HeadScan: A Wearable System for Radio-based Sensing of Head and Mouth-related Activities

Biyi Fang[‡], Nicholas D. Lane[†], Mi Zhang[‡], Fahim Kawsar[†]
[‡]Michigan State University, [†]Bell Labs

***Abstract*—The popularity of wearables continues to rise. However, possible applications, and even their raw functionality are constrained by the types of sensors that are currently available. Accelerometers and gyroscopes struggle to capture complex user activities. Microphones and image sensors are more powerful but capture privacy sensitive information. Physiological sensors are obtrusive to users as they often require skin contact and must be placed at certain body positions to function.**

In contrast, radio-based sensing uses wireless radio signals to capture movements of different parts of the body, and therefore provides a contactless and privacy-preserving approach to detect and monitor human activities. In this paper, we contribute to the search for new sensing modalities for the next generation of wearable devices by exploring the feasibility of mobile radio-based human activity recognition. We believe radio-based sensing has the potential to fundamentally transform wearables as we currently know them. As the first step to achieve our vision, we have designed and developed *HeadScan*, a first-of-its-kind wearable for radio-based sensing of a number of human activities that involve head and mouth movements. HeadScan only requires a pair of small antennas placed on the shoulder and collar and one wearable unit worn on the arm or the belt of the user. HeadScan uses the fine-grained CSI measurements extracted from radio signals and incorporates a novel signal processing pipeline that converts the raw CSI measurements into the targeted human activities. To examine the feasibility and performance of HeadScan, we have collected approximate 50.5 hours data from seven users. Our wide-ranging experiments include comparisons to a conventional skin-contact audio-based sensing approach to tracking the same set of head and mouth-related activities. Our experimental results highlight the enormous potential of our radio-based mobile sensing approach and provide guidance to future explorations.

I. Introduction

Wearable devices with various embedded sensors are increasingly becoming mainstream. Millions of people now wear commercial wearable devices, such as those from Fitbit and Jawbone, on a daily basis to track how many steps they take, how far they jog and how long they sleep. However, for wearables to have a much broader impact on our lives, the next generation of wearables must expand beyond the narrow set of predominately exercise-related physical activities and sleep behaviors currently being monitored.

We believe that a key bottleneck preventing a wider set of activities from being tracked is caused by the limitations of sensors present in existing wearable devices. For example, accelerometers and gyroscopes are constrained to only tracking the motion and rotation of the body parts to which they are attached. Microphones and cameras are rich in sensing capabilities but come with severe privacy concerns. Many physiological sensors such as respiration and electrocardiogram (ECG) sensors must be positioned at certain locations on the body (e.g., chest, neck, head), with some even requiring tight skin contact. The intrusiveness of these contact-based sensors makes people resistant to use them in practice.

Recently, radio-based sensing technologies have drawn significant attention as they provide a contactless and privacy-preserving approach to detect and monitor human activities [1], [2], [3], [4], [5], [6]. These technologies use radio transmitters and receivers deployed in the ambient environment to capture human activities that occur in the monitored area. Furthermore, many of them exploit the fine-grained PHY layer Channel State Information (CSI) extracted from the radio signals to infer the specific human activity. Compared to the traditional MAC layer Received Signal Strength Indicator (RSSI), that only provides a single power measurement of the received signal averaged over the entire channel bandwidth, CSI provides both the amplitude and phase measurements of the received signal for all the OFDM subcarriers within the channel bandwidth [7]. As such, by taking advantage of the fine-grained CSI measurements, radio-based sensing technologies have shown superior capabilities of capturing and recognizing not only activities that involve intense full body movements, such as walking [2] and falling down [3], but also activities that involve minute movements such as speaking (i.e., mouth movements) [4], typing on keyboards (i.e., hand/finger movements) [5], and even breathing and heartbeat (i.e., chest movements) [6].

Inspired by the contactless and privacy-preserving characteristics provided by radio-based infrastructure sensing, as well as the wide range of human activities that can be captured by the fine-grained CSI measurements, *we aim to contribute to the search for new sensing modalities for the next generation of wearable devices by exploring the feasibility of mobile radio-based human activity recognition*. We envision that wearable devices equipped with radio-based sensing could provide a comprehensive understanding of a wide range of human activities in a non-intrusive and privacy-preserving manner and thus have the significant potential to be widely adopted.

As the first step to achieve our vision, in this paper, we present *HeadScan*, a wearable system for radio-based sensing of a number of human activities that involve head and mouth movements. These activities include eating, drinking, coughing, and speaking. We target these four activities for two reasons. First, the continuous monitoring of these activities has considerable value for a wide variety of applications,

978-1-5090-0802-5/16/$31.00

such as those related to health and wellbeing [8] and social computing. Second, state-of-the-art wearable technologies for sensing these activities use microphones [9], [10], cameras [11], [12], [13], or capacitive sensors [14] and thus are either intrusive or have privacy concerns. In contrast, HeadScan uses wireless radio signals to sense the targeted activities and therefore provides a non-intrusive and privacy-preserving solution that overcomes the drawbacks of the existing wearable technologies. Specifically, HeadScan consists of two small unobtrusive commercial off-the-shelf (COTS) 5GHz antennas placed on the shoulder and collar of the user respectively, as well as one wearable unit that can be worn on the arm or the belt of the user. One antenna acts as a radio transmitter that continuously sends radio signals, while the other antenna acts as a radio receiver that continuously receives radio signals. The radio signals captured at the receiver are then forwarded to the wearable unit for storage and processing. The underpinning principle behind HeadScan is that movements of head and mouth caused by different targeted human activities generate different changes on the received radio signals. By analyzing those changes, the activity that causes the changes can be recognized. Based on this principle, HeadScan first extracts the CSI measurements from the received radio signals. It then incorporates a radio signal processing pipeline that converts the raw CSI measurements into the targeted human activities. To examine the feasibility and performance of HeadScan, we have conducted extensive experiments. Our experimental results show, for the first time, that radio as a sensing modality is feasible to sense and recognize the targeted eating, drinking, coughing, and speaking activities in the wearable setting with an overall recognition accuracy of 86.3%. Furthermore, we also highlight there remain a number of limitations of this approach. In this respect, our results not only demonstrate the feasibility and performance of radio-based sensing on wearable devices but also serve to guide future explorations by highlighting the areas of difficulty.

In sum, our work makes the following contributions:

- We introduce a first-of-its-kind wearable device that uses radio as a new sensing modality to detect and recognize human daily activities that involve head and mouth movements including eating, drinking, coughing, and speaking.
- We have designed a novel radio signal processing pipeline that converts raw CSI measurements into targeted human activities. The pipeline blends techniques from activity recognition (i.e., sliding window-based segmentation), CSI-based radio signal processing (i.e., PCA-based dimension reduction) and sparse representation-based classifier towards enabling radio-based human activity sensing to be noise-resistant in the challenging scenario of wearables.
- We have built a proof-of-concept wearable prototype that incorporates radio-based sensing using CSI measurements via a COTS radio chipset.
- We have conducted comprehensive experiments to examine the feasibility and performance of HeadScan on sensing and recognizing the four head and mouth-related activities. These include a quantitative comparison between our radio-based sensing approach with a conventional skin-contact audio-based sensing method, as well as experiments that examine the impact of a number of key factors on the activity recognition performance of radio-based sensing.

II. Design Considerations

In this section, we first list the design goals that HeadScan aims to achieve. We then describe the unique characteristics of radio-based human activity sensing in a wearable setting.

