensor. This increased distance was observed for a short period of time before it rapidly fell back to the baseline. However, we also notice that the duration of this effect was different from occasion to occasion as drinking patterns varied. Furthermore, the degree of tilting was different from person to person. We are currently developing a study design and methodology so that we can investigate the problem of drinking episode detection systematically and rigorously in the future.

8.2 Effect of Light Intensity on Sensor Data

A core component of our wearable necklace system is the VL6180X sensor. To reiterate, this device is a proximity sensor that calculates distance by measuring the time-of-flight of emitted light. Therefore, when used for eating episode detection, a reasonable question to ask is how much the sensor data is affected by the lighting conditions where the device is likely to be used. To answer this question, we conducted a focused study where we tested our system while varying the light intensity the proximity sensor was exposed to and measuring the intensity of light, as seen in Figure 10. For this study, we used a stand-alone digital lux meter to accurately measure the ambient light intensity. The signal at 2.2 lux was collected in our lab with all the lights turned off. The signal at 23.8 lux was obtained at a different location in the lab where it was slightly brighter due to the light from the outside. The signal at 256 lux was measured with one light switch on, and the signal at 530 lux was measured with two light switches on. The brightness measurement rapidly increased when the measurement was made outside the building where our laboratory is located, under sunny and cloudy conditions: 27300 lux and 37000 lux. In the signals collected outside, it is possible to see that the sensor saturates at times; the flat lines represent that the sensor value reached its maximum. Nonetheless, the sensor was still capable of detecting eating activity, as evidenced by the rhythmic patterns that are also captured. This was due to the fact that during eating, even under high light intensity overhead, the sensor is often shaded from light by the head.

8.3 Mechanical Challenges

Our intention with the wearable system was to make it practical for everyday use and highly wearable. By and large we succeeded at this goal; the device we built is lightweight, requires only a small proximity sensor, and does not require contact with the skin. It is attached to a necklace but otherwise it can move freely around the neck. However, these design requirements resulted in unanticipated challenges when it came to sensing. An

Effect of Ambient Light Intensity on Proximity Sensor Signal Response

Fig. 10. The bottom four plots were collected inside a building. The two plots from the top were collected outside.

important finding from the studies was that the performance of the device is dependent on the orientation of the sensor, since there is only one and it is so small. As expected, if the sensor was not pointing straight at the jawbone, it was not able to collect any data corresponding to jaw movement. This was relatively common in the studies as the sensor moved slightly as participants performed various activities. This was the case with P11 in the controlled field study; chewing was not captured because the sensor was pointing towards the neck more often than the jawbone. In addition, three participants from the wild study who reported to have engaged in sport activities (i.e., badminton and skateboarding) mentioned that the sensor was rotated around the neck after these activities, and they had to re-position the necklace. We plan to address these issues in a future iteration of the device in two ways. First, we will incorporate additional proximity sensors to the system in an array configuration such that a wider sensing area can be explored and the device is not so sensitive to misalignment due to motion. Secondly, we will explore methods and materials to prevent excessive movement of the sensor and facilitate custom fitting even as individuals are performing highly active physical activities. Ultimately, our goal is to reduce the size of the device further so that it can be integrated into existing jewelry and other wearables.

8.4 Mobility Confounds

A limitation we identified in our preliminary studies is that walking can be confused with chewing. This is highlighted further if a person walks in a straight line without turning their head. In other words, walking can give rise to a signal that has the similar characteristic frequency and amplitude as eating, which can result in a high false-positive rate. To circumvent this problem, the accelerometer from the accompanying smartphone was used. Given a certain amount of *energy* in the acceleration vector, it is assumed the individual is walking and not eating. In this case, the sensor data is discarded before it reaches the processing pipeline for efficiency reasons. This approach is simple and performed well in our studies, reducing the rate false positives, but its limitations are clear. For one, it does not account for scenarios where eating takes place while walking. And more importantly, it also fails to consider when a person is eating inside of a moving vehicle, which is a common

Table 4. The survey questions asked at the end of the wild study.

User Survey Questions
Q1. How comfortable was the device in scale from 1-5 with 5 being the most comfortable?
Q2. Were you aware of the device?
Q3. Will you attend a social event with the device on?

