MONITORING OF EATING BEHAVIOR

USING SENSOR-BASED

METHODS

by

DELWAR HOSSAIN

EDWARD SAZONOV, COMMITTEE CHAIR
FEI HU
AIJUN SONG
SEVGI ZUBEYDE GURBUZ
MEGAN ANN MCCRORY

A DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the Department of Electrical & Computer Engineering
in the Graduate School of
The University of Alabama

TUSCALOOSA, ALABAMA

2024

Copyright Delwar Hossain, 2024
ALL RIGHTS RESERVED

ABSTRACT

The essential physiological functions of the human body, including respiration, circulation, physical exertion, and protein synthesis, rely on energy from dietary constituents. Understanding eating behavior is crucial for overall health, as deviations in energy intake can lead to malnutrition-induced weight loss or obesity-related weight gain.

Traditionally, dietary intake assessment has relied on self-reporting methods, such as dietary records, 24-hour recalls, and food frequency questionnaires. While these methods help understand relationships between eating behavior and dietary intake, they lack the granularity needed to explore detailed food consumption processes. Therefore, there is a need for innovative solutions that enable objective, precise, and automated monitoring of eating behavior, especially in free-living conditions.

This dissertation investigates the application of wearable sensor systems for the automatic monitoring of eating behavior with minimal effort from subjects. First, a systematic review was conducted to identify available technology-driven methods for monitoring eating behavior. Then, a novel, contactless method for detecting and measuring eating behaviors such as chews and bites from eating videos was developed. Then an algorithm was devised to evaluate and compare different sensor modalities for identifying eating behavior, specifically focusing on chewing and chewing strength measurement. Four sensor modalities—Ear Canal Pressure Sensor, Piezoresistive Bend Sensor, Piezoelectric Strain Sensor, and EMG Sensor—were assessed. Results indicated comparable efficacy across all four systems in identifying chewing and chewing strength.

Next, a novel Ear Canal Pressure Sensor system was explored for monitoring eating behavior, particularly chewing, in free-living environments. The findings demonstrated accurate detection and estimation of chewing in both controlled and free-living settings. Finally, a machine learning model to estimate energy intake (EI) from food intake using sensor-captured eating behavior features was developed and evaluated in free-living settings. The results highlight the efficacy of the sensor-based EI model and the potential for improved accuracy by leveraging image assistance and automatic food item detection.

In conclusion, this research advances eating behavior monitoring using wearable sensor technologies. The findings hold promise for personalized nutrition interventions and mark a significant step forward in the objective assessment of eating habits.

DEDICATION

This work is dedicated to my beloved family, friends, and mentors.

ACKNOWLEDGEMENTS

I extend my sincere gratitude to the University of Alabama for providing an enriching environment that has fostered academic, personal, and spiritual growth during my tenure. This journey has been made possible through the unwavering support and encouragement of numerous individuals.

First and foremost, my deepest thanks go to my family, whose constant encouragement, inspiration, and unconditional love have been the bedrock of my journey. I am profoundly grateful to my advisor, Dr. Edward Sazonov, for his invaluable guidance, unwavering support, and immense knowledge, making this work possible. Dr. Sazonov's patience, motivation, and open-door policy for addressing doubts and encouragement were indispensable.

I extend my appreciation to my thesis committee members, Dr. Megan A. McCrory, Dr. Aijun Song, Dr. Fei Hu, and Dr. Sevgi Zubeyde Gurbuz, for their insightful comments and constructive criticism. Special acknowledgment goes to Dr. Muhammad Farooq, Dr. Sayeed Doulah, Dr. Masudul Haider Imtiaz, and Dr. Tonmoy Ghosh, my mentors during the early stages of my graduate studies along with Dr. Edward Sazonov.

I am grateful for the support and encouragement from the members of our research group, past and present, including, Rafi Islam, Billal Hossain, Siavash Esfandiari Fard, Dr. Viprav Raju, Dr. Mohamed Hassan, Dr. Kabir Hossain, and Dr. Prajakta Belsare whose discussions consistently led to improved solutions.

To my friends, who added richness to my studies, I express humble gratitude. Special thanks to Dr. Abu Horaira Banna, Dr. Abu Sufian, Tareque Sagar, Abdul Halim, Asaduzzaman

Nayan, Mustafiz Vai, Mohon, Roni Paul, and Himdary Das for their unwavering support during moments of uncertainty.

I would also like to acknowledge the BSA UA, Brothers sporting club at UA, University of Alabama Cricket Club (UACC), Tuscaloosa community members, University of Alabama International Student and Scholar Services, and Tuscaloosa International Friend for their kindness and friendliness, contributing to the wonderful experience of my graduate studies.

CONTENTS

ABSTRACT ii

DEDICATION iv

ACKNOWLEDGEMENTS v

LIST OF TABLES xi

LIST OF FIGURES xii

CHAPTER 1 – INTRODUCTION 1

1.1 Motivation 1

1.2 Overview of Dissertation 3

References 5

CHAPTER 2 – A SYSTEMATIC REVIEW OF TECHNOLOGY-DRIVEN METHODS FOR MEASUREMENT OF EATING BEHAVIOR 7

Abstract 7

2.1 Introduction 8

2.2 Review Methodology 11

2.2.1 Research Questions Identification 11

2.2.2 Databases 11

2.2.3 Search Strategy 11

2.2.4 Inclusion and Exclusion Criteria 14

2.2.5 Results 15

2.3 Review Findings 16

2.3.1 Taxonomy 16

2.3.2 Physiological/Environmental Phenomena and Computed Metrics: 17

2.3.3 Measurement technology/device 22
2.4 Discussion 53
2.5 Challenges and Future Trends 63
2.6 Conclusion 65
References 66
CHAPTER 3 – AUTOMATIC COUNT OF BITES AND CHEWS FROM VIDEOS OF EATING EPISODES 84
Abstract 84
3.1 Introduction 85
3.2 Methods and Materials 87
3.2.1 Data Collection Protocol 87
3.2.2 Face Detection to Extract ROI 89
3.2.3 Image Classification To Count Bites 92
3.2.4 Optical Flow to Count Number of Chews 94
3.2.5 Bite and Chew Counting 97
3.3 Results 100
3.4 Discussion 101
3.5 Conclusion 105
References 106
CHAPTER 4 – COMPARISON OF WEARABLE SENSORS FOR ESTIMATION OF CHEWING STRENGTH 111
Abstract 111
4.1 Introduction 112
4.2 Materials and Methods 115
4.2.1 Wearable Sensors 115
4.2.2 Test Foods 118

4.2.3 Objective Hardness Evaluation 118
4.2.3 Participants 119
4.2.4 Data Collection 119
4.2.5 Signal Processing 120
4.2.6 Statistical Analysis 122
4.3 Results 122
4.3.1 Objective Hardness Evaluation 122
4.3.2 Pressure Sensor 124
4.3.3 Piezoresistive Bend Sensor 125
4.3.4 Piezoelectric Strain Sensor 127
4.3.5 EMG Sensor 129
4.4 Discussion 130
4.5 Conclusions 133
References 135
CHAPTER 5 – EAR CANAL PRESSURE SENSOR FOR FOOD INTAKE DETECTION AND CHEW COUNTING IN FREE LIVING 139
Abstract 139
5.1 Introduction 140
5.2 Materials and Methods 144
5.2.1 Proposed ECPS (Ear Canal Pressure Sensor) 144
5.2.2 Study Protocol 146
5.2.3 Data Collection and Annotation 148
5.2.4 Food Intake Detection 149
5.3 Results 159
5.4 Discussion 162
5.5 Conclusion 166

References 168

CHAPTER 6 – DEVELOPMENT OF ENERGY INTAKE MODEL FROM SENSOR AND IMAGE CAPTURED EATING IN FREE LIVING 174

Abstract 174

6.1 Introduction 175

6.2 Methods 178

6.2.1 Data Collection Protocol 178

6.2.2 Food Intake and Eating Episode Detection 183

6.2.3 EI Model Training 183

6.2.4 EI Model Evaluation 185

6.3 Results 188

6.4 Discussion 190

6.5 Conclusion 194

References: 195

CHAPTER 7 – CONCLUSIONS AND FUTURE WORK 202

APPENDIX I PUBLICATIONS 207

APPENDIX II IRB APPROVAL FOR EVALUATION OF WEARABLE DEVICE FOR FOOD INTAKE AND INGESTIVE BEHAVIOR MONITORING 209

APPENDIX III IRB APPROVAL FOR WEIGHED RECORD FOOD INTAKE DATA COLLECTION USING A WEARABLE DEVICE TO DEVELOP MODELS FOR ENERGY INTAKE 210

LIST OF TABLES

Table 2-1: Search Query Strings to obtain candidate articles. 13

Table 2-2: Eligibility Criteria 14

Table 2-3: Summary of the measurement Devices for Biting 54

Table 2-4: Summary of measurement devices for Chewing 56

Table 2-5: Summary of measurement devices for swallowing 59

Table 2-6: Summary of measurement devices for Food Items. 60

Table 2-7: Summary of measurement devices for Eating Time/Rate 62

Table 2-8: Summary of measurement devices for Mass of food intake 63

Table 4-1: Objective hardness measurement record of test foods 123

Table 4-2: Summarized result of Single Factor ANOVA 132

Table 5-1: Study protocol 148

Table 5-2: Description of selected features for $CLP + ACC$ 154

Table 5-3: Precision, Recall (Sensitivity), and F1-Score for different epochs of Laboratory Experiments by classifier $CLP + ACC$ 160

Table 5-4: Precision, Recall (Sensitivity), and F1-Score for different epochs of Laboratory Experiments by classifier CLP 160

Table 5-5: Chew count of each subject using CLP classifier. 161

Table 5-6: Eating episode detection of free-living data by both classifiers. 161

Table 5-7: Manual review of captured images (Free-Living) 162

LIST OF FIGURES

Figure 2-1: Flow chart outlining the study selection process of the systematic review. 15

Figure 2-2: Taxonomy of quantifiable metrics and measurement technologies for food consumption. 17

Figure 2-3: Summary of findings: (a) Number of articles found for quantifiable metrics of eating behavior (b) Percentage of articles found for the measurement devices/technologies used to measure the eating behavior metrics. 53

Figure 3-1: Camera positions and Data collection setup 87

Figure 3-2: Manual Annotation of Bite and Chew from video (a) The three-button system for annotation of the video (b) Annotated bite and chews in a meal video (c) Annotated bite and chew of a chewing sequence. 88

Figure 3-3: Architecture of Faster R-CNN object detector 90

Figure 3-4: Structure and layers of face detector 90

Figure 3-5: Labeled training Image for face detector. 91

Figure 3-6: Cropped faces of an eating episode to detect a bite. 93

Figure 3-7: Structure and Layers of Image Classifier to classify bite. 95

Figure 3-8: Affine optical flow to find chewing. The first and last image show the jaw opening and the middle image shows the jaw closing movements during chewing. 96

Figure 3-9: Flowchart of developed bite and chew counting algorithm from meal video 98

Figure 3-10: Prediction of bite count from image classification. (Top figure represents raw prediction, middle two figures represent filtering and thresholding to remove false detection, the bottom figure represents the ground truth obtained from manual annotation). 99

Figure 3-11: Effect of bite and chew in optical flow values. 100

Figure 3-12: (a) False bite event detection because of gestures mimicking bite (b)Cumulative chew count from manual annotation and automatic count 101

Figure 3-13: 1-D optical flow values from cropped images of an eating episode 103

Figure 4-1: Four wearable sensor systems: (a) Ear Canal Pressure Sensor 116

Figure 4-2: Truncation of full signal for separate test food. 121

Figure 4-3: (a) Normalized pressure sensor signal (b) Signals of five chews of test foods (c)Variation of standard deviation in chewing test foods (d) Box plot of chewing strength estimation for three test foods. 124

Figure 4-4: Post-hoc analysis of variation using Tukey's test for pressure sensor. 125

Figure 4-5: (a) Normalized Piezoresistive bend sensor signal (b) Signals of five chews for test foods (c)Variation of standard deviation in chewing test foods (d) Box plot of chewing strength estimation for three test foods. 126

Figure 4-6: Post-hoc analysis of variation using Tukey's test for piezoresistive bend sensor. .. 126

Figure 4-7: (a) Normalized strain sensor signal (b) Signals of five chews for test foods (c)Variation of standard deviation in chewing test foods (d) Box plot of chewing strength estimation for three test foods 127

Figure 4-8: Post-hoc analysis of variation using Tukey's test for piezoelectric strain sensor.... 128

Figure 4-9: (a) Normalized EMG sensor signal (b) Signals of five chews for test foods (c)Variation of standard deviation in chewing test foods (d) Box plot of chewing strength estimation for three test foods. 129

Figure 4-10: Post-hoc analysis of variation using Tukey's test for EMG sensor. 130

Figure 5-1: A portable wearable device for monitoring food intake. The module is affixed to right side of a behind-neck headset. (a) Represents the ear canal deformation during chewing (b)represents the wearable sensor system (c) the individual wearing the ECPS prototype. 144

Figure 5-2: The electronics of ECPS. (Left) ECPS subsystems: Pressure Sensor, USB, MCU, Framebuffer, Camera; (right) Overview of the system with subsystems interconnections 145

Figure 5-3: The first row demonstrates the pressure sensor signal. Signals from the 3-axes accelerometer are shown on the second row. The foot pedal signal marked is shown on the third row. Here '1' indicates food intake and '0' indicates no food intake. 150

Figure 5-4: Three stages of the proposed algorithm to count the number of chews. 155

Figure 5-5: Chewing segment detection using classification labels and segmentation. The topfigure represents the filtered pressure sensor signal. The second top figure represents the label from segmentation. The third from the top figure represents the label from food intake classification. The bottom figure represents the detected chewing sequence. 157

Figure 5-6: This figure represents food intake detection in free-living using pressure sensor signal and accelerometer features for a 5-sec epoch size. The top figure represents the raw prediction. The epoch is labeled as '1' if it is predicted as a food intake epoch. Otherwise, epoch is labeled as '-1'. Kernel smoothing was performed to remove false positives applying fixed threshold. The middle and bottom figure represents the final prediction and the self-reported start and ends-time of the eating episode respectively. The epoch value is set as '1' if epoch is classified as food intake. Otherwise, the epoch value is set as -1. 158

Figure 5-7: Cumulative chew count of a meal. The blue line indicates result of the proposed chew counting algorithm and the red line represents the result of the manual annotation. 163

Figure 6-1: AIM-2 attached to the right temple of off-the-shelf eyeglasses to monitor intake. (a)original AIM-2 with flexible bend sensor contacts skin of the temporalis muscle to track movement [37] (previous study) (b) AIM-2 with an optical sensor that tracks temporalis muscle movement without any contact with the skin (current study). 179

Figure 6-2: Sensor signals of an eating event. Captured images are shown at the top. Signals from the 3-axes of the accelerometer are shown on the bottom. The middle segment demonstrates the optical sensor signal. 180

Figure 6-3: Steps of data processing for EI Estimation (a) represents the steps of the EI estimation from sensor signals and image information (b) graphical representations of EI estimation from the sensor signals and image information 182

Figure 6-4: (a) Bland-Altman plot showing mean difference and 2 SD limits of agreement (b)Scatter plot of EI (WFR) vs EI(Prediction) of eating episodes from LOSOCV 188

Figure 6-5: EER (Kcal/day) calculated from participant's information using equation. 189

Figure 6-6: Violin and Box plot of EI/day from blind EI estimation method, image-assisted EI estimation method, and EI estimation from manual annotation by nutritionists. 190

Figure 6-7: (a) Daily EI estimation by blind EI method. False eating episode detection from sensors causes over-estimation in prediction (b) Daily EI estimation by image-assisted EI method. False eating episodes are removed by manual image review. 191

CHAPTER 1 – INTRODUCTION

1.1 Motivation

Eating is a multifaceted interplay of physiological, emotional, social, cultural, environmental, and economic factors. Dietary intake refers to the types and amounts of food and drinks consumed over time, measured in calories, macronutrients, and micronutrients. Eating behavior involves how a person eats, including habits, patterns, meal timing, portion sizes, and food choices, influenced by various social and emotional factors. While related, dietary intake and eating behavior are distinct concepts; understanding both is crucial for promoting healthy eating habits and overall well-being. A deep understanding of eating behavior is vital for overall health, as deviations in energy intake can lead to conditions like malnutrition-induced weight loss or obesity-related gain [1]. Examining eating behavior is also crucial for identifying patterns associated with eating disorders such as anorexia nervosa, bulimia, and binge eating [2].

Assessing dietary intake in various age groups has historically relied on self-report methods such as 24-hour recalls, food records, and food frequency questionnaires [3]-[5]. While these tools aid in understanding the link between eating behavior and dietary intake, they lack the granularity needed to measure and explore crucial aspects of the food consumption process [6]-[8]. Technological progress has facilitated the creation of precise and objective systems for measuring dietary intake, with potential applications in assessing eating behaviors. Wearable sensors, detect chewing or swallowing by sensors placed in the head or neck [9] - [13], wrist-based inertial sensors for hand-to-mouth gestures [14] - [15], and camera-based methods

utilizing food images [16] - [19], offer innovative approaches. These advancements enable researchers to accurately identify consumed food items and estimate energy intake, contributing to a more comprehensive understanding of eating episodes.

In this study, available wearable systems to monitor and track dietary intake were explored to monitor and estimate eating behaviors. Most of the studies performed to explore dietary intake use eating episode videos to annotate eating behaviors such as chews, bites, etc. However little work has been done to make the annotation automatic and easy. A novel computer vision methodology was used to develop a contactless, non-invasive algorithm to detect and count these eating behaviors from eating videos which paved the way for future research to make this cumbersome and time-consuming annotation automatic in this dissertation. Though different wearable sensor systems can detect and count chewing, no previous work has been done to explore and compare these modalities in finding the efficacy of identifying chewing strength for different foods with different textures. In this dissertation, we explored four different sensor modalities in identifying the variation of chewing strength for consuming food with different textures. A novel wearable sensor system that uses ear canal pressure change while eating to detect and estimate chewing in free-living settings. Finally, a machine learning model to estimate EI consumption for eating episodes was developed from the sensor-captured eating behavior measures. The study underscores the success of a sensor-based Eating Intervention (EI) model, highlighting improved estimation accuracy through image assistance and automatic food item detection. This research advances eating behavior monitoring with wearable sensor technologies, holding promise for personalized nutrition interventions, and marking a significant step in the objective assessment of individuals' eating habits.

1.2 Overview of Dissertation

The systematic review to identify available technology-driven methodologies to monitor and estimate eating behavior is performed in Chapter 2. The potential future direction of research in eating behavior assessment and monitoring was evaluated considering the advantages and drawbacks of earlier approaches.

Chapter 3 includes the development of a novel, non-invasive methodology to monitor and estimate eating behavior by detecting and counting chews and bites from eating episode videos in the controlled laboratory setting. 84 eating episode videos from 28 participants were used to develop and evaluate the algorithm leveraging object detection, image classification, and computer vision algorithms. The findings of this study encourage the use of the proposed algorithm in feeding studies for annotation of eating behavior such as chews and bites instead of cumbersome and time-consuming manual annotation. Chapter 4 compares four sensor modalities including 1) Ear canal pressure sensor 2) Piezoresistive bend sensor 3) Piezoelectric strain sensor 4) EMG sensor in identifying and differentiating chewing strength in eating food with different textures and hardness. Data from 15 participants were used to perform statistical analysis such as ANOVA and Tukey's test to provide the conclusion. The results of this study indicate that all four sensor modalities can effectively identify the difference in chewing strength when eating food with different textures. Chapter 5 provides the development and validation of a novel wearable device that uses changes in ear canal pressure and the device's acceleration to detect food intake and estimate chewing in a free-living setting. Chewing is associated with the opening and closing of the mouth by moving the lower jaw, which causes the ear canal to expand as the mandible condyle slides forward and backward. The movement of the lower jaw creates a void that will be filled by the tissue surrounding the ear. This deformation of the ear canal associated

with chewing during solid food intake produces a change in pressure which can be used to monitor eating behavior. To validate the proposed device data were collected from 10 volunteers in a controlled environment and three volunteers in a free-living environment. Data from controlled experiments was used to train a Support Vector Machine (SVM) classifier to detect eating events and to develop a method to identify chewing segments and count the number of chews in each eating event. Results from this study highlighted that the proposed wearable device could detect eating events and count the number of chews with high accuracy and precision in free-living settings. Chapter 6 includes the development, evaluation, and improvement of an EI model from sensor-captured eating behavior features in free living. In this study, we used Automatic Ingestion Monitor V2 (AIM-2) to monitor eating behavior. AIM-2 uses an optical sensor and a 3-axis accelerometer to detect et eating events. The device also housed a 5Megapixel camera with a 120-degree wide-angle gaze-aligned lens to capture egocentric images when eating is detected. A training dataset including 30 participants wearing the device for 2 free-living days and providing weighted food records was used to train a regression model using eating behavior features such as chew count, chewing duration, pause duration, meal duration, etc. to estimate EI. The trained model was then evaluated in a separate evaluation dataset including 39 participants wearing the device for a week in a free-living environment. The findings highlight the efficacy of the sensor-based EI model and underscore the potential for improved estimation accuracy by leveraging image assistance and automatic food item detection. This research holds promise for advancing the field of dietary monitoring and facilitating personalized nutrition interventions based on eating behavior.

References

[1] Y. Schutz, “Macronutrients and energy balance in obesity,” Metabolism, vol. 44, pp. 7–11, Sep. 1995, doi: 10.1016/0026-0495(95)90311-9.

[2] B. T. Walsh, “The importance of eating behavior in eating disorders,” Physiol. Behav., vol. 104, no. 4, pp. 525–529, Sep. 2011, doi: 10.1016/j.physbeh.2011.05.007.

[3] S. S. Jonnalagadda et al., “Accuracy of Energy Intake Data Estimated by a Multiplepass, 24-hour Dietary Recall Technique,” J. Am. Diet. Assoc., vol. 100, no. 3, pp. 303–311, Mar. 2000, doi: 10.1016/S0002-8223(00)00095-X.

[4] N. Day, N. McKeown, M. Wong, A. Welch, and S. Bingham, “Epidemiological assessment of diet: a comparison of a 7-day diary with a food frequency questionnaire using urinary markers of nitrogen, potassium and sodium,” Int. J. Epidemiol., vol. 30, no. 2, pp. 309–317, Apr. 2001, doi: 10.1093/ije/30.2.309.

[5] D. A. Schoeller, L. G. Bandini, and W. H. Dietz, “Inaccuracies in self-reported intake identified by comparison with the doubly labelled water method,” Can. J. Physiol. Pharmacol., vol. 68, no. 7, pp. 941–949, Jul. 1990, doi: 10.1139/y90-143.

[6] D. A. Schoeller, “Limitations in the assessment of dietary energy intake by self-report,” Metabolism, vol. 44, pp. 18–22, Feb. 1995, doi: 10.1016/0026-0495(95)90204-X.

[7] L. S. Muhlheim, D. B. Allison, S. Heshka, and S. B. Heymsfield, “Do unsuccessful dieters intentionally underreport food intake?,” Int. J. Eat. Disord., vol. 24, no. 3, pp. 259–266, 1998, doi: 10.1002/(SICI)1098-108X(199811)24:3<259::AID-EAT3>3.0.CO;2-L.

[8] J. Macdiarmid and J. Blundell, “Assessing dietary intake: Who, what and why of under-reporting,” Nutr. Res. Rev., vol. 11, no. 2, pp. 231–253, Dec. 1998, doi: 10.1079/NRR19980017.

[9] “Automatic Detection of Swallowing Events by Acoustical Means for Applications of Monitoring of Ingestive Behavior | IEEE Journals & Magazine | IEEE Xplore.” Accessed: Feb. 27, 2024. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/5272275

[10] P. Lopez-Meyer, O. Makeyev, S. Schuckers, E. L. Melanson, M. R. Neuman, and E. Sazonov, “Detection of Food Intake from Swallowing Sequences by Supervised and Unsupervised Methods,” Ann. Biomed. Eng., vol. 38, no. 8, pp. 2766–2774, Aug. 2010, doi: 10.1007/s10439-010-0019-1.

[11] S. Paßler and W.-J. Fischer, “Food Intake Activity Detection Using a Wearable Microphone System,” in 2011 Seventh International Conference on Intelligent Environments, Jul. 2011, pp. 298–301. doi: 10.1109/IE.2011.9.

[12] “Food Intake Monitoring: Automated Chew Event Detection in Chewing Sounds | IEEE Journals & Magazine | IEEE Xplore.” Accessed: Feb. 27, 2024. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/6556940

[13] E. Sazonov et al., “Non-invasive monitoring of chewing and swallowing for objective quantification of ingestive behavior,” Physiol. Meas., vol. 29, no. 5, p. 525, Apr. 2008, doi: 10.1088/0967-3334/29/5/001.

[14] Y. Dong, J. Scisco, M. Wilson, E. Muth, and A. Hoover, “Detecting Periods of Eating During Free-Living by Tracking Wrist Motion,” IEEE J. Biomed. Health Inform., vol. 18, no. 4, pp. 1253–1260, Jul. 2014, doi: 10.1109/JBHI.2013.2282471.

[15] M. Mirtchouk, C. Merck, and S. Kleinberg, “Automated estimation of food type and amount consumed from body-worn audio and motion sensors,” in Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, in UbiComp ’16. New York, NY, USA: Association for Computing Machinery, Sep. 2016, pp. 451–462. doi: 10.1145/2971648.2971677.

[16] C. K. Martin, S. Kaya, and B. K. Gunturk, “Quantification of food intake using food image analysis,” in 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Sep. 2009, pp. 6869–6872. doi: 10.1109/IEMBS.2009.5333123.

[17] F. Zhu, M. Bosch, N. Khanna, C. J. Boushey, and E. J. Delp, “Multiple Hypotheses Image Segmentation and Classification With Application to Dietary Assessment,” IEEE J. Biomed. Health Inform., vol. 19, no. 1, pp. 377–388, Jan. 2015, doi: 10.1109/JBHI.2014.2304925.

[18] P. Pouladzadeh, S. Shirmohammadi, and R. Al-Maghrabi, “Measuring Calorie and Nutrition From Food Image,” IEEE Trans. Instrum. Meas., vol. 63, no. 8, pp. 1947–1956, Aug. 2014, doi: 10.1109/TIM.2014.2303533.

[19] F. Kong and J. Tan, “DietCam: Automatic dietary assessment with mobile camera phones,” Pervasive Mob. Comput., vol. 8, no. 1, pp. 147–163, Feb. 2012, doi: 10.1016/j.pmcj.2011.07.003.

CHAPTER 2 – A SYSTEMATIC REVIEW OF TECHNOLOGY-DRIVEN METHODS FOR MEASUREMENT OF EATING BEHAVIOR

Abstract

Behavioral aspects of eating may be characterized by the dynamic process of food consumption, such as the microstructure of eating, including eating rate, eating trajectory, and other metrics. This article presents a systematic review of the use of sensors to track the dynamic process of food consumption. PRISMA guidelines were used to review the full texts of 153 scientific manuscripts. The contributions of this review article are twofold: i) A taxonomy of sensors for quantifying various aspects of eating behavior is established, classifying the types of sensors used (such as Acoustic, Motion, Strain, Distance, Physiological, Cameras, and Others) ii) The accuracy of measurement devices and methods is assessed. The review highlights the advantages and limitations of methods that detect and track different metrics of eating using a combination of sensor modalities and machine learning and deep learning algorithms. Furthermore, the review emphasizes the importance of testing the methods outside of restricted laboratory conditions. The review concludes with a discussion of challenges and future trends in measurement devices.

2.1 Introduction

Eating is a complex interaction of a variety of physiological, emotional, social, cultural, environmental, and economic factors that influence the timing of an eating episode, amount of food intake, food choice and/or selection, and how food is consumed.

Dietary intake refers to the types and amounts of food and drinks that a person consumes over a given time, such as a day or a week. It can be measured in terms of calories, macronutrients (such as carbohydrates, protein, and fat), micronutrients (such as vitamins and minerals), and other nutritional components. Eating behavior, on the other hand, refers to the way a person eats, including their habits, patterns, and attitudes toward food. Eating behavior can include factors such as the timing of meals, portion sizes, food choices, eating speed, and social or emotional factors that influence eating. While dietary intake and eating behavior are related, yet they remain distinct concepts. A person's dietary intake is influenced by their eating behavior, but their eating behavior is influenced by a wide range of factors beyond just the types and amounts of food they consume. Understanding both dietary intake and eating behavior is important for developing effective strategies to promote healthy eating habits and support overall health and well-being. This review focuses on eating behavior.

Eating behaviors (i.e., food choices and motives, feeding practices, etc.) have very important roles in the development of chronic diseases, including type 2 diabetes, heart disease, and obesity [1], [2], [3]. As such, active research involving the automatic detection of eating episodes, recognition of the foods being consumed, and measurement of the quantity and the manner of consumption in an eating episode is rapidly escalating to understand human eating behavior. One approach to understanding the complex process of eating behavior is to track the important micro-level temporal patterns of an eating episode such as biting, chewing,

swallowing, type of food items of the eating episode, duration, speed/rate of eating, consumed mass, consumed energy, the environment of eating episode, etc. In this review, we focus on finding the technology and measurement devices explored to identify these eating behavior measures.

Assessment of dietary intake in both young and adult populations historically relies on self-report [4], [5] and tools such as 24-h recalls, food records (food diaries), and food frequency questionnaires (FFQ) [6], [7]. Although these can help understand relationships between eating behavior and dietary intake, important eating behavior measures and processes are not calculated nor can they be explored with these self-report methods due to the lack of granularity of measuring the process of food consumption [8].

Rapid advancements in technology have provided the tools to develop accurate and objective systems for measuring dietary intake with fine granularity. These dietary intake devices may potentially be used for the measurement of eating behaviors. Wearable sensors include devices with sensors on the head or neck to detect chewing or swallowing , [9], [10], [11], [12], [13], [14], [15], wrist-based inertial sensors to detect hand-to-mouth gestures as a proxy for bites , [16], [17] , and others [18], [19]. Camera-based methods [20], [21], [22], [23] use food images, often captured before and after a meal, to recognize the food items consumed and estimate energy intake (EI). The availability of image-capturing devices has enabled the researcher to identify what people eat during an eating episode. Image capture can be categorized into two methods - active (manual capture, typically before and after the eating episode) and passive (wearable cameras that can capture images continuously or at pre-determined intervals). Then the captured images are retrieved and analyzed either by the nutritionist manually [23], [24] or by computer vision algorithms [25], [26]. Weight scale devices are used to instrument dining

locations to continuously weigh food consumed [27], [28], [29] and thus eating behaviors can only be captured at the location of instrumented arrangement [30].

The main objective of this review paper is to provide a systematic evaluation of technology-based methods available for different measures of eating behavior such as biting, chewing, swallowing, food items, consumed mass, consumed energy, eating rate/speed, and meal environment. Each measure of eating behavior (such as chewing) can be measured using different quantifiable metrics (such as the number of chews, chewing rate/frequency, etc.) Most existing review papers, such as [31], [32] focus on the detection of eating events and dietary information rather than eating behavior. For example, computer vision and wearable sensor-based methods used for food intake detection were reviewed [32], [33], [34], [35]. A few review papers focused on particular food consumption metrics such as chewing and swallowing [36], the energy of food intake [37], [38], and portion size of food items [39]. However, there is a research gap to investigate available technology-driven measurement devices to detect and track eating behavior of eating episodes in a single review. In this paper, we systematically review the technology-driven methods of eating behavior monitoring by searching five major databases following the guidelines of Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) [40]. The contributions of this review can be described as twofold: 1) Taxonomy of quantifiable food consumption metrics was identified followed by a comprehensive review of technology-driven measurement devices for food consumption metrics conducted, and 2) Accuracy of the measurement devices and applicability to free-living were assessed.

The paper is organized as follows. First, the methodology of the systematic review is presented in Section II along with research questions (RQs) identification, search strategy, and the result of the database search. The review findings along with taxonomy are discussed in

Section III. In this section, a detailed description of the selected articles is presented. Sections IV, V, and VI are discussion, challenges/future direction, and conclusion, respectively.

2.2 Review Methodology

The systematic review protocol was developed following the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) guidance [40]. The titles and abstracts of the publications retrieved through the database search were independently screened by the authors and then the full-text review of all relevant studies was carried out. The following processes were used in this methodology:

2.2.1 Research Questions Identification

To guide this systematic review, two research questions were identified.

RQ1: What are the quantifiable metrics that describe the dynamic process of eating?

RQ2: What state-of-the-art sensors and technologies are employed to measure these quantifiable metrics?

2.2.2 Databases

Exhaustive electronic searches for relevant literature were performed across five reputed databases: ACM digital library, IEEE Explore, SCOPUS, PUBMED, and Science Direct from inception through 30th January 2024.

2.2.3 Search Strategy

Given the research topic of finding the available devices and technologies for describing the quantifiable metrics to track the dynamic process of eating behavior, the most related keywords were first identified. The following keywords were considered: ‘chewing’, ‘chewing cycle’, ‘chewing rate’, ‘chewing frequency’, ‘biting’, ‘bite rate’, ‘bite size’, ‘bite frequency’, ‘swallowing’, ‘swallowing cycle’, ‘swallowing rate’, ‘swallowing frequency’, ‘eating rate’,

‘eating speed’, ‘meal duration’, ‘food items’, ‘eating amount’, ‘eating environment’, ‘food intake curves’.

Three query strings were formed to use for the search. The search query strings are as follows:

1. (“Chewing " OR “Biting" OR " Swallowing " OR “Eating Amount" OR “Eating Environment” OR “Eating Speed”) AND ("Sensor OR Device OR Technology")

2. (“Chewing Rate" OR “Bite Rate" OR " Swallowing Rate" OR “Eating Rate" OR “Meal Environment” OR “Food Items”) AND ("Sensor” OR “Device” OR “Technology")

3. (“Chewing Cycle" OR “Chewing Frequency” OR “Bite Frequency” OR “Swallowing Frequency "OR “Meal Duration “OR “mealtime”) AND ("Sensor” OR “Device” OR “Technology")

Table 2-1: Search Query Strings to obtain candidate articles.

Search Strings	Databases					
	ACM	IEEE	PubMed	Science Direct	Scopus	Total
(“Chewing " OR “Biting" OR " Swallowing " OR “Eating Amount" OR “Eating Environment” OR “Eating Speed”) AND ("Sensor OR Device OR Technology")	60	252	35	462	135	944
(“Chewing Rate" OR “Bite Rate" OR " Swallowing Rate" OR “Eating Rate" OR “Meal Environment” OR “Food Items”) AND ("Sensor” OR “Device” OR “Technology")	114	125	78	301	231	849
(“Chewing Cycle" OR “Chewing Frequency” OR “Bite Frequency” OR “Swallowing Frequency "OR “Meal Duration “OR “mealtime”) AND ("Sensor” OR “Device” OR “Technology")	64	219	73	155	77	588

2.2.4 Inclusion and Exclusion Criteria

The search results were restricted to the English language. References from the selected primary full-text articles were further analyzed for relevant publications. The selection was further narrowed by applying the eligibility criteria described in Table 2-2.

Table 2-2: Eligibility Criteria

Inclusion Criteria	Exclusion Criteria
Articles published since 1st January 2000.	Articles not written in English
Articles published after peer-reviewed	Articles published as posters, abstracts, book chapters, database descriptions, and review articles
Articles must address a set of keywords chew, chewing rate, bite, bite rate, swallow, swallow rate, eating rate, meal duration, no. of food items, selection of food items, consumed mass, consumed energy, cumulative food intake curves.	Studies conducted on animals.
Papers that describe methods for the quantifiable metric of food consumption.	Papers that describe the eating process but with no quantifiable metrics.

The manual bibliographic search identified articles through sources outside the mentioned databases such as reference lists of articles. Cited references or citing references of articles included through the database search were examined, and suitable ones were added to the

survey. Articles that fulfilled the inclusion criteria were considered in this review and those that fulfilled the exclusion criteria were filtered out.

2.2.5 Results

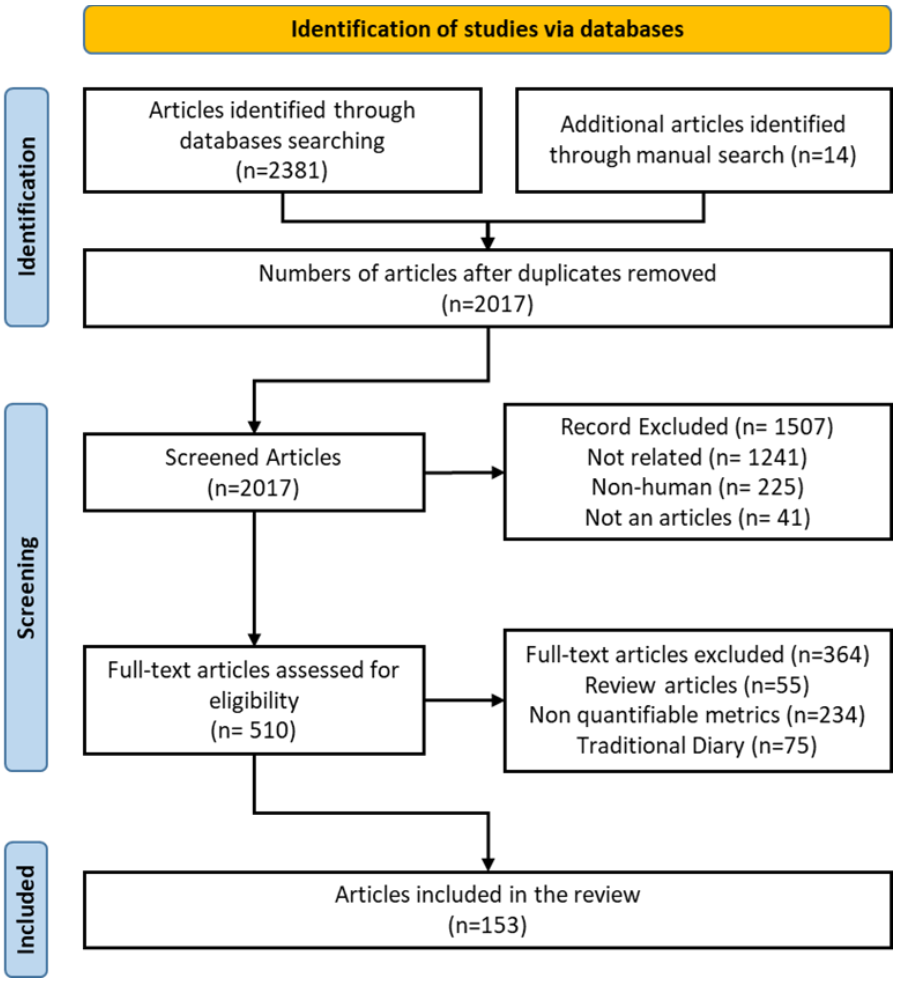


Figure 2-1: Flow chart outlining the study selection process of the systematic review.

The initial electronic database searches resulted in a total of 2381 publications. Table 2-1 shows the number of identified publications from each database for each search string. A manual bibliographic search also included 14 additional publications that qualified to be included. After removing the duplicate articles, a total of 2017 articles were set aside for screening based on abstracts and titles; of these 510 articles were selected for full-text review. Following the

exclusion criteria, 153 articles were included in this systematic review based on the PRISMA guideline process. The methodology and results of the review process are illustrated in Figure 2-1.

2.3 Review Findings

2.3.1 Taxonomy

First, the taxonomy of quantifiable measures that describe the dynamic process of eating behavior was identified to address RQ1 and RQ2. A classification scheme based on quantifiable metrics to describe different measures of eating behavior is proposed. The main measures of eating behavior found in this review were related to biting, chewing, swallowing, food items, eating time, the mass of food intake, and food intake environment. The main quantifiable metrics identified in this review were: metrics related to the process of biting, metrics related to the process of chewing, metrics related to the process of swallowing, metrics related to the food items being consumed, metrics related to the eating time, metrics related to the mass of food intake, metrics related to the energy of food intake, and, metrics related to the food intake environment. Each eating behavior measure can be represented using different metrics that can be recorded utilizing different devices, sensor modalities, or technology. A taxonomy that was established from the findings of this review is presented in Figure 2-2.

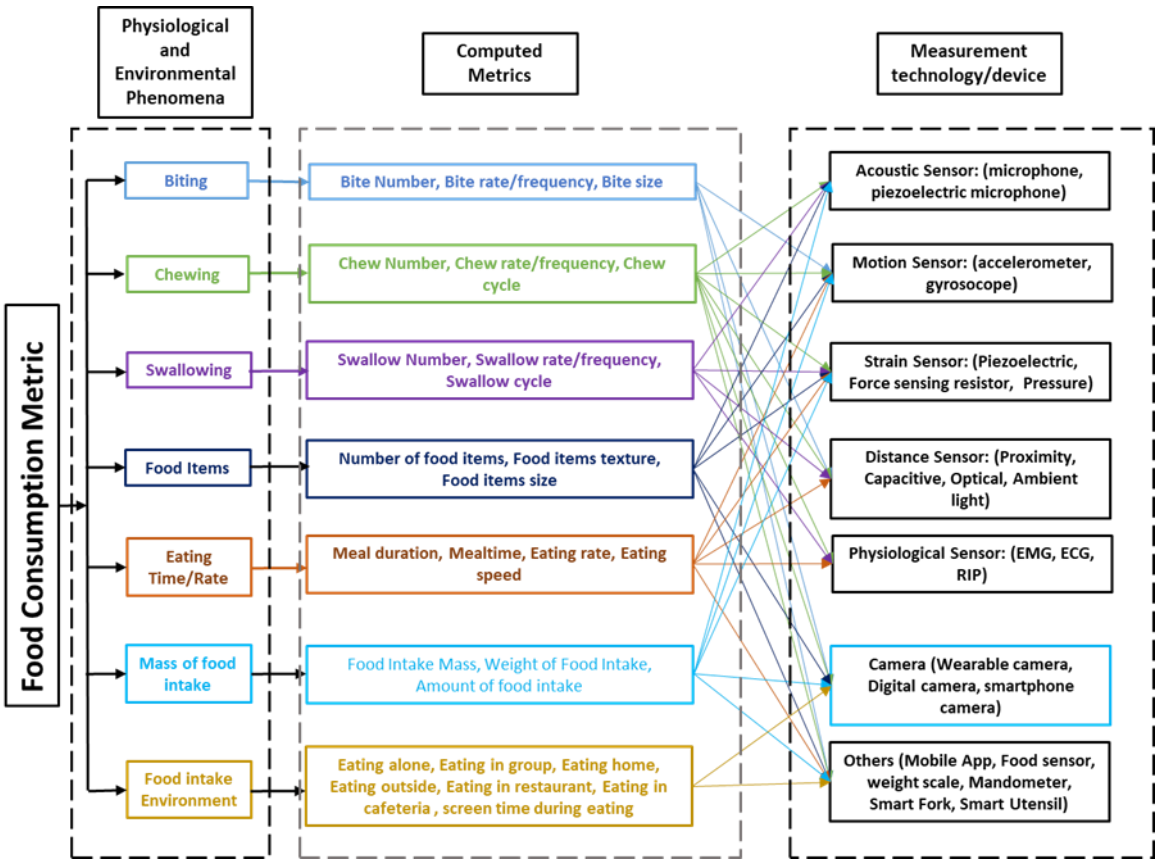


Figure 2-2: Taxonomy of quantifiable metrics and measurement technologies for food consumption.

2.3.2 Physiological/Environmental Phenomena and Computed Metrics:

We identified the physiological and environmental phenomena of food consumption and grouped them into eight major classes related to biting, chewing, swallowing, food items, eating time/rate, the mass of food intake, the energy of food intake, and the food intake environment. Figure 2-2 illustrates the taxonomy of food consumption metrics found in this review. Each eating behavior measure is described in detail below.

2.3.2.1 Metrics related to Biting

Biting is recognized as the very first stage of ingestion [41]. It involves two main steps: movement of the wrist and hand to transport food to the mouth, and movement of jaw muscles to take a bite. Researchers investigated several characteristics of biting phenomena of eating events, such as the number of bites [42], [43], [44], [45], bite rate or frequency [46], and bite size [47] to track the dynamic process of eating. The number of bites during eating refers to the count of individual processes of taking the liquid, semi-solid, or solid food by the person using a straw, glass, bottle, spoon, fork, or hand [45]. This metric is often studied in the context of eating behavior and portion control, as it can provide insights into how people consume their meals and the potential impact on their overall food intake[42], [48], [49]. Bite rate/frequency is considered a part of eating behavior because it can reflect the pace and rhythm of a person's eating. Research has shown that the speed at which a person eats, including their bite rate, can affect the amount of food they consume [46]. Bite size refers to the amount of food that an individual takes into their mouth with one bite typically reported in grams [48]. Research has shown that people tend to eat more when they take larger bite sizes, as they tend to consume more food. This can contribute to overeating and weight gain [47]. Feedback from these biting metrics such as the number of bites, bite rate or frequency, and bite size affects overall food consumption [42], [46], [47], [48], [49].

2.3.2.2 Metrics related to Chewing

Chewing or mastication is responsible for the physical degradation of food. The mastication process produces a bolus, which is later swallowed down into the esophagus and the stomach. Researchers investigated several characteristics of chewing phenomena of eating events such as the chewing cycles or bouts [50], [51], [52], the number of chews[45], [53], and chewing

rate or frequency[54], [55] to track eating behavior. A chewing cycle or bout is a repetitive sequence of jaw movements and muscle contractions involved in the process of breaking down food during mastication, and it is related to eating behavior as it influences the efficiency of food processing, eating pace, and overall meal satisfaction [50]. The number of chews refers to the count of masticatory movements taken while consuming food [45]. Chewing rate or frequency is the speed at which an individual performs these masticatory movements, typically measured in chews per minute[54]. These parameters, including the number of chews, chewing cycles or bouts, and chewing rate or frequency are closely related to eating behavior. Variations in these aspects can affect satiety, portion control, nutrient absorption, and overall enjoyment of meals, ultimately influencing an individual's dietary habits, digestion, and potential impact on health outcomes [56], [57].

2.3.2.3 Metrics related to Swallowing

Swallowing is an important part of eating behavior because it is the mechanism by which food is moved from the mouth to the stomach, allowing the body to digest and absorb the nutrients it needs to function. The speed and efficiency with which food is swallowed can impact the overall digestive process and the body’s ability to absorb nutrients[58]. Therefore, understanding the mechanics of swallowing and its role in the digestive process is an important aspect of eating behavior. Researchers investigated several characteristics of swallowing phenomena of eating events such as the swallowing cycle, the number of swallows, and swallowing rate or frequency to describe the dynamic process of eating [58], [59], [60], [61], [62]. The swallowing cycle refers to the coordinated sequence of muscle movements and reflexes involved in the act of swallowing food or liquid, starting from the mouth and ending in the stomach [58]. Number of swallows refers to the count of times a person swallows while

consuming food or beverages [62]. Swallowing rate or swallowing frequency represents the speed at which these swallowing actions occur, typically measured in swallows per minute [61]. These metrics provide insights into the pace of food consumption, potential overeating, and the efficiency of the swallowing process, which can be valuable for individuals seeking to manage their dietary habits, such as those looking to control portion sizes or address specific swallowing issues during meals.

2.3.2.4 Metrics related to Food Items

The type and variety of food items can be used as a measure of eating behavior and provide valuable information about an individual's dietary habits and patterns. Researchers investigated several metrics of consumed food items such as the number of food items, consumed food item type, food item textures, and size of the food item to investigate the dynamic process of eating [63], [64], [65], [66], [67]. The food items individual consumes can reflect their eating habits, preferences, and attitudes towards food. For example, an individual who frequently consumes high-calorie, high-fat foods may have a more indulgent eating behavior, while someone who regularly consumes lower-calorie, nutrient-dense foods may have a more health-conscious eating behavior [68].

2.3.2.5 Metrics related to Eating duration/Rate

Eating duration and rate are all aspects of eating behavior that characterize the dynamic process of eating [69], [70], [71], [72], [73]. Meal duration or mealtime refers to the duration of a meal or snack, from the first bite to the last [74]. Prolonged mealtime may be associated with greater enjoyment of the meal but may also contribute to overeating and poor dietary habits. In contrast, shorter mealtime may be associated with greater feelings of fullness and better portion control [75]. Eating rate/speed refers to how quickly an individual consumes a meal or snack

[76]. Eating too quickly can be associated with overeating, as the brain may not have time to register feelings of fullness before more food is consumed. In contrast, eating too slowly may be associated with disordered eating behaviors, such as anorexia nervosa [77].

2.3.2.6 Metrics related to the mass of food intake

The mass of the consumed food items is one important variable that is being explored by the researchers to describe the dynamic process of eating. Metrics related to the mass of food intake including the amount and weight of the consumed food item were found to be used to track the dynamic process of eating [67], [78], [79], [80], [81] in this review. The amount and weight of the consumed food items refer to the total weight or quantity of food and beverages consumed during a meal or eating episode typically measured in grams. The mass of food intake can be measured by continuous monitoring of weight or amount of food items [82], by using the body-worn sensor [83], or from images using computer vision or deep learning algorithm [67]. Monitoring these measures helps assess portion control, nutrient intake, and the overall balance between caloric and non-caloric components in one's diet, offering valuable information for understanding and managing eating behaviors[83].

2.3.2.7 Metrics related to the environment of food intake

The environment in which an individual consumes their food is an important aspect of their eating behavior. The environment can refer to a variety of factors, such as the physical surroundings, social setting, and cultural context in which the meal or snack is consumed [84]. The physical environment can include factors such as the lighting, temperature, and noise level of the dining area, as well as the type and quantity of food that is available [85]. The social setting is another important aspect of the environment of food intake. Eating in the company of others can influence an individual's food choices, portion sizes, and eating speed [86]. Overall,

the environment of food intake is an important aspect of eating behavior that can impact an individual's food choices, portion sizes, and eating speed. In this review, we identified metrics of environment such as eating alone or in a group, eating at home or a restaurant, and screen time during eating [86], [87], [88].

2.3.3 Measurement technology/device

Rapid advancements in technology have provided the tools to develop accurate and objective systems for measuring physiological and environmental phenomena related to eating behaviors. In this review, measurement devices are grouped into seven major categories based on the integrated sensor system. The seven major groups are i) Acoustic sensors (microphones) ii) Motion sensors (e.g., accelerometers) iii) Strain sensors (e.g. piezoelectric, force sensing resistor, pressure) iv) Distance sensors (proximity, capacitive, optical) v) Physiological sensors (EMG, ECG, RIP) vi) Cameras vii) Others (mobile application, food sensor, weight scale, mandometer, smart fork, smart utensil). Measurement devices for each quantifiable food consumption metric found in this review are described below. Most of these measurement devices can be used for assessing both dietary intake and eating behavior. For dietary intake monitoring, these devices are typically designed to measure the nutritional content of the food, such as calories, macronutrients (such as carbohydrates, fats, and proteins), and micronutrients (such as vitamins and minerals). In this review, we focused on the use of measurement devices for eating behavior, such as bite counters, chewing sensors, swallowing microphones, etc., to measure the physical and behavioral aspects of eating, such as bite rate, chewing rate, and swallowing rate, etc.

2.3.3.1 Biting

2.3.3.1.1 Motion Sensors

Motion sensors can be used to monitor eating behavior by tracking the movement of a person's head, mouth, jaw, and wrist. In [89], authors proposed a complete framework for automatically measuring eating behavior from in-the-wild collected inertial data (acceleration and orientation velocity) using a smartwatch. Their algorithm achieved a 0.92 F1 score in the detection of the bites. An orientation sensor (InertiaCube3) was placed on the wrist of the user and analyzed the rolling motion of the wrist to detect the pattern related to biting behavior [43]. To validate the device authors conducted a study of ten participants who ate a single meal. The device achieved 91% sensitivity in counting the number of bites in a meal. In [90] authors described a study to test the accuracy of using an accelerometer and gyroscope to track wrist motion during eating to detect and count bites. The authors conducted a study with 271 participants who ate a single meal while wearing the device consuming 374 different food and beverage items in total. Overall, the method for automatically detecting and counting the bites had a sensitivity of 75% with a positive predictive value of 89%. The same device is used in [91] to describe a novel approach of segmenting and classifying eating gestures from wrist motion using a deep-learning neural network. Their proposed neural network classifier recognized an average of 79.7% of ‘bite’ and 84.7% of ‘drink’ gestures correctly per meal and reported the bite rate or bite frequency per minute. Overall, 77.7% of all gestures were recognized correctly on average per meal. In [49], the authors described a wrist-worn bite counter, which uses a gyroscope to detect a pattern of motion, distinctive of hand-to-mouth gestures used to ingest foods and beverages. During eating, it displays the bite count for the current eating activity in real-time. They also provided a graphical representation of temporal patterns of biting, showing

variations in the number of cumulative bites, and bite rate/frequency per minute throughout the meal. Data from 94 volunteers shows that participants took fewer bites ($P<0.001$) with feedback but did not consume less. The progressive design of the gesture recognition module of an automated food journaling system – Annapurna, is described in [92]. Annapurna runs on a smartwatch and utilizes data from the accelerometer and gyroscope, to first identify eating gestures (hand to mouth), and then capture food images, which are presented to the user in the form of a food journal. Annapurna achieved a false positive rate of 6.5% and a false negative rate of 3.3% in detecting the number of bites. A real-time food intake monitoring system for mobile devices is introduced in [93]. The proposed system gets acceleration data from the sensor placed on the wrist of the user during a meal. Based on biting detection, real-time feedback regarding eating trends such as the total number of bites, and bite rate/frequency in minutes provided to the user. A bite-counting device consisting wrist-mounted sensor with an accelerometer and gyroscopic motion sensing capabilities is presented in [94]. The authors conducted a study with three participants who wore the device for a full day in free-living conditions. The device achieved an accuracy of 91.8% in counting the number of bites for a day. In [95] authors used the inertial signals (acceleration and orientation velocity) from an off-the-shelf smartwatch to track the hand-to-mouth gesture and detect food intake events (i.e., bite). The device achieved an F1 detection score of 0.91 in a leave-one-subject-out experiment containing 12 subjects in detecting biting events. A novel method using an SVM classifier to detect bite instances from 3-axis orientation data acquired by a smartwatch is described in [13]. They experimented with 8 participants and their method achieved an accuracy of over 90% in counting the number of bites. They also reported the bite rate/frequency for each participant to show the variation among different participants' biting behavior. In [44] authors described the validation of a bite counter

device in laboratory laboratory-controlled environment. It contains a tri-axial accelerometer, that detects an upward, arcing motion from the table to the mouth, known as a wrist roll. The bite counter was tested on 15 adults where it achieved an accuracy of 81.2% in counting the number of bites.

