Did you Remember to Brush? : A Noninvasive Wearable Approach to Recognizing Brushing Teeth for Elderly Care

Josh Cherian, Vijay Rajanna, Daniel Goldberg, Tracy Hammond
Sketch Recognition Lab. Dept. of Computer Science & Engineering
Texas A&M University, College Station, Texas
{jcherian14, vijayrajanna, daniel.goldberg, hammond}@tamu.edu

ABSTRACT
Failing to brush one's teeth regularly can have surprisingly serious health consequences, from periodontal disease to coronary heart disease to pancreatic cancer. This problem is especially worrying when caring for the elderly and/or individuals with dementia, as they often forget or are unable to perform standard health activities such as brushing their teeth, washing their hands, and taking medication. To ensure that such individuals are correctly looked after they are placed under the supervision of caretakers or family members, simultaneously limiting their independence and placing an immense burden on their family members and caretakers. To address this problem we developed a non-invasive wearable system based on a wrist-mounted accelerometer to accurately identify when a person brushed their teeth. We tested the efficacy of our system with a month-long in-the-wild study and achieved an accuracy of 94% and an F-measure of 0.82.

ACM Classification Keywords
H.5.2 Information Interfaces and Presentation: User Interfaces; J.3 Computer Applications: Life and Medical Sciences

Author Keywords
Activity recognition; brushing teeth; pervasive health; wearable solution; elderly care; dementia; machine learning; intervention

INTRODUCTION
From a young age most people are taught to brush their teeth on a regular basis. Even as adults, it is an activity that can seem like an inconvenience; however, failing to do so regularly can have surprisingly devastating health consequences. Failing to brush one's teeth regularly and maintaining poor oral hygiene in general can lead to a variety of systemic diseases including atherosclerosis, COPD, diabetes, and bacteremia as well as conditions such as brain abscess [20, 22, 23]. Periodontal disease, one of the most common results of poor oral hygiene is estimated to affect 10-15% of the world's population [18]. Individuals with this disease are 3.2 times more likely to die from Type 2 diabetes [28], 4.3 times more likely to have a cerebral ischemic stroke [10], have been shown to have a 25% increased risk of heart disease [6], and are at increased risk of developing Alzheimer's and pancreatic cancer [19, 29]. Furthermore, periodontal disease and poor oral hygiene in general has been shown to have a stronger association with total mortality than with coronary heart disease [6].

For the elderly and/or those experiencing the onset of dementia, this issue is compounded by the fact that they can fail to perform standard health activities such as brushing their teeth as often as they should. In the early stages of dementia, patients may need to be reminded that they need to brush their teeth and/or be supervised while doing so. In later stages, patients may be unable to brush their teeth, stop understanding that they need to clean their teeth, or lose interest in doing so [11]. Current methods of tracking whether or not a person performs standard cleanliness activities have serious limitations. One study found that of the nearly 70% of U.S. adults who track at least one health metric, half of them track this metric "in their heads" while one third keep track of this metric using some type of notebook or journal [9]. Depending on a person to recall if they brushed their teeth relies on a person's memory, while having them keep track of this in a notebook or journal relies on the user's diligent and accurate recording and recollection of where those records are kept. By leveraging the assortment of sensors within wearable devices today, it is entirely possible to not only accurately and securely track these metrics but also to develop more sophisticated and personalized applications and interventions based on these metrics [32, 33].

In this work, we describe a wearable system that can effectively identify when a person has brushed their teeth solely from a wrist-mounted accelerometer. By doing so, we aim to present the basis for a paradigm where physicians and/or family members can be alerted if activities are performed incorrectly or not at all. Our eventual goal is to develop a system to recognize this activity among elderly adults and those experiencing the onset of dementia. However, there is often significant diversity in the capabilities of individuals with motor and/or cognitive impairments due to the varying severity of their impairments. As a result, in this study, we sought to determine if it was possible to recognize this activity among individuals without impaired motor or cognitive functions as they form a homogeneous

PervasiveHealth '17, May 23–26, 2017, Barcelona, Spain

ACM ISBN 978-1-4503-6363-1/17/05…$15.00
https://doi.org/10.1145/3154862.3154866

group, which is the most appropriate target group to evaluate the performance of our initial system. Demonstrating that our initial system can recognize brushing teeth among individuals without an impairment allows us to further enhance the system to recognize brushing teeth among users with different kinds of impairments through iterative design and co-creation.

PRIOR WORK

Numerous studies have looked attempted to recognize the action of brushing one's teeth; however, it has commonly been investigated simultaneously with other ambulation and posture activities [15, 26]. Bao & Intille [4] placed bi-axial accelerometers on five different locations on the body to recognize 20 activities, finding that accelerometers placed on the wrist and arm were better for recognizing activities such as brushing teeth that required predominantly upper body movement. Lester et al. [17] developed a multimodal sensor board, which was used to discriminate between activities such as walking, walking up and down stairs, and brushing one's teeth with an accuracy of about 90%. Hong et al. [14] used a combination of accelerometers and radio-frequency identification (RFID) sensors to differentiate between many different ambulation and posture activities as well as activities such as brushing hair, reading, and brushing teeth. While these studies work well in laboratory settings, in real-life scenarios such models would be unable to distinguish between activities with repetitive wrist movements similar to brushing teeth. Furthermore the use of multiple sensors located on different locations on the body makes such systems unwieldy in real-life scenarios.

In recent years, activity recognition has solved the latter problem by taking advantage of the popularity of smartphones and the array of sensors that they contain. Kwapisz et al. [16] had users keep a smartphone in their front pant pocket and were able to discern between walking, jogging, ascending stairs, descending stairs, sitting, and standing with an overall accuracy of 91.7%. Anguita et al. [1] were also able to identify walking, standing, ascending stairs, descending stairs, sitting, and laying with a smartphone attached to a user's waist with an overall accuracy of around 89%. World of Workout, a mobile role-playing game (RPG), allowed the user to improve their character through a variety of exercises such as jumping jacks, crunches, and cycling, which were automatically recognized using the phone's accelerometer and GPS data [5]. In more recent years, several studies have utilized wrist-mounted devices [24, 25]. Thomaz et al. recognized eating movements [31] with an overall F-score of 76% using the 3-axis accelerometer in a Pebble smartwatch for both controlled and in-the-wild studies. Weiss et al. [32] recognized both hand-oriented (such as brushing teeth, eating different foods, folding clothes) and not hand-oriented activities (such as walking, jogging, and climbing stairs) using accelerometer and gyroscope data from both a LG G Watch smartwatch and a Samsung Galaxy S4 smartphone with an average accuracy of 91.9% when using the watch's accelerometer.