How comfortable was the device in a scale from 1-5 with 5 being the most comfortable?
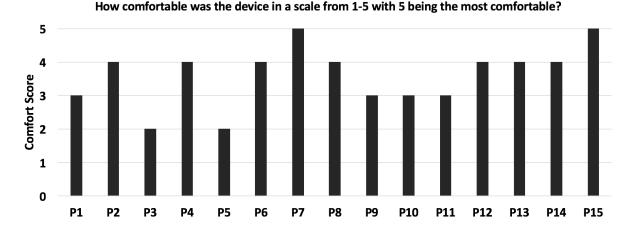

Fig. 11. One individual responded that the device was uncomfortable and two individuals responded that the device was very comfortable. The average score for comfort was 3.6 and it was considered moderately comfortable.

occurrence. Thankfully, there is significant prior work focused on the identification of transportation modes with smart phones sensors [22, 50]. In the future, we plan to incorporate these approaches into the system so that it can detect eating episodes even when acceleration is observed; the individual might be inside of a car or public transportation and having a snack.

8.5 Wearability and Comfort

The fifteen participants from the controlled field study and wild study were asked about their experience with the device at the conclusion of the experiments. The questions asked are shown in the Table 4. The comfort of the device was asked in scale from 1 to 5 with 5 being the most comfortable. The average comfort score from the fifteen participants was 3.6 and the standard deviation was 0.91 (See Figure 11). The maximum comfort value was 5 and the minimum comfort score was 2. The second question about whether the participants were aware of the device during the study. 86.7% of the fifteen participants responded that they were aware of the device throughout the study. The third question was about whether the participants were willing to attend a social event with the device on. 73.3% of the fifteen participants responded that they would not mind attending a social event with the device on.

9 CONCLUSION

We introduce a wearable device for eating episode detection that operates by capturing jawbone movements through a distance measure obtained with a proximity sensor. Compared to previous neck-based dietary monitoring techniques, a highlight of the proposed approach is that the sensor does not need to be in direct contact with skin. To test our system, we conducted an evaluation with a total of 32 participants comprised of three phases, a controlled laboratory study, a controlled field study, and an in-the-wild study. We were able to capture eating episodes with a precision of 91.2 % and recall of 92.6 % in the lab study and with the a precision of 95.2% and recall of 81.9% in the controlled field study. For the in-the-wild study, we obtained 78.2% and 72.5% for precision and recall, respectively. Although the system we developed has limitations, our studies have informed us about ways in which the identified shortcomings can be addressed in the future. Finally, to enable and encourage this research direction to be extended by the research community, we will be releasing the sensor data compiled in the studies upon the public availability of this work.

REFERENCES

[1] 2014. *Obesity and overweight*. Technical Report. World Health Organization. Fact Sheet 311.

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

[3] Oliver Amft and Gerhard Tröster. 2009. On-Body Sensing Solutions for Automatic Dietary Monitoring. *IEEE pervasive computing* 8, 2 (April 2009).

[4] Edward Archer, Gregory A Hand, and Steven N Blair. 2013. Validity of U.S. Nutritional Surveillance: National Health and Nutrition Examination Survey Caloric Energy Intake Data, 1971–2010. *PLoS ONE* 8, 10 (Oct. 2013), 76632.

[5] Yicheng Bai, Wenyan Jia, Zhi-Hong Mao, and Mingui Sun. 2014. Automatic eating detection using a proximity sensor. In *Bioengineering Conference (NEBEC), 2014 40th Annual Northeast*. IEEE, 1–2.

[6] Ronald J Baken. 1992. Electroglottography. *Journal of Voice* 6, 2 (1992), 98–110.

[7] Abdelkareem Bedri, Richard Li, Malcolm Haynes, Raj Prateek Kosaraju, Ishaan Grover, Temiloluwa Prioleau, Min Yan Beh, Mayank Goel, Thad Starner, and Gregory Abowd. 2017. EarBit: Using Wearable Sensors to Detect Eating Episodes in Unconstrained Environments. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (2017), 37.

[8] 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.

[9] Sheila A Bingham. 1987. *The dietary assessment of individuals; methods, accuracy, new techniques and recommendations*.

[10] Hsin-Chen Chen, Wenyan Jia, Yaofeng Yue, Zhaoxin Li, Yung-Nien Sun, John D Fernstrom, and Mingui Sun. 2013. Model-based measurement of food portion size for image-based dietary assessment using 3D/2D registration. *Measurement Science and Technology* 24, 10 (2013), 105701.

[11] Jingyuan Cheng, Oliver Amft, and Paul Lukowicz. 2010. Active capacitive sensing: Exploring a new wearable sensing modality for activity recognition. In *International Conference on Pervasive Computing*. Springer, 319–336.

[12] Jingyuan Cheng, Bo Zhou, Kai Kunze, Carl Christian Rheinländer, Sebastian Wille, Norbert Wehn, Jens Weppner, and Paul Lukowicz. 2013. Activity recognition and nutrition monitoring in every day situations with a textile capacitive neckband. In *Proceedings of the ACM conference on Pervasive and ubiquitous computing adjunct publication*. ACM Press, New York, New York, USA, 155.