2.3.3.1.2 Distance Sensors

Distance sensors can be used for eating behavior monitoring from biting phenomena detection by measuring the distance between the user's hand and mouth, necklace and jaw, etc. Several types of distance sensors can be used for this purpose, including proximity sensors, infrared sensors, and laser distance sensors. Each type of sensor uses a different technology to measure distance, but the general principle is the same - the sensor sends out a signal that reflects off an object (in this case, mouth, jaw, temporalis muscle), and the sensor measures the time it takes for the signal to return to the sensor. The detection of biting, using a wearable computer named, eButton is presented in [96]. The authors proposed a magnetic-field-based approach to detect the hand. In this approach, a miniature magnetic maker is worn on one finger of the hand. A magnetometer is embedded inside of the eButton which is worn in front of the chest. If the hand is engaged in eating, a sequence of signals is produced. By analyzing the patterns of the signals, the event of biting can be detected. They found that the proximity of the hand signal patterns corresponding to eating and other daily activities are different, which can be used to reduce the false detection rate. In [97], to evaluate the feasibility of using a neck-worn device for eating and chewing detection, the authors designed and developed a multi-sensor necklace prototype. The on-necklace ambient light sensor’s reading drops when the individual’s hand moves towards the mouth during a feeding gesture, thus allowing the detection of biting instances. The proximity sensor on the necklace is directed toward the chin and allows

monitoring of the variation in signal while an individual is biting. Data was collected from 10 volunteers (5 obese, 5 normal) who wore the device for two days in free-living conditions. The device achieved an accuracy of 82.2% in counting the number of bites. They also reported the average bite rate/frequency per meal to show the difference in biting between obese and normal volunteers.

2.3.3.1.3 Cameras

A method to detect biting using a depth camera of Kinect Xbox One by tracking the motion of a person's gesture and posture during food intake and reporting the biting instances described in [98]. The proposed system achieves an accuracy of 96.2% in 10 meals from 10 volunteers for counting the number of bites. In [45] authors proposed a method, to count the number of bites from videos of eating episodes, to provide an alternative solution to cumbersome manual annotation for food intake studies. The authors reported an accuracy of 85.4%±6.2% in counting the number of bites for 84 meal videos of 28 subjects. In [99] authors described a novel "Rapid Automatic Bite Detection" (RABiD) algorithm, that extracts and processes skeletal features from videos, trained in a video meal dataset (59 individuals; 85 meals; three different foods), to automatically measure meal duration, bite rate/frequency per minute and the number of bites. The proposed algorithm on meal videos from 30 volunteers achieved an F1 score of 0.948. In [100] authors propose a vision-based approach to first capture the food-sharing scenario with a 360-degree camera and then use a neural network to infer different eating states. The authors reported bite error percentage across 14 recorded panoramic videos is 26.2%. A wearable fish eye camera was used [42], to record sixteen participants' free-living eating behaviors, detect biting events, and report the number of bites, bite rate/frequency per minute as eating speed. A recent new video dataset that was collected by recording the meals of 264

participants in a real restaurant was described in [101]. Participants consumed 374 different types of foods and beverages. The camera was installed on the ceiling of the restaurant to minimize the effect of an experimental observer. To explore the feasibility of detecting intake gestures using their dataset and establish baseline performance, they implemented three state-of-the-art models, namely CNN-LSTM, SlowFast, and X3D-L. Results indicate that the best F1 scores were 0.899 and 0.778 on bite and drink gesture detection respectively.

2.3.3.1.4 Other sensor

A new method to monitor eating behavior is explored [42], by facilitating the electrical conductivity of food material to detect and count the number of bites. In [102] authors described an augmented fork for biting detection during a 15-week study. They provided real-time feedback based on the bite rate/frequency per minute, bite-size in grams. They reported that after a 4-week intervention period, people who received feedback on related biting-related metrics managed to achieve longer spacing between bites, leading to fewer fast bites, longer pauses between bites, fewer bites per meal, and significant weight loss. A tool in the form of a plate system, that automatically measures the consumed food during the meal, is proposed [103]. Weight sensors are integrated into the base, allowing for easy removal and cleaning of the plate. Data from 24 aging adults (ages 52–95), eating a single meal with chopsticks were used to train and evaluate the model. Out of a total of 836 true annotated bites, the algorithm detected 602 with a precision and recall of 0.78 and 0.76, respectively.

2.3.3.2 Chewing

2.3.3.2.1 Acoustic Sensors

The feasibility of detecting chewing activities based on audio and inertial sensor data obtained from an ear-worn device, eSense was investigated in [104]. The results of this study

including 5 volunteers indicate that fusing audio and inertial sensor modalities improves the accuracy to 97% in counting the number of chews. A wearable sensor platform was presented in [105], that autonomously provides detailed information regarding a subject's dietary habits. The sensor consists of a microphone and a camera and is worn discretely on the ear. Sound features are extracted in real-time and if a chewing activity is classified, the camera captures a video sequence for further analysis. They reported 80% accuracy in counting the number of chews using data from six volunteers consuming lunch. They also reported that whilst the auditory data perhaps gives a better estimation of the subject's chewing rate/frequency (per minute), the vision-based estimation is a better indicator of the actual rate of removal, and thus consumption, of food from the plate. In [106], authors presented an embedded hardware prototype to collect food intake sensor data using a high-fidelity throat microphone worn on the neck, to precisely record acoustic signals during eating in a noninvasive manner. The authors used hidden Markov models to detect chewing. Their method achieved an accuracy of 86.6% in counting the number of chews in the experiment consisting of 12 volunteers. A wearable device to detect chewing using a two-channel condenser microphone is proposed in [107]. The authors conducted a study with 18 subjects chewing gum, crackers (Ritz), and cabbage (shredded) under laboratory conditions. The methodology achieved an F score of 0.80 in detecting chewing events from the sound. In [108], authors explored detecting chewing cycles in the presence of restaurant background noise. The noisy signal was formed by the instantaneous addition of a clean throat microphone recording during eating and restaurant noise recorded with the same throat microphone. An F1 score of 71.4% was achieved among 12 subjects for detecting chewing cycles in a very low signal-to-noise ratio of -10 dB. In [109], authors used a custom earbud with two microphones along with an android smartwatch (9 axis IMU data) on each wrist, to introduce a novel training

method named Stochastic Variational Deep Kernel Learning (SVDKL), that classifies micro-events (e.g. chews) during macro events (e.g. meals) in time series data, and allows learning from data labeled at both levels of granularity. Their approach achieved 84% recall and 67% precision in detecting chewing among 6 volunteers. A laryngeal throat microphone is used in [110] to develop a methodology for the detection of chewing, from the recorded food intake sounds during a meal. The best accuracy results are attained as 0.712 and 0.783 for the subject-independent and subject-dependent settings, respectively among 8 subjects. A prototype of a wireless wearable in-ear microphone was developed in [111] to sense the chewing sound utilizing a common Bluetooth headset. The authors proposed a robust chewing number counting algorithm that consists of two recognition stages: “chew-like” signal detection and chewing verification stage. As a result, an average chewing number counting error rate of 1.93% was achieved from 1 participant. In [112], the authors presented an approach for chewing detection from chewing sounds during the process of eating, which was recorded by microphones applied to the outer ear canal of the user. The authors presented and evaluated eight different algorithms for automated chewing detection on two datasets. Their datasets contained 68 094 chews in around 18 h recording data. Precision and recall of over 80% were achieved by most of the algorithms. A reduction of the number of false detections by 28% on average was achieved using the noise reduction algorithm. They reported that their proposed algorithm can calculate chewing frequency in laboratory settings. The use of off-the-shelf Bluetooth headsets to unobtrusively monitor and detect users' eating episodes by analyzing the chewing sound, was proposed in [113]. The experimental results show that a Support Vector Machine (SVM) could achieve 94-95% classification accuracy in lab settings, though the detection performance quickly degraded to 65-76% for in-the-field testing. The authors then propose a novel deep learning-based

classification technique that drastically improved chewing cycle detection accuracy to 77-94% despite the existence of ambient noise. A skin contact microphone behind the wearer's ear, to capture the sound of a person chewing was investigated [114]. The device achieved an F1 score of 84.5% in detecting chewing in data collected from 10 children. In [115], authors proposed a system, composed of a cheap and small bone conduction microphone to collect intra-body sound signals and a smartphone that can process it and provide feedback based on chewing rate/frequency. Using medium Gaussian SVM (RBF kernel) authors achieved an accuracy 97.6% accuracy in counting the number of chews. In [116], the authors present a study on eating recognition, using audio (earbud microphone) sensors for 12 participants (6 from the lab study, 6 new to test generalizability). To automatically recognize a person's eating gestures (e.g., mastication) in real-time in a natural meal environment, authors in [117] added other sounds along with chewing sounds recorded from a bone conduction microphone. Using medium Gaussian SVM authors obtained 97.1% accuracy in chewing detection. Recently, the feasibility of using an off-the-shelf commercial device, named Razer Anzu, a smart-glass, equipped with stereo speakers and microphones to communicate with smartphones was explored to detect chewing [118]. Using a support vector machine and data from 5 participants they achieved a F score of 0.96 in detecting chewing activity.

2.3.3.2.2 Motion Sensors

A novel methodology, IDEA [119] was proposed, that performs accurate eating action identification by detecting chewing from an accelerometer. The proposed system achieved an accuracy of 92.1% in counting the number of chews among 36 volunteers. The device provides feedback to the user in every 2 min based on the computed chewing rate/frequency as eating speed. A novel chewing activity recognition system named CARE, which uses a single-axis

accelerometer attached to the temporalis, to detect the muscle bulge and recognize the user's chewing activity was proposed in [120]. CARE is also able to calculate the chewing rate/ frequency by recognizing the periodic pattern of acceleration data. To efficiently detect chewing, using jaw tracking, a prototype was developed [121] using a new ultra-miniaturized Inertial Measurement Unit (IMU) named WB-3. The WB-3 IMU was composed of a 3-axis gyroscope, 3-axis accelerometer, and 3-axis magnetometer, which can not only measure the acceleration and angular speed of jaw movement, but also measure mouth opening angle. The proposed system achieved an accuracy of 83.2% in counting the number of chews after detecting chewing cycles among 15 volunteers. The proposed system provides information about chewing activity such as the number of chewing cycles, number of chews, and chewing rate/frequency per minute.

2.3.3.2.3 Strain Sensors

A simple piezoelectric strain gauge sensor [122] was used, to capture the movement of the lower jaw to detect periods of food intake based on non-invasive monitoring of chewing. Data were collected from 20 volunteers during periods of quiet sitting, talking, and food consumption. Twenty-fold cross-validation demonstrated per-epoch classification accuracy of 80.98%. To compare the performance of the Support Vector Machine (SVM) and the Artificial Neural Network (ANN) for chewing cycle detection in [123], a combination of time domain (TD) and frequency domain (FD) features, extracted from signals captured using the same piezoelectric strain sensor, were used to train both types of classifiers. Data were collected from 12 subjects in free-living for 24 hours under unrestricted conditions. ANN achieved an average accuracy of 86.86 ± 6.5% whereas SVM (with the linear kernel) achieved an average classification accuracy of 81.93 ± 9.22%. The same piezoelectric strain sensor to record jaw movement during chewing was used in [53] to estimate the number of chews using linear

regression. The regression model achieved a mean absolute error of 9.66% in counting the number of chews among 30 different volunteers. A comparison between an off-the-shelf piezoelectric strain sensor and a plotter-drawn strain sensor for quantifying the number of chews for several food items was described in [124]. Data collected from 5 subjects shows that the piezoelectric strain sensors and plotter-drawn strain sensors were able to achieve absolute mean error rates of 8.09 ± 7.16% and 8.26 ± 7.51% respectively for estimating the number of chews. A fully automatic approach to measure chew count and chewing rate/frequency from the chewing cycle using a piezoelectric strain sensor was described in [54]. The sensor signal was divided into 5-s non-overlapping epochs. Leave-one-out cross-validation was used to train an artificial neural network (ANN) to classify epochs as “food intake” or “no intake” with an average F1 score of 91.09%. Chews were counted in epochs classified as food intake with a mean absolute error of 15.01% ± 11.06%. To detect chewing cycles even when the user is physically active and/or talking, a novel wearable device that consists of a piezoelectric strain sensor placed on the temporalis muscle, an accelerometer, and a data acquisition module connected to the temple of eyeglasses was introduced in [125]. Data from 10 participants was collected while they performed activities including quiet sitting, talking, eating while sitting, eating while walking, and walking. The device achieved an F score of 99.85% in detecting chewing. In [126], authors developed a novel wearable sensor system (Automatic Ingestion Monitor, AIM) for objective monitoring of ingestive behavior in free living. The device and the methodology were validated with data collected from 12 subjects wearing AIM for 24 hours. Results showed that the system was able to count the number of chews in a meal with an average accuracy of 92.1%. In [127], the same device was used to explore the use of several ensembles of weak classifiers to classify signals to detect food intake based on chewing. A novel algorithm for

the detection and characterization of chewing cycles from a piezoelectric strain sensor placed on the temporalis muscle was presented in [52]. The proposed data acquisition device was incorporated into the temple of eyeglasses. The system was tested by 10 participants in two-part experiments, one under controlled laboratory conditions and the other in unrestricted free living. A multivariate regression model was used to estimate chew counts from segments classified as chewing with an average mean absolute error of 3.83% on the participant level. In [128], the authors introduced Slowee, which consists of a necklace-shaped piezo sensor worn below the neck to detect chewing. The system achieved an average accuracy of 0.938 in detecting chewing.

2.3.3.2.4 Distance sensors

In [129], the authors presented a new wearable computing approach to detect chewing and report chewing rate/frequency per minute using proximity sensing. The approach centers on a discreet and lightweight instrumented necklace that captures head and jawbone movements without direct contact with the skin. An evaluation of the system with 32 participants comprised of three phases resulted in a chewing sequence detected with 95.2% precision and 81.9% recall in controlled studies and 78.2% precision and 72.5% recall in the free-living study. The proximity sensor on the necklace [97] is directed toward the chin and allows monitoring of the variation in signal while an individual is chewing. The device achieved an F1 score of 77.1% for chewing sequence detection in free-living conditions among 10 volunteers. The Acoustic Doppler sonar (ADS) chewing detection method was proposed [130], where a 40 kHz ultrasonic beam was focused on the lower jaw and neck, and movements of the chin and neck cause Doppler frequency shifts. Hence, it was possible to detect chewing and swallowing events using Doppler frequency shifts in the received ultrasound signals. The experimental results consisting of 10 subjects showed that the proposed ADS-based method yielded promising results with a

maximum accuracy of 91.4% for counting the number of chews. A small ear-hung wearable device for monitoring chewing rate/frequency using an infrared distance sensor and accelerometer was developed [55]. The device, worn on the ear pinna, allows the counting of the number of chewing cycles, and data were collected on a smartphone via Bluetooth. They achieved an accuracy of 88% in counting the number of chews among 22 participants. In [131], the authors proposed an approach, based on chewing detection, using a proximity sensor to capture the temporalis muscle movement during chewing. A wearable device of eyeglass is used with the proximity sensor of VCNL4040 in a form click board manufactured by Mikro-Elektronika (MikroE) was mounted at the right side of the eyeglass temple using a 3D printed housing. Data from six participants shows that the device achieved the highest accuracy of 97.6% in the classification stage and provided a chew count error rate of 2.69%. A recent methodology of chewing detection using a proximity sensor housed in the eyeglass, capturing the temporalis muscle movement during food intake was described [132]. Their proposed chewing detection classifies the chewing activity using a medium Gaussian support vector machine (SVM). They also proposed a novel chew count estimation based on particle swarm optimization (PSO). Their validation dataset includes 20 participants and eight different food types. Their algorithm achieved 96.4% accuracy in detecting chewing activity and 4.26% mean absolute error of chewing count estimation.

2.3.3.2.5 Physiological sensors

In [133], EMG was used to evaluate chewing behavior. They used bipolar surface electrodes positioned on the right and left masseter to collect EMG signals during teeth clenching. They achieved an accuracy of 86.4% in counting the number of chews among 56 individuals. To compare chewing features such as chewing rate/frequency between normal range

BMI individuals and overweight or obese young adults, an EMG sensor was used [134]. The experimental results showed that there are no relevant differences in the measured chewing rate/frequency between high (14 subjects) and normal BMI individuals (14 subjects). The sensor system described in [128], includes a headphone-shaped EMG(Electromyography) sensor attached to both sides of the chin to detect chewing. The system achieved an average accuracy of 0.938 in detecting the number of chewing cycles. In [135], the authors presented smart eyeglasses with EMG sensors for dietary monitoring, in particular, to sense food chewing and report chewing rate/frequency. The eyeglasses were used in a data recording study with eight participants eating different foods. The same smart eyeglasses were evaluated for the free-living condition [136]. The device achieved precision and recall of over 77% for the free living study including 10 volunteers in chewing detection. A combination of an in-ear microphone with a photoplethysmography (PPG) sensor placed in the ear concha, in a new high accuracy and low sampling rate prototype chewing detection system to report chewing rate/frequency was proposed [137]. The proposed method achieved accuracy up to 0.938 in counting the number of chews in the experiment including 14 volunteers. A chewing detection method that could assist self-report methods using electromyography sensors unobtrusively placed behind the right ear is tested [138]. Results from 15 volunteers' data showed that chewing could be detected with high accuracy (sensitivity, specificity >90%). In [139] chewing, detection was performed by measuring the myoelectric potential of the masticatory muscle using a Myoware myoelectric sensor to develop a chewing amount control system that adds visual information and auditory information at meals. They provided feedback about the number of chews and chewing rate/frequency throughout the meal. Data from 8 participants were analyzed and reported that this system is effective in promoting satiety.

2.3.3.2.6 Cameras

To achieve a detailed description of the pattern of chewing (number of chews, chewing rate/frequency) throughout a meal, a video recording of the maxillary–mandibular region of women eating food from a plate was analyzed using computer vision and deep learning algorithms[56]. They experimented on 11 women eating a meal. Their method achieved an accuracy of 73.3% in counting the number of chews. If a chewing sequence is detected by the signal collected from a microphone [105], the camera captures a video sequence for further analysis. The bite detection algorithm from meal videos explained in [45] is capable of counting chews from videos of eating using an optical flow algorithm within detected bites. The chew detection algorithm from meal videos achieved an accuracy of 88.9%. In [140], the authors used visual quasi-periodicity in chewing to detect chewing events and report chewing rate/frequency from videos using support vector machines. Their chew counting algorithm achieved an accuracy of 93% among 37 meal videos collected from 37 participants. In [141], the authors aim to design algorithms for determining the chew count from video recordings and providing food intake curves such as variation of chewing rate/frequency, and cumulative chew count throughout the meal. A novel algorithm based on image and signal processing techniques has been developed to continuously capture the area of interest from the video clips, determine facial landmarks, generate the chewing signal, and process the signal with two methods: low pass filter, and discrete wavelet decomposition. Peak detection was used to determine the chew count from the output of the processed chewing signal. The system was tested using recordings from 100 subjects at three different chewing speeds (i.e., slow, normal, and fast) without any constraints on gender, skin color, facial hair, or ambiance. The low pass filter algorithm achieved the best mean absolute percentage error of 6.48%, 7.76%, and 8.38% for the slow, normal, and fast

chewing speeds, respectively. A new device consisting regular dining tray equipped with a video camera and three separate built-in weighing stations, to assess food intake and eating behavior was developed [142]. The video camera positioned to the face records eating behavior characteristics (chews, bites), which are analyzed using artificial intelligence (AI) based automatic facial expression software.

2.3.3.2.7 Other sensors

To compare chewing strength variation using wearable sensors, in [143] authors used a pressure sensor, flexible bend sensor, piezoelectric strain sensor, and EMG to measure differences in chewing strength while eating a carrot, apple, and banana.

2.3.3.3 Swallowing

2.3.3.3.1 Acoustic sensors

A methodology of studying ingestive behavior by non-invasive monitoring of swallowing using a bone-conduction microphone was presented in [144]. In [62], the authors proposed and compared two methods of acoustical swallowing detection from sounds collected in [144], contaminated by motion artifacts, speech, and external noise. Methods based on the mel-scale fourier spectrum, wavelet packets, and support vector machines were studied considering the effects of epoch size, level of decomposition, and lagging on classification accuracy. The methodology was tested on a large dataset (64.5 h with a total of 9966 swallows) collected from 20 human subjects with various degrees of adiposity. 84.7% average weighted accuracy in the detection of swallowing events was achieved. A throat and an ambient microphone were used [145] to detect the swallowing cycle and report the number of swallows. Data was collected from 7 healthy volunteers who did not present any medical condition that would interfere with normal swallowing. The authors achieved an accuracy of 68.2% in counting the number of swallows. In

[146], authors used a throat microphone to develop a prototype of automated ingestion detection through swallow detection. The authors collected data from 7 volunteers where the swallows in the dataset consist of ad-libitum consumption of a peanut butter and jelly sandwich portion, ad-libitum drinking of a carbonated beverage, consumption of an apple and orange, and drinking of water. They achieved 90% classification accuracy in swallow detection. Swalloscope, a portable wearable system that can be used to detect swallowing in real-time using a neck-worn system with a microphone and provide feedback based on the number of swallows, and swallow rate/frequency was introduced in [147]. Their developed real-time algorithm for the detection of dry and water swallows was based on a template-matching approach. They analyzed the wavelet transformation-based spectral characteristics and the temporal characteristics of simultaneous synchronized video fluoroscopic swallow study (VFSS) and swallowing sound recordings of 25% barium mixed 3-ml water swallows of 70 subjects and the dry or saliva swallowing sound of 15 healthy subjects to establish the parameters of the template. With this algorithm, they achieved an overall accuracy of 79.3% (standard error: 4.2%) for the 92 water swallows. In [106] authors present an embedded hardware prototype to collect food intake sensor data using a high-fidelity throat microphone worn on the neck to precisely record acoustic signals during eating in a noninvasive manner. Authors used hidden Markov models to detect the swallow. Experimental data were collected from 12 participants who wore the device and ate 7 types of food items. They achieved an accuracy of 86.6% in counting the number of swallows. Spectrograms extracted from the sound signal recorded by a laryngeal throat microphone were used in the convolutional neural network (CNN) to detect swallow cycles in [110]. The best accuracy results in the number of swallow counts are attained as 0.712 and 0.783 for the subject-independent and subject-dependent settings, respectively among 8 subjects. To create a natural meal environment, authors

in [117] added other sounds along with swallowing sounds recorded from a bone conduction microphone. Using medium Gaussian SVM authors obtained 97.1% accuracy in swallow detection. In [115], the authors proposed a system that may support both consciousness improvement feedback in real-time and accurate quantified monitoring of swallowing. It is composed of a cheap and small bone conduction microphone to collect intra-body sound signals and a smartphone that can process it and provide feedback based on swallow rate/frequency. Data were collected among six volunteers who wore the device for the whole day. They reported an F1 score of 90.2% for the detection of swallow. They provided feedback based on the swallowing rate/frequency per minute through smartphone after swallowing cycle detection in real-time.

2.3.3.3.2 Strain sensors

A novel wearable system in the form of a necklace, which aggregates data from an embedded piezoelectric sensor capable of detecting skin motion in the lower trachea during ingestion was proposed [148] to detect swallow. The same necklace was used [149], to provide visual feedback to the user based on swallow rate/frequency per minute. In [60] authors presented a novel integrated wearable necklace that comprises two piezoelectric sensors vertically positioned around the neck, an inertial motion unit, and long short-term memory (LSTM) neural networks to detect and count swallows. In an in-lab study comprising confounding activities of 10 participants, results show a 3.34 RMSE of swallow count using LSTM, and a 76.07% average F-measure of swallow detection, outperforming the Random Forest classifier. In [150], a wearable system is designed using a piezoelectric sensor that will sense the trachea movement during food consumption and produce an output signal with variations in frequencies over time to detect the swallow cycle. A wearable sensor in the form of

a necklace embedded with a piezoelectric sensor, that detects skin movement from the lower trachea while eating has been presented recently [151]. Their proposed system fully exploits temporal amplitude-varying signals for optimal features, and thus classifies foods more accurately. Additionally, they presented a new approach in which the weight of solid food can be estimated from the swallow count, and the calorie count of food can be calculated from their estimated weight. Their system is based on a smartphone app that helps users live healthily by providing them with real-time feedback about their ingested food types, volume, and calorie count.

2.3.3.3.3 Physiological sensors

In [152] authors detected the swallow cycles through the difference between the normal breathing cycle and the breathing cycle with swallows, using a wearable chest belt. The chest-belt contains a piezo-respiratory belt for converting the changes of tension during breathing to a voltage signal. Data were collected from 3 volunteers who wore the device for one meal. They achieved an accuracy of 82.9% in counting the number of swallows after detecting swallow cycles using SVM classifier. In [153] authors used two Respiratory Inductance Plethysmography (RIP) belts on the chest and abdomen to detect swallow cycles through breathing signals. Support Vector Machine estimating the posterior probability is applied on the extracted features. Through experiments on six healthy subjects, the system is demonstrated to achieve a precision of 65.2% and a recall value of 80.3% in detecting swallow cycles. The development of an automated system to automatically detect swallowing events and provide real-time wristband haptic feedback to facilitate mindful eating using EMG sensors is recently described in [51]. They collected data from 16 participants. They extracted 18 features from the EMG sensor to

train a support vector machine classifier to detect swallows. Their algorithm achieved an F1 score of 0.87 in swallowing detection in leave-one-subject-out validation.

2.3.3.3.4 Distance Sensors

In [130], the authors proposed Acoustic Doppler Sonar (ADS) swallowing detection. Automatic detection of swallowing events was achieved via an artificial neural network. When a 40 kHz ultrasonic beam was focused on the lower jaw and neck, movements of the chin and neck caused Doppler frequency shifts and an amplitude envelope modulation of ultrasonic signals. Hence, it was possible to detect swallow cycles using Doppler frequency shifts in the received ultrasound signals. The experimental results from ten healthy volunteers showed that the proposed ADS-based food intake detection method yielded maximum recognition rates of 78.4% for swallowing detection.

2.3.3.4 Food Item

2.3.3.4.1 Acoustic sensors

AutoDietary, a wearable system to monitor and recognize food intake in daily life is proposed [106]. An embedded hardware prototype is developed to collect food intake sensor data, which is highlighted by a high-fidelity microphone worn on the subject's neck to precisely record acoustic signals during eating in a non-invasive manner. The acoustic data are preprocessed and then sent to a smartphone via Bluetooth, where food types are recognized. A lightweight decision-tree-based algorithm is adopted to recognize the type of food. Experiments involving 12 subjects eating 7 different food items showed that the accuracy of food-type recognition by AutoDietary is 84.9%, and those to classify liquid and solid food intakes are up to 97.6% and 99.7%, respectively. Data collected from an audio sensor placed around the neck is used in [154], to classify food types using a random forest classifier. They extracted 34

predictive features from the audio signal for each second. Then they calculated the average and standard deviation of the 34 features across 5 seconds (centered at each second), resulting in 68 final audio features. They also extracted 98 predictive features from the proximity, ambient-light signals. They trained a random forest classifier (RFC) (n=100 trees) to identify food types based on all the corresponding features extracted from the sensors. They applied a Leave One Food Out (LOFO) approach to each participant for evaluation. The classifier achieved an F score of 97.2% in food classification among 10 participants wearing the device and eating 4 different food items. A novel hardware iHearken, a headphone-like wearable sensor-based system to monitor eating activities and recognize food types in free living has been introduced recently [155]. For food type recognition they used bidirectional long short-term memory on the chewing sound extracted by the headphone. They collected data from 16 subjects and 20 different food items were used. Their proposed algorithm achieved an F score of 97.422% in food type recognition.

2.3.3.4.2 Motion sensors

In [78], authors proposed that multi-modal sensing (in-ear audio plus head and wrist motion) can be used to more accurately classify food type, as audio and motion features provide complementary information. They used features from audio (internal and external microphone in earbud) and motion (9-axis motion sensor from a smartwatch and 9-axis motion sensor from Google glass and trained a random forest classifier to classify food type. To test their method, the authors used data from people wearing audio and motion sensors eating 40 different food items. The proposed method achieved a classification accuracy of 82.7% with a combination of sensors (versus 67.8% for audio alone and 76.2% for head and wrist motion). The Random Forest Classifier resulted in an 86.6% F-measure compared to 77% and 78.1 % F-measures of

Bayesian Networks and KNN (with k=3 yielding the best result), respectively in detecting solid and liquid food items.

2.3.3.4.3 Strain sensors

A novel wearable system in the form of a necklace [148], which aggregates data from an embedded piezoelectric sensor capable of detecting skin motion in the lower trachea during ingestion was proposed to identify food type. The authors proposed an algorithm based on spectrogram analysis of piezoelectric sensor signals to accurately distinguish between food types such as liquid and solid, hot and cold drinks, and hard and soft foods. The necklace transmits data to a smartphone, which performs the processing of the signals, classifies the food type, and provides visual feedback to the user to assist the user in monitoring their eating habits over time. To test the proposed method authors experimented with 20 subjects eating three types of solid food: a meat-like veggie patty, a handful of mixed nuts, and two fun-sized snicker bars. They also consumed two types of liquid, an 8-ounce glass of room-temperature water and 8 ounces of hot tea. The Liquid class label precision and recall was 87.0% and 86.3%, respectively. The Solid class label produced a precision and recall of 86.4% and 87.1 %, respectively, which shows that the collected spectrogram features are good discriminants between liquids and solids.

2.3.3.4.4 Cameras

In [156], authors presented a novel and simple video-based algorithm, to objectively estimate the physical dimensions of objects on top of a table. Authors then utilized this method to estimate food dimensions in [64]. In [157] authors proposed an automatic food image recognition system for recording people's eating habits using the Multiple Kernel Learning (MKL) method to integrate several kinds of image features such as color, texture, and SIFT adaptively. In the experiment, authors achieved a 61.34% classification rate for 50 kinds of

foods. A mobile food recognition system is proposed in [63], where the image recognition algorithm is performed on a smartphone so that the system does not need to send images to a server and runs on an ordinary smartphone in a real-time way. In the experiments, the algorithm achieved an 81.55% classification rate for the top 5 category candidates when the ground-truth bounding boxes were given. In [158], authors described a food image recognition algorithm to propose a food calorie and nutrition measurement system, that can help patients and dietitians to measure and manage daily food intake. They achieved an accuracy of 92.1% in detecting food items among 15 different items. The authors described graph cut segmentation to improve the accuracy of food classification and recognition systems [159]. The results show an improvement in the accuracy of food recognition, especially mixed food where accuracy increases by 15% compared to their previous work. A mobile real-time eating action recognition system was proposed in [160]. The system implemented on a smartphone continuously monitors eating actions during mealtime. It detects the moment when a user eats foods, extracts food regions near the user's mouth, and classifies them. The proposed method achieved a 74.8% classification rate among 5 food items. MT-Diet demo, a fully automated diet assessment system based on a smartphone camera was introduced in [161]. They used a database of 80 frozen meals which contain several different types of foods so the actual total number of foods in their food database is 244 and the database has 33 different types of foods. The accuracy of food identification using a Support Vector Machine with a Radial Basis Function kernel based on color, texture, and histogram of oriented gradient features was 88.5%. Authors then described MT-Diet [162], which interfaces a thermal camera with a smartphone and identifies types of food consumed at the click of a button. Test results on 80 different types of cooked food showed that MT-Diet can isolate food parts with an accuracy of 97.5% and determine the type of food with an accuracy of

88.93%. FIT-EVE&ADAM, an armband-based diet monitoring system that provides food type analysis, triggered by a single hand gesture is proposed in [163]. The system collects the user's gestures using sensors such as an electromyogram embedded in the armband device, along with food image data using color and thermal cameras. The Fuzzy C-Means (FCM) algorithm is used for segmentation and the Sphere Shaped SVM classifier is used to classify the segmented food items in [164]. This method automatically identifies the food items and then calculates their calorie value. The proposed method shows 95% of accuracy in food classification among 100 different food items. In [165], the researcher attempts to design and develop an Android-based food recognition application, that could be used as a health awareness tool for non-health-conscious individuals. The application lets the user take a photo of the food and show its classification result. A novel deep learning-based visual food recognition algorithm, to achieve the best-in-class recognition accuracy was developed in [166]. A food photography application for smartphones, which can recognize 13 types of Thai food from photos, was proposed in [66]. For image recognition, the authors designed a deep convolutional neural network to learn from the dataset. In [65], image classification is performed on the Indian food dataset using different transfer learning techniques. The Indian food dataset of 20 classes with 500 images in each class is used for training and validation. The models used are InccceptionV3, VGG16, VGG19, and ResNet. After experimentation, it was found that Google InceptionV3 outperformed other models with an accuracy of 87.9% and a loss rate of 0.5893. A deep model-based food recognition and dietary assessment system, to study and analyze food items from daily meal images (e.g., captured by smartphone) was described in [167]. Specifically, the authors proposed a three-step algorithm to recognize multi-item (food) images by detecting candidate regions and using a deep convolutional neural network (CNN) for object classification.

2.3.3.4.5 Other sensors

A novel method that utilized optical, ion-selective electrical pH, and conductivity sensors to, sense and classify liquid in a cup in a practical way is described in [168]. Results from two experiments showed that this method is feasible and relatively accurate (up to 79% classification for 68 different drinks). The design, implementation, and validation of the Liquid Intake Detection System (LIDS) for real-time tracking of fluid intake type and volume was proposed [80]. The system devises a sensing module that is composed of ultrasonic, RGB color, temperature, and accelerometer sensors as well as a computational framework for machine-learning-based fluid intake type classification, volume estimation, and bottle-state recognition.

2.3.3.5 Eating Time/Rate:

2.3.3.5.1 Motion sensors

A real-time food intake monitoring system for mobile devices was proposed [93], which provides feedback based on the eating speed of the user. The proposed system gets acceleration data from the sensor placed on the wrist of the user which is sent to mobile devices via Bluetooth to be analyzed to detect bites. If the eating is too fast, the system warns the user in the form of both audio and text. A smartwatch app was presented [73], which is designed to help users eat meals at an appropriate speed or rate. The app measures the user's eating rate/speed in real-time and delivers appropriate feedback to slow down when eating fast. The authors explored four different types of feedback - graphics, text, clock, and vibration, to find the suitable type of feedback depending on the eating environment. A novel methodology, IDEA (Instant Detection of Eating Action) that performs accurate eating action identification and provides feedback on eating rate/speed is described in [169]. IDEA uses a single wristband with IMU sensors and functions without any manual intervention from the user.

2.3.3.5.2 Strain sensors

Using the device described in [53], [123], [124], time resolution needed to, accurately represent the microstructure of meals such as eating rate/speed, the duration of meal , the duration of actual ingestion, and the number of eating events was estimated in [72]. The results suggest that the desired time resolution of sensor-based food intake detection should be ≤5 sec to accurately detect meal microstructure. A necklace-shaped piezoelectric sensor [128] was used to detect meal duration and eating rate/speed and deliver the right feedback in the form of light in real-time.

2.3.3.5.3 Distance sensors

In [74], the authors described a technique for estimating meal times using an earphone-type wearable sensor. A small optical sensor composed of a light-emitting diode and phototransistor was inserted into the ear hole of a user and estimates the mealtimes of the user from the time variations in the amount of light received.

2.3.3.5.4 Physiological sensors

A wearable sensor system that can monitor breathing signals for estimating the time and duration of a meal was proposed in [170]. The system combines swallowing signatures from breathing signals with hand movement signatures from hand acceleration to train a hierarchical Support Vector Machine (SVM) classifier and a Hidden Markov Model (HMM), for mealtime and meal duration estimation. An armband-based diet monitoring system, FIT-EVE&ADAM, that provides eating speed was described in [163]. The system collects the user's gestures using sensors such as an electromyogram embedded in the armband device and estimates eating rate/speed.

2.3.3.5.5 Other sensors

In [71], the authors conducted a study to explore relationships among self-reported, laboratory-measured eating rate/speed using a Universal Eating Monitor (UEM). UEM is an apparatus that has been developed to permit the covert continuous weighing of a subject's plate or other food reservoir using a concealed electronic balance. By coupling the device with a digital computer, they were able to record precisely the amount consumed every 3 s throughout a single-course meal consisting of a relatively homogeneous mixture of foods. Whether the effect of eating rate/speed on energy intake is the same in normal-weight and overweight/obese subjects was examined in [70] using UEM. In [47], the authors explored the effect of portion size on bite size, eating rate/speed, deceleration rate, and meal duration using the Sussex Ingestion Pattern Monitor. A qualitative study to assess the acceptability, perceived efficacy, and user experience of a smart fork, to provide real-time feedback, based on eating rate/speed was described in [171]. Meal detection and meal-size estimation algorithms were developed in [172] using glucose sensors. The effect of real-time vibrotactile feedback delivered through an augmented fork on eating rate/speed is examined in [173]. The design of a new eating utensil that can reduce the eating rate/speed by interfering with the eater's ability to eat quickly was proposed in [174]. This utensil can change its rigidity and shape by deflating itself to interfere with eating. The development of a new sensing device for measuring the pace, time, order, and intake of meal consumption is described in [175]. The developed device is a portable sensor sheet consisting pressure sensor, which is suitable for use in homes because it is designed not to disturb meal consumption and can be used conveniently. Another smart eating utensil that can potentially increase the users' awareness of their eating rate/speed was proposed in [81].

2.3.3.6 Mass of Food Intake

2.3.3.6.1 Acoustic sensors

In [78], the authors used features from audio (internal and external microphone in earbud) sensors to develop a regression model to predict the amount of consumed food items from the internal and external microphones by customizing a standard earbud. They reported a 35.4% error in measuring the amount of consumed food by using sensor and image features from the annotation of 30 meals.

2.3.3.6.2 Motion sensors

A ubiquitous and unobtrusive system named FluidMeter, to track the amount of fluid intake leveraging the inertial sensors embedded in smartwatches, was described [176]. To achieve this, FluidMeter first separates the drinking activities from other human activities (playing, running, eating, etc.). Thereafter, it analyzes the sampled sensor data during the extracted drinking episodes to recognize the sequence of micro-activities (lift the bottle, sip, release the bottle) that constitute the drinking activity. Finally, it applies some machine learning algorithms on some features extracted from sampled sensor data during the sipping period to gauge the amount of fluid intake in the designated drinking episode. FluidMeter can estimate the overall amount of fluid intake in grams accurately with an estimation error limited to 15% using data from 72 volunteers. Another measurement of liquid intake using inertial sensors was described in [80]. The system devises a sensing module that is composed of ultrasonic, RGB color, temperature, and accelerometer sensors as well as a computational framework for machine-learning-based fluid intake type classification, volume estimation, and bottle-state

recognition. Their results for volume estimation show that the regression-based volume estimation supports a root-relative-squared error that ranges from 1.12% to 13.36%.

2.3.3.6.3 Strain sensors

Mom's Tray, a dietary monitoring system, that can measure the weight of the food intake, integrates a smart tray with a force sensing resistor, prearranged and RFID-tagged food packages, and a mobile app to provide real-time feedback on food ordering and consumption in a school cafeteria setup is described in [177]. In [178], the authors explored the potential use of a wearable system consisting piezoelectric strain system for providing just-in-time (JIT) feedback on the progression of a meal and tested its ability to reduce food intake. Experimental results showed that JIT feedback targeting a 25% reduction in chew counts resulted in a reduction in mass and energy intake.

2.3.3.6.4 Camera

The novel and challenging problem of predicting the relative amount of each ingredient from a food image is described in [67]. The authors proposed PITA, the Picture-to-Amount, deep learning architecture to solve the problem. A novel methodology for evaluating the weight of food by utilizing extracted features from images and training them through advanced boosting regression algorithms has been presented recently [179]. A unique dataset of 23,052 annotated food images of Mediterranean cuisine, including 226 different dishes with a reference object placed next to the dish, was used to train the proposed pipeline. Between the predicted and actual weight values for each food in the proposed dataset, the proposed model achieves a mean absolute percentage error of 3.73%.

2.3.3.6.5 Other sensors

In [180], authors described the development of a universal eating monitor (UEM) that permits the covert continuous weighing of a subject's plate using a concealed electronic balance. The effect of awareness that food intake is being measured by a universal eating monitor on the consumption amount is examined in [83]. Experimental results suggest that awareness of being monitored by the UEM does not affect the consumed mass. An algorithm for automatically processing continuous in-meal weight measurements using the Mandometer, to extract the clean and in-meal eating indicators, such as total food intake mass and food intake rate is described in [82]. To measure the weight of consumed food, the table had recessed scales that continuously measured plate weight at a rate of 15 Hz [181]. Here bites weight was measured as the change in plate weight between each bite of food. In [142], a regular dining tray was equipped with a video camera and three separate built-in weighing stations to continuously measure the weight of the bowl, plate, and drinking cup. To compare the performance of image-based dietary assessment, food intake mass was recorded from the weight scale [182].

2.3.3.7 Food Intake Environment

2.3.3.7.1 Cameras

A sensor device with a camera that automatically takes a photograph of the surroundings of a user's meal to examine the meal environment of the user was proposed [183]. The device is attached to a ceiling light in the dining room of the user's house. The utility of using wearable camera images to assess the context of eating episodes was investigated in [85]. Adult participants (N = 40) wore a SenseCam wearable camera for free-living conditions and had their diet assessed using three image-assisted multiple-pass 24-hour dietary recalls. Data for 107 days were used, with a total of 742 eating episodes considered for annotation. Most episodes were

seated at tables (27%) or sofas (26%) but eating standing (19%) or at desks (18%) was common. Social interaction was evident for 45% of episodes and media screens were viewed during 55% of episodes.

2.3.3.7.2 Other sensors

The influence of environmental conditions on energy intake at lunchtime in normal-weight versus overweight male adolescents was investigated in [87]. Healthy 15–17-year-old adolescents (19 normal-weight, 19 overweight) participated in lunch tests in laboratory settings. Four conditions were compared: eating in groups, eating alone, eating alone while viewing television, and eating alone while listening to music. The same menus (two solid foods, three drinks) were offered ad libitum. Significant group differences between meal conditions were observed for energy intake from solid foods, but not from drinks: normal-weight participants ate more solids while listening to music (5731 ± 426 kJ) than when eating alone (5012 ± 364 kJ, P = 0.026) or in groups (4974 ± 272 kJ, P = 0.049), whereas overweight participants ate more solid foods while viewing television (5806 ± 330 kJ) than when eating in groups (5208 ± 201 kJ, P = 0.014) or while listening to music (5288 ± 255 kJ, P = 0.035). On average, total energy intake at lunch (solids plus liquids) was found to be increased only in the overweight when viewing television (8527 ± 535 kJ) compared to eating in a group (7348 ± 445 kJ, P = 0.037) or while listening to music (7532 ± 435 kJ, P = 0.049). To determine whether there is an interaction between “small” (i.e., table setting) and “large” (i.e., eating location) scales of the eating environments in affecting food acceptance and consumption a study was performed in [88]. The authors' findings demonstrated that food acceptance and intake can vary according to whether the small (table setting) and large (eating location) scales of the eating environments are well-matched or not. FamilyLog - a practical system to log family mealtime activities using

smartphones and smartwatches was presented in [184]. FamilyLog automatically detects and logs details of activities during mealtime, including the occurrence and duration of the meal, conversations, participants, TV viewing, etc., in an unobtrusive manner. A new approach to studying the effect of the eating environment on food intake and eating behavior using virtual reality technology is described in [185]. Fifteen adults consumed pizza rolls in two virtual reality (VR) environments: a restaurant and a table in an empty room. This study suggested that VR could be developed as a new tool to study the effect of the eating environment on food intake and eating behavior.

2.4 Discussion

This paper intends to provide a systematic assessment of existing technology-driven measurement devices for different characteristics of eating.

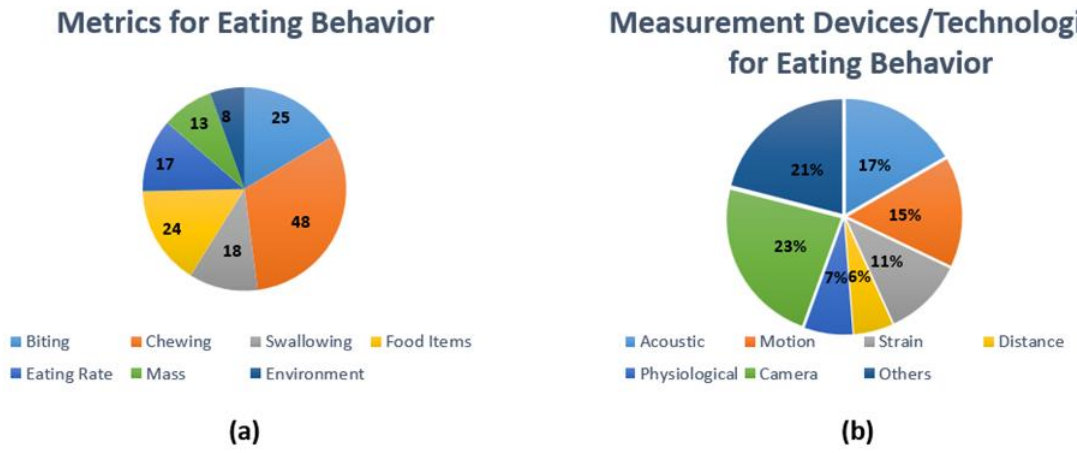


Figure 2-3: Summary of findings: (a) Number of articles found for quantifiable metrics of eating behavior (b) Percentage of articles found for the measurement devices/technologies used to measure the eating behavior metrics.

We identified the quantifiable physiological and environmental metrics of eating behavior and grouped them into seven major classes. The measurement devices/technologies used to measure

these eating behavior metrics were identified and grouped into seven groups. The summary of the review findings is shown in Figure 2-3.

Among the reviewed 153 full-text articles, 25 articles were found for the detection and tracking of dynamic characteristics of eating, known as biting. The summary of articles for bite characteristics is represented in Table 3. 44% (11) studies used motion sensor technology such as accelerometer, gyroscope etc. Distance sensors such as proximity sensors, capacitive sensors, and ambient light sensors were used in 12% (3) articles to detect biting. 28% (7) of studies used imaging technology such as depth cameras, action cameras, digital camcorders, fisheye cameras, etc. to track bite events. The remaining 16% (4) studies used weight sensors, food sensors, smart forks, etc. to detect and count the number of bites in eating episodes which are grouped into other categories of measurement devices. Among 25 studies, only 12 (50%) were found to test their methodology outside the laboratory and only 4 (16%) reported real-time feedback to the user to provide healthy eating suggestions

Table 2-3: Summary of the measurement Devices for Biting

Article	Device Group	Sensor	Position	Study Participants	Laboratory	Free-living	Real Time	Performance
[89]	Motion	Accelerometer, gyroscoe	Wrist	22	Yes	Yes	No	F1 Score = 0.923
[43]	Motion	Orientation	Wrist	10	Yes	No	No	Sensitivity =91%
[90]	Motion	Accelerometer Gyroscope	Wrist	271	Yes	Yes	No	Sensitivity =75%
[91]	Motion	Accelerometer Gyroscope	Wrist	276	Yes	Yes	No	Accuracy = 79.7%
[49]	Motion	Gyroscope	Wrist	99	Yes	Yes	Yes	NA
[92]	Motion	Accelerometer Gyroscope	Wrist	34	Yes	Yes	No	False positive rate = 6.5% False Negative Rate = 3.3%
[93]	Motion	Accelerometer	Wrist	1	Yes	No	Yes	NA
[94]	Motion	Accelerometer Gyroscope	Wrist	3	Yes	Yes	No	Accuracy = 91.8%
[95]	Motion	Accelerometer Gyroscope	Wrist	12	Yes			F1 Score= 0.91

[186]	Motion	Gyroscope	Wrist	8	Yes	No	No	Accuracy => 90%
[44]	Motion	Tri-axial accelerometer	wrist	15	Yes	No	No	Accuracy = 81.2%
[96]	Distance	Magnetic proximity	Neck	1	Yes	No	No	NA
[187]	Distance	Capacitive Sensing	Icecream Cone	NA	Yes	No	Yes	NA
[97]	Distance	Ambient light	Neck	20	Yes	Yes	No	F1 Score = 77.1%
[98]	Camera	Depth Camera (Kinect Xbox One)	In front of the user	1	Yes	No	No	Sensitivity= 96.2%
[45]	Camera	SJ4000 Action Camera (Black)	3 feet away from the user	28	Yes	No	No	Accuracy = 85.4%
[188]	Camera	Digital camcorder	1.5 meters away from the user	85	Yes	No	No	Recall =91.7%
[99]	Camera	Digital Camcorder	1.5 meters away from the user	85	Yes	No	No	F1 Score = 0.948
[100]	Camera	360 Degree Camera	In front of the table	4	Yes	Yes	No	Error = 26.2%
[101]	Camera	Digital camera	Ceiling	264	Yes	No	No	F1 score =0.899
[42]	Camera	Fish eye camera	Shoulder	16	Yes	Yes	No	NA
[189]	Others	Electrical Conductivity of Foods	Food Item	1	Yes	No	No	NA
[102]	Others	Augmented Fork	Fork	141	Yes	Yes	Yes	NA
[103]	Others	Weight Sensors	Plate	24	Yes	Yes	No	Precision =0.76 Recall = 0.76
[190]	Others	Voltage Divider	Embedded Fork	6	Yes	Yes	No	Accuracy = 77%

Chewing was detected and tracked in 48 (31%) studies found in this review process. The summary of articles on chewing is represented in Table 4. Among 48 studies, 14 (29%) used an acoustic sensor to detect and count the number of chews in eating episodes. Motion sensors were found to track the dynamic process of chewing in 7 (15%) studies. 8 (17%) studies used strain sensor technology for chewing detection. Different physiological sensors such as EMG, photoplethysmography (PPG), etc. were used in 7 (15%) of the studies. Distance sensor was used in 5 (10%) studies. The imaging technology was used in 6 (13%) of the studies to detect and

count the number of chews in eating episodes using a digital camera, action camera, and smartphone camera. Out of 48 only 14 (29%) studies tested their device in free-living conditions to track the dynamic process of chewing. Only 2 (4%) studies reported real-time feedback based on chewing rate to the user.

