Snacap: Snacking Behavior Monitoring with Smart Fabric Mechanomyography on the Temporalis

Bo Zhou
DFKI GmbH
Kaiserslautern
bo.zhou@dfki.de

Paul Lukowicz
DFKI GmbH and TU Kaiserslautern
Kaiserslautern
paul.lukowicz@dfki.de

ABSTRACT

This paper investigates the possibility of using soft smart textiles over the hair regions to detect chewing activities under episodes of snacking in a simulated scenario with everyday activities. The planar pressure textile sensors are used to perform mechanomyography of the temporalis muscles in the form of a cap. 10 participants contributed 30 recording sessions with time periods between 30 and 60 minutes. A frequency analysis method is developed to detect moments of snacking events with continuous sliding windows on 1-second time granularity. Our approach results in a baseline 80% accuracy, over 85% after outlier removal, and above 90% accuracy for some of the participants.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in ubiquitous and mobile computing**.

KEYWORDS

automated dietary monitoring, smart textile, mechanomyography, wearable sensing

ACM Reference Format:
Bo Zhou and Paul Lukowicz. 2020. Snacap: Snacking Behavior Monitoring with Smart Fabric Mechanomyography on the Temporalis. In *ISWC '20: ACM International Symposium on Wearale Computing, September 14–16, 2020, Virtual.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

Automated Dietary Monitoring (ADM) is a central topic in automated personal health and fitness management. The food consumption process can be understood as a process of ingestion, chewing, and swallowing [18]. In the past decade, ADM has been studied extensively in these three different stages. Inertial measurement units (IMU) were used to detect the hand gestures of bringing food to the mouth during the ingestion stage in [3, 7, 20].

Sensors placed around the neck and throat have been used to detect the swallowing moments [6, 22]. In [1, 11, 12], piezo sensors which generates electrical signal under planar force were proposed as necklace-like devices. Cheng, et al. studied the capacitance variation of the neck tissue during swallowing in [5]. Contact microphones were also demonstrated at the back of the neck in [19].

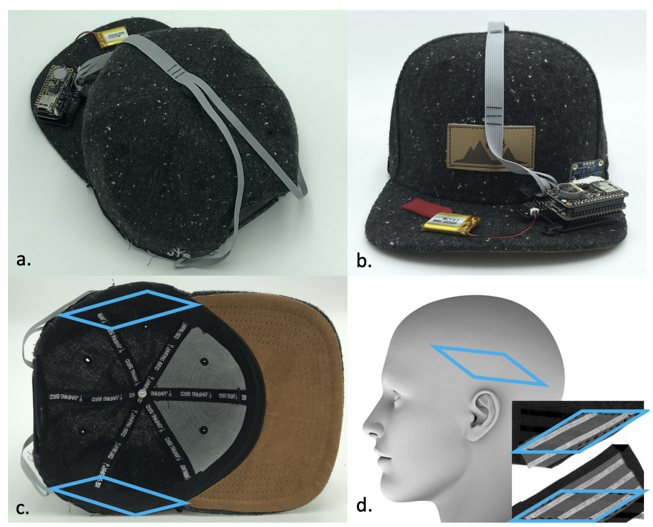

Figure 1: Experiment Apparatus. a.b. The outer view of the Snacap. c. The inner view of the Snacap. d. The fabric sensors with 2-by-3 points per patch and the position.

The chewing actions are mostly identified with the distinct sound which can be picked up by in-ear microphones [2, 15], and the jaw movements. Our jawbone moves from the skull around the temporomandibular joints as hinges with actions of adduction, abduction, protrusion and retraction during biting, chewing and grinding of food. The actions can be detected by mechanical changes around the temporomandibular joints and the ear, such as the studies in [4, 9, 10]

ISWC '20, September 14–16, 2020, Virtual

ACM ISBN 978-1-4503-XXXX-X/18/06…$15.00
https://doi.org/10.1145/1122445.1122456

with IMUs, [8] with piezofilms on a glasses frame, or [14] with a combination of photoplethysmography, audio and IMUs. The actions are controlled by the mastication muscles group, of which there are two superficial muscles that can be detected by unobtrusive methods: the temporalis (under the temple and the side of the scalp) and the masseter (under the cheek). The muscle activities have been studied with surface electromyography (EMG). In [16], generic wearable EMG recording devices were attached on the cheeks. Zhang, et al. custom-designed smart glasses with EMG electrodes in the legs, which covers the temples [21].