To recognize these activities, existing systems typically extract a subset of features–mainly from accelerometer data–designed for general purpose use and not for specific activities. Common features used in previous studies include average, standard deviation, correlation, energy and entropy. Prior to the work presented in this study we sought to test the efficacy of these common features in determining when someone was brushing their teeth, extracting 30 features traditionally used in activity recognition from our data to do so. However, with these features we were only able to achieve an accuracy of 87.8% and an F-measure of 0.68. To improve these results, we created and extracted 21 novel features.

In this work we present a noninvasive system designed to specifically determine when a user was brushing their teeth. It is worth noting that we chose this approach over developing an object-based system (e.g. placing sensors on the toothbrush, toothpaste, faucet, etc.) as an object-based system requires significantly more infrastructure, especially as more and more activities are sought to be recognized [30]. Through the evaluation of our system, we show that we can distinguish between activities that, in several cases, have only slight differences in orientation and vigor. Furthermore, we show that such recognition is not limited to a controlled setting but can realistically be done in a real-world setting.

METHODOLOGY

System Implementation

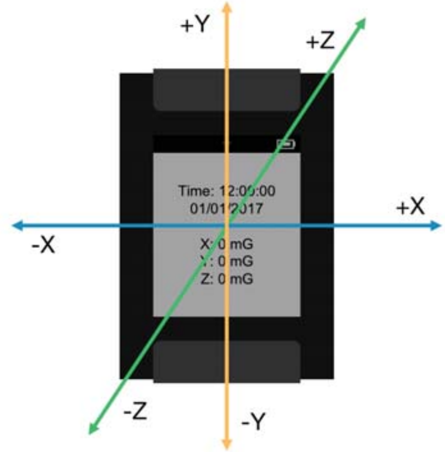

Figure 1. Pebble Accelerometer Axes

To collect data and to build a model for activity recognition, a system consisting of a Pebble smartwatch and an Android application was developed. While the Pebble currently does contain and allow access to both a 4G 3-axis accelerometer (shown in Figure 1) and a magnetometer, only the accelerometer data and corresponding timestamps were recorded and subsequently used for recognition. Data were recorded at a sampling rate of 25 Hz. Data for each activity were transmitted via Bluetooth using the Pebble SDK to the Android application. All collected data that did not fall within 90-110% of the sampling rate (22.5 Hz - 27.5 Hz) were not used in data analysis. In addition to facilitating the storage of this data, the Android application allowed users to start and stop data collection and select which activity was about to be performed. The latter feature was included to label the data to facilitate offline data analysis. Any data that were recorded while none of the six specified activities were being performed were labeled as “Inactive”.

Activities

We collected accelerometer data for six activities: brushing one’s teeth, combing one’s hair, scratching one’s chin, washing one’s hands, taking medication, and drinking. These activities were selected for at least one of the two following reasons:

1. The activity plays a role in the average person’s everyday health, and
2. The activity requires physical movements similar to those required by another activity being studied.

Activities such as brushing one’s teeth, washing one’s hands, and for some people taking medication are activities that are examples of the first reason as should be performed regularly and have an impact on our day-to-day health. When collecting data, taking medication was separated into two distinct activities, taking a pill (simulated with an M&M and a push and turn pill bottle) and drinking a glass of water.

Brushing one’s teeth, combing one’s hair, scratching one’s chin, and washing one’s hands are examples of the second reason, as they consist of back and forth movements of the wrist. In this study we sought to identify these movements, particularly those associated with brushing teeth. To that end, we limited our definition of each activity to the most literal definition of the activity. With brushing teeth this meant that we made no effort to recognize the other actions associated with the colloquial definition of brushing teeth such as putting toothpaste on the toothbrush, rinsing out one’s mouth, and cleaning one’s toothbrush. We did this because in general the physical movements associated with the most literal definition provide the clearest indication that the activity was actually performed.

ACTIVITY RECOGNITION FRAMEWORK

To build a system that could recognize brushing teeth in real time, we iteratively developed the system over three phases of user studies, where the duration of data collection increased from Phases 1 to 3. Each of these phases and the findings from each phases are described in the following sections. A summary of these phases are shown in Table 1.

Phase 1: Controlled Study

The first phase was a proof of concept study, i.e., we sought to determine if it was possible to identify brushing teeth when it was compared to other activities that looked quite similar to it. For this phase, 20 participants were recruited and were asked to complete each of the six activities consecutively. This was done in a controlled environment where the experiment facilitator labeled the data while the person performed the activity. Approximately 79 minutes of data were collected. Samples of this data can be seen in Figure 2.

Recognition

Data collected from these 20 participants was segmented into four-second sliding windows with a one-second overlap. From these windows we extracted a total of 51 features, which

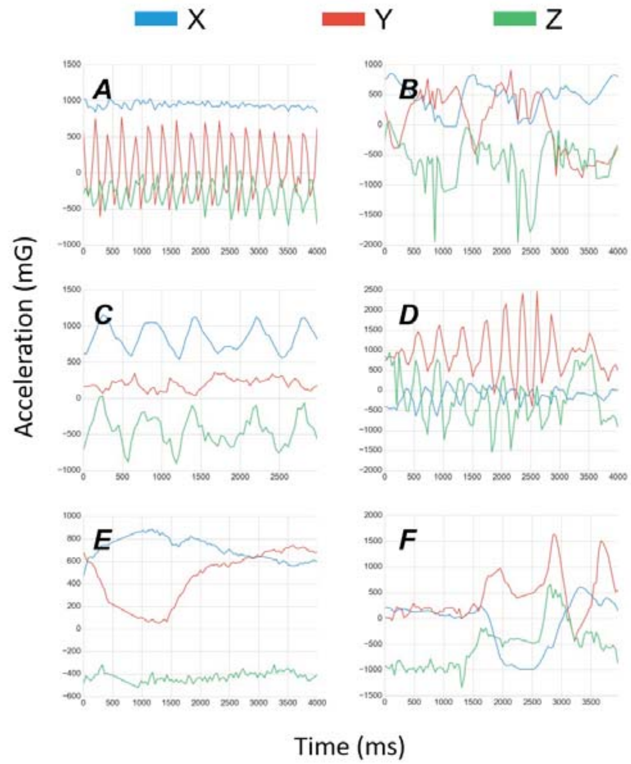

Figure 2. Sample graph of each activity where A = Brushing Teeth, B = Combing Hair, C = Scratching Chin, D = Washing Hands, E = Drinking, and F =Taking Medication

were run through several classification algorithms as shown in Table 3.