Table 2-4: Summary of measurement devices for Chewing

Article	Device	Sensor	Position	Study Participants	Laboratory	Free Living	Real Time	Performance
[104]	Acoustic, Motion	Accelerometer, Gyroscope, Microphone	Ear-bud	5	Yes	No	No	Accuracy = 97%
[105]	Acoustic, Camera	Microphone, camera	Ear	6	Yes	Yes	No	Accuracy =80%
[106]	Acoustic	Throat Microphone	Neck	12	Yes	No	No	Accuracy =86.6%
[107]	Acoustic	Two-channel condenser microphone	Under ear	18	Yes	No	No	F score = 0.8
[108]	Acoustic	Throat microphone	Neck	12	Yes	No	No	F score =71.4%
[109]	Acoustic, Motion	9 axis IMU, microphone	Wrist, Ear	6	Yes	Yes	No	Recall = 84% Precision =67%
[110]	Acoustic	Throat Microphone	Neck	8	Yes	No	No	Accuracy = 0.783
[111]	Acoustic	Microphone	Ear	NA	Yes	No	No	Error Rate =1.93%
[130]	Acoustic	Ultrasonic Doppler Sensor	Neck	10	Yes	No	No	Accuracy =91.4%
[112]	Acoustic	Microphone	Ear	55	Yes	No	No	Precision>80 Recall >80%
[113]	Acoustic	Bluetooth Headset	Ear	28	Yes	No	No	Accuracy =(77-94)%
[114]	Acoustic	Skin contact microphone	Neck	14	Yes	No	No	F score = 77.5%
[115]	Acoustic	Bone conduction microphone	Ear	6	Yes	No	Yes	Accuracy =97.6%
[116]	Acoustic, Motion	Microphone, 9-axis IMU, 9 axis motion sensor	Ear, Wrist, Head	6	Yes	Yes	No	Accuracy =85%
[117]	Acoustic	Bone Conduction Microphone	Ear	9	Yes	Yes	No	Accuracy =97.1%
[118]	Acoustic	Microphone	Eyeglass	5	Yes	No	No	F score =0.96
[119]	Motion, physiological	Accelerometer, Orientation, Gyroscope, EMG	Wrist	36	Yes	No	No	F score =0.92
[126]	Motion, Strain	Accelerometer, Hand Gesture sensor, Piezoelectric strain	Wrist, Wrist, Below Ear	12	Yes	Yes	No	Accuracy =89.8%
[120]	Motion	Single axis Accelerometer	Temporalis Muscle	4	Yes	No	No	Accuracy =97% F score =93%

[121]	Motion	3 axis accelerometer, 3 axis Gyroscope, 3 axis Magnetometer	Chin	13	Yes	No	No	NA
[122]	Strain	Piezoelectric Strain	Below ear	20	Yes	No	No	Accuracy =80.98%
[123]	Strain	Jaw motion sensor	Below ear	12	Yes	Yes	No	Accuracy = 86.86 ± 6.5%
[53]	Strain	Piezoelectric Strain Sensor	Below Ear	30	Yes	No	No	Error rate =9.66%
[124]	Strain	Piezoelectric Strain Sensor	Below ear	5	Yes	No	No	Error rate =8.09%
[54]	Strain	Piezoelectric Strain Sensor	Below Ear	30	Yes	Yes	No	Error rate =15.01%
[125]	Motion, Strain	Accelerometer, Piezoelectric Strain	Eyeglass temple, below ear	10	Yes	No	No	F score = 99.85%
[127]	Motion, Strain	Accelerometer, hand gesture sensor, Piezoelectric Strain	Wrist, Wrist, Below Ear	12	Yes	Yes	No	Accuracy =93%
[52]	Strain	Piezoelectric Strain	Temporalis Muscle	10	Yes	Yes	No	Error Rate =3.83%
[128]	Physiological, Strain	EMG, piezo	Chin, Neck	10	Yes	No	No	Accuracy =0.938
[97]	Distance	Proximity	Neck	20	Yes	Yes	No	F1 Score = 77.1%
[55]	Motion, Distance	Accelerometer, Infrared Distance Sensor	Ear pinna	22	Yes	No	No	Precision =85.3% Recall =84.5%
[129]	Distance	Proximity Sensor	Necklace	32	Yes	Yes	No	Precision =78.2&% Recall = 72.5%
[131]	Distance	Proximity Sensor	Right temple of eyeglass	10	Yes	No	No	Error rate =2.69%
[132]	Distance	Proximity Sensor	Eyeglass	20	Yes	No	No	Accuracy =96.4%
[133]	Physiological	EMG	Right and Left masseter, anterior temporalis miscle	37	Yes	No	No	NA
[57]	Physiological	EMG	Right and left masseter, Right and left temporalis muscle	13	Yes	No	No	NA
[134]	Physiological	Portable EMG	Center of masseter, mastoid	28	Yes	No	No	NA
[135]	Physiological	EMG	Eyeglass	8	Yes	No	No	Precision =80%
[65]	Acoustic, Physiological	Microphone, photoplethysmography (PPG)	Ear	22	Yes	Yes	No	Accuracy =0.938
[138]	Physiological	EMG	between the mastoid and the masseter muscle	15	Yes	Yes	No	Sensitivity >90%
[136]	Physiological	EMG	Eyeglass	10	Yes	Yes	No	Precision >77%

[139]	Physiological	Myoelectric sensor	Masseter muscle	8	Yes	No	Yes	NA
[56]	Camera	Digital camera	In front of user	6	Yes	No	No	NA
[45]	Camera	SJ4000 Action Camera (Black)	3 feet away from user	28	Yes	No	No	Accuracy = 88.9%
[140]	Camera	Digital camera	In front of the user	37	Yes	No	No	NA
[141]	Camera	Smartphone camera	In front of the user	100	Yes	No	No	Error Rate =7%
[142]	Camera	video camera	Dining tray	NA	Yes	No	No	NA
[143]	Others	EMG, Piezoelectric Strain Sensor, Piezoresistive Sensor, Pressure Sensor	Right temporalis muscle, left temporalis muscle, eyeglass,ear canel	15	Yes	No	No	NA

Swallowing, another dynamic characteristic of eating, was detected in 18 (12%) studies among 153 full-text articles. Among 18, 10 (66%) studies were found to detect swallows using an acoustic (microphone) sensor. A piezoelectric strain sensor placed on the throat was used in 5 (28%) studies. Piezo-respiratory belt, and respiratory inductance plethysmography (RIP) belt on the chest and abdomen, EMG were used in 3 (17%) of the studies to detect and count swallows. No studies were found in this review that tested their device outside a laboratory-controlled environment. 6 studies (33%) were found to provide real-time feedback based on swallow detection to the user. The summary of the articles found for swallowing is presented in Table 2-5.

Table 2-5: Summary of measurement devices for swallowing

Article	Device	Sensor	Position	Study Participants	Laboratory	Free Living	Real Time	Performance
[144]	Acoustic	Throat microphone	Throat	21	Yes	No	No	NA
[62]	Acoustic	Throat microphone	Throat	20	Yes	No	No	Accuracy =84.7%
[145]	Acoustic	Throat and Ambient Microphone	Throat	7	Yes	No	No	Recall >85%
[146]	Acoustic	Throat Microphone	Throat	7	Yes	No	No	Accuracy >94%
[147]	Acoustic	Microphone	Neck	85	Yes	No	Yes	Accuracy >79.3%
[106]	Acoustic	High fidelity microphone	Neck	12	Yes	No	Yes	Accuracy =86.6%
[130]	Acoustic	Ultrasonic Doppler Sensor	Neck	10	Yes	No	No	Accuracy =78.4%
[110]	Acoustic	Throat Microphone	Neck	8	Yes	No	No	Accuracy =0.712
[117]	Acoustic	Bone conduction microphone	Neck	9	Yes	No	Yes	Accuracy = 97.1%
[115]	Acoustic	Bone conduction microphone	Ear	6	Yes	No	Yes	Accuracy =97.6%
[148]	Strain	Piezoelectric Sensor	Lower Trachea	20	Yes	No	Yes	F score =80%
[149]	Strain	Piezoelectric Sensor	Lower Trachea	20	Yes	No	Yes	F score =91.2%
[60]	Strain	Piezoelectric Sensor, IMU	Neck	10	Yes	No	No	F score =76.07%
[150]	Strain	Piezoelectric Sensor	Lower Trachea	NA	Yes	No	No	NA
[151]	Strain	Piezoelectric Sensor	Lower Trachea	18	Yes	No	Yes	NA
[152]	Physiological	Piezo-respiratory belt	Chest	3	Yes	No	No	Accuracy =80%
[153]	Physiological	Respiratory Inductance Plethysmography (RIP) belt	Chest and Abdomen	6	Yes	No	No	Precision = 80%
[51]	Physiological	EMG	Masseter muscle	16	Yes	No	No	F1 score = 0.87

Among full-text articles, 24 (16%) studies were found to include food items to describe the dynamic process of eating. Only 3 studies were found that used acoustic means such as microphone to recognize food item types. Only 1 study was found that used a motion sensor to detect types and number of food items in an eating episode. A piezoelectric strain sensor attached to the lower trachea was used in one study to recognize food items. Most of the food item

recognition methodologies use different cameras such as smartphone cameras, thermal cameras, and digital cameras. 17 studies out of 23, were found to use the camera. For liquid food type and mass detection, ion selective PH sensor, conductive sensor, RGB color sensor, and temperature sensor were used in 2 studies. Among the 24 studies found in this review for food item characteristics of eating, none of them were tested in free-living conditions and only 5 (21%) studies provided real-time feedback to the user. The summary table of articles found for food item is presented in Table 2-6.

Table 2-6: Summary of measurement devices for Food Items.

Article	Device	Sensor	Position	Study Participants	Laboratory	Free Living	Real Time	Performance
[106]	Acoustic	High fidelity microphone	Neck	12	Yes	No	Yes	Accuracy = 84.9%
[148]	Strain	Piezoelectric Sensor	Lower Trachea	20	Yes	No	Yes	Precision =80%
[154]	Acoustic	Lavalier microphone (MAONO	Shirt collar	10	Yes	No	No	F score = 97.2%
[155]	Acoustic	Microphone	Over Ear	16	Yes	Yes	No	F score = 97.44%
[78]	Motion	Internal and External Microphone, 9-axis IMU in the wrist, 9-axis IMU in head	Ear, Wrist, Head	6	Yes	No	No	Accuracy = 82.7%
[156]	Camera	Digital camera	On top of table	NA	Yes	No	No	NA
[64]	Camera	Digital camera	On top of table	NA	Yes	No	No	NA
[157]	Camera	Cellular phone camera	Smartphone	NA	Yes	No	No	Accuracy =61.34%
[183]	Camera	Digital camera	Ceiling	NA	Yes	No	No	NA
[63]	Camera	Smartphone camera	Smartphone	NA	Yes	No	Yes	Accuracy = 81.55%
[158]	Camera	Smartphone camera	Smartphone	NA	Yes	No	No	Accuracy =92.1%
[159]	Camera	Smartphone Camera	Smartphone	NA	Yes	No	No	NA
[160]	Camera	Slight tilt in front of user	Smartphone	5	Yes	No	Yes	Classification rate =74.8%
[161]	Camera	Smartphone camera	Smartphone	NA	Yes	No	Yes	Accuracy = 88.5%
[162]	Camera	Thermal Camera	Smartphone	NA	Yes	No	No	Accuracy = 88.93%

[163]	Camera	Color, Thermal Cameras	Wrist	NA	Yes	No	No	NA
[164]	Camera	Camera	Smartphone	NA	Yes	No	No	Accuracy = 95%
[165]	Camera	Smartphone camera	Smartphone	NA	Yes	No	No	NA
[166]	Camera	Smartphone Camera	Smartphone	NA	Yes	No	No	NA
[66]	Camera	Smartphone Camera	Smartphone	NA	Yes	No	No	Accuracy =82%
[65]	Camera	Smartphone camera	Smartphone	NA	Yes	No	No	Accuracy =87.3%
[167]	Camera	Smartphone Camera	Smartphone	NA	Yes	No	No	NA
[168]	Others	Ion selective pH, conductivity	Cup	NA	Yes	No	No	Accuracy =79%
[80]	Others	Ultrasonic, RGB color, Temperature	Bottle	NA	Yes	No	No	Accuracy = (74.93% - 94.98%)

17 studies were identified in this review that reported eating rate/speed, to describe the dynamic process of eating. Among these 17 studies, 3 used motion sensor technology to report eating speed. 2 studies were found that described the use of a strain sensor to provide the eating speed of an eating episode.

Only 1 study was found that used distance sensing using an optical sensor to report the user's eating speed. EMG, RIP belts were used as a physiological sensors in 2 studies. Universal eating monitor, sussex ingestion pattern monitor, smart fork, and smart utensil were used in 9 of the studies to provide the eating speed of the participants. Out of 17 studies only 3 studies were reported to be tested outside a laboratory setting and 4 studies were found which provide real-time eating speed feedback to the user. The summary table of the articles found for eating rate/speed metric is shown in Table 2-7.

Table 2-7: Summary of measurement devices for Eating Time/Rate

Article	Device	Sensor	Position	Study Participants	Laboratory	Free Living	Real Time	Measured Metric
[93]	Motion	Accelerometer	Wrist	1	Yes	No	Yes	Eating speed
[128]	Strain	Piezoelectric	Neck	10	Yes	No	Yes	Eating speed
[163]	Physiological	EMG	Wrist	17	Yes	Yes	No	Eating speed
[73]	Motion	Accelerometer, IMU	Wrist	NA	Yes	No	Yes	Eating speed
[169]	Motion	IMU	Wrist	36	Yes	No	No	Eating Speed
[72]	Strain	Piezoelectric Strain	Temporalis Muscle	12	Yes	No	No	No. of meals, meal duration, Duration of actual ingestion
[74]	Distance	Optical Sensor	Ear	11	Yes	Yes	No	Mealtime
[170]	Physiological,	two Respiratory Inductance Plethysmography (RIP) belts	Chest and abdomen	14	Yes	No	No	Mealtime, duration
[71]	Others	Universal Eating Monitor	Table	60	Yes	Yes	No	Eating rate
[70]	Others	Weight Scale	Table	35	Yes	No	No	Eating rate
[47]	Others	Sussex Ingestion Pattern Monitor	Table	35	Yes	No	No	Eating rate, bite size, meal duration
[171]	Others	Smart Fork	Fork	11	Yes	Yes	Yes	Eating speed
[172]	Others	Glucose sensor	Artificial pancreas	30	Yes	No	No	Meal-size
[173]	Others	Smart Fork	Fork	128	Yes	No	No	Eating rate
[174]	Others	Smart utensil	Utensil	NA	Yes	No	No	Eating rate
[175]	Others	Pressure Sensor	Sheet	2	Yes	No	No	Mealtime, pace, duration
[81]	Others	Smart Utensil	Utensil	10	Yes	No	No	Eating rate

To describe the dynamic process of eating, 13 studies out of 153, reported food intake mass of eating episodes. Out of the 13 studies, only 1 study used acoustic sensing through a microphone to report the mass of food intake. 2 studies used an accelerometer and gyroscope signal to calculate the amount of food consumed in an eating episode. Strain sensing was used in 2 studies out of 12 to estimate food intake mass. A smartphone camera was used in one study.

Electronic balance, mandometer, and weight scales were used in 6 studies to estimate and produce cumulative food consumption. Among these 13 studies, only 3 of them were performed in free-living conditions and none of the studies reported real-time feedback to the user. The summary table of articles found for food intake mass metric is represented in Table 2-8.

Table 2-8: Summary of measurement devices for Mass of food intake

Article	Device	Sensor	Position	Study Participants	Laboratory	Free Living	Real Time	Performance
[78]	Acoustic	Internal and External Microphone, 9-axis IMU in wrist, 9-axis IMU in head	Ear, Wrist, Head	6	Yes	No	No	Error =35.4%
[80]	Motion	Accelerometer	Bottle	NA	Yes	No	No	Error =13.36%
[176]	Motion	Gyroscope, Accelerometer	Wrist	41	Yes	Yes	No	Accuracy =59.2%
[177]	Strain	Force sensing resistor	Tray	10	Yes	Yes	No	NA
[178]	Strain	Piezoelectric Strain	Temporalis Muscle	18	Yes	Yes	No	NA
[67]	Camera	Smartphone Camera	Smartphone	NA	Yes	No	No	NA
[179]	Camera	Smartphone Camera	Smartphone	NA	Yes	No	No	Error = 3.73%
[180]	Others	Electronic Balance	Table	26	Yes	No	No	NA
[83]	Others	Electronic Balance	Table	39	Yes	No	No	NA
[82]	Others	Mandometer	Table	77	Yes	No	No	Accuracy =0.69%
[181]	Others	Weight scale	Table	72	Yes	No	No	NA
[182]	Others	Weight scale	Table	84	Yes	No	No	NA
[142]	Others	Weight scale	Tray	NA	Yes	No	No	NA

2.5 Challenges and Future Trends

Though the technology-driven methods for detection and tracking different dynamic characteristics of eating made significant improvements over the past two decades, they are still not widely used in clinical and research practice. In this review, we found the use of stationary/portable sensor systems which are limited to the place of installation and thus, are not practical solutions to use for free-living conditions (where the food may be consumed at various

places throughout the day). On the other hand, non-invasive wearable sensors are a more practical choice to monitor eating behavior in free living. Two major issues need to be addressed in wearable sensor approaches i) sensor burden and ii) privacy concerns. In the case of sensor burden, it is important to evaluate whether the wearable device put any burden on the user, how the sensor looks when users wear the device, does the sensor cause any pain or tickling, does the sensor restrict the way of eating or movement, and finally the level of acceptability. Most of the studies did not answer those questions. To make wearable sensors acceptable to people, they should not put any sensor burdens on the users. More work is needed to develop wearable sensor methods that are not only accurate and precise to detect food intake but also good-looking, user-friendly, and widely acceptable by people. Few studies used off-the-shelf widely used devices such as a smartwatch, google glass, etc. to overcome the sensor burden issue, but did not overcome the privacy issue. As acoustic-based wearable sensor records sound and camera-based sensors capture images, both are subject to privacy concerns for the user if the device collects data continuously. More research is needed to eliminate non-food activity sounds or images so that the privacy of the user should be preserved. Another significant finding of the review is the very limited number of studies that tested their proposed system and algorithm outside the laboratory (free-living). More outside laboratory testing is necessary to prove its usability in day-to-day life. On top of that, the review identified that the proposed systems are tested for 1 or 2 days for each user. To evaluate the long-term effect of the eating monitoring system, it needs to be tested for a long duration (e.g., a week or a month for each user). As an indirect outcome of this review, we noticed that a significantly higher number of research publications target the detection of food intake and recognition of food being eaten. But eating is a highly dynamic behavior that occurs over time and in context. Research has repeatedly shown that contexts have

important impacts on eating patterns and health outcomes, including social contexts like eating with family, physical contexts like living in a rural food environment, and psychological contexts like chronic stress levels. Eating data from wearable sensors, in conjunction with other technologies that can detect these additional contextual features, can be used to eventually develop models of human eating behavior in real-time and in context, which can be leveraged to develop adaptive, personalized interventions to predict and change obesity-related behavior, and ultimately, improve population health.

2.6 Conclusion

This systematic review explored the technology-driven methodologies proposed in the past two decades to detect and track the dynamic process of eating. The paper contributed a taxonomy of measurement devices for dietary intake technologies that combined eating characteristics and sensor modalities. Current methodologies to detect and track different metrics of eating can be improved using a combination of sensor modalities and the implementation of powerful machine learning algorithms such as pattern recognition, classification, regression, deep learning, etc. Furthermore, the measurement devices need to be designed to use outside of restricted laboratory conditions. If these current challenges are addressed, food intake monitoring devices can be used in everyday life to promote healthy eating behavior.

References

[1] M. L. Neuhouser, “The importance of healthy dietary patterns in chronic disease prevention,” Nutrition Research, vol. 70, pp. 3–6, Oct. 2019, doi: 10.1016/j.nutres.2018.06.002.

[2] F. Jannasch, J. Kröger, and M. B. Schulze, “Dietary Patterns and Type 2 Diabetes: A Systematic Literature Review and Meta-Analysis of Prospective Studies,” The Journal of Nutrition, vol. 147, no. 6, pp. 1174–1182, Jun. 2017, doi: 10.3945/jn.116.242552.

[3] E. Tourlouki, A.-L. Matalas, and D. B. Panagiotakos, “Dietary habits and cardiovascular disease risk in middle-aged and elderly populations: a review of evidence,” Clin Interv Aging, vol. 4, pp. 319–330, 2009.

[4] F. E. Thompson, A. F. Subar, C. M. Loria, J. L. Reedy, and T. Baranowski, “Need for Technological Innovation in Dietary Assessment,” J Am Diet Assoc, vol. 110, no. 1, pp. 48–51, Jan. 2010, doi: 10.1016/j.jada.2009.10.008.

[5] J.-S. Shim, K. Oh, and H. C. Kim, “Dietary assessment methods in epidemiologic studies,” Epidemiol Health, vol. 36, p. e2014009, Jul. 2014, doi: 10.4178/epih/e2014009.

[6] A. Magarey et al., “Assessing dietary intake in children and adolescents: Considerations and recommendations for obesity research,” International Journal of Pediatric Obesity, vol. 6, no. 1, pp. 2–11, Feb. 2011, doi: 10.3109/17477161003728469.

[7] M. B. E. Livingstone, P. J. Robson, and J. M. W. Wallace, “Issues in dietary intake assessment of children and adolescents,” British Journal of Nutrition, vol. 92, no. S2, pp. S213–S222, Oct. 2004, doi: 10.1079/BJN20041169.

[8] D. A. Schoeller, “Limitations in the assessment of dietary energy intake by self-report,” Metabolism, vol. 44, pp. 18–22, Feb. 1995, doi: 10.1016/0026-0495(95)90204-X.

[9] S. Paßler and W.-J. Fischer, “Food Intake Activity Detection Using a Wearable Microphone System,” in 2011 Seventh International Conference on Intelligent Environments, Jul. 2011, pp. 298–301. doi: 10.1109/IE.2011.9.

[10] S. Päßler and W.-J. Fischer, “Food Intake Monitoring: Automated Chew Event Detection in Chewing Sounds,” IEEE Journal of Biomedical and Health Informatics, vol. 18, no. 1, pp. 278–289, Jan. 2014, doi: 10.1109/JBHI.2013.2268663.

[11] E. Sazonov et al., “Non-invasive monitoring of chewing and swallowing for objective quantification of ingestive behavior,” Physiol. Meas., vol. 29, no. 5, pp. 525–541, Apr. 2008, doi: 10.1088/0967-3334/29/5/001.

[12] E. S. Sazonov, O. Makeyev, S. Schuckers, P. Lopez-Meyer, E. L. Melanson, and M. R. Neuman, “Automatic Detection of Swallowing Events by Acoustical Means for Applications of Monitoring of Ingestive Behavior,” IEEE Transactions on Biomedical Engineering, vol. 57, no. 3, pp. 626–633, Mar. 2010, doi: 10.1109/TBME.2009.2033037.

[13] E. S. Sazonov, O. Makeyev, S. Schuckers, P. Lopez-Meyer, E. L. Melanson, and M. R. Neuman, “Automatic Detection of Swallowing Events by Acoustical Means for Applications of Monitoring of Ingestive Behavior,” IEEE Transactions on Biomedical Engineering, vol. 57, no. 3, pp. 626–633, Mar. 2010, doi: 10.1109/TBME.2009.2033037.

[14] P. Lopez-Meyer, O. Makeyev, S. Schuckers, E. L. Melanson, M. R. Neuman, and E. Sazonov, “Detection of Food Intake from Swallowing Sequences by Supervised and Unsupervised Methods,” Ann Biomed Eng, vol. 38, no. 8, pp. 2766–2774, Aug. 2010, doi: 10.1007/s10439-010-0019-1.

[15] O. Amft and G. Troster, “On-Body Sensing Solutions for Automatic Dietary Monitoring,” IEEE Pervasive Computing, vol. 8, no. 2, pp. 62–70, Apr. 2009, doi: 10.1109/MPRV.2009.32.

[16] Y. Dong, J. Scisco, M. Wilson, E. Muth, and A. Hoover, “Detecting Periods of Eating During Free-Living by Tracking Wrist Motion,” IEEE Journal of Biomedical and Health Informatics, vol. 18, no. 4, pp. 1253–1260, Jul. 2014, doi: 10.1109/JBHI.2013.2282471.

[17] J. N. Salley, A. W. Hoover, M. L. Wilson, and E. R. Muth, “Comparison between Human and Bite-Based Methods of Estimating Caloric Intake,” Journal of the Academy of Nutrition and Dietetics, vol. 116, no. 10, pp. 1568–1577, Oct. 2016, doi: 10.1016/j.jand.2016.03.007.

[18] T. Rahman et al., “BodyBeat: A Mobile System for Sensing NonSpeech Body Sounds,” in in MobiSys ’14: Proceedings of the 12th annual international, pp. 2–13.

[19] M. Mirtchouk, C. Merck, and S. Kleinberg, “Automated estimation of food type and amount consumed from body-worn audio and motion sensors,” in Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, in UbiComp ’16. New York, NY, USA: Association for Computing Machinery, Sep. 2016, pp. 451–462. doi: 10.1145/2971648.2971677.

[20] C. K. Martin, S. Kaya, and B. K. Gunturk, “Quantification of food intake using food image analysis,” in 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Sep. 2009, pp. 6869–6872. doi: 10.1109/IEMBS.2009.5333123.

[21] P. Pouladzadeh, S. Shirmohammadi, and R. Al-Maghrabi, “Measuring Calorie and Nutrition From Food Image,” IEEE Transactions on Instrumentation and Measurement, vol. 63, no. 8, pp. 1947–1956, Aug. 2014, doi: 10.1109/TIM.2014.2303533.

[22] F. Zhu, M. Bosch, N. Khanna, C. J. Boushey, and E. J. Delp, “Multiple Hypotheses Image Segmentation and Classification With Application to Dietary Assessment,” IEEE Journal of Biomedical and Health Informatics, vol. 19, no. 1, pp. 377–388, Jan. 2015, doi: 10.1109/JBHI.2014.2304925.

[23] F. Kong and J. Tan, “DietCam: Automatic dietary assessment with mobile camera phones,” Pervasive and Mobile Computing, vol. 8, no. 1, pp. 147–163, Feb. 2012, doi: 10.1016/j.pmcj.2011.07.003.

[24] C. K. Martin et al., “Validity of the Remote Food Photography Method (RFPM) for Estimating Energy and Nutrient Intake in Near Real-Time,” Obesity, vol. 20, no. 4, pp. 891–899, 2012, doi: 10.1038/oby.2011.344.

[25] Y. Deng, S. Qin, and Y. Wu, “An automatic food recognition algorithm with both shape and texture information,” in PIAGENG 2009: Image Processing and Photonics for Agricultural Engineering, SPIE, Jul. 2009, pp. 33–40. doi: 10.1117/12.836650.

[26] A. Meyers et al., “Im2Calories: Towards an Automated Mobile Vision Food Diary,” presented at the Proceedings of the IEEE International Conference on Computer Vision, 2015, pp. 1233–1241. Accessed: Mar. 21, 2022. [Online]. Available: https://openaccess.thecvf.com/content_iccv_2015/html/Meyers_Im2Calories_Towards_an_ICCV_2015_paper.html

[27] H. R. Kissileff, G. Klingsberg, and T. B. Van Itallie, “Universal eating monitor for continuous recording of solid or liquid consumption in man,” American Journal of Physiology-Regulatory, Integrative and Comparative Physiology, vol. 238, no. 1, pp. R14–R22, Jan. 1980, doi: 10.1152/ajpregu.1980.238.1.R14.

[28] K. Chang et al., “The Diet-Aware Dining Table: Observing Dietary Behaviors over a Tabletop Surface,” in Pervasive Computing, K. P. Fishkin, B. Schiele, P. Nixon, and A. Quigley, Eds., Berlin, Heidelberg: Springer, 2006, pp. 366–382. doi: 10.1007/11748625_23.

[29] V. Papapanagiotou, C. Diou, B. Langlet, I. Ioakimidis, and A. Delopoulos, “Automated Extraction of Food Intake Indicators from Continuous Meal Weight Measurements,” in Bioinformatics and Biomedical Engineering, F. Ortuño and I. Rojas, Eds., Cham: Springer International Publishing, 2015, pp. 35–46. doi: 10.1007/978-3-319-16480-9_4.

[30] V. Papapanagiotou, C. Diou, I. Ioakimidis, P. Södersten, and A. Delopoulos, “Automatic Analysis of Food Intake and Meal Microstructure Based on Continuous Weight Measurements,” IEEE Journal of Biomedical and Health Informatics, vol. 23, no. 2, pp. 893–902, Mar. 2019, doi: 10.1109/JBHI.2018.2812243.

[31] R. Paramastri et al., “Use of mobile applications to improve nutrition behaviour: A systematic review,” Computer Methods and Programs in Biomedicine, vol. 192, p. 105459, Aug. 2020, doi: 10.1016/j.cmpb.2020.105459.

[32] S. He et al., “A comprehensive review of the use of sensors for food intake detection,” Sensors and Actuators A: Physical, vol. 315, p. 112318, Nov. 2020, doi: 10.1016/j.sna.2020.112318.

[33] T. Vu, F. Lin, N. Alshurafa, and W. Xu, “Wearable Food Intake Monitoring Technologies: A Comprehensive Review,” Computers, vol. 6, no. 1, Art. no. 1, Mar. 2017, doi: 10.3390/computers6010004.

[34] H. Hassannejad, G. Matrella, P. Ciampolini, I. De Munari, M. Mordonini, and S. Cagnoni, “Automatic diet monitoring: a review of computer vision and wearable sensor-based methods,” International Journal of Food Sciences and Nutrition, vol. 68, no. 6, pp. 656–670, Aug. 2017, doi: 10.1080/09637486.2017.1283683.

[35] B. M. Bell et al., “Automatic, wearable-based, in-field eating detection approaches for public health research: a scoping review,” npj Digit. Med., vol. 3, no. 1, Art. no. 1, Mar. 2020, doi: 10.1038/s41746-020-0246-2.

[36] D. A. Schoeller and M. Westerterp, Advances in the Assessment of Dietary Intake. CRC Press, 2017.

[37] A. Doulah, M. A. Mccrory, J. A. Higgins, and E. Sazonov, “A Systematic Review of Technology-Driven Methodologies for Estimation of Energy Intake,” IEEE Access, vol. 7, pp. 49653–49668, 2019, doi: 10.1109/ACCESS.2019.2910308.

[38] M. C. Archundia Herrera and C. B. Chan, “Narrative Review of New Methods for Assessing Food and Energy Intake,” Nutrients, vol. 10, no. 8, Art. no. 8, Aug. 2018, doi: 10.3390/nu10081064.

[39] V. B. Raju and E. Sazonov, “A Systematic Review of Sensor-Based Methodologies for Food Portion Size Estimation,” IEEE Sensors Journal, vol. 21, no. 11, pp. 12882–12899, Jun. 2021, doi: 10.1109/JSEN.2020.3041023.

[40] D. Moher, A. Liberati, J. Tetzlaff, and D. G. Altman, “Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement,” International Journal of Surgery, vol. 8, no. 5, pp. 336–341, Jan. 2010, doi: 10.1016/j.ijsu.2010.02.007.

[41] D. Weathers, J. C. Siemens, and S. W. Kopp, “Tracking food intake as bites: Effects on cognitive resources, eating enjoyment, and self-control,” Appetite, vol. 111, pp. 23–31, Apr. 2017, doi: 10.1016/j.appet.2016.12.018.

[42] “Association of number of bites and eating speed with energy intake: Wearable technology results under free-living conditions | Elsevier Enhanced Reader.” Accessed: Mar. 08, 2022. [Online]. Available: https://reader.elsevier.com/reader/sd/pii/S0195666321005602?token=F1990B3CB2C90C6B3942B2B0A43C5EA9628FC3A6F4005CCB36877DCAB23BD1DDAE7DB52E4D4F8FB47AD8CE646F8D24F6&originRegion=us-east-1&originCreation=20220308210141

[43] Y. Dong, A. Hoover, and E. Muth, “A Device for Detecting and Counting Bites of Food Taken by a Person during Eating,” in 2009 IEEE International Conference on Bioinformatics and Biomedicine, Nov. 2009, pp. 265–268. doi: 10.1109/BIBM.2009.29.

[44] J. Desendorf, D. R. J. Bassett, H. A. Raynor, and D. P. Coe, “Validity of the Bite Counter device in a controlled laboratory setting.,” Eat Behav, vol. 15, no. 3, pp. 502–504, Aug. 2014, doi: 10.1016/j.eatbeh.2014.06.013.

[45] D. Hossain, T. Ghosh, and E. Sazonov, “Automatic Count of Bites and Chews From Videos of Eating Episodes,” IEEE Access, vol. 8, pp. 101934–101945, 2020, doi: 10.1109/ACCESS.2020.2998716.

[46] J. L. Scisco, E. R. Muth, Y. Dong, and A. W. Hoover, “Slowing bite-rate reduces energy intake: an application of the bite counter device.,” J Am Diet Assoc, vol. 111, no. 8, pp. 1231–1235, Aug. 2011, doi: 10.1016/j.jada.2011.05.005.

[47] E. Almiron-Roig, M. Tsiountsioura, H. B. Lewis, J. Wu, I. Solis-Trapala, and S. A. Jebb, “Large portion sizes increase bite size and eating rate in overweight women,” Physiology & Behavior, vol. 139, pp. 297–302, Feb. 2015, doi: 10.1016/j.physbeh.2014.11.041.

[48] J. A. Beatty, G. W. Greene, B. J. Blissmer, M. J. Delmonico, and K. J. Melanson, “Effects of a novel bites, steps and eating rate-focused weight loss randomised controlled trial intervention on body weight and eating behaviours.,” J Hum Nutr Diet, vol. 33, no. 3, pp. 330–341, Jun. 2020, doi: 10.1111/jhn.12704.

[49] P. W. Jasper, M. T. James, A. W. Hoover, and E. R. Muth, “Effects of Bite Count Feedback from a Wearable Device and Goal Setting on Consumption in Young Adults.,” J Acad Nutr Diet, vol. 116, no. 11, pp. 1785–1793, Nov. 2016, doi: 10.1016/j.jand.2016.05.004.

[50] K. I. Sari and A. Rafisa, “Chewing and Swallowing Patterns for Different Food Textures in Healthy Subjects,” International Journal of Dentistry, vol. 2023, p. e6709350, Jun. 2023, doi: 10.1155/2023/6709350.

[51] B. Nicholls et al., “An EMG-based Eating Behaviour Monitoring system with haptic feedback to promote mindful eating,” Computers in Biology and Medicine, vol. 149, p. 106068, Oct. 2022, doi: 10.1016/j.compbiomed.2022.106068.

[52] M. Farooq and E. Sazonov, “Segmentation and Characterization of Chewing Bouts by Monitoring Temporalis Muscle Using Smart Glasses With Piezoelectric Sensor,” IEEE Journal of Biomedical and Health Informatics, vol. 21, no. 6, pp. 1495–1503, Nov. 2017, doi: 10.1109/JBHI.2016.2640142.

[53] M. Farooq and E. Sazonov, “Linear regression models for chew count estimation from piezoelectric sensor signals,” presented at the Proceedings of the International Conference on Sensing Technology, ICST, 2016. doi: 10.1109/ICSensT.2016.7796222.

[54] M. Farooq and E. Sazonov, “Automatic Measurement of Chew Count and Chewing Rate during Food Intake.,” Electronics (Basel), vol. 5, no. 4, 2016, doi: 10.3390/electronics5040062.

[55] K. Hori et al., “Reliability of a novel wearable device to measure chewing frequency.,” J Prosthodont Res, Jan. 2021, doi: 10.2186/jpr.JPR_D_20_00032.

[56] I. Ioakimidis, M. Zandian, L. Eriksson-Marklund, C. Bergh, A. Grigoriadis, and P. Södersten, “Description of chewing and food intake over the course of a meal,” Physiology & Behavior, vol. 104, no. 5, pp. 761–769, Oct. 2011, doi: 10.1016/j.physbeh.2011.07.021.

[57] H. J. Smit, E. K. Kemsley, H. S. Tapp, and C. J. K. Henry, “Does prolonged chewing reduce food intake? Fletcherism revisited,” Appetite, vol. 57, no. 1, pp. 295–298, Aug. 2011, doi: 10.1016/j.appet.2011.02.003.

[58] A. Okada, M. Honma, S. Nomura, and Y. Yamada, “Oral behavior from food intake until terminal swallow,” Physiology & Behavior, vol. 90, no. 1, pp. 172–179, Jan. 2007, doi: 10.1016/j.physbeh.2006.09.032.

[59] J. Cheng et al., “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, in UbiComp ’13 Adjunct. New York, NY, USA: Association for Computing Machinery, 2013, pp. 155–158. doi: 10.1145/2494091.2494143.

[60] Dzung Tri Nguyen, E. Cohen, M. Pourhomayoun, and N. Alshurafa, “SwallowNet: Recurrent neural network detects and characterizes eating patterns,” in 2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), Mar. 2017, pp. 401–406. doi: 10.1109/PERCOMW.2017.7917596.

[61] O. Makeyev, P. Lopez-Meyer, S. Schuckers, W. Besio, and E. Sazonov, “Automatic food intake detection based on swallowing sounds,” Biomedical Signal Processing and Control, vol. 7, no. 6, pp. 649–656, Nov. 2012, doi: 10.1016/j.bspc.2012.03.005.

[62] E. S. Sazonov, O. Makeyev, S. Schuckers, P. Lopez-Meyer, E. L. Melanson, and M. R. Neuman, “Automatic Detection of Swallowing Events by Acoustical Means for Applications of Monitoring of Ingestive Behavior,” IEEE Transactions on Biomedical Engineering, vol. 57, no. 3, pp. 626–633, Mar. 2010, doi: 10.1109/TBME.2009.2033037.

[63] Y. Kawano and K. Yanai, “Real-Time Mobile Food Recognition System,” in 2013 IEEE Conference on Computer Vision and Pattern Recognition Workshops, Jun. 2013, pp. 1–7. doi: 10.1109/CVPRW.2013.5.

[64] M. Sun et al., “Determination of food portion size by image processing,” in 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Aug. 2008, pp. 871–874. doi: 10.1109/IEMBS.2008.4649292.

[65] J. R. Rajayogi, G. Manjunath, and G. Shobha, “Indian Food Image Classification with Transfer Learning,” in 2019 4th International Conference on Computational Systems and Information Technology for Sustainable Solution (CSITSS), Dec. 2019, pp. 1–4. doi: 10.1109/CSITSS47250.2019.9031051.

[66] U. Tiankaew, P. Chunpongthong, and V. Mettanant, “A Food Photography App with Image Recognition for Thai Food,” in 2018 Seventh ICT International Student Project Conference (ICT-ISPC), Jul. 2018, pp. 1–6. doi: 10.1109/ICT-ISPC.2018.8523925.

[67] J. Li, F. Han, R. Guerrero, and V. Pavlovic, “Picture-to-Amount (PITA): Predicting Relative Ingredient Amounts from Food Images,” in 2020 25th International Conference on Pattern Recognition (ICPR), Jan. 2021, pp. 10343–10350. doi: 10.1109/ICPR48806.2021.9412828.

[68] P. A. Hall, “Executive-Control Processes in High-Calorie Food Consumption,” Curr Dir Psychol Sci, vol. 25, no. 2, pp. 91–98, Apr. 2016, doi: 10.1177/0963721415625049.

[69] R. Hubel, R. G. Laessle, S. Lehrke, and J. Jass, “Laboratory measurement of cumulative food intake in humans: Results on reliability,” Appetite, vol. 46, no. 1, pp. 57–62, Jan. 2006, doi: 10.1016/j.appet.2005.10.006.

[70] M. Shah, J. Copeland, L. Dart, B. Adams-Huet, A. James, and D. Rhea, “Slower Eating Speed Lowers Energy Intake in Normal-Weight but not Overweight/Obese Subjects,” Journal of the Academy of Nutrition and Dietetics, vol. 114, no. 3, pp. 393–402, Mar. 2014, doi: 10.1016/j.jand.2013.11.002.

[71] A. J. Petty, K. J. Melanson, and G. W. Greene, “Self-reported eating rate aligns with laboratory measured eating rate but not with free-living meals,” Appetite, vol. 63, pp. 36–41, Apr. 2013, doi: 10.1016/j.appet.2012.12.014.

[72] A. Doulah et al., “Meal Microstructure Characterization from Sensor-Based Food Intake Detection.,” Front Nutr, vol. 4, p. 31, 2017, doi: 10.3389/fnut.2017.00031.

[73] J. Kim and B.-C. Bae, “A Smartwatch-Based Feedback System for Eating Rate Guidance,” in Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers, in UbiComp ’18. New York, NY, USA: Association for Computing Machinery, 2018, pp. 384–387. doi: 10.1145/3267305.3267669.

[74] K. Taniguchi, H. Chiaki, M. Kurosawa, and A. Nishikawa, “A Novel Earphone Type Sensor for Measuring Mealtime: Consideration of the Method to Distinguish between Running and Meals.,” Sensors (Basel), vol. 17, no. 2, Jan. 2017, doi: 10.3390/s17020252.

[75] R. G. Laessle, S. Lehrke, and S. Dückers, “Laboratory eating behavior in obesity,” Appetite, vol. 49, no. 2, pp. 399–404, Sep. 2007, doi: 10.1016/j.appet.2006.11.010.

[76] J. kim, M. Lee, K.-J. LEE, T. Lee, B.-C. Bae, and J.-D. Cho, “An Eating Speed Guide System Using a Wristband and Tabletop Unit,” in Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct, in UbiComp ’16. New York, NY, USA: Association for Computing Machinery, 2016, pp. 121–124. doi: 10.1145/2968219.2971460.

[77] D. M. Garner and P. E. Garfinkel, “The Eating Attitudes Test: an index of the symptoms of anorexia nervosa,” Psychological Medicine, vol. 9, no. 2, pp. 273–279, May 1979, doi: 10.1017/S0033291700030762.

[78] M. Mirtchouk, C. Merck, and S. Kleinberg, “Automated Estimation of Food Type and Amount Consumed from Body-Worn Audio and Motion Sensors,” in Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, in UbiComp ’16. New York, NY, USA: Association for Computing Machinery, 2016, pp. 451–462. doi: 10.1145/2971648.2971677.

[79] T. Stütz, R. Dinic, M. Domhardt, and S. Ginzinger, “Can mobile augmented reality systems assist in portion estimation? A user study,” in 2014 IEEE International Symposium on Mixed and Augmented Reality - Media, Art, Social Science, Humanities and Design (ISMAR-MASH’D), Sep. 2014, pp. 51–57. doi: 10.1109/ISMAR-AMH.2014.6935438.

[80] M. Pedram et al., “LIDS: Mobile System to Monitor Type and Volume of Liquid Intake,” IEEE Sensors Journal, pp. 1–1, 2021, doi: 10.1109/JSEN.2021.3081012.

[81] Z. Zhang et al., “A Smart Utensil for Detecting Food Pick-up Gesture and Amount While Eating,” in Proceedings of the 11th Augmented Human International Conference, in AH ’20. New York, NY, USA: Association for Computing Machinery, 2020. doi: 10.1145/3396339.3396361.

[82] V. Papapanagiotou, C. Diou, I. Ioakimidis, P. Södersten, and A. Delopoulos, “Automatic Analysis of Food Intake and Meal Microstructure Based on Continuous Weight Measurements,” IEEE Journal of Biomedical and Health Informatics, vol. 23, no. 2, pp. 893–902, Mar. 2019, doi: 10.1109/JBHI.2018.2812243.

[83] J. M. Thomas, C. T. Dourish, and S. Higgs, “Effects of awareness that food intake is being measured by a universal eating monitor on the consumption of a pasta lunch and a cookie snack in healthy female volunteers,” Appetite, vol. 92, pp. 247–251, Sep. 2015, doi: 10.1016/j.appet.2015.05.034.

[84] A. D. M. Sawyer et al., “Dynamics of the complex food environment underlying dietary intake in low-income groups: a systems map of associations extracted from a systematic umbrella literature review,” Int J Behav Nutr Phys Act, vol. 18, no. 1, p. 96, Jul. 2021, doi: 10.1186/s12966-021-01164-1.

[85] L. Gemming, A. Doherty, J. Utter, E. Shields, and C. Ni Mhurchu, “The use of a wearable camera to capture and categorise the environmental and social context of self-identified eating episodes,” Appetite, vol. 92, pp. 118–125, Sep. 2015, doi: 10.1016/j.appet.2015.05.019.

[86] H. K. Ruddock, J. M. Brunstrom, and S. Higgs, “The social facilitation of eating: why does the mere presence of others cause an increase in energy intake?,” Physiology & Behavior, vol. 240, p. 113539, Oct. 2021, doi: 10.1016/j.physbeh.2021.113539.

[87] A. Mekhmoukh, D. Chapelot, and F. Bellisle, “Influence of environmental factors on meal intake in overweight and normal-weight male adolescents. A laboratory study,” Appetite, vol. 59, no. 1, pp. 90–95, Aug. 2012, doi: 10.1016/j.appet.2012.03.021.

[88] P. García-Segovia, R. J. Harrington, and H.-S. Seo, “Influences of table setting and eating location on food acceptance and intake,” Food Quality and Preference, vol. 39, pp. 1–7, Jan. 2015, doi: 10.1016/j.foodqual.2014.06.004.

[89] K. Kyritsis, C. Diou, and A. Delopoulos, “A Data Driven End-to-End Approach for In-the-Wild Monitoring of Eating Behavior Using Smartwatches,” IEEE Journal of Biomedical and Health Informatics, vol. 25, no. 1, pp. 22–34, Jan. 2021, doi: 10.1109/JBHI.2020.2984907.

[90] Y. Shen, J. Salley, E. Muth, and A. Hoover, “Assessing the Accuracy of a Wrist Motion Tracking Method for Counting Bites Across Demographic and Food Variables,” IEEE Journal of Biomedical and Health Informatics, vol. 21, no. 3, pp. 599–606, May 2017, doi: 10.1109/JBHI.2016.2612580.

[91] Y. Y. Luktuke and A. Hoover, “Segmentation and Recognition of Eating Gestures from Wrist Motion using Deep Learning,” in 2020 IEEE International Conference on Big Data (Big Data), Dec. 2020, pp. 1368–1373. doi: 10.1109/BigData50022.2020.9378382.

[92] S. Sen, V. Subbaraju, A. Misra, R. K. Balan, and Y. Lee, “Experiences in Building a Real-World Eating Recogniser,” in Proceedings of the 4th International on Workshop on Physical Analytics, in WPA ’17. New York, NY, USA: Association for Computing Machinery, 2017, pp. 7–12. doi: 10.1145/3092305.3092306.

[93] E. Mendi, O. Ozyavuz, E. Pekesen, and C. Bayrak, “Food intake monitoring system for mobile devices,” in 5th IEEE International Workshop on Advances in Sensors and Interfaces IWASI, Jun. 2013, pp. 31–33. doi: 10.1109/IWASI.2013.6576082.

[94] C. Fortuna, C. Giraud-Carrier, and J. West, “Hand-to-Mouth Motion Tracking in Free-Living Conditions for Improved Weight Control,” in 2016 IEEE International Conference on Healthcare Informatics (ICHI), Oct. 2016, pp. 341–348. doi: 10.1109/ICHI.2016.62.

[95] K. Kyritsis, C. DIou, and A. Delopoulos, “Modeling Wrist Micromovements to Measure In-Meal Eating Behavior from Inertial Sensor Data,” IEEE Journal of Biomedical and Health Informatics, vol. 23, no. 6, pp. 2325–2334, 2019, doi: 10.1109/JBHI.2019.2892011.

[96] C. Li, Y. Bai, W. Jia, and M. Sun, “Eating Event Detection by Magnetic Proximity Sensing,” in 2013 39th Annual Northeast Bioengineering Conference, Apr. 2013, pp. 15–16. doi: 10.1109/NEBEC.2013.85.

[97] S. Zhang et al., “NeckSense: A Multi-Sensor Necklace for Detecting Eating Activities in Free-Living Conditions,” Proc. ACM Interact. Mob. Wearable Ubiquitous Technol., vol. 4, no. 2, Jun. 2020, doi: 10.1145/3397313.

[98] M. F. Kassim, M. N. H. Mohd, M. R. M. Tomari, N. S. Suriani, W. N. W. Zakaria, and S. Sari, “A non-invasive and non-wearable food intake monitoring system based on depth sensor,” Bulletin of Electrical Engineering and Informatics, vol. 9, no. 6, pp. 2342–2349, 2020, doi: 10.11591/eei.v9i6.2256.

[99] D. Konstantinidis, K. Dimitropoulos, B. Langlet, P. Daras, and I. Ioakimidis, “Validation of a Deep Learning System for the Full Automation of Bite and Meal Duration Analysis of Experimental Meal Videos,” Nutrients, vol. 12, no. 1, Art. no. 1, Jan. 2020, doi: 10.3390/nu12010209.

[100] J. Lei, J. Qiu, F. P.-W. Lo, and B. Lo, “Assessing Individual Dietary Intake in Food Sharing Scenarios with Food and Human Pose Detection,” in Pattern Recognition. ICPR International Workshops and Challenges, A. Del Bimbo, R. Cucchiara, S. Sclaroff, G. M. Farinella, T. Mei, M. Bertini, H. J. Escalante, and R. Vezzani, Eds., in Lecture Notes in Computer Science. Cham: Springer International Publishing, 2021, pp. 549–557. doi: 10.1007/978-3-030-68821-9_45.

[101] Z. Tang and A. Hoover, “A New Video Dataset for Recognizing Intake Gestures in a Cafeteria Setting,” in 2022 26th International Conference on Pattern Recognition (ICPR), Aug. 2022, pp. 4399–4405. doi: 10.1109/ICPR56361.2022.9956550.

[102] S. Hermsen, M. Mars, S. Higgs, J. H. Frost, and R. C. J. Hermans, “Effects of eating with an augmented fork with vibrotactile feedback on eating rate and body weight: A randomized controlled trial,” International Journal of Behavioral Nutrition and Physical Activity, vol. 16, no. 1, 2019, doi: 10.1186/s12966-019-0857-7.

[103] G. Mertes, L. Ding, W. Chen, H. Hallez, J. Jia, and B. Vanrumste, “Measuring and Localizing Individual Bites Using a Sensor Augmented Plate During Unrestricted Eating for the Aging Population,” IEEE Journal of Biomedical and Health Informatics, vol. 24, no. 5, pp. 1509–1518, May 2020, doi: 10.1109/JBHI.2019.2932011.

[104] R. Lotfi, G. Tzanetakis, R. Eskicioglu, and P. Irani, “A Comparison between Audio and IMU Data to Detect Chewing Events Based on an Earable Device,” in Proceedings of the 11th Augmented Human International Conference, in AH ’20. New York, NY, USA: Association for Computing Machinery, 2020. doi: 10.1145/3396339.3396362.

[105] J. Liu et al., “An Intelligent Food-Intake Monitoring System Using Wearable Sensors,” in 2012 Ninth International Conference on Wearable and Implantable Body Sensor Networks, May 2012, pp. 154–160. doi: 10.1109/BSN.2012.11.

[106] Y. Bi, M. Lv, C. Song, W. Xu, N. Guan, and W. Yi, “AutoDietary: A Wearable Acoustic Sensor System for Food Intake Recognition in Daily Life,” IEEE Sensors Journal, vol. 16, no. 3, pp. 806–816, Feb. 2016, doi: 10.1109/JSEN.2015.2469095.

[107] A. Nakamura, H. Mineno, M. Nishimura, T. Saito, D. Ikeda, and K. Ohta, “Automatic Detection of the Chewing Side Using Two-channel Recordings under the Ear,” in 2020 IEEE 2nd Global Conference on Life Sciences and Technologies (LifeTech), Mar. 2020, pp. 82–83. doi: 10.1109/LifeTech48969.2020.1570618941.

[108] T. Olubanjo, E. Moore, and M. Ghovanloo, “Detecting food intake acoustic events in noisy recordings using template matching,” in 2016 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI), Feb. 2016, pp. 388–391. doi: 10.1109/BHI.2016.7455916.

[109] M. Mirtchouk and S. Kleinberg, “Detecting Granular Eating Behaviors From Body-worn Audio and Motion Sensors,” in 2021 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI), Jul. 2021, pp. 1–4. doi: 10.1109/BHI50953.2021.9508519.

[110] M. A. Tuğtekin Turan and E. Erzin, “Detection of Food Intake Events From Throat Microphone Recordings Using Convolutional Neural Networks,” in 2018 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), Jul. 2018, pp. 1–6. doi: 10.1109/ICMEW.2018.8551492.

[111] J. Nishimura and Tadahiro Kuroda, “Eating habits monitoring using wireless wearable in-ear microphone,” in 2008 3rd International Symposium on Wireless Pervasive Computing, May 2008, pp. 130–132. doi: 10.1109/ISWPC.2008.4556181.

[112] S. Päßler and W. Fischer, “Food Intake Monitoring: Automated Chew Event Detection in Chewing Sounds,” IEEE Journal of Biomedical and Health Informatics, vol. 18, no. 1, pp. 278–289, Jan. 2014, doi: 10.1109/JBHI.2013.2268663.

[113] Y. Gao et al., “iHear Food: Eating Detection Using Commodity Bluetooth Headsets,” in 2016 IEEE First International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE), Jun. 2016, pp. 163–172. doi: 10.1109/CHASE.2016.14.

[114] S. Bi et al., “Measuring children’s eating behavior with a wearable device,” in 2020 IEEE International Conference on Healthcare Informatics (ICHI), Dec. 2020, pp. 1–11. doi: 10.1109/ICHI48887.2020.9374304.

[115] T. Kondo, H. Shiro, A. Yokokubo, and G. Lopez, “Optimized Classification Model for Efficient Recognition of Meal-Related Activities in Daily Life Meal Environment,” in 2019 Joint 8th International Conference on Informatics, Electronics & Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision & Pattern Recognition (icIVPR), Jun. 2019, pp. 146–151. doi: 10.1109/ICIEV.2019.8858526.

[116] M. Mirtchouk, D. Lustig, A. Smith, I. Ching, M. Zheng, and S. Kleinberg, “Recognizing Eating from Body-Worn Sensors: Combining Free-Living and Laboratory Data,” Proc. ACM Interact. Mob. Wearable Ubiquitous Technol., vol. 1, no. 3, Sep. 2017, doi: 10.1145/3131894.

[117] T. Kondo, H. Kamachi, S. Ishii, A. Yokokubo, and G. Lopez, “Robust Classification of Eating Sound Collected in Natural Meal Environment,” in Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers, in UbiComp/ISWC ’19 Adjunct. New York, NY, USA: Association for Computing Machinery, 2019, pp. 105–108. doi: 10.1145/3341162.3343780.

[118] V. Papapanagiotou, A. Liapi, and A. Delopoulos, “Chewing Detection from Commercial Smart-glasses,” in Proceedings of the 7th International Workshop on Multimedia Assisted Dietary Management on Multimedia Assisted Dietary Management, Lisboa Portugal: ACM, Oct. 2022, pp. 11–16. doi: 10.1145/3552484.3555746.

[119] J. Lee, P. Paudyal, A. Banerjee, and S. K. S. Gupta, “A User-Adaptive Modeling for Eating Action Identification from Wristband Time Series,” ACM Trans. Interact. Intell. Syst., vol. 9, no. 4, Oct. 2019, doi: 10.1145/3300149.

[120] S. Wang, G. Zhou, L. Hu, Z. Chen, and Y. Chen, “CARE: Chewing Activity Recognition Using Noninvasive Single Axis Accelerometer,” in Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers, in UbiComp/ISWC’15 Adjunct. New York, NY, USA: Association for Computing Machinery, 2015, pp. 109–112. doi: 10.1145/2800835.2800884.

[121] Z. Lin, M. Zecca, S. Sessa, H. Ishii, and A. Takanishi, “Development of an ultra-miniaturized inertial measurement unit for jaw movement analysis during free chewing,” Journal of Computer Science, vol. 6, no. 8, pp. 896–903, 2010, doi: 10.3844/jcssp.2010.896.903.

