Food Watch: Detecting and Characterizing Eating Episodes through Feeding Gestures

Shibo Zhang
Preventive Medicine Dept.
Northwestern, Chicago
shibo.zhang@northwestern.edu

Rawan Alharbi
Preventive Medicine Dept.
Northwestern, Chicago
rawan.alharbi@northwestern.edu

William Stogin
Preventive Medicine Dept.
Northwestern, Chicago
wstogin@northwestern.edu

Mohamad Pourhomayun
Computer Science Dept.
CSULA, Los Angeles
mpourho@calstatela.edu

Bonnie Spring
Preventive Medicine Dept.
Northwestern, Chicago
bspring@northwestern.edu

Nabil Alshurafa
Preventive Medicine Dept.
Northwestern, Chicago
nabil@northwestern.edu

ABSTRACT

Obesity, caused primarily by overeating, is a preventable chronic disease yielding staggering healthcare costs. To detect overeating passively, a machine learning framework was designed to detect and accurately count the number of feeding gestures during an eating episode to characterize each eating episode with a feeding gesture count. With the ubiquitous nature of wrist-worn sensors, existing literature has focused on detecting eating-related gestures and eating episodes that are at least f ve minutes long. In this paper, our objective is to show the potential of commercial smartwatches to be used in detection of eating episodes with short durations confounded by other activities of daily living in order to truly capture all eating episodes in the field The effect of time-series segmentation and sensing configuration on the accuracy of detecting and characterizing feeding gestures is then analyzed. Finally, the effects of personalized and generalized machine learning models in predicting feeding gestures are compared. Results demonstrate the large within-subject variability of eating, where a generalized user-independent model yields a 75.7% average F-measure, whereas a personalized user-dependent model yields a 85.7% average F-measure. This shows the effects of personalized clustering on feeding gesture count, resulting in a root mean square error of 8.4.

CCS Concepts

•Computing methodologies → Machine learning; •Applied computing → Consumer health;

Keywords

Wrist-worn sensors, wearables, hand-to-mouth gestures, overeating, inertial sensors

1. INTRODUCTION

Eating is essential to human life, not made or distributed for profit or commercial advantage and thatbut overeating relative to need is not. Unfortunately, once bad eating habits are formed, they become challenging to overcome. People overeat for many reasons, such as loss of control [14], impulsivity due to cues [16], or heightened emotional state as a result of stress [15] or negative affect [8], or even positive affect [6]. Being able to passively detect overeating in real time will enable researchers to understand the antecedents and causes of overeating.

Prior to detecting overeating, researchers have been focused on detecting eating compared to other confounding activities of daily living. However, many of the developed eating detection systems are challenging for participants to adhere to throughout the day, requiring multiple on-body sensors [4] or specialized neck-worn [1], ear-based [3, 5], or chest-based sensors [23]. Moreover, the eff cacy of these systems to detect eating in a variety of real-life settings is yet to be validated.

Prior research on eating using commercial wearable products has focused primarily on detecting the act of eating [24], but fall short of characterizing eating episodes to enable behavioral scientists to properly understand the causal determinants of problematic eating behaviors. Due to the ubiquitous nature and acceptability of wrist-worn sensors [12], they provide a viable solution to seamlessly and objectively detecting eating episodes. Gathering this data will empower nutrition researchers to develop improved systems of dietary recall, and enable behavioral scientists to test timely interventions to successfully understand and prevent problematic eating.

Prior literature has begun examining eating episodes in the wild through feeding gestures [24] and counting of bites [9]. However, much of the eating episodes involve participants sitting and continuously eating, uninterrupted by the realities of daily life. While these systems prove to work well in lab settings, they are unable to consistently and accurately detect short eating episodes, let alone characterize them with accurate counts of feeding gestures. In this paper, these efforts are further expanded upon by analyzing more realistic eating episodes, where participants make phone calls, walk around to answer the door, type on their laptops, and converse with people in a room.

Objectively detecting physical activity through passive sensing of wearable sensors has been researched extensively [17]. It is currently feasible, within reasonable error, and using passive sensing from wearable sensors, to detect bouts and minutes of sedentary, moderate, and vigorous physical activity, and ultimately classify a subset of activities of daily living. The problem of passively sensing and characterizing eating episodes, however, is yet to manifest in behavioral interventions. This paper attempts to address important practical challenges in deploying such a system.