Features

Once the collected data were segmented into windows, a total of 51 features were extracted. We selected a combination of features that were discerned from visualizing and comparing the data for each activity and features that have been shown to be effective in activity recognition previously. Most features were calculated for each axis or from a combination of axes.

Of the features that have been previously shown to be effective, we chose a combination of features in the time domain and in the frequency domain. Time domain features are advantageous due to their lack of computational complexity, and as a result can provide a basic yet practical quantitative sense of the data. For our purposes we calculated the average (X, Y, Z), standard deviation (X, Y, Z), average jerk (X, Y, Z), average distance between axes (XY, XZ, YZ), correlation (XY, XZ, YZ), number of peaks (X, Y, Z), number of valleys (X, Y, Z), and root mean square (RMS) (X, Y, Z). Frequency domain features are advantageous for their ability to highlight the repetitiveness of signals. These features included energy (X, Y, Z) and entropy (X, Y, Z). Energy has been shown to be effective in differentiating between activities by intensity, while entropy has been shown to be capable of further distinguishing between activities that have similar energy levels. The formulas for energy and entropy can be seen in Equations 1 and 2, where n is the size of the window and a and b are the real and imaginary

Table 1. Summary of methodology

	Phase 1	Phase 2	Phase 3
Description	Controlled Study[a]	Hour-long Study	Naturalistic Study
Number of Participants	20	10[b]	12[c]
Features	51	51	51
Feature Subset	24 (Table 2)	13 (Table 5)	13 (Table 5)
Activities	6	2[d]	2[d]
Algorithm	Tier I Recognition	Tier I Recognition	Tiers I and II Recognition
Results	Table 3 & 4	Tables 6	Table 7

[a]Proof of concept study
[b]Six unique users
[c]None of the users participated in Phase 1 or 2
[d]The five activities that were not brushing teeth were re-labeled as inactive

components respectively of each data point in the window after they have been converted to the frequency domain using a discrete Fast Fourier Transform (FFT).

$$energy = \sum_{j=1}^{n} \frac{a_j^2 + b_j^2}{n} \quad (1)$$

$$entropy = \sum_{j=1}^{n} c_j * log(c_j), c_j = \frac{\sqrt{a_j^2 + b_j^2}}{\sum_{k=1}^{n} \sqrt{a_k^2 + b_k^2}} \quad (2)$$

These two features have commonly been used together in activity recognition [7, 13].

While these features had been shown to effectively distinguish between a variety of activities, none of them are specific to the problem of differentiating exclusively between activities consisting primarily of back and forth wrist and forearm movements. Therefore to recognize brushing teeth, we sought to account for the unique motion and characteristics of this activity by extracting several additional features by observing and quantifying the differences between the data for each activity.

As can be seen in Figure 2, while many of the activities featured the see-saw pattern in the accelerometer data that expected from a back and forth physical gesture, the frequency and amplitude of this pattern varried between activities. Specifically for brushing teeth we noticed that the lengths of the segments connecting the peaks and valleys present in the data were fairly consistent. As a result, we calculated the average and standard deviation of the side heights for each axis. Furthermore, the peaks and valleys were fairly consistent as well so we calculated the average and standard deviation of both. Additionally we noticed that the x-axis rarely crossed the other two axes; to capture this quality we calculated the number of times the axes crossed each other. All together this produced an additional 21 features. To determine the optimal subset of the 51 calculated features, we used the CfsSubset Evaluator with the BestFirst search method found in the Weka Data Mining Toolkit [12]. The 24 selected features are shown in Table 2.

Table 2. Feature subset for discerning between the six activities

Features
Average Jerk: Z
Average Height: X, Z
Standard Deviation Height: X
Energy: Z
Entropy: Z
Average: X, Y, Z
Average Distance Between Axes: XZ, YZ
Standard Deviation: Y, Z
Axis Order: YZ
Number of Peaks: X, Z
Average Peaks: X, Y, Z
Standard Deviation of Peaks: Y
Average Valleys: X, Z
Standard Deviation Valleys: Y, Z

Results

The extracted features were run through several different classifiers with 10-fold cross-validation; the results can be seen in Table 3. Table 4 shows the confusion matrix of the best classifier, Random Forest, that recognizes each activity with an accuracy of 83.4% and a F-measure of 0.83.

Table 3. Performance of classifiers for distinguishing between the six activities

Classifier	Overall Accuracy (%)	Overall F-measure
C4.5	71.0	0.71
KNN (K=5)	75.0	0.75
Multilayer Perceptron	66.5	0.66
Random Tree	70.5	0.704
Random Forest	83.4	0.83

Phase 2: Hour-long Study

While Phase 1 demonstrated the system's ability to recognize activities that are similar, it did not take into account what the user would be doing in between these activities, which is vital to being able to recognize brushing teeth over the course of a day. Thus, the second phase of data collection sought to introduce unrelated or "inactive" data into the model. This phase of the study was conducted over the course of an hour, during which participants went about their normal everyday routine, completing each of the activities at least

Table 4. Confusion matrix for discerning between the six activities using Random Forest

Activity	Classified As					
	Brush Teeth	Comb Hair	Drinking	Scratch Chin	Take Meds	Wash Hands
Brush Teeth	93.5	0.36	0.71	0.27	2.14	3.03
Comb Hair	22.9	64.0	3.56	2.77	3.95	2.77
Drinking	6.14	1.28	88.5	1.54	1.28	1.28
Scratch Chin	13.9	3.01	2.41	73.5	4.22	3.01
Take Meds	19.6	0.42	5.42	1.67	70.0	2.92
Wash Hands	19.51	0.45	1.79	0.22	2.92	75.1

once at random times during the hour. This was done over two days, with data being collected from six participants on the first day, and data being collected from four of these six participants again on the second day.