Contribution

In this paper, we elaborate smart textiles in the scope of detecting chewing activities from the temporalis muscle. The muscle activities are detected by surface pressure mechanomyography (MMG), which has been shown possible over the forehead with textile headbands in [23]. Compared to EMG approaches which usually come with strict skin-electrode contact requirements, the smart textile MMG sensors can be completely isolated from the user's skin with normal fabrics. In our evaluation, we also show that our approach can cope with different hairstyles between the participant's skin and the sensors, even changing hairstyles of the same person.

We implemented our approach in the prototype 'Snacap' with the form of a cap, which enables convenient data collection. The prototype is evaluated with 10 participants in simulated everyday scenarios. Our data analysis method investigates the time-frequency features of sliding windows during continuous periods, and shows above 85% accuracy with leave-person-out and leave-session-out validation conditions, with the highest individual accuracy of 95%.

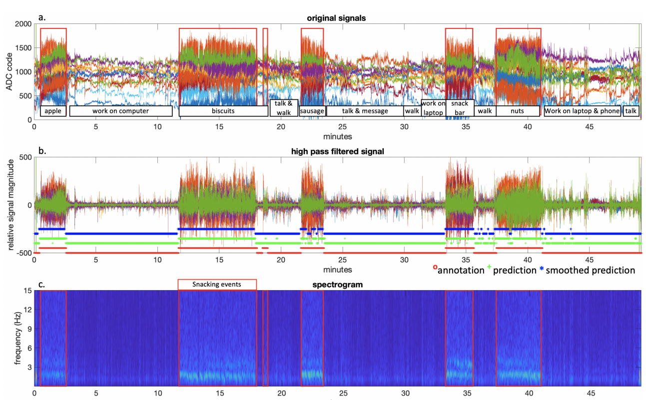

Figure 2: Signal example, annotation and prediction of a session. a. 12-channel raw signals with detailed events. b. 12-channel high-pass filtered signals with annotation, prediction and temporal smoothed predictions, Class 1 is distinguished by marker at higher positions. c. spectrogram of the average high-pass filtered signals.

2 EXPERIMENT METHODS

Apparatus

The smart fabric sensor is made of Sefar SimpleSkin fabric with silver stripes (0.7cm width with 1.5cm pitch) and Carbotex (carbonpolymer fabric). The Carbotex is sewn in the middle of two SimpleSkin layers as shown in Figure 1d. The top and bottom layers of SimpleSkin are positioned with approximate 40 degrees, which allows the sensor to cover longer area with the same amount of wires. The electrical resistance between the metal stripes from the top and bottom layer each is sensitive to the pressure force applied at each cross-points. The cap has a folded back flap as the inner rim, which makes it comfortable to wear. The smart fabric sensors are sewn under the flap rim, thus there is no direct skin contact between the sensing element and the wearer.

An Arduino board (Adafruit Huzzah32) controls the data acquisition system. For every 3-by-2 sensor patch, the side with 3 metal stripes are connected to 3 analog input channels of the Arduino board with a $1k\Omega$ grounding resistor each. The side with 2 stripes are connected to 2 digital output pins, to power one row at a time as the other pulled to ground. The hardware scans all 12 sensor channels at approximately 30Hz, as the data is saved to an SD card. All rigid electronics are secured on top of the rigid tongue of the cap.