[122] E. S. Sazonov and J. M. Fontana, “A Sensor System for Automatic Detection of Food Intake Through Non-Invasive Monitoring of Chewing,” IEEE Sensors Journal, vol. 12, no. 5, pp. 1340–1348, May 2012, doi: 10.1109/JSEN.2011.2172411.

[123] M. Farooq, J. M. Fontana, A. F. Boateng, M. A. Mccrory, and E. Sazonov, “A Comparative Study of Food Intake Detection Using Artificial Neural Network and Support Vector Machine,” in 2013 12th International Conference on Machine Learning and Applications, Dec. 2013, pp. 153–153. doi: 10.1109/ICMLA.2013.33.

[124] M. Farooq and E. Sazonov, “Comparative testing of piezoelectric and printed strain sensors in characterization of chewing,” in 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Aug. 2015, pp. 7538–7541. doi: 10.1109/EMBC.2015.7320136.

[125] M. Farooq and E. Sazonov, “A Novel Wearable Device for Food Intake and Physical Activity Recognition.,” Sensors (Basel), vol. 16, no. 7, Jul. 2016, doi: 10.3390/s16071067.

[126] J. M. Fontana, M. Farooq, and E. Sazonov, “Automatic Ingestion Monitor: A Novel Wearable Device for Monitoring of Ingestive Behavior,” IEEE Transactions on Biomedical Engineering, vol. 61, no. 6, pp. 1772–1779, Jun. 2014, doi: 10.1109/TBME.2014.2306773.

[127] M. Farooq and E. Sazonov, “Detection of chewing from piezoelectric film sensor signals using ensemble classifiers,” in 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Aug. 2016, pp. 4929–4932. doi: 10.1109/EMBC.2016.7591833.

[128] J. Kim et al., “Slowee: A Smart Eating-Speed Guide System with Light and Vibration Feedback,” in Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems, in CHI EA ’16. New York, NY, USA: Association for Computing Machinery, 2016, pp. 2563–2569. doi: 10.1145/2851581.2892323.

[129] K. S. Chun, S. Bhattacharya, and E. Thomaz, “Detecting Eating Episodes by Tracking Jawbone Movements with a Non-Contact Wearable Sensor,” Proc. ACM Interact. Mob. Wearable Ubiquitous Technol., vol. 2, no. 1, Mar. 2018, doi: 10.1145/3191736.

[130] K. Lee, “Food Intake Detection Using Ultrasonic Doppler Sonar,” IEEE Sensors Journal, vol. 17, no. 18, pp. 6056–6068, Sep. 2017, doi: 10.1109/JSEN.2017.2734688.

[131] N. A. Selamat and S. H. M. Ali, Analysis of Chewing Signals Based on Chewing Detection Using Proximity Sensor for Diet Monitoring, vol. 12665 LNCS. in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 12665 LNCS. 2021, p. 616. doi: 10.1007/978-3-030-68821-9_48.

[132] N. A. Selamat, S. H. Md. Ali, K. N. B. Minhad, S. A. Ahmad, and J. Sampe, “A Novel Peak Detection Algorithm Using Particle Swarm Optimization for Chew Count Estimation of a Contactless Chewing Detection,” IEEE Transactions on Instrumentation and Measurement, vol. 71, pp. 1–12, 2022, doi: 10.1109/TIM.2022.3220281.

[133] A. Rey, R. González, J. L. Martínez-de-Juan, J. Benedito, and A. Mulet, “EMG assessment of chewing behaviour for food evaluation: Influence of personality characteristics,” Food Quality and Preference, vol. 18, no. 3, pp. 585–595, Apr. 2007, doi: 10.1016/j.foodqual.2006.08.002.

[134] A. K. White et al., “A comparison of chewing rate between overweight and normal BMI individuals,” Physiology & Behavior, vol. 145, pp. 8–13, Jun. 2015, doi: 10.1016/j.physbeh.2015.03.028.

[135] R. Zhang, S. Bernhart, and O. Amft, “Diet eyeglasses: Recognising food chewing using EMG and smart eyeglasses,” in 2016 IEEE 13th International Conference on Wearable and Implantable Body Sensor Networks (BSN), Jun. 2016, pp. 7–12. doi: 10.1109/BSN.2016.7516224.

[136] R. Zhang and O. Amft, “Monitoring Chewing and Eating in Free-Living Using Smart Eyeglasses,” IEEE Journal of Biomedical and Health Informatics, vol. 22, no. 1, pp. 23–32, Jan. 2018, doi: 10.1109/JBHI.2017.2698523.

[137] V. Papapanagiotou, C. Diou, L. Zhou, J. van den Boer, M. Mars, and A. Delopoulos, “A Novel Chewing Detection System Based on PPG, Audio, and Accelerometry,” IEEE Journal of Biomedical and Health Informatics, vol. 21, no. 3, pp. 607–618, May 2017, doi: 10.1109/JBHI.2016.2625271.

[138] J. Blechert, M. Liedlgruber, A. Lender, J. Reichenberger, and F. H. Wilhelm, “Unobtrusive electromyography-based eating detection in daily life: A new tool to address underreporting?,” Appetite, vol. 118, pp. 168–173, Nov. 2017, doi: 10.1016/j.appet.2017.08.008.

[139] Y. Sugita, K. Zempo, Y. Ando, Y. Kakutani, K. Mizutani, and N. Wakatsuki, “Diet Gamification toward Chewing Amount Control via Head Mounted Display,” in SIGGRAPH Asia 2018 Posters, in SA ’18. New York, NY, USA: Association for Computing Machinery, 2018. doi: 10.1145/3283289.3283369.

[140] S. Cadavid, M. Abdel-Mottaleb, and A. Helal, “Exploiting visual quasi-periodicity for real-time chewing event detection using active appearance models and support vector machines,” Pers Ubiquit Comput, vol. 16, no. 6, pp. 729–739, Aug. 2012, doi: 10.1007/s00779-011-0425-x.

[141] S. Alshboul and M. Fraiwan, “Determination of Chewing Count from Video Recordings Using Discrete Wavelet Decomposition and Low Pass Filtration,” Sensors, vol. 21, no. 20, Art. no. 20, Jan. 2021, doi: 10.3390/s21206806.

[142] M. P. Lasschuijt et al., “Concept development and use of an automated food intake and eating behavior assessment method,” Journal of Visualized Experiments, vol. 2021, no. 168, 2021, doi: 10.3791/62144.

[143] D. Hossain, M. H. Imtiaz, and E. Sazonov, “Comparison of Wearable Sensors for Estimation of Chewing Strength,” IEEE Sensors Journal, vol. 20, no. 10, pp. 5379–5388, May 2020, doi: 10.1109/JSEN.2020.2968009.

[144] E. Sazonov et al., “Non-invasive monitoring of chewing and swallowing for objective quantification of ingestive behavior,” Physiol. Meas., vol. 29, no. 5, pp. 525–541, Apr. 2008, doi: 10.1088/0967-3334/29/5/001.

[145] J. M. Fontana, P. L. Melo, and E. S. Sazonov, “Swallowing detection by sonic and subsonic frequencies: A comparison,” in 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Sep. 2011, pp. 6890–6893. doi: 10.1109/IEMBS.2011.6091735.

[146] W. P. Walker and D. K. Bhatia, “Automated Ingestion Detection for a Health Monitoring System,” IEEE Journal of Biomedical and Health Informatics, vol. 18, no. 2, pp. 682–692, Mar. 2014, doi: 10.1109/JBHI.2013.2279193.

[147] D. Jayatilake et al., “Smartphone-Based Real-time Assessment of Swallowing Ability From the Swallowing Sound,” IEEE Journal of Translational Engineering in Health and Medicine, vol. 3, pp. 1–10, 2015, doi: 10.1109/JTEHM.2015.2500562.

[148] N. Alshurafa, H. Kalantarian, M. Pourhomayoun, S. Sarin, J. J. Liu, and M. Sarrafzadeh, “Non-invasive monitoring of eating behavior using spectrogram analysis in a wearable necklace,” in 2014 IEEE Healthcare Innovation Conference (HIC), Oct. 2014, pp. 71–74. doi: 10.1109/HIC.2014.7038877.

[149] N. Alshurafa et al., “Recognition of Nutrition Intake Using Time-Frequency Decomposition in a Wearable Necklace Using a Piezoelectric Sensor,” IEEE Sensors Journal, vol. 15, no. 7, pp. 3909–3916, Jul. 2015, doi: 10.1109/JSEN.2015.2402652.

[150] C. Mahesh, P. Kingston Stanley, P. Vijaydaniel, and A. Saneevi Gandhi, “Embedded design consideration for the implementation of food intake monitoring system: Validated using LabVIEW software,” International Journal of Engineering and Technology(UAE), vol. 7, no. 2, pp. 47–51, 2018, doi: 10.14419/ijet.v7i2.8.10322.

[151] G. Hussain, B. Ali Saleh Al-rimy, S. Hussain, A. M. Albarrak, S. N. Qasem, and Z. Ali, "Smart Piezoelectric-Based Wearable System for Calorie Intake Estimation Using Machine Learning," Applied Sciences, vol. 12, no. 12, Art. no. 12, Jan. 2022, doi: 10.3390/app12126135.

[152] B. Dong and S. Biswas, "Wearable diet monitoring through breathing signal analysis," in 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Jul. 2013, pp. 1186–1189. doi: 10.1109/EMBC.2013.6609718.

[153] B. Dong, S. Biswas, R. Gernhardt, and J. Schlemminger, "A mobile food intake monitoring system based on breathing signal analysis," in Proceedings of the 8th International Conference on Body Area Networks, in BodyNets '13. Brussels, BEL: ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), Sep. 2013, pp. 165–168. doi: 10.4108/icst.bodynets.2013.253586.

[154] S. Zhang, D. Nguyen, G. Zhang, R. Xu, N. Maglaveras, and N. Alshurafa, "Estimating Caloric Intake in Bedridden Hospital Patients with Audio and Neck-Worn Sensors," in 2018 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE), Sep. 2018, pp. 1–2. doi: 10.1145/3278576.3278577.

[155] M. I. Khan, B. Acharya, and R. K. Chaurasiya, "iHearken: Chewing sound signal analysis based food intake recognition system using Bi-LSTM softmax network," Computer Methods and Programs in Biomedicine, vol. 221, p. 106843, Jun. 2022, doi: 10.1016/j.cmpb.2022.106843.

[156] N. Yao, R. J. Sclabassi, Q. Liu, and M. Sun, "A video-based algorithm for food intake estimation in the study of obesity," in 2007 IEEE 33rd Annual Northeast Bioengineering Conference, Mar. 2007, pp. 298–299. doi: 10.1109/NEBC.2007.4413396.

[157] Taichi Joutou and Keiji Yanai, "A food image recognition system with Multiple Kernel Learning," in 2009 16th IEEE International Conference on Image Processing (ICIP), Nov. 2009, pp. 285–288. doi: 10.1109/ICIP.2009.5413400.

[158] P. Pouladzadeh, S. Shirmohammadi, and R. Al-Maghrabi, "Measuring Calorie and Nutrition From Food Image," IEEE Transactions on Instrumentation and Measurement, vol. 63, no. 8, pp. 1947–1956, Aug. 2014, doi: 10.1109/TIM.2014.2303533.

[159] P. Pouladzadeh, S. Shirmohammadi, and A. Yassine, "Using graph cut segmentation for food calorie measurement," in 2014 IEEE International Symposium on Medical Measurements and Applications (MeMeA), Jun. 2014, pp. 1–6. doi: 10.1109/MeMeA.2014.6860137.

[160] K. Okamoto and K. Yanai, "Real-time eating action recognition system on a smartphone," in 2014 IEEE International Conference on Multimedia and Expo Workshops (ICMEW), Jul. 2014, pp. 1–6. doi: 10.1109/ICMEW.2014.6890718.

[161] J. Lee, A. Banerjee, and S. K. S. Gupta, "MT-diet demo: Demonstration of automated smartphone based diet assessment system," in 2016 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops), Mar. 2016, pp. 1–3. doi: 10.1109/PERCOMW.2016.7457078.

[162] J. Lee, A. Banerjee, and S. K. S. Gupta, “MT-Diet: Automated smartphone based diet assessment with infrared images,” in 2016 IEEE International Conference on Pervasive Computing and Communications (PerCom), Mar. 2016, pp. 1–6. doi: 10.1109/PERCOM.2016.7456506.

[163] J. Lee, P. Paudyal, A. Banerjee, and S. K. S. Gupta, “FIT-Eve and ADAM: Estimation of velocity and energy for automated diet activity monitoring,” presented at the Proceedings - 16th IEEE International Conference on Machine Learning and Applications, ICMLA 2017, 2017, pp. 1071–1074. doi: 10.1109/ICMLA.2017.000-7.

[164] S. J. Minija and W. R. S. Emmanuel, “Food image classification using sphere shaped — Support vector machine,” in 2017 International Conference on Inventive Computing and Informatics (ICICI), Nov. 2017, pp. 109–113. doi: 10.1109/ICICI.2017.8365314.

[165] A. B. Ocay, J. M. Fernandez, and T. D. Palaoag, “NutriTrack: Android-based food recognition app for nutrition awareness,” in 2017 3rd IEEE International Conference on Computer and Communications (ICCC), Dec. 2017, pp. 2099–2104. doi: 10.1109/CompComm.2017.8322907.

[166] C. Liu et al., “A New Deep Learning-Based Food Recognition System for Dietary Assessment on An Edge Computing Service Infrastructure,” IEEE Transactions on Services Computing, vol. 11, no. 2, pp. 249–261, Apr. 2018, doi: 10.1109/TSC.2017.2662008.

[167] L. Jiang, B. Qiu, X. Liu, C. Huang, and K. Lin, “DeepFood: Food Image Analysis and Dietary Assessment via Deep Model,” IEEE Access, vol. 8, pp. 47477–47489, 2020, doi: 10.1109/ACCESS.2020.2973625.

[168] J. Lester, D. Tan, S. Patel, and A. J. B. Brush, “Automatic classification of daily fluid intake,” in 2010 4th International Conference on Pervasive Computing Technologies for Healthcare, Mar. 2010, pp. 1–8. doi: 10.4108/ICST.PERVASIVEHEALTH2010.8906.

[169] J. Lee, P. Paudyal, A. Banerjee, and S. K. S. Gupta, “IDEA: Instant Detection of Eating Action Using Wrist-Worn Sensors in Absence of User-Specific Model,” in Proceedings of the 26th Conference on User Modeling, Adaptation and Personalization, in UMAP ’18. New York, NY, USA: Association for Computing Machinery, 2018, pp. 371–372. doi: 10.1145/3209219.3209265.

[170] B. Dong and S. Biswas, “Meal-time and duration monitoring using wearable sensors,” Biomedical Signal Processing and Control, vol. 32, pp. 97–109, 2017, doi: 10.1016/j.bspc.2016.09.018.

[171] S. Hermsen, J. H. Frost, E. Robinson, S. Higgs, M. Mars, and R. C. J. Hermans, “Evaluation of a Smart Fork to Decelerate Eating Rate,” Journal of the Academy of Nutrition and Dietetics, vol. 116, no. 7, pp. 1066–1067, 2016, doi: 10.1016/j.jand.2015.11.004.

[172] S. Samadi, K. Turksoy, I. Hajizadeh, J. Feng, M. Sevil, and A. Cinar, “Meal Detection and Carbohydrate Estimation Using Continuous Glucose Sensor Data,” IEEE Journal of Biomedical and Health Informatics, vol. 21, no. 3, pp. 619–627, May 2017, doi: 10.1109/JBHI.2017.2677953.

[173] R. C. J. Hermans, S. Hermsen, E. Robinson, S. Higgs, M. Mars, and J. H. Frost, “The effect of real-time vibrotactile feedback delivered through an augmented fork on eating rate, satiation, and food intake.,” Appetite, vol. 113, pp. 7–13, Jun. 2017, doi: 10.1016/j.appet.2017.02.014.

[174] Z. Zhang, J. Kim, Y. Sakamoto, T. Han, and P. Irani, Applying a pneumatic interface to intervene with rapid eating behaviour, vol. 257. in Studies in Health Technology and Informatics, vol. 257. 2019, p. 519. doi: 10.3233/978-1-61499-951-5-513.

[175] T. Watanabe, T. Shimokakimoto, D. Jayatilake, M. Inoue, and K. Suzuki, “A portable sensor sheet for measuring the eating pace in meal assistance care,” in 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Jul. 2019, pp. 4297–4300. doi: 10.1109/EMBC.2019.8856883.

[176] T. Hamatani, M. Elhamshary, A. Uchiyama, and T. Higashino, “FluidMeter: Gauging the Human Daily Fluid Intake Using Smartwatches,” Proc. ACM Interact. Mob. Wearable Ubiquitous Technol., vol. 2, no. 3, Sep. 2018, doi: 10.1145/3264923.

[177] G. Jung, A. Zarzycki, and J.-H. Lee, “Mom’s Tray: Real-Time Dietary Monitoring System,” in SIGGRAPH Asia 2017 Mobile Graphics & Interactive Applications, in SA ’17. New York, NY, USA: Association for Computing Machinery, 2017. doi: 10.1145/3132787.3139196.

[178] M. Farooq, M. A. McCrory, and E. Sazonov, “Reduction of energy intake using just-in-time feedback from a wearable sensor system.,” Obesity (Silver Spring), vol. 25, no. 4, pp. 676–681, Apr. 2017, doi: 10.1002/oby.21788.

[179] F. S. Konstantakopoulos, E. I. Georga, and D. I. Fotiadis, “A novel approach to estimate the weight of food items based on features extracted from an image using boosting algorithms,” Sci Rep, vol. 13, no. 1, Art. no. 1, Nov. 2023, doi: 10.1038/s41598-023-47885-0.

[180] H. R. Kissileff, G. Klingsberg, and T. B. Van Itallie, “Universal eating monitor for continuous recording of solid or liquid consumption in man.,” Am J Physiol, vol. 238, no. 1, pp. R14-22, Jan. 1980, doi: 10.1152/ajpregu.1980.238.1.R14.

[181] J. N. Salley, A. W. Hoover, and E. R. Muth, “Between- and Within-Subjects Predictors of the Kilocalorie Content of Bites of Food,” Journal of the Academy of Nutrition and Dietetics, vol. 119, no. 7, pp. 1109–1117, Jul. 2019, doi: 10.1016/j.jand.2018.12.009.

[182] R. Naaman et al., “Assessment of Dietary Intake Using Food Photography and Video Recording in Free-Living Young Adults: A Comparative Study,” Journal of the Academy of Nutrition and Dietetics, vol. 121, no. 4, pp. 749-761.e1, Apr. 2021, doi: 10.1016/j.jand.2020.09.040.

[183] T. Maekawa, “A sensor device for automatic food lifelogging that is embedded in home ceiling light: A preliminary investigation,” presented at the Proceedings of the 2013 7th International Conference on Pervasive Computing Technologies for Healthcare and Workshops, PervasiveHealth 2013, 2013, pp. 405–407. doi: 10.4108/icst.pervasivehealth.2013.252128.

[184] C. Bi et al., “FamilyLog: Monitoring Family Mealtime Activities by Mobile Devices,” IEEE Transactions on Mobile Computing, vol. 19, no. 8, pp. 1818–1830, 2020, doi: 10.1109/TMC.2019.2916357.

[185] J. H. Oliver and J. H. Hollis, “Virtual Reality as a Tool to Study the Influence of the Eating Environment on Eating Behavior: A Feasibility Study.,” Foods, vol. 10, no. 1, Jan. 2021, doi: 10.3390/foods10010089.

[186] C. Maramis, V. Kilintzis, and N. Maglaveras, “Real-Time Bite Detection from Smartwatch Orientation Sensor Data,” in Proceedings of the 9th Hellenic Conference on Artificial Intelligence, in SETN ’16. New York, NY, USA: Association for Computing Machinery, 2016. doi: 10.1145/2903220.2903239.

[187] Y. Wang, Z. Li, R. Jarvis, R. A. Khot, and F. “Floyd” Mueller, “IScream!: Towards the Design of Playful Gustosonic Experiences with Ice Cream,” in Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems, in CHI EA ’19. New York, NY, USA: Association for Computing Machinery, 2019, pp. 1–4. doi: 10.1145/3290607.3313244.

[188] D. Konstantinidis, K. Dimitropoulos, I. Ioakimidis, B. Langlet, and P. Daras, “A Deep Network for Automatic Video-Based Food Bite Detection,” in Computer Vision Systems, D. Tzovaras, D. Giakoumis, M. Vincze, and A. Argyros, Eds., in Lecture Notes in Computer Science. Cham: Springer International Publishing, 2019, pp. 586–595. doi: 10.1007/978-3-030-34995-0_53.

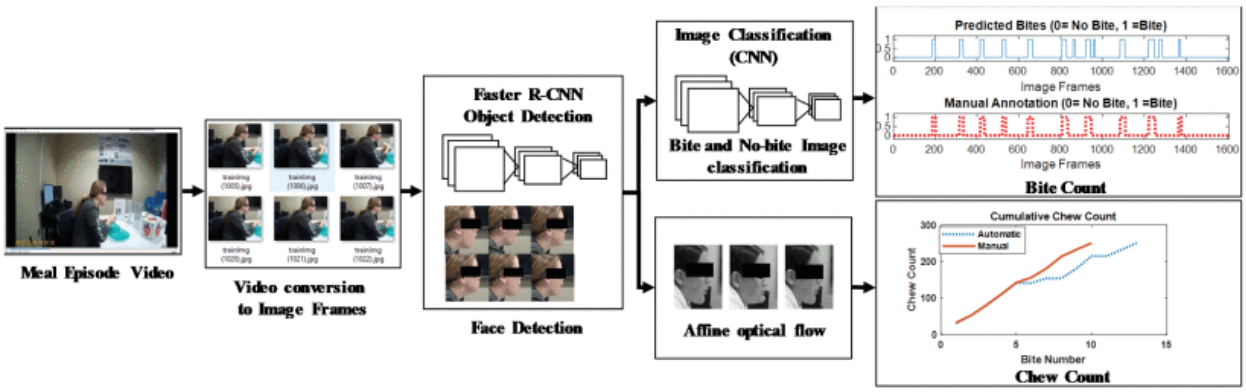
[189] P. Punpongsanon and H. Ishizuka, “Edisensor: Facilitating food electricity for eating habits analysis,” presented at the LifeTech 2021 - 2021 IEEE 3rd Global Conference on Life Sciences and Technologies, 2021, pp. 104–105. doi: 10.1109/LifeTech52111.2021.9391948.

[190] A. Kadomura, C.-Y. Li, K. Tsukada, H.-H. Chu, and I. Siio, “Persuasive Technology to Improve Eating Behavior Using a Sensor-Embedded Fork,” in Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing, in UbiComp ’14. New York, NY, USA: Association for Computing Machinery, 2014, pp. 319–329. doi: 10.1145/2632048.2632093.

CHAPTER 3 – AUTOMATIC COUNT OF BITES AND CHEWS FROM VIDEOS OF EATING EPISODES

Abstract

Methods for measuring eating behavior (known as meal microstructure) often rely on manual annotation of bites, chews, and swallows on meal videos or wearable sensor signals. The manual annotation may be time-consuming and erroneous, while wearable sensors may not capture every aspect of eating (e.g. chews only). This study aims to develop a method to detect and count bites and chews automatically from meal videos. The method was developed on a dataset of 28 volunteers consuming unrestricted meals in the laboratory under video observation. First, the faces in the video (regions of interest, RoI) were detected using Faster R-CNN. Second, a pre-trained AlexNet was trained on the detected faces to classify images as a bite/no bite image. Third, affine optical flow was applied in consecutively detected faces to find the rotational movement of the pixels in the RoIs. The number of chews in a meal video was counted by converting the 2-D images to 1-D optical flow parameter and finding peaks. The developed bite and chew count algorithm was applied on 84 meal videos collected from 28 volunteers. A mean accuracy (±STD) of 82.65% (±8.70%) in respect to manual annotation was obtained for the number of bites and 88.64% (± 5.29%) for the number of chews. The proposed method for automatic bite and chew counting shows promising results which can be used as an alternative solution to manual annotation.



3.1 Introduction

The study of eating behavior of individuals is important for researchers to understand eating patterns of people suffering from obesity and eating disorders. In order to understand the complex process of dietary habits, it is necessary to track the dynamic process of eating episodes which is known as meal microstructure [1]–[3]. Meal microstructure combines factors of food intake behavior such as total eating episode duration (start of the food intake to the end including pauses), true ingestion duration (time spent in chewing), eating event number (a bite which is potentially followed by a segment of chewing and swallowing), ingestion rate, frequency and efficiency of chewing, and size of bite [4]. As the meal microstructure is used to characterize eating behavior of individual, studies of meal microstructure may provide new insights and findings in the treatment of obesity. In [5], authors claimed that a reduction in bite-rate results lower intake. Other studies have shown that there is a clear relationship between food intake rate and energy intake [6]–[9]. In [9], authors suggested that to fight against obesity, interventions to improve chewing activity could be a necessary tool. In [6], authors claimed that eating rate feedback can be helpful to aid on intervention in eating disorder treatment. Health researchers are applying the number of chews and the rate of chewing as variables in their models for estimation of ingested mass and energy intake. Counts of chews and swallows were used to

develop energy intake estimation model (CCS) in [10]. In [11]authors concluded that for solid food, recordings of chewing sound can be used to predict bite weight.

Recently, many health researchers are exploring several wearable devices to monitor food intake behavior and record meal microstructure in eating episodes [12]. Different sensor modalities such as acoustic sensors [13], [14], accelerometers [15], [16], surveillance video cameras [17], smartphone cameras [18], piezoelectric sensors [19], [20], motion sensors [5], hand gesture sensors [21] are being explored in development of these wearable devices. To train and validate sensor-based models, most of the studies of food intake monitoring systems [6]–[8], [10], [20], [22]–[25] apply manual annotation of chew count by either participants or the investigators. Although video observation and manual annotation is an accepted gold standard for these studies, the manual chew count can be inaccurate, time-consuming and a burden to the investigators when the data size is large. In [26], authors used both manual annotation of videos and sensor-driven features to quantify energy intake. They mentioned that manual analysis is not realistic in large scale. Hence an algorithm to make the annotation automatic in feeding studies is necessary.

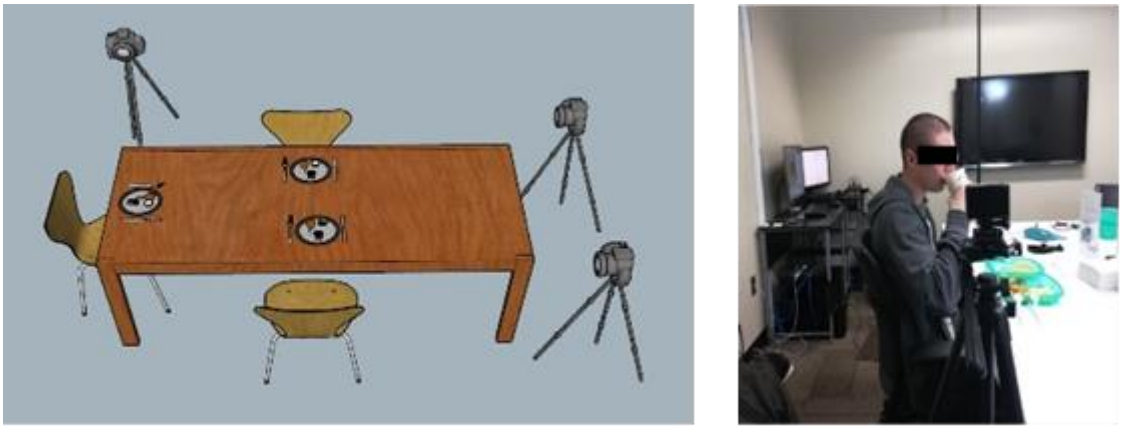
Little work has been done so far to make the annotation from eating episode videos automatic. In [27], authors used face detection and a variable-intensity template modeling to track changes in luminance of the left and right cheeks and chin areas during chewing. They achieved an accuracy of 83% on a small dataset (480 chews) and detected faces only from front on a limited number of foods, whereas chewing depends upon food type. In [17], authors used visual quasi-periodicity in chewing to detect chewing event from videos using support vector machines. This model worked only to detect chewing events (classification) while not counting the number of chews. Recently, in [28] authors developed a deep network using features from

both face and body to detect bite instances and count number of bites, but did not count the number of chews in the eating episodes. They developed a novel “Rapid Automatic Bite Detection” (RABid) [29] algorithm that extracts and process skeletal features to measure meal duration and bites from meal videos using using long short-term memory (LSTM) network. None of the above mentioned methods can be used as a fully automatic alternative to human meal-video annotations for the experimental analysis of human eating behavior, as they don’t calculate bites and chews of the meal together.

In this paper, we propose a novel contactless bite and chew counting method from video based on object detection, image classification, and optical flow. The proposed method combines automatic face detection, and bite and chew counting to provide a fully automatic approach for annotation of food consumption documented on video.

3.2 Methods and Materials

3.2.1 Data Collection Protocol



(a) Data collection setup (b) Real Participant

Figure 3-1: Camera positions and Data collection setup

A total of 28 volunteers were recruited (17 males and 11 females, average age 29.03 ± 12.20 years, range 19-41 years, average body mass index (BMI) 27.87 ± 5.51 kg/m2, range 20.5 to

41.7). Participants were asked to consume three free meals (breakfast, lunch, and dinner) in the laboratory where the eating was recorded on video. Participants were taken to on-campus food courts at the University of Alabama to self-select their meals. There was no restriction on the choice of food items. Several participants (up to 3) were invited to the laboratory at the same time to simulate social eating. This study was approved by The University of Alabama's Institutional Review Board. Recruited participants did not show any medical conditions that would hinder their normal eating or chewing. All participants read and signed an Informed Consent before participating in the study. The meal episodes were recorded using a SJCAM SJ4000 Action Camera (Black)**.** This camera takes 1080p video at 30 frames per second. Figure 3-1 shows the data collection setup in the laboratory. The cameras were positioned 3 feet away from the participants.

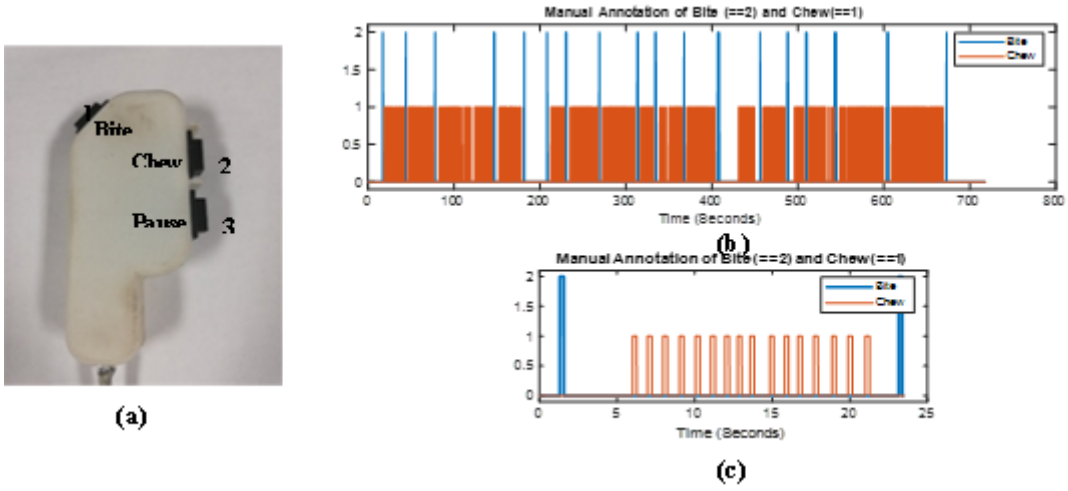


Figure 3-2: Manual Annotation of Bite and Chew from video (a) The three-button system for annotation of the video (b) Annotated bite and chews in a meal video (c) Annotated bite and chew of a chewing sequence.

The cameras were positioned to take video in the profile view to make the tracking of jaw movement easier for manual annotation. The subjects were instructed to eat in a natural way which

can include talking and gestures, and with a self-determined pace. The video dataset contained a total of 19 Hours and 26 Minutes videos of 84 meals from 28 different subjects. Because the highest frequency of chewing is about 2Hz, the recorded videos were down-sampled into 6fps image frames. The dataset contained 419737 image frames, 1288 bites, and 29328 chews that were counted using manual annotation of the video. The manual annotation process was performed using a 3-button system and custom-made LabView software as described in [30]. The three-button system shown in Figure 3-2(a) contains dedicated buttons to mark bite and chew events. The meal video was played on a monitor with a 5x slower speed and bite and chewing events were marked by pressing the button. One complete meal video annotation is shown in Figure 3-2(b) and a short sequence of a bite followed by chewing is shown in Figure 3-2(c). All the computation and processing were performed offline in MATLAB 2018b (Mathworks Inc., Natick, MA, USA).

3.2.2 Face Detection to Extract ROI

As the first step of processing, we extracted the region of interest (RoI) from the image frames of the meal video. As the goal of this project is to automatically count bites and chews, only the face of the participant is the RoI. To detect the face from images, an object detection algorithm based on deep learning was used. In recent years, deep Convolutional Neural Networks (CNNs) have become very popular in computer vision applications. For object detection, region-based CNN detection methods are becoming the primary choice for researchers. To bring more speed and accuracy in object detection, three generations of region-based CNN detection models, from the R-CNN [31], to the Fast R-CNN [32], and finally the Faster R-CNN [33], have been proposed within the space of a few years. In this project, Faster R-CNN was used to detect faces in image frames. Figure 3-3. shows the architecture of the Faster-RCNN object detection algorithm. Typically, any RCNN-based object detection [34] algorithm consists of two stages: 1) a Region Proposal Network (RPN) consists of a fully

convolutional network which is responsible for generation of specific region proposals for each image, and 2) a deep convolution network which is used for pooling of region proposals and object detection. As most of the computation of RPN stage and object detection network stage is shared, a single deep network (e.g. CNN) can be trained to serve in both stages and produce high-quality proposals, pooling of the proposals, and object detection.

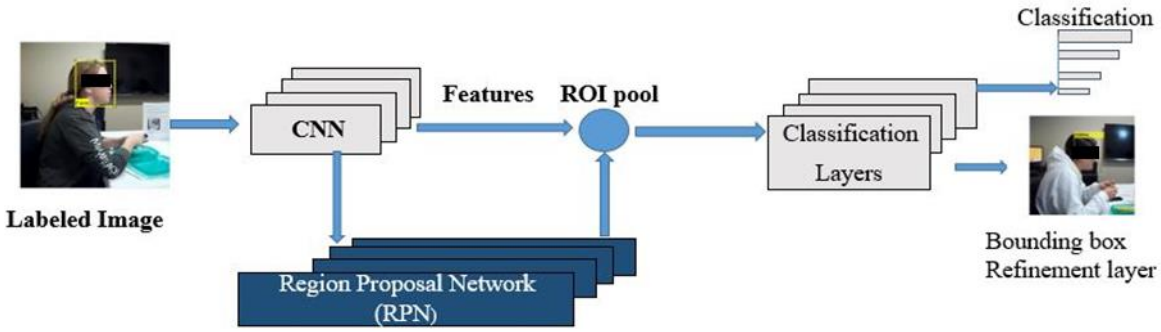


Figure 3-3: Architecture of Faster R-CNN object detector

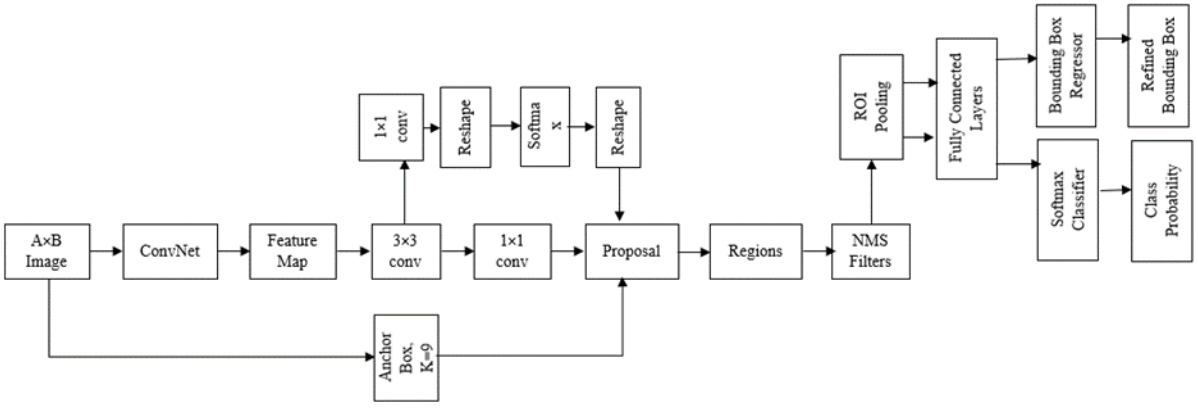


Figure 3-4: Structure and layers of face detector

Figure 3-4 represents the structure and layers of the trained object detection network. We used a transfer learning approach with a pre-trained ResNet-50, a network consisting of 50 convolution layers. To produce a fixed-size feature vector, the final layer of ResNet-50 was

followed by a 3 × 3 convolutional layer [34]. The region proposal network produced 9 proposals with three different aspect ratios (1:1,1:2, 2:1) and three different scales (32, 64, and 128 pixels), at each place which was parameterized relative to an anchor box called a reference box. For classification and regression, the output was then passed to the 1×1 class score layer and regression layer. The regression layer provided a 4×9-dimensional output which included coordinates of the corner for all anchors of the bounding box. The class score layer provided a 2×9-dimensional output which included a score for both the object and background. To predict the class and bounding box refinement, the region proposals generated by RPN and the convolutional features, both were passed into the RoI (region of interest) pooling layer.

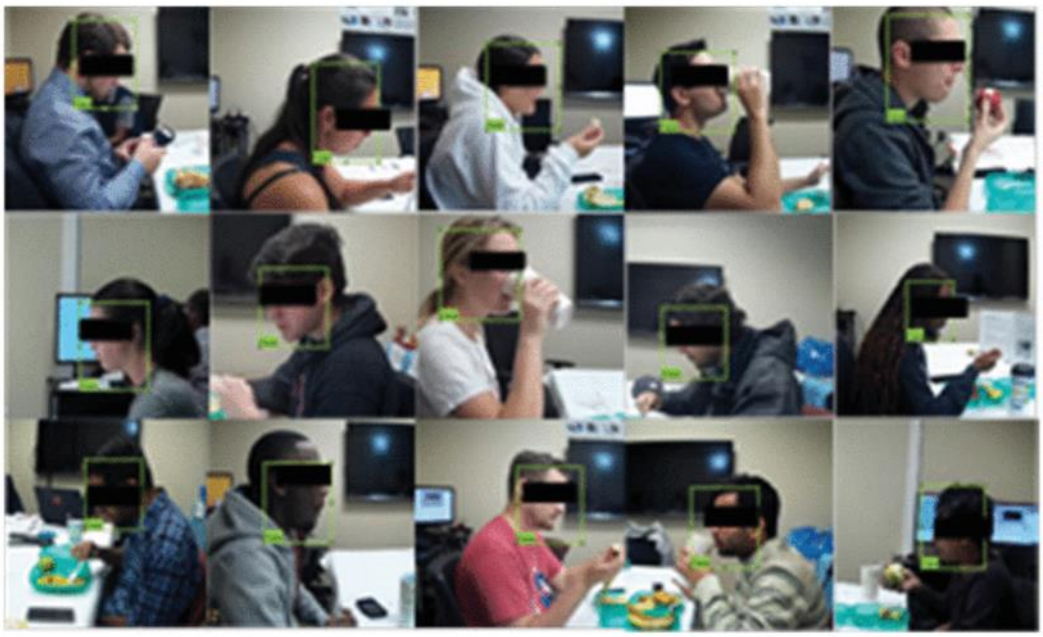


Figure 3-5: Labeled training Image for face detector.

The meal videos were recorded with a resolution of 1920×1080. To remove the surroundings, the image frames were cropped with fixed coordinates to a size of 840×760. To train the object detector, 300 images of each participant were randomly selected, which provided

a total of 8400 training images. To label the training image, ImageLabeler API [35] was used on the image set and a human rater manually marked the boundary (the bounding box) of the face as the object to be identified. Figure 3-4 represents the labeled training images. In the training stage, a 4-fold cross-validation method was implemented. The images of 21 subjects (6300 images) were used for training at once and the rest of the 7 subject's images (2100 images) were used in validation. To increase the size of the dataset, image augmentation, and resizing procedures were implemented on the whole training image set. The image set was augmented by a) a Y-reflection mirror, b) 2-D Gaussian smoothing (blurring), c) brightness adjustment, d) orientation change to portrait mode, and e) crop and rotation by ±10 degrees. The post-augmentation image set contained 42700 training images at each fold. Learning rate, batch size, and the number of epochs were chosen as 0.001, 1, and 4 respectively which were empirically selected from grid search procedures. This training was performed on a MATLAB 2018b environment installed on a Windows 10 computer equipped with an Intel core i7 9th Generation CPU with 16GB DDR4 RAM, and an NVIDIA GeForce GTX 1070 GPU with 8GB memory.

3.2.3 Image Classification To Count Bites

A bite can be defined as the process of taking liquid, semi-solid, or solid food by the person using a straw, glass, bottle, spoon, fork, or hand [36], [37]. We used detected faces from image frames to count the number of bites in an eating episode by detecting faces with straw/glass/bottle/hand/spoon/fork and food in the field of view (defined as 'bite' image). The images containing only faces were considered as 'non-bite' images. Fig 6, shows a sequence of cropped faces of an eating episode with the 'bite' images highlighted by the box. We used a pre-trained AlexNet [38] to train a binary image classifier to classify cropped faces as 'bite' images or 'non-bite' images. The structure and layers of AlexNet are presented in Figure 3-7. The

structure of AlexNet comprises 23 layers which include an input layer, 5 convolution layers, 5 rectified linear units (ReLu), 2 layers for normalization, 3 max-pooling layers, 3 fully connected layers, 2 drop-out layers, a probabilistic layer with softmax units and a classification layer. For bite image classification, the pre-trained AlexNet consisting of 23 layers was used with two additional ReLu layers. The first 22 layers of the pretrained network were transferred and then a fully connected layer, softmax layer, and classification layer were added. As the cropped faces were classified into the bite or non-bite image, thus, the fully connected layer had two outputs. The input layer of the network was fed by a 227×227×3 color image. The first convolution layer consisted of 96 kernels of size 11×11×3 with a stride of 4 pixels which passed the input to the second convolution layer

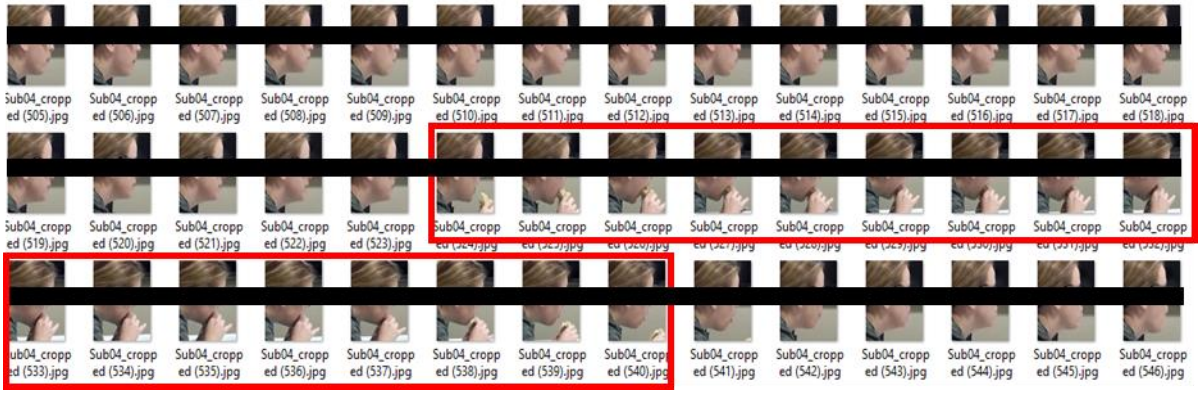


Figure 3-6: Cropped faces of an eating episode to detect a bite.

through normalization and pooling. The second, third, fourth, and fifth convolution layers had 256 kernels of size 5*5*48, 384 kernels of size 3*3*256, 384 kernels of size 3*3*192, and 256 kernels of size 3*3*192 respectively. The fully connected layers had 4096 neurons each. The images with cropped faces resized to 227×227×3 were used to retrain the last three layers of the described network. During training, 4-fold cross-validation scheme was implemented. For example, in the first fold, the first 21 subjects cropped faces from meal videos were used and the

rest of the 7 subjects cropped faces were used in validation. For each participant, 100 images of bite events and 200 images of non-bite events were randomly chosen as training images which resulted in a total of 6300 training images (2100 bite images and 4200 non-bite images) and 2800 validation images (700 bites images and 2100 non-bite images) in a single fold. To reduce overfitting, both image augmentation and dropout schemes were implemented. The image set was augmented by a) Y-reflection mirror, b) 2-D Gaussian smoothing (blurring), c) brightness adjustment, d) orientation change to portrait mode, and e) crop and rotation by ±10 degrees. The two dropout layers drop neurons that have a probability of 0.5 or less. Learning rate, batch size and the number of epochs hyperparameters were chosen as 0.001, 10 and 6 respectively which were empirically selected from grid search procedures. This training was performed on a MATLAB 2018b environment installed on a Windows 10 computer equipped with an Intel core i7 9th Generation CPU with 16GB DDR4 RAM, and an NVIDIA GeForce GTX 1070 GPU with 8GB memory.

3.2.4 Optical Flow to Count Number of Chews

In order to track the movement of jaw while chewing, we used optical flow, a computer vision approach which estimates the apparent motion of surfaces, edges, and objects between images [39]–[42]. The motion between two image frames which are taken at times t and t+∇t can be interpreted by calculating the motion at every pixel location. The images are represented as 3-D vector fields, $I(x, y, t)$ where (x, y) and t denotes the spatial coordinates and time, respectively. For a specific time instance pair of images, the spatial motion, $\boldsymbol{u} = (u_x, u_y)$, is estimated for every pixel.

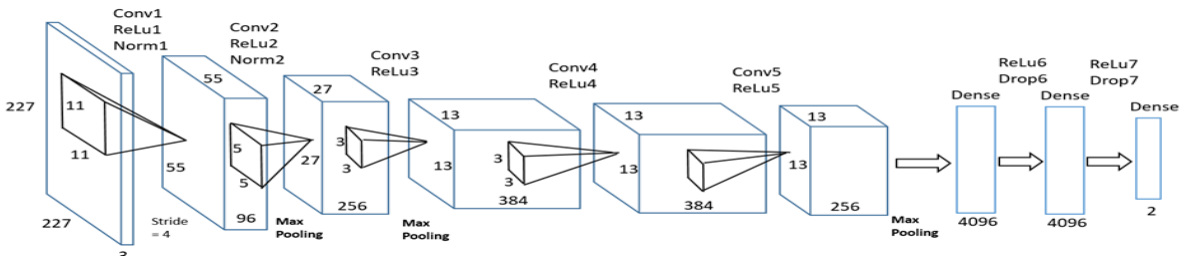


Figure 3-7: Structure and Layers of Image Classifier to classify bite.

The basic assumption behind the idea of optical flow is the brightness constancy assumption, which states that the brightness (intensity) of a pixel remains constant during its motion, and for a specific pixel, it is expressed as

$$I_x u_x + I_y u_y + I_t = 0$$

where I_x, I_y, and I_t denote the partial derivatives of the image function in the x-axis, y-axis, and time, respectively. As the equation has two unknowns (u_x, u_y), it does not have any unique solution which is the mathematical consequence of the aperture problem, which is defined as the lack of information in a small area to determine motion. To solve the problem, additional constraints need to be introduced. Several approaches bring additional constraints to finding the optical flow, such as Lucas-Kanade method [39], Horn-Schunk method [41], Buxton-Buxton [43], and Black-Jepson method [44]. The constraints imposed by Lucas-Kanade method suggest that the displacement of the contents of two nearby image frames is quite small and can be considered constant within a neighborhood of the point p under consideration. Hence the local image flow (velocity) vector ((u_x, u_y) must satisfies for all pixels within a window centered at p. If we use a 5×5 window, that gives us 25 equations per pixel.

Ad= b, where $A_{25\times2}=\begin{bmatrix} I_x(p_1) & I_y(p_1) \\ I_x(p_1) & I_y(p_1) \\ \cdot & \cdot \\ \cdot & \cdot \\ I_x(p_{25}) & I_y(p_{25}) \end{bmatrix}$ $d_{2\times2} = \begin{bmatrix} u \\ v \end{bmatrix}$ and, $b_{25\times1} = -\begin{bmatrix} I_t(p_1) \\ I_t(p_2) \\ \cdot \\ \cdot \\ I_t(p_{25}) \end{bmatrix}$

Now the system has way more equations than unknowns and thus is commonly over-determined. The Lucas-Kanade method obtains a solution by using least squares principle $Ad = b$; minimize $\|ad - b\|^2$. The minimum least squares solution : $(A^TA)d = A^Tb$ which can be written as following

$$\begin{bmatrix} u \\ v \end{bmatrix} = -\begin{bmatrix} \sum I_xI_x & \sum I_xI_y \\ \sum I_xI_y & \sum I_yI_y \end{bmatrix}^{-1}\begin{bmatrix} \sum I_xI_t \\ \sum I_yI_t \end{bmatrix}$$

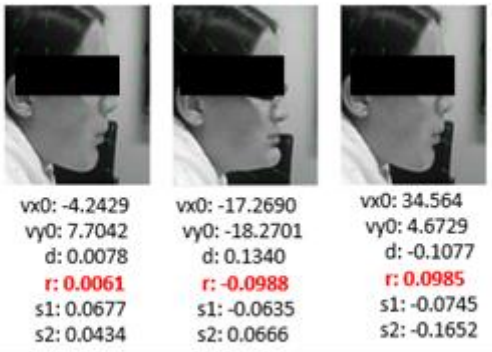


Figure 3-8: Affine optical flow to find chewing. The first and last image show the jaw opening and the middle image shows the jaw closing movements during chewing.

To detect chewing, we used an affine (or first-order) optical flow model which has 6 parameters: two to describe the translational movement in x (*vxo)* and y direction (*vy0*); one to describe rotation (*r*); two to describe angular sheer movement (*s1* and *s2*); and one for directional movement (*d*). A least-squares fit of the parameters, which is an extension of Lucas-Kandle

method [45], was implemented to estimate the spatial and temporal grey-level gradients. In order to apply optical flow, each image frame was passed through the face detector. The detected face was cropped, converted to a grayscale image, and resized to 150×120 pixels. Figure 3-8 shows the effect of optical flow in consecutive images while chewing. Chewing is associated with the opening and closing movement of jaw muscles during mastication[46]. The values of the rotational parameter (*r*) of optical flow are the highest during jaw opening and the lowest during jaw closing. This phenomenon suggests that the rotational parameter of optical flow can be used to explain chewing. The translational movement in the y-direction (*vy0*) also shows similar characteristics. But as the grinding movement of the jaw during chewing can be best described by the rotational movement, we used the rotational parameter.

3.2.5 Bite and Chew Counting

The algorithm for counting number of bites and chews from videos of eating episodes is presented in the Figure 3-9. First, the recorded 30 fps meal video was converted to 6fps image frames. Each frame then passed through the face detector. If no face was detected or the confidence level of detection was below 0.8, then frame was ignored and new frame was taken as input. The detected faces were cropped, resized and then passed through two processing steps 1) bite/non-bite image classification, and 2) calculation of affine optical flow parameter to be used in chew counting. The counting of chews took several considerations into account.

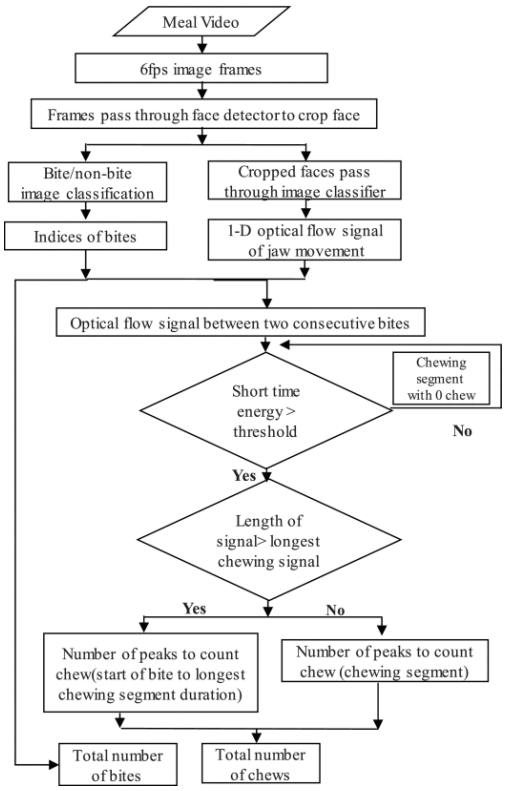


Figure 3-9: Flowchart of developed bite and chew counting algorithm from meal video

A typical sequence of solid food consumption consists of a bite followed by a chewing sequence. Consumption of fluids consists of a bite (placing liquid in mouth) that is not followed by chewing. Also, gestures such as use of napkin to wipe mouth may be detected as a bite but followed by no chewing. Figure 3-10 shows the bite recognition on an eating episode. The top figure shows the raw prediction from cropped faces (bite images labeled as '1' and non-bite images labeled as '0'). In order to remove false detection, the classification labels were smoothed by a moving average filter with span of 10 images. The smoothed classification results were binarized by a threshold with a value of 0.5. This false detection may cause overprediction in chew counting. Thus, chewing segments were identified by calculating the short-time energy of the signal between consecutive bites and estimating a dynamic threshold (T) as described in [47].