Since this dataset was particularly imbalanced, we combined the data from this phase with the data from the previous phase so as to increase the amount of data representative of the six activities. Since the focus of this work is on identifying when someone is brushing their teeth, data that were not labeled as brushing teeth were relabeled as inactive. In other words we sought to classify data as either brushing teeth or inactive. From this data we extracted the same set of 51 features. As the labels had changed we again used the CfsSubset Evaluator with the BestFirst search method to determine the optimal subset of features. The 13 selected features are shown in Table 5.

Table 5. Feature subset for identifying brushing teeth

Features
Entropy: Z
Average: X, Z
Axis Order: YZ
Number of Peaks: Y
Standard Deviation Peaks: X, Y, Z
Number of Valleys: X, Y
Standard Deviation Valleys: X, Y, Z

Results

The extracted features were run through the same series of algorithms that were used in Phase 1; the results can be seen in Table 6. Notably, although the accuracies are quite high, the F-measure values are quite low. This discrepancy is due to two specific issues: the definition of brushing teeth, and the nature of F-measure. As mentioned earlier, we sought to recognize solely the physical back and forth movements of brushing one's teeth. As participants in Phase 2 were not instructed to specifically indicate when they were literally brushing their teeth much of the data labeled as brushing teeth included the physical movements associated with the other actions involved in the general definition, reducing the accuracy of the classifiers.

Furthermore since our dataset was imbalanced in favor of Inactive data, the majority of the data within the dataset resulted in true negatives, skewing the accuracy to be quite high. However, since the F-measure does not take true negatives into account, it was more indicative of the classifier's ability to identify windows of brushing teeth.

Table 6. Performance of classifiers for identifying four-second windows of brushing teeth

Classifier	Overall Accuracy (%)	F-measure
C4.5	96.1	0.62
KNN (K=5)	96.1	0.636
Multilayer Perceptron	96.0	0.594
Random Tree	94.4	0.560
Random Forest	96.5	0.657

Phase 3: Naturalistic Study

The goal of Phase 3 was two-fold: to mitigate the issue of ambiguously labeled data and to determine if we could identify when an individual brushed their teeth over the course of a day. To that end, we developed a two-tier recognition system.

Two-Tier Recognition

In order to identify instances of brushing teeth we developed a two-tier recognition system. In Tier I we followed the approach described in the previous two phases, where data was segmented into four-second sliding windows with a one-second overlap and features extracted from these windows were inputted into several classifiers. In Tier II we constructed a dynamic window comprised of the individual windows classified as brushing teeth or inactive in Tier I. To inform the inclusion of windows into this dynamic window we established three thresholds: a *minimum activity duration*, a *maximum inactivity duration*, and an *activity percentage*.

Given that the recommended duration for brushing teeth is 2-3 minutes, we initially considered setting the *minimum activity duration* duration to 60 seconds, or in other words considered users to have brushed their teeth when they were predicted to have been brushing their teeth for at least 60 consecutive windows [2]. However, since studies have shown that people on average spend less than this recommended time brushing their teeth, we lowered this threshold to 45 seconds [3].

The *maximum inactivity duration* threshold was established to account for small groups of windows classified as inactive within larger groups of windows classified as brushing teeth. If the number of sequential windows classified as inactive did not exceed this threshold before a window was classified as brushing teeth, these windows classified as inactive were included in the total count of windows that were part of the instance of brushing teeth. If the number of sequential windows classified as inactive did exceed this threshold, the count of the number of windows towards an instance of brushing teeth was reset. In this study this threshold was set to 15 seconds as setting this metric to be much less than this led to single instances of brushing teeth being divided into separate instances.

The *activity percentage* threshold was included to ensure that most of the windows in the series were classified as brushing teeth, thereby ensuring that situations where sequences of windows classified as inactive interspersed by short sequences of windows classified as brushing teeth were not classified as instances of brushing teeth. We found that setting this metric around 75% worked well; any value significantly lower than this introduced a number of false positives. A simplified example of this system can be seen in Figure 3.

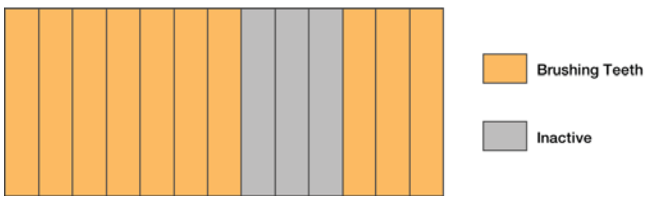

Figure 3. A series of windows where all but three windows are classified as brushing teeth in the middle of the image. If the threshold for inactive windows is greater than three, these windows will be treated as misclassified brushing teeth and will be included in the count towards the number of windows towards this instance of brushing teeth. If the duration of this series is greater than the minimum duration threshold this section will be considered an instance of brushing teeth, as more than 75% of the windows were classified as brushing teeth.

Data Collection

To evaluate the efficacy of this recognition system, we decided to evaluate the models built on the data from Phases 1 & 2 (shown in Table 6) on a more representative data set. For this study, 12 users who had not participated in either Phase 1 or 2 were recruited. Data was collected over the course of a month, with data being collected on average for 6.25 days per user over the 31 day period. For this phase, data collection was run for an average of 4.9 hours per day per user for a total of 367.5 hours of data. Users were asked to perform each of the six activities at least once for each day of the study.