Experiment Design

10 people (3 females, 7 males) between 20 and 31 years of age participated in the experiment, which took part in an office building. The goal of the experiment is to have the participants consume various snacks in a period of time as they perform their everyday activities, and detect the moments when they are eating. In this experiment, we only consider snacks that require chewing. The girth of the cap is adjusted for each participant to be comfortably secure fit. The participant consumed them with breaks between different snacks. To simulate daily life scenarios, during the breaks, the participants were instructed to walk around the building, talk with other people, watch videos or work on the computer, all with wearing the cap. All participants preferred to eat the snacks while seating, either in the office or the social area with sofas. Although there were short instances when they were finishing up the last few bites of a snack standing up and walking. Every session lasted between 30 and 60 minutes. The participants gave informed consent in accordance with the policies of the University of Kaiserslautern Committee for the Protection of Human Subjects, which approved the protocol. The recording is video logged for annotation. As the activities of the session is labeled as in Figure 2a., we define the snacking events as Class 1, everything else as Class 0.

Table 1: Participant Information

Person	Sessions	hairstyle	ACC(%)[1]	ACC(%)[2]
1	5	short	85.54	95.70
2	3	mid-short	57.73	68.33
3	2	mid-short	90.82	91.68
4	5	multiple long[3]	81.35	81.38
5	2	long curl	85.58	81.52
6	4	long curl	83.69	82.47
7	1	short	69.80	71.55
8	3	mid-short curl	86.33	85.70
9	2	mid-short	88.21	88.25
10	3	mid-short	86.71	85.88

1: leave person out cross validation
2: leave session out cross validation
3: Person 4 has changed the hairstyle among straight long, tied-back curls, box braids, tied-back box braids in different sessions, flat hair clips were under the sensor for the tied-back hairstyles.

Ultimately, the snacks included in the experiment are: apples (sliced or whole), bananas, biscuits, chewing gums, chocolates, cookies, gummy bears, nachos, nuts, muffins, potato chips, protein bars, salami mini sausages, snack bars of nuts, tangerines, waffles, yogurts with cereal.

The participants' details are summarized in Table 1, as overall 30 sessions were recorded.

3 DATA ANALYSIS METHOD

Signal Processing

The raw signals $S_n(t)|n \in [1, 12]$ of the 12 channels from a session of participant 1 are plotted in Figure 2a. First, we can observe that during snacking events, there are clear high frequency patterns, which can be better visualized in a zoomed-in figure in Figure 3. Our following signal processing thus is aimed at exploring the frequency domain. Also, among different channels there are obvious bias. The bias persists and varies among different participants and sessions. We assume this is caused by the contour of the head structure or the hairstyle. To remove the bias, the raw signals are processed with a zero-phase, IIR Butterworth highpass filter with 0.1Hz stop frequency and 1Hz pass frequency. The highpass filtered signals $S_{hp,n}(t) = filter_{hp}(S_n(t))$ are shown in Figure 2b. The individual highpassed signals are then averaged $M_{hp} = \sum_{n=1}^{12} S_{hp,n}(t)$. The spectrogram of M_{hp} is shown in Figure 2c. We observe from the spectrograms of all sessions, that during snacking events, the major magnitude is approximately in the range of 1 to 3 Hz, which coincides with the study on the cheek masseter EMG signals during chewing actions in [16]. Since there may be cancellations between positive and negative values during the average calculation, we also calculate the average of the absolute values of the highpass filtered signals for further processing $L_{hp} = \sum_{n=1}^{12} |S_{hp,n}(t)|$.

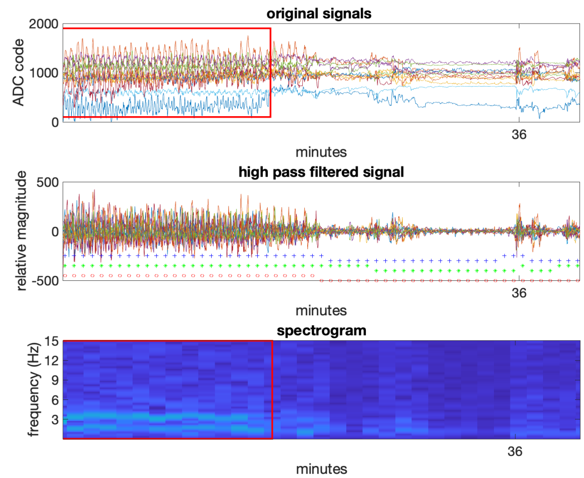

Figure 3: The zoomed signal of Figure 2 around 36 min.

