University of Memphis

University of Memphis Digital Commons

Electronic Theses and Dissertations

8-17-2023

AI Modeling Approaches for Detecting, Characterizing, and Predicting Brief Daily Behaviors such as Toothbrushing using Wrist Trackers.

Sayma Akther

Follow this and additional works at: https://digitalcommons.memphis.edu/etd

Recommended Citation

Akther, Sayma, "AI Modeling Approaches for Detecting, Characterizing, and Predicting Brief Daily Behaviors such as Toothbrushing using Wrist Trackers." (2023). *Electronic Theses and Dissertations*. 2987.
https://digitalcommons.memphis.edu/etd/2987

This Dissertation is brought to you for free and open access by University of Memphis Digital Commons. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of University of Memphis Digital Commons. For more information, please contact khggerty@memphis.edu.

AI MODELING APPROACHES FOR DETECTING, CHARACTERIZING, AND PREDICTING BRIEF DAILY BEHAVIORS SUCH AS TOOTHBRUSHING USING WRIST TRACKERS

by

Sayma Akther

A Thesis

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

Major: Computer Science

The University of Memphis

August 2023

©Copyright 2023 Sayma Akther

Partial rights reserved

ACKNOWLEDGEMENTS

First and foremost, all praise and gratitude belong to Allah. Over the past seven years, I have received unwavering support and encouragement from numerous individuals. I would like to express my deepest gratitude to Dr. Santosh Kumar, my advisor, who has served as an extraordinary mentor, consistently motivating and guiding me during this fulfilling journey. He recognized potential in me that I was unaware of, and for that, I am truly thankful. When I initially joined the MD2K lab seven years ago, I had no knowledge of sensor data-related work. Therefore, when my advisor, Dr. Santosh Kumar, suggested that I commence my research in the realm of smartwatch motion sensors, the transition appeared daunting at first. The way he presents problem statements and constructs robust issues broadened my intellectual horizons, fostering critical and consistent thinking, and ultimately enabling me to make significant contributions. I will always cherish his constructive and considerate suggestions that have influenced my life decisions, and I have been profoundly impacted by his philosophy. I could not have asked for a more exceptional advisor.

I had the privilege of collaborating with Dr. Vivek Shetty from UCLA School of Dentistry, who has been incredibly supportive and encouraging throughout my journey. I was pleasantly surprised by his generosity, as he took the time to provide quick feedback despite his busy schedule. His recognition of my hard work and words of congratulations served as a great inspiration to me. I would also like to express my gratitude to Dr. Emre Ertin from Ohio State University, for developing the wrist sensor and offering valuable insights. Additionally, I am thankful to the rest of my doctoral committee, Dr. Myounggyu Won and Dr. Weizi Li, for their invaluable guidance, feedback, and suggestions for my dissertation.

I had the privilege of collaborating with a multitude of brilliant researchers, and I am incredibly fortunate for that. Firstly, I would like to express my deep gratitude to my colleague, Dr. Nazir Saleheen, for his valuable contribution on my mORAL and mTeeth papers. I am also thankful to my another colleague, Mithun Saha, for his significant input in my mTeeth paper. Furthermore, I would like to express my appreciation to Dr. Soujanya

Chatterjee, who has been more than just a colleague, but also a supportive friend. I am grateful for the guidance and inspiration provided by my senior colleagues, Dr. Mahbubur Rahman and Dr. Rummana Bari, whose valuable suggestions have been instrumental in my work. During my time in Memphis, the support and companionship of my fellow graduate colleagues, particularly those in the MD2K group, have been of immense benefit. I am sincerely grateful for their encouragement and friendship.

I express my heartfelt gratitude to the University of Memphis Computer Science Faculty and the Graduate School for their unwavering support throughout my doctoral journey. Additionally, I extend my appreciation to the study coordinator, Shahin Alan Samiei, who collaborated with me during the ROBAS phase 1 data collection. His warm welcome and our discussions at his office about various aspects, such as data collection design, IRB approval, and participant recruitment procedures, were invaluable. I am also thankful to the participants of the ROBAS study for their involvement and contribution. Furthermore, I am grateful to Joseph Biggers, Cheryl Hayes, and Lyndsey Rush in our lab for their assistance, which made our administrative work more streamlined.

A special debt of gratitude goes to my parents and brothers. Their unwavering belief in my capabilities provided me with the strength and motivation to persist during moments of doubt. Their prayers and encouragement were invaluable sources of inspiration along this challenging path. I am incredibly grateful to my loving husband, Nazir Saleheen, who has been my rock throughout this entire journey. His constant support, guidance, and unwavering belief in my abilities have been paramount. It has been a remarkable experience for us to pursue our doctoral journey together. My heart is filled with overwhelming joy and gratitude as I celebrate the arrival of our beautiful baby, Hrid. Throughout my doctoral journey, the presence of my child has brought immeasurable happiness and purpose to my life. Hrid has been a constant reminder of what truly matters, and his smiles have fueled my determination during the toughest times.

My dissertation received support from various funding sources. The National Institutes of Health (NIH) provided funding under award R01DE024244 from the National Institute of Dental and Craniofacial Research, as well as awards P41EB028242, R01CA224537, R01MD010362, R01CA190329, R24EB025845, U01CA229437, and U54EB020404 through the trans-NIH Big Data-to-Knowledge (BD2K) initiative. Additionally, the National Science Foundation (NSF) supported my work through awards IIS-1722646, ACI-1640813, and CNS-1823221. I would also like to express my appreciation to Hansjoerg Reick and Ingo Vetter from Oral-B/Procter & Gamble for their valuable material and support.

ABSTRACT

Akther, Sayma. Ph.D. The University of Memphis. June 2023.
AI Modeling Approaches for Detecting, Characterizing, and Predicting Brief Daily Behaviors such as Toothbrushing using Wrist Trackers.
Major Professor: Dr. Santosh Kumar

Continuous advancements in wrist-worn sensors have opened up exciting possibilities for real-time monitoring of individuals' daily behaviors, with the aim of promoting healthier, more organized, and efficient lives. Understanding the duration of specific daily behaviors has become of interest to individuals seeking to optimize their lifestyles. However, there is still a research gap when it comes to monitoring short-duration behaviors that have a significant impact on health using wrist-worn inertial sensors in natural environments. These behaviors often involve repetitive micro-events that last only a few seconds or even microseconds, making their detection and analysis challenging. Furthermore, these micro-events are often surrounded by non-repetitive boundary events, further complicating the identification process. Effective detection and timely intervention during these short-duration behaviors are crucial for designing personalized interventions that can positively impact individuals' lifestyles.

To address these challenges, this dissertation introduces three models: mORAL, mTeeth, and Brushing Prompt. These models leverage wrist-worn inertial sensors to accurately infer short-duration behaviors, identify repetitive micro-behaviors, and provide timely interventions related to oral hygiene. The dissertation's contributions extend beyond the development of these models. Firstly, precise and detailed labels for each brief and micro-repetitive behavior are acquired to train and validate the models effectively. This involved meticulous marking of the exact start and end times of each event, including any intervening pauses, at a second-level granularity. A comprehensive scientific research study was conducted to collect such data from participants in their free-living natural environments. Secondly, a solution is proposed to address the issue of sensor placement variability. Given the different positions of the sensor within a wristband and variations in wristband placement on the

wrist, the model needs to determine the relative configuration of the inertial sensor accurately. Accurately determining the relative positioning of the inertial sensor with respect to the wrist is crucial for the model to determine the orientation of the hand.

Additionally, time synchronization errors between sensor data and associated video, despite both being collected on the same smartphone, are addressed through the development of an algorithm that tightly synchronizes the two data sources without relying on an explicit anchor event. Furthermore, an event-based approach is introduced to identify candidate segments of data for applying machine learning models, outperforming the traditional fixed window-based approach. These candidate segments enable reliable detection of brief daily behaviors in a computationally efficient manner suitable for real-time.

The dissertation also presents a computationally lightweight method for identifying anchor events using wrist-worn inertial sensors. Anchor events play a vital role in assigning unambiguous labels in a fixed-length window-based approach to data segmentation and effectively demarcating transitions between micro-repetitive events. Significant features are extracted, and explainable machine learning models are developed to ensure reliable detection of brief daily and micro-repetitive behaviors. Lastly, the dissertation addresses the crucial factor of the opportune moment for intervention during brief daily behaviors using wrist-worn inertial sensors. By leveraging these sensors, users can receive timely and personalized interventions to enhance their performance and improve their lifestyles.

Overall, this dissertation makes substantial contributions to the field of real-time monitoring of short-duration behaviors. It tackles various technical challenges, provides innovative solutions, and demonstrates the potential for wrist-worn sensors to facilitate effective interventions and promote healthier behaviors. By advancing our understanding of these behaviors and optimizing intervention strategies, this research has the potential to significantly impact individuals' well-being and contribute to the development of personalized health solutions.

TABLE OF CONTENTS

Contents **Pages**

LIST OF FIGURES

Figures **Pages**

LIST OF TABLES

Tables **Pages**

Chapter 1

Introduction

"An Ounce of Prevention Is Worth a Pound of Cure."

—Benjamin Franklin

In our day-to-day lives, we engage in a wide range of everyday behaviors that shape our overall well-being. These behaviors vary in duration, from mere milliseconds to several hours, and can have direct implications for our health. It is essential to follow the recommendations of healthcare providers to maintain a healthy lifestyle, which often includes specific behavioral practices. For instance, brushing our teeth correctly at least twice a day is recommended to prevent preventable oral diseases, while adhering to certain dietary guidelines and avoiding specific foods is necessary to prevent obesity. However, healthcare spending in the United States has become a major concern, projected to consume nearly 20 percent of the economy by 2026. The unsustainable nature of this spending has sparked a growing shift from reactive and expensive healthcare focused on treating illnesses to proactive, preventive approaches that address the underlying health behaviors that contribute to disease risk.

One key aspect of understanding and addressing these health behaviors lies in the accurate detection of these human behaviors. By precisely estimating the start and end times of these behaviors, we can gain valuable insights into individuals' overall behavioral patterns. This knowledge is crucial for designing effective interventions and personalized feedback systems that empower individuals to take greater ownership of their health and well-being. Moreover, the detection of specific short-duration behaviors, such as drinking, smoking, or toothbrushing, can provide valuable data for assessing and improving these activities. For example, analyzing each bite during eating helps estimate overall calorie intake, which is an important metric in avoiding obesity and maintaining a healthy weight. Similarly, the detection of micro-events within toothbrushing activities, such as individual brush strokes, allows for a more comprehensive evaluation of brushing quality and identifies areas where improvements can be made.

However, detecting these micro-events and short-duration behaviors is not without its challenges. Some of these behaviors occur within a few seconds or even microseconds, making their accurate detection more complex. Despite these challenges, advances in technology, such as sensor-enabled devices, provide promising avenues for precise behavior detection and monitoring. By leveraging these technologies, we can obtain a deeper understanding of these daily behaviors and develop interventions that are tailored to individuals' specific needs and characteristics.

Furthermore, it is worth noting that many daily behaviors are preceded by unique non-repetitive events, which can serve as crucial moments for designing appropriate interventions. For example, detecting the act of putting toothpaste on the brush head before tooth brushing, initial portion of tooth brushing or lighting up a cigarette before smoking provides opportunities to deliver timely interventions. These interventions can potentially prevent lapses and encourage individuals to adopt healthier habits. By capitalizing on these about-to-event or tooth brushing initiation moments, we can design targeted interventions that address specific behavioral challenges and promote positive changes in individuals' lives.

In conclusion, the detection of daily behaviors, ranging from brief micro-events to short-duration activities, plays a vital role in understanding individuals' behaviors and promoting healthier lifestyles. Accurate detection allows for personalized feedback, intervention design, and monitoring of progress. By embracing technological advancements and innovative approaches, we can unlock new opportunities for behavior detection and analysis, ultimately leading to improved health outcomes and a more proactive approach to overall well-being.

1.1 Motivating Application: Detecting, Characterizing, and Predicting Daily Toothbrushing Behaviors

Dental diseases, such as caries (tooth decay) and periodontal disease (gum disease), are pervasive chronic conditions with significant consequences for individuals and society as a whole [1]. In the United States, the prevalence of periodontal disease is alarmingly high, affecting approximately half of all adults, while over fifty-three million people live with untreated tooth decay in their permanent teeth [2]. The impact of these dental diseases extends far beyond oral health, as they can impair essential functions like eating, speaking, and socializing. Furthermore, they can give rise to local and systemic infections, causing significant discomfort and compromising an individual's overall well-being. The associated healthcare costs are substantial, and the burden of these expenses weighs heavily on individuals who lack adequate insurance coverage.

Notably, many dental diseases are preventable and closely linked to the performance of simple yet vital oral hygiene behaviors, such as regular tooth brushing and flossing. The American Dental Association (ADA) recommends that individuals brush their teeth at least twice daily and floss at least once a day, especially after meals [3]. However, research has shown that a significant proportion of the population fails to meet these recommendations. For instance, studies have revealed that 33% of men brush their teeth only once a day, while 59% of women regularly skip brushing at bedtime. These inadequate oral hygiene practices contribute to the persistence of dental diseases, as they often result in insufficient cleaning of each tooth surface. Certain areas may be missed entirely, while excessive time may be devoted to other areas. Over time, the accumulation of dental plaque—a colorless, sticky biofilm containing bacteria—can lead to gum disease, tooth decay, and eventual tooth loss.

In recent years, there has been a growing interest in utilizing mobile health (mHealth) approaches to measure and optimize oral hygiene behaviors, specifically tooth brushing and flossing. Researchers have focused on leveraging the sensing capabilities of electronic toothbrushes or developing smart toothbrushes to enhance oral care practices. However, these approaches often come with practical limitations. Users are required to actively activate data collection mechanisms and maintain specific positions relative to cameras during brushing, making these systems burdensome and less suitable for everyday use. To date, there is a lack of computational models that can passively and accurately detect tooth brushing and flossing behaviors using regular toothbrushes in real-world settings.

Addressing this gap in research, the development of a robust computational model capable of reliably detecting and characterizing tooth brushing and flossing behaviors in natural field environments becomes essential. Such a model would enable continuous monitoring and provide individuals with valuable feedback on their oral hygiene practices. By passively capturing data on brushing duration, technique, and frequency, it could facilitate personalized interventions, encouraging individuals to adopt and maintain effective oral hygiene habits. Ultimately, the aim is to improve oral health outcomes, reduce the prevalence of dental diseases, and enhance overall well-being for individuals across diverse populations.

1.2 Proposed Setup and Approach

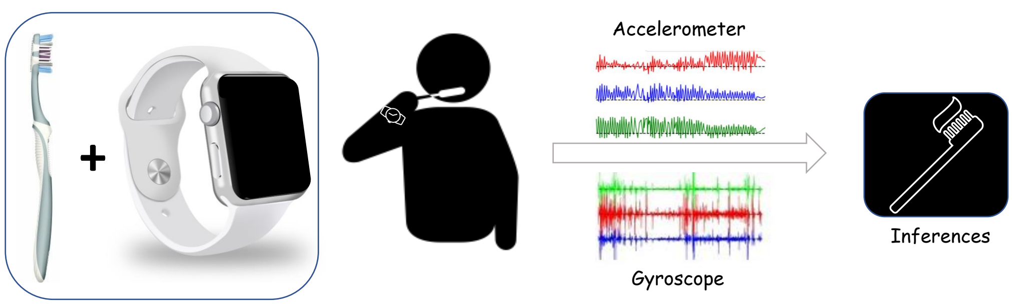


Fig. 1.1: Overview of the proposed approach

Significant hand movement is required to perform brief, micro, and boundary behaviors related to oral hygiene. These behaviors encompass actions such as brushing teeth, flossing, and other related activities. In recent years, there has been a growing interest in utilizing wearable devices, particularly wrist wearables like smartwatches and wristbands, due to their increasing prevalence among the general population. According to projections, the sales of wrist wearables were expected to reach an impressive 200 million units by 2022 [4].

The appeal of wrist wearables lies in their ability to gather various types of data, including vital signs (e.g., heart rate, skin temperature) and activity data (e.g., movements). This wealth of information can be harnessed to infer and understand a range of health behaviors that occur in real-world settings. Moreover, this data can serve as a foundation for developing engaging and persuasive health messages and interventions [5].

To leverage the capabilities of wrist wearables for monitoring oral hygiene behaviors, it is crucial to develop robust computational models capable of extracting relevant features from the signals captured by the inertial sensors embedded within these devices. These signals may include parameters such as speed, wrist rotation, and arm displacement. By analyzing and interpreting these features, we can effectively detect and classify oral hygiene behaviors, such as brushing and flossing, within the natural environment.

Our approach focuses on utilizing wearable sensor technology, with specific emphasis on wrist-worn devices. Given that over 80% of the population employs a manual toothbrush for oral hygiene practices [6], we have chosen to concentrate on toothbrushing with a manual toothbrush as our primary behavior of interest. Notably, smartwatches have become increasingly integrated into our day-to-day lives, making them an ideal source of information for capturing relevant data. In our methodology, we utilize the data collected from the 3-axis accelerometer and 3-axis gyroscope sensors embedded within the wristwatch. These sensors provide valuable time-series data that will be utilized in the development of machine learning models for accurate detection and classification of oral hygiene behaviors.