A. Design Goals

The design of our HeadScan wearable aims to achieve the following goals:

- **Automated Tracking of Head and Mouth-related Human Activities:** HeadScan is designed to be able to automatically tracking a set of head and mouth-related human daily activities including eating, drinking, coughing, and speaking. Automatically tracking these activities has significant value for applications related to healthcare and wellbeing [8], as well as social computing. Current practice requires users to self-report those activities via either journaling or taking pictures of their food. HeadScan aims to automatically track those activities without user involvement using a wearable system.
- **Enabling Contactless and Non-intrusive Sensing:** HeadScan is designed to be able to provide a contactless and non-intrusive solution to capture the movements of the head and mouth caused by the targeted human activities. Accelerometer and gyroscope-based solutions require users to wear two wearables (i.e., one on each wrist) to capture eating and drinking activities [15]. Microphone-based solutions such as [9], [10] suffer if they do not maintain close contact with the skin. Capacitive sensor-based solutions such as [14] require users to wear the capacitive sensor around the neck. HeadScan aims to examine radio-based sensing that neither relies on skin contact nor requires the radio sensor to be placed at a specific location on the body, thus providing a truly contactless and non-intrusive solution.
- **Protection of Privacy:** HeadScan is designed to be able to protect privacy while still being able to provide rich information about the targeted human activities. Microphones and cameras are the most commonly used sensing modalities in existing wearables to track the targeted specific activites we target [11], [9], [10], [12], [13]. However, these powerful sensors also bring unwanted negative privacy side effects (e.g. capturing speech and images of surrounding people). HeadScan aims to examine radio-based sensing that is unable to reconstruct speech (e.g., [16]) or images yet still has the resolution to capture and recognize eating, drinking, coughing, and speaking activities.

To the best of our knowledge, no currently used wearable technology can satisfy all of the above criteria. This motivates us to design and develop HeadScan to fill this critical gap.

B. Unique Characteristics of Wearable Radio-based Sensing

Most of the existing research on radio-based human activity sensing considers monitoring users from radio antennas deployed in the ambient indoor environments. Performing radio-based human activity sensing from wearable devices with antennas located on the human body presents a number of unique characteristics that are different from the infrastructure-based scenarios. We summarize these differences below:

- **Sensing Distance:** In infrastructure-based scenarios, radio transmitters and receivers are placed at a distance (typically a few meters) from the user being monitored. In a wearable setting, the antennas are placed directly on the body of the user. As such, the radio signals are much more sensitive to the movements caused by human activities due to the short sensing distances. This high sensitivity amplifies the movements of interests as well as unwanted noise.
- **Radio Transmitter and Receiver Location:** The deployment locations of radio transmitters and receivers, under infrastructure-based scenarios, have a significant impact on the activity sensing and recognition performance due to the variability of indoor conditions (e.g., fittings, furniture) found in different locations [1]. In a wearable setting, although the antennas are placed on the body, they can be placed at different on-body locations as needed. The choice of on-body location plays a key role in determining what type of body movement, or more broadly activity, can be recognized and at what level of accuracy.
- **Non-user Interference:** In infrastructure-based scenarios, radio signals can be contaminated by radio frequency interference (RFI) as well as interference caused by other people in the monitored area. In a wearable setting, although both sources of interferences exist, those interferences are often dwarfed by the activity performed by the user since the antennas are worn on the body of the user.

Examining the unique characteristics of radio-based sensing under a wearable scenario, such as those listed above, are at the heart of this work. We have developed radio signal processing techniques at different stages of the pipeline to filter out unwanted wearable-centric noise. We have also adopted empirical methods through extensive experiments to identify the best on-body locations to wear antennas as well as quantify the impact of non-user interference on activity sensing and recognition performance.

III. Wearable Prototype

Our HeadScan wearable prototype consists of two small (21 x 21 mm; 5 g) unobtrusive commercial off-the-shelf (COTS) 5GHz antennas, that can be worn on the shoulder and the collar respectively, as well as one wearable unit (85.0 x 60.0 x 36.0 mm; 390 g) that can be worn on the arm or the belt of the user. The two antennas are wired to the wearable unit. Figure 1 provides a conceptual illustration of how HeadScan is worn while it tracks the movements of head and mouth of the user.

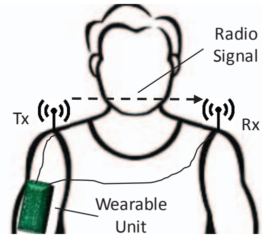


Fig. 1: The illustration of HeadScan in operation.

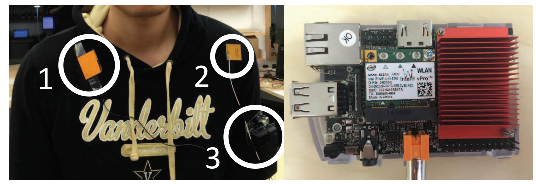


Fig. 2: Our HeadScan wearable prototype consists of: 1) one transmitter antenna (Tx), 2) one receiver antenna (Rx) and 3) one wearable unit that contains two HummingBoard Pro (HMB) mini-computers (Left). Each HMB is connected with one Intel WiFi Link 5300 card for measuring CSI (Right).

The wearable unit consists of two HummingBoard Pro (HMB) devices [17], each powered by one 3000 mAH battery. The HMB is a low-cost ARM-based mini-computer that contains an on-board 1.2 GHz ARM Cortex-A9 processor and 1GB RAM. Each HMB is connected with an antenna as well as an Intel WiFi Link 5300 card [18] via the micro PCIe socket for measuring CSI (Figure 2). In our prototype, two HMB are needed because a single HMB can not support two Intel WiFi Link 5300 cards at the same time. One HMB acts as the radio transmitter that sends packets periodically and the other HMB acts as the radio receiver that continuously receives the packets. We installed the modified firmware released by the CSI measurement tool [19] through the Debian Cubox-i Linux OS running on the receiver HMB, enabling the Intel WiFi Link 5300 card to extract CSI measurements from the received packets on the receiver HMB. During operation, we used a custom case to hold the wearable unit and attach it to the arm of the user with an elastic armband for data collection. The collected CSI measurements are exported to a desktop computer for offline processing.

IV. Radio Signal Processing Pipeline

In this section, we first provide the background knowledge of Channel State Information (CSI). We then describe the details of our radio signal processing pipeline that converts the raw CSI measurements into the targeted human activities.

A. Background of Channel State Information (CSI)

Modern wireless devices that support IEEE 802.11a/g/n/ac standards use Orthogonal Frequency Division Multiplexing (OFDM) in the PHY layer [7]. OFDM divides the entire wireless channel into multiple narrowband subcarriers, with each subcarrier having a different carrier frequency. The Channel State Information (CSI) contains both the amplitude frequency response and phase frequency response at the granularity of

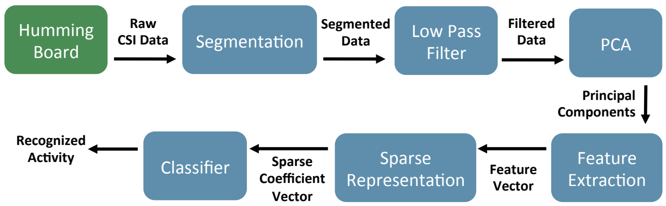
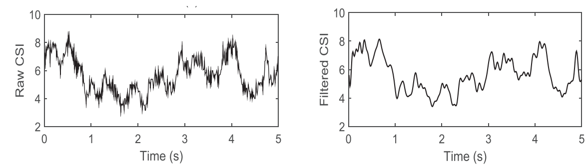


Fig. 3: Overview of the radio signal processing pipeline of HeadScan.



(a) Raw CSI measurements (b) Filtered CSI measurements

Fig. 4: Raw and low-pass filtered CSI measurements.

each subcarrier [20]. Head and mouth movements affect the propagation of radio signals via reflections, diffractions and scattering, and thus cause different amplitude and phase frequency responses at each subcarrier. Such fine-grained differences at the subcarrier level are captured in the CSI measurements, but are smoothed out by the RSSI measurement which is a single measurement averaged over the entire channel [21]. As such, analyzing the fine-grained CSI measurements at each subcarrier provides a great opportunity to capture the minute head and mouth movements caused by eating, drinking, coughing and speaking activities. Therefore, HeadScan takes advantage of the fine-grained CSI measurements provided by commercial wireless devices to sense and recognize those activities. It should be noted that the CSI phase measurements are reported to be unreliable due to the low-cost hardware components used in commercial wireless devices [22]. As such, HeadScan only uses the CSI amplitude measurements at all the subcarriers.