2. RELATED WORKS

The widespread availability of embedded wearable accelerome-

BODYNETS 2016, December 15-16, Turin, Italy

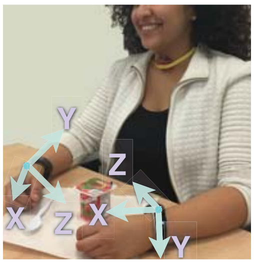

Figure 1: A participant wearing two Microsoft Band 2's (one on each wrist) prepared to begin eating.

ters and gyroscopes has enabled a new area of research to detect eating passively through on-body inertial sensors. The focus of this research effort is on detecting and characterizing eating through feeding gestures.

The problem of identifying hand-to-mouth gestures has also been studied to detect smoking activities in order to predict smoking relapse [21, 22]. PuffMarker successfully detects the timing of a relapse using multiple sensors, where detecting hand-to-mouth gestures and hand orientation (using roll and pitch angles) are part of the smoking lapse system [21]. RisQ also leverages multiple inertial measurement units (IMUs) placed on a person's body together with 3D animation to detect hand-to-mouth gestures, and while the focus of these systems is primarily smoking, their systems have shown preliminary success in detecting feeding gestures [19].

Dong et al. shows correlation between bites and caloric intake and measure intake via automated tracking of wrist motion [10, 11]. Some of their limitations include requiring the user to turn the device on and off, and they focus primarily on detecting the start and end of an eating episode throughout the day, as opposed to characterizing a given eating episode.

Thomaz et al. present a framework for detecting eating episodes using wrist-worn accelerometer data [24] and density-based spatial clustering of applications with noise (DBSCAN) to identify eating episodes throughout the day. Assumed in the study under primary consideration in this paper are that the beginning and end of an eating episode are defined interspersed with activities of daily living, and this study takes the next step of accurately characterizing the eating episode through feeding gestures. The loop is thus closed on the practical aspects of deploying an eating gesture system, providing insight into the impact of slow, fine-graine and fast coarse-grained segmentation, and report on using both personalized and generalized classificatio models and their impact on counting the number of feeding gestures.

The following are the contributions of this effort: 1) Showing the correlation and association of feeding gestures and eating duration with caloric intake; 2) Presenting accuracy of counting feeding gestures; 3) The impact of the non-dominant hand, sensor, and feature combinations on detecting hand-to-mouth gestures; 4) The benefit of a personalized compared to a generalized machine learning model using clustering techniques on detecting feeding gestures; and 5) Characterizing eating episodes, interspersed with confounding activities, by accurately detecting and counting feeding gestures. In the following sections our system and results are described in more detail.

3. DATA COLLECTION

3.1 Eating Study

Participants in an in-lab experiment were requested to wear one

		Breakfast	
Activity	*Object*	*Task Method*	*Calories*
Eating	Fruit Cup	Fork	60
Non-Eating	Glasses	Wear & adjust	-
Eating	Pancake	Fork & Knife	100
Non-Eating	Conversation	Talk naturally	-
Drinking	Water	Hands	0
Non-Eating	Phone	Call	-
Eating	Yogurt	Spoon	140
Drinking	Tea	Hands	0

Table 1: Activities and food consumed during breakfast.

		Lunch	
Activity	*Object*	*Task Method*	*Calories*
Eating	Sandwich	Hands	340
Non-Eating	Keyboard	Type	-
Eating	Soup	Spoon	230
Non-Eating	Walk	Walk naturally	-
Drinking	Sparkling Water	Hands	0
Eating	Chips	Hands	150

Table 2: Activities and food consumed during lunch.

wrist-worn sensor on each hand (Microsoft Band 2s), while following an eating and activity protocol. Fifteen participants were recruited, 7 male and 8 female. The mean age of the participants was 31.5 years (ranging from 21 - 63 years, std=12.6). Two participants were left-handed. The participants were scheduled to come into the lab on two separate occasions, once for breakfast and another for lunch. Tables 1 and 2 provide descriptions of the tasks performed during each meal.

Participants were asked to perform activities that involved eating using different methods, both using utensils (fork and knife, fork alone, and spoon alone) and eating with their hands. Intermittently they were asked to perform other activities such as: have a conversation, make a phone call, put on and off a pair of sun glasses, walk around the room, and type on the keyboard. The participants were requested to focus on eating, unless they felt discomfort, and in the event that a participant was full, they were given the opportunity to take a break and continue after the break. If they could not eat more food, the remaining food was weighed to estimate the amount of food consumed during the meal.