By harnessing the power of wrist-worn wearable devices and employing advanced computational techniques, we aim to enhance the understanding and monitoring of oral hygiene behaviors in a naturalistic setting. This research has the potential to contribute valuable insights into individuals' oral health habits, paving the way for personalized interventions and improvements in oral hygiene practices.

1.3 Research Gap

This section discusses several research gaps that motivate us to work on this problem and rigorously investigate different aspects of the problem.

This section explores various research gaps that serve as the driving force behind our work, compelling us to thoroughly investigate different facets of the problem at hand.

1.3.1 Existing Activity for Daily Living (ADL) Detection “only" Demonstrates Feasibility

The detection of Activity for Daily Living (ADL) behaviors using wrist-worn inertial sensors has emerged as a rapidly growing research field. These sensors have proven to be effective in detecting a wide range of ADL activities, including walking, sleeping, eating, combing hair, dressing, climbing stairs, sitting, standing, and cooking [7, 8, 9, 10, 11, 12].

Some previous studies [7, 13] have explored the feasibility of using wrist-worn inertial sensors to detect toothbrushing behavior by analyzing hand gestures. However, it is important to note that these studies primarily focused on demonstrating the potential of detecting a broad range of ADL activities and, therefore, trained their models using data collected in controlled or scripted settings. As a result, their models exhibited a false positive rate exceeding 15% for each class, including toothbrushing. This high false positive rate makes these models unsuitable for practical, real-world applications, particularly for passive detection scenarios where the aim is to minimize erroneous detections.

To address this limitation, our research aims to develop a robust and reliable toothbrushing detection model that can be seamlessly integrated into daily routines. Unlike previous studies, we recognize the importance of training our model using data collected in naturalistic environments to ensure its effectiveness in real-world settings. By leveraging datasets acquired from individuals performing toothbrushing tasks in their own homes, we can capture the true variability and nuances of toothbrushing behaviors. This approach allows us to develop a model that exhibits higher accuracy and specificity, resulting in significantly reduced false positive rates compared to previous methods.

By focusing on refining toothbrushing detection using wrist-worn inertial sensors, we aim to bridge the research gap in reliable oral hygiene behavior monitoring. This advancement will contribute to the development of innovative technologies and interventions that promote better oral health practices and ultimately enhance overall well-being.

Table 1.1: This table shows the number of false positive events produced per day by a model for specific false positive rates. It is assumed that sensors are worn for 16 hours per day and the toothbrushing event lasts an average of 2 minutes.

False positive rate	**15%**	**10%**	**5%**	**1%**	**0.1%**	**0.01%**
False positives per day	72	48	24	4	$\frac{1}{2}$	$\frac{1}{20}$

1.3.2 Limitation of Existing Micro-repetitive Modeling

Developing behavior-specific models has become essential to ensure a more reliable and continuous detection of target behaviors in natural field settings. However, applying existing models designed for different behaviors poses a significant challenge when it comes to detecting specific behaviors like brushing or flossing. Behavior-specific models need to be developed for each distinct behavior, as models designed for activities such as walking [12, 14] cannot be directly applicable to eating, smoking [15, 16], or oral hygiene behaviors like brushing or flossing. Consequently, the reliable and passive detection of brushing behaviors using wrist-worn sensors remains an open problem.

One primary issue is that most existing models are not tested on continuous data collected in real-world field settings, which is the typical scenario for everyday use. Only a few researchers have tested their models on continuous passive data collected from wrist sensors, specifically for behaviors such as eating [15] or smoking [16, 17].

Moreover, the sporadic and transient nature of toothbrushing presents a unique challenge. This activity typically lasts for just a few minutes within a day that spans 12 to 16 hours. When applying existing models during the waking hours of the day, it leads to a significant number of false positives. Even with a reported 1% false positive rate, this would result in four false positive detections per day (as illustrated in Table 1.1). Hence, effectively detecting toothbrushing behaviors passively using wrist-worn sensors requires novel approaches and tailored models that address the specific characteristics and challenges of this behavior.

1.3.3 Limitation of Existing Toothbrushing Methods

Maintaining proper oral hygiene is essential for overall dental health, and advancements in technology have paved the way for innovative solutions. Smart toothbrushes, instrumented toothbrushes, and audio/video-based monitoring systems have emerged as promising approaches. These systems offer real-time feedback, motion tracking, and even acoustic or visual analysis to enhance oral hygiene practices. However, they still face limitations that hinder their widespread adoption and effectiveness. Addressing these limitations is crucial to develop reliable and unobtrusive methods for monitoring oral hygiene behaviors.

Smart Toothbrush based Solutions:

In recent years, significant efforts have been directed towards the detection of oral hygiene behaviors (OHBs) by capitalizing on the sensing capabilities of advanced electronic toothbrushes. Smart toothbrushes have emerged as a diverse range of solutions, each offering unique features to enhance oral care practices. These innovative devices are designed to detect and provide users with real-time feedback on various aspects of brushing. Smart toothbrushes have emerged as a diverse range of solutions, each offering unique features to enhance oral care practices. These innovative devices are designed to detect and provide users with real-time feedback on various aspects of brushing. For instance, some smart toothbrushes can accurately measure the pressure applied to the teeth during brushing, aiding in the promotion of optimal brushing techniques [18]. Others are equipped with miniature cameras integrated into the toothbrush head, allowing for the detection of plaque and precise monitoring of brushing motions [19]. Moreover, implantable assistive brushing devices have been developed to support individuals with special needs, such as children or those with disabilities, in maintaining proper oral hygiene [20].

Researchers have also explored the integration of advanced sensors into smart toothbrushes to enable more detailed monitoring. In studies like [21] and [22], smart toothbrushes were embedded with a 3-axis accelerometer and a magnetic sensor, enabling the tracking of specific groups of teeth being brushed at any given moment. These innovative solutions have paved the way for commercially available smart toothbrushes that guide users in achieving effective brushing habits. For instance, well-known brands like [23] have introduced smart toothbrushes with brushing heads capable of providing real-time feedback based on the applied pressure. Paired with smartphone applications, these devices offer visual displays that help users identify and target specific tooth surfaces during brushing.

Instrumented Toothbrush based Solutions:

As an alternative to conventional smart toothbrushes, researchers in [24] introduced a novel approach using smartwatches for recognizing and evaluating brushing quality. They devised a system that involved attaching magnets to a regular toothbrush, enabling the collection of inertial data from wrist-worn sensors. By analyzing the arm motion patterns captured by the sensors in real-time, brushing gestures could be detected and assessed. Similarly, in [25], a unique solution was proposed by attaching a 3D colored ball to the end of a toothbrush. Through spatial analysis of the ball's position and orientation during brushing, the system could estimate which areas of the teeth were being targeted. These innovative methods showcase the potential of leveraging wearable devices and creative attachments to monitor oral hygiene behaviors, providing valuable insights into brushing techniques. Such approaches contribute to the development of advanced technologies for promoting effective oral care practices.

Audio and Video based Solutions:

In an initial study [26], the evaluation of brushing techniques was conducted by analyzing acoustic signals captured by a smartphone positioned near the sink. Similarly, in [27], a tooth brushing monitoring system was proposed, relying on acoustic inputs. The system incorporated an asymmetrical sound-field detector, comprising a Bluetooth earphone and a throat microphone, to capture audio data from the surrounding environment and the user's body, respectively.

Other approaches utilize image analysis to detect tooth surfaces. For instance, in [28], a computer-based webcam was employed to identify the position of a smart toothbrush. The system featured a visual feedback mechanism, utilizing a physical avatar with LED teeth, enabling real-time tracking of children's tooth-brushing activities. Another study [25] utilized the front camera of a smartphone to detect both the toothbrush and the user's face. By employing a face tracker and replacing the captured image with that of an avatar on the smartphone's display, a "virtual mirror" effect was achieved. The avatar was capable of accurately mimicking the user's gestures and expressions, while also providing guidance on correct brushing movements.

These innovative approaches demonstrate the diverse ways in which acoustic signals and image analysis can be utilized to monitor and enhance tooth-brushing practices. By leveraging smartphones and computer vision technologies, these systems contribute to the advancement of oral hygiene monitoring and offer personalized feedback for improving brushing techniques.

In summary, while there have been notable advancements in the field of smart and instrumented toothbrushes, these solutions have certain limitations when it comes to automatically detecting brushing or flossing events using wrist-worn inertial sensors in a day-long wearing scenario. These existing approaches typically rely on user-initiated actions, such as pressing a button, to indicate the start and end of brushing activity, which may not be suitable for continuous and passive monitoring. Additionally, many of these methods require some level of instrumentation, whether it be sensors embedded in the toothbrush or environmental setup for audio/video analysis, to detect tooth surfaces.

To develop a more unobtrusive and reliable method for monitoring oral hygiene behaviors, it is necessary to explore alternative approaches that leverage wrist-worn inertial sensors without the need for user interaction or additional instrumentations. By addressing these challenges, researchers can pave the way for innovative solutions that seamlessly integrate into daily routines and provide accurate monitoring of oral hygiene practices.

1.4 Problem Formulation

In this section, we aim to address the research gap by providing a detailed overview of the problem at hand. Our goal is to tackle the challenges associated with monitoring oral hygiene behaviors using wrist-worn inertial sensors. To achieve this, we divide the problem into three distinct subproblems, each highlighting a specific aspect of the overall problem.

1.4.1 Detection Problem

The detection problem involves accurately pinpointing the timing of brief oral hygiene events, such as brushing or flossing, from the continuous time-series data collected by wrist-worn inertial sensors. With access to data from sensors like the 3-axis accelerometer and 3-axis gyroscope, our objective is to determine the precise start and end times of these events. Since brushing and flossing activities are relatively transient, lasting only a few minutes out of the approximately 960 awake minutes per day, it is crucial for our detection algorithm to closely match the actual event's duration and timing.

1.4.2 Characterization Problem

The characterization problem revolves around obtaining detailed insights into the brushing process by determining the duration of brushing on different surfaces. This involves dividing the entire teeth region into meaningful and distinct surfaces and assigning them appropriate labels. Once the surfaces are defined, the challenge is to identify the specific time intervals during the brushing event when each surface is being brushed. These micro-detections can be aggregated to generate a comprehensive summary of the entire brushing event. While detecting brushing events provides an overall understanding of brushing duration, surface detection enables a more precise assessment of brushing quality.

1.4.3 Prediction Problem

The prediction problem revolves around leveraging the insights gained from previous brushing events to effectively guide and improve future oral hygiene practices. By accurately identifying about-to-brushing events or the commencement of brushing, we can intervene in a timely manner and provide guidance based on the summarized information gathered from past brushing events. About-to-brushing events pertain to actions like applying toothpaste to the toothbrush immediately before brushing, while the beginning of brushing refers to the initiation or initial strokes of the toothbrushing event. By detecting these pivotal moments, we can deliver timely interventions and tailored guidance, drawing from the aggregated data of toothbrushing activities from previous days. This approach encourages the development of consistent and precise brushing habits. The ability to predict and intervene before a brushing event begins plays a vital role in establishing effective oral hygiene practices.

Our objective is to tackle these three interrelated subproblems in order to create robust and efficient approaches for automatically monitoring oral hygiene behaviors utilizing wrist-worn inertial sensors. By making advancements in the areas of detection, characterization, and prediction, we aim to bridge the divide between technology and oral health. This endeavor holds the potential to significantly enhance oral hygiene practices and contribute to overall dental well-being.

1.5 Overview of the Proposed Solution

This dissertation presents a comprehensive approach aimed at tackling three interconnected subproblems (Fig. 1.2), with a specific focus on toothbrushing as the case study. Our overarching goal is to develop a practical and effective model capable of accurately detecting short brushing events using wrist-worn inertial sensors. To achieve this, the model needs to exhibit a high recall, capturing the majority of brushing events, while maintaining a high precision to minimize false positives.

Once the start and end times of brushing episodes are obtained from the brushing detection model, our research extends to building another model specifically designed to detect the brushing surface using data gathered from the accelerometer and gyroscope sensors.

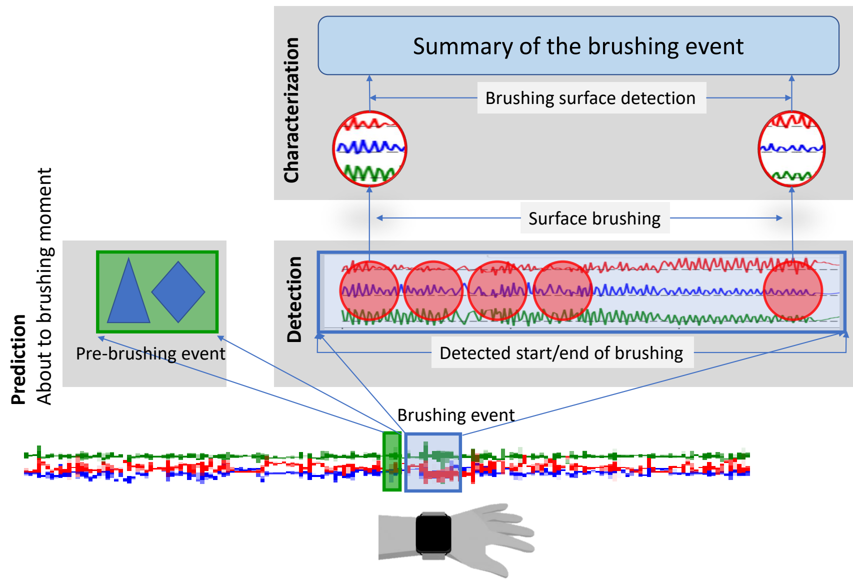


Fig. 1.2: Overview of the proposed comprehensive solution

By analyzing the duration of each brushing surface, we can assess the overall quality of the brushing event, providing valuable insights for oral hygiene evaluation.

To further support the maintenance of a healthy oral hygiene routine, we recognize the importance of scheduling timely interventions to enhance the quality of the overall brushing event. Therefore, we delve into developing two models that detect the crucial "about-to-brush" moment or the initiation of toothbrushing moment. These moments serves as a key component of our proposed approach, enabling us to provide real-time guidance and interventions based on previous brushing events.

Through our research, we aim to bridge the gap between technology and oral health, ultimately leading to improved oral hygiene practices and better overall dental health. By addressing these subproblems, we contribute to the development of comprehensive and effective methods for automatic monitoring of oral hygiene behaviors using wrist-worn inertial sensors. Our work holds the potential to revolutionize oral health practices by leveraging advancements in detection, surface characterization, and prediction techniques, fostering the cultivation of consistent and accurate brushing habits.

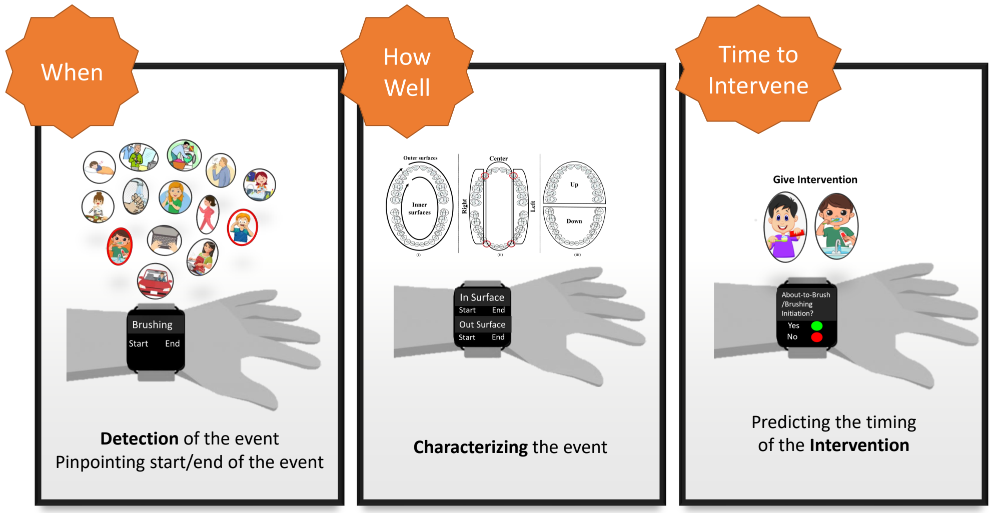


Fig. 1.3: Three main contributions of this work.

1.6 Summary of Key Contributions

To reliably develop our models, we first design and conduct a research study where we collect wrist-worn sensor data for seven continuous days from 30 participants. For preciously labeling ground truth, we also asked the participants to record videos while performing tooth-brushing activities. We make this dataset open publicly to research purposes.

As shown in Figure 1.3, this dissertation makes three technical contributions. First, *when does the event occur?* That is detecting the start and end of the tooth-brushing event. Second, *How well does the user perform the event?* That is finding the quality of the brushing event or characterization of the detected brushing event. This information can be delivered to the user as an intervention to guide the user towards correct brushing habits. Finally, *predicting the timing of the intervention.* The summary of the contributions is described in the following.