Table 7. Performance of Tier II recognition for identifying four-second windows of brushing teeth

Classifier	Overall Accuracy (%)	F-measure
C4.5	93.6	0.824
KNN (K=5)	88.4	0.645
Multilayer Perceptron	90.5	0.710
Random Tree	77.6	0.333
Random Forest	91.7	0.777

Results

In evaluating our model we defined instances of brushing teeth, as found by our two-tier recognition system as true positives, while we defined the time in between these instances as true negatives. The results of using our two-tier recognition system are shown in Table 7. C4.5 achieved the best results with an accuracy of 94% and an F-measure of 0.82. As mentioned earlier, using the 30 features commonly used in activity recognition led to an accuracy of 87.8% and an F-measure of 0.68, showing that the inclusion of the 21 novel features had a clear positive impact on the system's ability to recognize instances of brushing teeth. This system did produce five false negatives and one false positive. Each of these cases occurred for slightly different reasons, and as such are described and addressed individually below.

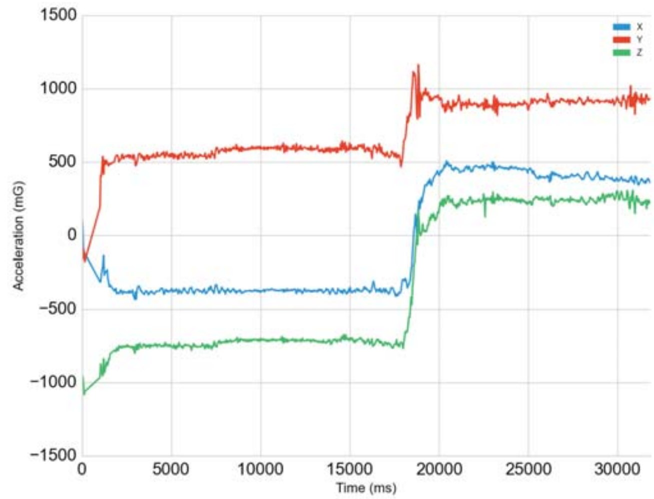

Figure 4. The image above shows a 28-second period consisting of one false negative. Although the user indicated that they were brushing their teeth during this time, the data seems to indicate that they were not in fact doing so and that they may have mistakenly labeled it as if they were.

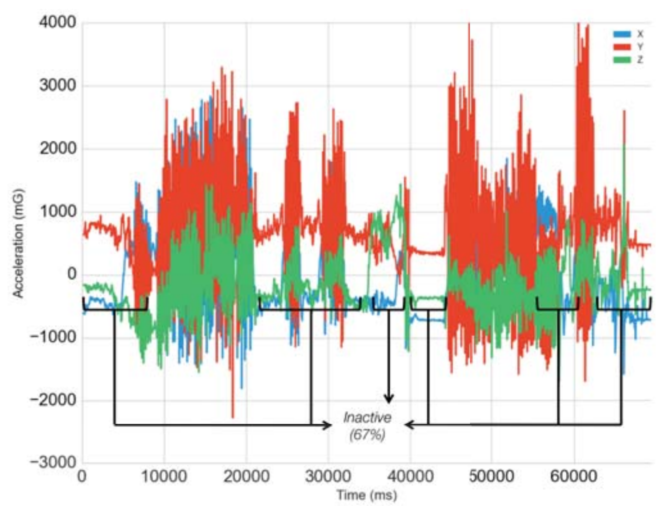

Figure 5. The image shows an approximately 60-second period during which only 67% of the windows within the period were classified as brushing teeth producing a false negative. The areas that were misclassified are indicated by the bracketed regions in the figure.

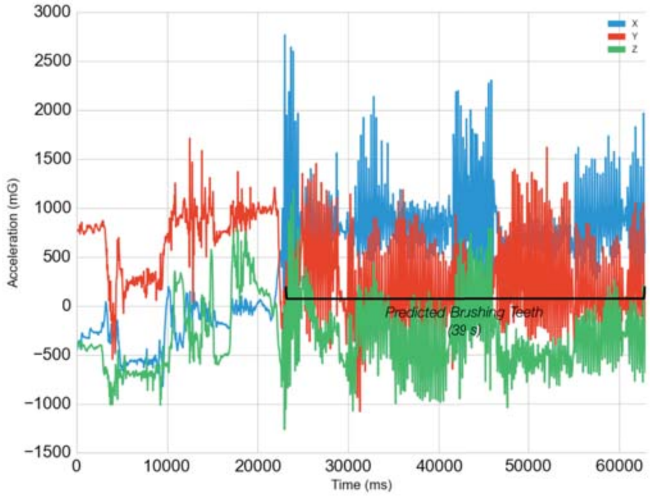

Figure 6. The image above shows an approximately 60-second period of which only 39 seconds were recognized as brushing teeth.

False Negative 1

This false negative occurred because the user indicated that they only brushed their teeth for 28 seconds. Within this period our system did not classify any of the windows within that period as brushing teeth. Looking at the data from this period (shown in Figure 4) it does not seem like the user brushed their teeth at all during this period, and might have mistakenly labeled the data as if they had.

False Negative 2

In the case of one false negative, the system only achieved an activity percentage of 67%. Figure 5 shows this instance of brushing teeth with the sections that our model classified as inactive specifically annotated. The data seems to indicate that the user paused during these times, which reduced the percentage of time the model classified as brushing teeth. This issue could be averted by reducing the minimum activity percentage; however, this introduces many false positives. If the number of false positives could be reduced, reliable recognition of smaller sections of brushing teeth could be recognized, put together if occurring within a short period of time, and/or even serve as the basis for motivational interventions encouraging participants to brush their teeth for longer periods of time.

False Negative 3

In the case of one false negative, the user indicated they brushed their teeth for 60 seconds, however, our system only identified the user as having brushed their teeth for 39 seconds. Interestingly, our system misclassified the first 21 seconds of the period, and looking at the data shown in Figure 6, it seems that the user spent the first 21 seconds of the labeled period of time doing something other than literally brushing their teeth, e.g., perhaps putting toothpaste on their toothbrush and/or running the toothbrush under the faucet.

False Negative 4

As with the previous false negative, in this instance the user brushed their teeth for at least 45 seconds, but the algorithm recognized a duration shorter than this threshold, as can be seen in Figure 7. In this case, the user officially brushed their teeth for 49 seconds while the algorithm predicted they only brushed their teeth for only 34 seconds. Looking at the figure, there does seem to be a five second gap from about 40 to 45 seconds into the gesture where the user does not seem to be brushing their teeth. Given that analysis was done using four-second windows with a one-second overlap it is possible that this five-second period altered the feature values for the windows around it enough for the model to not recognize those surrounding windows as brushing teeth.