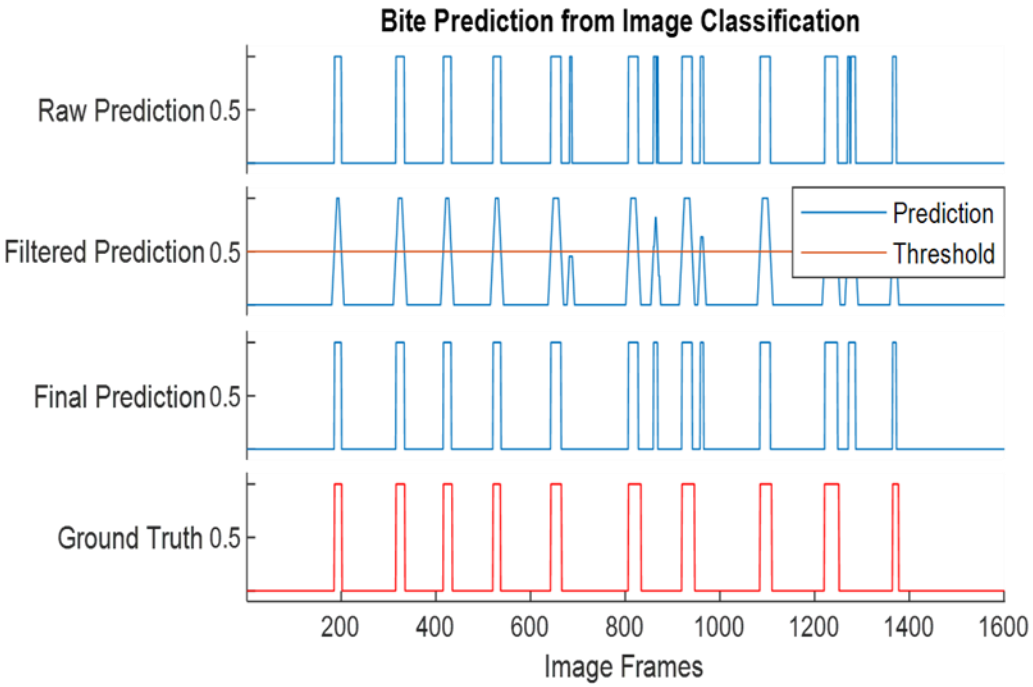


Figure 3-10: Prediction of bite count from image classification. (Top figure represents raw prediction, middle two figures represent filtering and thresholding to remove false detection, the bottom figure represents the ground truth obtained from manual annotation).

If the short-time energy of the optical flow signal between consecutive bites was below the threshold, then that segment was not used in counting the chews. Another consideration is that food consumption is not a continuous process. People often take rest after taking a bite and chewing the food. The movement of the head during these rest periods can also produce a false chew count. From the experimental data, we determined the that longest observed chewing sequence was 52 seconds in duration.

Thus, at the most 52 seconds of optical flow signal was used in the calculation of chew counting. The chew counts for all identified chewing segments were computed by finding peaks and counting them. A peak was defined as the highest point around which there are points lower by X (threshold=0.5) on both sides. The number of peaks in the chewing segment represents the number of chews. Figure 3-11 Shows the optical flow parameter values corresponding to bites

and chews. The figure shows a bite manifesting as a large spike in the value of optical flow, followed by a chewing sequence that produces oscillating values of optical flow.

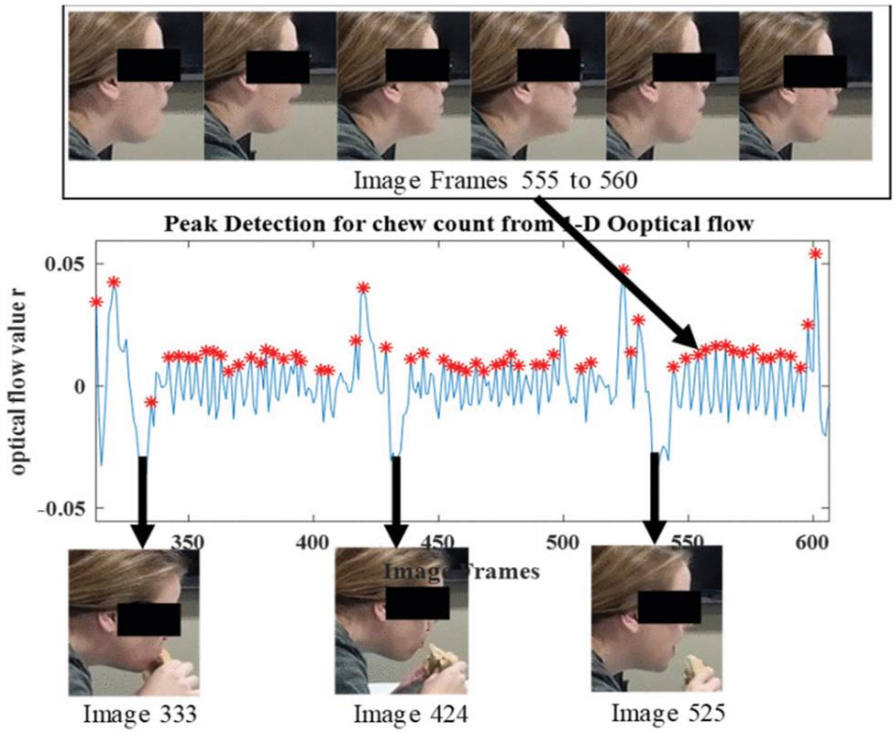


Figure 3-11: Effect of bite and chew in optical flow values.

3.3 Results

The performance evaluation of the face detector utilized IoU (Intersection over Union) and mAP (mean Average Precision) metrics. IoU was calculated from the area of overlap between the predicted bounding box and ground-truth bounding box, and the area encompassed by both the predicted bounding box and the ground-truth bounding box using the following formula.

$$IoU = \frac{Area\ of\ Overlap}{Area\ of\ Union}$$

mAP is the average AP (Average Precision) calculated for each class where AP is the area under the precision-recall curve.

$$AP = \int_0^1 p(r)dr$$

$mAP = \frac{1}{N}\sum_{i=1}^{N} AP_i$ The trained face detector achieved an average IoU of 0.97±0.01 and an average mAP of 0.91±0.02 in 4-fold validation.

The classifier achieved an average F1-score (±standard deviation) of 98.5% ± 1.5%, average precision of 98.5% ± 1.5%, and sensitivity 99.1% ± 0.9% in training. Using the image classifier, we achieved an accuracy of 82.65%±8.70% in bite counts. The chew counting algorithm achieved an accuracy of 88.64% ± 5.29% in 84 meal videos. Chews were counted by detecting peaks in the chewing segment of affine optical flow rotational values.

3.4 Discussion

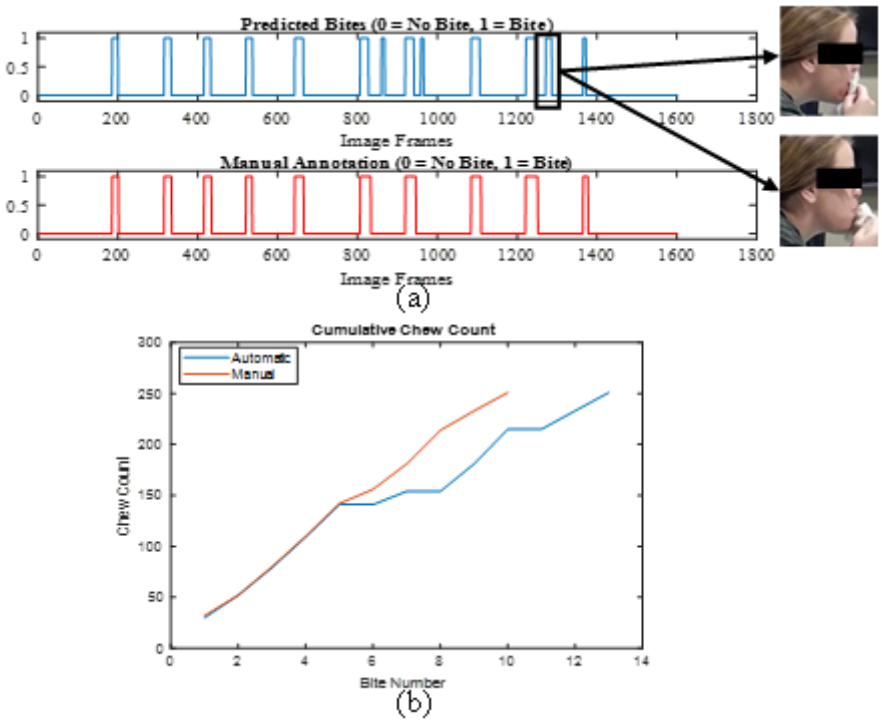


Figure 3-12: (a) False bite event detection because of gestures mimicking bite
(b)Cumulative chew count from manual annotation and automatic count

In this manuscript, we proposed a method for automatic measurement of meal microstructure (bites and chews) from the video of eating episodes. The proposed method detects faces, identifies face images with a bite taking place, identifies chewing segments between consecutive bites, and counts chews from affine optical flow parameters. A bite is a phenomenon of placing the food (liquid/solid) in the mouth. During chewing the jaw makes a grinding motion which can be inferred from the movement of pixels of the face in consecutive image frames. In finding both bite and chew in images, the face is our region of interest (ROI). To extract the ROI from image frames of meal videos, we applied Faster-RCNN object detection in each image frame. To reduce false detection, the threshold for confidence level was set as 0.80. The face detector was able to correctly identify the face in 385892 images among 388074 image frames. In cases, when participants move away from the camera's field of view or blurriness in the images caused by the movement of participants then, the face detector fails to detect the face. The number of bites in the meal was identified from the detected faces using the trained image classifier. Using the image classifier, we achieved an accuracy of 85.4%±6.2% in counting bites for 84 meal videos of 28 subjects. The total number of predicted bites was 1903, whereas the total number of manually annotated bites was 1887 in the dataset. The mean number of bites in an eating episode from manual annotation was 22.5±10.3 (manual annotation) bites and 22.7±11.1 (image detection) bites. Gestures that involve bringing hand toward the mouth such as the use of a napkin to wipe the mouth during eating contribute in the false prediction of bite count. Figure. 3-12(a) shows the false bite detection because of gestures mimicking the bite. The top graph shows the automatic bite count from image classification and the bottom graph represents the manual annotation of bite/non-bite images in the eating episode.

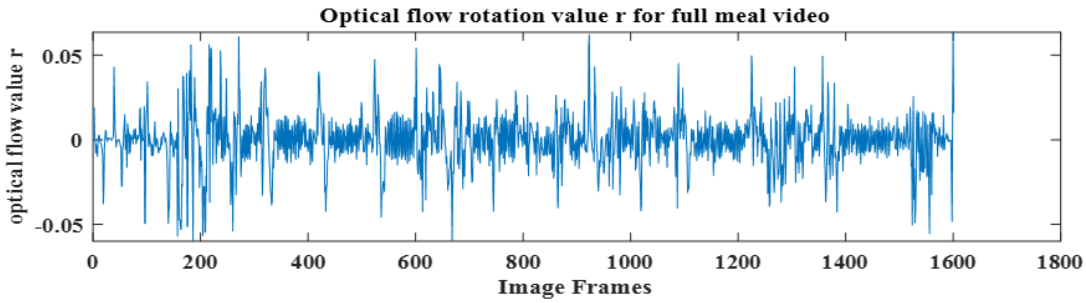


Figure 3-13: 1-D optical flow values from cropped images of an eating episode

The movement of the jaw during chewing can be inferred from the movement of pixels in the face in consecutive image frames. To find the movement of pixels in consecutive cropped faces, the affine optical flow was calculated, and the rotational parameter of pixels' movement was used to describe the chewing. Figure 3-13 shows the calculated rotational parameter of affine optical flow in an eating episode. To remove the effect of false prediction of bites from gestures mimicking a bite and drinking which are not followed by chewing, short-time energy of the optical flow signal between consecutive bites was calculated and thresholded to identify the chewing segment. Talking and movement of the head during these non-chewing segments often contribute enough in short time energy to cross the threshold and hence cause false chewing segment detection.

The developed chew counting algorithm achieved an accuracy of 88.9% ±7.4% in counting the number of chews in 84 meal videos from 28 subjects. The total number of manually counted chews is 40057 and the total number of predicted chews from the proposed method is 41224. The average number of chews in an eating episode as per manual annotation is 476.9±228.7. The mean number of predicted chews by the automatic method in an eating episode is 490.8±248.4. False detection of bite and chewing segment because of gestures, contribute to overprediction of chew count. Also, several participants (up to 3) were invited to

take a meal at a time and allowed to talk while eating, which contributed to false chewing detection and overcounting of chews. Fig. 14(b) represents the cumulative chew count of an eating episode for both manual and automatic annotation.

The proposed method of counting bites and chews achieved an accuracy of 85.4%±6.2% in counting bites and 88.9% ±7.4% in counting chews in 84 videos of eating episodes. These results show promising encouragement in further enrichment of the proposed method to apply as an alternative to manual annotation. The manual annotation of meal microstructures such as bites and chews from video consumes an immense amount of time for researchers in large feeding studies. In this study, videos were played 5X slower than the original speed while annotating bites and chews. An eating episode spanning 10 min and 6 seconds took around 50 min and 30 seconds to annotate whereas the proposed method took around 22 minutes and 55 minutes to convert the videos into image frames, pass the image frames through the face detector, save the cropped faces, apply image classification to count the number of bites and optical flow to count the number of chews in MATLAB 2018b environment installed on a Windows 10 computer equipped with an Intel Core i7 9th Generation CPU with 16GB DDR4 RAM, and an NVIDIA GeForce GTX 1070 GPU with 8GB memory. Note that the time required in the manual annotation is spent by the human annotator but a single click on the mouse is all required by the human in the computer-based annotation.

Thus, the proposed automatic method of bite and chew counting will save both effort and time and increase the reliability considerably than manual annotation of feeding study. The video of eating episodes was recorded from the side of the participant to keep the movement of the jaw during chewing most visible. Future work should concentrate on developing a more robust algorithm that will be able to count the number of bites and chews in an eating episode

from the video of any angle. In this study, the calculation of bites and chews is done offline. Future works may focus on implementing the process in real-time, providing meal microstructure information (number of bites, number of chews, duration of chewing segments) immediately after the meal.

These results show promising encouragement in further enrichment of the proposed method to apply as an alternative to manual annotation.

3.5 Conclusion

We present a novel computer vision-based method of finding total number of bites and chews in an eating episode from the recorded video. To the best of our knowledge, the developed method is the first computer vision-based method that is applied to count both the number of bites and chews in the same eating episode. The contribution of this work is threefold. (I) Development of a face detector based on a Faster-RCNN object detection scheme to detect faces (regions of interest, ROI) in recorded meal videos. (II) Development of a simple image classifier based on pre-trained AlexNet, to classify the detected faces as 'bite' images and 'non-bite' images and hence calculate the total number of bites in the eating episode. (III) To count the number of chews in an eating episode by applying optical flow in consecutive cropped faces to find the movement of the jaw. The developed method was applied to a very large dataset that contains 19 hours and 26 minutes of recorded meal videos of 84 eating episodes from 28 subjects. A mean accuracy of 88.9% ±7.4% to manual annotation in counting the number of bites and a mean accuracy of 88.64 ±5.29% for the number of chews in the 84 meal videos shows promising results which pave the way for further exploration to make the annotation of bites and chews in eating episodes fully automatic, accurate, less time consuming and easy to use.

References

[1] A. Doulah *et al.*, "Meal Microstructure Characterization from Sensor-Based Food Intake Detection," *Front. Nutr.*, vol. 4, 2017, doi: 10.3389/fnut.2017.00031.

[2] F. Bellisle and J. Le Magnen, "The structure of meals in humans: Eating and drinking patterns in lean and obese subjects," *Physiology & Behavior*, vol. 27, no. 4, pp. 649–658, Oct. 1981, doi: 10.1016/0031-9384(81)90237-7.

[3] F. Bellisle, F. Lucas, R. Amrani, and J. Le Magnen, "Deprivation, paiatability and the micro-structure of meals in human subjects," *Appetite*, vol. 5, no. 2, pp. 85–94, Jun. 1984, doi: 10.1016/S0195-6663(84)80027-6.

[4] J. L. Guss and H. R. Kissileff, "Microstructural analyses of human ingestive patterns: from description to mechanistic hypotheses," *Neuroscience & Biobehavioral Reviews*, vol. 24, no. 2, pp. 261–268, Mar. 2000, doi: 10.1016/S0149-7634(99)00079-2.

[5] J. L. Scisco, E. R. Muth, Y. Dong, and A. W. Hoover, "Slowing Bite-Rate Reduces Energy Intake: An Application of the Bite Counter Device," *Journal of the American Dietetic Association*, vol. 111, no. 8, pp. 1231–1235, Aug. 2011, doi: 10.1016/j.jada.2011.05.005.

[6] M. Zandian, I. Ioakimidis, C. Bergh, U. Brodin, and P. Södersten, "Decelerated and linear eaters: Effect of eating rate on food intake and satiety," *Physiology & Behavior*, vol. 96, no. 2, pp. 270–275, Feb. 2009, doi: 10.1016/j.physbeh.2008.10.011.

[7] C. Lepley, G. Throckmorton, S. Parker, and P. H. Buschang, "Masticatory Performance and Chewing Cycle Kinematics—Are They Related?," *The Angle Orthodontist*, vol. 80, no. 2, pp. 295–301, Nov. 2009, doi: 10.2319/061109-333.1.

[8] Y. Zhu and J. H. Hollis, "Increasing the Number of Chews before Swallowing Reduces Meal Size in Normal-Weight, Overweight, and Obese Adults," *Journal of the Academy of Nutrition and Dietetics*, vol. 114, no. 6, pp. 926–931, Jun. 2014, doi: 10.1016/j.jand.2013.08.020.

[9] J. Li *et al.*, "Improvement in chewing activity reduces energy intake in one meal and modulates plasma gut hormone concentrations in obese and lean young Chinese men–," *Am J Clin Nutr*, vol. 94, no. 3, pp. 709–716, Sep. 2011, doi: 10.3945/ajcn.111.015164.

[10] J. M. Fontana *et al.*, "Energy intake estimation from counts of chews and swallows," *Appetite*, vol. 85, pp. 14–21, Feb. 2015, doi: 10.1016/j.appet.2014.11.003.

[11] O. Amft*, M. Kusserow, and G. TrÖster, "Bite Weight Prediction From Acoustic Recognition of Chewing," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 6, pp. 1663–1672, Jun. 2009, doi: 10.1109/TBME.2009.2015873.

[12] D. Hossain, M. H. Imtiaz, and E. Sazonov, “Comparison of wearable sensors for estimation of chewing strength,” *IEEE Sensors Journal*, pp. 1–1, 2020, doi: 10.1109/JSEN.2020.2968009.

[13] S. P. sler, M. Wolff, and W.-J. Fischer, “Food intake monitoring: an acoustical approach to automated food intake activity detection and classification of consumed food,” *Physiol. Meas.*, vol. 33, no. 6, pp. 1073–1093, May 2012, doi: 10.1088/0967-3334/33/6/1073.

[14] O. Amft*, M. Kusserow, and G. TrÖster, “Bite Weight Prediction From Acoustic Recognition of Chewing,” *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 6, pp. 1663–1672, Jun. 2009, doi: 10.1109/TBME.2009.2015873.

[15] S. Wang, G. Zhou, L. Hu, Z. Chen, and Y. Chen, “CARE: Chewing Activity Recognition Using Noninvasive Single Axis Accelerometer,” in *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers*, New York, NY, USA, 2015, pp. 109–112, doi: 10.1145/2800835.2800884.

[16] M. Farooq and E. Sazonov, “Accelerometer-Based Detection of Food Intake in Free-Living Individuals,” *IEEE Sensors Journal*, vol. 18, no. 9, pp. 3752–3758, May 2018, doi: 10.1109/JSEN.2018.2813996.

[17] S. Cadavid, M. Abdel-Mottaleb, and A. Helal, “Exploiting Visual Quasi-periodicity for Real-time Chewing Event Detection Using Active Appearance Models and Support Vector Machines,” *Personal Ubiquitous Comput.*, vol. 16, no. 6, pp. 729–739, Aug. 2012, doi: 10.1007/s00779-011-0425-x.

[18] F. Kong and J. Tan, “DietCam: Automatic dietary assessment with mobile camera phones,” *Pervasive and Mobile Computing*, vol. 8, no. 1, pp. 147–163, Feb. 2012, doi: 10.1016/j.pmcj.2011.07.003.

[19] H. Kalantarian, N. Alshurafa, T. Le, and M. Sarrafzadeh, “Monitoring eating habits using a piezoelectric sensor-based necklace,” *Computers in Biology and Medicine*, vol. 58, pp. 46–55, Mar. 2015, doi: 10.1016/j.compbiomed.2015.01.005.

[20] M. Farooq, J. M. Fontana, A. F. Boateng, M. A. Mccrory, and E. Sazonov, “A Comparative Study of Food Intake Detection Using Artificial Neural Network and Support Vector Machine,” in *2013 12th International Conference on Machine Learning and Applications*, Miami, FL, USA, 2013, pp. 153–153, doi: 10.1109/ICMLA.2013.33.

[21] Y. Dong, A. Hoover, J. Scisco, and E. Muth, “A New Method for Measuring Meal Intake in Humans via Automated Wrist Motion Tracking,” *Appl Psychophysiol Biofeedback*, vol. 37, no. 3, pp. 205–215, Sep. 2012, doi: 10.1007/s10484-012-9194-1.

[22] J. M. Fontana, M. Farooq, and E. Sazonov, “Automatic Ingestion Monitor: A Novel Wearable Device for Monitoring of Ingestive Behavior,” *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 6, pp. 1772–1779, Jun. 2014, doi: 10.1109/TBME.2014.2306773.

[23] M. Farooq and E. Sazonov, “Segmentation and Characterization of Chewing Bouts by Monitoring Temporalis Muscle Using Smart Glasses With Piezoelectric Sensor,” *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 6, pp. 1495–1503, Nov. 2017, doi: 10.1109/JBHI.2016.2640142.

[24] M. Farooq and E. Sazonov, “Automatic Measurement of Chew Count and Chewing Rate during Food Intake,” *Electronics*, vol. 5, no. 4, p. 62, Dec. 2016, doi: 10.3390/electronics5040062.

[25] M. Farooq, P. Chandler-Laney, M. Hernandez-Reif, and E. Sazonov, “A Wireless Sensor System for Quantification of Infant Feeding Behavior,” in *Proceedings of the Conference on Wireless Health*, New York, NY, USA, 2015, pp. 16:1–16:5, doi: 10.1145/2811780.2811934.

[26] X. Yang *et al.*, “Statistical models for meal-level estimation of mass and energy intake using features derived from video observation and a chewing sensor,” *Scientific Reports*, vol. 9, no. 1, p. 45, Jan. 2019, doi: 10.1038/s41598-018-37161-x.

[27] A. Fujimoto, T. Ohkawauchi, J. Yamato, and J. Ohya, “An Image Processing Based Method for Chewing Detection Using Variable-intensity Template,” *Electronic Imaging*, vol. 2018, no. 9, pp. 237-1-237–6, Jan. 2018, doi: 10.2352/ISSN.2470-1173.2018.09.IRIACV-237.

[28] D. Konstantinidis, K. Dimitropoulos, I. Ioakimidis, and B. Lang, “A Deep Network for Automatic Video-based Food Bite Detection,” p. 11.

[29] D. Konstantinidis, K. Dimitropoulos, B. Langlet, P. Daras, and I. Ioakimidis, “Validation of a Deep Learning System for the Full Automation of Bite and Meal Duration Analysis of Experimental Meal Videos,” *Nutrients*, vol. 12, no. 1, p. 209, Jan. 2020, doi: 10.3390/nu12010209.

[30] M. Farooq, A. Doulah, J. Parton, M. A. McCrory, J. A. Higgins, and E. Sazonov, “Validation of Sensor-Based Food Intake Detection by Multicamera Video Observation in an Unconstrained Environment,” *Nutrients*, vol. 11, no. 3, p. 609, Mar. 2019, doi: 10.3390/nu11030609.

[31] R. Girshick, J. Donahue, T. Darrell, and J. Malik, “Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation,” in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH, USA, 2014, pp. 580–587, doi: 10.1109/CVPR.2014.81.

[32] “Girshick_Fast_R-CNN_ICCV_2015_paper.pdf.” .

[33] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks,” *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017, doi: 10.1109/TPAMI.2016.2577031.

[34] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks,” in *Advances in Neural Information Processing Systems 28*, C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett, Eds. Curran Associates, Inc., 2015, pp. 91–99.

[35] J. Budiarto, J. Qudsi, and M. Tajuddin, “Counting vehicle using javascript framework library with internet of things implementation concept to lighten up object detection,” *J. Phys.: Conf. Ser.*, vol. 1211, p. 012028, Apr. 2019, doi: 10.1088/1742-6596/1211/1/012028.

[36] N. Zijlstra, R. de Wijk, M. Mars, A. Stafleu, and C. de Graaf, “Effect of bite size and oral processing time of a semisolid food on satiation,” *The American Journal of Clinical Nutrition*, vol. 90, no. 2, pp. 269–275, Aug. 2009, doi: 10.3945/ajcn.2009.27694.

[37] R. A. de Wijk, N. Zijlstra, M. Mars, C. de Graaf, and J. F. Prinz, “The effects of food viscosity on bite size, bite effort and food intake,” *Physiology & Behavior*, vol. 95, no. 3, pp. 527–532, Oct. 2008, doi: 10.1016/j.physbeh.2008.07.026.

[38] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” in *Advances in Neural Information Processing Systems 25*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2012, pp. 1097–1105.

[39] J.-Y. Bouguet, “Pyramidal implementation of the lucas kanade feature tracker,” 2000.

[40] B. K. P. Horn and B. G. Schunck, “Determining optical flow,” *Artificial Intelligence*, vol. 17, no. 1, pp. 185–203, Aug. 1981, doi: 10.1016/0004-3702(81)90024-2.

[41] J. L. Barron, D. J. Fleet, and S. S. Beauchemin, “Performance of optical flow techniques,” *Int J Comput Vision*, vol. 12, no. 1, pp. 43–77, Feb. 1994, doi: 10.1007/BF01420984.

[42] S. S. Beauchemin and J. L. Barron, “The Computation of Optical Flow,” *ACM Comput. Surv.*, vol. 27, no. 3, pp. 433–466, Sep. 1995, doi: 10.1145/212094.212141.

[43] B. Buxton and H. Buxton, “Computation of optic flow from the motion of edge features in image sequences,” *Image and Vision Computing*, vol. 2, no. 2, pp. 59–75, May 1984, doi: 10.1016/0262-8856(84)90001-5.

[44] S. S. Beauchemin and J. L. Barron, “The computation of optical flow,” *ACM Comput. Surv.*, vol. 27, no. 3, pp. 433–466, Sep. 1995, doi: 10.1145/212094.212141.

[45] S. Baker and I. Matthews, “Lucas-Kanade 20 Years On: A Unifying Framework,” *International Journal of Computer Vision*, vol. 56, no. 3, pp. 221–255, Feb. 2004, doi: 10.1023/B:VISI.0000011205.11775.fd.

[46] “Jaw Movement - an overview | ScienceDirect Topics.” [Online]. Available: https://www.sciencedirect.com/topics/medicine-and-dentistry/jaw-movement. [Accessed: 08-Oct-2019].

[47] M. Farooq and E. Sazonov, “Segmentation and Characterization of Chewing Bouts by Monitoring Temporalis Muscle Using Smart Glasses With Piezoelectric Sensor,” *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 6, pp. 1495–1503, Nov. 2017, doi: 10.1109/JBHI.2016.2640142.

CHAPTER 4 – COMPARISON OF WEARABLE SENSORS FOR ESTIMATION OF CHEWING STRENGTH

Abstract

This paper presents wearable sensors for detecting differences in chewing strength while eating foods with different hardness (carrot as a hard, apple as moderate and banana as soft food). Four wearable sensor systems were evaluated. They were: (1) a gas pressure sensor measuring changes in ear pressure proportional to ear canal deformation during chewing, (2) a flexible, curved bend sensor attached to the right temple of eyeglass measuring the contraction of the temporalis muscle, (3) a piezoelectric strain sensor placed on the temporalis muscle, and (4) an electromyography sensor with electrodes placed on the temporalis muscle. Data from 15 participants, wearing all four sensors at once were collected. Each participant took and consumed 10 bites of carrot, apple, and banana. The hardness of foods was measured by a food penetrometer. Single-factor ANOVA found a significant effect of food hardness on the standard deviation of signals for all four sensors (P-value < .001). Tukey's multiple comparison test with 5% significance level confirmed that the mean of the standard deviations were significantly different for the provided test foods for all four sensors. Results of this study indicate that the wearable sensors may potentially be used for measuring chewing strength and assessing food hardness.

4.1 Introduction

Mastication is a very complex process which depends upon many factors. Physical properties of food such as hardness, flavor, appearance, size, etc. play an important role in the selection of food. The masticatory function is related to sensory-motor activities involving teeth, muscles of jaw, lips, tongue, cheeks, and the salivary secretions production. This is why, the interaction among food perception, oral physiology, and food properties is an intriguing field of research. The perceived food texture or hardness is one of the three main essential factors (appearance, flavor, and texture) of food. Previous research brought into attention that food texture influences the chewing process because changes in food texture or hardness cause masticatory changes [1], [2]. Most of the previous studies which were performed to investigate chewing strength change due to change in food hardness used artificial and nonfood test materials instead of natural food [3]–[8]. In[4], authors used silicon rubber as model food material to investigate effects of sample hardness on human chewing force. In [7], authors used four separate visco-elastic model foods which were identical in shape and size but different in hardness to examine effects of increased hardness on jaw movement. However, no artificial characterization of food hardness can completely simulate the complex oral process during chewing [9]. One notable difference is that within the first few chews of real food, mechanical properties change drastically. The time require to change the properties of real food is very short (as little as 0.2s to 0.3s of first chew)[10], thus, detecting these changes depends on the sensitivity of the sensors used to measure chewing strength. In this study, we aim to clarify the relationship between estimation of chewing strength measured by a variety of sensors and food hardness by conducting a study among healthy individuals consuming real food.

In order to analyze the relationship of chewing strength estimation and food hardness or texture, several sensory modalities have been used. Researchers were using Gnathodynamometer to measure bite force for a long time [11], [12]. Dynamometer is another choice for estimation of chewing force [13]–[15]. Another way of accurately estimating the chewing strength is using electrical strain gauges[16]–[20]. In order to find the mean of maximum bite force in adults with normal occlusion, a strain gssssssauge force transducer was developed [21] and positioned between first molars. In[22], authors used strain-gauge bite transducer for assessment of chewing force on single tooth. More recently researchers used strain-gauge to develop a low-cost bite force measurement device [23]. A multiple point pressure sensor sheet was used to detect the biting pressure for crackers [10], raw carrots and fish gels [24], and cucumbers [25]. Force Sensitive Resistor (FSR) is another common sensor, which is becoming popular for measurement of bite force [26], [27]. The off the shelf availability, low cost, and ease of use of FSR are the reasons for their wide application in chewing strength estimation and bite force measurement [28]–[30]. A combination of sensors is also used by the researchers to record bite force. Recently both FSR and strain gauge were used to develop a novel human bite force recorder [31]. Surface EMG is another type of sensor which is being used by researchers for estimation of chewing strength and bite force measurement [24], [32], [33]. These devices were positioned inside mouth, placed on tooth and jaws, which might interfere or disturb the smooth movements of jaw during chewing. Over the past 40 years, several methods have been proposed to evaluate the chewing strength and bite force in mastication process using different chewing simulators [34]–[36]. But the complex phenomena of intra-oral process in chewing cannot be fully orchestrated by any simulator yet. In this study, four wearable sensor systems were used which do not

interfere with jaw movement during chewing, making the estimation of chewing strength differences among foods with different hardness more reliable.

In this recent era of wearable devices, several approaches were developed for detection of chewing, both in laboratory and free-living environment [37]–[41], but never been explored to estimate strength of chewing. In [42], authors have developed a wearable device, which can automatically count chews using a piezoelectric film sensor placed below the ears to detect periods food intake based on the chewing signal. The chewing sound captured through an earpiece was used to develop automatic chewing detection in [41]] to monitor food intake behavior. In [43], authors used a combination of microphone and piezoelectric strain sensor to detect swallowing and chewing to estimate energy intake. Chewing event can be detected via chewing sounds [41], movement of muscle using strain sensor [37], [38], [42], [44], or EMG and force sensors [45]. As these wearabale sensors are used to detect chewing and monitor food intake, energy or caloric intake so far, it is necessary to further explore these chewing sensors to estimate strength of chewing. In this study, we used four wearable chewing sensors which were already developed to test the feasibility of using them to estimate strength of chewing and compare the performance among them.

Chewing strength can be defined as the resultant of all the forces applied between the maxillary and mandibular teeth resulting from contraction of the masticatory elevator muscles during food intake. The sensors used so far to estimate chewing strength are directly placed inside the mouth placed on teeth or jaws [11]–[32] which can interfere with the complex process of mastication and can lead to erroneous results. Detection of chewing, on the other hand, can be defined as the detection of any physical changes associated with the chewing such as pressure change in air canal, contraction of temporalis muscle, sound of chewing, etc. Nowadays different

wearable sensors such as wearable microphone to detect chewing sound, printed strain sensors attached on jaw and temporalis muscle, accelerometers, etc. are being used to detect chewing event [37]–[42]. But these wearable sensors were never explored to estimate strength of chewing after detection of chewing event. Again most of the studies performed to estimate strength of chewing used artificial test foods which fails to completely simulate the complex oral process during chewing. In this study, we explored the feasibilities of four wearable sensor modalities which are capable of detecting chewing event to estimate chewing strength related to each bite using real edible test foods.

This study aims to explore the possibility of using wearable sensors to estimate strength of chewing real foods. Air pressure sensor, piezoresistive bend sensor, piezoelectric strain sensor, and surface EMG sensor are the four sensor modalities that were used in this study. Using these four wearable sensors, the relationship between food hardness and estimated chewing strength was explored.

4.2 Materials and Methods

4.2.1 Wearable Sensors

In this study, four wearable sensors were used to estimate the chewing strength during consumption of test foods.

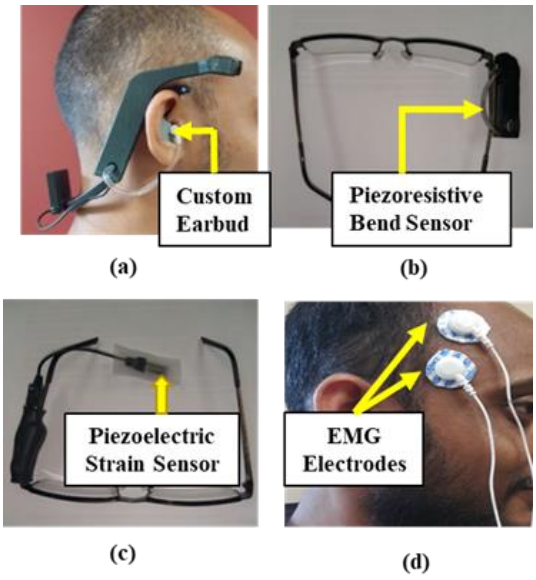


Figure 4-1: Four wearable sensor systems: (a) Ear Canal Pressure Sensor
(b) Piezoresistive Bend Sensor (c) Piezoelectric Strain Sensor (d) EMG Sensor.

4.2.1.1 Ear Canal Pressure Sensor

The sensor measures changes in ear canal pressure during chewing due to the grinding motion of the lower jaw. The position of the human ear canal is between the Mandible Condyle (the upper tip of the mandible) and the Mastoid (the back part of the temporal bone), and it extends into the Mastoid. During chewing, the lower jaw moves to open and close the mouth, causing the ear canal to expand as the Mandible Condyle slides forward and backward, creating a void that will be filled by the tissue surrounding the ear. This expansion and contraction of the ear canal during chewing produce a change in pressure which can be used to monitor food intake. This wearable device consists of an air pressure sensor (SM9541, from Silicon Microstructures) system and a custom-molded earbud (Figure 4-1a). Commercially available Sharkfin Self-Molding Earbud silicon rubber was used for making custom-molded earbuds for each participant. The data from the sensor was sampled at a rate of 128 Hz by STM32L476 microprocessor and stored in an SD card.

4.2.1.2 Piezoresistive Bend Sensor

The sensor was implemented as an attachment to the right side of the eyeglass frame (Figure 4-1b). A short piezoresistive sensor, sensitive to bending (Spectra Symbol 2.2” sensor) produced a signal proportional to the force applied by the temporalis muscle to the frame of the glasses. The contraction of the temporalis muscle during chewing caused the bend sensor to change the resistance. A layer of polyurethane coating was applied to the sensor to protect the sensor from sweat. The device was attached to the temple of the eyeglasses using double-sided tape covered with strong adhesive. Data from the sensor were sampled at a rate of 128 Hz by a microprocessor (MSP430F2418, Texas Instrument, TX, USA) and stored in an SD card.

4.2.1.3 Piezoelectric Strain Sensor

This sensor relies on the movement of the temporalis muscle during various activities. The temporalis epidermis is forced to contract and relax by the up-down and side-to-side grinding movements of the lower jawbone (mandible) during the chewing cycle. This sensor utilizes the oscillatory movements of the temporalis muscle during chewing for detection of chewing events. The piezoelectric sensor (LDT0-028K from Measurement Specialties Inc. VA) was attached to the left temporalis muscle by medical tape (Fig. 1c). This device also consists a low power 3D accelerometer (ADXL335 from Analog Devices, Norwood, MA, USA) and a Bluetooth module (RN-42). The signal produced by the piezoelectric strain sensor was created by the contraction of the temporalis muscle. The data collected by the sensor were sampled at a rate of 1 kHz using MSP4 30F2417 processor and sent via Bluetooth to an Android phone for storage. The detailed information about the sensor can be found in [37].

4.2.1.4 Electromyography Sensor

The activation of the temporalis muscle caused by the grinding movement of the lower jaw (Mandible) during chewing can also be recorded using an electromyography (EMG) sensor (Fig. 1d). The EMG surface electrodes were placed on the right temporalis muscle and the EMG signal was captured by commercially available BITalino (r) evaluation board. This evaluation board consists of onboard EMG, ECG, EDA, ACC, and LUX sensors. The hardware of this device contains a control block with ATMega328P microcontroller, a data transfer block that consists of Class II Bluetooth v2.0, and two connectivity blocks that enable RJ22 pugs to be added to the device. The data from the ECG sensor were sampled at a rate of 1KHz using ATMega328 microcontroller and sent over Bluetooth to an Android phone for storage using OpenSignal application.

4.2.2 Test Foods

Three types of natural foods with different textures and properties were used in the study: raw carrots (8-10 inches in length) as hard food, apple (Gala) as food with medium hardness and banana (organic, medium ripe) as soft food. All these food items were purchased as a single batch from a supermarket (Publix). All foods were preserved in the refrigerator when not in use. The study was performed within two days of the purchase to reduce variation in food properties due to ripening.

4.2.3 Objective Hardness Evaluation

The hardness of the test foods were measured using a hand-held Magness-Taylor pressure tester known as a Fruit Hardness Tester. In this study, GY-3 Fruit Hardness Tester from Vetus Instrument was used. This device has 2 pressure tips (8mm and 11mm diameter) for measuring the hardness of fruits with different textures. The 8mm pressure tip was used to measure the

hardness of the apple and banana. A custom-made 3.5mm diameter pressure tip was used to record the hardness of the carrot. Approximately 2mm of the skin of the apple was peeled. The hardness measurements were made halfway between the stem and calyx ends of the fruit. Using the 8mm probe, two readings were taken from two sides of the apple. The rest of the fruit was sliced and served to a participant. For the carrot, after peeling approximately 2mm of skin in the middle, two readings were taken using the 3.5mm probe. The rest of the carrot was sliced (approximately 0.5 inch) and served to the participant. For bananas, two bananas were served to each participant and one reading was taken from each. The banana was cut in half, the skin was peeled and hardness was measured using 8mm probe. To validate the measurement of a custom-made 3.5mm pressure head, a pilot study was performed where 10 samples from 5 apples (2 measurements for each) were taken using both 8mm and 3.5mm pressure tip. A single factor ANOVA (F-value .03, p-value 0.87) shows that there is no significant difference between 3.5mm and 8mm pressure head readings of the hardness of the apples.

4.2.3 Participants

A total of 15 participants (8 males, 7 females), aged 24.1±4.7 years, were recruited to participate in this study. The average body mass index is 22.7 $\pm 2.5\ kg/m^{-2}$. Participants did not present any existing tooth and jaw-muscle pain or medical condition that affected chewing. This study was approved by the Institutional Review Board at The University of Alabama. Participants read and signed an informed consent document before participating in the study.

4.2.4 Data Collection

The full experiment consisted of two visits to the laboratory. During the initial visit, participants were briefed about the study and signed informed consent. In addition, a custom-molded earbud was also made, and an eyeglass frame was selected for individual participants.

The placement of the bend sensor on the frame was also determined in the initial visit. In the second visit, upon arrival, the wearable sensors were given to the participant to wear, all at the same time. The participants consumed all of the food items while seated. Before serving each food item, a measurement of food hardness was recorded using a fruit hardness tester. The participants took at least 10 bites of each food item. The subjects were asked to chew at a natural and self-determined pace on their preferred chewing side. The sequence of test foods was randomly assigned among the subjects; the order of the food items was recorded. Subjects were instructed to provide a pause of approximately 30 seconds in chewing between two consecutive food items.

4.2.5 Signal Processing

The signals recorded by each sensor system were processed offline in MATLAB (Mathworks Inc., Natick, MA, USA). As the characteristic frequencies that define chewing phenomena normally fall within the range of 1.25-2.5 Hz[44], pressure sensor, piezoresistive bend sensor, and piezoelectric strain sensor signals were filtered using a low pass filter with the cut off frequency of 3 Hz. Raw EMG signal was used for analysis. Each recorded sensor signal was z-normalized using the following equation before processing.

$$z = x - \bar{x}/S \quad (1)$$

Where,

$\bar{x}$= mean of the sample

S=standard deviation of the sample

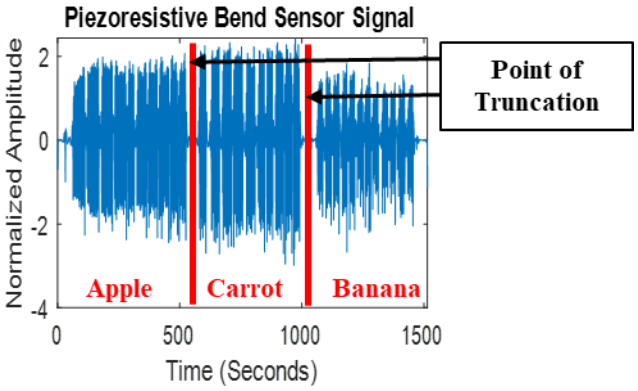


Figure 4-2: Truncation of full signal for separate test food.

After being normalized, each sensor signal was manually truncated for separate test foods using the pause showed as the vertical red line in Figure 4-2 and the record of order of served test foods. The first five chews of each bite were then truncated manually. The figures 4-3(b), 4-5(b), 4-7(b), and 4-9(b) represent the manual truncation of the first five chews of each bite. The standard deviation of each bite was recorded for analysis. Piezoresistive bend sensor, piezoelectric strain sensor and EMG sensors used in this study utilize the movement of temporalis muscle while eating to detect chewing. The pressure sensor uses the change in air pressure inside ear canal while chewing to detect eating episodes. As the sensors are in contact with the skin or use air canal pressure rather than placed inside the mouth, we can only estimate the strength of chewing rather than directly measure the force on the interface between teeth. In[49], the authors used standard deviation as an estimator of bite force variability among different positions of dental arch. In[4], authors conducted a study using silicone rubber as model food to find the effects of sample hardness on human chewing force and that the standard deviation of the chewing cycle is proportional to the peak force (N) for both incisors and molars. Therefore, the standard deviation of the sensor signal was taken as an estimator of the chewing force in this study.

4.2.6 Statistical Analysis

Statistical analysis was performed using MATLAB (ver. 2018b for windows) (MathWorks, Inc.). A single-factor analysis of variance was performed to test the effect of sample hardness on the chewing strength. When the hardness effect was significant (P<.001), Tukeys multiple comparison test was applied as post-hoc analysis. The null hypothesis, there is no significant variation in chewing strength estimation among test foods, against the alternative hypothesis, there is significant variation in chewing strength estimation among test foods, is tested using a Single-Factor ANOVA. A post-hoc analysis of variation (Tukey's test) was performed to find out whether the variation is significant or not among the estimators of chewing strength for different test foods after rejecting the null hypothesis.

4.3 Results

4.3.1 Objective Hardness Evaluation

Objective hardness of the test foods was measured using GY-3 fruit penetrometer before serving the foods to individual subject. For each test food item, two readings were taken and an average reading was recorded. Table 4-1 represents the records of objective hardness measurement of test foods. The average objective hardness measured for carrot is 29.4±1.3 $kgcm^{-2}$, for apple is 6.5±0.3 $kgcm^{-2}$ and for banana is 1.3±0.1 $kgcm^{-2}$.

Table 4-1: Objective hardness measurement record of test foods

Subject	**Carrot hardness** $(^{kg}/_{cm^2})$	**Apple hardness** $(^{kg}/_{cm^2})$	**Banana hardness** $(^{kg}/_{cm^2})$
01	29.5	6.75	1.3
02	29.8	6.35	1.1
03	27.4	6.5	1.2
04	30.8	6.25	1.3
05	30	7	1.35
06	29.5	6.2	1.2
07	29.25	6.6	1.35
08	30	6.6	1.2
09	28	6.75	1.5
10	30.3	6.5	1.3
11	31.3	7	1.2
12	28.75	6.1	1.1
13	30	6.6	1.1
14	26.6	6.35	1.3
15	30	6.1	1.2

The small standard deviation in readings for each test food objective hardness measurement suggests that there was no significant variation in the hardness of similar food served to participants. Again, the variation of objective hardness among different test foods is significant.

4.3.2 Pressure Sensor

Pressure sensor signal of the full experiment (eating of carrot, apple, and banana) is shown in Figure 4-3(a). The first five chews of a bite of three different test foods are shown in Figure 4-3(b). The magnitude of the sensor signal of these three test foods is quite separable. Figure 4-3(c) represents the variation of chewing estimation for three test foods among all participants. The estimated chewing strength is well separable for different foods. The box plot of the standard

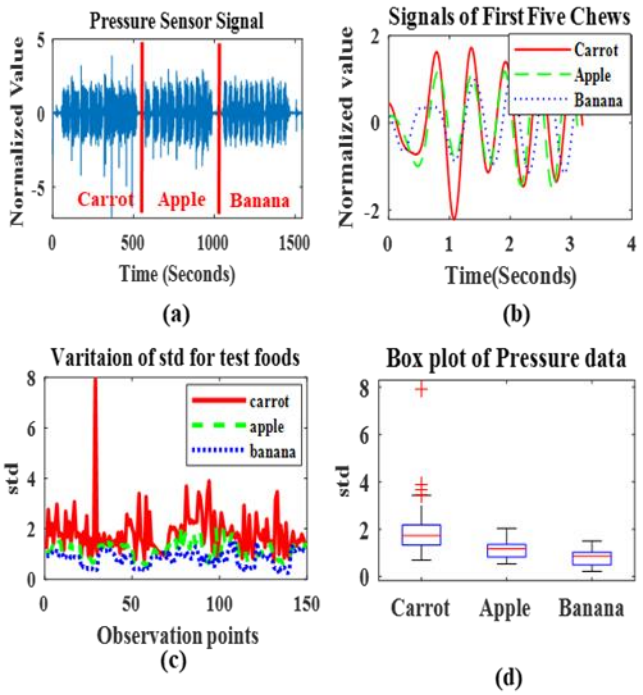


Figure 4-3: (a) Normalized pressure sensor signal (b) Signals of five chews of test foods (c)Variation of standard deviation in chewing test foods (d) Box plot of chewing strength estimation for three test foods.

deviation of chewing signals represented in Figure 4-3(d) also indicates enough variation of chewing strength estimation among test foods. The single-factor ANOVA analysis with the 150 bites of each test food provides a p-value, $< .001$ and F-value, 133.3 which clearly indicates that there is significant variation of chewing strength among different test foods. Hence, we

reject the null hypothesis. Tukey's test shown in Figure 4-4, also confirms that the variation of estimated chewing strength among test foods is significant enough (with $\propto = .05$). The three groups (test foods) have means of chewing strength significantly different from each other. These tests also showed the chewing strength estimated by the pressure sensor is highest for carrot and lowest for banana.

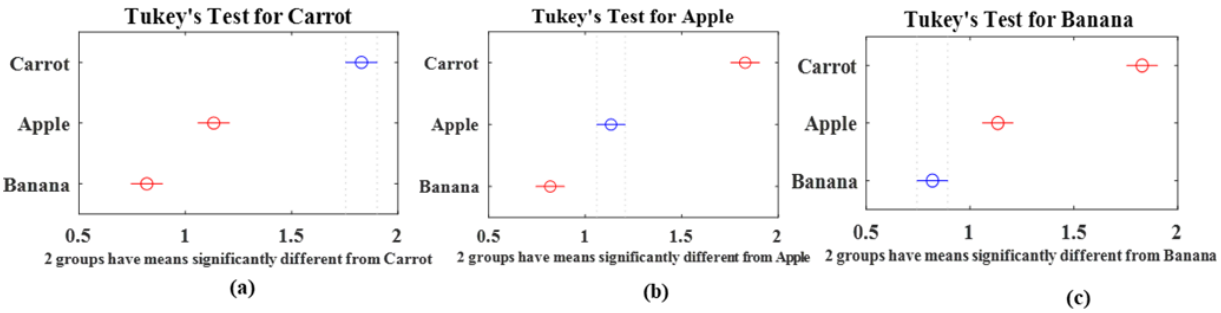


Figure 4-4: Post-hoc analysis of variation using Tukey's test for pressure sensor.

4.3.3 Piezoresistive Bend Sensor

Figure 4-5(a) represents the piezoresistive bend sensor signal of the full experiment. When there is no chewing, the sensor signal has values around zero. While chewing, the sensor signal oscillates which represents the movement of the temporalis muscle. Figure 4-5(b) represents the signals of first five chews of a bite for three different test foods. From the signal, the variation in estimation of chewing strength is quite visible. Carrot has higher peak to peak amplitude, apple has moderate and banana has the lowest amplitude. Figure 4-5(c) represents the variation of standard deviation of the signals for all 15 participants. Figure 4-5(d) also represents the box plot of standard deviation readings of 150 data points for each test food item.

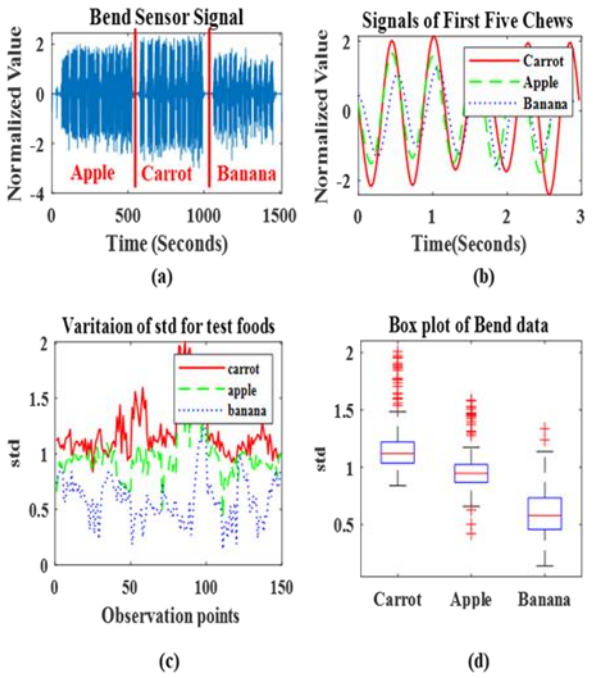


Figure 4-5: (a) Normalized Piezoresistive bend sensor signal (b) Signals of five chews for test foods (c)Variation of standard deviation in chewing test foods (d) Box plot of chewing strength estimation for three test foods.

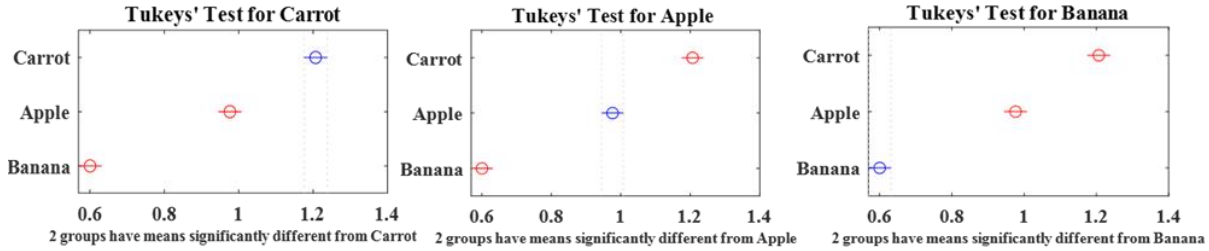


Figure 4-6: Post-hoc analysis of variation using Tukey's test for piezoresistive bend sensor.

These figures indicate that chewing strength estimated by the piezoresistive bend sensor for foods with different hardness is well separable. Single Factor ANOVA analysis with the 150 data points for each test food provides a p-value of $< .001$ and F-value of 269.3, which suggests that there is significant variation of chewing strength estimation among different test foods. Tukey's test (Figure 4-6) also confirms that the variation of estimation of chewing strength among test foods is significant enough (with $\propto = .05$).The three groups (test foods) have means

of estimated chewing strength significantly different from each other. This test also proved that the estimated chewing strength by the piezoresistive bend sensor is highest for carrot, lowest for banana and moderate for apple. This confirms that the wearable piezoresistive bend sensor system can be used for estimation of chewing strength variation among foods with different hardness.