B. Radio Signal Processing Pipeline Overview

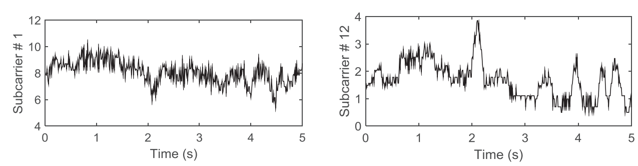
Figure 3 provides an overview of the radio signal processing pipeline of HeadScan. As illustrated, the pipeline consists of six stages to convert the raw CSI measurements into the targeted activities. Specifically, in stage one, a sliding window is used to segment the streaming raw CSI measurements extracted from the PHY layer radio signals into a sequence of fixed-length windows. In this work, the window is set to be five seconds long. The window is long enough so that all the important information of each targeted activity is contained inside each window. In stage two, a low-pass filter is applied to the raw CSI measurements in each window to remove noise that is not caused by head and mouth movements of the targeted activities. In stage three, Principal Component Analysis (PCA) is performed on the filtered CSI measurements of all subcarriers to extract the most sensitive changes caused by head and mouth movements. In stage four, features that capture the unique characteristics of each activity are extracted. In stage five, the noise-robust sparse representation framework is used to train an overcomplete dictionary [23]. Finally, in stage six, the sparse coefficient vector extracted from the sparse representation framework is used to build a classifier to recognize the targeted activity. We use the sparse representation framework because it has been proved to be robust to the noise caused by radio frequency interference (RFI) in radio signal processing [2].

C. Noise Removal

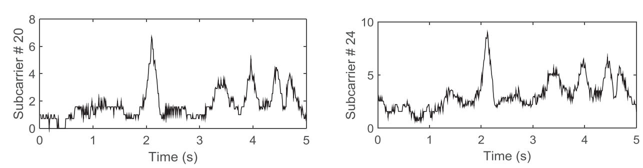
The raw CSI measurements extracted from the PHY layer radio signals are very noisy. To use CSI measurements for recognizing head and mouth-related activities, it is necessary to begin by eliminating as much of this noise as possible. Since the changes of CSI measurements caused by head and mouth movements lie at the lower end of the frequency spectrum, HeadScan uses a Butterworth low-pass filter to capture the changes caused by head and mouth movements while discarding noise that lies at the higher end of the frequency spectrum. Based on our experiments, we observed that the changes caused by head and mouth movements of our targeted activities are below 20Hz. Therefore, we set the cut-off frequency of the Butterworth low-pass filter to 20Hz and use it to filter out CSI measurement noise across all subcarriers in the wireless channel. Figure 4(a) and Figure 4(b) show the raw and filtered CSI measurements of subcarrier #12 for the eating activity respectively. As illustrated, the Butterworth low-pass filter successfully filters out high-frequency noise while preserving the head and mouth movement information of the eating activity.

D. PCA-Based Dimension Reduction

Changes of CSI measurements caused by head and mouth movements occur across all the subcarriers in the wireless channel. As an illustration, Figure 5 plots the CSI measurements of four subcarriers. From these plots, we have two key observations. First, we observed that changes of CSI measurements are highly correlated among all the subcarriers. Second, we observed that different subcarriers have different

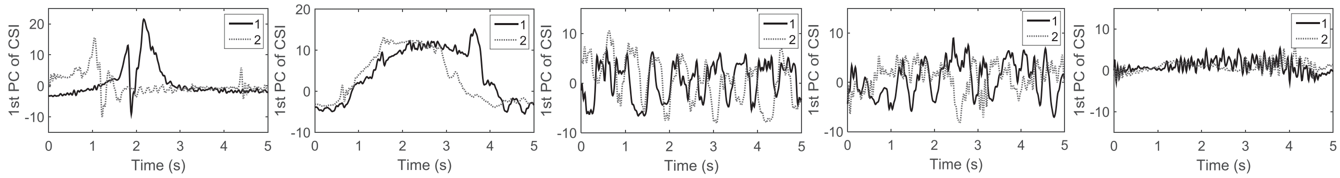


(a) Subcarrier #1 (b) Subcarrier #12



(c) Subcarrier #20 (d) Subcarrier #24

Fig. 5: Correlations and sensitivity differences among different subcarriers.



(a) *Coughing* (b) *Drinking* (c) *Eating* (d) *Speaking* (e) *Idle*

Fig. 6: The first principal component (PC) of the four targeted activities, as well as the idle activity, performed by two subjects.

sensitivity to head and mouth movements. Based on these two observations, in order to remove the redundant information due to the high correlation while still capturing the most sensitive changes caused by head and mouth movements, HeadScan applies Principal Component Analysis (PCA) on the filtered CSI measurements across all the subcarriers to extract principal components (PCs) that capture the dominant changes caused by head and mouth movements. The PCs extracted are ranked by the magnitude of their variance. We select the top two PCs because we empirically find the top two PCs already contain the majority of the total variance and thus preserve most of the information of the targeted activities.

E. Feature Extraction

To recognize our targeted activities, we need to extract features that can capture the differences across them. Figure 6 illustrates the first PC of all four targeted activities as well as the baseline idle activity (i.e., no activity) performed by two subjects. As illustrated, different activities have distinctive shapes and periodicities. Specifically, the shape of coughing exhibits a sharp change in amplitude within a short period of time, which reflects the captured head movement during coughing. The shape of drinking exhibits a sharp increase at the beginning, a plateau in the middle, and a sharp decrease in the end. This is because our head moves up when we start to drink and our head moves back after we finish drinking. For both eating and speaking, we observe the waveforms exhibit periodic changes that reflect the periodic mouth movements when eating and speaking. Between these two periodic activities, eating exhibits a relatively lower frequency than speaking because the speed of mouth movement when eating is slower compare to that of speaking. Furthermore, the shape of idle exhibits the least variance in amplitude and most randomness in frequency due to little head and mouth movements. Finally, we can see that the two samples of the same activity performed by two subjects exhibit high similarity, indicating the consistency across subjects.

Based on these observations, we extract features that characterize the shape and periodicity of the waveforms. Specifically, to capture the shape information, we extract features including mean, median, standard deviation, range, skewness, kurtosis, and interquartile range values of the waveforms. To capture the periodicity information, we extract features including mean crossing rate, dominant frequency, and spectral entropy values of the waveforms. Table I summarizes the list of features. All these features are extracted from the top two PCs and are then concatenated into a single feature vector with the dimensionality of 20.

Mean	Median	Standard Deviation
Range	Skewness	Kurtosis
Dominant Frequency	Interquartile Range	Mean Crossing Rate
Spectral Entropy		

TABLE I: Head and mouth activity features.

F. Training via an Overcomplete Dictionary Construction

As the first step of the sparse representation framework, we construct an overcomplete dictionary using training samples. Let k denote the number of activity classes (k equal to five in this work) and n_i denote the number of training samples from class i, $i \in [1, 2, \dots k]$. Each training sample is represented as an m-dimensional feature vector (m equal to 20 in this work). We arrange the n_i training samples from class i as columns of a data matrix $\boldsymbol{D_i} = [\boldsymbol{x}_{i,1}, \boldsymbol{x}_{i,2}, \dots, \boldsymbol{x}_{i,n_i}] \in \mathbb{R}^{m \times n_i}$. Next, we define a new matrix $\boldsymbol{A}$ which stacks the training samples from all the activity classes as

$$\begin{aligned}\boldsymbol{A} &= [\boldsymbol{D_1}, \boldsymbol{D_2}, \dots, \boldsymbol{D_k}] \in \mathbb{R}^{m \times n} \\ &= [\boldsymbol{x}_{1,1}, \dots, \boldsymbol{x}_{1,n_1}, \boldsymbol{x}_{2,1}, \dots, \boldsymbol{x}_{2,n_2}, \dots, \boldsymbol{x}_{k,1}, \dots, \boldsymbol{x}_{k,n_k}]\end{aligned} \tag{1}$$

where $n = n_1 + n_2 + \ldots + n_k$. Given sufficient training samples from each class, the matrix $\boldsymbol{A}$ can be seen as an overcomplete dictionary. As such, a test sample $\boldsymbol{y}$ from class i can be expressed using the overcomplete dictionary $\boldsymbol{A}$ as