1.6.1 Detection of a Brief Daily Behavior

We develop a robust ML model, named mORAL[29], for detecting brushing with manual toothbrushes and flossing behaviors from wrist-worn inertial sensors. We show that for detecting brief daily events such as toothbrushing, adopting a model that is based on identifying candidate windows based on events rather than fixed-length time blocks, leads to significantly higher performance. Towards the development of this model, we solve a well-known open problem, named virtual orientation, which identifies correct sensor mounting on the device and proper placement of the device on the wrist. We annotated more than 100 videos that are used as ground truth. Trained and tested on 2,797 hours of sensor data collected over 192 days on 25 participants (using video annotations for ground truth labels), our brushing model achieves 100% median recall with a false positive rate of one event in every nine days of sensor wearing.

1.6.2 Characterization of the Detected Brief Daily Behavior

After detecting the brushing event, our task is to characterize the brushing event. For that, we develop the mTeeth model [30] to detect teeth surfaces being brushed with a manual toothbrush in the natural free-living environment using wrist-worn inertial sensors. We solve another open problem of detecting brushing strokes from a wrist-worn accelerometer sensor. To unambiguously label sensor data corresponding to different surfaces and capture all transitions that last only milliseconds, detected brushing strokes cleanly demarcate transitions among brushing surfaces. We annotate 10,230 instances of brushing on different surfaces for training and testing and evaluate the impact of wide between-person and within-person between-episode variability on the machine learning model's performance for brushing surface detection. Despite high person variability, the model can summarize brushing time on each surface with more than 92.5% accuracy.

1.6.3 Detection of About-to-Brush Moment/Brushing Initiation

We introduce two novel approaches: detecting about-to-brush moments and detecting the initial portion of tooth brushing. These methods leverage synthetic data generation and a Stroke Detection and Clustering algorithm to enhance the accuracy of detecting brushing behavior. By addressing the timing of interventions and improving detection techniques, we open up a design a system that provides intervention within three seconds to users to remember to brush the right surface. By using audio or vibration a system can design to provide intervention within three seconds where as current smart brush give intervention in 30 seconds.

1.7 Thesis Organization

The structure of this dissertation is outlined as follows: Chapter 2 introduces the ROBAS study, providing an in-depth understanding of its methodology. In Chapter 3, we address the wrist mounting problem, discussing the challenges and solutions related to wearing the device. The *mORAL* model is presented in Chapter 4, where we delve into its development and implications. Moving on to Chapter 5, we introduce the *mTeeth* model, focusing on its design and functionality. Chapter 6 is dedicated to the intervention, showcasing its implementation and impact. Lastly, Chapter 7 provides a concise overview of the dissertation's contributions and future goals.

Chapter 2

Labeled Data Collection: ROBAS Phase 1 Research Study

We conducted a real-time field study, referred to as Remote Oral Behaviors Assessment System (ROBAS), aimed at collecting data on regular oral health routines. The study employed wearable sensors to measure wrist movement during daily activities, while also utilizing a smartphone to gather ground truth data. Participants were instructed to perform their usual tooth brushing, flossing, and oral rinsing routines while wearing the sensors and keeping the study smartphone with them throughout the day. To ensure data preservation, a software system was employed to store both the sensor data and ground truth information. The primary objective of the study was to detect oral health behaviors using wrist-worn inertial sensors. The study was carefully designed to fulfill all our goals and requirements. Prior approval for this study, with a protocol number 4274, was obtained from the Institutional Review Board (IRB) at the University of Memphis, and written informed consent was obtained from all participants. This chapter provides a comprehensive discussion of the overall data collection process, the devices used, the total amount of collected data, the participant enrollment protocol, and the software employed for data storage.

2.1 Study Requirements

Our study design encompasses the necessary elements to fulfill multiple requirements in the production of sensor data, associated labels for model development, and software updates for data capture. Firstly, we focus on wearable devices and sensor measurements, specifically wrist-worn inertial sensors, to capture hand movement during various daily oral hygiene behaviors in natural field environments. Secondly, we employ video measurement to precisely locate the start and end times of each brief and micro event, requiring the capture of corresponding videos during these events. Additionally, we update the software, specifically utilizing *mCerebrum*, to accommodate the storage of large-scale sensor data while ensuring the capture of videos for ground-truth annotations. In the subsequent section, we will provide a comprehensive description of these aspects.

((a)) Wristband sensors

((b)) Smartphone

((c)) SmartBrush

Fig. 2.1: Wearable wristband sensors, study smartphone and SmartBrush used in the study

2.1.1 Study Design for Capturing Wrist-Worn Sensor Data

This section outlines the study design aimed at capturing wrist-worn inertial sensor data during real-life daily behaviors, with a particular focus on oral hygiene behaviors that involve significant hand movement. The study utilized wearable devices and a smartphone carried by each participant to capture the necessary data.

1. **Wristband Sensors:** Participants were equipped with wristbands on both wrists (left and right) to capture their wrist movements during daily activities, as shown in Figure 2.1(a). These wristbands were equipped with a suite of sensors, including a 3-axis accelerometer sampled at 16 Hz and a 3-axis gyroscope sampled at 32 Hz. The sen-

sors were worn by participants throughout their waking hours, excluding bathing and swimming, to capture wrist movement during various activities. The accelerometer measured the linear acceleration along the three axes, while the gyroscope measured the angular velocity. Participants engaged in teeth brushing twice a day and flossing once a day, during which wrist movements were captured by the wrist-worn inertial sensors (Figure 2.1). There were no specific time instructions given for brushing and flossing, allowing participants to perform these activities naturally according to their habitual routines. Consequently, participants were instructed to wear the wristbands throughout their waking hours, except during bathing and swimming.

2. **Smartphone Integration:** Each participant was provided with an Android smartphone as part of the study. The smartphone served multiple purposes, including communication with, data reception, and timestamping from the sensor suites. In addition to its primary functions, the phone collected data through its internal sensors, which included a 3-axis acceleration sensor, a 3-axis gyroscope, GPS traces for geo-location data, battery state information, and user interaction data.

3. **SmartBrush Incorporation:** As a form of compensation for their participation, participants were provided with a commercially available Bluetooth-enabled toothbrush known as "Oral-B." Throughout the study, participants were instructed to incorporate three oral hygiene behaviors into their daily routine while wearing the sensors. Specifically, they were asked to use their personal (manual) toothbrush once daily, utilize the SmartBrush once daily, and perform flossing at least once daily.

2.1.2 Video-Based Ground Truth Generation for Precise Detection

In order to accurately classify brushing and flossing activities from participants' daily behaviors, it was crucial to obtain precise information regarding the start and end times of these oral hygiene events. However, relying solely on participants' self-reporting or recollection often led to inaccuracies and added extra burden to the participants. To address

Fig. 2.2: Setup for collecting toothbrushing video data: participants capturing videos during toohbrushing

Fig. 2.3: Dental flossing with string

this issue, a method was employed to alleviate the reporting burden and ensure accurate timing estimation. Participants utilized the front-facing camera of the study smartphone to record videos of themselves while performing their oral health routines, including brushing, flossing, and using oral rinse (Figure 2.2 and 4.8). By storing the timing of the start and end of brushing and flossing events within the recorded videos, the smartphone provided ground truth information for accurately detecting these behaviors from the inertial sensor data captured by the wristband.

2.2 A Software Platform for High-Frequency Sensor Data Collection

The detection and validation of daily behaviors in research studies rely on the collection of high-frequency sensor data, specifically from accelerometers and gyroscopes. To accommodate these data collection requirements, an ideal software platform was paramount. The software needed to establish a synchronous connection to handle the influx of high-frequency sensor data. In order to cater to the data collection requirements, the software platform should encompass the following key features:

- **Synchronous Connection and Sensor Compatibility:** The software must support a synchronous connection for high-frequency sensor data and have the flexibility to connect new sensors seamlessly. This ensures that data from multiple sensors can be synchronized accurately.

- **Robust Data Reception and Handling:** Given the large volume of frequently arriving sensor data, the software must have a reliable operating system capable of efficiently receiving and processing data. Data loss can compromise the integrity of the study, making it crucial for the software to handle data streams effectively.

- **Video Recording Functionality:** The software should include built-in capabilities to record self-video data. This feature is essential for generating ground truth annotations of oral hygiene behaviors, enabling accurate validation of the detected behaviors.

- **User-Friendly Interface:** To ensure participant convenience and compliance, the software should have a user-friendly interface. Participants should be able to easily navigate and operate the software, allowing them to attach and manage the sensors without feeling burdened or overwhelmed.

By fulfilling these requirements, the software platform can effectively support the collection of high-frequency sensor data and facilitate the accurate detection and validation of daily behaviors, particularly in the context of oral hygiene routines. In our study, we utilized our lab-made software platform called mCerebrum [31] to facilitate the collection of sensor data. We made necessary updates to the software to accommodate the collection of wrist-worn inertial signals during oral hygiene behaviors, while also capturing videos of these behavior events in real-time.

2.2.1 Storage of High-Frequency Sensor Data: The mCerebrum Approach

One of the key challenges in mobile sensing is the storage and management of high-frequency sensor data. To address this challenge, mCerebrum utilizes a distributed data storage architecture that enables efficient storage and retrieval of sensor data. The platform leverages the storage capabilities of both the mobile device and cloud infrastructure to ensure scalable and reliable data storage. Sensor data is collected at high sampling rates, often ranging from tens to hundreds of Hertz, resulting in large volumes of data. To optimize storage efficiency, mCerebrum employs compression techniques and data reduction algorithms that selectively store relevant features or summary statistics instead of raw sensor data. This approach reduces the storage footprint while preserving the necessary information for subsequent analysis and interpretation.

In addition to compression and data reduction techniques, mCerebrum incorporates data partitioning strategies to efficiently store and retrieve sensor data. The platform intelligently partitions the data based on time intervals, allowing for faster access to specific segments of interest during analysis. By dividing the data into smaller units, mCerebrum enables targeted retrieval and processing, minimizing the computational overhead associated with handling large datasets.

To ensure data integrity and reliability, mCerebrum incorporates robust error-checking mechanisms and redundancy measures. Data is checksummed and verified during storage and retrieval processes to detect any potential corruption or data loss. Furthermore, redundant copies of the sensor data can be stored across multiple storage locations, such as the mobile device, cloud servers, and external storage devices, providing an additional layer of data protection.

mCerebrum also supports seamless synchronization and data transfer between the mobile device and cloud servers. This enables continuous backup and synchronization of sensor data, ensuring that valuable data is not lost in the event of device failure or loss. Moreover, the platform offers data encryption options to protect sensitive user information, adhering to strict privacy and security protocols.

By employing these techniques and strategies, mCerebrum effectively addresses the challenges associated with high-frequency sensor data storage, providing a robust and efficient platform for the development and validation of digital biomarkers and interventions.

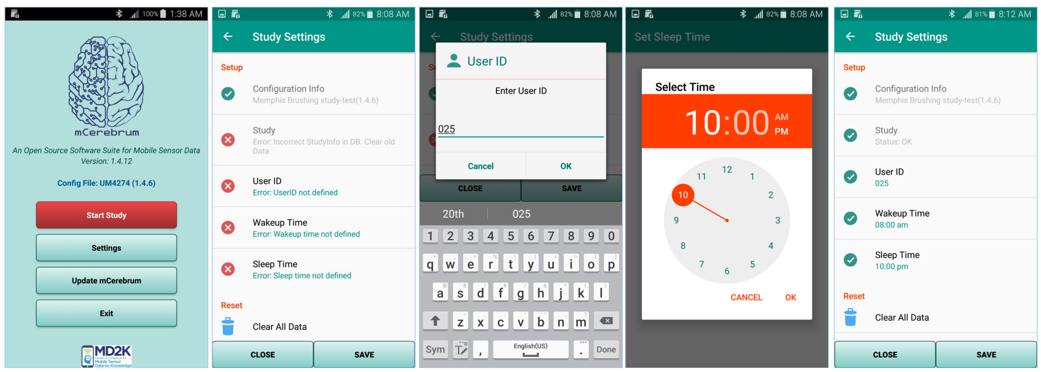


Fig. 2.4: Configuration of study setting in the mCerebrum software prior to study initiation

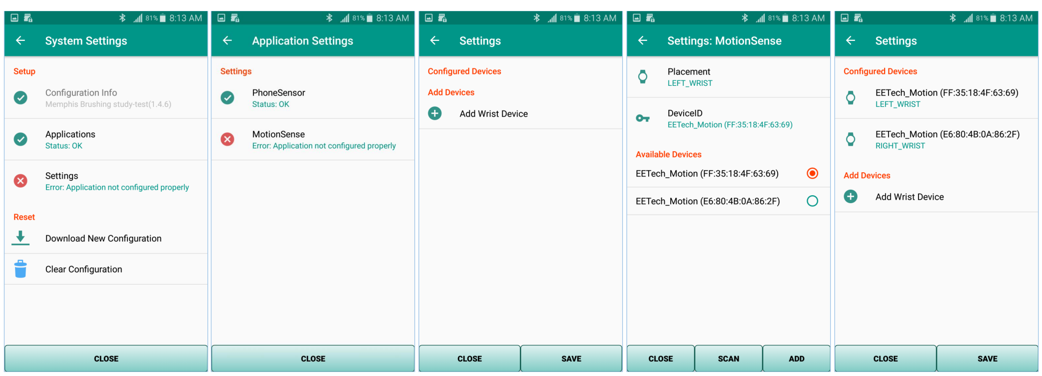


Fig. 2.5: Setup of inertial sensors on the left and right wrist

2.2.2 Enhanced mCerebrum for ROBAS Data Collection

The updated version of mCerebrum introduces new capabilities to support high-rate data collection from multiple sensors, offering real-time assessment of data quality. Its scalable storage architecture ensures efficient handling of rapidly growing data volumes while maintaining optimal performance. Figure 2.4 provides an overview of the study

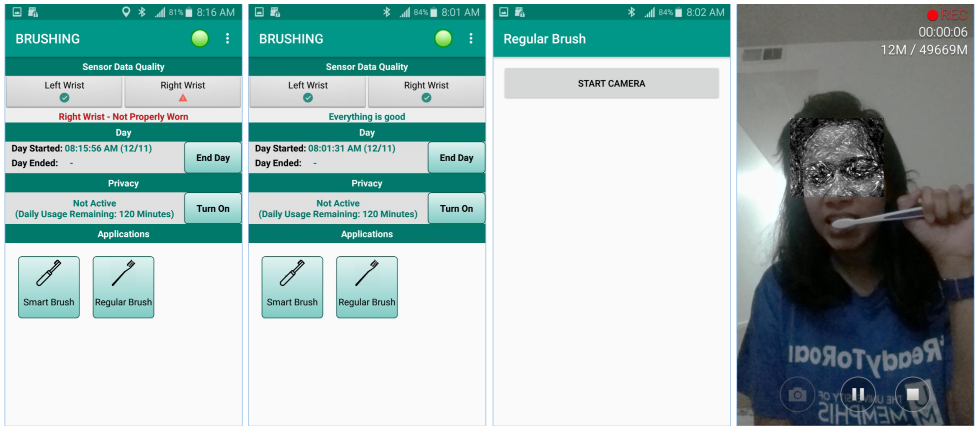


Fig. 2.6: Interface for monitoring sensor connection status and capturing brushing videos during the study

setup, showcasing the configuration of wrist sensors, software integration, participant ID, and sleep duration monitoring. Figure 2.5 illustrates the connection status of the left and right wrist sensors, along with indicators for active and idle sensor states.

To enable seamless data collection and storage, modifications were made to the mCerebrum software, which was then installed on the study smartphone. This empowered participants to engage with the system and initiate video recordings while ensuring accurate time synchronization between the captured videos and wristband data. Figure 2.6 exemplifies the precise synchronization during self-recorded brushing moments, displaying the connection status of the left and right wrist sensors.

In addition to technical enhancements, the software underwent user-oriented modifications to provide an intuitive and user-friendly interface. This optimized interface allows participants to navigate and operate the software with ease, effortlessly managing the attachment and configuration of the sensors. By reducing any potential feelings of burden or overwhelm, the user-friendly design promotes seamless engagement and facilitates optimal participation.

With its upgraded functionality and user-centric improvements, mCerebrum empowers researchers to enhance ROBAS data collection, ensuring both accuracy and a positive user experience for study participants.

2.3 Study Protocols and Participant Engagement: A Comprehensive Overview

During the initial phase of participant recruitment, individuals were surveyed regarding their oral hygiene behaviors (OHBs) and their willingness to engage in brushing and flossing routines. It was specifically inquired if they were committed to brushing at least twice daily and flossing at least once daily. Following enrollment, phone interviews were conducted to verify the eligibility of participants for an in-person screening. Subsequently, two in-person visits were arranged, each lasting approximately one hour. To gain a comprehensive understanding of the study environment and procedures, a combination of qualitative methods, including interviews, focus groups, observations, and case studies, was employed. During the first visit, a detailed explanation of the study procedures was provided, along with a comprehensive overview of potential risks and discomforts. The second in-person visit involved soliciting participant feedback and responses regarding their experiences and attitudes toward the overall study and the utilized devices.