4.3.4 Piezoelectric Strain Sensor

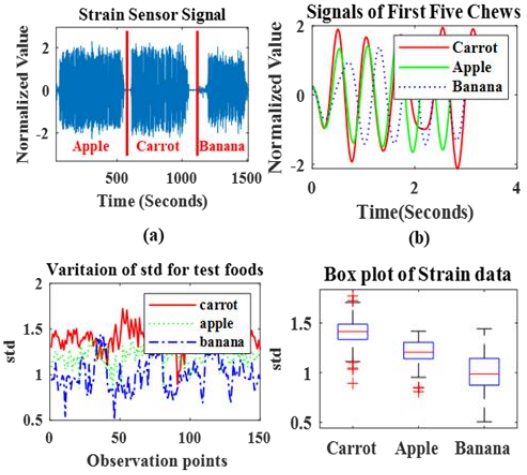


Figure 4-7: (a) Normalized strain sensor signal (b) Signals of five chews for test foods (c)Variation of standard deviation in chewing test foods (d) Box plot of chewing strength estimation for three test foods

The strain sensor signal in Figure 4-7(a) represents that this sensor can effectively detect food intake, separate the bites, and also can count the number of chews in an eating episode. The signal of first five chews of three test food items, Figure 4-7(b), shows that the variation in chewing strength estimation for three different food items is quite visible. Carrot has higher peak to peak amplitude, apple has moderate and banana has the lowest amplitude. This indicates that standard deviation in the first five chews of a bite can be considered as an estimator of chewing

strength. Figure 4-7(c) represents the variation of standard deviation in chewing signal for different test foods for all 15 participants.

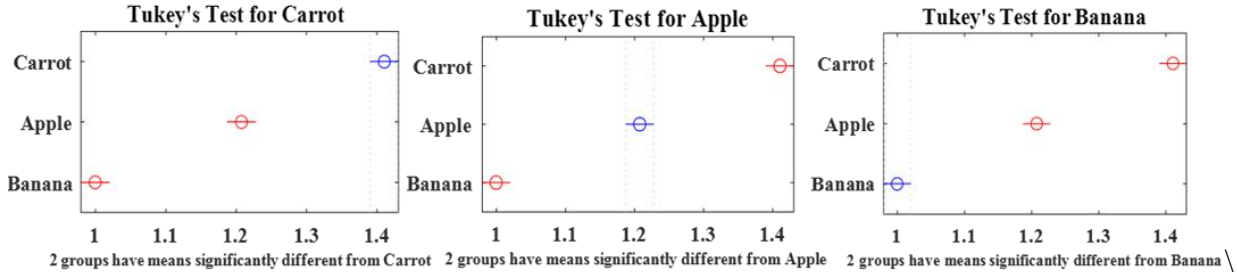


Figure 4-8: Post-hoc analysis of variation using Tukey's test for piezoelectric strain sensor.

The boxplot of the readings in Figure 4-7(d), shows that the means of the standard deviation of chewing signals are different for the test foods. Single Factor ANOVA using the 150 data points of estimated chewing strength of each test food item provides a very small p-value $< .001$ and large F-value (289.8) which suggests that the strain sensor can identify the variation of chewing strength estimation among different test foods. Post-hoc analysis by Tukey's test, Figure 4-8, also confirms that the variation of estimated chewing strength among test foods is significant enough (with $\propto = .05$). The three groups (test foods) have means of chewing strength estimation significantly different from each other. This post-hoc test also showed that variation in estimated chewing strength from strain sensor is highest for carrot, lowest for banana and moderate for apple. Both the ANOVA and Tukey's test confirms that this wearable strain sensor can be used for estimation of chewing strength variation among foods with different hardness.

4.3.5 EMG Sensor

The surface EMG sensor was used as a base for comparison of proposed wearable sensor performance. The signal of EMG sensor, Figure 3-9(a), suggests that the chewing of each test foods can be successfully identified by this sensor. Also, from the signals of the EMG, each bite can be separated, and number of chews can be counted.

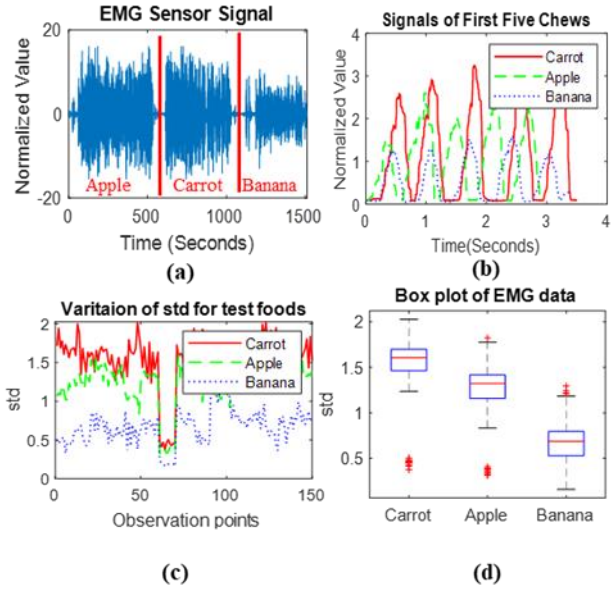


Figure 4-9: (a) Normalized EMG sensor signal (b) Signals of five chews for test foods (c)Variation of standard deviation in chewing test foods (d) Box plot of chewing strength estimation for three test foods.

The signals of first five chews of three test food items, Figure 4-9(b), shows that the variation in sensor signal for three different food items is quite visible, which confirms that the standard deviation in the first five chews of a bite can be considered as an estimator of chewing strength for EMG sensor. The variation of standard deviation in chewing different test foods for all 15 participants in Figure 4-9(c) and the boxplot of the readings in Figure 4-9(d), indicates that the surface EMG sensor can effectively record the differences in the estimation of chewing strength for the three different test foods. Again, the null hypothesis that there is no significant variation in chewing strength estimation among test foods against the alternative hypothesis,

there is significant variation in chewing variable among test foods is tested using a Single-Factor ANOVA for EMG. A post-hoc analysis of variation (Tukey's test) was performed to find out whether the variation is significant or not among the mean of estimated chewing strength for different test foods after rejecting the null hypothesis.

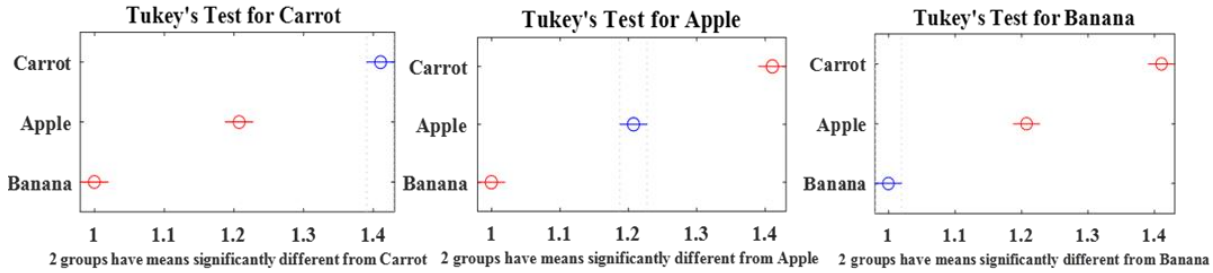


Figure 4-10: Post-hoc analysis of variation using Tukey's test for EMG sensor.

Single-factor ANOVA analysis using the estimated chewing strength of 150 bites of each test food provides a small p-value ($< .001$) and large F-value (344.6) which indicates that there is significant variation in chewing strength estimation among different test foods. Tukey's test (Figure 4-10) also confirms that the variation of estimation of chewing strength among test foods is significant enough (with $\propto = .05$). The three groups (test foods) having means of estimated chewing strength significantly different from each other. This test also showed the estimated chewing strength from the EMG sensor is highest for carrot and lowest for banana and moderate for apple.

4.4 Discussion

Wearable sensors were never explored to detect and record chewing strength to the best of our knowledge. In this study, 4 different wearable sensor modalities were tested for the first time to estimate chewing strength while chewing three different test foods with hardness significantly different from each other.

Four different wearable sensors were developed to detect chewing using temporalis muscle movement (piezoresistive bend sensor, piezoelectric strain sensor, EMG sensor) and change in air pressure in the ear canal (pressure sensor) while eating. As the piezoresistive bend sensor and piezoelectric strain sensor devices were attached to an eyeglass and different subjects may have different head sizes, the force acting on the eyeglass would be different if a fixed eyeglass were used. These two sensor systems were attached to the subject's eyeglasses (6 of the 15 subjects). If any subject was not habituated to wearing eyeglasses a choice of reading glasses (overall frame width from 125mm to 145mm) were provided and subjects chose the one they felt comfortable wearing.

As the mechanical properties of food change dramatically within a short period of the first chew, only the first five chews of a bite were used in this study. The signals of the first five chews of a bite for all four sensors showed that there is variation among signals for three different test foods. The piezoresistive bend sensor (sampling frequency 128 Hz), piezoelectric strain sensor (sampling frequency 1 KHz), and Surface EMG (Sampling frequency 1 KHz) used in this study are validated sensor systems that were used in previous chewing studies. The novel air canal pressure sensor is developed to record data at 128 Hz sampling frequency considering power consumption, signal strength and effect of noise on the signal etc. As the standard deviation of the sensor signal of first five chews of each bite was taken as an estimator of the chewing force in this study, the difference in sampling frequencies of the sensor systems do not contribute any difference in estimation values.

Analysis of Variance using the estimated chewing strength for different test foods showed that all four-sensor systems can effectively (p-value$<.001$) identify the variation in estimated chewing strength because of the difference in the texture of test foods. The post-hoc

Tukey's test showed that the variation in chewing strength estimation is significant to categorize these test foods into three different categories (soft, medium, and hard food). Results showed that the estimated chewing strength is proportional to the food hardness.

Table 4-2: Summarized result of Single Factor ANOVA

Sensor Type	P-Value	F-Value
Pressure Sesnor	$< .001$	133.3
Piezoresistive bend Sensor	$< .001$	269.3
Piezoelectric Strain Sensor	$< .001$	289.9
EMG	$< .001$	344.5

A total of 15 subjects were recruited for this study. Each subject was provided with three test foods and was instructed to have at least 10 bites of individual test food items. The standard deviation of the first five chews of each bite was calculated and used as an estimator of chewing strength (data point), which resulted in a total of (15×10) 150 data points for individual sensor systems and individual test food items. The conclusion was drawn from Single Factor ANOVA analysis with the 150 data points for each test food.

The result of Single Factor ANOVA shows that EMG performs best in finding variation in the estimation of chewing strength for different test foods. However, the P-values for other sensors are so small and the F-values are quite large which indicates that their performance in finding a variation of estimation of chewing strength is comparable with EMG (Table 4-2).
A limitation of this study is that the chewing strength was estimated rather than measured directly. However, the use of direct measurement would either make the cost dramatically higher (if instrumenting the teeth) or prevent the use of real foods (use of food substitutes instrumented

with sensors). A contribution of the proposed approach is that it allows to conduct of inexpensive studies with real foods.

Another limitation of this study could be the small sample size and limited test food items. Also, the GY-3 fruit penetrometer can only be used on fruits to measure the objective hardness. There are many commercially available hardness testers for different food items such as FCA-DSV-50N for bread, jam, cookies, sausages, etc., which are costlier. As the current study is a feasibility study of the wearable sensors to estimate the strength of chewing, for simplicity the sample size of the study was kept small and the test foods were selected to represent 3 different types of food (carrot with hard texture, apple which represents food with medium-hard texture, and banana to represent foods with softer texture). With the knowledge of the current feasibility study, a study with a larger population and an expanded variety of test foods can be conducted in the future to explore the effect of chewing habits, age, gender, etc. on chewing strength.

In the future, the findings of this study will help develop a predictive model to identify food hardness from estimated chewing strength using wearable sensors. Identifying the physical properties of food (hardness) using wearable sensors may provide valuable information (mass, caloric value) to build a more robust model of energy estimation in the wild. The relationship between chewing number, chewing duration, and chewing time with food hardness can be explored in the future using these wearable sensors.

4.5 Conclusions

To record bite force or estimate chewing strength while chewing foods with different hardness, wearable sensors were never explored. This paper presented an idea of detecting of variations in chewing strength for foods with different hardness using wearable sensor systems.

The proposed pressure sensor, piezoresistive bend sensor attached to the temporalis muscle, piezoelectric strain sensor, and EMG sensor are feasible to estimate the chewing strength in the masticatory process. All four sensor modalities can effectively detect the variation in chewing strength while eating foods with different textures. The estimated strength of chewing is proportional to the hardness of the food. Hard food provides a larger value of estimated chewing strength, while soft food results smaller value in chewing strength estimation. The detection of variation in chewing strength estimation for different foods using these low-cost, ease-of-use devices, may be useful for understanding the complex food intake process.

References

[1] M. Bourne, “Relation Between Texture and Mastication,” *Journal of Texture Studies*, vol. 35, no. 2, pp. 125–143, 2004.

[2] L. Mioche, “Mastication and Food Texture Perception: Variation with Age,” *Journal of Texture Studies*, vol. 35, no. 2, pp. 145–158, 2004.

[3] K. D. Foster, A. Woda, and M. A. Peyron, “Effect of Texture of Plastic and Elastic Model Foods on the Parameters of Mastication,” *Journal of Neurophysiology*, vol. 95, no. 6, pp. 3469–3479, Jun. 2006.

[4] K. Kohyama, E. Hatakeyama, T. Sasaki, H. Dan, T. Azuma, and K. Karita, “Effects of sample hardness on human chewing force: a model study using silicone rubber,” *Archives of Oral Biology*, vol. 49, no. 10, pp. 805–816, Oct. 2004.

[5] L. Mioche and M. A. Peyron, “Bite force displayed during assessment of hardness in various texture contexts,” *Archives of Oral Biology*, vol. 40, no. 5, pp. 415–423, May 1995.

[6] J. Paphangkorakit and J. W. Osborn, “Effects on human maximum bite force of biting on a softer or harder object,” *Archives of Oral Biology*, vol. 43, no. 11, pp. 833–839, Nov. 1998.

[7] M. Peyron, C. Lassauzay, and A. Woda, “Effects of increased hardness on jaw movement and muscle activity during chewing of visco-elastic model foods,” *Exp Brain Res*, vol. 142, no. 1, pp. 41–51, Jan. 2002.

[8] A. Tarkowska, L. Katzer, and M. O. Ahlers, “Assessment of masticatory performance by means of a color-changeable chewing gum,” *Journal of Prosthodontic Research*, vol. 61, no. 1, pp. 9–19, Jan. 2017.

[9] M. R. Heath and P. W. Lucas, “MASTICATION: THE NEED FOR COLLABORATIVE RESEARCH,” *Journal of Texture Studies*, vol. 18, no. 2, pp. 111–123, Jun. 1987.

[10] K. Kohyama and M. Nishi, “Direct Measurement of Biting Pressures for Crackers Using a Multiple-Point Sheet Sensor,” *Journal of Texture Studies*, vol. 28, no. 6, pp. 605–617, 1997.

[11] V. Rane, S. Hamde, and A. Agrawal, “Development of computerized masticatory force measurement system,” *Journal of Medical Engineering & Technology*, vol. 41, no. 1, pp. 65–71, Jan. 2017.

[12] A. E. Rowlett, “The Gnathodynamometer and Its Use in Dentistry,” *Proceedings of the Royal Society of Medicine*, p. 9.

[13] A. S. Guimaraes, G. E. Carlsson, and S. K. N. Marie, “Bite force and handgrip force in patients with molecular diagnosis of myotonic dystrophy,” *Journal of Oral Rehabilitation*, vol. 34, no. 3, pp. 195–200, 2007.

[14] E. M. Kogawa, P. S. Calderon, J. R. P. Lauris, C. R. P. Araujo, and P. C. R. Conti, “Evaluation of maximal bite force in temporomandibular disorders patients,” *Journal of Oral Rehabilitation*, vol. 33, no. 8, pp. 559–565, 2006.

[15] M. Palinkas *et al.*, “Age and gender influence on maximal bite force and masticatory muscles thickness,” *Archives of Oral Biology*, vol. 55, no. 10, pp. 797–802, Oct. 2010.

[16] C. P. Fernandes, P.-O. J. Glantz, S. A. Svensson, and A. Bergmark, “A novel sensor for bite force determinations,” *Dental Materials*, vol. 19, no. 2, pp. 118–126, Mar. 2003.

[17] D. J. Anderson, “A Method of Recording Masticatory Loads,” *Journal of Dental Research*, vol. 32, no. 6, pp. 785–789, Dec. 1953.

[18] A. H. Howell and F. Brudevold, “Vertical Forces Used During Chewing of Food,” *Journal of Dental Research*, vol. 29, no. 2, pp. 133–136, Apr. 1950.

[19] A. H. Howell and R. S. Manly, “An Electronic Strain Gauge for Measuring Oral Forces,” *Journal of Dental Research*, vol. 27, no. 6, pp. 705–712, Dec. 1948.

[20] X. Xu, P. S. Wright, M. P. Hector, M. R. Heath, and A. M. Ferman, “Force, Rate and Work Used During Incisor Penetration on Different Textural Foods,” *Journal of Texture Studies*, vol. 39, no. 2, pp. 115–128, 2008.

[21] M. Bonakdarchian, N. Askari, and M. Askari, “Effect of face form on maximal molar bite force with natural dentition,” *Archives of Oral Biology*, vol. 54, no. 3, pp. 201–204, Mar. 2009.

[22] V. F. Ferrario, C. Sforza, G. Serrao, C. Dellavia, and G. M. Tartaglia, “Single tooth bite forces in healthy young adults,” *Journal of Oral Rehabilitation*, vol. 31, no. 1, pp. 18–22, 2004.

[23] J. Fastier-Wooller *et al.*, “Novel Low-Cost Sensor for Human Bite Force Measurement,” *Sensors*, vol. 16, no. 8, p. 1244, Aug. 2016.

[24] K. Kohyama, E. Hatakeyama, H. Dan, and T. Sasaki, “Effects of Sample Thickness on Bite Force for Raw Carrots and Fish Gels,” *Journal of Texture Studies*, vol. 36, no. 2, pp. 157–173, 2005.

[25] K. Kohyama, A. Kato-Nagata, H. Shimada, Y. Kazami, and F. Hayakawa, “Texture of Sliced Cucumbers Measured by Subjective Human-Bite and Objective Instrumental Tests,” *Journal of Texture Studies*, vol. 44, no. 1, pp. 1–11, 2013.

[26] A. Diaz Lantada, C. González Bris, P. Lafont Morgado, and J. Sanz Maudes, "Novel System for Bite-Force Sensing and Monitoring Based on Magnetic Near Field Communication," *Sensors*, vol. 12, no. 9, pp. 11544–11558, Sep. 2012.

[27] V. A. Bousdras *et al.*, "A novel approach to bite force measurements in a porcine model in vivo," *International Journal of Oral and Maxillofacial Surgery*, vol. 35, no. 7, pp. 663–667, Jul. 2006.

[28] K.-R. Lin, C.-H. Chang, T.-H. Liu, S.-W. Lin, and C.-H. Lin, "Experimental and numerical estimations into the force distribution on an occlusal surface utilizing a flexible force sensor array," *Journal of Biomechanics*, vol. 44, no. 10, pp. 1879–1884, Jul. 2011.

[29] D. Flanagan, "Bite force and dental implant treatment: a short review," *Med Devices (Auckl)*, vol. 10, pp. 141–148, Jun. 2017.

[30] K. Lin, T. Liu, S. Lin, C. Chang, and C. Lin, "Occlusive bite force measurement utilizing flexible force sensor array fabricated with low-cost multilayer ceramic capacitors (MLCC)," in *TRANSDUCERS 2009 - 2009 International Solid-State Sensors, Actuators and Microsystems Conference*, 2009, pp. 2238–2241.

[31] M. S. Khaghaninejad, A. Peyravi, A. Khosravifard, E. Peyravi, H. R. Eftekharian, and M. R. Peyravi, "Novel User-friendly Device for Human Bite Force Measurement," *Journal of Dental Biomaterials*, vol. 4, no. 4, pp. 475-483–483, Dec. 2017.

[32] A. Shimada, Y. Yamabe, T. Torisu, L. Baad-Hansen, H. Murata, and P. Svensson, "Measurement of dynamic bite force during mastication," *Journal of Oral Rehabilitation*, vol. 39, no. 5, pp. 349–356, 2012.

[33] M. Steiner, M. E. Mitsias, K. Ludwig, and M. Kern, "In vitro evaluation of a mechanical testing chewing simulator," *Dental Materials*, vol. 25, no. 4, pp. 494–499, Apr. 2009.

[34] K. Kohyama, T. Sasaki, and F. Hayakawa, "Characterization of Food Physical Properties by the Mastication Parameters Measured by Electromyography of the Jaw-Closing Muscles and Mandibular Kinematics in Young Adults," *Bioscience, Biotechnology, and Biochemistry*, vol. 72, no. 7, pp. 1690–1695, Jul. 2008.

[35] T. W. P. Korioth and A. Versluis, "Modeling the Mechanical Behavior of the Jaws and Their Related Structures By Finite Element (Fe) Analysis," *Critical Reviews in Oral Biology & Medicine*, vol. 8, no. 1, pp. 90–104, Jan. 1997.

[36] C. Salles *et al.*, "Development of a chewing simulator for food breakdown and the analysis of in vitro flavor compound release in a mouth environment," *Journal of Food Engineering*, vol. 82, no. 2, pp. 189–198, Sep. 2007.

[37] J. M. Fontana, M. Farooq, and E. Sazonov, “Automatic Ingestion Monitor: A Novel Wearable Device for Monitoring of Ingestive Behavior,” *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 6, pp. 1772–1779, Jun. 2014.

[38] M. Farooq and E. Sazonov, “A Novel Wearable Device for Food Intake and Physical Activity Recognition,” *Sensors*, vol. 16, no. 7, p. 1067, Jul. 2016.

[39] S. A. Rahman, C. Merck, Y. Huang, and S. Kleinberg, “Unintrusive Eating Recognition Using Google Glass,” in *Proceedings of the 9th International Conference on Pervasive Computing Technologies for Healthcare*, ICST, Brussels, Belgium, Belgium, 2015, pp. 108–111.

[40] Y. Dong, J. Scisco, M. Wilson, E. Muth, and A. Hoover, “Detecting Periods of Eating During Free-Living by Tracking Wrist Motion,” *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 4, pp. 1253–1260, Jul. 2014.

[41] S. Päßler and W. Fischer, “Food Intake Monitoring: Automated Chew Event Detection in Chewing Sounds,” *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 1, pp. 278–289, Jan. 2014.

[42] M. Farooq and E. Sazonov, “Automatic Measurement of Chew Count and Chewing Rate during Food Intake,” *Electronics*, vol. 5, no. 4, p. 62, Dec. 2016.

[43] J. M. Fontana *et al.*, “Energy intake estimation from counts of chews and swallows,” *Appetite*, vol. 85, pp. 14–21, Feb. 2015.

[44] E. S. Sazonov and J. M. Fontana, “A Sensor System for Automatic Detection of Food Intake Through Non-Invasive Monitoring of Chewing,” *IEEE Sensors Journal*, vol. 12, no. 5, pp. 1340–1348, May 2012.

[45] K. Fueki, T. Sugiura, E. Yoshida, and Y. Igarashi, “Association between food mixing ability and electromyographic activity of jaw-closing muscles during chewing of a wax cube,” *Journal of Oral Rehabilitation*, vol. 35, no. 5, pp. 345–352, 2008.

CHAPTER 5 – EAR CANAL PRESSURE SENSOR FOR FOOD INTAKE DETECTION AND CHEW COUNTING IN FREE LIVING

Abstract

This paper presents a novel Ear Canal Pressure Sensor (ECPS) for objective detection of food intake, chew counting, and food image capture in both controlled and free-living conditions. The contribution of this study is threefold: 1) Development and validation of a novel wearable sensor that uses changes in ear canal pressure and the device's acceleration as an indicator of food intake, 2) A method to identify chewing segments and count the number of chews in each eating episode, and 3) Facilitation of egocentric image capture only during eating by triggering camera from sensor detection thus reducing power consumption, privacy concerns, as well as storage and computational cost. To validate the device, data were collected from 10 volunteers in a controlled environment and 3 volunteers in a free-living environment. During the controlled activities, each participant wore the device for approximately 1 hour, and during the free living for approximately 12 hours. The food intake of the participants was not restricted in any way in both part of the experiment. Subject-independent Support Vector Machine classifiers were trained to identify periods of food intake from the features of both the pressure sensor and accelerometer and features only from the pressure sensor. Results from leave-one-out cross-validation showed an average 5 sec-epoch classification F-score of 87.6% using only pressure sensor features and 88.6% using features from both pressure sensor and accelerometer in the controlled environment. For the free-living environment, both classifiers accurately detected all

eating episodes. The wearable sensor achieves 95.5% accuracy in counting the number of chews with respect to manual annotation from the videos of the eating episodes using a pressure sensor classifier in the controlled environment. The manual review of the images found that only 3.7% of captured images belonged to the detected eating episodes, suggesting that sensor-triggered camera capture may facilitate reducing the number of captured images and power consumption of the sensor.

5.1 Introduction

The primary source of energy and nutrients necessary to maintain human life is food intake. Monitoring of daily food intake and ingestive behavior is an emerging research area that has direct implications on human health, as inadequate or excessive food intake may lead to the development of medical conditions such as malnutrition and underweight, or overweight and obesity, respectively [1]. Monitoring ingestive behavior and understanding an individual's diet is also key in the diagnosis and treatment of eating disorders such as anorexia, bulimia, and binge eating [2]. To monitor food intake and identify the dietary restrictions that should be followed for a healthy life, technology-driven wearable food intake monitoring tools are gaining popularity.

Food intake monitoring tools can be divided into two major categories: a) traditional and b) automatic monitoring tools [3]. Traditional food intake monitoring tools include methods such as frequency questionnaires of food items [4], food recall of 24 hours [5], and food records [3], which largely depend on the self-report of the individual. Though these traditional methods are simple and low cost, they can be tedious, as lots of input is required to fill, compliance rates are poor, and produce inaccurate data due to misreporting by the individual. With the advancement in technology, over the past decade, researchers are developing automatic food intake monitoring systems mainly using different physical sensors to alleviate the inaccuracy of

self-reporting methods. Thus, providing easy-to-use, reduction in the burden of the individual, more accurate, and real-time assessment. But up to this day, there is no commercially available automatic food intake monitoring device, which is a research field that remains for exploration [6], [7]. The modern automatic approach to food intake monitoring can be divided into several categories, such as food intake detection, food type classification, food mass estimation, and others. Food intake detection can be considered the first and one of the important stages of the development of the ingestive monitoring device. Several sensor modalities have been explored using different physiological characteristics of eating, such as hand movement during eating, chewing sound, temporalis muscle contraction, head movement, and swallowing sound [8]. Using the pattern of hand movement during eating several hand/wrist-worn wearable devices have been introduced including accelerometers, gyroscopes, and smartwatches [9]–[11]. Food intake can also be detected using the chewing sound of solid and semi-solid food [12]–[14]. The temporalis muscle movement using eyeglasses during chewing was also explored to detect food intake [15], [16]. To monitor food intake using swallowing, sound researchers implemented microphones in the ear or surface electromyography on the throat [17]–[19]. A three-axis accelerometer is also being extensively used by researchers in food intake studies [20], [21]. The eating behavior of an individual is characterized using meal microstructure such as chew count, chewing rate, bite-size, eating episode duration, etc. Studies of meal microstructure may provide new insights and findings in the treatment of obesity. In [22], the authors claimed that a reduction in chewing rate results in the individual's lower energy intake. Other studies have shown that there is a clear relationship between chewing rate and energy intake [23]–[26]. In [26], the authors suggested that to fight against obesity, interventions to improve chewing activity could be a necessary tool. In [23], the authors claimed that eating rate feedback can be

helpful to aid in intervention in eating disorder treatment. Health researchers are applying the number of chews and the rate of chewing as variables in their models for the estimation of ingested mass and energy intake. Counts of chews and swallows were used to develop the energy intake estimation model [27]. In [28] authors concluded that for solid food, recordings of chewing sounds can be used to predict bite weight. Several modalities of wearable sensors and physiological phenomena are being explored by researchers to detect and count the number of chews. Temporalis muscle movement during chewing using a strain sensor is being explored in [15], [29], [30]. Chewing was also detected using chewing sound [31]–[34], EMG and force sensors [35]–[37].

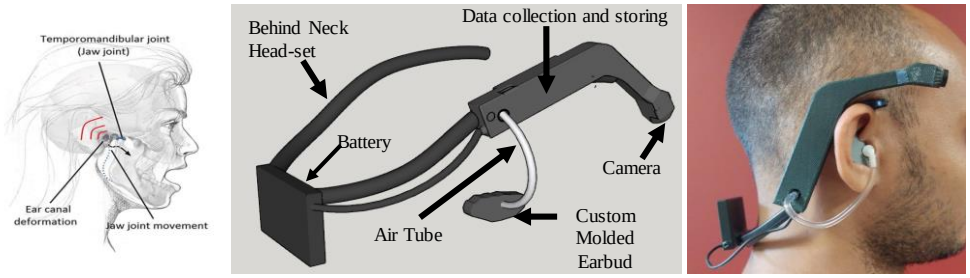
Chewing of solid and semi-solid food causes the ear canal to contract by the opening of the jaw. If the ear canal is closed, for example, by an earbud, the contraction will cause the ear canal pressure to change which can be used as an indicator of food intake. In this study, we use an Ear Canal Pressure Sensor (ECPS) to detect food intake in both controlled and free-living environments. This sensor can be implemented as a part of other devices, that use earbuds such as headphones, hearing aid, etc. The capability of recording the scenes in front of the wearer continuously and automatically by an egocentric wearable camera has also been explored for its potential implementation in objective eating behavior monitoring [38]–[41]. Development in the field of computer vision and image processing allowed these egocentric wearable cameras to be introduced in food intake monitoring applications. Generally, the image-based monitoring method requires images of food intake episodes. These images can be supplied in two ways: 1. Using a smartphone by the individual called active capture, 2. Using a wearable camera that captures images continuously or at a predetermined interval is called passive capture. Most of the modern image-based methods deploy passive capture to make the process less burdensome and

fully automatic. Different wearable egocentric camera systems such as SenseCam [42], AIM (Automatic Ingestion Monitor) [43], and the eButton [44] have been developed using passive capture mode, which results in a bulk number of images when worn in free-living and a longer period. These studies demonstrated that the egocentric wearable camera may be used to study the eating environment and behavior. However, these devices continuously capture images throughout the day, potentially capturing unwanted pictures of the wearer and the people in front of the wearer, which can raise privacy and ethical concerns [45]. As most of the image-based methods of eating behavior monitoring rely on manual review or annotation, privacy becomes a concern when the images of the wearer are observed by another person [46], [47]. As another consideration, continuous capture of images throughout the day requires a larger battery which may affect the weight and comfort of the device. The food intake detection by an ECPS system may be exploited to trigger the camera to start capturing images and to stop at the end of the eating episode. Thus, capturing and storing images that are only needed, provides more privacy to the individual, reducing power consumption, storage, and computational cost.

In this study we propose a novel wearable ECPS system that can be implemented in earbuds as part of other devices such as headphones to monitor food intake in the free-living environment. It accurately detects chewing segments and counts the number of chews in the eating episodes. This device can potentially be explored to control the egocentric camera to capture images only during food intake thus reducing power consumption, privacy concern, storage, and computational cost.

5.2 Materials and Methods

5.2.1 Proposed ECPS (Ear Canal Pressure Sensor)

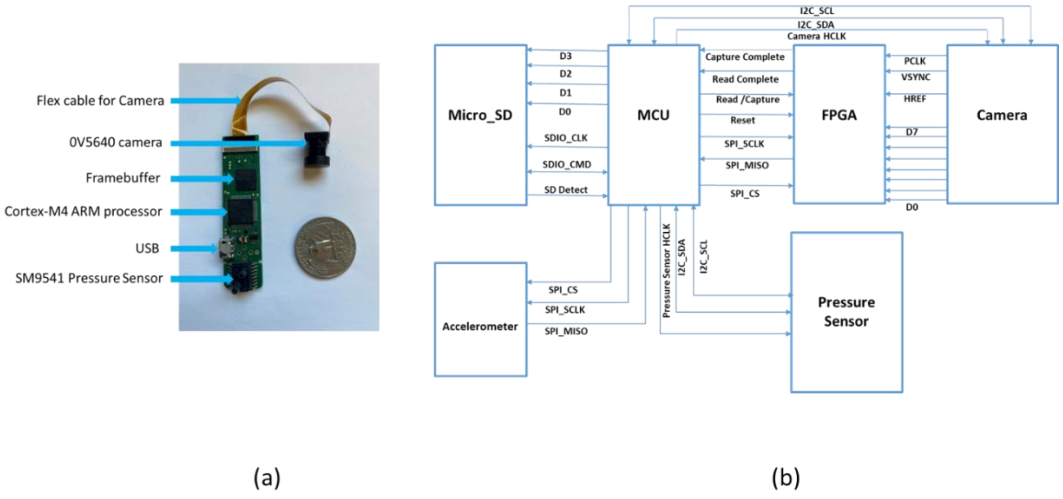


(a) (b) (c)

Figure 5-1: A portable wearable device for monitoring food intake. The module is affixed to right side of a behind-neck headset. (a) Represents the ear canal deformation during chewing (b)represents the wearable sensor system (c) the individual wearing the ECPS prototype.

Figure 5-1(a) represents the working principle of the proposed ECP sensor. The ear canal is positioned between the Mandible Condyle (the upper tip of the mandible) and the Mastoid (the back part of the temporal bone) which extends into the Mastoid. Chewing is associated with the opening and closing of the mouth by moving the lower jaw, which causes the ear canal to expand as the Mandible Condyle slides forward and backward. The movement of the lower jaw creates a void that will be filled by the tissue surrounding the ear. This deformation of the ear canal associated with chewing during solid food intake produces a change in pressure which can be used to monitor ingestive behavior. The ear canal pressure changes during chewing were recorded by a pressure sensor through an air tube connected to an earbud. As a different person has a different size of the ear, an off-the-self-earbud may not feel comfortable for a long time use and can cause poor signal strength. Custom molded earbud for each participant was made using commercially available Sharkfin Self-Molding Earbud silicon rubber. Figure 5-1(b) represents the assembly of the proposed wearable sensor systems. Figure 5-1(c) shows the designed prototype of ECPS worn by an individual. It is important to note that the system can be manufactured in smaller sizes and more convenient but as the aim of this study was limited to the

validation of ECPS for food intake monitoring, the system was developed considering time, effort, and cost. The ECPS device used in this study consisted of a digital air pressure sensor (SM9541, from Silicon Microstructures), a low-power 3D accelerometer (ADXL332 from Analog Devices, Norwood, MA, USA), and a miniature wide-angle lens camera (OV5640, Omnivision Inc.). The sensors and the camera were interfaced with a Cortex-M4 ARM processor (STM32L476RG, ST Microelectronics, Geneva, Switzerland, 80 MHz CPU, 39 µA/MHz). The camera used its DVP interface to output a JPEG compressed image. An FPGA (Actel, Microsemi)-based framebuffer was employed to receive 8-bit image data from the camera, store in a 1Mb external memory chip, and transfer to the MCU via a serial (SPI) interface.



(a) (b)

Figure 5-2: The electronics of ECPS. (Left) ECPS subsystems: Pressure Sensor, USB, MCU, Framebuffer, Camera; (right) Overview of the system with subsystems interconnections

Data from the pressure sensor was sampled at 14-bit/128 Hz using the I2C interface. Accelerometer data were sampled at 16-bit/ 128 Hz via the SPI interface. The device was designed to capture food images in front of the wearer with a gaze-aligned wide-angle lens at a rate of one image per 10 seconds. The components along with communication and storage units of ECPS were installed on an 8.0x1.9x1.5 cm PCB covered by a plastic enclosure with a total

weight of 15g. The enclosed device was attached to the right side of a commercially available off-the-shelf behind-neck headset (Sony MDR-G45LP – Black). A commercially available acoustic air tube was used to convey the ear pressure onto the sensor. The enclosure of the ECPS has an opening to connect an air tube with the pressure sensor. The other side of the air tube goes into the custom-molded earbud to convey ear canal pressure. For user comfort, the earbuds from the headset were removed. A Li-Ion cylindrical battery of 1000mAh capacity, 22g weight, and 50 × 34 × 6mm size were used in this system as the external power source. The device can record sensor data and images for around 20.5 hours. The egocentric camera was used in this device to record images in front of the wearer. The manual review of the recorded images was used for validation of the developed food intake detection models. Also, the recorded images were analyzed to compute the percentage of images related to food intake in a daily setting. The captured images related to food intake can provide more information (food type, nutrition, portion size, etc.) which will be used in future work. Figure 5-2 represents the subsystems of the proposed ECPS.

5.2.2 Study Protocol

The experiment was divided into two parts: 1. Laboratory controlled environment, and 2. Free-living environment. 1. Laboratory-controlled environment: In this phase, the experiment was conducted in a controlled environment, such as a laboratory, where the conditions could be controlled and manipulated. During this phase, the data was collected and used to develop the classification models. 2. Free-living environment: In this phase, the experiment was conducted in a real-world setting, such as the participants' natural habitat. This environment is typically more complex and less controlled than a laboratory environment. The data collected during this phase was used to test the classification models developed in the first phase to evaluate their performance in a more realistic scenario.

In the laboratory-controlled experiment, the participants were requested to come to the laboratory in the morning without having breakfast. Upon arrival, participants were given the wearable sensor system. The participants were then accompanied to get the meal from the school cafeteria. The food choice was not restricted in any way. After purchasing the meal participants were accompanied back to the laboratory to perform the activities listed in Table 1. The participant was encouraged to talk with the research assistant while eating to simulate a real-life environment. The order of the activities was randomized among participants. There was no restriction on the time required to finish the meal. The duration of the controlled experiment for each participant was approximately 1 hour. For the free-living experiment, participants were requested to come to the human participant laboratory in the morning without taking any meals. Upon arrival, the participants were provided the ECPS to wear and then instructed to continue day-to-day life throughout the day. The participants were instructed to perform daily activities with virtually no restriction except during the activities that have a chance of contact of water with the device and during sleeping/taking naps. The duration of the free-living experiment for each participant was approximately 12 hours.

Table 5-1: Study protocol

Activity	Description	Duration (Approximately)
Laboratory Part	Sit silently: sit in a comfortable position	5 min
	Sit while talking: read a document aloud	5 min
	Eat a meal: eat a meal (eating episode -1)	20 - 40 min
	Walking while silent: walk on a treadmill at a self-selected comfortable speed	5 min
	Walking while talking: walk on a treadmill at a self-selected comfortable speed and talk with the research assistant	5 min
	Sit silently: sit in a comfortable position	5 min
	Activities of daily living: shelving /stacking items	5 min
Free-living Part	Carry out usual daily activities and eat a usual diet	Full day

5.2.3 Data Collection and Annotation

A total of 10 volunteers were recruited for the controlled environment experiment of this study (6 males and 4 females, average age 29.03 ± 12.20 years, range 19-41 years, average body mass index (BMI) 27.87 ± 5.51 kg/m2, range 20.5 to 41.7). For the free-living environment experiment of this study, data were collected from 3 volunteers (2 males and 1 female, average age 25.33 ± 3.05 years, range 22-28 years, average body mass index (BMI) 22.8 ± 1.4 kg/m2, range 20 to 24). The study was approved by The University of Alabama's Institutional Review Board. Participants did not report any medical conditions that would impact their natural

chewing or eating. During the controlled experiment, participants used a foot pedal to mark chewing sequences while consuming the food items to record the ground truth of chewing sequences [48]. To provide the chew count of the meal the eating episodes were video recorded using an SJCAM SJ4000 Action camera. The camera was positioned such that it records 1080p video at 30fps in profile view, to make the manual annotation of chew count from jaw movement accurate. More elaboration on the annotation of chew count can be found here [49]. During free living, participants were asked to report major activities such as sleeping, sedentary, and physical activities including eating episodes activity through aTimeLogger mobile application [50].

5.2.4 Food Intake Detection

5.2.4.1 Sensor Signal Processing and Feature Extraction

The sensor data from both the accelerometer and pressure sensor were processed to develop a pattern recognition algorithm. To remove the DC component of the accelerometer signal, a high-pass filter with a cutoff frequency of 0.1 Hz was applied. The pressure sensor signal was filtered by applying a low-pass filter with a cut-off frequency of 3 Hz to remove unwanted signals as typical chewing is associated with a frequency in the range of 0.94 to 2.17 Hz [51]. Both sensor signals were z-normalized using the following equation to adjust inter-subject variability.

$$z = x - \bar{x}/S \quad (1)$$

Where,

$\bar{x}$= mean of the sample

S=standard deviation of the sample

Next, the signals were divided into different non-overlapping fixed time segments called ‘epoch’ (1sec, 2 sec, 3 sec, 5 secs, 10 sec, 15 sec, and 30 sec). Figure 5-3 shows the accelerometer

signals, pressure sensor signals, and corresponding foot pedal signals (self-annotation) during eating episodes. The foot pedal signal is ‘pressed or 1’ during food intake and ‘released or 0’ during no food intake. To detect food intake, a set of 20 time and frequency domain features reported in our previous studies [48], [52] was utilized in this work. Because the spectrum of chewing frequencies is mostly contained in the range of 0.94 to 2.17 Hz, physical activity in the range of 2.5-10 Hz, and speech in the range of 10-300 Hz, band filters were applied to the signals to extract activity-specific features [53].

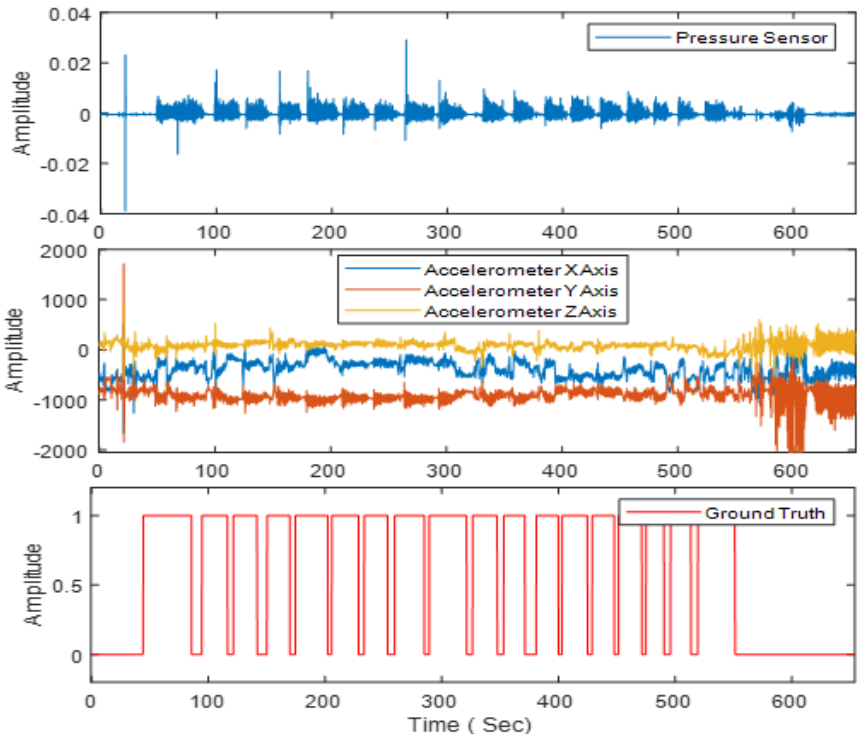


Figure 5-3: The first row demonstrates the pressure sensor signal. Signals from the 3-axes accelerometer are shown on the second row. The foot pedal signal marked is shown on the third row. Here ‘1’ indicates food intake and ‘0’ indicates no food intake.

For each of the 3 axes of the accelerometer, and pressure sensor signal, these 20 features were computed on different epoch lengths (1 sec, 2 sec, 3 sec, 5 sec, 10 sec, 15 sec, and 30 sec). Aggregating all features from both the pressure sensor (20) and accelerometer (60), a total

number of 80 features were obtained from each epoch of different lengths. The list of the feature set is given below.

Time domain:

Mean absolute value: Absolute value of the signal in the epoch

RMS value: RMS (Root Mean Square) value of the signal in the epoch

Maximum absolute value: Maximum absolute value of the signal in the epoch

Median value: Median of the absolute value of the signal in the epoch.

Mean to RMS ratio: The ratio between mean absolute value and RMS value of the signal in the epoch.

Maximum to RMS ratio: The ratio between Maximum absolute value and RMS value of the signal in the epoch.

Mean to Maximum ratio: The ratio between Mean absolute value and Maximum absolute value of the signal in the epoch.

Median to RMS ratio: The ratio between Median and RMS value of the signal in the epoch.

Number of zero crossings: The number of zero-crossings in a signal is the number of times the signal crosses the horizontal axis (zero) during a given time period. We used circular shift to find the number of zero crossing. The circular shift operation shifts the elements of a vector cyclically by a specified number of positions. By multiplying the vector with a circularly shifted version of itself (i.e., the same vector shifted by one position), the zero-crossing points of the vector can be identified.

Mean time between zero crossings: Average time difference between the zero crossings in the epoch

Waveform length: Waveform length is a feature used to describe the shape or morphology of a signal. It measures the total length of the waveform over a given time period, and is computed by summing the absolute differences between adjacent samples of the signal.

$$WL(x) = \sum_{i=1}^{n-1} |x(i+1) - x(i)|$$

Frequency domain:

Spectrum energy: Spectrum energy is a feature used to describe the energy distribution of a signal across different frequency components. It measures the total amount of energy contained in the signal's frequency spectrum and can be calculated using the formula $SE(x) = \sum_{i=1}^{i=n} |X(i)|^2$.

Chewing energy: Chewing energy is computed by first computing the frequency spectrum of the signal using Fast Fourier Transform (FFT) for the chewing frequency range (1.25-2.5 Hz). The chewing energy is then computed for this frequency range, which is the sum of the squared magnitudes of the frequency components within the range. This feature can be useful for detecting chewing activity in signals, as chewing typically produces a characteristic pattern of energy in the frequency domain.

Chewing energy to spectrum energy ratio: The ratio between the spectrum energy and chewing energy.

Walking energy: Walking energy is computed by first computing the the frequency spectrum of the signal using Fast Fourier Transform (FFT) for the walking frequency range (2.5-10 Hz). The walking energy is then computed for this frequency range, which is the sum of the squared magnitudes of the frequency components within the range.

Walking energy to spectrum energy ratio: The ratio between spectrum energy and walking energy

Talking energy: Talking energy is computed by first computing the frequency spectrum of the signal using Fast Fourier Transform (FFT) for the talking frequency range (100-300 Hz). The talking energy is then computed for this frequency range, which is the sum of the squared magnitudes of the frequency components within the range.

Talking energy to spectrum energy ratio: The ratio between spectrum energy and talking energy.

Peak frequency of chewing spectrum: Maximum frequency value of the chewing spectrum.

Peak frequency of walking spectrum: Maximum frequency value of the walking spectrum.

5.2.4.2 Feature Selection and Classification

To classify each epoch as food intake and no food intake Linear Support Vector Machine (SVM) models were chosen due to their speed and generalization ability [54]. Two separate SVM models were trained using the controlled experiment data for comparison: 1. CL_{P+ACC}: Classifier using all 80 features from both pressure sensor and accelerometer, 2. CL_P : Classifier using 20 features from only a pressure sensor. To reduce the computational burden and redundancy in the computed features, the feature selection procedure was used before training. The feature selection using minimum Redundancy and Maximum Relevance (mRMR) was applied to the computed features [55] to select the top-ranked 12 features the CL_{P+ACC}. mRMR is a feature selection algorithm that seeks to identify the most relevant features for a given task while reducing the redundancy among them. The selected top-ranked 12 features are listed in Table 5-2. The top-ranked 8 features out of 20 were selected for the CLP. Training of the models was performed using the Classification Learner tool in MATLAB 2019b (from Mathworks Inc.). The foot pedal signals were used to assign labels to each epoch during training. If more than 50%

of an epoch belonged to food intake, the epoch i was assigned a label Ep_i = '1' (food intake), otherwise the label Ep_i = '0' (no food intake).

Table 5-2: Description of selected features for CL_{P+ACC}

Feature Number	Feature Name	Sensor
1.	Mean Time Between Zero Crossing	Pressure
2.	Maximum Absolute Value	AccZ
3.	Chewing Energy to Spectrum Energy Ratio	AccY
4.	Chewing Energy	AccY
5.	Mean absolute value	Pressure
6.	Maximum Chewing Spectrum	Pressure
7.	Median absolute value	AccY
8.	Mean to RMS ratio	Pressure
9.	Spectrum Energy	AccZ
10.	Walking Energy to Spectrum Energy Ratio	AccZ
11.	Chewing Energy to Spectrum Energy Ratio	AccZ
12.	Maximum Chewing Spectrum	AccY

The SVM models were trained using a leave-one-out cross-validation procedure where the classifier was trained with data from 9 participants, and data from the remaining participant was used for testing the trained SVM model. This procedure was repeated 10 times such that each participant was used for testing once. The free-living data collected from the participants were never used in the training of the classifiers. F1 score was computed for each iteration as:

$$F1 = \frac{2 * Precision * Recall}{Precision + Recall} \quad (2)$$

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

where TP, FP, and FN denote true positives, false positives, and false negatives, respectively.

5.2.4.3 Chew Count Estimation

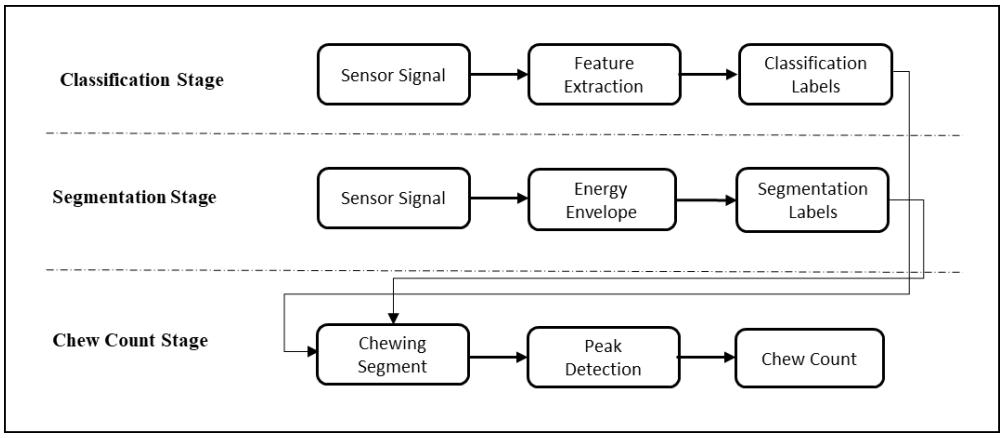


Figure 5-4: Three stages of the proposed algorithm to count the number of chews.

The three stages in the proposed method are depicted in Figure 5-4. In the classification stage, if an epoch is classified as food intake it is labeled as 1, and 0 otherwise. Here the epoch has a fixed length such as 1 sec, 3 sec, 5 sec, etc. But the chewing period of humans has a variable length. As chewing sequences are of various lengths of time, the use of fixed-length epochs may result in partial chewing or non-chewing at all. Therefore, to detect the boundaries of chewing segments of variable length, segmentation of the signal using short-time energy may be used. However, these high-energy segments can also be produced by physical activities such as talking, walking, etc.; therefore, the classification labels were used to differentiate chewing segments from other activities. The number of chews was calculated by detecting the pressure

signal peaks in the chewing segments. In the segmentation stage, first, the pressure sensor signal was filtered by a low-pass filter with a cutoff of 3 Hz to remove unwanted noise. Second, the energy envelope of the signal was computed using a sliding window (Hanning) to identify the boundaries of the signal segments with high energy (potential chewing events). The Hanning window is a type of tapering function that smoothly transitions from zero at the beginning and end of the segment to one in the middle. This smooth transition reduces the amount of spectral leakage, which occurs when energy from the signal leaks into adjacent frequency bins in the frequency domain. The Hanning window also has relatively low side lobes, which helps to reduce the amount of energy that is spread out from the main spectral peak. As the chewing frequency is located in the frequency band of 0.94 to 2.17 Hz, to capture the energy of chewing a window size of 1 second with an overlap of 0.5 seconds was used to compute the short-time energy. The value of each window was then stored in a vector and a threshold value of short-time energy (T) was dynamically calculated using the formula $T = \frac{W.M1+M2}{W+1}$ where M1 and M2 are the positions of the first and second local maxima respectively, and W is a user-defined parameter. Larger values of W lead threshold values closer to the first local maxima, M1. The ideal value of W (w=1) was empirically found from the grid search experiment. Successive windows for which the computed energies were higher than T, were combined to form high-energy segments that were considered candidate chewing segments labeled as 1 and successive windows with lower energies compared to T were combined to form low-energy segments and labeled as 0. To remove high energy segments produced by physical activities such as talking, walking, logical AND operation was performed between labels from the food intake classification prediction stage and labels from the segmentation stage. The identified chewing segments were then finally labeled as 1 for chewing segments and 0 for non-chewing segments.