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{\alpha} + \boldsymbol{e} \tag{2}$$

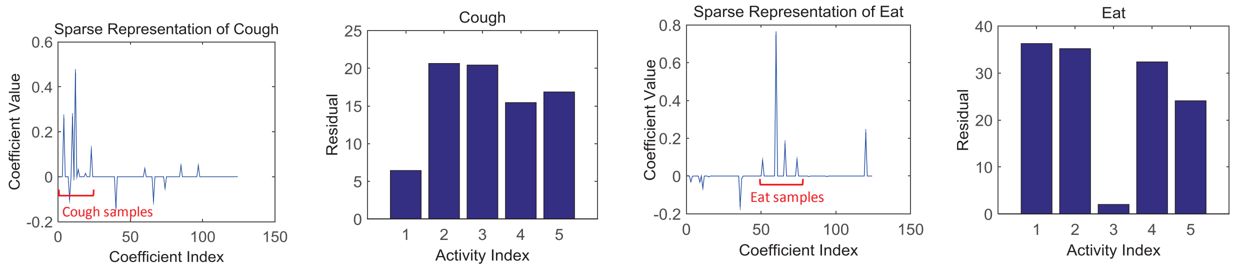
where

$$\boldsymbol{\alpha} = [0, \dots, 0, \alpha_{i,1}, \alpha_{i,2}, \dots, \alpha_{i,n_i}, 0, \dots, 0]^T \tag{3}$$

is a sparse coefficient vector whose entries are zero except those associated with class i, and $\boldsymbol{e}$ is the noise term with bounded energy $\|\boldsymbol{e}\|_2 < \epsilon$. Therefore, $\boldsymbol{\alpha}$ can be regarded as a sparse representation of $\boldsymbol{y}$ based on the overcomplete dictionary $\boldsymbol{A}$. More importantly, the entries of $\boldsymbol{\alpha}$ encode the identity of $\boldsymbol{y}$. In other words, we can infer the class membership of $\boldsymbol{y}$ by finding the solution of equation (2).

G. Classification via ℓ^1 Minimization

The solution of equation (2) depends on the characteristic of the matrix $\boldsymbol{A}$. If $m > n$, $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{\alpha} + \boldsymbol{e}$ is overdetermined and there is one unique solution. However, in most real world applications, the number of prototypes in the overcomplete



(a) The sparse coefficient solution recovered via ℓ^1 minimization for one test sample from *Coughing*.

(b) The residual values with respect to the five activity classes. The test sample is correctly classified as *Coughing* (index 1). The ratio between the two smallest residual values is $1 : 2.62$.

(c) The sparse coefficient solution recovered via ℓ^1 minimization for one test sample from *Eating*.

(d) The residual values with respect to the five activity classes. The test sample is correctly classified as *Eating* (index 3). The ratio between the two smallest residual values is $1 : 11.9$.

Fig. 7: Sparse representation solutions and the corresponding residual values of two test samples from *Coughing* and *Eating*.

dictionary is typically much larger than the dimensionality of the feature representation (i.e., $m << n$). In this scenario, equation (2) is underdetermined and has no unique solution.

Recent research in the field of compressed sensing [24] [23] has shown that if $\boldsymbol{\alpha}$ is sufficiently sparse, it can be recovered by solving the ℓ^1 minimization problem:

$$\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}} \|\boldsymbol{\alpha}\|_1 \qquad \text{subject to} \quad \|\boldsymbol{A}\boldsymbol{\alpha} - \boldsymbol{y}\|_2 \leq \epsilon. \tag{4}$$

where $\|.\|_1$ denotes the ℓ^1 norm. This optimization problem, also known as Basis Pursuit (BP), is built on a solid theoretical foundation and can be solved very efficiently with traditional linear programming techniques whose computational complexities are polynomial in n [25].

Given a new test sample $\boldsymbol{y}$ in the form of an m-dimensional feature vector from one of the k activity classes, we first compute its sparse coefficient vector $\hat{\boldsymbol{\alpha}}$ by solving (4). To identify the class membership of $\boldsymbol{y}$, we compare how well the various parts of the coefficient vector $\hat{\boldsymbol{\alpha}}$ associated with different activity classes can reproduce $\boldsymbol{y}$, where the reproduction error is measured by the residual value. Specifically, the residual of class i is defined as

$$r_i(\boldsymbol{y}) = \|\boldsymbol{y} - \boldsymbol{A}\delta_i(\hat{\boldsymbol{\alpha}})\|_2 \tag{5}$$

where $\delta_i(\hat{\boldsymbol{\alpha}})$ is a characteristic function that selects only the coefficients in $\hat{\boldsymbol{\alpha}}$ associated with class i. Therefore, $r_i(\boldsymbol{y})$ measures the difference between the true solution $\boldsymbol{y}$ and the approximation using only the components from class i. Using the residual value as the classification criterion, the test sample $\boldsymbol{y}$ is classified as the activity class c that generates the smallest residual value:

$$c = \arg\min_i r_i(\boldsymbol{y}) \tag{6}$$

Figure 7(a) and Figure 7(c) illustrate the two coefficient vectors recovered by solving (4) with the noise tolerance $\epsilon = 0.02$ for two test samples: one from coughing and the other from eating. As shown, both of the recovered coefficient vectors are sparse and the majority of the large coefficients are associated with the training samples belonging to the same activity class. Figure 7(b) and Figure 7(d) illustrate the corresponding residual values with respect to the five targeted activity classes. As shown, both test samples are correctly classified since the smallest residual value is associated with the true activity class. To examine the robustness of our residual-based classifier, we calculate the ratios between the two smallest residuals for each test sample. The classification result is robust if the ratio value is large. In the examples of Figure 7(b) and Figure 7(d), the ratios between the two smallest residuals are $1 : 2.62$ for coughing and $1 : 11.9$ for eating. We obtain similar results for the other three activities.

V. Evaluation

A. Experimental Methodology

To examine the feasibility and performance of HeadScan, we conduct extensive experiments. We breakdown the overall evaluation into seven experiments and one usability study. State-of-the-art wearable sensing systems use microphone to sense and recognize eating, drinking, speaking, and coughing [9], [10]. The objective of the first two experiments is to examine the feasibility of radio-based activity sensing and recognition in a wearable setting, as well as to make a quantitative comparison between radio- and audio-based sensing approaches. To achieve this objective, we replicated one state-of-the-art audio-based wearable system called BodyScope from [9] (Figure 8). Since both BodyScope and HeadScan target the same activities, we believe a fair comparison between radio- and audio-based sensing in the wearable setting is achieved. To fully examine the performance in different

Fig. 8: The hardware for audio-based sensing. (1) A chestpiece of a clinical Elite 4000 stethoscope. (2) A unidirectional condenser microphone. (3) The chestpiece attached to skin using an elastic band.

Activity	Radio (time)	Audio (time)
Coughing (C)	52m 40s	28m 40s
Drinking (D)	52m 50s	29m 00s
Eating (E)	72m 30s	28m 30s
Speaking (S)	57m 10s	28m 10s
Idle (I)	54m 20s	28m 20s
Total Amount	289m 30s	142m 40s

TABLE II: List of activities (activity abbreviations) and the amount of labeled radio and audio data.

wearable conditions, the first experiment is conducted in a clean environment, and the second experiment is conducted in a noisy (i.e., with the existence of interferences) environment. Besides comparing performance to audio-based sensing, we conduct another five experiments to examine the impact of various factors on the activity recognition performance of our radio-based sensing approach. These include classification performance of a personalized model and the impact of: training dataset size; radio signal transmission rates; on-body locations of the radio transmitter and receiver; and, the interference caused by nearby people.