In addition to the qualitative approach, quantitative data collection procedures were implemented. Participants were instructed to wear wristbands consistently throughout their waking hours for a continuous seven-day period. These wristbands collected valuable data from the wrist sensors, including measurements from accelerometers and gyroscopes. The continuous usage of wristbands enabled seamless collection of wrist sensor data, providing precise insights into participants' physical movements and gestures. Furthermore, participants were directed to use the study smartphone daily to record videos of themselves engaging in OHBs, such as brushing and flossing, within the comfort of their homes. It was crucial for participants to wear the wrist sensors during these activities, as depicted in Figure 2.2. By combining accelerometer and gyroscope measurements from the wrist sensors with the recorded videos, a comprehensive analysis of participants' oral hygiene practices was conducted. The dedicated commitment of participants to wearing the wristbands and recording OHBs videos played a vital role in generating accurate and comprehensive data, ultimately contributing to the overall success of the study.

2.3.1 Effective Recruitment Strategies and Participant Criteria

We successfully employed flyers as a strategic method for recruiting research volunteers by strategically placing them on bulletin boards throughout various university buildings. Furthermore, we utilized word-of-mouth techniques within different community networks and friend groups to effectively disseminate information about the research study. Our primary objective was to attract individuals who demonstrated a strong interest in actively participating in scientific research and contributing to valuable scientific advancements. To achieve this, we sought volunteers who met specific criteria, including:

1. Being enthusiastic about participating in research studies and making valuable contributions to scientific advancements.

2. Having a strong commitment to maintaining good oral hygiene habits, such as brushing their teeth at least twice daily and flossing at least once daily.

3. Being comfortable with wearing and using mobile sensors, such as wristbands, study smartphones, and sensor-enabled toothbrushes, for the duration of the study period.

4. Possessing the willingness to record videos of themselves engaging in oral health behaviors, such as brushing or flossing, as part of the data collection process.

5. Being available to attend two in-person lab sessions at the beginning and end of the study period, each lasting approximately one hour.

6. Falling within the age range of 18 to 64 and being in good overall health.

By specifically targeting individuals who met these defined criteria, our aim was to ensure the formation of a participant group that was highly motivated, actively engaged, and aligned with the study's objectives.

2.3.2 Initial Phone Screening

During the initial phone screening, we engaged potential participants in a conversation to assess their suitability for the research study. In addition to gathering information about their oral hygiene habits, age, and oral health conditions, we also explored their level of comfort with wearing and using mobile sensor devices. This step was crucial in identifying individuals who were not only committed to maintaining good oral hygiene but also comfortable with the study's requirements, such as wearing wrist sensors and using a sensor-enabled toothbrush. By conducting a comprehensive phone screening, we aimed to ensure that the selected participants were well-informed about the study's expectations and motivated to actively participate. The screening process served as a valuable opportunity to establish a strong foundation for the subsequent in-person screening and ultimately shape a participant group that aligned with the study's objectives.

2.3.3 In Person Interview: Establishing Connection and Gathering Information

Interviews may be structured and conducted under controlled conditions, or they may be conducted with a loose set of questions asked in an open-ended manner. When gathering demographic data, such as age, interview questions can also be quantitative in nature. We collected written informed consent A from the participant.

Prior to the initiation of the study, a face-to-face interview was conducted to establish a meaningful connection with the participants and gather essential information. These interviews were designed to be flexible, allowing for structured or semi-structured conversations to capture valuable data. Through these personal interactions, participants were encouraged to openly express their thoughts, experiences, and expectations related to the study. This interview process served as a platform for collecting qualitative insights and additional details that contributed to a comprehensive understanding of the participants' perspectives. Additionally, participants received a comprehensive overview of the study procedures, including its purpose, potential risks, and benefits. The informed consent A process was carefully conducted to ensure participants were fully informed of their rights and responsibilities.

The demographics questionnaire A played a pivotal role in collecting relevant information about the participants. It aimed to gather demographic data, such as gender, age, race, yearly income, and oral history. Furthermore, the questionnaire sought to explore participants' oral habits, including the frequency of brushing and flossing, any existing oral health conditions, and their previous experiences with oral health studies or interventions. This comprehensive data collection enabled us to develop a holistic understanding of the participants' background and oral health profiles. The insights obtained from the demographics questionnaire played a significant role in the overall analysis and interpretation of the study results. Moreover, it facilitated tailoring the research study to address the specific needs and characteristics of the participants, ensuring a personalized approach to data collection and analysis.

By conducting in-person interviews, our aim was to establish a strong rapport with the participants and foster a collaborative partnership throughout the research study. This approach not only facilitated data collection but also emphasized the importance of participant engagement and involvement in the study.

2.3.4 The Exit Interview: Insights and Reflections

The exit interview served as a crucial component of the study, allowing us to gather valuable feedback and insights from participants at the conclusion of their involvement. This interview was conducted during the final in-person visit, providing an opportunity for participants to reflect on their experiences and attitudes toward the study and the devices used. Through open-ended questions A and prompts, we encouraged participants to share their thoughts, opinions, and suggestions regarding various aspects of the study, including the study procedures, the usability of the devices, and any challenges they encountered. Participants were also invited to provide feedback on their overall satisfaction with the study and its impact on their oral hygiene behaviors. This exit interview not only provided valuable qualitative data but also offered participants a platform to voice their experiences, contributing to a deeper understanding of the study's outcomes and potential areas for improvement. The insights gained from the exit interviews helped refine future iterations of the study and enhance the overall participant experience.

2.4 Comprehensive Description and Analysis of ROBAS Dataset

This section focuses on the comprehensive description of the collected data and observations conducted in the ROBAS dataset phase 1 study, aimed at examining oral health behaviors. In this dissertation, we address three distinct problems: detection, characterization, and prediction. Each problem utilizes the ROBAS dataset. However, we encountered various challenges when attempting to use all of the participants' data during each problem-solving stage.

2.4.1 Data Set Description: Demographics, Sensor, and Ground-truth data

A diverse dataset was collected from a total of 30 participants, encompassing variations in age, gender, and employment. The participant group consisted of 15 males and 15 females, with an average age of 28.5±10.6 years. Notably, among the participants, 2 individuals were left-handed, and they represented a range of backgrounds, including undergraduate students, graduate students, administrative staff, software engineers, business people, and homemakers.

Over the course of the study, a substantial amount of sensor data was collected, amounting to 3,117 hours, equivalent to approximately 180,000,000 sensor data points, spanning a duration of 215 days. In order to establish a reference standard, 412 videos were gathered and meticulously coded, with an average duration of 4 minutes per video, to annotate Oral Health Behaviors (OHBs).

Throughout the study period, a total of 188 brushing events involving a normal brush and 177 brushing events utilizing a SmartBrush were observed. Additionally, 155 instances of flossing were recorded. Among these flossing events, one-third were associated with normal brushing, while the remaining two-thirds were connected to the use of a SmartBrush.

In terms of the flossing technique employed, the majority (85%) of the instances involved the use of string, while the remaining portion utilized picks.

2.4.2 Data Set: Detection Problem

During the detection problem focused on toothbrushing and flossing, we had a total of 25 participants' data available. Since our objective was to detect manual toothbrushing, we excluded the data from smartbrush usage. For the detection task, we considered sensor data as usable only when wrist-worn accelerometer and gyroscope data were present. Specifically, we utilized 192 days of sensor data from 21 participants to detect manual toothbrushing using the wrist-worn inertial sensor. In the case of flossing detection, we exclusively analyzed data from participants who used string for flossing, resulting in 95 string flossing episodes from 16 participants recorded over 125 days.

2.4.3 Data Set: Characterization Problem

In the characterization problem, we expanded our dataset by collecting data from an additional five participants, bringing the total to 30 participants over a period of 215 days. Due to challenges faced during ground truth extraction from videos, we had to discard some data. We focused on usable sensor data and manual toothbrushing data to characterize the surfaces of teeth being brushed with a manual toothbrush. After data curation, we utilized 114 manual brushing episodes from 19 participants for characterization purposes.

2.4.4 Data Set: Prediction Problem

The prediction problem posed a significant challenge due to the limited amount of data available. Our prediction task involved to detect about-to-moment, and when we initially began data collection in 2016, our focus was solely on brushing and flossing detection. However, as our research progressed, we introduced a second problem, which necessitated the expansion of our dataset by including data from an additional five participants, resulting in a total of 30 participants. Throughout the data collection phase, participants were instructed to record their episodes of brushing, flossing, and rinsing. At that time, we had no foresight regarding the subsequent problems we would tackle. Given that data collection concluded in 2021, we were unable to augment the participant pool by introducing new instructions. Thankfully, we were able to secure data from 12 participants, encompassing a total of 70 brushing episodes, to address the prediction problem. Importantly, for this particular problem, we considered both manual toothbrush and smartbrush data.

In conclusion, the data collection chapter of this study focused on gathering comprehensive and valuable information about participants' oral hygiene behaviors. Various methods were employed, including the recruitment of volunteers through flyers and word-of-mouth, conducting phone screenings and in-person interviews, and administering demographic questionnaires. Both qualitative and quantitative approaches were utilized to gain a deeper understanding of participants' experiences and perspectives. Participants were actively engaged in wearing wristbands and recording videos of their oral health behaviors, allowing for the collection of accelerometer and gyroscope data. The combination of these methods provided a multifaceted analysis of toothbrushing and flossing habits, facilitating a comprehensive evaluation of participants' routines. The dedication and commitment of the participants played a vital role in generating accurate and valuable data. Overall, the data collection phase was essential in achieving the objectives of the study and lays the foundation for further analysis and insights into oral hygiene behaviors.

Chapter 3

Virtual Orientation: Wrist Mounting Correction

In the free-living natural environment, and due to diversity in devices, mounting and wearing, configurations are unknown and can even dynamically change each time the device is taken off and put back on. Hence, we need to determine the orientation of the three axes with respect to the wrist using only sensor data, i.e., accelerometer and gyroscope traces.

In the realm of wearable technology, understanding the orientation of devices in relation to the wearer's body is crucial for accurate and meaningful data interpretation and inference. In this chapter, we delve into the concept of virtual orientation, which refers to the determination of device orientation on the user's wrist solely through sensor data.

In real-world scenarios, wearable devices are subject to a myriad of factors that can affect their configuration and positioning. Factors such as the diversity of devices available, different mounting methods, and the dynamic nature of wearing and removing devices all contribute to uncertainties in their orientation. Traditional approaches relying on fixed and calibrated reference points are not viable in free-living environments where devices can be worn interchangeably or adjusted throughout the day.

To overcome these challenges, we focus on leveraging the data captured by the device's built-in sensors, namely the accelerometer and gyroscope traces. These sensors provide raw measurements of acceleration and rotational motion, which can be processed and analyzed to determine the relative orientation of the device's axes with respect to the wearer's body. By carefully analyzing the sensor data and applying sophisticated algorithms, we can infer the orientation of the device in real-time or during post-processing.

The accurate determination of virtual orientation opens up a realm of possibilities for wearable technology applications. It enables precise tracking of physical activities, allowing for more accurate quantification of steps, calories burned, and exercise intensity. Furthermore, it facilitates gesture recognition, enabling users to interact with devices through natural hand and arm movements. In augmented reality applications, virtual orientation is essential for aligning virtual objects with the wearer's surroundings, creating immersive and interactive experiences.

Throughout this chapter, we will explore the intricacies of virtual orientation and the techniques employed to accurately determine device orientation using sensor data. We will delve into signal processing methodologies, sensor fusion techniques, and machine learning approaches that enable robust and reliable orientation estimation. By understanding and harnessing the power of virtual orientation, we can unlock the full potential of wearable devices, revolutionizing the way we interact with technology and enhancing our daily lives.

3.1 Notations and Definitions

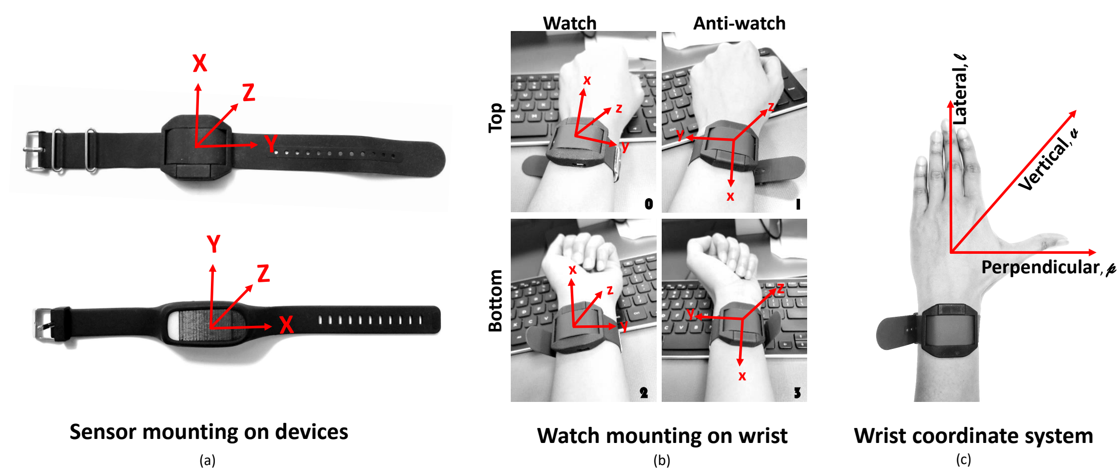


Fig. 3.1: (a) Lateral (l), perpendicular (p), and vertical (v) axes of wrist coordinate system; (b) Variation in sensor mounting on the wrist-worn devices (c) Four sensor positions on the wrist, referred to as Configuration c (for $c \in \{0, 1, 2, 3\}$)

Sensor data: Let $a_s(t) = (a_x, a_y, a_z)$ denote the accelerometer data and $\omega_s(t) = (\omega_x, \omega_y, \omega_z)$ denote the gyroscope data at time t.

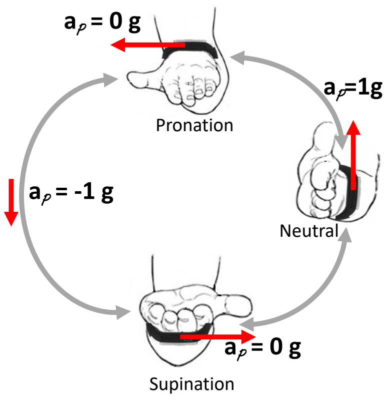
Wrist coordinate system: In the three-dimensional real space, all inertial sensors are inherently three-dimensional, with each accelerometer and gyroscope sensor data point containing measurements along the three axes. However, the triaxial inertial sensors' placement can vary among wrist-worn devices. For instance, the y-axis in one sensor may correspond to the x-axis in another sensor (see Figure 4.3a). Similarly, even for the same device, it can be worn in multiple orientations (see Figure 4.3b). Consequently, models designed to infer wrist orientation and hand gestures must be independent of specific sensor placements.

To address this challenge, we introduce a novel coordinate system known as *the wrist coordinate system*. This system defines three axes: lateral (l), perpendicular (p), and vertical (v) axes. The lateral axis aligns with the arm, the perpendicular axis aligns with the thumb, and the vertical axis corresponds to the direction of gravity when the palm is parallel to the Earth's surface (see Figure 4.3c). We represent the corresponding accelerometer data in the wrist coordinate system as $a(t) = (a_l, a_p, a_v)$ and the gyroscope data as $\omega(t) = (\omega_l, \omega_p, \omega_v)$.

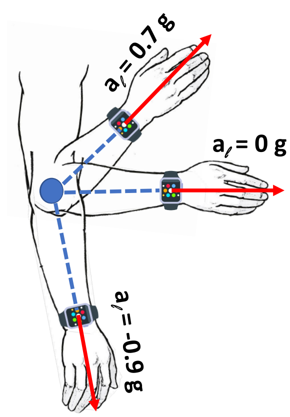
Sensor orientation configurations: The configuration of a wrist-worn sensor refers to its current orientation relative to the wrist. It describes the mapping between the sensor's axes coordinates and the wrist coordinate system. Typically, the z-axis is perpendicular to the sensor's surface. Thus, we assume the vertical axis always aligns with the z-axis. For example, in Figure 4.3b, the mapping for the top left placement is $(l, p, v) = (x, y, z)$. This particular configuration is considered the *base* configuration, as it occurs most frequently.

3.2 Virtual Orientation

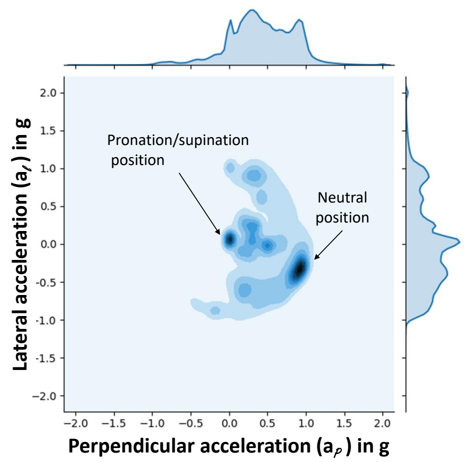
The inertial sensor itself can be mounted within the device in a total of $3! \times 2^3 = 6 \times 8 = 48$ possible configurations. This is because each of the sensor axes can be matched to one of the three real-world dimensions (in $3! = 6$ ways), and each axis can be pointing upwards or downwards (in $2 \times 2 \times 2 = 8$ ways). However, for any given wrist-worn device, the mounting of the inertial sensors is predetermined during the design phase. Consequently, the device can be worn in four different configurations (see Figure 4.3b). We can utilize a configuration file to specify this mounting, thereby reducing the number of dynamic changes in orientation to four. Therefore, the problem of virtual orientation revolves around determining which of these four configurations best represents the data being collected. It



((a)) Prevalent wrist orientations



((b)) Prevalent wrist positions



((c)) Kernel density function to show prevalent wrist positions and orientations

Fig. 3.2: (a) *Prevalent wrist orientations*; the value of a_p varies with the movements of rotating the forearm (b) *Prevalent wrist positions*; the value of a_l varies for different positions of wrist relative to elbow (c) *Kernel density function* to show prevalent wrist positions and orientations. The marginal distributions, at the top for perpendicular axis and at the right for the lateral axis, show the distribution of different measurements of the accelerometer sensor over a day (using $g = 9.8ms^{-2}$).

is important to note that while it is possible for the sensor orientation to change during the day (e.g., due to slippage), we assume that such changes in orientation are temporary, and participants reorient the device shortly after. Additionally, we note that as gyroscopes and accelerometers are on the same chipset, usually the orientation of accelerometers also determines the orientation of the gyroscopes.