[13] N V Dhurandhar, D Schoeller, A W Brown, S B Heymsfield, D Thomas, T I A Sorensen, J R Speakman, M Jeansonne, and D B Allison. 2014. Energy balance measurement: when something is not better than nothing. *International Journal of Obesity* (Nov. 2014).

[14] Yujie Dong. 2012. Tracking Wrist Motion to Detect and Measure the Eating Intake of Free-Living Humans. *Thesis (Ph. D.) Clemson University* (May 2012), 1–106.

[15] Yujie Dong, Jenna Scisco, Mike Wilson, E Muth, and A Hoover. 2013. Detecting periods of eating during free living by tracking wrist motion. *IEEE Journal of Biomedical Health Informatics* (Sept. 2013).

[16] Martin Ester, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu, and others. 1996. A density-based algorithm for discovering clusters in large spatial databases with noise.. In *Kdd*, Vol. 96. 226–231.

[17] Muhammad Farooq, Juan M Fontana, and Edward Sazonov. 2014. A novel approach for food intake detection using electroglottography. *Physiological measurement* 35, 5 (2014), 739.

[18] Muhammad Farooq and Edward Sazonov. 2015. Comparative testing of piezoelectric and printed strain sensors in characterization of chewing. In *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*. IEEE, 7538–7541.

[19] Muhammad Farooq and Edward Sazonov. 2016. Detection of chewing from piezoelectric film sensor signals using ensemble classifiers. In *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the.* IEEE, 4929–4932.

[20] Muhammad Farooq and Edward Sazonov. 2016. A novel wearable device for food intake and physical activity recognition. *Sensors* 16, 7 (2016), 1067.

[21] Vay Liang W Go, Christine T H Nguyen, Diane M Harris, and Wai-Nang Paul Lee. 2005. Nutrient-gene interaction: metabolic genotype-phenotype relationship. *The Journal of nutrition* 135, 12 Suppl (Dec. 2005), 3016S–3020S.

[22] Samuli Hemminki, Petteri Nurmi, and Sasu Tarkoma. 2013. Accelerometer-based Transportation Mode Detection on Smartphones. In *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems (SenSys '13).* ACM, New York, NY, USA, Article 13, 14 pages. DOI:http://dx.doi.org/10.1145/2517351.2517367

[23] D R Jacobs. 2012. Challenges in research in nutritional epidemiology. *Nutritional Health* (2012), 29–42.

[24] Wenyan Jia, Yaofeng Yue, John D Fernstrom, Zhengnan Zhang, Yongquan Yang, and Mingui Sun. 2012. 3D localization of circular feature in 2D image and application to food volume estimation. In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE.* IEEE, 4545–4548.

[25] Holger Junker, Oliver Amft, Paul Lukowicz, and Gerhard Tröster. 2008. Gesture spotting with body-worn inertial sensors to detect user activities. *Pattern Recognition* 41, 6 (June 2008), 2010–2024.

[26] H Kalantarian, N Alshurafa, and M Sarrafzadeh. 2014. A Wearable Nutrition Monitoring System. In *Wearable and Implantable Body Sensor Networks (BSN), 2014 11th International Conference on.* 75–80.

[27] Nathaniel Kleitman. 1963. *Sleep and wakefulness.* The University of Chicago Press, Chicago.

[28] Flegal KM, Graubard BI, Williamson DF, and Gail MH. 2005. Excess deaths associated with underweight, overweight, and obesity. *JAMA* 293, 15 (2005), 1861–1867. DOI:http://dx.doi.org/10.1001/jama.293.15.1861 arXiv:/data/Journals/JAMA/4972/JOC50018.pdf

[29] N Lifson and Ruth McClintock. 1966. Theory of use of the turnover rates of body water for measuring energy and material balance. *Journal of theoretical biology* 12, 1 (1966), 46–74.

[30] Jindong Liu, E Johns, L Atallah, C Pettitt, B Lo, G 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 Computer Society, 154–160.

[31] Oleksandr Makeyev, Paulo Lopez-Meyer, Stephanie Schuckers, Walter Besio, and Edward Sazonov. 2012. Biomedical Signal Processing and Control. *Biomedical Signal Processing and Control* 7, 6 (Nov. 2012), 649–656.

[32] Christopher Merck, Christina Maher, Mark Mirtchouk, Min Zheng, Yuxiao Huang, and Samantha Kleinberg. 2016. Multimodality Sensing for Eating Recognition. *Proceedings of PervasiveHealth* (March 2016), 1–8.