False Negative 5 & False Positive 1

Our algorithm also produced one false positive, which occurred right after a false negative as can be seen in Figure 8. Although the user only annotated the first 43 seconds as brushing their teeth (producing the false negative), in all likelihood the user actually brushed their teeth for 75 seconds, took a one minute break, and then continued brushing their teeth for another 67 seconds. Clearly this approximately three minute period can be treated as one instance of the user brushing their teeth over the course of a day.

FUTURE WORK

Having proven that brushing one's teeth can be identified over the course of a day through the offline processing techniques outlined above, one of the next steps is to determine whether this model can be used to identify brushing teeth in real time.

Additionally, we will seek to develop similar systems for recognizing activities other than brushing teeth. Two activities that are particularly attractive candidates for recognition are washing hands and taking medication, as they are important indicators of cleanliness and medication nonadherence, respectively. To this end, a question that remains to be answered is how well our model will generalize to these other domains. It does seem possible to establish thresholds specific to other activities, e.g., the CDC recommends that people wash their hands for at least 20 seconds; however, there is a good chance that the issue of people doing the activity in much less than the recommended time will arise [8]. Furthermore, recognition of other activities could require modified algorithms; recognizing taking medication will most likely require looking for a specific series of events rather than looking for repetitive patterns.

The eventual goal of this work is to develop intelligent personalized interventions based on this recognition. We envision a few potential use cases. One use case is developing a system for people with dementia, as patients in both the early and late stages of dementia can struggle with performing such basic activities. [11]. Another use case is medication noadherance, a widespread and expensive problem worldwide [21, 27]. By recognizing these activities we can provide users with interventions such as reminders, encouragement, and suggestions, or even alert the user's physician and/or family members if something might be wrong.

DISCUSSION

Through a three phase iterative development, we have created a noninvasive wearable solution that recognizes when someone

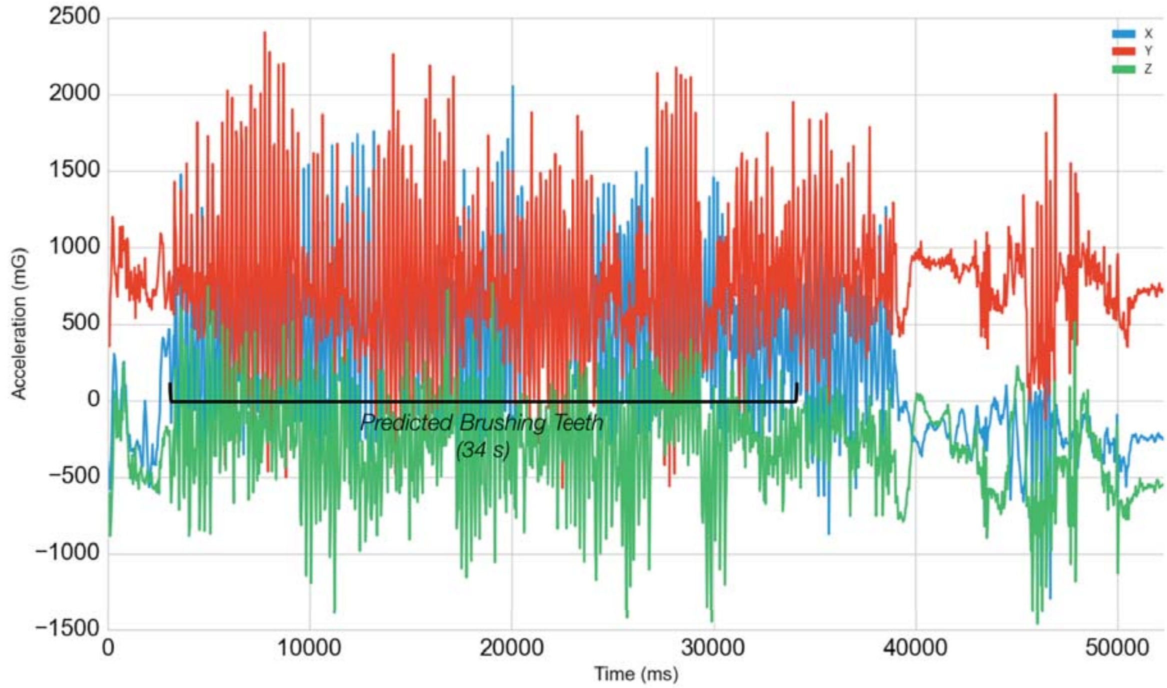

Figure 7. The image above shows a three minute period consisting of one false negative. The section labeled "Predicted Brushing Teeth" contained continuous instances of four-second windows classified as brushing teeth, but because the entire length was less than one minute, the entire time was not ultimately classified as an instance of brushing teeth during the second pass. This type of error represents one of the five false negatives. One solution to this type of error would be to reduce the minimum duration required for recognition, but that increases the number of false positives.

brushed their teeth. There are three primary contributions of this work. Firstly, while existing systems that recognize brushing teeth have been shown to work in laboratory studies, they primarily distinguish this activity from dissimilar activities and remain unproven in real-world scenarios. Our solution identifies brushing teeth not only from amongst other gesturally similar activities, but also from full days of normal activity. Secondly, to improve the efficacy of our system, we added new features to the existing body of knowledge in activity recognition. Existing features have been shown to effectively aid in discriminating between brushing teeth and predominantly ambulatory and posture activities, but have not been shown to effectively inform accurate distinction between activities requiring similar movements. To remedy this, we developed 21 novel features by observing unique patterns in and characteristics of the data representative of brushing one's teeth. Thirdly, we presented a two-tier recognition system for recognizing when someone brushed their teeth in the wild. In Tier I, four-second windows of data are classified as brushing teeth or not brushing teeth, while in Tier II these windows are combined to find instances of brushing teeth.