In order to develop and evaluate a solution for the virtual orientation problem, which

Algorithm 1 Pseudo code for finding the correct configuration of a wristband

Require: $\vec{a} = \{a_s(t) = (a_x, a_y, a_z)\}, D$
Ensure: $c \in \{0, 1, 2, 3\}$
for each $c \in \{0, 1, 2, 3\}$ **do**
  $\vec{a_c} = \vec{a} \cdot M_c$
  $S_c = d(\vec{a_c}, D)$
end for
$c = \arg\min_{c}\{S_c\}$
return c

involves determining the current configuration of the sensor among the four possible options, labeled training data is essential. To construct such a dataset, we leverage the video captures taken twice daily during the performance of Oral Hygiene Behaviors (OHBs). By carefully observing the position of the sensor on the wrist in these videos, we label a day's worth of data as belonging to a specific configuration. If we consistently observe the same configuration in both videos, we assign that day's data to that particular configuration.

3.3 Dataset description

In our dataset, a significant portion of the data (77.08%, or 148 out of 192 days) corresponds to Configuration 0, which we designate as the base configuration (refer to Figure 4.3). This configuration represents the most frequently occurring orientation. We compile a dedicated database, denoted as D, comprising all the data associated with Configuration 0, encompassing a total of 148 person-days' worth of data.

3.4 Detection of Correct Orientation

We denote $\vec{a}$ as a time series of data points within a single window of unknown configuration. To explore various scenarios, we consider multiple window sizes, such as half an hour, 1 hour, 1.5 hours, 2 hours, 2.5 hours, and so on. By examining data within different window sizes, we can gain insights into the behavior and characteristics of the sensor's orientation over varying time intervals.

Our approach aims to determine the configuration among the four possibilities that produce data most similar to the base configuration when transformed through matrix multiplication. We denote the correct wearing configuration as $c \in 0, 1, 2, 3$. We can translate the data from the given configuration to the wrist coordinate system by applying matrix multiplication. For each time point t, the transformed data is given by $a_c(t) = a(t) \cdot M_c$, where M_c represents the corresponding transformation matrix for each configuration.

The transformation matrices M_c for the four configurations are as follows:

$$M_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}; M_1 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}; M_2 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}; M_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

For example, for any accelerometer sample, (a_x, a_y, a_z), of Configuration 0, the corresponding value in the wrist coordinate system is $(a_l, a_p, a_v) = (a_x, a_y, a_z) * M_0 = (a_x, a_y, a_z)$.

For a given window of accelerometer sensor data $\vec{a}$, if $c \in 0, 1, 2, 3$ represents the correct configuration, the transformed data $\vec{a_c} = \vec{a} * M_c$ corresponds to the data in the wrist coordinate system. Our assumption is that the similarity between $\vec{a_c}$ and the database D is greater than the similarity between any other transformed data $\vec{a}_{c'}$ (where $c' \neq c$). Based on this assumption, we create four transformations of $\vec{a}$, namely $\vec{a_0}$, $\vec{a_1}$, $\vec{a_2}$, and $\vec{a_3}$. Next, we compute the similarity index between each transformed data $\vec{a_c}$ and the database D. However, computing the similarity between $\vec{a_c}$ and D efficiently is crucial due to the potentially large number of data points in $\vec{a_c}$ and D.

In order to investigate the potential for significant data reduction, we conducted an analysis of our dataset's distribution. Figure 3.2(c) displays the joint density, illustrating the different positions and orientations of the wrist along the lateral and perpendicular axes. The marginal distributions, situated at the top for the perpendicular axis and at the right for the lateral axis, demonstrate the distribution of sensor measurements throughout a day. Upon examination, we identified two distinct clusters in the joint density plot, representing the three most commonly observed wrist orientations: pronation, supination, and neutral. This observation guided us in developing two concise representations of the data—probability distribution and principal components. Consequently, we considered two corresponding distance indices, denoted as $d(\vec{a}, \vec{D})$, to further analyze the dataset.

1. Distribution distance ($d_{dist}(\vec{a}, D)$): By calculating the probability distributions of the data points in $\vec{a}$ and D, denoted as P and Q respectively, we can determine the dissimilarity between these distributions using the earth moving distance. This distance metric, denoted as $d_{dist}(\vec{a}, D)$, provides insight into the dissimilarity between the two distributions.

2. Distance of principal components ($d_{PCA}(\vec{a}, D)$): To compute the major directions of the data, we perform principal component analysis (PCA) on $\vec{a}$ and D. This analysis yields the vectors $pc\vec{a}_a$ and $pc\vec{a}_D$, representing the three major components of all data points in $\vec{a}$ and D, respectively. The Euclidean distance is then used to quantify the dissimilarity between $pc\vec{a}_a$ and $pc\vec{a}_D$, resulting in the value $d_{PCA}(\vec{a}, D)$.

We compute these distances for all values of $\vec{a}$ and $\vec{D}$ across each configuration c. The configuration c that yields the minimum distance is assigned to the window of data $\vec{a}$. The complete process is described in Algorithm 1.

In summary, our algorithm aims to determine the correct configuration of a wristband by comparing the transformed data with a database, using distribution distance and distance of principal components as similarity measures. This approach leverages matrix multiplication to translate the data to the wrist coordinate system and efficiently computes distances between the transformed data and the database.

3.5 Performance of Virtual Orientation

To identify the virtual orientation, we conducted experiments to determine the optimal window size and method that achieves the best performance. We explored three methods: using a default configuration, applying Principal Component Analysis (PCA), and utilizing distribution similarity.

To create the labeled dataset for the experiments, we divided the entire dataset into

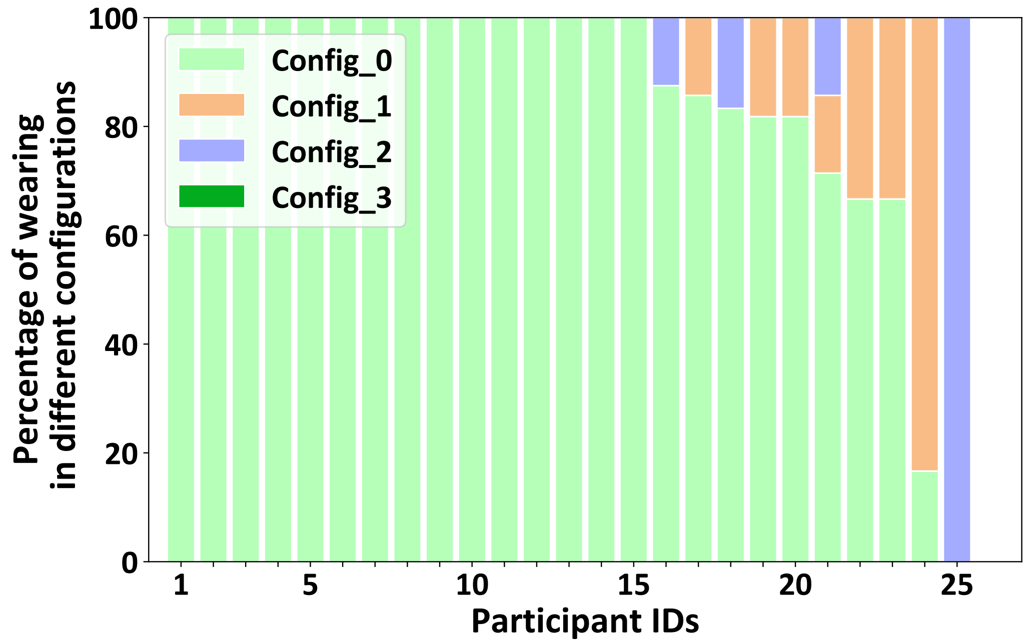


Fig. 3.3: Use of different configurations by participants

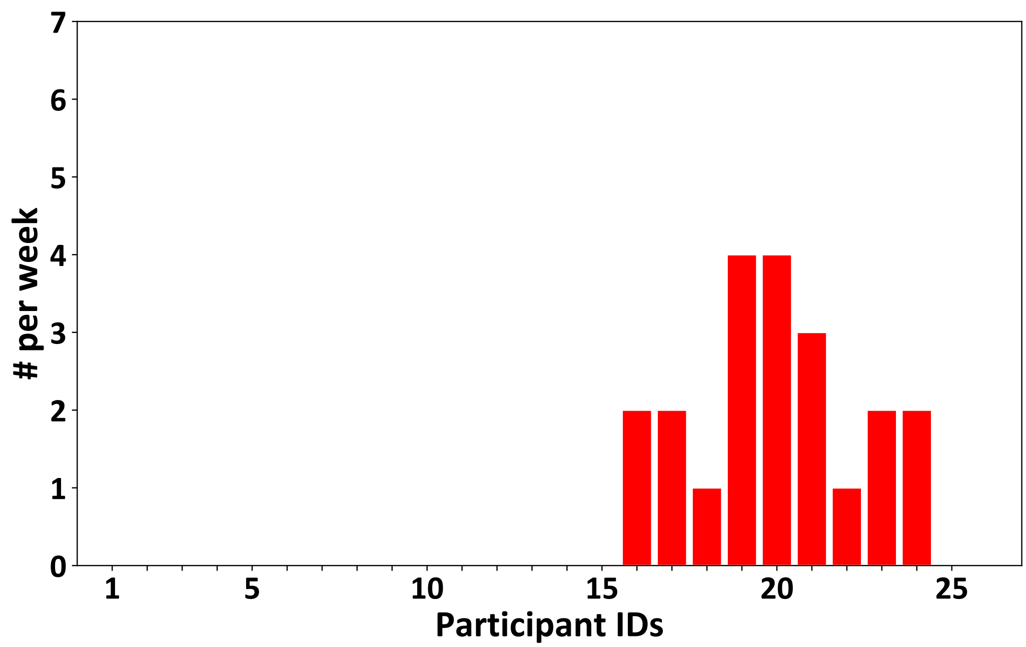


Fig. 3.4: Frequency of configuration changes by participants

specific window sizes. Each window of data was labeled according to the configuration observed from two video recordings for that day. This ensured that the labeled data accurately represented the configuration during the corresponding time intervals.

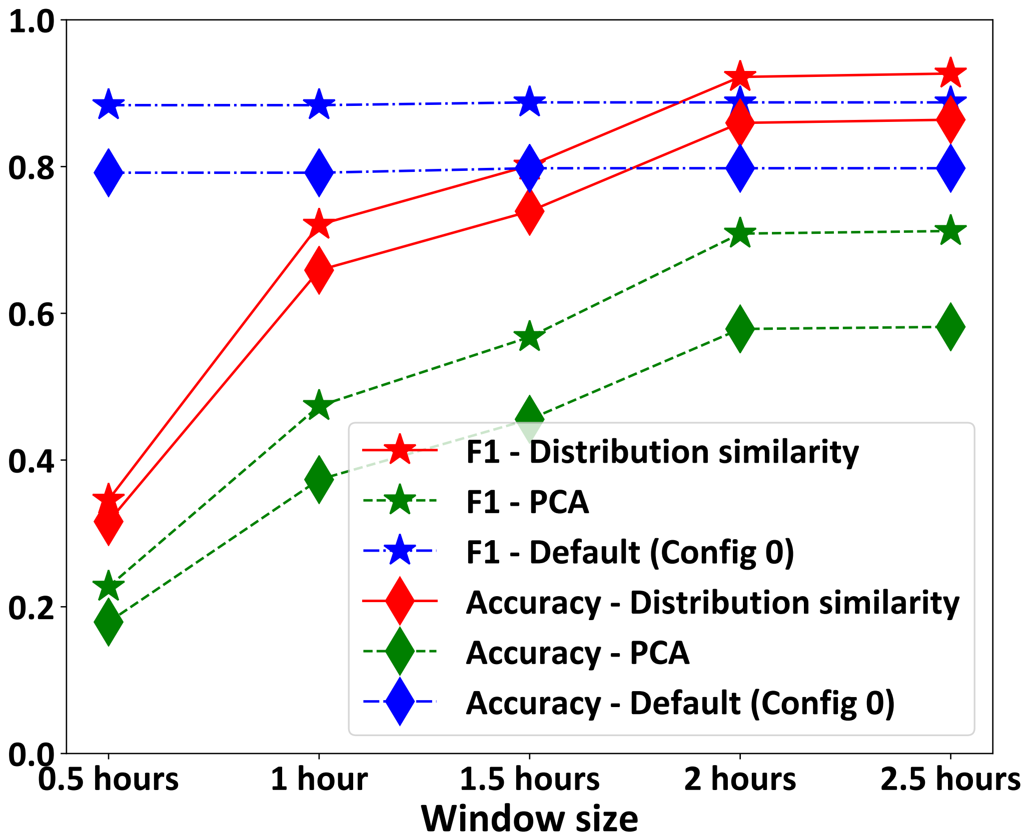


Fig. 3.5: F1-score and accuracy for virtual orientation

We used a dataset comprising 192 days of sensor data collected from 25 participants. Figure 3.3 illustrates the distribution of wristband configurations, indicating the percentage of time participants wore the sensor in different configurations. We observed that 15 out of 25 participants consistently wore the sensor in Configuration 0, while nine participants switched configurations across days, and one participant always used Configuration 2. Based on this pattern, we established Configuration 0 as the default configuration.

We evaluated the performance of different window sizes, including half an hour, 1 hour, 1.5 hours, 2 hours, and 2.5 hours. From the results shown in Figure 3.5, we observed that the performance reached saturation after a window size of 2 hours. Therefore, we selected a window size of 2 hours for further analysis. Notably, the choice of window size did not significantly affect the performance of the default configuration.

Among the three methods, the results of our experiments clearly indicate that the distribution-based method outperforms both the default configuration and PCA-based methods in accurately identifying virtual orientation configurations. The distribution-based method achieved an impressive accuracy of 86% and an F1-score of 93%, indicating its ability to accurately classify the sensor data into the correct configurations.

On the other hand, the default configuration method, which relies on a fixed configuration, achieved an accuracy of 79% and an F1-score of 88%. While this method still provides reasonable accuracy, it falls short compared to the distribution-based approach. This is expected, as the default configuration method assumes that the majority of participants wear the sensor in Configuration 0, which may not be true for all individuals. Therefore, personalizing the virtual orientation based on individual data improves the accuracy significantly.

The PCA-based method, which utilizes principal component analysis, exhibited the lowest performance among the three methods, with an accuracy of 58% and an F1-score of 71%. This suggests that the PCA approach alone is not sufficient to capture the variations and nuances in the sensor data that are crucial for accurate virtual orientation identification. The distribution-based method, on the other hand, considers the probability distributions of the data points and utilizes the earth moving distance to measure dissimilarity. This approach proves to be more effective in capturing the characteristics of different configurations and accurately discriminating between them.

In real-life deployment scenarios, one can initially rely on the default configuration and switch to the distribution-based method after collecting two hours of data. This personalized approach can further improve the accuracy of virtual orientation identification for individual users.

Overall, our experiments demonstrate the effectiveness of the distribution-based method and emphasize the importance of selecting an appropriate window size for virtual orientation analysis.

3.6 Chapter Summary

Throughout this chapter, we tackled the issue of virtual orientation in wrist-worn devices and proposed a solution that relies on analyzing sensor data. Our framework was developed to accurately identify the device configuration by utilizing a labeled dataset created from video recordings.

We compared three different methods: default configuration, PCA-based, and distribution-based. Out of the three, the distribution-based method was the most successful, achieving an accuracy rate of 86% and an F1-score of 93%. This method makes use of probability distributions and captures the inherent patterns in the data.

The distribution-based approach creates a personalized virtual orientation, which enhances the user's experience and application performance. Furthermore, it has practical implications for the real-life deployment of wrist-worn devices as it provides accurate configuration identification.

Our research emphasizes the importance of taking sensor measurement distributions into account for virtual orientation. The distribution-based method captures configuration characteristics and enables precise discrimination.

In conclusion, our solution effectively identifies virtual orientation in wrist-worn devices. The distribution-based method provides superior accuracy and personalized orientation, advancing virtual orientation techniques and improving user experiences in various applications.