B. Human Participants and Data Collection

We recruit seven participants (5 males and 2 females) who volunteer to help collect data and conduct evaluation experiments. The participants are university students and researchers with an age ranging from 20 to 35 (μ = 28.9; σ = 5.2), a weight ranging from 49 kg to 82 kg (μ = 74 kg; σ = 9.8 kg) and are between 158 cm and 182 cm tall (μ = 172 cm; σ = 7.0 cm). The experiments were conducted in a laboratory environment. For experiments that aim to make quantitative performance comparisons between radio and audio sensing modalities, participants wear both HeadScan and the replicated audio-based sensing system to collect both radio and audio data simultaneously. As such, a fair comparison between these two sensing modalities is possible. For experiments that aim to solely examine the impact of various factors on the activity recognition performance of radio-based sensing, participants only wore HeadScan.

During data collection, the participants were instructed to perform a sequence of the four targeted activities: coughing, eating (potato chips), drinking (a cup of water), and speaking (reading a book aloud); as well as being idle (no activity). For radio data collection, by default, participants clip the Tx antenna on one of the collars and the Rx antenna on the shoulder of the other side. For audio data collection, we followed the data collection procedure of the BodyScope study [9] and ask the participants to attach the chestpiece to the skin using an elastic band (Figure 8). In total, approximate 50.5 hours of data was collected. Since data labeling is very time consuming, we labeled 7.2 hours of the collected data and use the labeled data to evaluate HeadScan. Table II lists the five activities and the amount of data labeled for each activity for both radio and audio sensing modalities respectively.

	C	D	E	I	S	Recall
C	65	5	0	0	0	92.9%
D	15	48	7	0	0	68.6%
E	1	6	62	1	0	88.6%
I	1	2	8	59	0	84.3%
S	1	1	0	0	68	97.1%
Precision	78.3%	77.4%	80.5%	98.3%	100%	

TABLE III: The confusion matrix of radio-based activity recognition in a clean environment (see Table II for activity abbreviations).

	C	D	E	I	S	Recall
C	51	8	0	1	10	72.8%
D	19	48	2	1	0	68.5%
E	2	11	55	1	1	78.6%
I	1	6	0	63	0	90.0%
S	0	0	1	0	69	98.6%
Precision	69.8%	64.8%	94.8%	95.5%	86.2%	

TABLE IV: The confusion matrix of audio-based activity recognition in a clean environment (see Table II for activity abbreviations).

C. Performance Comparison between Radio- and Audio-based Wearable Sensing in a Clean Environment

Objective: In this experiment, we evaluate the activity recognition performance of our radio-based wearable sensing system and compare it with a state-of-the-art audio-based approach in a clean environment.

Experimental Setup: We have created a clean environment for participants to collect radio and audio data simultaneously. To create a clean environment for radio data collection, all other wireless devices in the environment are turned off and only the participant was in the environment performing activities. To create a clean environment for audio data collection at the same time, the environment was kept quiet as well. Each participant performs 10 trials for each activity. We examine the activity recognition performance of both radio- and audio-based sensing using leave-one-subject-out validation.

Results and Implication: Figure 9 illustrates the activity classification accuracy of both radio- and audio-based sensing in a clean environment. The average classification accuracy across all participants are 86.3% for radio- sensing and 81.7% for audio-based sensing. To provide a more detailed view of the results, the confusion matrices of radio-based and audio-based sensing with data accumulated from all participants are presented in Table III and Table IV respectively. This result demonstrates that HeadScan is feasible of sensing and recognizing the targeted four activities that involve head and mouth movements. It also shows that our radio-based wearable sensing system HeadScan is able to achieve very competitive recognition performance compared to the audio-based wearable sensing system in a clean environment.

D. Performance Comparison between Radio- and Audio-based Wearable Sensing in a Noisy Environment

Objective: In this experiment, we evaluate the activity recognition performance of our radio-based wearable sensing system and compare it with a state-of-the-art audio-based approach under a noisy environment.

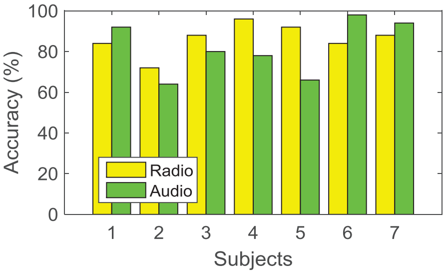


Fig. 9: The activity classification accuracy of radio- and audio-based sensing in a clean environment.

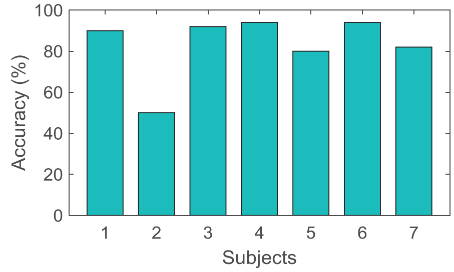


Fig. 10: The activity classification accuracy of a personalized model.

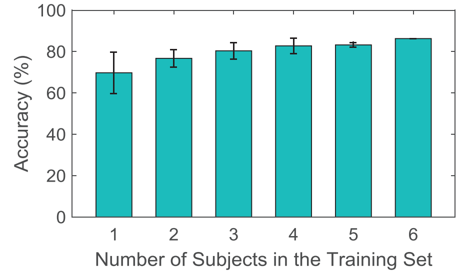


Fig. 11: The impact of the training dataset size on activity classification accuracy.

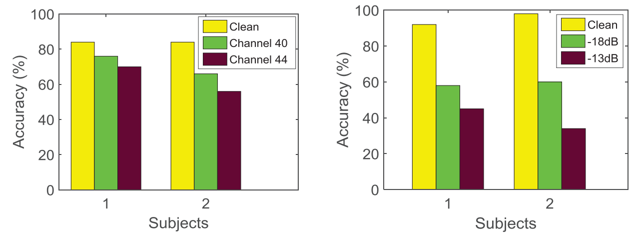
Experimental Setup: We have created a noisy environment for participants to collect the radio and audio data respectively. To create a noisy environment for radio data collection, we generate radio frequency interference (RFI) by collecting in an environment where wireless communication was set up at the same channel that HeadScan uses (channel #44) as well as an adjacent channel (channel #40). We place the Tx and Rx that generates RFI approximately one meter away from the participants and configured the transmission rate at 2 packets per second. To create a noisy environment for audio data collection, we generated audio noise using a computer speaker in the background such that the noise levels we measure at our audio-based wearable sensing system are -18 dB and -13 dB respectively (we use the Audacity software to perform noise level measurements). The noise level at -18 dB is an average measurement in the acoustic environment where students kept talking to each other in a whisper-like voice. The noise level at -13 dB is an average measurement in the acoustic environment where students kept talking to each other at a more natural volume. We examine the activity recognition performance of both radio and audio-based sensing using leave-one-subject-out validation.

Results and Implication: Figure 12(a) and Figure 12(b) illustrate the activity classification accuracy of radio- and audio-based sensing in a noisy environment. For illustrative purposes, only results from the first two participants are shown, with the other five participants having similar results. As shown, both radio and audio sensing modalities are affected by noise but with a different sensitivity. In particular, the performance of radio-based sensing drops from an average of 84% (clean) to 72.2% (channel #40) and 62% (channel #44), a loss of 11.8% and 22% respectively. In comparison, the performance of audio-based sensing drops from an average of 95% (clean) to 59% (-18 dB) and 39.5% (-13 dB), with a loss of 36% and 55.5% respectively.

E. Performance of Personalized Model

Objective: In this experiment, we evaluate the activity recognition performance of a personalized classification model.

Experimental Setup: We use the data of each participant for both training and testing. We examine the activity recognition performance of the personalized model using leave-one-trial-out validation.



(a) Radio (b) Audio

Fig. 12: The activity classification accuracy of radio- and audio-based sensing in a noisy environment.

Results and Implication: Figure 10 illustrates the activity classification accuracy of each personalized model. As illustrated, among seven participants, Subject #2, #5, and #7 achieve slightly inferior performance compared to the general model developed using leave-one-subject-out validation (see Figure 9). This indicates the same activity can be performed very differently even by the same participant. This within-subject variance makes recognition very challenging. On the other hand, in leave-one-subject-out, more training data is used to build the overcomplete dictionary in the sparse representation framework. This indicates that within the sparse representation framework, the classification performance can be improved if more training data is available, even if the training data is collected from other participants.