Overall through this work, we showed that by extracting unique features that are specific to an activity and processing the extracted features through intelligent algorithms that incorporate human knowledge about the activity, the activity of interest can be accurately recognized. Our results demonstrate that even activities that have subtle differences in the orientation and vigor can be accurately recognized using just the accelerometer data from a wrist-worn device. Most importantly, we achieve this through a noninvasive wearable solution that is lifestyle compatible, as smartphones and smartwatches are becoming increasingly affordable and ubiquitous. While it is entirely possible that similar results could be achieved by augmenting the objects commonly used for performing the activity with sensors, such a system would require much more infrastructure, especially as the number of activities needing to be recognized increases. We believe, further research in recognizing activities with subtle differences should not only focus on extracting the unique features, but also focus on developing multi-tier recognition systems that are built on an understanding of the various sub-activities that constitute an activity.

CONCLUSION

In order to expand the utility of current wearable devices it is vital to broaden their functionality. In this paper we present a system that can differentiate between six activities using accelerometer data from a Pebble smartwatch worn on the wrist using a combination of features derived from literature and from observing unique patterns in the data. The six chosen activities all represented activities chosen for either their gestural similarity to other chosen activities and/or their role in an average person's everyday health. On top of this, we presented a two-tier recognition system for identifying when a user is brushing their teeth. Through this work we

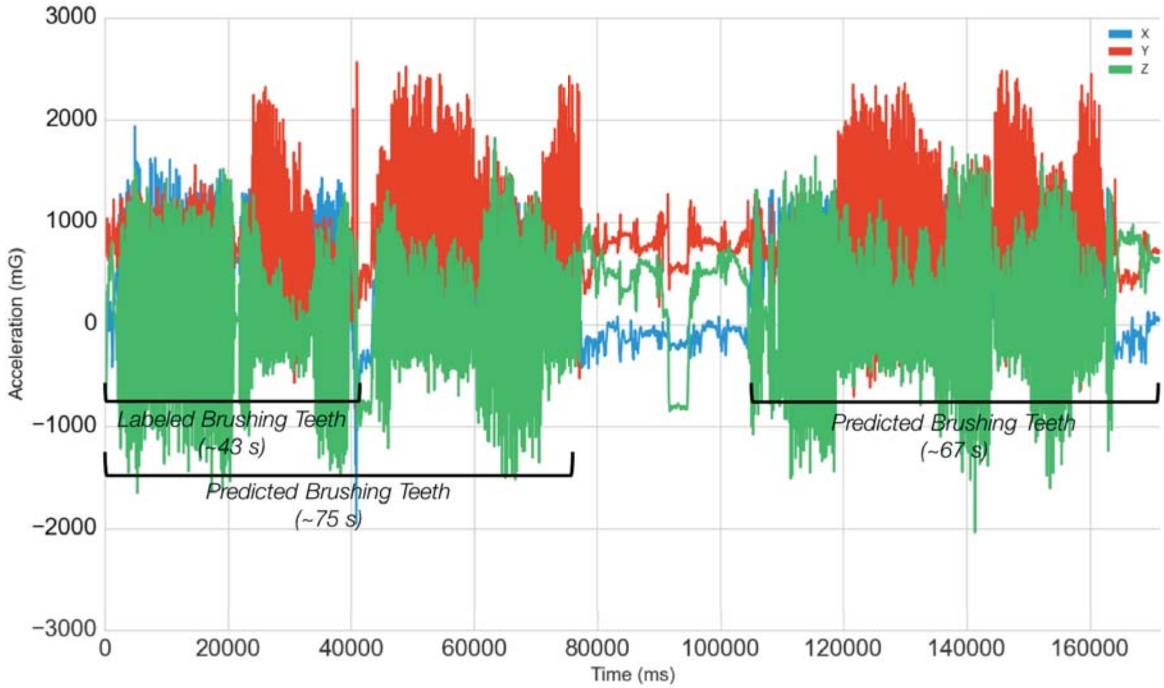

Figure 8. The image above shows a three minute period consisting of both a true positive and a false positive.As can be seen there are two continuous periods of time that were predicted to be instances of brushing teeth. However, during this three-minute period, the user only indicated they brushed their teeth for the first 43 seconds. Following this 43-second period there is a small gap during which it is likely the user indicated via the Android app that they were no longer brushing their teeth. Looking at the data we believe there is a high likelihood that this user continued to brush their teeth after indicating they had stopped.

aim to facilitate the development of personalized intelligent interventions that can improve the everyday health of a wider population.

ACKNOWLEDGMENTS

The authors thank the members of the Sketch Recognition Lab and TEES AggiE_Challenge as well as Aaron Harmon for their perspectives on this work. This work is funded in part by TEES AggieChallenge and Future Position X. Furthermore the authors gratefully acknowledge the grant from NSF (#SMA-1560106)

REFERENCES

1. Davide Anguita, Alessandro Ghio, Luca Oneto, Xavier Parra, and Jorge L Reyes-Ortiz. 2012. Human activity recognition on smartphones using a multiclass hardware-friendly support vector machine. In *Ambient assisted living and home care*. Springer, 216–223.
2. American Dental Association. 2016. Brushing Your Teeth. *http://www.mouthhealthy.org/en/az-topics/b/brushing-your-teeth* (2016).
3. T Attin and E Hornecker. 2005. Tooth brushing and oral health: how frequently and when should tooth brushing be performed? *Oral health & preventive dentistry* 3, 3 (2005).
4. Ling Bao and Stephen S Intille. 2004. Activity recognition from user-annotated acceleration data. In *Pervasive computing*. Springer, 1–17.
5. Joey Bartley, Jonathon Forsyth, Prachi Pendse, Da Xin, Garrett Brown, Paul Hagseth, Ashish Agrawal, Daniel W Goldberg, and Tracy Hammond. 2013. World of workout: a contextual mobile RPG to encourage long term fitness. In *Proceedings of the Second ACM SIGSPATIAL International Workshop on the Use of GIS in Public Health*. ACM, 60–67.
6. Frank DeStefano, Robert F Anda, Henry S Kahn, David F Williamson, and Carl M Russell. 1993. Dental disease and risk of coronary heart disease and mortality. *Bmj* 306, 6879 (1993), 688–691.
7. Davide Figo, Pedro C Diniz, Diogo R Ferreira, and João MP Cardoso. 2010. Preprocessing techniques for context recognition from accelerometer data. *Personal and Ubiquitous Computing* 14, 7 (2010), 645–662.
8. Centers for Disease Control and Prevention. 2015. Handwashing: Clean Hands Save Lives, When & How to Wash Your Hands. *http://www.cdc.gov/handwashing/when-how-handwashing.html* (2015).