Chapter 4

Detection of Brief Daily Behaviors:

A Case Study on Detecting Toothbrushing and Flossing

4.1 Introduction

The act of brushing and flossing one's teeth is a quick daily routine. Brushing typically occurs once or twice a day and can last for a mere two minutes. For optimal oral health, it is recommended that brushing lasts for at least two minutes, covers all tooth surfaces adequately with appropriate pressure, and is complemented by flossing. Interestingly, dental diseases can largely be prevented and are closely tied to the insufficient practice of simple oral hygiene habits such as regular toothbrushing and flossing. The American Dental Association (ADA) advises everyone to brush their teeth at least twice a day and floss at least once, especially after meals. However, studies have revealed that a significant portion of the population fails to follow these recommendations. Around 33% of men brush only once a day, and 59% of women frequently skip brushing before bedtime [3]. This discrepancy between health guidelines and actual behavior has led to the argument for the development and widespread use of mHealth approaches to help individuals improve their oral health practices, ultimately reducing the burden of dental diseases on public health [32].

The expenditure on healthcare is projected to consume nearly 20% of the United States economy by 2026 [1]. This unsustainable spending has prompted a shift from reactive and costly healthcare that focuses on treating illnesses to proactive and preventive approaches aimed at addressing the underlying health behaviors that contribute to disease risk. Our focus in this study is dental diseases such as cavities and periodontal disease, which are common chronic conditions with significant consequences [1]. In the United States, half of adults suffer from periodontal disease, and over fifty-three million people have untreated tooth decay in their permanent teeth. A significant subset of this population experiences the debilitating effects of advanced periodontal disease [2]. Aside from the pain and suffering

[1]https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/NationalHealthAccountsProjected.html

caused, oral health problems also impact essential functions like eating, speaking, and socializing, and they can lead to local and systemic infections. These health implications not only diminish an individual's overall health and well-being but also result in substantial personal and societal costs. Poor oral hygiene has also been linked to conditions such as heart disease, stroke, diabetes, pneumonia, respiratory diseases, preterm births, and low birth-weight babies [33].

To achieve optimal oral health, it is crucial to brush for a minimum of two minutes, ensuring that all tooth surfaces are adequately cleaned with appropriate pressure, and flossing should complement brushing. The industry has introduced smart toothbrushes with various features to assist users in tracking their brushing habits. These toothbrushes can monitor brushing duration and provide feedback to users about the pressure applied during brushing [18]. Researchers have explored adding newer sensors for improving the detection of brushing surfaces [24], instrumenting miniature cameras into the toothbrush head for detecting plaque [19], and developing implantable assistive brushing devices [34, 20] to assist children and individuals with disabilities. Other efforts involve the concurrent use of video cameras and mobile applications to capture brushing movements [35, 36, 37, 38]. However, these systems can be burdensome and impractical as they require users to activate data collection and restrict their movements relative to the camera during brushing.

While these advancements benefit users of smart toothbrushes, the majority of the population (over 80% in the United States) still use manual toothbrushes and cannot take advantage of these technological innovations [2]). The ability to infer timing, duration, and quality of brushing behaviors using manual toothbrushes would extend these benefits to the general population. This development could pave the way for personalized oral health management, individual and population-level disease risk assessment, hybrid health insurance programs [32], personalized feedback, and engagement through rewards or gamification [37, 38].

[2]https://www.statista.com/statistics/278116/us-households-usage-of-manual-toothbrushes/

In addition to helping users maintain oral hygiene, the detection of brushing and flossing events can have multiple other utilities. Dentists can use them to provide encouragement and feedback. Insurance companies can use it to offer discounts and reinforce oral hygiene behaviors. ADA recommends replacing the manual toothbrush every three to four months [3]. Detection of brushing can be used to remind when to change toothbrushes and to reorder toothpaste automatically. The dental insurance company can provide some discount based on the habit of oral hygiene. Also, the dentist can create the profile for his patients based on their oral hygiene practice. Most importantly, all these approaches focus primarily on the use of instrumented toothbrushes to monitor brushing behaviors and require attentional effort by the user. Clearly, we need alternate systems that leverage data streams captured passively by alternative technologies (e.g., wrist wearables) to create digital phenotypes of brushing behaviors and use them as the basis for improving brushing behaviors, linking naturalistic oral hygiene behaviors to disease outcomes, stratifying dental disease risk, and titrating treatment resources to actual need instead providing the same set of preventive services to all.

Furthermore, automated inference of brushing and flossing behaviors can have a wide range of applications in research studies. It can help identify predictors of dental disease outcomes, inform treatments, and shape public health policies. Dentists can provide feedback and encouragement, while insurance companies can offer discounts and reinforce good oral hygiene habits. The American Dental Association recommends changing manual toothbrushes every three to four months [4], and detection of brushing behaviors can remind individuals to replace their toothbrushes and order toothpaste automatically. This data can also be used by dental insurance companies to determine discounts and by dentists to create personalized patient profiles.

Achieving this vision necessitates the development of robust computational models capable of accurately detecting toothbrushing and flossing behaviors in natural field environ-

[3] https://www.mouthhealthy.org/en/az-topics/b/brushing-your-teeth
[4] https://www.mouthhealthy.org/en/az-topics/b/brushing-your-teeth

ments. For this purpose, we utilize inertial sensors embedded in wrist-worn devices, which are already used for detecting sleep, activity, eating, smoking, and various other daily behaviors. A crucial initial step in this direction is the creation of reliable computational models that can extract distinct features of oral hygiene behaviors from signals generated by inertial sensors in wrist wearables. These features, including speed, wrist rotation, and arm displacement, can then be used to detect oral hygiene behaviors such as brushing and flossing in real-world settings.

4.1.1 Challenges

Inferring oral hygiene behaviors (OHBs) primarily from inertial sensors on the wrist presents several technical challenges.

Variability in sensor mounting: The placement of the sensor in a wristband or the position of the wristband on the wrist can vary between devices and even within the same device for different wearing episodes. This variation, such as palm-facing or back-palm-facing configurations (as shown in Figure 4.3), results in significantly different signals. The model must determine the relative configuration of the inertial sensor to the wrist, considering that the signal from the axis parallel to the hand is crucial for determining whether the hand is facing upward or not.

Reliable detection of rare daily behaviors: Brushing and flossing are salient but relatively brief events that occupy only a small fraction of the approximately 960 awake minutes per day. Events that take less than 1% of the assessment time require stringent recall requirements and even more stringent control over false positives. A false positive rate of just 1% can lead to four false positive events per day.

Access to fine-grained accurate labels for model training/evaluation: Obtaining precise labels for each brushing event is necessary to train and validate OHB models. Detailed markings, at a second-level granularity, are needed to identify the exact start and end times of each event, as well as any intervening pauses. However, collecting such carefully labeled data from the natural environment where individuals freely engage in oral hygiene behaviors can be challenging.

Precise estimation of duration and start/end times: To be clinically useful, the detected oral hygiene event must closely match the actual event in terms of duration and start/end times. Achieving accurate estimation of these parameters is essential for practical applications.

4.2 Related Works

Wrist-worn inertial sensors have been utilized to detect various Activities for Daily Living (ADL), such as walking, sleeping, eating, combing hair, dressing, climbing stairs, sitting, standing, and cooking [7, 8, 9, 10, 11, 12]. Some studies [7, 13] have shown the feasibility of detecting brushing through hand gestures while focusing on a wide range of ADL activities. However, these studies mostly collected data in controlled environments and aimed to demonstrate feasibility rather than practical application, resulting in a false positive rate of over 15% for toothbrushing. Such models are not suitable for passive detection in real-life settings, as they would generate a high number of false positive events per day (72 according to Table 4.1).

To develop a more reliable model for passive detection using continuous data collected in natural settings, researchers have focused on specific target behaviors. For example, some studies have focused on eating [15], while others have focused on smoking [16, 17]. Behaviors like toothbrushing, smoking, and eating are brief events that last only a few minutes, but continuous data collection throughout an entire day yields 16 hours of data during the awake period. Thus, stringent accuracy requirements for both recall and false positives are necessary. For instance, a model with a 5% false positive rate for detecting brushing would produce 24 false positive events per day (as shown in Table 4.1).

The main challenge in using existing models to detect brushing or flossing lies in the need for behavior-specific models. For example, the model for detecting physical activity like walking [12, 14] is not directly applicable to detecting eating. Similarly, the models

Table 4.1: The table shows the number of false positive events produced per day by a model for specific false positive rates. It is assumed that sensors are worn for 16 hours per day, and the toothbrushing event lasts an average of 2 minutes.

False positive rate	**15%**	**10%**	**5%**	**1%**	**0.1%**	**0.01%**
False positives per day	72	48	24	4	$\frac{1}{2}$	$\frac{1}{20}$

for eating [15] or smoking [16] are not directly applicable to detecting brushing or flossing. Consequently, reliable passive detection of brushing using wrist-worn sensors remains an unresolved problem.

Previous efforts to detect oral hygiene behaviors have predominantly focused on instrumented or smart toothbrushes. Some researchers have utilized commercial smart toothbrushes [39, 40, 37], while others have instrumented manual toothbrushes by adding inertial sensors [24] or attaching mini smartphones [41]. These works often aim to recognize the brushed tooth surfaces [37, 39, 24, 25]. Acoustic-based approaches have also been explored, where audio data collected by a smartphone app or microphone is used to evaluate toothbrushing performance [42, 43, 44]. Other studies have analyzed audio signals of brushing strokes using machine learning techniques [26], or combined a microphone placed on the neck with an earphone placed in the ear to monitor toothbrushing [27].

Although investigations using smart or instrumented toothbrushes do not provide a reliable solution for automatically detecting toothbrushing or flossing events from wrist-worn inertial sensors, they offer valuable insights that can be applied once a model for detecting these events and identifying their start/end times is developed. The sensor data corresponding to brushing and flossing events can be used for further analysis, such as detecting brushing surfaces, monitoring pressure, and developing engagement apps to improve oral hygiene.

4.3 Data Set for Detection of Toothbrushing and Flossing

Participants in the study were instructed to record videos of themselves engaging in their typical oral health routine, capturing moments of toothbrushing and flossing. These videos served as the ground truth, providing reliable labels for identifying and analyzing toothbrushing and flossing events. To ensure data security, the phone used by participants was designed to store this sensitive information on an encrypted microSD card. Additionally, the sensor data (excluding video and GPS data) was periodically uploaded to a secure server through a dedicated HTTPS connection. This data transfer process allowed researchers to collect and analyze the sensor data remotely while maintaining privacy and data integrity. Finally, at the conclusion of the study, the recorded videos and GPS data were retrieved from the participants' phones for further analysis and evaluation

4.3.1 Annotation of Oral Health Behaviors from Video Data

To generate labeled data for the development and evaluation of the model, we conducted video annotations to capture the timing of each oral health behavior (OHB). A total of 362 videos were collected, with an average duration of 3.12 minutes. During the annotation process, we carefully marked the start and end times of brushing and flossing events, as well as any pauses within these events, based on the video footage.

One of the key challenges we encountered was correctly identifying the pause segments within each brushing and flossing event. This required precise marking to ensure that only data corresponding to active brushing or flossing were considered for model training and testing. For brushing events, we paid special attention to the interval when the hand holding the brush moved away from the mouth and then returned, which we designated as a pause.

In addition to marking the start and end times of each brushing and flossing event, we recorded other important details related to the OHBs. These included the orientation configuration of both wrists, identifying the brushing wrist (left or right), specifying whether a manual brush or a smart brush was used, determining the flossing method (string or picks), noting the flossing wrist (left or right), and capturing any video pause times.

To ensure the accuracy and consistency of the labeled data, two coders independently labeled all the video ground truth data. In cases where there were discrepancies in their coding, both coders jointly reviewed the segments in question and reached a consensus on the labeling of the event. This collaborative approach helped enhance the reliability of the annotations.

Table 4.2 provides a comprehensive overview of our video annotation protocol, which outlines the meaning and description of each labeled event along with their corresponding characteristics. For instance, "bn_st" represents the start time of brushing with a manual brush, while "fs_ed" denotes the end time of flossing with string. Pause events are marked with "p_st" for the start time and "p_ed" for the end time, depending on the specific OHB. We also included the orientation of the left and right wrists, the brushing wrist (left or right), the flossing wrist (left or right), and the type of flossing (string or picks) in our annotations.

Through meticulous video annotation, we were able to create a valuable dataset with detailed information on each OHB event. This labeled dataset serves as a crucial resource for training and evaluating our model, enabling us to develop a robust and accurate system for detecting and analyzing oral health behaviors.

4.3.2 Participant Study and Data Collection

A comprehensive study was conducted with a total of 25 participants (12 males, 13 females; mean age 28.5 ± 7.6 years) who successfully completed the study. The data collection period spanned 290 days, during which a substantial amount of sensor data, totaling 3,886 hours, was collected. For the purpose of model development and testing, we focused on a subset of this data, specifically 192 days (equivalent to 2,797 hours of sensor data) for which video data was available.

Within this dataset, we observed a total of 160 brushing events with the manual brush and 164 brushing events with the SmartBrush. These brushing events were distributed across the study period and provided valuable insights into participants' oral health habits. Furthermore, a total of 137 flossing events were recorded, with one-third of these events being associated with normal brushing and two-thirds associated with the SmartBrush.

When considering the modality of flossing, it was found that 81% of the flossing events were performed using string, highlighting its popularity among the participants. The re-

Table 4.2: Video annotation protocol for identifying OHB events and labeling their start and end times

Label	Meaning	Description and event characteristics
bn_st	brushing start time with manual brush	The first video frame when the participant's hand is close to the mouth and brushing begins
bn_ed	brushing end time with manual brush	When the participant stops brushing and wrist holding the manual brush is going down from the mouth
bo_st	brushing start time with SmartBrush	Similar to the bn_st event, but for SmartBrush
bo_ed	brushing end time with SmartBrush	Similar to the bn_ed event, but for SmartBrush
fs_st	flossing start time with string	When the participant actively starts flossing with string (preparation time before flossing such as string winding around fingers) is excluded
fs_ed	flossing end time with string	When both of the participant's wrists move down from the mouth with string floss without subsequent resumption of flossing
fp_st	flossing start time with picks	When the participant flosses with a pick, and one wrist with pick moves up to the mouth and flossing begins
fp_ed	flossing end time with picks	When participant ends flossing with pick, and wrist moves down from the mouth
p_st	pause start	For brushing, pause begins when the participant temporarily stops brushing (e.g., to spit out the accumulated toothpaste or to rinse) and moves the wrist away from the mouth. For string flossing, pause begins when both wrists move away from the mouth. For pick flossing, pause begins when the wrist holding the pick moves away from the mouth
p_ed	pause end	Following p_st, pause ends for brushing when the brushing wrist goes up to the mouth. For string flossing, pause ends when both wrists go up to resume flossing. For pick flossing, pause ends when the wrist holding the pick moves back up to the mouth to resume flossing
pv	video pause time	A few participants paused the video when they paused brushing or flossing (e.g., to spit out the accumulated toothpaste or saliva from mouth), or sometimes between the brushing and flossing events. During these times, we missed all ground truth video. We annotate those events as video pause time.
ori_left	orientation of left wrist	We observe the device orientation on the left wrist in video and label the configuration as one of $\{1, 2, 3, 4\}$
ori_right	orientation of right wrist	We observe the device orientation on the right wrist in video and label the configuration as one of $\{1, 2, 3, 4\}$
bn_wt	brushing wrist	The brushing wrist is marked as left or right
fl_wt	flossing wrist	If flossing with pick, we mark the flossing wrist as left or right – otherwise mark it as both wrists
fl_tp	flossing type	Flossing type is marked as string or pick

maining flossing events were conducted using picks, representing an alternative approach to oral hygiene.

The comprehensive dataset obtained from this study, comprising extensive sensor data and video recordings, serves as a valuable resource for the development and evaluation of our model. By leveraging this dataset, we aim to enhance our understanding of oral health behaviors and improve the accuracy and effectiveness of oral health monitoring systems.