The entire session was recorded using a camera to annotate the activities and serve as ground truth for our models. Different colored stickers were placed on the chin, hands, and throat of the participants to aid in annotating the video. Figure 1 shows a participant wearing the sensors.

3.2 Video Labeling

A Logitech C615 HD webcam camera was placed on the side of the non-dominant hand in order to see the plate, both hands, and the participants face. We used Chronoviz [13] to label data. Although in this paper we focus on detecting primarily feeding gestures, the variety of our labels includes eating, drinking, non-eating, food-to-mouth (F-M), bite, back-to-rest (B-R), chewing duration, and swallows. Figure 2 illustrates the system used to annotate the recorded data with time-synchronized ground truth labels.

3.3 Devices and System

An Android application was designed to act as an information gateway to the wrist bands and our backend database. The data is stored in an sql-lite relational database, and can both be transmitted to a back-end or accessed directly from the smartphone through local storage. The application allows researchers to specify several parameters, including participant ID, study ID, device name, and the option to turn on a combination of sensors (accelerometer, gy-

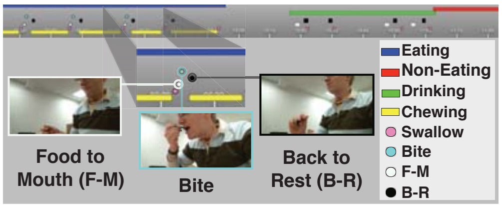

Figure 2: Labeling data using Chronoviz to enable the detection of feeding gestures.

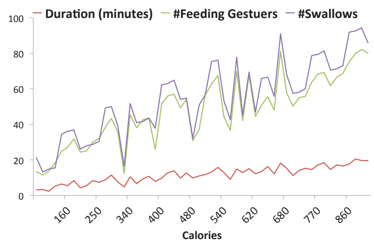

Figure 3: Relationship between Calories and feeding gestures, swallows, and eating duration for 54 foods consumed.

roscope, heart-rate, skin temperature, etc.) and their corresponding sample rates. The local database makes it possible to readily upload the data to a back end where the data is analyzed for in-the-wild studies.

One essential challenge was streaming data disconnecting from the Microsoft Band 2. As a result, a heart beat was added to the Android application that checks intended sensor connectivity every one minute, and reestablishes connectivity if necessary.

3.4 Features and Calories

While passive sensing cannot currently capture caloric overeating, determining how and when people eat still conveys information about increased caloric consumption. Despite the fact that a single bite of food can vary drastically in kilocalorie content, it is consistently true that the more someone puts into their mouth, the more kilocalories they are consuming. Figure 3 shows the average of hand-to-mouth gestures, swallows, and eating duration based on an average of all 15 participants, showing 54 combinations of foods (from Table 1 and 2). While the trend of the number of feeding gestures in a meal combination does dip around 340 kilocalories (that specifi meal included the sandwich which did not require as many feeding gestures as other meals), the general trend is the number of feeding gestures linearly increasing with kilocalories. The Pearson product-momentis used to examine the association between kilocalories and each of number of feeding gestures, number of swallows, and eating duration (in minutes).

A significan association between Calories and each of feeding gestures ($r = .86, P < .0001$), swallows ($r = .87, P < .0001$), and eating duration ($r = .91, P < .0001$) was found.

4. MODEL DEVELOPMENT

In order to build a model that will be used to detect an intended outcome, data quality and reliability must be assessed and improved. Figure 4 presents this study's approach to developing a model that detects and characterizes moments of eating through feeding gestures, including algorithms that perform data preprocessing, segmentation, and classification in order to build personalized and generalized models to count feeding gestures.

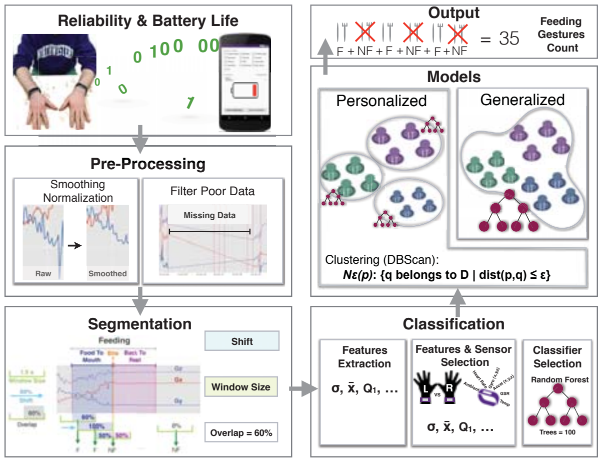

Figure 4: Framework for building a model that detects feeding gestures.