Feature Extraction

Sliding windows are then used to sweep through the data and calculate features f_i. The temporal features are inspired by previous studies with the similar textile mechanomyography, such as the study on the quadriceps in [24, 25]. Within every window T, we first calculate the statistical presentation: f_1 average, f_2 variance, f_3 range, f_4 skewness, f_5 kurtosis, f_6 waveform length [13], f_7 sum of values greater than mean.

Then we calculate the power spectrum density of the window (PSD) with the fast Fourier transform to quantify the frequency characteristics. Then we calcualte the average magnitude of the PSD as feature f_8 and the mean frequency (the weighted center of the PSD as f_9. Then the PSD is divided to 5 equal frequency bands, the average values of each band is calculated as $f_{10} \sim f_{14}$ (from low to high frequency).

Wavelet transform is also an effective method for frequency domain analysis. We then perform fast wavelet transform using the LTFAT toolbox [17] with J=4 filterbank iterations and 'Daubechies 8' as the mother wavelet. Every filterbank iteration ($j \in [0, 4]$) generates a coefficient vector of varying lengths. We then calculate statistical presentations of each coefficient vector: the mean value (f_{15}, f_{20}, f_{25}, f_{30}, f_{35}), the variance (f_{16}, f_{21}, f_{26}, f_{31}, f_{36}), the range (f_{17}, f_{22}, f_{27}, f_{32}, f_{37}) the skewness (f_{18}, f_{23}, f_{28}, f_{33}, f_{38}) and the kurtosis (f_{19}, f_{24}, f_{29}, f_{34}, f_{39}).

Thus, 39 features are calculated separately for M_{hp} and L_{hp}; overall 78 features are calculated from every time window.

Machine Learning

We use the Matlab®Classification Learner to evaluate classifier models. First, the entire dataset with 25% hold out validation is used to search for the best performing classifier model, which resulted in the support vector machine model (SVM) with the Gaussian kernel function. Obviously, random partitioning with the entire dataset does not represent real-world applications. Therefore, we perform leave-session-out and leave-person-out with the exported SVM classifier fitting function. In leave-session-out, one session from each participant is left out for testing while the remaining are in the training data. Five iterations are performed to have every session in the testing data at least once.[1] In leave-person-out, every iteration leaves out all sessions from one participant for testing, while all the remaining participants' data are used for training the model.[2] We vary the sliding window's parameter in the leave-session-out scheme to have the optimal accuracy, which resulted in 10 seconds window size and 1 second window step.

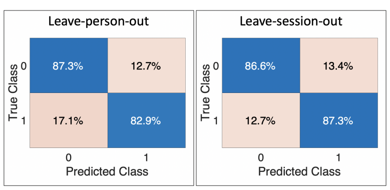

Figure 4: The confusion matrices of detecting snacking moments after removing the two outlier participants.

4 RESULTS AND DISCUSSION

The individual leave-session-out and leave-person-out accuracy values are shown in Table 1, with the average accuracy 81.73% for leave-person-out and 83.25% for leave-session-out. Except for Person 2 and 7 (who only recorded one session before the lock down), every participant has above 80% accuracy. We notice that for Person 2, the frontal distance between their temples are narrower than the span of the rigid tongue of the cap, leaving a gap between their temples and the sensing fabric. This can be addressed by customized headwear design in the future. We still consider our exploratory method with no skin contact, soft fabric sensors advantageous in terms of user comfort and acceptance. Notably, with Person 4, who has changed 3 hairstyles in different sessions, has also shown above 80% accuracy. If we consider the two participants as outliers and remove their results, the average accuracy is then 85.93% for leave-person-out and 86.57% for leave-session-out.