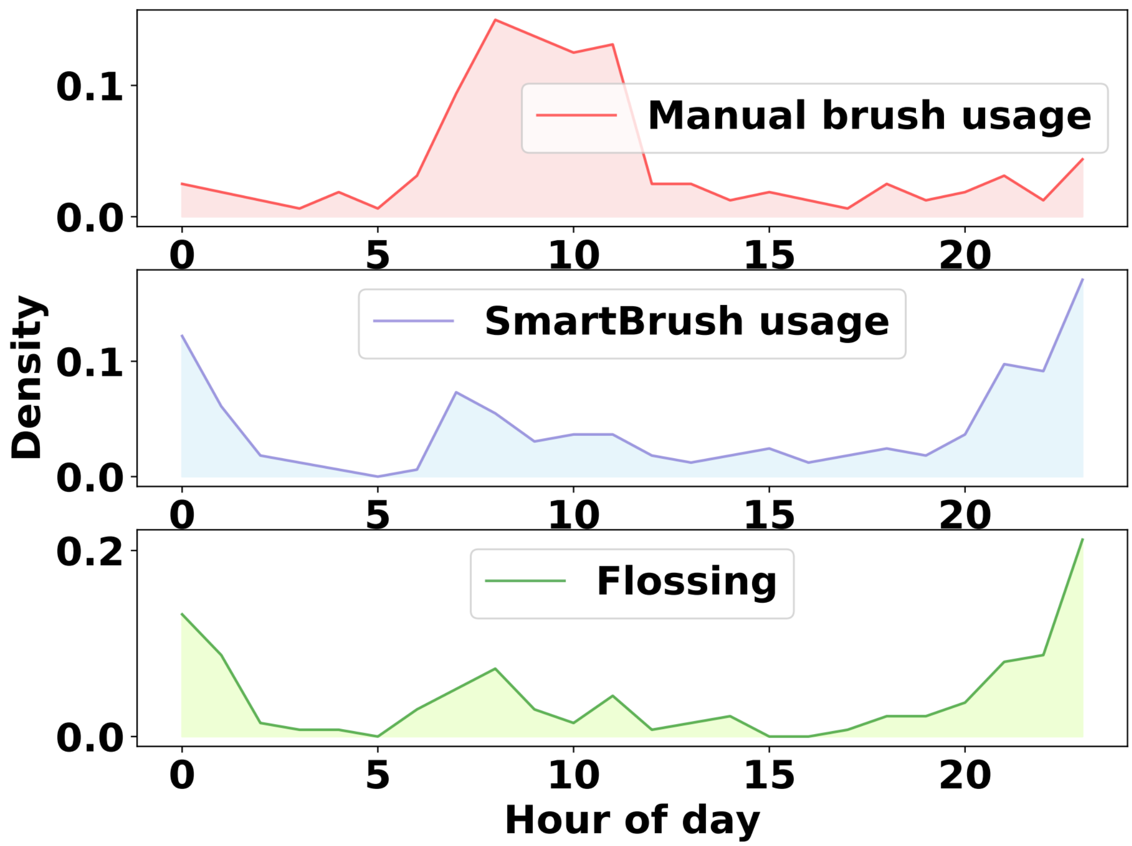


Fig. 4.1: Distribution of brushing and flossing Events throughout the day: morning preference for manual toothbrush and nighttime preference for SmartBrush

4.3.3 Insights from Video Annotations: Timing and Duration of Oral Health Behaviors

Through our comprehensive analysis of video annotations of Oral Health Behaviors (OHBs), we have uncovered several intriguing insights. These observations shed light on

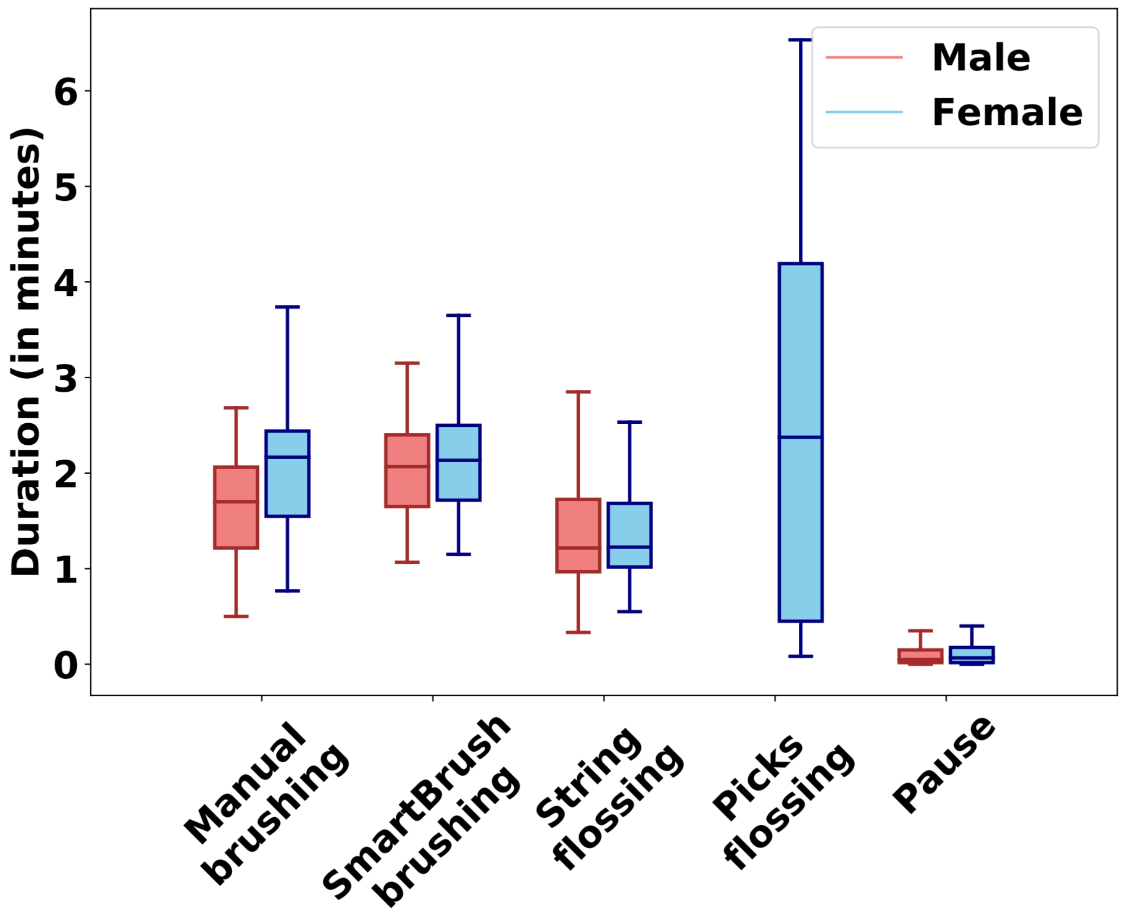


Fig. 4.2: Analysis of oral hygiene event durations obtained from video annotations.

the timing and duration of OHBs, providing valuable information for understanding oral health habits.

Firstly, we examined the temporal distribution of OHBs throughout the day (see Figure 4.1). Our findings revealed that OHBs occurred at various times, with a noticeable concentration in the morning and evening. These peak periods indicate higher engagement in oral health routines during these specific times. Specifically, we observed a higher occurrence of manual brushing in the morning, suggesting a proactive approach to oral care at the start of the day. Conversely, the use of the SmartBrush was more prevalent during nighttime hours, implying a preference for convenience and efficiency during the winding down of the day. Additionally, we noted an increased frequency of flossing during the evening, highlighting the importance individuals place on thorough oral hygiene practices before bedtime.

Secondly, we delved into the duration of different OHB events (see Figure 4.2), offering insights into the time individuals allocate to maintain oral health. Remarkably, despite participants being aware of video recording, the duration of brushing with a manual toothbrush was marginally lower compared to brushing with the SmartBrush. The SmartBrush, equipped with vibrating features to indicate the recommended 2-minute brushing duration as advised by the American Dental Association (ADA), seemed to promote compliance with brushing guidelines, as evidenced by its median duration of 2.1 minutes. Notably, we observed a slight gender difference, with females exhibiting a slightly longer brushing duration when utilizing a manual toothbrush, suggesting a potential inclination towards more meticulous oral care practices.

Finally, our analysis revealed that flossing events had a median duration of 1.22 minutes, reflecting the average time individuals dedicated to this essential aspect of oral hygiene. Interestingly, we noted that pause durations within OHB events were generally short. This observation suggests that individuals tended to minimize interruptions or breaks during their oral health routines, maintaining a continuous focus on effective brushing and flossing practices.

The comprehensive insights gained from our video annotation analysis provide valuable knowledge about the temporal patterns and duration of OHBs. These findings enhance our understanding of individuals' oral health habits and contribute to the development of personalized and effective oral care strategies.

4.4 Notations, Definitions, and Observations of Sensor Signals during Oral Hygiene Behaviors

Before we delve into the details of our modeling approaches aimed at reliably detecting brushing (with a manual toothbrush) and flossing using wrist-worn inertial sensors, it is crucial to establish a solid foundation by introducing a set of notations and definitions. In this section, we lay the groundwork by presenting key notations and definitions that will enable a better understanding of our subsequent discussions on the modeling methodologies for accurate detection of brushing and flossing behaviors. This section also presents the notations, definitions, and observations of sensor signals during oral hygiene behaviors. It establishes the terminology and symbols used throughout the study while providing insights into the sensor data collected during brushing and flossing activities. By combining these two perspectives, we gain a comprehensive understanding of the data and lay the foundation for further analysis and detection algorithms.

4.4.1 Establishing Terminology and Symbols

Sensor data: Let $a_s(t) = (a_x, a_y, a_z)$ be accelerometer data and $\omega_s(t) = (\omega_x, \omega_y, \omega_z)$ be gyroscope data at time t.

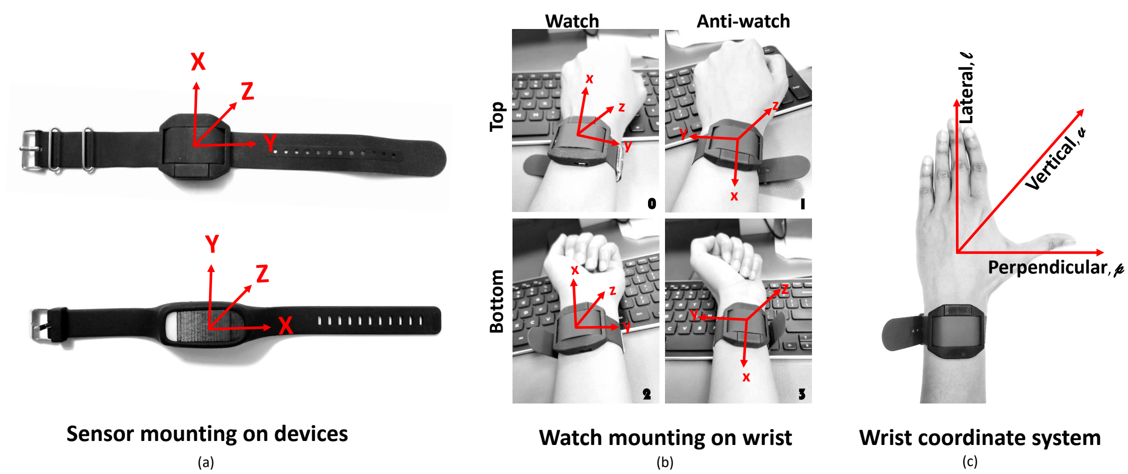


Fig. 4.3: (a) Variation in sensor mounting on the wrist-worn devices; (b) There are four positions on the wrist where the sensor can be placed, denoted as Configuration c (where $c \in 0, 1, 2, 3$); (c) The axes of the wrist coordinate system are referred to as the lateral axis (l), perpendicular axis (p), and vertical axis (v).

Wrist coordinate system: Different wrist-worn devices have different placements of the triaxial inertial sensors. For example, y-axis in one sensor may correspond with the x-axis in another sensor (see Figure 4.3b). Similarly, the same device can be worn in different ways (see Figure 4.3c). Therefore, models for inferring wrist orientation and hand gestures, need to be independent of the sensor placement.

To accomplish this, we define a new coordinate system, called *the wrist coordinate system*. The three axes in the wrist coordinate system are defined as lateral (l), perpendicular (p), and vertical (v) axes. Here, lateral axis is aligned with the arm, perpendicular axis is aligned with the thumb, and vertical axis is the gravity axis when the palm is parallel to the earth's surface (see Figure 4.3a). We denote $a(t) = (a_l, a_p, a_v)$ to be the corresponding accelerometer data in the wrist coordinate system and $\omega(t) = (\omega_l, \omega_p, \omega_v)$ to be the corresponding gyroscope data in the wrist coordinate system.

Sensor orientation configurations: The configuration of a wrist-worn sensor is defined as the current orientation of the sensor relative to the wrist. In other words, it specifies the mapping between sensor's axes coordinate and the wrist coordinate system. Usually, the z-axis is perpendicular to the surface of the sensor. Hence, we assume that the vertical axis is always aligned with the z-axis. For example, in Figure 4.3c, mapping for the top left placement is $(l, p, v) = (x, y, z)$. We regard this as the *base* configuration, since it occurs most frequently.

4.4.2 Observations of Sensor Signals during Oral Hygiene Behaviors

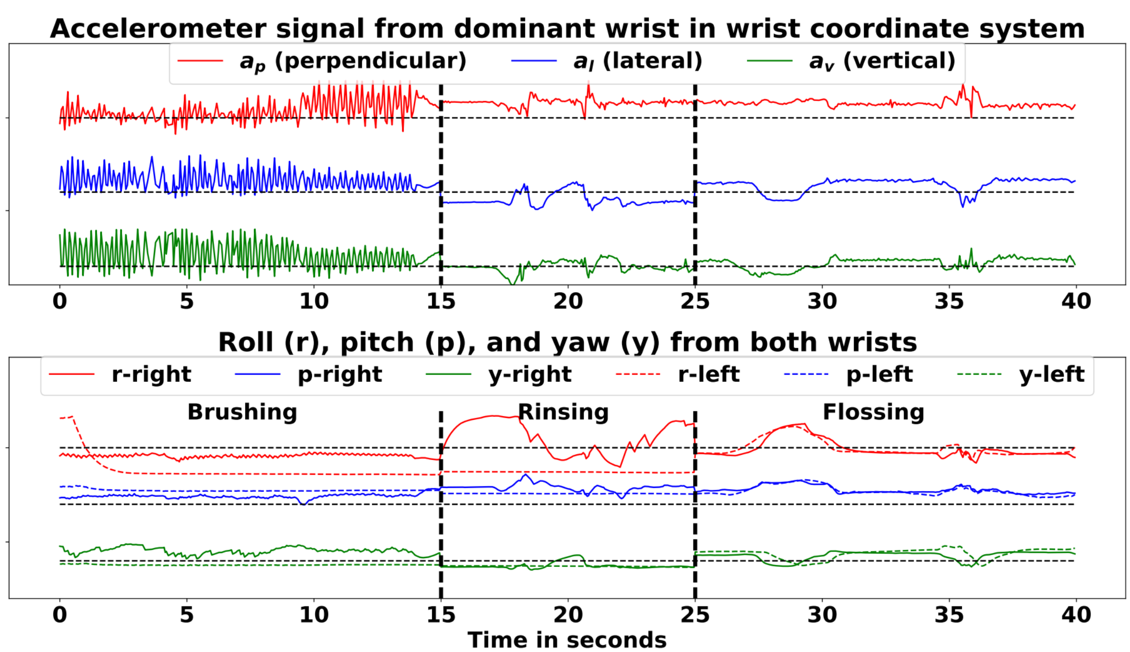


Fig. 4.4: Wrist sensor signals during brushing, rinsing, and flossing with string: analyzing accelerometer patterns.

Figure 4.4 provides visual representations of the accelerometer, gyroscope, and orientation signals captured during different oral hygiene behaviors, specifically brushing with a manual toothbrush, rinsing, and flossing with a string. These observations serve as valuable insights for our model development process.

First and foremost, during brushing, we observe that the wrist is positioned above the elbow, indicated by the positive values of the lateral axis in the accelerometer. The brushing hand exhibits continuous movement, either in an up-down or left-right direction, leading to noticeable periodicity in the accelerometer signals (as shown in the left segment of the top figure). On the other hand, during flossing with a string, there is synchronized motion of both wrists, as depicted in the right segment of the bottom figure.

Furthermore, we note the repetitive nature of wrist motion during these oral hygiene behaviors, which is evident through the significant activity observed in all three accelerometer axes. This repetitive motion characteristic aids in distinguishing brushing and flossing actions from other daily activities.

However, flossing with a string presents certain challenges due to the relatively low magnitude of motion. This low magnitude makes it more difficult to reliably differentiate flossing from other common behaviors. Additionally, during flossing with string, the synchronized motion of both wrists becomes more pronounced across all three orientation axes.

These observations provide important insights into the unique patterns and characteristics of oral hygiene behaviors, which we leverage in developing our model. In the subsequent sections, we will describe our approach to addressing the specific challenges posed by these behaviors

4.5 Overview of mORAL Approach for Reliable Detection of Oral Health Behaviors

We break down the overall problem of reliably detecting oral health behaviors from wrist-worn inertial sensors into several sub-problems to provide a clearer understanding of the process. Firstly, in order to process the data accurately, the signal processing unit requires knowledge of the placement configuration of the wrist-worn device. This necessitates transforming the data into the wrist coordinate system before proceeding with further analysis.

Secondly, our goal is to detect infrequent oral health behavior (OHB) events from a continuous stream of sensor data. To mitigate the computational load on the final machine-learning model, we employ a strategy of identifying candidate windows that are highly likely to contain OHB events. By focusing on these specific windows, we can significantly reduce the amount of sensor data that needs to be processed in the subsequent steps.

Thirdly, after locating the candidate windows, our attention turns to identifying, selecting, and computing features that are relevant for training a classifier. This feature extraction process plays a crucial role in capturing the distinctive patterns and characteristics of oral health behaviors, enabling the model to distinguish them from other activities.

Finally, the utility of detecting these OHB events relies on correctly determining the start and end points of each event. Accurate boundary identification is essential for precisely capturing the duration and timing of oral health behaviors. Therefore, it is imperative to develop techniques that can effectively identify the boundaries of the detected OHB events.

By breaking down the problem into these sub-problems, we can tackle each aspect individually and develop robust solutions that collectively contribute to the reliable detection and analysis of oral health behaviors using wrist-worn inertial sensors.

4.6 Time Series Segmentation: Candidate Identification

Our objective is to develop an efficient model that can automatically detect brushing and flossing events by analyzing continuous sensor data. Despite the sensor being worn for approximately 16 hours each day (during the awake period), the actual events of interest, such as brushing or flossing, typically last only around 4 minutes. Consequently, the majority of the data represents the negative class, making it crucial to devise a highly efficient model capable of real-time processing to filter out non-event data. By doing so, we can significantly reduce the volume of data that needs to be processed by a more complex model. We refer to these filtered data segments as "candidate segments," and this staged detection approach also helps to minimize false positives.