The stages of the proposed algorithm are depicted in Figure 5-4. The chew counts for all identified chewing segments were computed by finding peaks and counting them. A peak was defined as the highest point around which there are points lower by X (threshold=0.5, found from empirical results) on both sides. The number of peaks in the chewing segment represents the number of chews. Consequently, adding chews of all chewing segments of an eating episode provide the total count of chewing during the meal. Figure 5-5, represents the candidate chewing sequence identification process. The second top figure represents the segmentation labels. It can be seen that physical activities such as talking, and walking produce false candidate chewing sequences which are removed using the food intake classification labels represented in the second from the bottom graph.

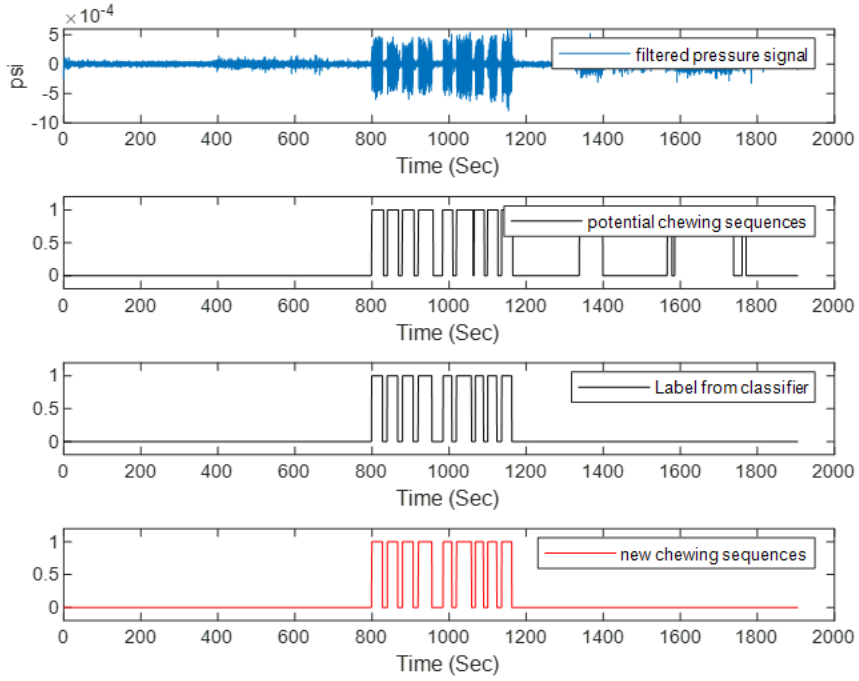


Figure 5-5: Chewing segment detection using classification labels and segmentation. The topfigure represents the filtered pressure sensor signal. The second top figure represents the label from segmentation. The third from the top figure represents the label from food intake classification. The bottom figure represents the detected chewing sequence.

5.2.4.4 Meal Detection

To detect meals from food intake detection in free-living, the predictions were kernel smoothed with a width of 50 seconds and applied a threshold value of 0.2. The value of the threshold was empirically selected from the controlled environment food intake detection results. The width and the threshold value of kernel smoothing were carefully selected as described above to detect meals and remove false positives but not to remove short eating episodes (approximately 1-minute duration). Figure 5-6 represents the prediction of the classifier for a free-living dataset and finding the boundaries of the eating episodes using CL_{P+ACC}.

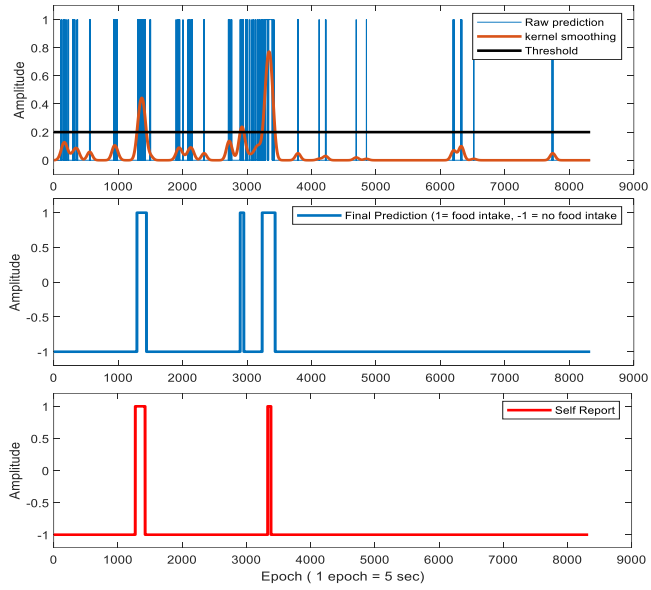


Figure 5-6: This figure represents food intake detection in free-living using pressure sensor signal and accelerometer features for a 5-sec epoch size. The top figure represents the raw prediction. The epoch is labeled as '1' if it is predicted as a food intake epoch. Otherwise, epoch is labeled as '-1'. Kernel smoothing was performed to remove false positives applying fixed threshold. The middle and bottom figure represents the final prediction and the self-reported start and ends-time of the eating episode respectively. The epoch value is set as '1' if epoch is classified as food intake. Otherwise, the epoch value is set as -1.

5.2.4.5 Image Capture

The long-term goal of the proposed ECPS system is to automatically, without any active intervention from the wearer, capture and store images only during eating. The sensor detection of food intake may facilitate the capacity of capturing and storing images only during food intake by triggering the camera. In this study for validation, the images were captured continuously every 10 secs throughout the day. The recorded images were reviewed manually to identify the true eating episodes including the start and end time of the meal in the free-living environment to match with the self-reported events. Subjects were allowed to review and delete any picture before being reviewed by a research assistant. When no images of food items were detected within a sensor-detected eating episode then it is counted as a false eating episode. The review of captured images also provides information about the number of images captured only during food intake and the number of images captured throughout the day. This information provides an idea of how much power, storage, and computational cost can be avoided by triggering the camera from the food intake detection.

5.3 Results

Results of the leave-one-out cross-validation procedure for the described two classifiers are shown in the following Tables. Table 5-3 represents the performance of the trained CL_{P+ACC} for the different epochs. For a 5-sec epoch length, the CL_{P+ACC} achieved a mean F1-score of 88.7%. Table 5-4 represents the performance matrices of the trained CL_P for different epoch sizes. For an epoch size of 5 sec, the classifier CL_P achieved a mean F1-score of 87.6%. Table 5-5 represents the performance of the chew counting algorithm used CL_P in controlled experiments. The chew counting algorithm achieved an accuracy of 95.5%.

Table 5-3: Precision, Recall (Sensitivity), and F1-Score for different epochs of Laboratory Experiments by classifier CL_{P+ACC}

Epoch (Sec)	Precision	Sensitivity	Specificity	Accuracy	Fscore
1	0.855	0.954	0.954	0.858	0.846
2	0.873	0.957	0.957	0.876	0.865
3	0.888	0.960	0.960	0.885	0.875
5	0.893	0.959	0.959	0.899	0.887
10	0.881	0.952	0.952	0.889	0.876
15	0.909	0.960	0.960	0.918	0.908
30	0.931	0.964	0.964	0.938	0.928
Mean	0.890	0.958	0.958	0.895	0.884
Std	0.023	0.004	0.004	0.024	0.025

Table 5-4: Precision, Recall (Sensitivity), and F1-Score for different epochs of Laboratory Experiments by classifier CL_P

Epoch (Sec)	Precision	Sensitivity	Specificity	Accuracy	Fscore
1	0.869	0.959	0.959	0.846	0.834
2	0.865	0.960	0.960	0.834	0.820
3	0.899	0.967	0.967	0.862	0.851
5	0.896	0.964	0.964	0.886	0.876
10	0.861	0.953	0.953	0.871	0.863
15	0.891	0.961	0.961	0.902	0.897
30	0.906	0.958	0.958	0.914	0.910
Mean	0.884	0.960	0.960	0.873	0.865
Std	0.018	0.004	0.004	0.029	0.033

Table 5-6 indicates the food intake episodes detection by using both types of classifiers in the free-living environment. Table 5-7, represents the analysis of manual image review in the free-living experiment. Only 3.7% of the captured images are relevant to food intake episodes in the free-living environment.

Table 5-5: Chew count of each subject using CL_P classifier.

Subject	Manual Chew Count	Automatic Chew Count	Accuracy %
1	673	709	94.7
2	1002	1064	93.8
3	522	557	93.8
4	545	576	93.3
5	384	364	94.3
6	985	1042	94.8
7	874	883	94.2
8	454	445	98.97
9	200	206	97
10	761	773	98.423
Mean	660.67 ± 273.6	686 ± 291.4	95.5 ± 2.09

Table 5-6: Eating episode detection of free-living data by both classifiers.

		Image Review	Self-report (atime-logger)	Using CL_P			Using CL_{P+ACC}		
Sub	Duration (Hour)	No. of meal	No. of meal	No. of (only pressure)	Missed Meal	False detection	No. of Meal	Missed Meal	False Detection
sub1	10.7	2	2	2	0	0	2	0	0
sub2	11.8	2	2	2	0	0	3	0	1
sub3	8.8	2	2	2	0	0	2	0	0

Table 5-7: Manual review of captured images (Free-Living)

Participant	data collected (hours)	Total number of Images	Food Images	% of food images
1	10.7	3852	134	3.48 %
2	11.8	4248	158	3.72 %
3	8.8	3168	125	3.95 %
Total	31.3	11268	417	3.7 %

5.4 Discussion

The necessity of the development of an accurate, easy-to-use, ingestive behavior monitoring device has remained an area of research to improve as the number of diseases caused by obesity, overweight, and eating disorders is increasing day by day. In this study, the developed ECPS food intake monitoring device was validated in both controlled and free-living environments. Features from both pressure sensor and accelerometer were collected and two separate classifier models were trained. Feature selection based on mutual information using mRMR was performed before training for both classifiers to reduce computation burden, redundancy, and overfitting. During training leave, one cross-validation approach was adopted which enables data from test participants to never be seen during training. The classifier CL_{P+ACC}uses both pressure sensor and accelerometer features, achieving the largest F-score of 0.92 with a 30-sec epoch size and the smallest F-score of 0.846 with a 1-sec epoch size. As during free-living our primary goal is to record all eating episodes successfully, a 5-second epoch size was selected to use in free-living considering precision, sensitivity, and F-score. Table 5-6 represents the classifier CL_{P+ACC} performance in free living. The classifier CL_P trained with

only pressure sensor achieved the highest F1-score of 0.91 with a 30-sec epoch size and the lowest F1-score of 0.82 for a 2-sec epoch size. A 5-sec epoch size was selected to use in free-living considering precision, sensitivity, and F-score. The classifier CL_P accurately detected all 6 eating episodes from 3 free-living days of data.

Using the food intake labels from the classifier and labels from the segmentation stage, chewing segments were detected and the number of chews was counted. Table 5-5 represents the performance of the chew counting algorithm using the classifier CL_P with an epoch size of 5 sec. The proposed chew counting algorithm achieved an accuracy of 95.5% using only the pressure sensor signal. Figure 5-7 represents the cumulative chew count for both the proposed automatic chew counting algorithm and manual annotation of one controlled environment meal.

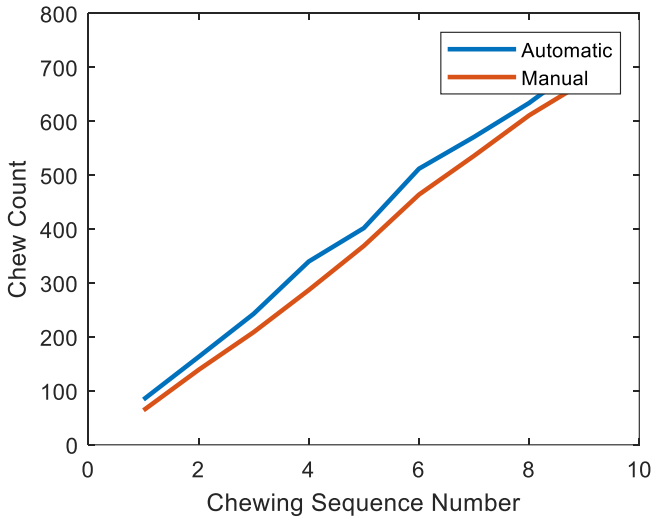


Figure 5-7: Cumulative chew count of a meal. The blue line indicates result of the proposed chew counting algorithm and the red line represents the result of the manual annotation.

The manual image review of the captured images of free-living shows that a total of 11268 images were captured during free-living: among them, only 417 images were captured during food intake which is a mere 3.7%.

The designed ECPS system has some limitations such as both sensor signals depend on chewing which detects food intake only during solid and semi-solid food consumption. Also, if the camera captures images continuously, the wearer and the people surrounding the wearer become somewhat concerned about their privacy. There could be some sources of interference such as speaking, facial movement, head movement, air tube movement, touching the earbud, etc. To remove these interferences, the pressure sensor signal was filtered by applying a low-pass filter with a cut-off frequency of 3 Hz as typical chewing is associated with a frequency in the range of 0.94 to 2.17 Hz. Moreover, these interferences are random and have different characteristics than chewing. The design of the air tube minimizes motion artifacts by maintaining the circular cross-section of the tube under moderate tension, compression, and bending of the tube due to head motion. The cross-section diameter of the acoustic tube (2mm) nearly perfectly matched the cross-section of the pressure sensor (2mm), which together with virtually zero flow in the tube allows to avoid pressure variations across the tube due to Bernoulli's principle.

As the goal of this study was to develop a novel Ear Canal Pressure sensor and to validate the sensor system for objective detection of food intake, the study was limited to only 10 volunteers. Using the data, we were able to show that the noble Ear Canal Pressure Sensor can detect food intake events with high precision and count the number of chews in the meal with high accuracy. A relatively small number of participants is a limitation of this study; however, the sensor was validated in unrestricted free-living conditions, thus supporting the rigor of the proposed sensor. A study with a larger sample size to further validate the device after further miniaturization of the sensor should be done in the future.

LOSO-CV is a cross-validation technique that is particularly useful in human subject research, where data is often correlated with individuals. It evaluates model performance on a test set that includes data from a single subject while training on data from all other subjects. This allows for a more accurate estimate of model generalization performance and is more sensitive to individual differences in the data. However, LOSO-CV can be computationally expensive and performance estimates can have high variance. The choice of validation technique depends on the specific needs of the research question and the nature of the data.

In this study, all the processing was done offline in MATLAB 2019b (from Mathworks Inc.) installed on a Windows 10 computer equipped with an Intel Core i7 9th Generation CPU with 16GB DDR4 RAM, and an NVIDIA GeForce GTX 1070 GPU with 8GB memory. The trained SVM classifier CL_{P+ACC} contains 77 support vectors. The Classifier CL_{P+ACC} takes 12 time and frequency features to predict. The average execution time to calculate 12 features and make inference in a 5 second epoch is 0.0022 second. Calculating these features and providing prediction from the micro-controller in real-time will introduce delay as the wearable device has limited computational resources. Another way of providing real-time feedback can be done from the server. Date from the ECPS may be sent to the server via Bluetooth which will be then processed by the server where the trained models will be stored and implemented. As, the aim of this study was to validate the novel ECPS for food intake monitoring and chew counting, real-time application was not considered in this study. Future work should be done to implement the ECPS for food intake monitoring and providing feedback in real-time.

As the goal of this study to validate the developed ECPS (Ear Canal Pressure Sensor) for food intake monitoring and chew counting, the food choice of was not restricted in any way. But

to develop future models such as caloric energy estimation using regression, the food items including the consumed weight was recorded.

Future work should implement image-based detection and combine it with sensor-based food intake detection to provide more reliable and accurate ingestive behavior monitor including liquid consumption. Meal microstructure such as chewing rate, bite count, bite rate, etc. should be calculated to perform an estimation of caloric consumption. The sensor detection may be useful in triggering the camera to capture images when food intake is detected by the sensor thus increasing the battery, storage, and computation capability.

The developed novel wearable sensor system can be implemented in other devices such as headphones and shows promising results in food intake detection in both controlled and wild environments. Further, the proposed wearable sensor system can be used to trigger the camera to start capturing images when food intake is detected and stop capturing after an eating episode. Thus, providing a significantly less number of captured images which reduces: the sensor system power consumption, privacy concern of the wearer and surrounding people, storage cost, and computational cost.

5.5 Conclusion

In this paper, we propose a novel wearable sensor system that can be implemented in other devices with an earbud that detects all eating episodes using features from only the ear canal pressure sensor and using both pressure sensor and accelerometer features. In Leave One Subject Out (LOSO) cross-validation experiments, an average of 88.4% F-score in food intake detection using both pressure sensor and accelerometer was achieved for 5 seconds epoch size in a controlled environment and detected all six eating episodes in free-living with only one false eating episode. The SVM classifier trained using features from only pressure sensors achieved an average of 86.5% F-score in a controlled environment and detected all six eating episodes in

free-living. The proposed chew counting algorithm achieved an accuracy of 95.5% in a controlled environment using only a pressure sensor signal. A manual review of the continuously captured egocentric images showed that a mere 3.7% of the captured images were related to food intake. These results show that the proposed wearable sensor system can effectively detect food intake episodes and count the number of chews in both controlled and free-living conditions, which needs further exploration to be used as an ingestive monitoring tool and sensor-based triggered passive image capture tool.

References

[1] C. M. Hales, “Prevalence of Obesity and Severe Obesity Among Adults: United States, 2017–2018,” no. 360, p. 8, 2020.

[2] D. E. Wilfley, M. B. Schwartz, E. B. Spurrell, and C. G. Fairburn, “Using the eating disorder examination to identify the specific psychopathology of binge eating disorder,” International Journal of Eating Disorders, vol. 27, no. 3, pp. 259–269, 2000, doi: 10.1002/(SICI)1098-108X(200004)27:3<259::AID-EAT2>3.0.CO;2-G.

[3] N. E. Day, N. McKeown, M. Y. Wong, A. Welch, and S. Bingham, “Epidemiological assessment of diet: a comparison of a 7-day diary with a food frequency questionnaire using urinary markers of nitrogen, potassium and sodium,” Int J Epidemiol, vol. 30, no. 2, pp. 309–317, Apr. 2001, doi: 10.1093/ije/30.2.309.

[4] D. A. Schoeller, L. G. Bandini, and W. H. Dietz, “Inaccuracies in self-reported intake identified by comparison with the doubly labelled water method,” Can. J. Physiol. Pharmacol., vol. 68, no. 7, pp. 941–949, Jul. 1990, doi: 10.1139/y90-143.

[5] “Accuracy of Energy Intake Data Estimated by a Multiplepass, 24-hour Dietary Recall Technique - ScienceDirect.” https://www.sciencedirect.com/science/article/pii/S000282230000095X?casa_token=Q0wLgsCgxQYAAAAA:Wm_MBJk_47ybintX4BMZrAHNsrkfjIUKOtBIIzkeELqwiKcfTlMCFaeh2jFEmp2ZSQ3zVoC7 (accessed Apr. 13, 2020).

[6] S. Bi et al., “Auracle: Detecting Eating Episodes with an Ear-mounted Sensor.” Association for Computing Machinery, Sep. 18, 2018. Accessed: Apr. 13, 2020. [Online]. Available: https://doi.org/10.1145/3264902

[7] H. Heydarian, M. Adam, T. Burrows, C. Collins, and M. E. Rollo, “Assessing Eating Behaviour Using Upper Limb Mounted Motion Sensors: A Systematic Review,” Nutrients, vol. 11, no. 5, Art. no. 5, May 2019, doi: 10.3390/nu11051168.

[8] D. Hossain, M. H. Imtiaz, and E. Sazonov, “Comparison of Wearable Sensors for Estimation of Chewing Strength,” IEEE Sensors Journal, vol. 20, no. 10, pp. 5379–5388, May 2020, doi: 10.1109/JSEN.2020.2968009.

[9] Y. Dong, A. Hoover, J. Scisco, and E. Muth, “A New Method for Measuring Meal Intake in Humans via Automated Wrist Motion Tracking,” Appl Psychophysiol Biofeedback, vol. 37, no. 3, pp. 205–215, Sep. 2012, doi: 10.1007/s10484-012-9194-1.

[10] Y. Dong, J. Scisco, M. Wilson, E. Muth, and A. Hoover, “Detecting Periods of Eating During Free-Living by Tracking Wrist Motion,” IEEE Journal of Biomedical and Health Informatics, vol. 18, no. 4, pp. 1253–1260, Jul. 2014, doi: 10.1109/JBHI.2013.2282471.

[11] H. Kalantarian and M. Sarrafzadeh, “Audio-based detection and evaluation of eating behavior using the smartwatch platform,” Computers in Biology and Medicine, vol. 65, pp. 1–9, Oct. 2015, doi: 10.1016/j.compbiomed.2015.07.013.

[12] M. H. Rahman et al., “Food Volume Estimation in a Mobile Phone Based Dietary Assessment System,” in 2012 Eighth International Conference on Signal Image Technology and Internet Based Systems, Nov. 2012, pp. 988–995. doi: 10.1109/SITIS.2012.146.

[13] W. Jia et al., “Accuracy of food portion size estimation from digital pictures acquired by a chest-worn camera,” Public Health Nutrition, vol. 17, no. 8, pp. 1671–1681, Aug. 2014, doi: 10.1017/S1368980013003236.

[14] S. Päßler and W.-J. Fischer, “Food Intake Monitoring: Automated Chew Event Detection in Chewing Sounds,” IEEE Journal of Biomedical and Health Informatics, vol. 18, no. 1, pp. 278–289, Jan. 2014, doi: 10.1109/JBHI.2013.2268663.

[15] J. M. Fontana, M. Farooq, and E. Sazonov, “Automatic Ingestion Monitor: A Novel Wearable Device for Monitoring of Ingestive Behavior,” IEEE Transactions on Biomedical Engineering, vol. 61, no. 6, pp. 1772–1779, Jun. 2014, doi: 10.1109/TBME.2014.2306773.

[16] M. Farooq and E. Sazonov, “Comparative testing of piezoelectric and printed strain sensors in characterization of chewing,” in 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Aug. 2015, pp. 7538–7541. doi: 10.1109/EMBC.2015.7320136.

[17] E. Sazonov et al., “Non-invasive monitoring of chewing and swallowing for objective quantification of ingestive behavior,” Physiol. Meas., vol. 29, no. 5, Art. no. 5, Apr. 2008, doi: 10.1088/0967-3334/29/5/001.

[18] O. Makeyev, P. Lopez-Meyer, S. Schuckers, W. Besio, and E. Sazonov, “Automatic food intake detection based on swallowing sounds,” Biomedical Signal Processing and Control, vol. 7, no. 6, Art. no. 6, Nov. 2012, doi: 10.1016/j.bspc.2012.03.005.

[19] J. M. Fontana and E. S. Sazonov, “Evaluation of Chewing and Swallowing Sensors for Monitoring Ingestive Behavior,” Mar. 2013. https://www.ingentaconnect.com/content/asp/senlet/2013/00000011/00000003/art00019 (accessed Apr. 14, 2020).

[20] Y. Dong, J. Scisco, M. Wilson, E. Muth, and A. Hoover, “Detecting Periods of Eating During Free-Living by Tracking Wrist Motion,” IEEE Journal of Biomedical and Health Informatics, vol. 18, no. 4, Art. no. 4, Jul. 2014, doi: 10.1109/JBHI.2013.2282471.

[21] M. Farooq and E. Sazonov, “Accelerometer-Based Detection of Food Intake in Free-Living Individuals,” IEEE Sensors J., vol. 18, no. 9, Art. no. 9, May 2018, doi: 10.1109/JSEN.2018.2813996.

[22] J. L. Scisco, E. R. Muth, Y. Dong, and A. W. Hoover, “Slowing Bite-Rate Reduces Energy Intake: An Application of the Bite Counter Device,” Journal of the American Dietetic Association, vol. 111, no. 8, pp. 1231–1235, Aug. 2011, doi: 10.1016/j.jada.2011.05.005.

[23] M. Zandian, I. Ioakimidis, C. Bergh, U. Brodin, and P. Södersten, “Decelerated and linear eaters: Effect of eating rate on food intake and satiety,” Physiology & Behavior, vol. 96, no. 2, pp. 270–275, Feb. 2009, doi: 10.1016/j.physbeh.2008.10.011.

[24] C. Lepley, G. Throckmorton, S. Parker, and P. H. Buschang, “Masticatory Performance and Chewing Cycle Kinematics—Are They Related?,” Angle Orthod, vol. 80, no. 2, pp. 295–301, Mar. 2010, doi: 10.2319/061109-333.1.

[25] H. A. Raynor et al., “Partial Meal Replacement Plan and Quality of the Diet at 1 Year: Action for Health in Diabetes (Look AHEAD) Trial,” Journal of the Academy of Nutrition and Dietetics, vol. 115, no. 5, pp. 731–742, May 2015, doi: 10.1016/j.jand.2014.11.003.

[26] “Improvement in chewing activity reduces energy intake in one meal and modulates plasma gut hormone concentrations in obese and lean young Chinese men1–3 | The American Journal of Clinical Nutrition | Oxford Academic.” https://academic.oup.com/ajcn/article/94/3/709/4411851 (accessed Jun. 02, 2020).

[27] M. Farooq and E. Sazonov, “Segmentation and Characterization of Chewing Bouts by Monitoring Temporalis Muscle Using Smart Glasses With Piezoelectric Sensor,” IEEE Journal of Biomedical and Health Informatics, vol. 21, no. 6, pp. 1495–1503, Nov. 2017, doi: 10.1109/JBHI.2016.2640142.

[28] O. Amft*, M. Kusserow, and G. TrÖster, “Bite Weight Prediction From Acoustic Recognition of Chewing,” IEEE Transactions on Biomedical Engineering, vol. 56, no. 6, pp. 1663–1672, Jun. 2009, doi: 10.1109/TBME.2009.2015873.

[29] E. S. Sazonov and J. M. Fontana, “A Sensor System for Automatic Detection of Food Intake Through Non-Invasive Monitoring of Chewing,” IEEE Sensors Journal, vol. 12, no. 5, pp. 1340–1348, May 2012, doi: 10.1109/JSEN.2011.2172411.

[30] A. Bedri, A. Verlekar, E. Thomaz, V. Avva, and T. Starner, “Detecting Mastication: A Wearable Approach,” in Proceedings of the 2015 ACM on International Conference on Multimodal Interaction, Seattle, Washington, USA, Nov. 2015, pp. 247–250. doi: 10.1145/2818346.2820767.

[31] E. Thomaz, C. Zhang, I. Essa, and G. D. Abowd, “Inferring Meal Eating Activities in Real World Settings from Ambient Sounds: A Feasibility Study,” in Proceedings of the 20th International Conference on Intelligent User Interfaces, Atlanta, Georgia, USA, Mar. 2015, pp. 427–431. doi: 10.1145/2678025.2701405.

[32] O. Amft, “A wearable earpad sensor for chewing monitoring,” in 2010 IEEE SENSORS, Nov. 2010, pp. 222–227. doi: 10.1109/ICSENS.2010.5690449.

[33] A. Natarajan, G. Angarita, E. Gaiser, R. Malison, D. Ganesan, and B. M. Marlin, "Domain adaptation methods for improving lab-to-field generalization of cocaine detection using wearable ECG," in Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, Heidelberg, Germany, Sep. 2016, pp. 875–885. doi: 10.1145/2971648.2971666.

[34] M. Shuzo et al., "Wearable Eating Habit Sensing System Using Internal Body Sound," Journal of Advanced Mechanical Design, Systems, and Manufacturing, vol. 4, no. 1, pp. 158–166, 2010, doi: 10.1299/jamdsm.4.158.

[35] K. Kohyama, E. Hatakeyama, T. Sasaki, T. Azuma, and K. Karita, "Effect of sample thickness on bite force studied with a multiple-point sheet sensor," Journal of Oral Rehabilitation, vol. 31, no. 4, pp. 327–334, 2004, doi: 10.1046/j.1365-2842.2003.01248.x.

[36] V. A. Bousdras et al., "A novel approach to bite force measurements in a porcine model in vivo," International Journal of Oral and Maxillofacial Surgery, vol. 35, no. 7, pp. 663–667, Jul. 2006, doi: 10.1016/j.ijom.2006.01.023.

[37] K. Fueki, T. Sugiura, E. Yoshida, and Y. Igarashi, "Association between food mixing ability and electromyographic activity of jaw-closing muscles during chewing of a wax cube," Journal of Oral Rehabilitation, vol. 35, no. 5, pp. 345–352, 2008, doi: 10.1111/j.1365-2842.2008.01849.x.

[38] T. Vu, F. Lin, N. Alshurafa, and W. Xu, "Wearable Food Intake Monitoring Technologies: A Comprehensive Review," Computers, vol. 6, no. 1, Art. no. 1, Mar. 2017, doi: 10.3390/computers6010004.

[39] A. Beltran et al., "Adapting the eButton to the Abilities of Children for Diet Assessment," Proc Meas Behav 2016 (2016), vol. 2016, pp. 72–81, May 2016.

[40] L. Gemming et al., "Wearable cameras can reduce dietary under-reporting: doubly labelled water validation of a camera-assisted 24 h recall," British Journal of Nutrition, vol. 113, no. 2, pp. 284–291, Jan. 2015, doi: 10.1017/S0007114514003602.

[41] V. Bruno, S. Resende, and C. Juan, "A Survey on Automated Food Monitoring and Dietary Management Systems," J Health Med Inform, vol. 8, no. 3, 2017, doi: 10.4172/2157-7420.1000272.

[42] W. Jia et al., "Automatic food detection in egocentric images using artificial intelligence technology," Public Health Nutrition, vol. 22, no. 7, Art. no. 7, May 2019, doi: 10.1017/S1368980018000538.

[43] J. M. Fontana, M. Farooq, and E. Sazonov, "Automatic Ingestion Monitor: A Novel Wearable Device for Monitoring of Ingestive Behavior," IEEE Transactions on Biomedical Engineering, vol. 61, no. 6, Art. no. 6, Jun. 2014, doi: 10.1109/TBME.2014.2306773.

[44] M. Sun et al., “eButton: A Wearable Computer for Health Monitoring and Personal Assistance,” in Proceedings of the 51st Annual Design Automation Conference, San Francisco, CA, USA, Jun. 2014, pp. 1–6. doi: 10.1145/2593069.2596678.

[45] P. Kelly et al., “An Ethical Framework for Automated, Wearable Cameras in Health Behavior Research,” American Journal of Preventive Medicine, vol. 44, no. 3, pp. 314–319, Mar. 2013, doi: 10.1016/j.amepre.2012.11.006.

[46] A. R. Doherty et al., “Wearable Cameras in Health: The State of the Art and Future Possibilities,” American Journal of Preventive Medicine, vol. 44, no. 3, pp. 320–323, Mar. 2013, doi: 10.1016/j.amepre.2012.11.008.

[47] S. Safavi and Z. Shukur, “Conceptual Privacy Framework for Health Information on Wearable Device,” PLoS ONE, vol. 9, no. 12, p. e114306, Dec. 2014, doi: 10.1371/journal.pone.0114306.

[48] A. Doulah, T. Ghosh, D. Hossain, M. H. Imtiaz, and E. Sazonov, “‘Automatic Ingestion Monitor Version 2’ – A Novel Wearable Device for Automatic Food Intake Detection and Passive Capture of Food Images,” IEEE Journal of Biomedical and Health Informatics, vol. 25, no. 2, pp. 568–576, Feb. 2021, doi: 10.1109/JBHI.2020.2995473.

[49] D. Hossain, T. Ghosh, and E. Sazonov, “Automatic Count of Bites and Chews From Videos of Eating Episodes,” IEEE Access, vol. 8, pp. 101934–101945, 2020, doi: 10.1109/ACCESS.2020.2998716.

[50] M. H. Imtiaz, D. Hossain, V. Y. Senyurek, P. Belsare, S. Tiffany, and E. Sazonov, “Wearable Egocentric Camera as a Monitoring Tool of Free-Living Cigarette Smoking: A Feasibility Study,” Nicotine & Tobacco Research, vol. 22, no. 10, pp. 1883–1890, Oct. 2020, doi: 10.1093/ntr/ntz208.

[51] J. M. C. Po, J. A. Kieser, L. M. Gallo, A. J. Tésenyi, P. Herbison, and M. Farella, “Time-Frequency Analysis of Chewing Activity in the Natural Environment,” J Dent Res, vol. 90, no. 10, pp. 1206–1210, Oct. 2011, doi: 10.1177/0022034511416669.

[52] T. Ghosh, D. Hossain, M. Imtiaz, M. A. McCrory, and E. Sazonov, “Implementing Real-Time Food Intake Detection in a Wearable System Using Accelerometer,” in 2020 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES), Mar. 2021, pp. 439–443. doi: 10.1109/IECBES48179.2021.9398760.

[53] M. Farooq and E. Sazonov, “Accelerometer-Based Detection of Food Intake in Free-Living Individuals,” IEEE Sensors Journal, vol. 18, no. 9, pp. 3752–3758, May 2018, doi: 10.1109/JSEN.2018.2813996.

[54] C. Cortes and V. Vapnik, “Support-vector networks,” Mach Learn, vol. 20, no. 3, Art. no. 3, Sep. 1995, doi: 10.1007/BF00994018.

[55] Hanchuan Peng, Fuhui Long, and C. Ding, “Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, no. 8, Art. no. 8, Aug. 2005, doi: 10.1109/TPAMI.2005.159.ER

CHAPTER 6 – DEVELOPMENT OF ENERGY INTAKE MODEL FROM SENSOR AND IMAGE CAPTURED EATING IN FREE LIVING

Abstract

This study presents a novel methodology for training and evaluating the energy intake estimation model using the Automatic Ingestion Monitor v2 (AIM-2), a wearable device equipped with a novel contactless optical sensor to track temporalis muscle movement and accelerometer mounted on eyeglasses. The research contribution encompasses three key aspects: 1) the Development of an Energy Intake (EI) estimation model based on accelerometer and novel optical sensor signals using data from a free-living study, 2) the evaluation of the developed EI model using an independent free-living dataset, and 3) Improvement of the EI estimation accuracy through identification of eating events from images.

The EI model was trained using a dataset comprising 30 volunteers who consumed unrestricted meals while maintaining a Weighted Food Record (WFR) for two days. The developed regression model ($R^2 = 0.47$) reported a mean error of 42.2% ± 53.4% in leave-one-subject-out validation against WFR.

Subsequently, the EI estimation model was evaluated on 3 days of data collected from 39 subjects engaged in unrestricted free-living activities over a week. The blind EI estimation method (based only on sensor signals) yielded an estimated mean/std EI of 2423.66 ± 1416.04$Kcal/day$. In comparison, EI estimated from image review by trained nutritionists resulted in 1218.51 ± 586.26 $Kcal/day$. The image-assisted EI estimation where eating episode

detection by sensors was further refined through a manual image review resulted in 2041.85 ± 1048.22 $Kcal/day$. The blind and image-assisted EI estimation methods provided a closer match than the image review to the Estimated Energy Requirement (EER) calculated from participants' demographic information (age, sex, height, and weight) with the mean value of 2663.57 ± 398.86 $Kcal/day$.

The findings highlight the feasibility of the sensor-based EI estimation and underscore the potential for the combination of sensor and image information in EI estimation.

6.1 Introduction

Despite rising life expectancy, the global increase in obesity has led to a significant spike in deaths from non-communicable diseases like diabetes [1]. Intensive lifestyle intervention programs have shown that even a moderate weight loss of 5% can bring notable health benefits [2]. These interventions typically involve dietary changes, increased physical activity, and behavior modification [3], [4], [5], [6]. By adopting a hypocaloric diet and engaging in regular exercise, individuals can achieve meaningful weight loss and improve their overall health outcomes [7], [8], [9]. Monitoring dietary intake and energy intake (EI) is crucial for effective management of obesity and related conditions [2], [10].

Common methods for assessing dietary intake include food records, food frequency questionnaires, and 24-hour recalls [11]. While these methods offer insights into food consumption patterns, they rely on individuals' memory and may lead to inaccuracies [12], [13], [14]. Analysis of data from the National Health and Nutrition Examination Survey (NHANES) 2003–2012, revealed that 25.1% of the US adult population (aged 20 and above) misreported their EI in self-report [15]. While 24-hour recalls (24HR) can mitigate some of these limitations, they too suffer from recall bias, including food omission, forgetfulness, and errors in portion size estimation [16]. Smartphone apps like MyFitnessPal [17], LoseIt [18], and MyNetDiary [19] aim

to improve accuracy, manual input, and potential user fatigue remain challenges [20]. Critics argue that these methods hinder scientific progress in obesity and nutrition research, emphasizing the need for innovative solutions [21].

Doubly labeled water [22] (DLW) is a powerful method used to estimate EI in scientific research. It involves the ingestion of specially labeled water and tracking the rate at which these labels are eliminated from the body. By measuring the changes in the body's isotopic composition over a specific period, researchers can calculate energy expenditure and, by extension, EI. While highly accurate, DLW is not practical for everyday use due to its expense, the need for specialized equipment, and the collection of biological samples, making it primarily suitable for controlled research settings rather than real-life, day-to-day applications.Top of Form As an alternative to conventional self-reporting methods and expensive scientific procedures such as DLW, there have been proposals to adopt approaches involving food images and wearable sensors placed on the body. Some research studies have explored the use of smartphones or handheld cameras to capture images of food, followed by the analysis of these images by expert nutritionists offline to estimate EI [23], [24], [25]. Yet, even these techniques depend on users taking the initiative to capture images themselves and, as a result, share limitations with traditional self-reporting methods [26]. Studies utilizing deep learning algorithms using images for food item recognition have shown promising results in recent years. These algorithms leverage neural networks to analyze images to identify and recognize food items consumed by individuals [27], [28], [29], [30], [31], [32]. In [31] authors developed a method to detect food items using Faster R-CNN from the captured egocentric images and integrated it with sensor food intake detection by a hierarchical classification to reduce false food intake detection from the sensor alone. A novel transformer-based architecture was designed to

caption egocentric dietary images by recognizing food images from captured egocentric images to monitor eating behavior [32].

Rapid advancements in technology have provided the tools to develop accurate and objective systems for measuring dietary intake using wearable sensors. These wearable devices for dietary intake monitoring include sensors on the head or neck to detect chewing or swallowing [25], [33], [34], [35], [36], [37], wrist-based inertial sensors to detect hand-to-mouth gestures as a proxy for bites [38], [39], [40], [41] and other [42]. Wearable sensors for assessing EI have the potential to supply rich information about eating behavior patterns of food intake by capturing timing, duration, rate of ingestion, and ingested mass without self-reporting while reducing user burden. Using the sensor features related to the eating behavior of individuals researchers are trying to develop different algorithms to estimate EI from food intake events. The bite counter [40] using a gyroscope to automatically track wrist motion was used for the development of an equation correlating kilocalories consumed per bite with demographic and physical factors, enabling the estimation of EI. In a study involving 280 participants consuming a single meal in a cafeteria, the device outperformed human estimates of EI.

The study of [44], presented an EI estimation model from counts of chews and swallows (CCS model) from a throat microphone and piezoelectric strain sensor placed on the jaw muscle in the laboratory setting. The development and validation of the mathematical EI model was based on stationary equipment and could never be used in free-living conditions. In [42], authors described a neck-worn device to estimate EI from a proximity sensor placed around the neck. Multiple linear regression models were then trained for each consumed food type. However, due to the need for separate regression models for each food type, this method isn't practical for most people in real-life situations. Hence, there's a need for a more accurate sensor-based algorithm

for EI estimation in free-living conditions, reducing user burden and combating obesity effectively.

The original Automatic Ingestion Monitor Version V2 (AIM-2) was based on a flexible bend sensor in contact with the temporalis muscle and an accelerometer mounted on frame of eyeglasses. The AIM-2 detects food intake and uses its camera to capture egocentric images of food eaten [37]. In our previous work, we developed EI model from the bend sensor-derived features. Data were collected from 30 volunteers wearing the device for one day of pseudo-free-living conditions where participants attended the laboratory to consume breakfast, lunch, and dinner with WFR. The sensor-derived EI from regression modeling (R^2 = 0.331) showed close agreement with EI from WFR [43].

This study presents several important contributions: First, the EI model was developed from the data collected by the latest generation of AIM-2 device that relies on a novel contactless optical sensor as opposed to the bend sensor used in the original device. Second, the sensor-based EI model was developed entirely on unconstrained free-living data from 2 days of device wear. Third, two methods of implementation ("blind" and image-assisted) using the EI model were evaluated on data from an independent week-long free-living study. The proposed methods were compared to manual image review by trained nutritionists and EER calculated from participants' demographic information. Finally, this study explores the potential for improved estimation accuracy by leveraging information from food images.

6.2 Methods

6.2.1 Data Collection Protocol

6.2.1.1 Wearable Sensor: AIM-2 (Automatic Ingestion Monitor V2)

Figure 6-1: AIM-2 attached to the right temple of off-the-shelf eyeglasses to monitor intake. (a)original AIM-2 with flexible bend sensor contacts skin of the temporalis muscle to track movement [37] (previous study) (b) AIM-2 with an optical sensor that tracks temporalis muscle movement without any contact with the skin (current study).

In this study, a wearable sensor system with an optical sensor (Automatic Ingestion Monitor, AIM-2) was used to monitor eating episodes. During the act of eating, the temporalis muscle plays a pivotal role in the process of mastication. As food is ingested, signals from the brain prompt the contraction of the temporalis, leading to the elevation of the mandible against the maxilla [45]. This rhythmic and coordinated movement, along with contributions from muscles like the masseter, facilitates the grinding and crushing of food between the teeth. Previously the AIM-2 used a flex sensor (SpectraSymbol 2.2" flex sensor) which was in contact with the skin to track the temporalis muscle movement and a 3-axis accelerometer to track head movement in recognizing eating activity. This contact with the temporalis muscle may cause irritation and discomfort to the participant from long wear. In this study, we used a contactless way of tracking the temporalis muscle by using an optical sensor. In addition, AIM-2 includes a low-power 3D accelerometer (ADXL332 from Analog Devices, Norwood, MA, USA) to track head movement, and a miniature wide-angle lens camera (OV5640, Omnivision Inc.) to capture egocentric images. The sensor signals of a small eating event are shown in Figure 6-2. The

description of the AIM2 sensor systems can be found in our previously published study [37]. The device was designed to capture food images in front of the wearer with a gaze-aligned wide-angle lens at a rate of one image per 10 seconds while eating. To trigger the camera from the sensor (accelerometer) detected eating episodes and record images only during eating events to reduce privacy concerns, the microcontroller was programmed with a decision tree classifier [46]. All collected sensor signals and captured images were stored on an SD card and processed offline in Python (version 3.9).

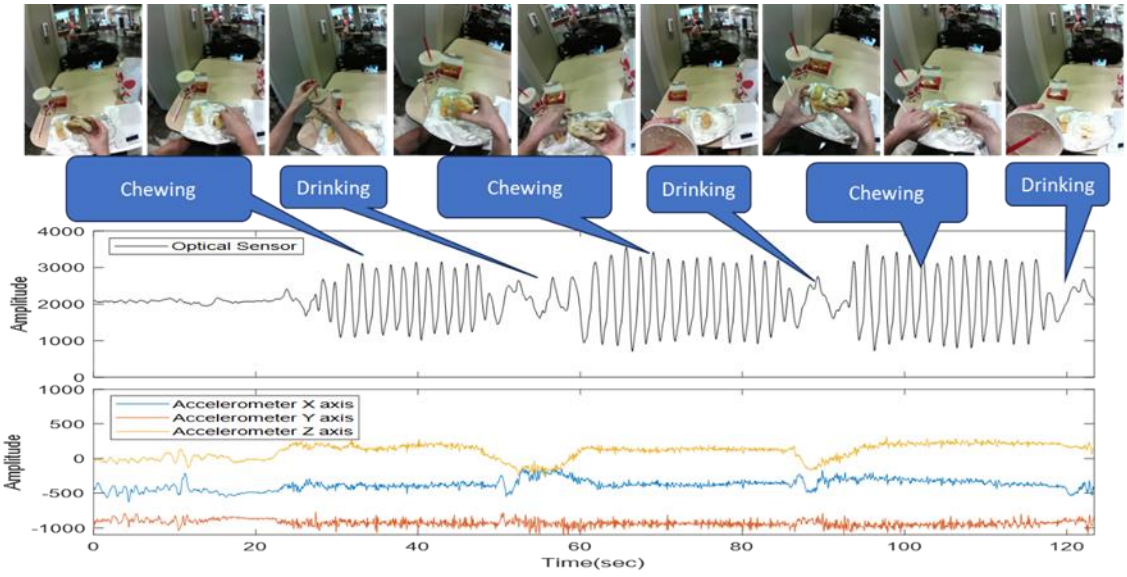


Figure 6-2: Sensor signals of an eating event. Captured images are shown at the top. Signals from the 3-axes of the accelerometer are shown on the bottom. The middle segment demonstrates the optical sensor signal.

6.2.1.2 Training Dataset

The model training dataset was gathered from 30 healthy participants aged 18–45 with a BMI range of 16.5-45.4 kg/m^2 at the University of Alabama. Recruited participants had no medical conditions that affected their ability to eat or chew food. The study received approval from the University of Alabama's Institutional Review Board (IRB), and all participants

provided informed consent for the utilization of recorded data (sensor and image) for presentations, publications, or other forms of dissemination. During an initial visit, participants were provided the opportunity to practice collecting WFR data from four training meals prepared using food replicas. These participants wore an AIM-2 device for two days, providing detailed WFR of each food item for their consumed meals and snacks in a diary. The consumed EI from the WFR diary provided by the participants was computed using I2N (Image to Nutrients), a desktop application, which is a sensor-guided semi-automated tool for nutrition analysis of eating episodes [47] by the research assistant. I2N can offer multiple food and nutrient databases (e.g., USDA-SR, FNDDS, USDA Global Branded Food Products Database) for annotating eating episodes and food items. The computed EI from WFR served as the ground truth for the development of the sensor-based regression model for EI estimation.

6.2.1.3 Evaluation Dataset

The evaluation dataset was collected at the Miriam Hospital in Providence, Rhode Island, and Boston University in Boston, Massachusetts. Approval for the studies was obtained from the Institutional Review Board at Miriam Hospital, and all participants provided written, informed consent before participating. The study included 39 healthy, non-smoking individuals aged 18–65 years, with a BMI range of 27.1 - 44.2 kg/m2, and without medical conditions affecting their eating or chewing abilities. Participants wore the AIM-2 device during waking hours for one week. Three days, ideally comprising two weekdays and one weekend day, were randomly selected for annotating the eating episodes for nutrition information from the AIM-2 captured images using the I2N semi-automated tool. Each participant was assigned a predetermined order of weekdays and weekend days for analysis, ensuring an equitable distribution across the entire dataset. To qualify for image review, a day needed to meet specific criteria: the participant must

have worn the AIM-2 device for at least 8 hours, and there should have been at least two separate eating episodes recorded. If these criteria weren't met initially, subsequent days were considered until three eligible days were analyzed.

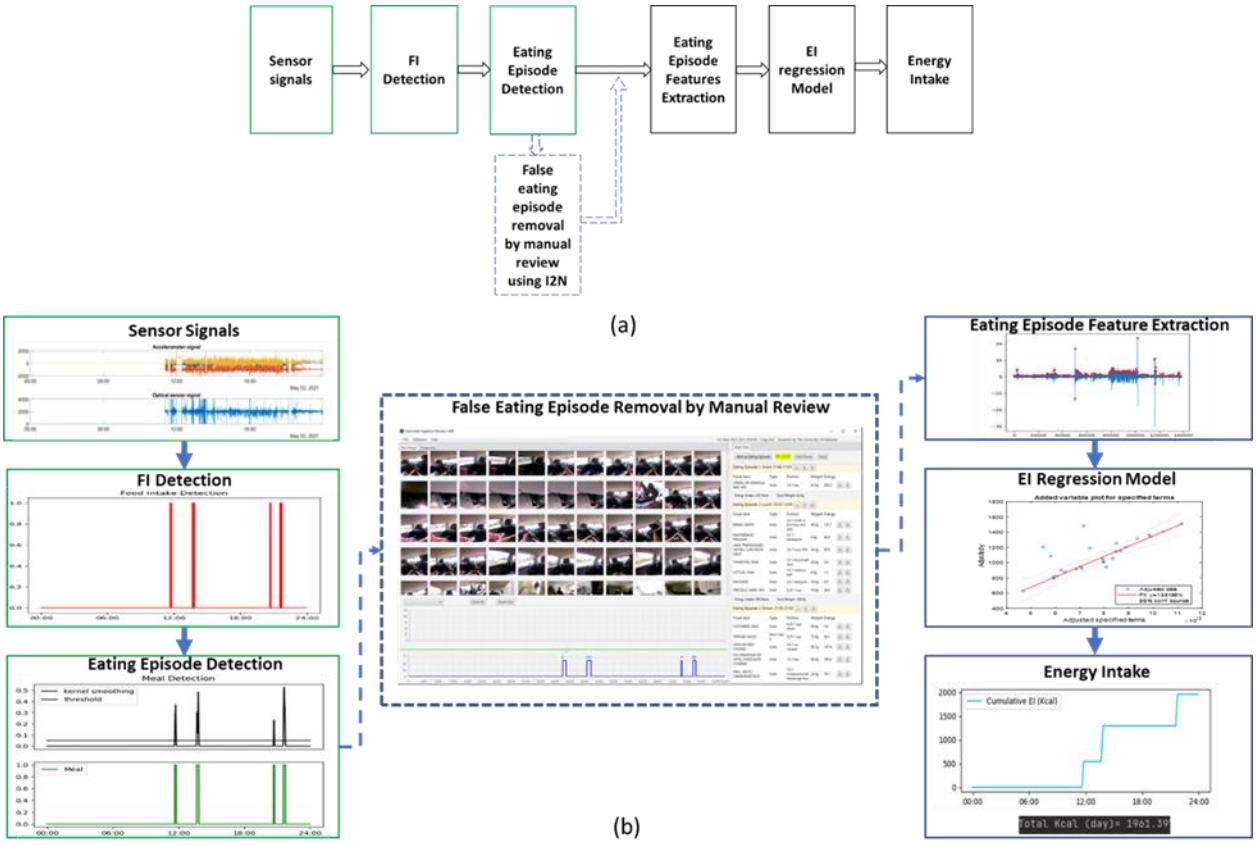


Figure 6-3: Steps of data processing for EI Estimation (a) represents the steps of the EI estimation from sensor signals and image information (b) graphical representations of EI estimation from the sensor signals and image information

The steps of data processing for EI estimation from sensor signals are shown in Figure 6-3. In the first step, AIM2 sensor signals (Accelerometer, Optical, Image) are extracted. In the next step, FI episodes are detected from accelerometer and optical sensor signals. The boundaries of the sensor-detected eating episodes are then identified from FI detection labels. These three steps of data processing are similar for both training and evaluation of the EI

estimation model. Two methods for evaluation of EI estimation was implanted. In one method, the selected feature set (found in training) for EI regression model is extracted blindly (without any information from image) from the sensors detected eating episodes. After that the sensor feature set (found in training) are extracted within the detected eating episodes. In other method, any false eating episodes detected by sensors are removed using a manual image review of the AIM2 captured images.

6.2.2 Food Intake and Eating Episode Detection

The first stage of processing is food intake detection from the raw accelerometer and sensor signal. First, the signals were partitioned into non-overlapping 8s fixed time segments referred to as epochs. Then the epochs were classified and labeled as '1' for food intake and '0' for no food intake by a classifier described in [48].

To identify eating episodes and determine both the start and end of each episode derived from food intake detection, the predictions underwent kernel smoothing as described in [43], [49].

6.2.3 EI Model Training

The steps of training the EI estimation model using AIM-2 sensor signals are described below.

6.2.3.1 Feature extraction from AIM-2 signals

Before extracting features from sensor signals to train the EI models, initial processing was applied to the raw accelerometer and optical sensor data. A high-pass filter with a cutoff frequency of 0.1 Hz was used to eliminate the DC component and to adjust inter-subject variations, the signals were then normalized. To eliminate unwanted noise from the optical sensor, particularly given the typical chewing frequency range of 0.94 to 2 Hz [50], a low-pass filter with a cutoff frequency of 3 Hz was employed. Then the AIM-2 sensor signals were divided into nonoverlapping 8-second-long time segments called "epochs". For each detected

eating episode a set of features was computed from AIM-2 sensor signals for each epoch. The feature set includes 43 time and frequency domain features (Supplemental Table 1), 10 correlation features (Supplemental Table 2), 22 features from pitch and roll (Supplemental Table 3), and 10 eating behavior features (Supplemental Table 4), resulting in 259 features for each epoch. Mean values of features were computed over each eating episode for time and frequency domain features, correlation features, and pitch and roll features. The eating behavior features were summed over each eating episode and utilized to train the regression model.

6.2.3.2 Feature Selection

To train the EI model, feature selection was applied to reduce the computational burden and redundancy in the computed features.

First, to rank the feature based on importance, an algorithm called Least Absolute Shrinkage and Selection Operator (LASSO) [51] was applied to the computed features, identifying the 30 most important features. LASSO simplifies regression models by selecting features through regularization, controlled by a tuning parameter called lambda. The optimal lambda value, determining feature elimination, was found via cross-validation, selecting the largest lambda with mean square error within one standard error of the minimum MSE across a range of lambda values.

Second, a Sequential Forward Feature Selection (SFFS) [52] was applied to the most important 30 features selected by LASSO to get a further reduced feature set. SFFS assesses individual features' impact on model performance, adding the one that maximizes performance improvement during each iteration. This process continues until a stopping criterion, usually minimum mean squared error, is met. SFFS streamlines the linear regression model by

progressively incorporating the most informative features, enhancing interpretability and avoiding overfitting.