4.1 Data Preprocessing

Prior to processing the data, we analyze its reliability to ensure correct sample rate and negligible missing data. Tri-axial accelerometer and gyroscope data is collected at a sampling rate of 31Hz (or 31 samples per second, on average), a setting offered by the Microsoft Band 2.

The firs step is to remove poor data from the training models. For one participant about 23 seconds of data was lost and excluded from our model training. The second step of data preprocessing is to ensure that the inertial signals' intended measurements were captured, primarily by smoothing to reduce noise and normalization. The premise of smoothing data is that one is measuring a variable that is both slowly varying and corrupted by random noise. Consequently replacing each data point with an average of surrounding points reduces the level of noise, while hopefully not biasing the values. To smooth the gyroscope and accelerometer data a rolling mean (*pandas.rolling_mean*) with a window size of 100 points (approximately 3 seconds, which we empirically set) is applied. Then the data is normalized to unit norm which can improve the accuracy of results when quantifying the similarity of signals across samples.

4.2 Segmentation

A feeding gesture is define through two subfeeding moments: food-to-mouth, and back-to-rest. Figure 5 provides an illustration of each part of a gyroscope signal during a feeding gesture. Prior to processing values for three variables are set: 1) *window_size*; 2) overlap threshold (*overlap*); and 3) sliding window shift (*shift*). When processing the data a fi ed time subdivision of the data is continuously analyzed in order to decide if it is a feeding gesture or not. *Window_size* is the fi ed time subdivision of the signal (1.5 seconds in Figure 5), and the *shift* represents how fast the window is being slid; a 50% shift is shown in Figure 5, where the greater the percentage the faster the window is slid and the less overlap there is between consecutive data samples. The *overlap* is the threshold of overlap between the window and the true feeding gesture label that determines whether the segment is labeled as a feeding gesture or not.

When processing the data, however, as shown in Figure 5 a window is identifie as a feeding gesture (as a 'F' in Figure 5) only if the window overlaps with the food-to-mouth segment or the back-to-rest segment more than the overlap threshold. For example, in Figure 5, since the overlap threshold is set to 60%, segment 3 will be

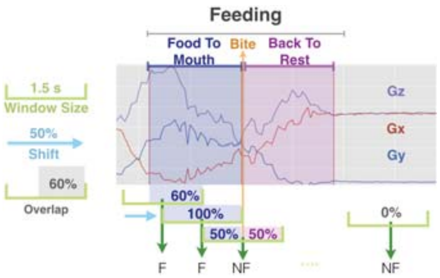

Figure 5: Segmentation process of Gyro data with window size, overlap and shift.

characterized as a non-feeding gesture. As a result, it will identify the bite section as a non-feeding gesture. This is important because some people spend a longer time biting, which might impede the detection of feeding gestures. Dynamic Time Warping (DTW) may be able to account for this problem [18] in detecting similarity between signals that vary in time.

4.3 Feature Extraction

Following data preprocessing of the raw signal, it is important to determine what features to collect on the raw signal that will be predictive of one's outcome. Due to the high variability across signals that represent the same activity, and to ensure that the system is capable of running in real-time, 11 statistical features are extracted on fi ed time subdivisions of the data that are known to be useful in detecting activity [2] and eating [24], including: mean, median, max, min, standard deviation, kurtosis, interquartile range, quartile 1, quartile 3, skewness, and root mean square (RMS). Running each statistical feature on each axis of the inertial sensors generates 132 features, creating samples with 132 dimensions in the $\mathbb{R}^{132}$ feature space.

4.4 Fine-tuning Signal Processing Parameters

A subsequent step to preprocessing is fine-tunin the signal processing parameters. It is important to test whether a slow-moving fine-graine (small window size, small shift, and high overlap) or fast-moving coarse (large window size, large shift, and low overlap) segmentation of the data provides improvements in the analysis of feeding gestures. While several prior efforts in detecting feeding gestures overlook this important step, this can drastically impact the results of classificatio and can provide insight into the detection of feeding gestures. We tested window sizes ranging from 1 to 2.5 seconds, sliding window shifts from 30% to 70%, and signal overlap thresholds from 50% to 80%.