With the sliding window approach along the continuous periods of time, we can also perform fine time granularity spotting. In Figure 2b., the green markers show the snacking moments that are detected by the SVM model trained with the leave-person-out approach. For the scope of this study, the spotting performance can be considered the same as the leave out validation.

In the related studies that investigate hand gestures, the achieved accuracy is typically between 75% and 85% [3, 7, 20]. As the sensors are positioned directly at the neck, the accuracy values are around 85% to 90% [1, 5, 11, 12]. Around 90% accuracy values have been reported in the studies that place various sensors at the ear and the temporomandibular joints [4, 9, 10, 14]. While the highest accuracy values are reported in studies with sensor placements above the ear and around the temples. The studies in [8] with piezo films in a glass frame and in [21] at the temples achieved 99% accuracy/F1-score in detecting chewing activities. It is worth nothing that the classification in [8] is performed on pre-segmented epochs of activities under controlled chewing activities in a laboratory setting; while our study is based on continuous sliding windows along a period of time under a simulated scenario with everyday activities. The F1-score in [21] is also the measure of complete eating events.

We can observe a trend that wearable sensing approaches are better at detecting eating activities when the sensors are closely coupled with the anatomical structures that control the chewing actions. Overall, our textile approach with hairs of varying thickness between the sensors and the skin provides comparable recognition accuracy with related studies.

5 CONCLUSION AND OUTLOOK

This study have explored using smart fabric in normal headwears over the hair region to detect snacking activities. This is done by perform surface pressure mechanomyography at the temporalis muscles which are responsible for pulling up the jaw during chewing actions. The prototype Snacap takes the form of a smart cap, which utilizes the existing structure of the headwear to accommodate the sensing hardware.

In our future work, we would first look into integrating the hardware more unobtrusively into the headwear. A possible direction is to have flattened electronics which can be embedded inside the cap tongue. Other headwear such as glasses and sweatbands will also be evaluated. With more integrated and robust system, we would also perform longer period user studies.

REFERENCES

[1] Nabil Alshurafa, Haik Kalantarian, Mohammad Pourhomayoun, Jason J Liu, Shruti Sarin, Behnam Shahbazi, and Majid Sarrafzadeh. 2015.

[1]Since the participants' sessions are not even, some sessions are left out more than once in different iterations.

[2]For Person 7, the only recorded session is always in the testing data for leave-session-out and leave-person-out. The difference is for leave-session-out, one session from each participant is not present in the training data.

Recognition of nutrition intake using time-frequency decomposition in a wearable necklace using a piezoelectric sensor. *IEEE sensors journal* 15, 7 (2015), 3909–3916.

[2] Oliver Amft, Mathias Stäger, Paul Lukowicz, and Gerhard Tröster. 2005. Analysis of chewing sounds for dietary monitoring. In *International Conference on Ubiquitous Computing*. Springer, 56–72.

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

[4] Abdelkareem Bedri, Richard Li, Malcolm Haynes, Raj Prateek Kosaraju, Ishaan Grover, Temiloluwa Prioleau, Min Yan Beh, Mayank Goel, Thad Starner, and Gregory Abowd. 2017. EarBit: using wearable sensors to detect eating episodes in unconstrained environments. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies* 1, 3 (2017), 1–20.

[5] Jingyuan Cheng, Bo Zhou, Kai Kunze, Carl Christian Rheinländer, Sebastian Wille, Norbert Wehn, Jens Weppner, and Paul Lukowicz. 2013. Activity recognition and nutrition monitoring in every day situations with a textile capacitive neckband. In *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication*. 155–158.

[6] Keum San Chun, Sarnab Bhattacharya, and Edison Thomaz. 2018. Detecting eating episodes by tracking jawbone movements with a non-contact wearable sensor. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 1 (2018), 1–21.

[7] Yujie Dong, Jenna Scisco, Mike Wilson, Eric Muth, and Adam Hoover. 2013. Detecting periods of eating during free-living by tracking wrist motion. *IEEE journal of biomedical and health informatics* 18, 4 (2013), 1253–1260.