F. Impact of Training Dataset Size

Objective: In this experiment, we examine the impact of the training dataset size on the activity recognition performance.

Experimental Setup: We set up different training dataset sizes by including data from different numbers of participants. Since we have seven participants, we considered six different sizes by including data from one, two, three, four, five, and six participants. We examined the performance of each size using leave-one-subject-out cross validation.

Results and Implication: Figure 11 illustrates the activity classification accuracy of all six different sizes of training dataset. As shown, the classification performance increases as the size of the training set increases. This result again demonstrates that within the sparse representation framework, the classification performance can be improved if more training data is included.

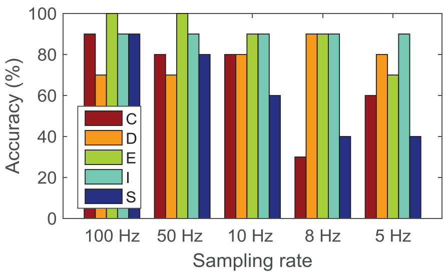


Fig. 13: The impact of radio signal transmission rates on activity classification accuracy.

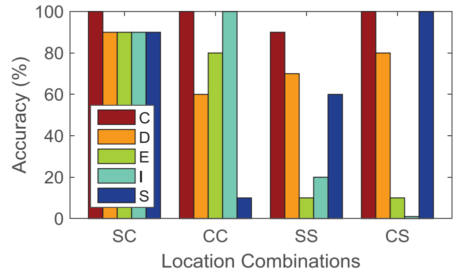


Fig. 14: The impact of radio transmitter and receiver on-body locations on activity classification accuracy.

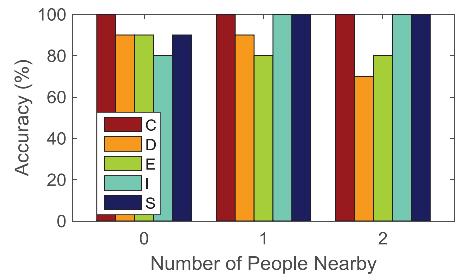


Fig. 15: The impact of interference caused by nearby people on activity classification accuracy.

G. Impact of Radio Signal Transmission Rate

Objective: In this experiment, we examine the impact of radio signal transmission rate on the activity recognition performance. There is a trade-off between transmission rate and the power consumption of the wearable system. Increasing the transmission rate provides higher resolution of the captured head and mouth movements but also increases power consumption. Another drawback of the increase of the transmission rate is that it increases the traffic load of the wireless channel and thus reduces the bandwidth for regular data transmission. The objective of this experiment is to find the optimal transmission rate that achieves a balance between recognition performance and wireless bandwidth.

Experimental Setup: We examine five radio signal transmission rates including 100Hz, 50Hz, 10Hz, 8Hz, and 5Hz. We collect approximately 50 minutes of data at these five transmission rates in total from one participant. For each transmission rate, the participant performs 10 trials for each activity. We examine the performance of each transmission rate using leave-one-trial-out validation.

Results and Implication: Figure 13 illustrates the activity classification accuracies of all five radio signal transmission rates. The average classification accuracies across all five activities for transmission rates at 100Hz, 50Hz, 10Hz, 8Hz, and 5Hz are 86.2%, 85.7%, 80.6%, 67.4%, and 63.7% respectively. This result shows that as the transmission rate decreases, the classification performance drops. This indicates higher resolution provided by higher transmission rate improves classification performance. Moreover, there is a significant performance drop when the transmission rate decreases from 10Hz to 8Hz. This indicates that a minimum 10Hz of radio signal transmission rate is needed to provide enough resolution to capture the distinctive head and mouth movements across different activities. Finally, it is worthwhile to note that 10Hz only takes a tiny part of the overall wireless bandwidth, indicating high performance activity recognition can be achieved without sacrificing the regular wireless data transmission.

H. Impact of Radio Transmitter/Receiver On-body Locations

Objective: Under a wearable setting, the locations where the radio transmitter antenna (Tx) and receiver antenna (Rx) are placed on the human body have a significant impact on the radio propagation. The objective of this experiment is to examine the impact of the on-body locations of transmitter and receiver antennas on the activity recognition performance and then identify the best locations to wear each antennas.

Experimental Setup: We examine four different location combinations to wear the radio Tx and Rx antennas: (1) Rx on the shoulder, Tx on the collar (SC); (2) both Tx and Rx on the collar (CC); (3) both Tx and Rx on shoulder (SS); and (4) Rx on the collar, Tx on the shoulder (CS). We select these location combinations because compared to other candidate locations, they are less intrusive while still being close to the head and mouth. As such, the movements of the head and mouth are captured while the movements of other parts of the human body have limited influence on the radio signals. We collect approximately 35 minutes of data in four location combinations in total from one participant. At each location combination, the participant performs 10 trials for each activity. We examine the performance of each location combination using leave-one-trial-out validation.

Results and Implication: Figure 14 illustrates the activity classification accuracy of all four location combinations. As shown, the location combination SC performs the best among all four locations, achieving an average classification accuracy of 92% across all five activities. In contrast, the location combination CS achieves the lowest average classification accuracy (68%). This is because Rx is much more sensitive to movements that occur at a closer location to Rx. As such, by placing Rx on the shoulder which is farther from mouth and head compared to collar, Rx is less sensitive to the movements which act as noise that degrades the classification performance.

I. Impact of Interference Caused by Nearby People

Objective: In a wearable setting, there might be other people nearby moving around. The movements of nearby people may act as a source of interference to the radio signals. The objective of this experiment is to examine the robustness of HeadScan to the interference caused by nearby people.

Experimental Setup: We examine three scenarios where we have zero, one, and two people moving nearby respectively. During the experiment, people are instructed to stay within a one meter range of the participant wearing HeadScan, performing some common activities such as speaking, typing on the keyboard, and walking around. We collected approximate 30 minutes of data for the three scenarios from one participant. For each scenario, the participant performs 10 trials of each activity. We examine the performance of each scenario using leave-one-trial-out cross validation.

Results and Implication: Figure 15 illustrates the activity classification accuracies when zero, one, and two people are moving nearby. As shown, the classification performance, on average, does not decrease when there are one and two people moving nearby. This result indicates that our HeadScan wearable system is robust to the interference caused by nearby people. This is because in the wearable setting, although the movements of nearby people may cause interferences, those interferences are dwarfed by the activity performed by the user.

J. Usability Study

Objective: Finally, we conduct a user study on all seven participants about the usability of our radio-based wearable system HeadScan and the audio-based wearable system we replicated based on BodyScope. The objective of the user study is to understand whether HeadScan has the potential to be widely adopted by users in real world applications.

Experimental Setup: After completing the evaluation experiments, all the participants are asked to finish a survey which includes six questions related to comfort, privacy, safety and form-factor of the two wearable systems. We use a Likert Scale to code the answers of the participants. This scale is one of the most widely used approaches for scaling responses in survey research [26]. It adopts a five point scale ranging between: -2 (strongly disagree), -1 (disagree), 0 (neutral), 1 (agree), and 2 (strongly agree).

Results and Implication: Table V lists the six questions and the corresponding average point and standard deviation across all seven participants. As listed in the table, the results of this preliminary study show strong potential for our radio-based wearable sensing system in all dimensions. In terms of comfort, due to the tight skin contact of the stethoscope, most participants do not enjoy wearing the audio-based system. In contrast, most participants agree that they feel comfortable with antennas worn on their bodies because they require no tight skin contact. In terms of privacy, participants in general have privacy concerns about the audio-based system although their opinions vary significantly. In contrast, participants have a consensus that privacy is not a concern for the radio-based system. In terms of safety, participants also have a consensus that radio exposure from the HeadScan system is not really a concern. Finally, in terms of form-factor, participants in general feel the current prototype is rather bulky for a wearable. This result indicates our radio-based sensing approach exhibits significant advantages over the audio-based sensing approach by providing a truly non-intrusive and privacy-preserving solution. In the meantime, the size and dimensions of our current prototype needs to be improved to meet the requirements of real world usage.