In order to test the performance of each signal processing parameter, data was used from the 13 right-handed participants and each participant's data was balanced into equal feeding and non-feeding gestures (total of 520 samples). Random Forest (with number of trees n=100) was used to build predictive models from the training set prior to testing [7].

Each parameter setting was tested using 10-fold Cross Validation (CV) and Leave One Subject Out Cross Validation (LOSOCV). To ensure stability 10-fold CV was run ten times using a different random seed and average results.

4.5 Selecting the Best Feature Subset

After identifying the optimal signal processing parameters, the dimensionality of the problem space is further reduced in order to identify the primary predictive features. To ensure there is no bias, feature selection is performed on a different dataset than the one

	Precision	Recall	F-measure
Avg	0.859	0.858	0.857
Std	0.041	0.042	0.042

Table 3: Results of the 67/33 train/test split on data of breakfast and lunch combined.

used to train the predictive models in order to prevent over-fittin of the training data, and then select the top six features. A filte -based chi-squared statistical method is used for feature selection to assign a score to each feature with respect to the outcome.

4.6 Selecting the Optimal Classifier

Once the optimal feature subset and signal processing algorithms are identified a varied set of classificatio algorithms is tested in order to ensure selection of the optimal classifie . The tested algorithms include Logi stic Regressi on (LogisticRe g), AdaBoostClassifie (AdaBoost), C4.5 Decision Trees (DecisionTree), Gaussian Naive Bayes (GaussianNB), Linear Support Vector Classifie (LinearSVC), and Random Forest (RF) with n=100 trees. Both LOSOCV and 10-fold CV (averaged 10 times) were tested. The variability across the different runs was also calculated to ensure that not only was the classifie with the highest F-measure (harmonic mean of precision and recall) selected, but the one with the lowest variability (i.e. highest consistency) across the runs. This is essential to show the consistency of the classifier when tested on different subjects or a different test set.

4.7 Personalization

It was desirable to test the effect of personalized models on the accuracy of detecting feeding gestures, compared to generalized models. As a result the breakfast and lunch were combined and a 67/33 split was performed between training and testing data. The results of a purely personalized model are tested on each participant and F-measure, precision, and recall are reported.

4.8 Clustering to Detect Feeding Gestures

When detecting feeding gestures, it is important to perform post-processing of the data, in order to cluster feeding gestures. Since there is overlap between window sizes, and because some feeding gestures are longer than others, it is important to identify small clusters of feeding gestures, and filte any isolated feeding gestures. Density-based spatial clustering of applications with noise (DBSCAN) is applied to group together the samples that are close together (high density), while marking as outliers the points that are lower density. A range of values is tested for the two parameters used by the DBSCAN algorithm: 1) ε is used to create an neighborhood of points to assess whether a cluster is worthy of being formed (a range from 2-4 is tested), and 2) *minPts* is used to calculate the minimum number of points required to form a dense region (a range from 1-3 is tested).

5. RESULTS

5.1 Fine-tuning Signal Processing Parameters

The average feeding gesture in the training set used was about 2.8 seconds. However, when looking purely at the LOSOCV (Figure 6) the results seem like there are several optimal parameter combinations (the blue regions have the highest F-measure). However, a very important feature in analyzing performance is variability across the different runs. Thus the optimal results from 10-fold CV (Figure 7) and LOSOCV are combined, and then the results that exhibited

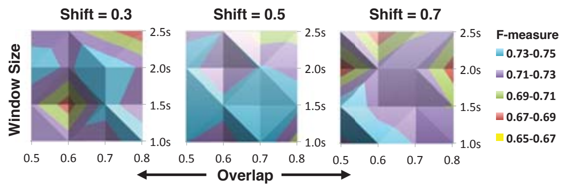

Figure 6: LOSOCV results varying signal processing parameters.

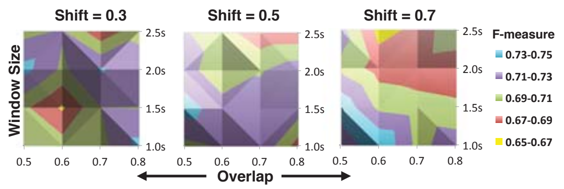

Figure 7: 10FCV results varying signal processing parameters.

the lowest variability, or highest consistency in performance across the test set, are selected.