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

[9] Muhammad Farooq and Edward Sazonov. 2018. Accelerometer-based detection of food intake in free-living individuals. *IEEE sensors journal* 18, 9 (2018), 3752–3758.

[10] Juan M Fontana, Muhammad Farooq, and Edward Sazonov. 2014. Automatic ingestion monitor: A novel wearable device for monitoring of ingestive behavior. *IEEE Transactions on Biomedical Engineering* 61, 6 (2014), 1772–1779.

[11] Ghulam Hussain, Kamran Javed, Jundong Cho, and Juneho Yi. 2018. Food intake detection and classification using a necklace-type piezoelectric wearable sensor system. *IEICE TRANSACTIONS on Information and Systems* 101, 11 (2018), 2795–2807.

[12] Haik Kalantarian, Nabil Alshurafa, and Majid Sarrafzadeh. 2014. A wearable nutrition monitoring system. In *2014 11th International Conference on Wearable and Implantable Body Sensor Networks*. IEEE, 75–80.

[13] Zeeshan O Khokhar, Zhen G Xiao, and Carlo Menon. 2010. Surface EMG pattern recognition for real-time control of a wrist exoskeleton. *Biomedical engineering online* 9, 1 (2010), 41.

[14] Vasileios Papapanagiotou, Christos Diou, Lingchuan Zhou, Janet van den Boer, Monica Mars, and Anastasios Delopoulos. 2016. A novel chewing detection system based on ppg, audio, and accelerometry. *IEEE journal of biomedical and health informatics* 21, 3 (2016), 607–618.

[15] Sebastian Passler and Wolf-Joachim Fischer. 2011. Food intake activity detection using a wearable microphone system. In *2011 Seventh International Conference on Intelligent Environments*. IEEE, 298–301.

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

[17] Zdeněk Prša, Peter L Søndergaard, Nicki Holighaus, Christoph Wiesmeyr, and Peter Balazs. 2013. The large time-frequency analysis toolbox 2.0. In *International Symposium on Computer Music Multidisciplinary Research*. Springer, 419–442.

[18] Giovanni Schiboni and Oliver Amft. 2018. Automatic dietary monitoring using wearable accessories. In *Seamless healthcare monitoring*. Springer, 369–412.

[19] Jaemin Shin, Seungjoo Lee, and Sung-Ju Lee. 2019. Accurate Eating Detection on a Daily Wearable Necklace. In *Proceedings of the 17th Annual International Conference on Mobile Systems, Applications, and Services*. 649–650.

[20] 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*. 1029–1040.

[21] Rui Zhang and Oliver Amft. 2020. Retrieval and Timing Performance of Chewing-Based Eating Event Detection in Wearable Sensors. *Sensors* 20, 2 (2020), 557.

[22] Shibo Zhang, Yuqi Zhao, Dzung Tri Nguyen, Runsheng Xu, Sougata Sen, Josiah Hester, and Nabil Alshurafa. 2020. NeckSense: A Multi-Sensor Necklace for Detecting Eating Activities in Free-Living Conditions. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 2 (2020), 1–26.

[23] Bo Zhou, Tandra Ghose, and Paul Lukowicz. 2020. Expressure: Detect Expressions Related to Emotional and Cognitive Activities Using Forehead Textile Pressure Mechanomyography. *Sensors* 20, 3 (2020), 730.

[24] Bo Zhou and Paul Lukowicz. 2019. TPM Feature Set: a Universal Algorithm for Spatial-Temporal Pressure Mapping Imagery Data. In *IARIA. International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies (UBICOMM-2019), September 22-26, Porto, Portugal*. IARIA XPS Press.

[25] Bo Zhou, Mathias Sundholm, Jingyuan Cheng, Heber Cruz, and Paul Lukowicz. 2017. Measuring muscle activities during gym exercises with textile pressure mapping sensors. *Pervasive and Mobile Computing* 38 (2017), 331–345.