Survey Questions	Mean	Standard Deviation
1. User feels comfortable with contact sensing	-0.86	1.07
2. User feels comfortable with HeadScan	0.88	0.35
3. Audio recording is not privacy-intrusive	-0.29	1.50
4. Radio recording is not privacy-intrusive	1.29	0.76
5. User does not worry about radio exposure	0.71	0.49
6. User feels HeadScan bulky	1.00	1.15
-2: Strongly Disagree, -1: Disagree, 0: Neutral, 1: Agree, 2: Strongly Agree		

TABLE V: List of survey questions and the corresponding Likert Scale average and standard deviation across all participants.

VI. Discussion

The evaluation results of our HeadScan system strongly indicate radio is a very promising sensing modality for detecting and recognizing head and mouth-related activities. We highlight the most prominent results from our experiments.

Competitive Recognition Performance: The activity classification accuracy of HeadScan exceeds its audio counterpart when classifying the targeted head and mouth-related activities (i.e., eating, drinking, coughing, and speaking) with an average accuracy of 86.3% compared to 81.7%.

Low Bandwidth Requirement: The classification performance of HeadScan remains above 80% when the radio signal transmission rate is lowered to 10Hz. Therefore, high performance activity recognition is achieved without sacrificing the regular wireless data transmission.

Robustness to Sensor Interference: HeadScan not only performs better in a clean environment, it also demonstrates stronger robustness to radio interference than audio-based sensing (e.g., RFI for HeadScan and background acoustic noise for audio-based sensing). Under reasonable level of interference, HeadScan classification results drops only 22% while audio-based sensing drops 55.5%.

Robustness to Nearby People: The classification accuracy of HeadScan remains relatively unaffected even when there are moving people in the testing environment. This suggests HeadScan can still function despite the potential non-user interferences caused by nearby people in the real world.

VII. Limitations

While our experiments have demonstrated a number of positive results, it is important to acknowledge the limitations of what we have findings, especially for drawing conclusions. We highlight key limitations here that we plan to address in future work.

Scripted Data Collection: Since this was the first time the targeted head and mouth-related activities had been attempted to be detected and recognized using radio-based sensing on wearables, we adopted a conservative scripted data collection protocol that enabled many factors to be controlled during later data analysis stages. As a result, participants followed our data collection protocol and did not deliberately confuse our system (e.g., excessively move or shake their head wildly) during experiments. In addition, we plan to conduct a prolonged real world field study to examine the performance of HeadScan.

Number of Participants: In this work, we recruited only seven participants to contribute to the dataset of HeadScan. Further verification of the results we have reported in this paper requires a larger and diverse set of participants.

Prototype Limitation: Although our wearable prototype is a proof of concept that is adequate for experiments, its dimensions, weight, and battery life can be further improved. We are currently investigating the use of the Intel Edison mini-computer that is built into a much smaller form factor (35.5 x 25.0 x 3.9 mm; 68 g). We estimate the new wearable prototype based on the Edison will be $10\times$ smaller and $2\times$ lighter. For battery life, our current wearable prototype can support approximately five hours of continuous CSI measurement sampling. We plan to are investigate techniques, such as duty cycling, to reduce the power consumption.

Antenna Locations: In our experiments, we have tested four different on-body antenna location combinations. There are numerous other location combinations where users can potentially wear the antenna pair. We plan to investigation other on-body location combinations once we have completed our new prototype design.

RFI Limitation: In our experiments, we have investigated the impact of RFI on the activity sensing and recognition performance when an interference source was placed one meter away from the participants. More experiments are needed to fully understand the impact of RFI on our wearable-based radio sensing approach when the interference source is placed at different distances from the participants.

VIII. Related Work

The design of HeadScan is closely related to two research areas: (1) wearable and mobile sensing systems and (2) radio-based activity recognition. In this section, we provide a brief review of the most relevant state-of-the-art research and compare them with HeadScan.

Wearable and Mobile Sensing Systems: With advances in MEMS technologies, the past decade has witnessed a surge of smartphones and wearable devices. These devices are equipped with miniature sensors that can be used to track a wide range of human activities. Among them, the accelerometer is one of the most widely used sensors. Accelerometer-based sensing systems have been developed for physical exercise tracking [8], fall detection [27] as well as activities of daily living (ADL) monitoring [28]. The Microphone is another extensively explored sensor because of the rich information contained in audio signals. In [29], Hao *et al.* developed iSleep that detects and recognizes sounds of body movement, coughing and snoring to infer sleep quality. In [30], Georgiev *et al.* developed DSP.Ear that recognized stress and emotion (amongst others) via the microphone using only the DSP of a smartphone. In [31], Nirjon *et al.* developed MusicalHeart that integrated a microphone into an earphone to extract heartbeat information from audio signals. Finally, in [9], Yatani *et al.* developed BodyScope that used a condenser microphone with a stethoscope head to capture a variety of sounds produced from activities that involve the mouth and throat such as eating, drinking, speaking, laughing and coughing. Our work is similar to BodyScope in the sense that we target a similar set of activities. However, since BodyScope uses the microphone as a sensing modality, it brings unwanted negative privacy side-effects. Moreover, BodyScope has to be worn on a user's neck in order to capture the subtle sounds generated by eating and drinking. The requirement of physical contact to a user's neck makes BodyScope intrusive to the user's daily life. In contrast, HeadScan resolves both privacy and skin contact issues in that it uses contactless radio signals to recognize those activities.

Radio-Based Activity Recognition: Recently, radio-based sensing systems have emerged and drawn considerable attention as they provide a non-intrusive solution to detect and recognize human activities. In [1], Wang *et al.* developed E-eyes that exploits CSI measurements extracted from radio signals to detect and recognize a variety of household activities such as cooking, washing dishes, and walking inside rooms. In [2], Wei *et al.* also leveraged the CSI measurements from radio signals and developed recognition algorithms that are robust to radio frequency interference (RFI) to recognize a set of location-oriented stationary and walking activities. In [32], Kaltiokallio *et al.* demonstrated the feasibility of using a single pair of radio transmitter and receiver to accurately measure an individual's respiration rate. The fundamental difference between these works and HeadScan is that they use radio transmitters and receivers deployed in the ambient environment to monitor stationary and moving activities that involve whole-body movements; in comparison, our work uses a radio transmitter and receiver worn on the human body to capture activities that involve mouth and head movements. Finally, in [33], Li *et al.* developed Tongue-n-Cheek, a radio-based wearable system for tongue gesture recognition. Our work is similar to Tongue-n-Cheek in the sense that we both explore the potential of radio as a new sensing modality for wearable devices. However, Tongue-n-Cheek uses 24GHz radio signals generated from an array of dual-channel radar sensors and leverages the principle of Doppler effect to detect and recognize tongue movements. In contrast, HeadScan uses fine-grained CSI measurements extracted from 5GHz radio signals from a single pair of radio transmitters and receivers and adopts a novel signal processing pipeline based on a noise-robust sparse representation framework to detect and recognize activities involving head and mouth movements.

IX. Conclusion

In this paper, we presented the design, implementation and evaluation of *HeadScan*, a wearable system that uses radio as a sensing modality to capture and recognize head and mouth-related activities including eating, drinking, coughing and speaking. Through extensive experiments, we have demonstrated that our radio-based sensing approach achieves competitive performance compared to a state-of-the-art audio-based sensing approach. We have also empirically identified the best on-body locations to wear HeadScan and demonstrated the robustness of the system to non-user interferences. HeadScan represents our first exploration of radio-based human activity recognition in the context of a wearable device. We envision radio-based sensing has the potential to fundamentally transform wearables as we currently know them, allowing them to have a much broader impact on our lives.

Immediate future work include plans to develop the next generation of our wearable prototype, with a much smaller form factor and reduced power consumption. This will enable large-scale and prolonged real world experiments. We also plan to broadly investigate other potential use cases of radio-based sensing on wearables including sensing human activities that involve body and limb movements, as well as physiological signals.