Two parameters were discovered to yield optimal results (both high F-measure and low variability). One combination represents a slow fine-graine approach (window size = 1.0sec, overlap = 0.8, and shift = 0.3) and resulted in LOSOCV F-measure = 0.746 and variance = 0.008, and the other is more of a fast, coarse-grained approach (window size = 1.5 sec, overlap = 0.7, and shift = 0.5) and resulted in LOSOCV F-measure = 0.732, and variance = 0.02. The slow, fine-graine approach identifie feeding gestures with the following pattern "FFNFF" or "FFNNFF," where the bite in the middle of the gesture was captured as a non-feeding gesture moment (N), and the food-to-mouthand back-to-rest were detected as feeding gestures (F).

5.2 Optimal Feature Subset

The top six features (Gx_qurt1, Gx_mean, Gz_irq, Gx_stdev, and Gx_min) that were selected involved the gyroscope z- and x-axes, which correspond to the pitch and roll, respectively. Figure 5 shows the gyroscope data (the rise in G_x and drop in G_z prior to the bite) of the individual eating yogurt with a spoon.

5.3 Selecting the Optimal Classifier

It was decided to test the predictive power of multiple classifica tion algorithms to assess which classifie will outperform the other in detecting feeding gestures using the subset of features. The performance using LOSO CV was compared. It was found that while the AdaBoost classifie outperforms other classifier in LOSOCV with an F-measure of 75.7%, the Random Forest Classifie outperforms all algorithms in 10-fold CV with an F-measure of 75.2% and produces a comparable F-measure of 75.3% in LOSOCV (See Figure 8). The interesting factor arises in the fact that LOSOCV often outperforms 10-fold CV (averaged 10 times), which shows that the within-subject variability may be so high that even when the training data contains test-subject data, it remains challenging to predict feeding gestures.

5.4 Personalization

As a result of the large within-subject variability experienced within each subject, the logical next step was seen as building models based purely on each participant's data. This experiment would show the results of an individual exhaustively trained in an in-lab setting.

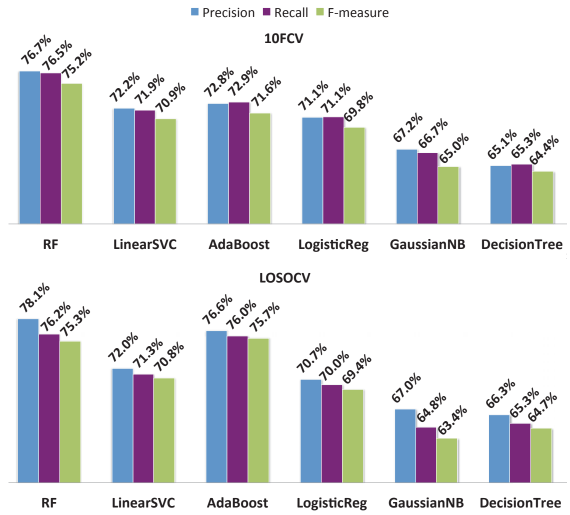

Figure 8: Optimal classifier selection.

We see that the results of a 67/33 split of train/test data yields high recall and F-measure as shown in Table 3. Subject 8 seems the most consistent in her eating pattern, and subject 14 seems to exhibit the most variability in feeding gestures.

5.5 Clustering to Count Feeding Gestures

Without clustering we realized that our classificatio model would overestimate the number of feeding gestures. This could be a result of the overlap in the sliding window. As a result we felt that applying clustering algorithms would improve our feeding gesture count. By testing a range of ε and $minPts$ values, however, we achieved poor feeding gesture count for a generalized model (RMSE = 30.2).

As a result, we tested a more personalized cluster model setting for each individual, where the model was trained on lunch data and tested on breakfast. The results for counting feeding gestures are presented in Figure 9, and show an RMSE of 8.43.

6. DISCUSSION

The results show promise in the potential to detect feeding gestures using wrist-worn sensors (the Microsoft Band 2). Challenges do exist with streaming multiple sensor data, but by adding a heart beat to reestablish connectivity a high reliability of data is ensured. While feeding gestures may not provide the granularity of caloric intake required to assess weight loss, it can provide an indicator of overeating, which can potentially be a trigger for intervention.