6.2.3.3 Training Linear Regression Model

The optimal minimal feature set obtained from the SFFS-based feature selection stage was utilized to fit a Multivariate Ordinary Least Squares Linear regression model with selected features of eating episodes as independent variables and EI from WFR of eating episodes as the dependent variable. Leave-One-Subject-Out Cross-Validation (LOSOCV)[45], [53] was incorporated, repeating process 30, where the model was trained on data from 29 participants and tested on the 30th participant in each iteration so that the model was tested on the participant's data who is not in training. The resulting errors from these iterations were averaged to report the mean error. Finally, for model reporting, a multivariate ordinary least squares regression model was trained using the 30 participants with selected minimal subset features for estimating the EI of each eating episode.

6.2.4 EI Model Evaluation

The evaluation dataset was used to verify the performance of the EI models. Being a completely free-living and unrestricted study, the dataset lacks the ground truth data (WFR). Thus, several methods were used for comparison. First, trained nutritionists annotated 3 days of data by reviewing food images captured by the AIM-2. Second, the Estimated Energy Requirement (EER) was computed for each participant. Third, a "blind" sensor-based EI estimation (not using any image information) was performed on the same days as manual annotation. Finally, image-assisted EI estimation from sensor data was performed. The details of each method are described below.

6.2.4.1 EI estimation from manual image review

Annotation of AIM2 recorded images of the three out of seven days of each participant from the evaluation study was performed by the nutritionists at the University of Colorado Clinical and Translational Sciences Institute, University of Colorado Anschutz Medical Campus. The annotators used the I2N to annotate true and false eating episodes using an operating procedure described in [55]. For each true eating episode, the annotator finds food items consumed from the recorded AIM-2 images. The annotator selects the desired food item from the database and its corresponding portion size (e.g., cup, tablespoon, gram, etc.). Next, the annotator estimates the initial and leftover portion size of the selected food item from the recorded AIM-2 images. After getting all these inputs from the annotator, I2N calculates the amount consumed and the energy and nutrient contents of each eating episode. These annotated project files were used to calculate daily EI consumption from manual image review which was used to evaluate the proposed two EI estimation methods.

6.2.4.2 Estimated Energy Requirement (EER)

The body's vital functions, such as respiration, circulation, physical work, and protein synthesis, require energy derived from dietary sources like carbohydrates, proteins, fats, and, to a lesser extent, alcohol. Maintaining an energy balance relies on an individual's dietary EI and expenditure. The EER, a key metric in this context, represents the average predicted dietary EI to sustain energy balance in adults based on various factors [56]. The computed EER was used as point of comparison with the EI estimates. Due to the lack of an objective measure for physical activity in the evaluation study, a 'low active' level (Physical Activity Level $\geq 1{\cdot}4$ to $< 1{\cdot}6$) [58] was used during the calculation of EER as reported in [15].

6.2.4.3 Blind EI Estimation Method

In this method, EI is estimated using only optical sensor and accelerometer signals. Any information from the AIM-2 captured images is not used (thus 'Blind' EI method). After detecting eating episodes from the sensor signals, selected features (found in training) within the detected eating episodes were extracted and provided as input into the regression model for EI estimation. The cumulative EI curve is then generated for the day by summing the estimated EI from all sensor-detected eating episodes of the day. In this method, any false food intake detection can contribute to false eating detection which results to over-estimation in cumulative EI of the day. To reduce this over-estimation, we can use the AIM-2 recorded image information to remove the false eating events.

6.2.4.4 Image-Assisted EI Estimation Method

In this method, the AIM-2 images are used to remove the false eating events (Figure 6-3). I2N software was used to mark false and true eating episodes [47] [54] based on whether images of foods or beverages were present in the detected eating events. I2N uses sensor detection of eating thus allowing to review only the images that belong to an eating event and ignore images captured in other activities of daily living. Some of these sensors detected eating episodes may be false eating episodes, meaning the food intake detection algorithm incorrectly detected food intake. If the annotator finds images that represent eating within the sensor-detected eating episode, then the annotator labels the eating episode as a true eating episode. On the other hand, if the annotator does not find any image that represents eating, then the annotator labels the eating episode as a false eating episode. The annotation of the true and false eating episodes is saved into the annotated project files. In this method, only the features from annotated true eating episodes are used to estimate daily EI.

6.3 Results

Training Results:

The optimal minimal feature set obtained from the two-stage feature selection procedure consists of 5 features. The feature set includes the number of chews, chewing duration (sec), pause duration (sec), eating episode duration (sec), and percentage of chewing in the eating episode. Using the minimal feature set, the regression model achieved a R^2 value of 0.47 in training. The equation of the regression model comprising the selected five features is as follows:

$$EI\ (Kcal) = -70.12 + 0.23 * chew_{count} + 0.25 * chewing_{dur}\ (\text{sec}) - 0.15 * \text{pause}_{\text{dur}}(\text{sec}) + 0.4 * \text{eating episode}_{\text{dur}}(\text{sec}) + 6.26 * \text{percent_chewing}$$

The Bland–Altman plots representing the mean differences and 2 SD limits of agreement between EI from WFR and EI from the prediction in LOSO-CV during training are shown in the following Figure 6-4. The figure also includes a scatter plot of EI from WFR and EI from prediction in LOSO-CV.

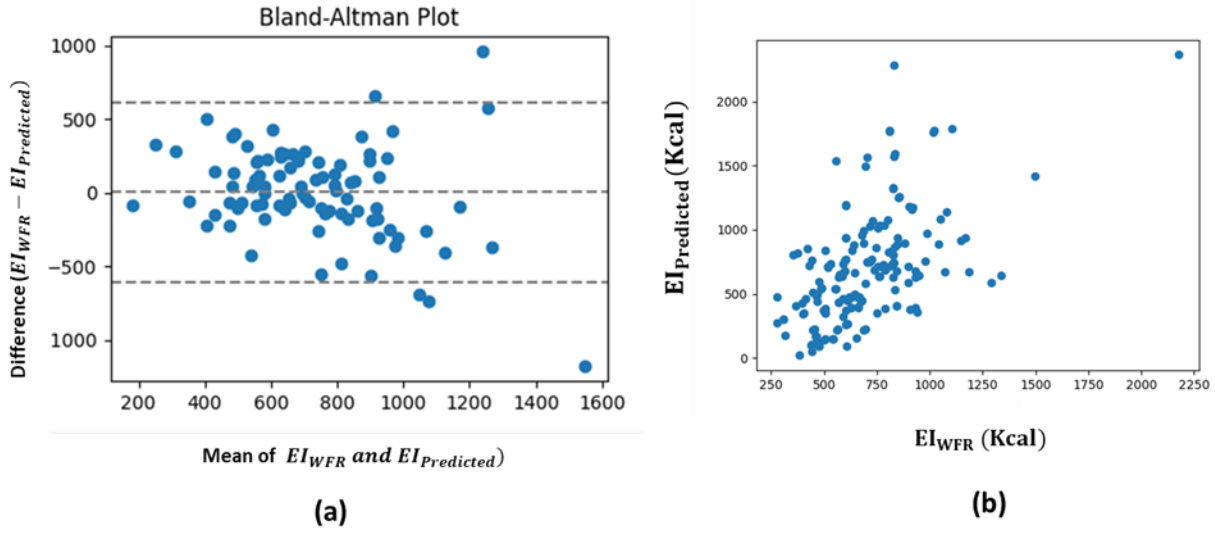


Figure 6-4: (a) Bland-Altman plot showing mean difference and 2 SD limits of agreement (b)Scatter plot of EI (WFR) vs EI(Prediction) of eating episodes from LOSOCV

Evaluation Results:

Two separate methods of EI estimation i) Blind EI estimation method ii) Image-assisted EI estimation method were implemented in the evaluation study data. These two methods were evaluated against i) EI estimation by manual image review by trained nutritionists and ii) EER calculated from participant's demographic information.

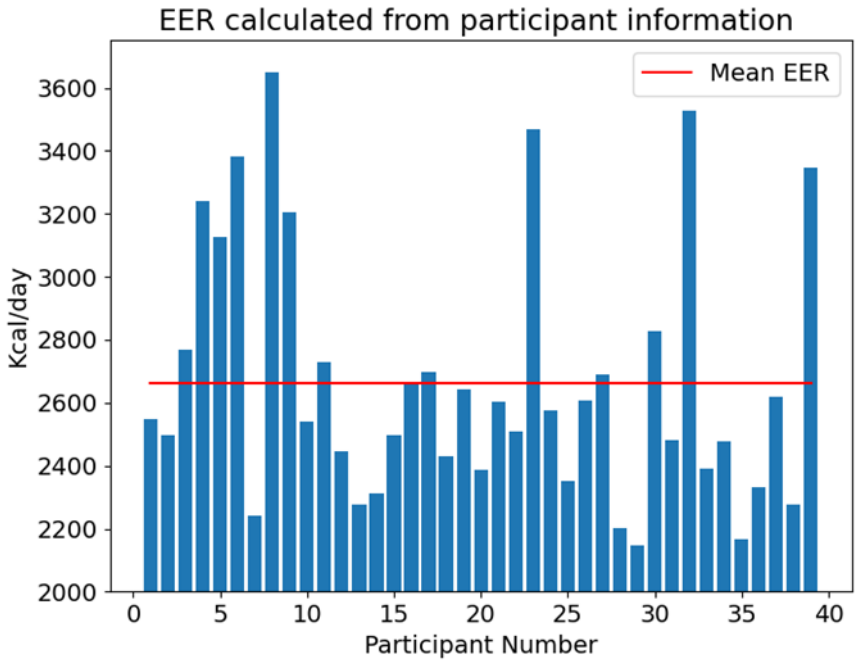


Figure 6-5: EER (Kcal/day) calculated from participant's information using equation.

The mean ± std EI estimated per day from the manual annotation of images was 1218.51 ± 586.26. The calculated EER was 2663.57 ± 398.86. The EI estimated by blind EI method was 2423.66 ± 1416.04. Image-assisted estimated EI was 2041.85 ± 1048.22. The results are presented by the violin plot and boxplot in Figure 6-6.

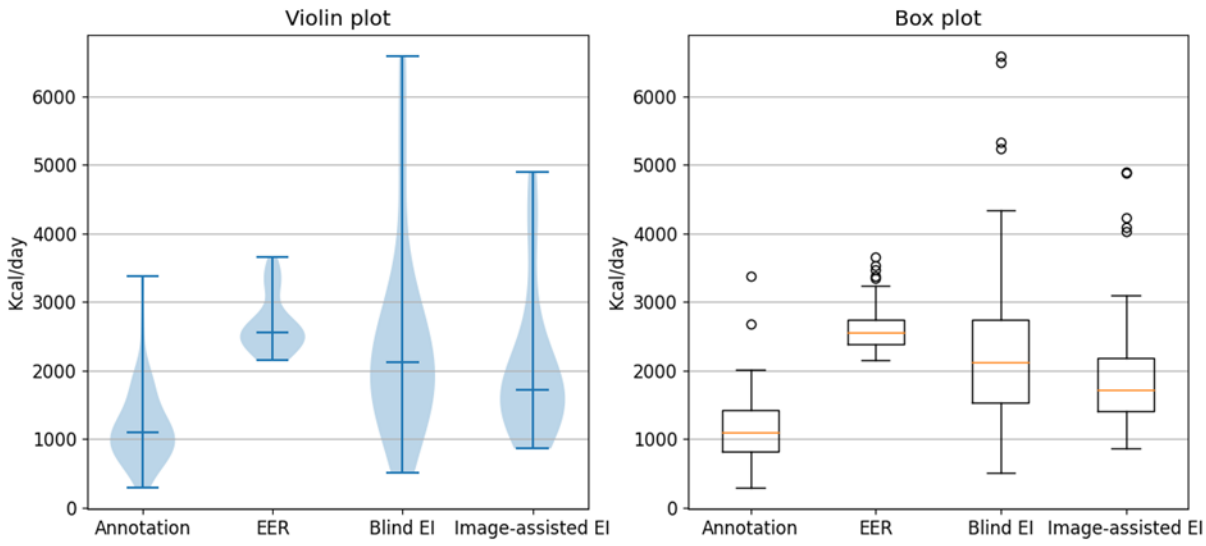


Figure 6-6: Violin and Box plot of EI/day from blind EI estimation method, image-assisted EI estimation method, and EI estimation from manual annotation by nutritionists.

6.4 Discussion

In this study, we present the development of a sensor-based regression model and evaluation of two methods of implementation, i) the Blind EI method ii) the Image-assisted EI method, to estimate consumed EI from eating episodes detected by the AIM-2 sensor system. The model was trained using data from a free-living training study with WFR as the ground truth. The selected subset of features included 5 features out of computed 259 features. All selected features were related to eating behavior. With the selected five features the mean absolute training error was 42.2% ± 53.4%. The sensor-based regression model achieved a R^2 value of 0.47 with the selected 5 features in LOSO-CV during training.

The two methods (Blind EI method and image-assisted EI method) of EI estimation were implemented and evaluated in an independent free-living evaluation study. Since WFR was not available for the evaluation study, manual image annotation and EER were used as points of comparison. The sensor-derived EI estimation methods (blind EI, and image-assisted EI) provide

a closer match to the EER. EI from manual image review by nutritionist (1218.51 ± 586.26) is significantly less than EER (2663.57 ± 398.86).

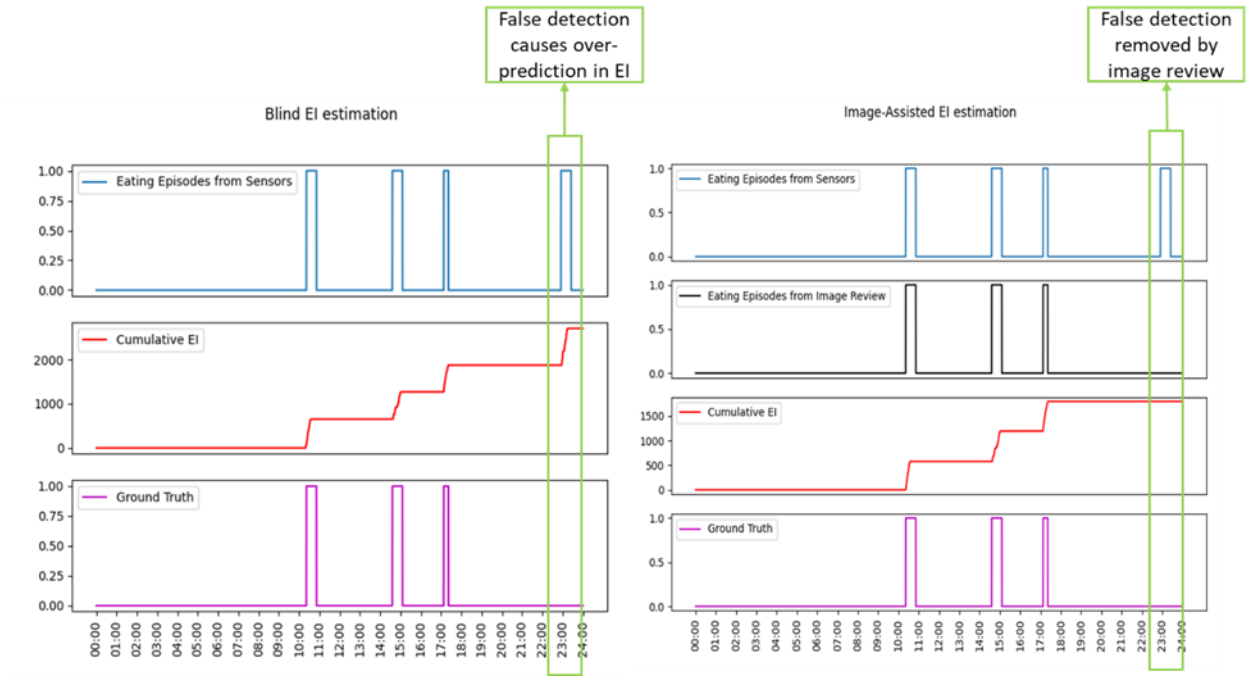


Figure 6-7: (a) Daily EI estimation by blind EI method. False eating episode detection from sensors causes over-estimation in prediction (b) Daily EI estimation by image-assisted EI method. False eating episodes are removed by manual image review.

The blind EI method estimated EI solely from the food intake detection performance from the accelerometer and optical sensor of AIM-2. Any false food intake detection can contribute to false eating detection resulting in overprediction in EI estimation. Figure 6-7(a) represents an example of daily EI estimation from sensor-detected eating episodes using the blind EI estimation method. A false eating episode was detected which contributed to the over-estimation of EI. Figure 6-7(b) represents the daily EI estimation using the image-assisted EI estimation method. In this method, AIM-2 captured images within the sensor-detected eating episodes are reviewed manually and annotated as false and true eating episodes. The EI

estimation model is then applied to the eating episodes which are annotated as true, thus reducing overprediction in estimation.

The EI estimation from manual image review using AIM-2 captured images is very challenging as reported in our previous study [43]. The annotators must identify correct food items and correct portion size of the consumed foods from AIM-2 recorded images of a true eating episode to correctly estimate EI. As AIM-2 automatically captures images, finding representative images may be difficult due to head movement, poor lighting, and contextual issues such as the existence of extraneous foods in the images. If the annotator can see that food is being eaten within a sensor-detected eating episode but cannot find representative images to correctly identify food items and portion size then, the annotator marks the sensor-detected eating episode as a True eating episode and does not assign any food items or portion size for that eating episode, resulting EI estimation of 0 Kcal from manual annotation. Furthermore, unlike traditional photographic food diaries, the images are captured as food is naturally consumed rather than being neatly staged, which makes the portion size estimation of the consumed food items more difficult. After identifying consumed food items and portion sizes, the annotators must find an exact match of the consumed food item in the database for correct EI estimation. Lack of representative images to correctly identify food items and portion size in AIM-2 captured images, mismatch in portion size estimation, and food item selection in the database can be the reason behind under-prediction in EI estimation by manual annotation.
In this study, a novel contactless optical sensor was tested to monitor eating by tracking the temporalis muscle for user comfort in long wear. The results obtained from the optical sensor are comparable with the contact flexible “bend” sensor used in a previous study [43]. However, the optical sensor significantly increases user comfort by eliminating a pressure point from the

sensor on the temple and does not need individual fitting to the frame size and participant's anatomy.

One of the major strengths of this study is that the proposed sensor-based regression model for EI estimation was trained on data collected from a free-living environment. The developed EI estimation model was then evaluated using data from a separate free-living study against EI from manual image annotation and EER computed from participants' information. Furthermore, to reduce the effect of false eating episode detection by the sensor, information from images by the manual review was incorporated to remove the false detection and improve the model accuracy in the image-assisted EI estimation method.

A limitation of this study is that while food intake detection was done automatically, the review of images was performed manually to remove false eating episode detection to improve the blind EI estimation performance. In the future, instead of manual image review, deep learning-based computer vision techniques could potentially be used for automatic food recognition from the AIM-2 captured images thus providing a robust automatic eating episode detection algorithm by incorporating both sensor and image information without manual intervention. Another limitation of this study is the lack of ground truth in the evaluation study to compare with the findings. Due to the lack of actual ground truth, the EI estimation methods are evaluated against EER and manual image review by trained nutritionists which are approximations rather than actual values. Further studies need to be performed in the future, where doubly labeled water could be used to find the actual EER for comparing the EI estimation by sensor-based methods.

6.5 Conclusion

This study introduces a novel approach for the training and evaluation of an Energy Intake (EI) estimation model utilizing the AIM-2 wearable sensor. The developed regression model, using novel contactless optical sensor and accelerometer signal, was trained with data from 30 participants who wore AIM-2 for two days in a free-living environment. Subsequently, the developed regression model underwent evaluation using data collected from 39 participants who wore AIM-2 over a week in a free-living setting. The evaluation focused on three randomly selected days, including two weekdays and one weekend day. Two distinct methods for EI estimation were implemented using the sensor-based regression model and evaluated: 1) the Blind EI estimation method and 2) the Image-assisted EI estimation method. The latter method integrated AIM-2 captured image information through manual review to eliminate false eating episode detections, thereby enhancing the performance and accuracy of the blind EI method. Both methods were assessed against 1) EI estimation derived from manual reviews of AIM-2 captured images by trained nutritionists, and 2) EER calculated from participants' demographic information, including age, sex, weight, and height. The sensor-derived EI estimation methods (blind EI, and image-assisted EI) provide a closer match to the EER. EI from manual image review by trained nutritionists is significantly less than EER. The findings underscore the effectiveness of the sensor-based EI model, emphasizing its potential to advance dietary monitoring and enable personalized nutrition interventions. Further studies with available ground truth such as doubly labeled water need to be performed to provide further evaluation of the sensor-based EI estimation.

References:

[1] H. Wang *et al.*, “Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: a systematic analysis for the Global Burden of Disease Study 2015,” *The Lancet*, vol. 388, no. 10053, pp. 1459–1544, Oct. 2016, doi: 10.1016/S0140-6736(16)31012-1.

[2] “Fifty Years of Behavioral/Lifestyle Interventions for Overweight and Obesity: Where Have We Been and Where Are We Going? - Williamson - 2017 - Obesity - Wiley Online Library.” Accessed: Oct. 29, 2023. [Online]. Available: https://onlinelibrary.wiley.com/doi/full/10.1002/oby.21914

[3] S. K. Das *et al.*, “Body-composition changes in the Comprehensive Assessment of Long-term Effects of Reducing Intake of Energy (CALERIE)-2 study: a 2-y randomized controlled trial of calorie restriction in nonobese humans,” *Am J Clin Nutr*, vol. 105, no. 4, pp. 913–927, Apr. 2017, doi: 10.3945/ajcn.116.137232.

[4] T. D. P. P. (DPP) R. Group, “The Diabetes Prevention Program (DPP): Description of lifestyle intervention,” *Diabetes Care*, vol. 25, no. 12, pp. 2165–71, Dec. 2002, doi: 10.2337/diacare.25.12.2165.

[5] “Eight-year weight losses with an intensive lifestyle intervention: The look AHEAD study - - 2014 - Obesity - Wiley Online Library.” Accessed: Oct. 29, 2023. [Online]. Available: https://onlinelibrary.wiley.com/doi/full/10.1002/oby.20662

[6] “Weight-Loss Diets, Adiponectin, and Changes in Cardiometabolic Risk in the 2-Year POUNDS Lost Trial | The Journal of Clinical Endocrinology & Metabolism | Oxford Academic.” Accessed: Oct. 29, 2023. [Online]. Available: https://academic.oup.com/jcem/article/101/6/2415/2804777

[7] B. K. Sharma, M. Patil, and A. Satyanarayana, “Negative Regulators of Brown Adipose Tissue (BAT)-Mediated Thermogenesis,” *Journal of Cellular Physiology*, vol. 229, no. 12, pp. 1901–1907, 2014, doi: 10.1002/jcp.24664.

[8] D. H. Ryan and M. D. Jensen, “What the new obesity guidelines will tell us,” *Current Opinion in Endocrinology, Diabetes and Obesity*, vol. 20, no. 5, p. 429, Oct. 2013, doi: 10.1097/01.med.0000432612.80860.34.

[9] S. N. Blair, “Physical inactivity: the biggest public health problem of the 21st century,” *British Journal of Sports Medicine*, vol. 43, no. 1, pp. 1–2, Jan. 2009.

[10] A. F. Subar *et al.*, “Addressing Current Criticism Regarding the Value of Self-Report Dietary Data,” *J Nutr*, vol. 145, no. 12, pp. 2639–2645, Dec. 2015, doi: 10.3945/jn.115.219634.

[11] R. K. Johnson, “Dietary intake--how do we measure what people are really eating?,” *Obes Res*, vol. 10 Suppl 1, pp. 63S-68S, Nov. 2002, doi: 10.1038/oby.2002.192.

[12] E. Archer, G. Pavela, and C. J. Lavie, “The Inadmissibility of What We Eat in America and NHANES Dietary Data in Nutrition and Obesity Research and the Scientific Formulation of National Dietary Guidelines,” *Mayo Clin Proc*, vol. 90, no. 7, pp. 911–926, Jul. 2015, doi: 10.1016/j.mayocp.2015.04.009.

[13] N. V. Dhurandhar *et al.*, “Energy balance measurement: when something is not better than nothing,” *Int J Obes*, vol. 39, no. 7, Art. no. 7, Jul. 2015, doi: 10.1038/ijo.2014.199.

[14] J. P. A. Ioannidis, “Implausible results in human nutrition research,” *BMJ*, vol. 347, p. f6698, Nov. 2013, doi: 10.1136/bmj.f6698.

[15] K. Murakami and M. B. E. Livingstone, “Prevalence and characteristics of misreporting of energy intake in US adults: NHANES 2003–2012,” *British Journal of Nutrition*, vol. 114, no. 8, pp. 1294–1303, Oct. 2015, doi: 10.1017/S0007114515002706.

[16] P. P. Basiotis, S. O. Welsh, F. J. Cronin, J. L. Kelsay, and W. Mertz, “Number of days of food intake records required to estimate individual and group nutrient intakes with defined confidence,” *J Nutr*, vol. 117, no. 9, pp. 1638–1641, Sep. 1987, doi: 10.1093/jn/117.9.1638.

[17] “Calorie Tracker & BMR Calculator to Reach Your Goals | MyFitnessPal.” Accessed: Oct. 29, 2023. [Online]. Available: https://www.myfitnesspal.com

[18] “Lose It! - Calorie counting made easy.” Accessed: Oct. 29, 2023. [Online]. Available: https://www.loseit.com/

[19] “MyNetDiary - Free Calorie Counter and Diet Assistant,” MyNetDiary. Accessed: Oct. 29, 2023. [Online]. Available: https://www.mynetdiary.com/

[20] F. Cordeiro *et al.*, *Barriers and Negative Nudges*. 2015. doi: 10.1145/2702123.2702155.

[21] F. E. Thompson, A. F. Subar, C. M. Loria, J. L. Reedy, and T. Baranowski, “Need for Technological Innovation in Dietary Assessment,” *J Am Diet Assoc*, vol. 110, no. 1, pp. 48–51, Jan. 2010, doi: 10.1016/j.jada.2009.10.008.

[22] D. A. Schoeller, “Measurement of energy expenditure in free-living humans by using doubly labeled water,” *J Nutr*, vol. 118, no. 11, pp. 1278–1289, Nov. 1988, doi: 10.1093/jn/118.11.1278.

[23] S. Gill and S. Panda, “A Smartphone App Reveals Erratic Diurnal Eating Patterns in Humans that Can Be Modulated for Health Benefits,” *Cell Metabolism*, vol. 22, no. 5, pp. 789–798, Nov. 2015, doi: 10.1016/j.cmet.2015.09.005.

[24] C. Höchsmann and C. K. Martin, “Review of the validity and feasibility of image-assisted methods for dietary assessment,” *Int J Obes*, vol. 44, no. 12, Art. no. 12, Dec. 2020, doi: 10.1038/s41366-020-00693-2.

[25] J. M. Fontana *et al.*, “Energy intake estimation from counts of chews and swallows,” *Appetite*, vol. 85, pp. 14–21, Feb. 2015, doi: 10.1016/j.appet.2014.11.003.

[26] J.-S. Shim, K. Oh, and H. C. Kim, “Dietary assessment methods in epidemiologic studies,” *Epidemiol Health*, vol. 36, p. e2014009, 2014, doi: 10.4178/epih/e2014009.

[27] K. Yanai and Y. Kawano, “Food image recognition using deep convolutional network with pre-training and fine-tuning,” in *2015 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*, Jun. 2015, pp. 1–6. doi: 10.1109/ICMEW.2015.7169816.

[28] N. Martinel, G. L. Foresti, and C. Micheloni, “Wide-Slice Residual Networks for Food Recognition,” in *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, Mar. 2018, pp. 567–576. doi: 10.1109/WACV.2018.00068.

[29] W. Min, L. Liu, Z. Luo, and S. Jiang, “Ingredient-Guided Cascaded Multi-Attention Network for Food Recognition,” in *Proceedings of the 27th ACM International Conference on Multimedia*, in MM ’19. New York, NY, USA: Association for Computing Machinery, Oct. 2019, pp. 1331–1339. doi: 10.1145/3343031.3350948.

[30] J. Qiu, F. P.-W. Lo, Y. Sun, S. Wang, and B. Lo, “Mining Discriminative Food Regions for Accurate Food Recognition.” arXiv, Jul. 08, 2022. doi: 10.48550/arXiv.2207.03692.

[31] T. Ghosh *et al.*, “Integrated image and sensor-based food intake detection in free-living,” *Sci Rep*, vol. 14, no. 1, p. 1665, Jan. 2024, doi: 10.1038/s41598-024-51687-3.

[32] J. Qiu *et al.*, “Egocentric Image Captioning for Privacy-Preserved Passive Dietary Intake Monitoring,” *IEEE Transactions on Cybernetics*, vol. 54, no. 2, pp. 679–692, Feb. 2024, doi: 10.1109/TCYB.2023.3243999.

[33] “Food Intake Activity Detection Using a Wearable Microphone System | IEEE Conference Publication | IEEE Xplore.” Accessed: Oct. 29, 2023. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/6063400

[34] S. Päßler and W.-J. Fischer, “Food Intake Monitoring: Automated Chew Event Detection in Chewing Sounds,” *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 1, pp. 278–289, Jan. 2014, doi: 10.1109/JBHI.2013.2268663.

[35] E. Sazonov *et al.*, “Non-invasive monitoring of chewing and swallowing for objective quantification of ingestive behavior,” *Physiol. Meas.*, vol. 29, no. 5, p. 525, Apr. 2008, doi: 10.1088/0967-3334/29/5/001.

[36] “On-Body Sensing Solutions for Automatic Dietary Monitoring | IEEE Journals & Magazine | IEEE Xplore.” Accessed: Oct. 29, 2023. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/4814940

[37] A. Doulah, T. Ghosh, D. Hossain, M. H. Imtiaz, and E. Sazonov, “‘Automatic Ingestion Monitor Version 2’ – A Novel Wearable Device for Automatic Food Intake Detection and Passive Capture of Food Images,” *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 2, pp. 568–576, Feb. 2021, doi: 10.1109/JBHI.2020.2995473.

[38] M. Mirtchouk, C. Merck, and S. Kleinberg, “Automated estimation of food type and amount consumed from body-worn audio and motion sensors,” in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, in UbiComp ’16. New York, NY, USA: Association for Computing Machinery, Sep. 2016, pp. 451–462. doi: 10.1145/2971648.2971677.

[39] Y. Dong, J. Scisco, M. Wilson, E. Muth, and A. Hoover, “Detecting Periods of Eating During Free-Living by Tracking Wrist Motion,” *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 4, pp. 1253–1260, Jul. 2014, doi: 10.1109/JBHI.2013.2282471.

[40] Y. Dong, A. Hoover, J. Scisco, and E. Muth, “A New Method for Measuring Meal Intake in Humans via Automated Wrist Motion Tracking,” *Appl Psychophysiol Biofeedback*, vol. 37, no. 3, pp. 205–215, Sep. 2012, doi: 10.1007/s10484-012-9194-1.

[41] J. N. Salley, A. W. Hoover, M. L. Wilson, and E. R. Muth, “Comparison between Human and Bite-Based Methods of Estimating Caloric Intake,” *Journal of the Academy of Nutrition and Dietetics*, vol. 116, no. 10, pp. 1568–1577, Oct. 2016, doi: 10.1016/j.jand.2016.03.007.

[42] S. Zhang, D. T. Nguyen, G. Zhang, R. Xu, N. Maglaveras, and N. Alshurafa, “Estimating caloric intake in bedridden hospital patients with audio and neck-worn sensors,” in *Proceedings of the 2018 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies*, in CHASE ’18. New York, NY, USA: Association for Computing Machinery, Jan. 2020, pp. 1–2. doi: 10.1145/3278576.3278577.

[43] “Energy intake estimation using a novel wearable sensor and food images in a laboratory (pseudo-free-living) meal setting: quantification and contribution of sources of error | International Journal of Obesity.” Accessed: Feb. 23, 2024. [Online]. Available: https://www.nature.com/articles/s41366-022-01225-w

[44] J. M. Fontana *et al.*, “Energy intake estimation from counts of chews and swallows,” *Appetite*, vol. 85, pp. 14–21, Feb. 2015, doi: 10.1016/j.appet.2014.11.003.

[45] S. Wang *et al.*, “Eating detection and chews counting through sensing mastication muscle contraction,” *Smart Health*, vol. 9–10, pp. 179–191, Dec. 2018, doi: 10.1016/j.smhl.2018.07.004.

[46] T. Ghosh, D. Hossain, M. Imtiaz, M. A. McCrory, and E. Sazonov, “Implementing Real-Time Food Intake Detection in a Wearable System Using Accelerometer,” in *2020 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, Mar. 2021, pp. 439–443. doi: 10.1109/IECBES48179.2021.9398760.

[47] T. Ghosh *et al.*, “I2N: image to nutrients, a sensor guided semi-automated tool for annotation of images for nutrition analysis of eating episodes,” *Front Nutr*, vol. 10, p. 1191962, Jul. 2023, doi: 10.3389/fnut.2023.1191962.

[48] A. Doulah, T. Ghosh, D. Hossain, M. H. Imtiaz, and E. Sazonov, “‘Automatic Ingestion Monitor Version 2’ – A Novel Wearable Device for Automatic Food Intake Detection and Passive Capture of Food Images,” *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 2, pp. 568–576, Feb. 2021, doi: 10.1109/JBHI.2020.2995473.

[49] D. Hossain, T. Ghosh, M. Haider Imtiaz, and E. Sazonov, “Ear canal pressure sensor for food intake detection,” *Frontiers in Electronics*, vol. 4, 2023, Accessed: Feb. 23, 2024. [Online]. Available: https://www.frontiersin.org/articles/10.3389/felec.2023.1173607

[50] J. M. C. Po, J. A. Kieser, L. M. Gallo, A. J. Tésenyi, P. Herbison, and M. Farella, “Time-Frequency Analysis of Chewing Activity in the Natural Environment,” *J Dent Res*, vol. 90, no. 10, pp. 1206–1210, Oct. 2011, doi: 10.1177/0022034511416669.

[51] R. Tibshirani, “Regression Shrinkage and Selection Via the Lasso,” *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, Jan. 1996, doi: 10.1111/j.2517-6161.1996.tb02080.x.

[52] F. J. Ferri, P. Pudil, M. Hatef, and J. Kittler, “Comparative study of techniques for large-scale feature selection*,” in *Machine Intelligence and Pattern Recognition*, vol. 16, E. S. Gelsema and L. S. Kanal, Eds., in Pattern Recognition in Practice IV, vol. 16. , North-Holland, 1994, pp. 403–413. doi: 10.1016/B978-0-444-81892-8.50040-7.

[53] S. Bi *et al.*, “Auracle: Detecting Eating Episodes with an Ear-mounted Sensor,” *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 2, no. 3, pp. 1–27, Sep. 2018, doi: 10.1145/3264902.

[54] V. Raju and E. Sazonov, “Processing of Egocentric Camera Images from a Wearable Food Intake Sensor,” in *2019 SoutheastCon*, Apr. 2019, pp. 1–6. doi: 10.1109/SoutheastCon42311.2019.9020284.

[55] Z. Pan *et al.*, “Improvement of Methodology for Manual Energy Intake Estimation From Passive Capture Devices,” *Frontiers in Nutrition*, vol. 9, 2022, Accessed: Feb. 23, 2024. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fnut.2022.877775

[56] *Dietary Reference Intakes for Energy*. Washington, D.C.: National Academies Press, 2023. doi: 10.17226/26818.

[57] *Dietary Reference Intakes for Energy, Carbohydrate, Fiber, Fat, Fatty Acids, Cholesterol, Protein, and Amino Acids*. Washington, D.C.: National Academies Press, 2005. doi: 10.17226/10490.

[58] P. Trumbo, S. Schlicker, A. A. Yates, and M. Poos, “Dietary reference intakes for energy, carbohydrate, fiber, fat, fatty acids, cholesterol, protein and amino acids. (Commentary).,” *Journal of the American Dietetic Association*, vol. 102, no. 11, pp. 1621–1631, Nov. 2002.

Supplemental Table 1: List of extracted time and frequency domain features

No.	Feature Name	No.	Feature Name	No.	Feature Name
1	Mean Absolute Value	15	Peaks to zero crossing	29	Chewing energy to walking energy
2	RMS value	16	Zero crossing to peaks	30	Talking energy to spectrum energy ratio
3	Max absolute value	17	Wavelength	31	Chewing entropy to walking entropy ratio
4	Median value	18	Std to mean ratio	32	Chewing energy to talking energy
5	Mean to RMS ratio	19	Slope sign change	33	Chewing energy entropy to talking energy entropy
6	RMS to maximum ratio	20	Spectrum energy	34	Walking energy to talking energy
7	Mean to Max ratio	21	Chewing energy	35	Walking energy entropy to talking energy entropy
8	Median to Max	22	Chewing energy to spectrum energy ratio	36	Max of chewing frequency
9	Entropy	23	Entropy of chewing	37	Max of walking frequency
10	Number of zero crossing	24	Walking energy	38	Max of talking frequency
11	Time between zero crossing	25	Entropy of walking energy	39	Spectral centroid
12	Number of peaks	26	Walking energy to spectrum energy ratio	40	Spectral spread
13	Mean peak value	27	Talking energy	41	skewness
14	Distance between peaks	28	Entropy talking energy	42	kurtosis
43	Standard deviation				

Supplemental Table 2: List of extracted correlation features

No.	Feature Name	No.	Feature Name
1	AccX corr AccY	**6**	AccZ corr AccNET
2	AccX corr AccZ	**7**	AccX corr OPT
3	AccX corr AccNET	**8**	AccY corr OPT
4	AccY corr AccZ	**9**	AccZ corr OPT
5	AccY corr AccNET	**10**	Acc NET corr OPT

Supplemental Table 3: List of extracted features from roll and pitch

No.	Feature Name	No.	Feature Name
1	Mean of pitch	**12**	Mean of roll
2	Std of pitch	**13**	Std of roll
3	Rms of pitch	**14**	Rms of roll

4	Shape factor of pitch	**15**	Shape factor of roll
5	Kurtosis of pitch	**16**	Kurtosis of roll
6	Skewness of pitch	**17**	Skewness of roll
7	Max of pitch	**18**	Max of roll
8	Impulse factor of pitch	**19**	Impulse factor of roll
9	Clearance factor of pitch	**20**	Clearance factor of roll
10	Crest factor of pitch	**21**	Crest factor of roll
11	Median of pitch	**22**	Median of roll

Supplemental Table 4: List of extracted eating behavior features

No	**Feature Name**	**No.**	**Feature Name**
1	Chew count	**6**	Actual chewing duration
2	No of chew bouts	**7**	Pause duration
3	Meal duration	**8**	% Pause duration =(pause duration/meal duration) × 100
4	Chewing rate	**9**	%Chewing = (chewing duration/meal duration) × 100
5	Chewing period	**10**	Chew count to chewing duration

CHAPTER 7 – CONCLUSIONS AND FUTURE WORK

This dissertation contributes to the work on automatically monitoring eating behavior using wearable sensors. The main contribution of this work is developing and testing new signal processing, pattern recognition, and estimation algorithms for objective, automatic, and accurate monitoring of eating behavior from wearable sensor systems. Another important contribution of this work is the development of a machine-learning EI estimation model using sensor-captured eating behavior features which can be used to provide feedback based on the user energy consumption.

Eating videos were explored in this dissertation for automatic detection and estimation of chewing and biting for contactless, non-invasive eating behavior detection. This study presents a significant advancement in leveraging computer vision for the automated detection and estimation of chewing and biting during eating episodes. The study contributes three key components: a face detection model utilizing the Faster-RCNN object detection scheme to identify faces as regions of interest (ROI) in recorded meal videos; an image classifier based on pre-trained AlexNet, facilitating the classification of detected faces into 'bite' and 'non-bite' categories and thereby calculating the total number of bites during an eating episode; and the introduction of a novel approach for counting chews by applying optical flow analysis to consecutive cropped faces to discern jaw movements. The achieved promising accuracy in counting both chews and bites from recorded videos underscores the potential for further exploration, aiming to fully automate and enhance the accuracy, efficiency, and user-friendliness

of bite and chew annotations in eating episodes. Gestures such as speaking, facial movements, head movements, use of napkin to wipe the mouth can also affect the bite and chew count accuracy. The major limitation of this study is that eating behaviors, such as chews and bites, were detected and counted in laboratory conditions rather than in real-world scenarios. The video documentation of eating episodes was intentionally captured from the participant's lateral perspective to enhance the visibility of jaw movements during chewing. Future research should focus on refining the algorithm to increase its robustness, enabling it to accurately quantify the number of bites and chews in eating episodes regardless of the video angle. While the current study involved offline calculations, future investigations could implement a real-time process to offer immediate insights into eating behavior, including the number of bites, chews, and the duration of chewing segments post-meal. One potential application of the proposed algorithm could be the use of smartphone cameras to monitor eating behavior, allowing users to record their eating events and receive real-time information on chews, bites, eating rate, and eating duration as the meal progresses. Though the proposed method opens new avenues for research in eating behavior monitoring using cameras, it is currently best suited for annotating eating behavior from video in laboratory-controlled experiments, thereby easing the manual annotation process.

An innovative approach was also explored in this dissertation to record bite force and estimate chewing strength during the consumption of foods with varying hardness levels, leveraging wearable sensors—an area not previously explored in this context. This study proposes the utilization of a combination of sensors, including a pressure sensor, a piezoresistive bend sensor attached to the temporalis muscle, a piezoelectric strain sensor, and an electromyography (EMG) sensor, for the detection of variations in chewing strength in response to different food textures.

Results demonstrate that all four sensor modalities effectively capture changes in chewing strength across foods with varying hardness. The estimated chewing strength correlates proportionally with the hardness of the food; harder textures yield higher estimated chewing strength values, whereas softer textures yield lower values. This study's findings pave the way for developing a predictive model to recognize food items from sensor signals based on chewing strength estimation. It also opens new avenues for integrating food property identification into sensor-based energy estimation. Linking physical food properties, such as texture and hardness, can potentially enhance the accuracy of sensor-based food intake mass and energy consumption predictive algorithms. Future research should involve a larger number of participants and a wider variety of test foods to further investigate the performance of chewing strength estimation using wearable sensor systems.

Further, the use of ear canal pressure change during chewing was explored to detect eating and estimate the number of chews in eating episodes in free-living settings. Controlled experiments provided the dataset for training a Support Vector Machine (SVM) classifier in this study, enabling the detection of eating events and the development of a method to identify chewing segments and count chews within each eating event. The outcomes underscored the efficacy of the proposed wearable device, demonstrating its high accuracy and precision in detecting eating events and quantifying the number of chews, particularly in real-life, free-living settings. The designed Ear Canal Pressure Sensor (ECPS) system has several limitations. Both sensor signals depend on chewing, which restricts detection to solid and semi-solid food consumption. Additionally, continuous image capture by the camera raises privacy concerns for both the wearer and those nearby. Interferences such as speaking, facial movements, head movements, air tube movements, and touching the earbud can also affect sensor accuracy. This study aimed to

develop and validate the ECPS system for objective food intake detection and was limited to 10 volunteers. Future studies should include a larger number of participants in free-living environments to assess the usability and comfort of the proposed sensor system in daily life. This dissertation further explored a new method for training and assessing an Energy Intake (EI) estimation model utilizing the AIM-2 wearable sensor. The regression model, incorporating eating behavior features from an optical sensor and accelerometer data, was trained on a two-day free-living dataset from 30 participants and evaluated using a week-long dataset from 39 participants. Two EI estimation methods were employed: the Blind Sensor Method and the Image-assisted Sensor Method. The latter integrated AIM-2 images, enhancing accuracy through manual review to eliminate false eating episode detections. Both methods were compared to manual image reviews by nutritionists and Total Energy Expenditure (TEE) derived from participant demographics. Results emphasize the effectiveness of the sensor-based EI model, showcasing its promise for advancing dietary monitoring and personalized nutrition interventions. One limitation of this study lies in the manual review of images to eliminate false eating episode detections, despite automated food intake detection. To enhance the blind sensor model's performance in the future, deep learning-based computer vision techniques could replace manual image review for automatic food recognition from AIM-2 captured images, creating a robust eating episode detection algorithm that integrates sensor and image information seamlessly, without manual intervention. Another limitation of this study is the lack of ground truth in the evaluation study to compare with the findings. Due to the lack of actual ground truth, the EI estimation methods are evaluated against EER and manual image review by trained nutritionists which are approximations rather than actual values. Further studies need to be

performed in the future, where doubly labeled water could be used to find the actual EER for comparing the EI estimation by sensor-based methods.

A prospective avenue for research involves the integration of real-time feedback mechanisms for AIM-2 users, aimed at promoting and sustaining healthy eating behaviors. This would entail the seamless transmission of AIM-2 data to a server via Bluetooth, where trained models for food intake detection, energy intake (EI) estimation, and physical activity, utilizing the embedded accelerometer, could be stored, and implemented. The ongoing development of AIM-2 holds the potential to comprehensively track both energy intake and expenditure, thereby facilitating the monitoring of individual energy balance. This signifies the prospective role of AIM-2 in delivering real-time dietary intervention protocols and guiding users regarding optimal eating times, portion sizes, and food choices.

APPENDIX I PUBLICATIONS

Published:

1. Ghosh, Tonmoy, Yue Han, Viprav Raju, **Delwar Hossain**, Megan A. McCrory, Janine Higgins, Carol Boushey, Edward J. Delp, and Edward Sazonov. "Integrated image and sensor-based food intake detection in free-living." Scientific Reports 14, no. 1 (2024): 1665.
2. Fard, Siavash Esfandiari, Tonmoy Ghosh, **Delwar Hossain**, Megan A. McCrory, Graham Thomas, Janine Higgins, Wenyan Jia et al. "Development of a Method for Compliance Detection in Wearable Sensors." In 2023 International Conference on Electrical, Computer and Energy Technologies (ICECET), pp. 1-5. IEEE, 2023.
3. **Hossain, Delwar**, Tonmoy Ghosh, Masudul Haider Imtiaz, and Edward Sazonov. "Ear Canal Pressure Sensor for Food Intake Detection." Frontiers in Electronics 4: 1173607.
4. Raju, Viprav B., **Delwar Hossain**, and Edward Sazonov. "Estimation of Plate and Bowl Dimensions for Food Portion Size Assessment in a Wearable Sensor System." IEEE Sensors Journal 23, no. 5 (2023): 5391-5400.
5. Doulah, Abul, Tonmoy Ghosh, **Delwar Hossain**, Tyson Marden, Jason M. Parton, Janine A. Higgins, Megan A. McCrory, and Edward Sazonov. "Energy intake estimation using a novel wearable sensor and food images in a laboratory (pseudo-free-living) meal setting: quantification and contribution of sources of error." International Journal of Obesity 46, no. 11 (2022): 2050-2057.
6. Simon, Stacey L., Zhaoxing Pan, Tyson Marden, Wenru Zhou, Tonmoy Ghosh, **Delwar Hossain**, J. Graham Thomas, Megan A. McCrory, Edward Sazonov, and Janine Higgins. "Body mass index and variability in meal duration and association with rate of eating." Frontiers in Nutrition 9 (2022): 941001.
7. Pan, Zhaoxing, Dan Forjan, Tyson Marden, Jonathan Padia, Tonmoy Ghosh, **Delwar Hossain**, J. Graham Thomas, Megan A. McCrory, Edward Sazonov, and Janine A. Higgins. "Improvement of methodology for manual energy intake estimation from passive capture devices." Frontiers in Nutrition 9 (2022): 877775.
8. Liao, Sarah, Tonmoy Ghosh, **Delwar Hossain**, Abul Doulah, Tyson Marden, Janine Higgins, Edward Sazonov, and Megan McCrory. "Effect of Gradually Reduced Portion Size Estimation Training on Portion Size and Daily Energy Intake Estimation Accuracy." Current Developments in Nutrition 6 (2022): 773.
9. Ghosh, Tonmoy, **Delwar Hossain**, and Edward Sazonov. "Detection of Food Intake Sensor's Wear Compliance in Free-Living." IEEE sensors journal 21, no. 24 (2021): 27728-27735.
10. Ghosh, Tonmoy, **Delwar Hossain**, Masudul Imtiaz, Megan A. McCrory, and Edward Sazonov. "Implementing real-time food intake detection in a wearable system using accelerometer." In 2020 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES), pp. 439-443. IEEE, 2021.

11. **Hossain, Delwar**, Qian Mao, Immanuel Manohar, and Fei Hu. "Jamming attacks and countermeasures in UAV networks." UAV Swarm Networks: Models, Protocols, and Systems (2020): 207-222.
12. Imtiaz, Masudul H., **Delwar Hossain**, Volkan Y. Senyurek, Prajakta Belsare, and Edward Sazonov. "PACT CAM: wearable sensor system to capture the details of cigarette smoking in free-living." In 2020 IEEE SENSORS, pp. 1-4. IEEE, 2020.
13. Imtiaz, Masudul H., **Delwar Hossain**, Volkan Y. Senyurek, Prajakta Belsare, Stephen Tiffany, and Edward Sazonov. "Wearable egocentric camera as a monitoring tool of free-living cigarette smoking: a feasibility study." Nicotine and Tobacco Research 22, no. 10 (2020): 1883-1890.
14. **Hossain, Delwar**, Masudul Haider Imtiaz, Tonmoy Ghosh, Viprav Bhaskar, and Edward Sazonov. "Real-time food intake monitoring using wearable egocnetric camera." In 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), pp. 4191-4195. IEEE, 2020.
15. **Hossain, Delwar**, Tonmoy Ghosh, and Edward Sazonov. "Automatic count of bites and chews from videos of eating episodes." Ieee Access 8 (2020): 101934-101945.
16. Doulah, A., Ghosh, T., **Hossain, D**., Imtiaz, M.H. and Sazonov, E., 2020. "Automatic ingestion monitor version 2"–a novel wearable device for automatic food intake detection and passive capture of food images. IEEE journal of biomedical and health informatics, 25(2), pp.568-576.
17. **Hossain, Delwar**, Masudul Haider Imtiaz, and Edward Sazonov. "Comparison of wearable sensors for estimation of chewing strength." IEEE sensors journal 20, no. 10 (2020): 5379-5388.

Under Preparation:

1. A systematic review of technology-driven methods for measurement of eating behavior
2. Development of the energy intake model from sensor and image captured eating in free-living

APPENDIX II IRB APPROVAL FOR EVALUATION OF WEARABLE DEVICE FOR FOOD INTAKE AND INGESTIVE BEHAVIOR MONITORING

IRB Project #: 17-005-ME

UNIVERSITY OF ALABAMA
INSTITUTIONAL REVIEW BOARD FOR THE PROTECTION OF HUMAN SUBJECTS
REQUEST FOR APPROVAL OF RESEARCH INVOLVING HUMAN SUBJECTS

I. Identifying information

	Principal Investigator	Second Investigator	Third Investigator
Names:	Edward S. Sazonov		
Department:	Electrical and Computer Engineering		
College:	Engineering		
University:	University of Alabama		
Address:	342 HM Comer, Tuscaloosa, AL 35487-0286		
Telephone:	(205) 348-1981		
FAX:			
E-mail:	esazonov@eng.ua.edu		

Title of Research Project: Evaluation of a Wearable Camera for Food Intake and Ingestive Behavior Monitoring

Date Submitted: 6/14/17
Funding Source: NIH - UA OSP # 14-0438

Type of Proposal	☒ New	☐ Revision	☐ Renewal Please attach a renewal application	☐ Completed	☐ Exempt
			Please attach a continuing review of studies form		
			Please enter the original IRB # at the top of the page		

UA faculty or staff member signature: _____

II. NOTIFICATION OF IRB ACTION (to be completed by IRB):
Type of Review: ___✓___ Full board _____ Expedited

IRB Action:
___ Rejected Date: _____
___ Tabled Pending Revisions Date: _____
___ Approved Pending Revisions Date: _____
___✓___ Approved-this proposal complies with University and federal regulations for the protection of human subjects.
Approval is effective until the following date: 8-9-18 e.s.
Items approved: ___ Research protocol (dated_____)
___ Informed consent (dated_____)
(dated_____)
(dated_____)

Approval signature ___ Date 8/22/2017

APPENDIX III IRB APPROVAL FOR WEIGHED RECORD FOOD INTAKE DATA COLLECTION USING A WEARABLE DEVICE TO DEVELOP MODELS FOR ENERGY INTAKE

THE UNIVERSITY OF ALABAMA | Research & Economic Development
Office for Research Compliance

September 27, 2022

To: Edward Sazonov, Ph.D.
Professor
Electrical & Computer Engineering
College of Engineering
Box 870286

From: Debra Pettit Bruns PhD, MSN, MSPH
Chair, Medical IRB

Re: **Notice of Approval**

IRB Application #:	21-04-4517-R1-A
Project Title:	"Weighed Record Food Intake Data Collection Using a Wearable Device to Develop Models for Energy Intake"
Submission Type:	Minor Revision
Approval Date:	September 27, 2022
Expiration Date:	August 10, 2023
Funding Source:	NIH/OSP #: 19-0180
Review Category:	Minor Expedited Revision for Full Board Protocol
Approved Documents:	Informed Consent, Waiver of Documentation of Informed Consent

Dear Dr. Sazonov:

The University of Alabama Medical Institutional Review Board has approved your proposed research. Therefore, your application has been approved according to 45 CFR part 46.

The approval for your application will lapse, as noted above. If your research will continue beyond this date, please submit the Continuing Review to the IRB as University policy requires before the lapse. Please note any modifications made in research design, methodology, or procedures must be submitted to and approved by the IRB before implementation. Please submit a final report form when the study is complete.

Please use reproductions of the stamped IRB-approved informed consent form to obtain consent from your participants.

All the best with your research.

166 Rose Administration Building | Box 870127 | Tuscaloosa, AL 35487-0127 | 205-348-8461
Fax 205-348-8882 | Toll Free 1-877-820-3066 | rscompliance@research.ua.edu