References

[1] Y. Wang, J. Liu, Y. Chen, M. Gruteser, J. Yang, and H. Liu, "E-eyes: device-free location-oriented activity identification using fine-grained wifi signatures," in *Proceedings of the 20th annual international conference on Mobile computing and networking*. ACM, 2014, pp. 617–628.

[2] B. Wei, W. Hu, M. Yang, and C. T. Chou, "Radio-based device-free activity recognition with radio frequency interference," in *Proceedings of the 14th International Conference on Information Processing in Sensor Networks*, ser. IPSN '15. New York, NY, USA: ACM, 2015, pp. 154–165. [Online]. Available: http://doi.acm.org/10.1145/2737095.2737117

[3] C. Han, K. Wu, Y. Wang, and L. M. Ni, "Wifall: Device-free fall detection by wireless networks," in *INFOCOM, 2014 Proceedings IEEE*. IEEE, 2014, pp. 271–279.

[4] G. Wang, Y. Zou, Z. Zhou, K. Wu, and L. M. Ni, "We can hear you with wi-fi!" in *Proceedings of the 20th annual international conference on Mobile computing and networking*. ACM, 2014, pp. 593–604.

[5] K. Ali, A. X. Liu, W. Wang, and M. Shahzad, "Keystroke recognition using wifi signals," in *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*. ACM, 2015, pp. 90–102.

[6] J. Liu, Y. Wang, Y. Chen, J. Yang, X. Chen, and J. Cheng, "Tracking vital signs during sleep leveraging off-the-shelf wifi," in *Proceedings of the 16th ACM International Symposium on Mobile Ad Hoc Networking and Computing*. ACM, 2015, pp. 267–276.

[7] J. Heiskala and J. Terry, *OFDM wireless LANs: A theoretical and practical guide*. Sams, 2001.

[8] N. D. Lane, M. Lin, M. Mohammod, X. Yang, H. Lu, G. Cardone, S. Ali, A. Doryab, E. Berke, A. T. Campbell, and T. Choudhury, "Bewell: Sensing sleep, physical activities and social interactions to promote wellbeing," *Mobile Networks and Applications*, vol. 19, no. 3, pp. 345–359, 2014. [Online]. Available: http://dx.doi.org/10.1007/s11036-013-0484-5

[9] K. Yatani and K. N. Truong, "Bodyscope: A wearable acoustic sensor for activity recognition," in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, ser. UbiComp '12. New York, NY, USA: ACM, 2012, pp. 341–350. [Online]. Available: http://doi.acm.org/10.1145/2370216.2370269

[10] T. Rahman, A. T. Adams, M. Zhang, E. Cherry, B. Zhou, H. Peng, and T. Choudhury, "Bodybeat: A mobile system for sensing non-speech body sounds," in *Proceedings of the 12th Annual International Conference on Mobile Systems, Applications, and Services*, ser. MobiSys '14. New York, NY, USA: ACM, 2014, pp. 2–13. [Online]. Available: http://doi.acm.org/10.1145/2594368.2594386

[11] S. Reddy, A. Parker, J. Hyman, J. Burke, D. Estrin, and M. Hansen, "Image browsing, processing, and clustering for participatory sensing: lessons from a dietsense prototype," in *Proceedings of the 4th workshop on Embedded networked sensors*. ACM, 2007, pp. 13–17.

[12] E. Thomaz, A. Parnami, J. Bidwell, I. Essa, and G. D. Abowd, "Technological approaches for addressing privacy concerns when recognizing eating behaviors with wearable cameras," in *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*. ACM, 2013, pp. 739–748.

[13] E. Thomaz, A. Parnami, I. Essa, and G. D. Abowd, "Feasibility of identifying eating moments from first-person images leveraging human computation," in *Proceedings of the 4th International SenseCam & Pervasive Imaging Conference*. ACM, 2013, pp. 26–33.

[14] J. Cheng, O. Amft, G. Bahle, and P. Lukowicz, "Designing sensitive wearable capacitive sensors for activity recognition," *Sensors Journal, IEEE*, vol. 13, no. 10, pp. 3935–3947, 2013.

[15] S. Sen, V. Subbaraju, A. Misra, R. K. Balan, and Y. Lee, "The case for smartwatch-based diet monitoring," in *Pervasive Computing and Communication Workshops (PerCom Workshops), 2015 IEEE International Conference on*. IEEE, 2015, pp. 585–590.

[16] G. Wang, Y. Zou, Z. Zhou, K. Wu, and L. M. Ni, "We can hear you with wi-fi!" in *Proceedings of the 20th Annual International Conference on Mobile Computing and Networking*, ser. MobiCom '14. New York, NY, USA: ACM, 2014, pp. 593–604. [Online]. Available: http://doi.acm.org/10.1145/2639108.2639112

[17] "HummingBoard Pro," http://wiki.solid-run.com/doku.php?id=products:imx6:hummingboard.

[18] "Intel 5300," http://www.intel.com/content/www/us/en/wireless-products/ultimate-n-wifi-link-5300-brief.html.

[19] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, "Tool release: Gathering 802.11n traces with channel state information," *ACM SIGCOMM CCR*, vol. 41, no. 1, p. 53, Jan. 2011.

[20] Y. Li, L. J. Cimini Jr, and N. R. Sollenberger, "Robust channel estimation for ofdm systems with rapid dispersive fading channels," *Communications, IEEE Transactions on*, vol. 46, no. 7, pp. 902–915, 1998.

[21] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, "Tool release: Gathering 802.11 n traces with channel state information," *ACM SIGCOMM Computer Communication Review*, vol. 41, no. 1, pp. 53–53, 2011.

[22] R. Nandakumar, B. Kellogg, and S. Gollakota, "Wi-fi gesture recognition on existing devices," *arXiv preprint arXiv:1411.5394*, 2014.

[23] D. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, April 2006.

[24] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, February 2006.

[25] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Review*, vol. 43, pp. 129–159, January 2001. [Online]. Available: http://dl.acm.org/citation.cfm?id=588736.588850

[26] T. J. Maurer and H. R. Pierce, "A comparison of likert scale and traditional measures of self-efficacy." *Journal of applied psychology*, vol. 83, no. 2, p. 324, 1998.

[27] M. Zhang and A. A. Sawchuk, "Context-aware fall detection using a bayesian network," in *Proceedings of the 5th ACM International Workshop on Context-Awareness for Self-Managing Systems*. ACM, 2011, pp. 10–16.

[28] C. Spenkelink, M. M. Hutten, H. Hermens, and B. O. Greitemann, "Assessment of activities of daily living with an ambulatory monitoring system: a comparative study in patients with chronic low back pain and nonsymptomatic controls," *Clinical Rehabilitation*, vol. 16, no. 1, pp. 16–26, 2002.

[29] T. Hao, G. Xing, and G. Zhou, "isleep: unobtrusive sleep quality monitoring using smartphones," in *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems*. ACM, 2013, p. 4.

[30] P. Georgiev, N. D. Lane, K. K. Rachuri, and C. Mascolo, "Dsp.ear: Leveraging co-processor support for continuous audio sensing on smartphones," in *Proceedings of the 12th ACM Conference on Embedded Network Sensor Systems*, ser. SenSys '14. New York, NY, USA: ACM, 2014, pp. 295–309. [Online]. Available: http://doi.acm.org/10.1145/2668332.2668349

[31] S. Nirjon, R. F. Dickerson, Q. Li, P. Asare, J. A. Stankovic, D. Hong, B. Zhang, X. Jiang, G. Shen, and F. Zhao, "Musicalheart: A hearty way of listening to music," in *Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems*. ACM, 2012, pp. 43–56.

[32] O. Kaltiokallio, H. Yigitler, R. Jantti, and N. Patwari, "Non-invasive respiration rate monitoring using a single cots tx-rx pair," in *Information Processing in Sensor Networks, IPSN-14 Proceedings of the 13th International Symposium on*. IEEE, 2014, pp. 59–69.

[33] Z. Li, R. Robucci, N. Banerjee, and C. Patel, "Tongue-n-cheek: non-contact tongue gesture recognition," in *Proceedings of the 14th International Conference on Information Processing in Sensor Networks*. ACM, 2015, pp. 95–105.