9. Susannah Fox and Maeve Duggan. 2013. *Tracking for health*. Pew Research Center's Internet & American Life Project.

10. Armin J Grau, Heiko Becher, Christoph M Ziegler, Christoph Lichy, Florian Buggle, Claudia Kaiser, Rainer Lutz, Stefan Bültmann, Michael Preusch, and Christof E Dörfer. 2004. Periodontal disease as a risk factor for ischemic stroke. *Stroke* 35, 2 (2004), 496–501.

11. Janet Griffiths and Vicki Jones. 2015. *Tracking for health*. Alzheimer's Society.

12. Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H Witten. 2009. The WEKA data mining software: an update. *ACM SIGKDD explorations newsletter* 11, 1 (2009), 10–18.

13. Joyce Ho. 2005. *Interruptions: using activity transitions to trigger proactive messages*. Ph.D. Dissertation. Massachusetts Institute of Technology.

14. Yu-Jin Hong, Ig-Jae Kim, Sang Chul Ahn, and Hyoung-Gon Kim. 2010. Mobile health monitoring system based on activity recognition using accelerometer. *Simulation Modelling Practice and Theory* 18, 4 (2010), 446–455.

15. Tzu-Ping Kao, Che-Wei Lin, and Jeen-Shing Wang. 2009. Development of a portable activity detector for daily activity recognition. In *Industrial Electronics, 2009. ISIE 2009. IEEE International Symposium on*. IEEE, 115–120.

16. Jennifer R Kwapisz, Gary M Weiss, and Samuel A Moore. 2011. Activity recognition using cell phone accelerometers. *ACM SigKDD Explorations Newsletter* 12, 2 (2011), 74–82.

17. Jonathan Lester, Tanzeem Choudhury, and Gaetano Borriello. 2006. A practical approach to recognizing physical activities. In *Pervasive Computing*. Springer, 1–16.

18. Mara S Meyer, Kaumudi Joshipura, Edward Giovannucci, and Dominique S Michaud. 2008. A review of the relationship between tooth loss, periodontal disease, and cancer. *Cancer Causes & Control* 19, 9 (2008), 895–907.

19. Dominique S Michaud, Kaumudi Joshipura, Edward Giovannucci, and Charles S Fuchs. 2007. A prospective study of periodontal disease and pancreatic cancer in US male health professionals. *Journal of the National Cancer Institute* 99, 2 (2007), 171–175.

20. Cindy L Munro and Mary Jo Grap. 2004. Oral health and care in the intensive care unit: state of the science. *American Journal of critical care* 13, 1 (2004), 25–34.

21. Lars Osterberg and Terrence Blaschke. 2005. Adherence to medication. *New England Journal of Medicine* 353, 5 (2005), 487–497.

22. Poul Erik Petersen and Tatsuo Yamamoto. 2005. Improving the oral health of older people: the approach of the WHO Global Oral Health Programme. *Community dentistry and oral epidemiology* 33, 2 (2005), 81–92.

23. Janette C Rahamat-Langendoen, Marit GA van Vonderen, Lex J Engström, Willem L Manson, Arie Jan Van Winkelhoff, and Esther ANM Mooi-Kokenberg. 2011. Brain abscess associated with Aggregatibacter actinomycetemcomitans: case report and review of literature. *Journal of clinical periodontology* 38, 8 (2011), 702–706.

24. Vijay Rajanna, Folami Alamudun, Daniel Goldberg, and Tracy Hammond. 2015. Let Me Relax: Toward Automated Sedentary State Recognition and Ubiquitous Mental Wellness Solutions. In *Proceedings of the 5th EAI International Conference on Wireless Mobile Communication and Healthcare (MOBIHEALTH'15)*. ICST, Brussels, Belgium, 28–33. DOI: http://dx.doi.org/10.4108/eai.14-10-2015.2261900

25. Vijay Rajanna, Raniero Lara-Garduno, Dev Jyoti Behera, Karthic Madanagopal, Daniel Goldberg, and Tracy Hammond. 2014. Step Up Life: A Context Aware Health Assistant. In *Proceedings of the Third ACM SIGSPATIAL International Workshop on the Use of GIS in Public Health (HealthGIS '14)*. ACM, New York, NY, USA, 21–30. DOI:http://dx.doi.org/10.1145/2676629.2676636

26. Nishkam Ravi, Nikhil Dandekar, Preetham Mysore, and Michael L Littman. 2005. Activity recognition from accelerometer data. In *AAAI*, Vol. 5. 1541–1546.

27. Eduardo Sabaté. 2003. *Adherence to long-term therapies: evidence for action*. World Health Organization.

28. Aramesh Saremi, Robert G Nelson, Marshall Tulloch-Reid, Robert L Hanson, Maurice L Sievers, George W Taylor, Marc Shlossman, Peter H Bennett, Robert Genco, and William C Knowler. 2005. Periodontal disease and mortality in type 2 diabetes. *Diabetes Care* 28, 1 (2005), 27–32.

29. Pamela Sparks Stein, Michelle J Steffen, Charles Smith, Gregory Jicha, Jeffrey L Ebersole, Erin Abner, and Dolph Dawson. 2012. Serum antibodies to periodontal pathogens are a risk factor for AlzheimerâĂŹs disease. *Alzheimer's & Dementia* 8, 3 (2012), 196–203.

30. Emmanuel Munguia Tapia, Stephen S Intille, and Kent Larson. 2004. *Activity recognition in the home using simple and ubiquitous sensors*. Springer.

31. 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.

32. Gary M Weiss, Jessica L Timko, Catherine M Gallagher, Kenichi Yoneda, and Andrew J Schreiber. 2016. Smartwatch-based activity recognition: A machine learning approach. In *2016 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI)*. IEEE, 426–429.

33. Natalie Wilde, Katrin Hänsel, Hamed Haddadi, and Akram Alomainy. 2015. Wearable Computing for Health and Fitness: Exploring the Relationship between Data and Human Behaviour. *arXiv preprint arXiv:1509.05238* (2015).