The results of this study show that statistical features surrounding gyroscope data are most informative, and further support prior research surroundingthe effectiveness of pitch and roll in predicting feeding gestures [21, 10]. We also show that optimizing time-series segmentation parameters can provide further insight in improving the accuracy of feeding gestures. However, the effect of f ne- and coarse-grained data analysis is shown to yield similar results. Nonetheless, personalizing segmentation parameters for each individual can still optimize the detection of feeding gestures.

The system studied shows promise in detecting feeding gestures despite several confounding factors that include talking on the phone, walking and playing with glasses. Thomaz et al. [24] show promise in designing a generalized eating detection algorithm that detects eating at a coarse level over 60-minute window sizes with an F-measure of 76.1%, however, being able to provide generalized fine grained characterization of an eating episode remains a challenge.

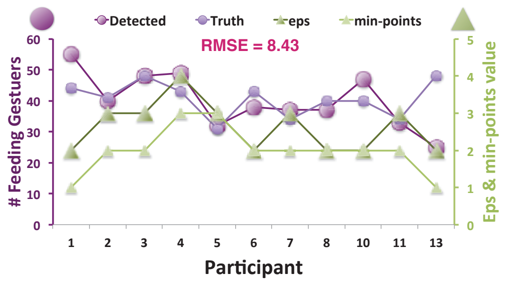

Figure 9: DBSCAN personalized clustering results.

This effort shows that within-subject variability is large and that testing a personalized model to detect feeding gestures for one individual over long periods of time will be a strong contribution to the study of eating habits and behavior.

The ultimate goal of this research is to predict problematic eating episodes. Prior to prediction, we are developing features that characterize these eating episodes in an attempt to predict them. Rahman et al. have started analyzing "about-to-eat" moments in order to trigger a just-in-time eating intervention, and attempt to estimate the time until the next eating event [20]. We hope to further expand on this effort and not only predict eating moments, but also predict problematic eating episodes.

7. CONCLUSION AND FUTURE WORK

These finding show the challenge and potential of detecting feeding gestures in a confounded eating setting. The results of a generalized and personalized model to detect feeding gestures is reported on in detail and show a significan correlation between caloric intake and feeding gestures (r=0.86). Also shown is that given the large within-subject variability in eating, a personalized machine learning and density-based clustering model can characterize eating episodes with feeding gestures with an RMSE of 8.4. To further assess our models, future testing assess whether the models used in this study will hold with participants in the wild. Future work will test different clustering methods, as well as the effect of different methods of coding feeding gestures. Moreover, combining wrist-worn sensors with other sensing modalities, such as neck-worn or ear-based sensors, has the potential to increase the accuracy of characterizing eating episodes.

8. ADDITIONAL AUTHORS

Additional authors: Angela F. Pfammatter (Northwestern University, email: angela@northwestern.edu).

9. REFERENCES

[1] N. Alshurafa, H. Kalantarian, M. Pourhomayoun, S. Sarin, J. Liu, Jason, and M. Sarrafzadeh. Non-invasive monitoring of eating behavior using spectrogram analysis in a wearable necklace. In *Healthcare Innovation Conference (HIC), 2014 IEEE*, pages 71–74, Oct 2014.

[2] N. Alshurafa, W. Xu, J. J. Liu, M.-C. Huang, B. Mortazavi, M. Sarrafzadeh, and C. Roberts. Robust human intensity-varying activity recognition using stochastic approximation in wearable sensors. In *2013 IEEE International Conference on Body Sensor Networks*, pages 1–6, May 2013.

[3] O. Amft, M. Stäger, P. Lukowicz, and G. Tröster. *Analysis of Chewing Sounds for Dietary Monitoring*, pages 56–72. UbiComp'05. Springer-Verlag, Berlin, Heidelberg, 2005.

[4] O. Amft and G. Träster. On-body sensing solutions for automatic dietary monitoring. *IEEE Pervasive Computing*, 8(2):62–70, 2009.

[5] A. Bedri, A. Verlekar, E. Thomaz, V. Avva, and T. Starner. A wearable system for detecting eating activities with proximity sensors in the outer ear. In *Proc. of the 2015 ACM Int. Symposium on Wearable Computers*, ISWC '15, pages 91–92, New York, NY, USA, 2015. ACM.

[6] P. Bongers, A. Jansen, R. Havermans, A. Roefs, and C. Nederkoorn. Happy eating: The underestimated role of overeating in a positive mood. *Appetite*, 67:74–80, 2013.

[7] L. Breiman. Random forests. *Machine Learning*, 45(1):5–32, 2001.