[33] K B Michels. 2001. A renaissance for measurement error. *International journal of epidemiology* 30, 3 (June 2001), 421–422.

[34] 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.

[35] Sebastian Päßler, Matthias Wolff, and Wolf-Joachim Fischer. 2012. Food intake monitoring: an acoustical approach to automated food intake activity detection and classification of consumed food. *Physiological Measurement* 33, 6 (May 2012), 1073–1093.

[36] JMC Po, JA Kieser, LM Gallo, AJ Tésenyi, P Herbison, and M Farella. 2011. Time-frequency analysis of chewing activity in the natural environment. *Journal of dental research* 90, 10 (2011), 1206–1210.

[37] Tauhidur Rahman, Mary Czerwinski, Ran Gilad-Bachrach, and Paul Johns. 2016. Predicting "About-to-Eat" Moments for Just-in-Time Eating Intervention. In *DH '16: Proceedings of the 6th International Conference on Digital Health Conference.* Cornell University, ACM.

[38] Ahmad Safari and E Koray Akdogan. 2008. *Piezoelectric and acoustic materials for transducer applications.* Springer Science & Business Media.

[39] Edward Sazonov, Stephanie Schuckers, Paulo Lopez-Meyer, Oleksandr Makeyev, Nadezhda Sazonova, Edward L Melanson, and Michael Neuman. 2008. Non-invasive monitoring of chewing and swallowing for objective quantification of ingestive behavior. *Physiological Measurement* 29, 5 (April 2008), 525–541.

[40] Schoeller. 1995. Limitations in the assessment of dietary energy intake by self-report. *Metabolism* 44 (Feb. 1995), 5–5.

[41] Donna Spruijt-Metz and Wendy Nilsen. 2014. Dynamic models of behavior for just-in-time adaptive interventions. *IEEE Pervasive Computing* 3, 13 (2014), 13–17.

[42] E Stellar and E E Shrager. 1985. Chews and swallows and the microstructure of eating. *The American journal of clinical nutrition* 42, 5 (1985), 973–982.

[43] Mingui Sun, Lora E Burke, Zhi-Hong Mao, Yiran Chen, Hsin-Chen Chen, Yicheng Bai, Yuecheng Li, Chengliu Li, and Wenyan Jia. 2014. eButton: a wearable computer for health monitoring and personal assistance. In *Design Automation Conference (DAC), 2014 51st ACM/EDAC/IEEE.* IEEE, 1–6.

[44] Edison Thomaz, GD Abowd, and Irfan Essa. 2015. A Practical Approach for Recognizing Eating Moments with Wrist-Mounted Inertial Sensing. In *UbiComp '15: Proceedings of the 2015 ACM international joint conference on Pervasive and ubiquitous computing.* 1–12.

[45] Edison Thomaz, Irfan Essa, and Gregory D Abowd. 2015. A practical approach for recognizing eating moments with wrist-mounted inertial sensing. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing.* ACM, 1029–1040.

[46] Edison Thomaz, Cheng Zhang, Irfan Essa, and Gregory D Abowd. 2015. Inferring Meal Eating Activities in Real World Settings from Ambient Sounds. In *the 20th Intelligent User Interfaces Conference (IUI)*. ACM Press, New York, New York, USA, 427–431.

[47] D Turner and others. 1940. The estimation of the patient's home dietary intake. *Journal of the American Dietetic Association* 16 (1940), 875–881.

[48] EM Widdowson. 1936. A study of English diets by the individual method: Part I. Men. *Journal of Hygiene* 36, 03 (1936), 269–290.

[49] EM Widdowson and RA McCance. 1936. A study of English diets by the individual method: Part II. Women. *Journal of Hygiene* 36, 03 (1936), 293–307.

[50] Peter Widhalm, Philippe Nitsche, and Norbert Brändie. 2012. Transport mode detection with realistic smartphone sensor data. In *Pattern Recognition (ICPR), 2012 21st International Conference on*. IEEE, 573–576.

[51] Dorothy G Wiehl. 1942. Diets of a group of aircraft workers in Southern California. *The Milbank Memorial Fund Quarterly* (1942), 329–366.

[52] Dorothy G Wiehl and Robert Reed. 1960. Development of new or improved dietary methods for epidemiological investigations. *American Journal of Public Health and the Nations Health* 50, 6_Pt_1 (1960), 824–828.

[53] Koji Yatani and Khai N Truong. 2012. BodyScope: a wearable acoustic sensor for activity recognition. *UbiComp '12: Proceedings of the 2012 ACM Conference on Ubiquitous Computing* (2012), 341–350.