[8] J. Cools, D. E. Schotte, and R. J. McNally. Emotional arousal and overeating in restrained eaters. 101(2):348–351, 1992.

[9] Y. Dong, A. Hoover, and E. Muth. A device for detecting and counting bites of food taken by a person during eating. In *Proc. of the 2009 IEEE Int. Conf. on Bioinformatics and Biomedicine*, pages 265–268. IEEE, 2009.

[10] Y. Dong, A. Hoover, J. Scisco, and E. R. Muth. A new method for measuring meal intake in humans via automated wrist motion tracking. *Applied Psychophysiology and Biofeedback*, 37(3):205–215, Sep 2012.

[11] Y. Dong, J. Scisco, M. Wilson, E. R. Muth, and A. Hoover. Detecting periods of eating during free-living by tracking wrist motion. *IEEE Journal of Biomedical Health Informatics*, 18(4):1253–1260, Jul 2014.

[12] J. M. Fisher, N. Y. Hammerla, L. Rochester, P. Andras, and R. W. Walker. Body-Worn Sensors in Parkinson's Disease: Evaluating Their Acceptability to Patients. *Telemed J E Health*, 22(1):63–69, Jan 2016.

[13] A. Fouse, N. Weibel, E. Hutchins, and J. D. Hollan. Chronoviz: A system for supporting navigation of time-coded data. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '11, pages 299–304, New York, NY, USA, 2011. ACM.

[14] A. B. Goldschmidt, M. Jones, J. L. Manwaring, K. H. Luce, M. I. Osborne, D. Cunning, K. L. Taylor, A. C. Doyle, D. E. Wilfl y, and C. B. Taylor. The clinical significanc of loss of control over eating in overweight adolescents. *International Journal of Eating Disorders*, 41(2):153–158, 2008.

[15] C. G. Greeno and R. R. Wing. Stress-induced eating. 115(3):444–464, 1994.

[16] C. Herman and J. Polivy. External cues in the control of food intake in humans: The sensory-normative distinction. *Physiology and Behavior*, 94(5):722–728, 2008.

[17] O. D. Lara and M. A. Labrador. A survey on human activity recognition using wearable sensors. *Communications Surveys & Tutorials, IEEE*, 15(3):1192–1209, 2013.

[18] M. Müller. Dynamic time warping. *Information retrieval for music and motion*, pages 69–84, 2007.

[19] A. Parate, M.-C. Chiu, C. Chadowitz, D. Ganesan, and E. Kalogerakis. Risq: Recognizing smoking gestures with inertial sensors on a wristband. In *Proc. of the 12th Annual Int. Conf. on Mobile Systems, Applications, and Services*, MobiSys '14, pages 149–161, New York, NY, USA, 2014. ACM.

[20] T. Rahman, M. Czerwinski, R. Gilad-Bachrach, and P. Johns. Predicting "about-to-eat" moments for just-in-time eating intervention. In *Proc. of the 6th Int. Conf. on Digital Health Conference*, DH '16, pages 141–150, New York, NY, USA, 2016. ACM.

[21] N. Saleheen, A. A. Ali, S. M. Hossain, H. Sarker, S. Chatterjee, B. Marlin, E. Ertin, M. al'Absi, and S. Kumar. puffmarker: A multi-sensor approach for pinpointing the timing of firs lapse in smoking cessation. In *Proc. of the 2015 ACM Int. Joint Conf. on Pervasive and Ubiquitous Computing*, UbiComp '15, pages 999–1010, New York, NY, USA, 2015. ACM.

[22] E. Sazanov, P. Lopez-Meyer, and T. Stephen. A wearable sensor system for monitoring cigarette smoking. *Journal of Studies on Alcohol and Drugs*, 74(6):956–964, Nov 2013.

[23] M. Sun, L. E. Burke, Z.-H. Mao, Y. Chen, H.-C. Chen, Y. Bai, Y. Li, C. Li, and W. Jia. ebutton: A wearable computer for health monitoring and personal assistance. In *Proc. of the 51st Annual Design Automation Conf.*, DAC '14, pages 16:1–16:6. ACM, 2014.

[24] E. Thomaz, I. Essa, and G. D. Abowd. A practical approach for recognizing eating moments with wrist-mounted inertial sensing. In *Proc. of the 2015 ACM Int. Joint Conf. on Pervasive and Ubiquitous Computing*, UbiComp '15, pages 1029–1040, New York, NY, USA, 2015. ACM.