To achieve this, we explore two approaches for candidate identification: the window-based approach and the event-based approach. In the window-based approach, we segment the time series into equal-sized windows and rapidly extract relevant features to quickly determine whether they should be considered as candidate segments. On the other hand, the event-based approach focuses on identifying markers within the time series that may indicate the start of an event of interest. We have observed that during brushing and flossing, the wrist position tends to be higher than the elbow, as illustrated in Figure 3.2(b). By detecting specific patterns of upward and downward wrist movements, we can effectively identify and isolate segments of data that are more likely to contain the events we are interested in.

4.6.1 Toothbrushing Candidate Identification Based on Fixed Windows

In the approach based on fixed window sizes, we experiment with a range of window durations, varying from 2 seconds to 60 seconds, to determine the optimal size. For each window size, we compute both time domain and frequency domain features, aiming to identify the most informative ones. The distribution of filtered windows across different features and window sizes is illustrated in Figure 4.5. Among the considered features, we find that the mean value of the lateral axis of the accelerometer consistently yields the highest rejection rate, making it our primary filtering criterion. Another promising feature is the standard deviation of the accelerometer magnitude, which consistently achieves a rejection rate exceeding 60%. Interestingly, both of these features exhibit stability across various window sizes. When examining window sizes of 15 seconds or longer, most features exhibit a stable rejection rate. However, to accurately capture pause events within brushing episodes, we prefer a smaller window size. Therefore, we settle on a window size of 15 seconds. Furthermore, we explore the potential synergistic effect of combining two fea-

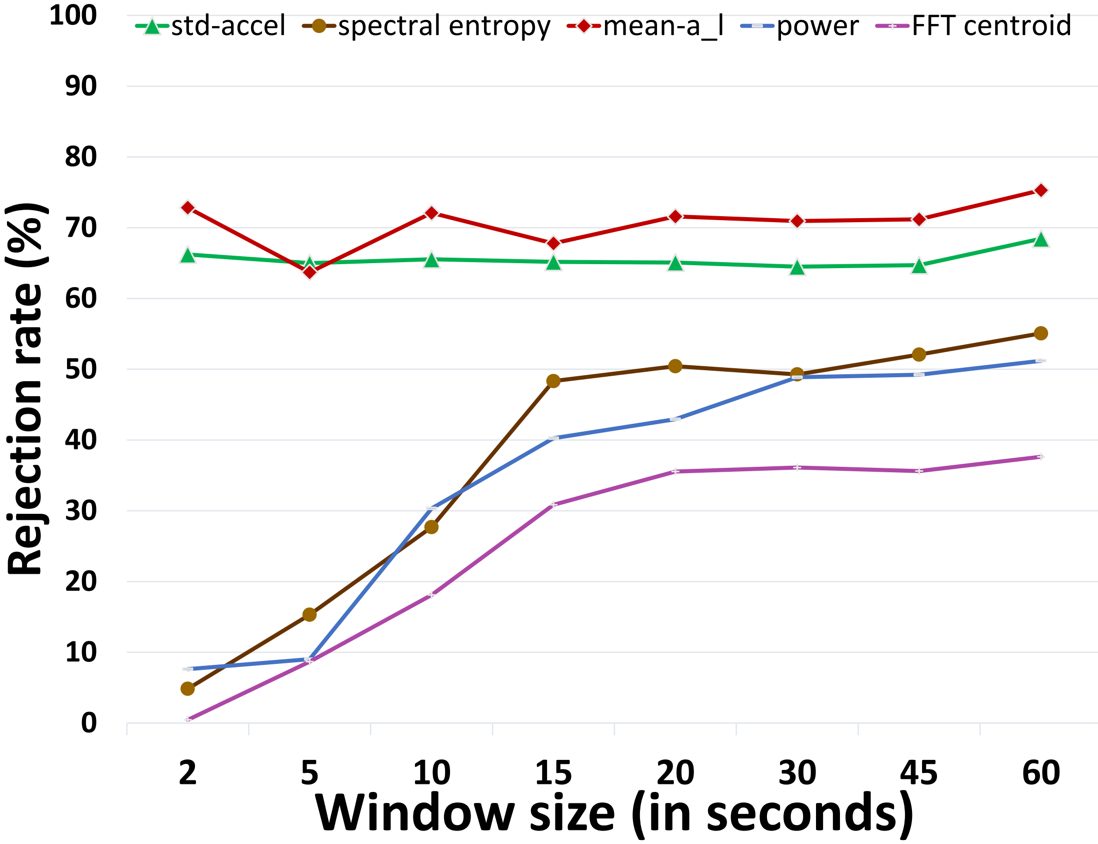


Fig. 4.5: The rejection rate for manual brushing varies as different features are used and the window size is adjusted.

tures to assess whether it leads to a substantial improvement in the rejection rate. Notably, the combination of the top two features enhances the rejection rate from below 70% (using the top feature alone) to an improved rate of 75.6% for a window size of 15 seconds.

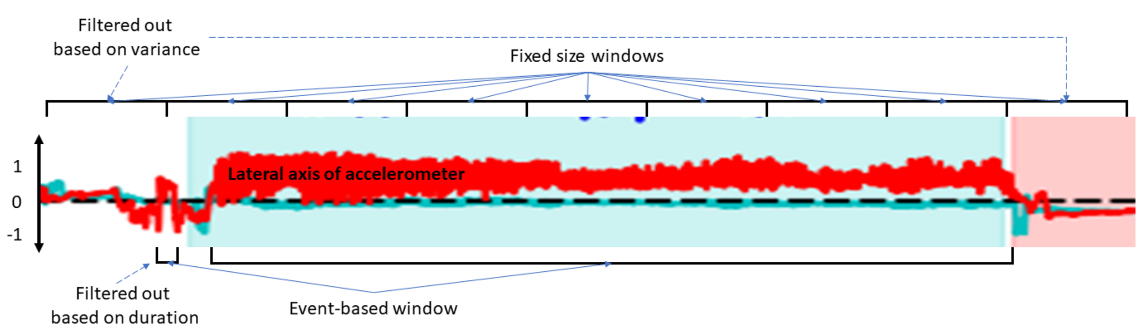


Fig. 4.6: Various methods are employed to identify candidates for brushing events

4.6.2 Toothbrushing Candidate Identification Based on Event

During brushing and flossing, it is observed that the wrist position tends to be higher than the elbow. However, the method used in PuffMarker [16] for detecting hand-to-mouth gestures, which relies on stationary periods of the wrist, is not applicable for identifying candidates for oral hygiene behaviors (OHBs) since the hand is in continuous motion during brushing.

To generate candidate segments for brushing events, we employ a different approach based on detecting upward and downward wrist movements, as illustrated in Figure 3.2(b) and Figure 4.6. Initially, a threshold (TH_l) is determined for the lateral axis of the accelerometer, which filters out samples below this threshold. The optimal threshold value is found to be 0.24 g (using $g = 9.8ms^2$), resulting in filtering 81% of the data. However, this method introduces some temporal clusters in the data. To create candidate segments, clusters are merged if the time difference between the corresponding retained samples is less than 1 second, as small gaps could be caused by jerks or spikes.

To further refine the candidate segments, a filtering process based on their time duration is applied. Optimal values for the minimum (min_{dur}) and maximum (max_{dur}) duration of the segments are determined to include all positive events while filtering out false candidates. For our dataset, we find that $min_{dur} = 11$ seconds and $max_{dur} = 2.5$ minutes yield the best results, filtering an additional 10% of the data. Overall, this method rejects 91% of the data, resulting in an average of 100 candidate segments per day. In comparison, the window-based approach with a 75% rejection rate generates approximately 1,000 candidate segments during 16 hours of sensor wearing.

4.6.3 Flossing Candidate Identification Based on Fixed Size Window

Similar to our approach for candidate identification in brushing, as depicted in Figure 4.7, we explore varying window sizes from 2 seconds to 60 seconds. For each window size, we calculate a range of correlation-based features, leveraging the similarity in wrist orientations during flossing with string. Figure 4.7 illustrates the percentage of filtered

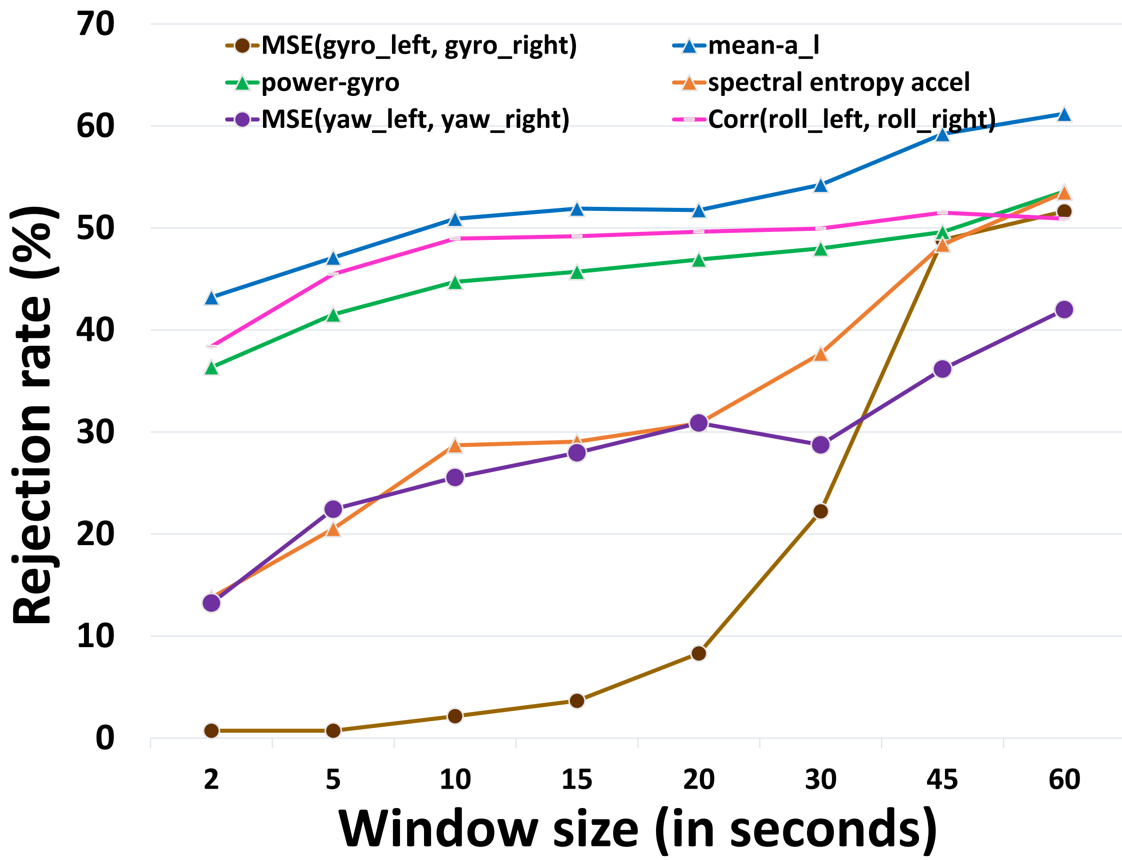


Fig. 4.7: Variation of Rejection Rate with Different Features as Window Size Changes for Flossing with String

windows for each feature and window size. Notably, the top-performing feature, namely the mean value of the lateral axis of the accelerometer, demonstrates a substantial 10% increase in the rejection rate. However, the second and third performing features exhibit only marginal improvements. Recognizing that flossing events are often brief and punctuated by pauses, we opt for a window size of 15 seconds. Furthermore, to strengthen our filtering criteria, we combine the top two features, resulting in an impressive rejection rate of 62.3% for the chosen 15-second window size.

4.6.4 Flossing Candidate Identification Based on Event

Figure 4.8 displays multiple images capturing the process of flossing with string. It is evident from the images that both wrists are consistently oriented in an upward direction during flossing. This observation allows us to apply a similar approach used for brushing

Fig. 4.8: During the act of flossing, the orientation of both wrists is directed upwards.

candidate identification to identify candidates for flossing as well. However, in the case of flossing, we can further refine the identification process since it typically involves both wrists. We utilize the same filtering method as in brushing candidate identification, but we compute separate sets of candidates from each wrist. Only candidate segments that exhibit overlapping data from both wrists are retained. By employing the same duration threshold as used in brushing candidate identification, we discard any candidate segment with an overlapped duration falling outside the specified range. This approach effectively rejects 95% of the data, resulting in an average of 70 candidate segments per day. It represents a substantial increase of 32% in the rejection rate compared to the fixed window-based approach.

4.7 Feature Computation and Selection for Candidate Classification

After identifying the candidates, we compute several time domain, frequency domain, multi-sensor fusion, and cross-wrist features from accelerometer and gyroscope data in each window. Table 4.3 summarizes all the computed features. For *time domain features*, we compute the mean, median, standard deviation, quartile deviation, skew, kurtosis, and zero crossing rate of three axes of accelerometer and gyroscope. For *frequency domain features,* Fourier transformation is applied on the window of data before calculating common frequency-domain features [45]. For *wrist orientation features*, we compute roll, pitch, and yaw that provide information about the orientation of the wrist with respect to gravity.

To obtain robust measurement of orientation features from noisy inertial sensor data, we perform the following processing steps [46, 47]. Accelerometer, $a(t)$, gives a good indication of orientation in static conditions. Gyroscope, $\omega(t)$, provides a good indication of tilt in dynamic conditions, but drifts in the long term. The value of $a(t)$ is noisy, but over longer intervals is useful, as it is more robust to drift [46, 47]. By passing the accelerometer signals through a low-pass filter, passing the gyroscope signals through a high-pass filter, and combining the resultant signals, we compute a final rate function. The idea behind *complementary filter* is to take slow moving signals from $a(t)$, fast moving signals from $\omega(t)$, and combine them. This method combines the strengths of both sensor signal streams. From Figure 4.4, we observe that change in the orientation of both wrists is similar, so we compute correlation and root mean square error of orientation of left wrist and orientation of right wrist. We compute these *cross-wrist features* only for flossing because flossing with string requires both wrists.

In total, we obtain more than 100 features. But, to avoid overfitting (as there are only 139 brushing and 136 flossing events), we use selected features for modeling. The idea behind feature selection is to remove non-informative features. Our goal is to find a subset of features that are a) mutually not correlated but b) highly correlated to the OHBs. In this work, we used the Correlation-based Feature Selection (CFS) [48] to select a subset of the features (25) as in other detection based works [49, 50]. CFS selects features that are mutually uncorrelated but highly indicative of the OHB classes. We describe feature selection further in Section 4.10.2.

4.8 Model Selection and Training

We explore a range of models for toothbrushing and flossing classification, employing grid search to optimize the hyperparameters of each model. The following models are considered:

- **Naive Bayes classifier (NB)**: Naive Bayesian networks are very simple Bayesian networks which are composed of directed acyclic graphs with only one parent (repre-

Table 4.3: Summary of features extracted from each segment of the data

Time-domain (td)	**Frequency-domain (fd)**	**Multi-sensor fusion orientation (ori)**	**Cross-wrist (only for flossing)**
mean	Maxima	Roll	corr(roll-left, roll-right)
median	Energy	Pitch	corr(pitch-left, pitch-right)
standard deviation	Spectral centroid	Yaw	corr(yaw-left, yaw-right)
quartile deviation	Spectral Flux		root Mean Square Error (rMSE)
skewness	Spectral roll-off		
kurtosis			
zero-crossing			
power			

senting the unobserved node) and several children (corresponding to observed nodes) with a strong assumption of independence among child nodes in the context of their parent [51].

- **Random forest classifier (RF)**: Random forests are a combination of tree predictors such that each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest. The generalization error for forests converges to a limit as the number of trees in the forest becomes large. The generalization error of a forest of tree classifier depends on the strength of the individual trees in the forest and the correlation between them [52]. We use Random forests with 100 trees and 1,000 trees.

- **Ensemble method (Ens)**: Ensemble classifiers [53, 54] consist of a set of many individual classifiers (called base-learners) where those decisions are combined to output a single class label. Ensemble learning helps improve machine learning results by combining several models. This approach allows for the production of better predictive performance compared to a single model. Several different methods are proposed to address data imbalance issues by using ensemble methods [55, 56]. Ensemble models also tend to generalize better, which makes this approach easy to handle. In our experiments, we use a combination of Decision tree, K-nearest neighbors (KNN), and Naive Bayes [57].

- **Ada-Boosting (AB)**: AdaBoost, short for *Adaptive Boosting*, is the first practical boosting algorithm proposed by [58]. It focuses on classification problems and aims to convert a set of weak classifiers into a strong one. This is also an ensemble method that learns models on subsets of the training data and boosts the weights of misclassified instances, which allows models to focus on those for improving classification performance.

These models provide a diverse set of techniques for toothbrushing and flossing classification, each with its own strengths and characteristics. Through our experiments, we assess the performance of these models and determine the most effective approach for our specific task.

4.9 Comprehensive Data Processing Pipeline for Oral Health Behavior Detection

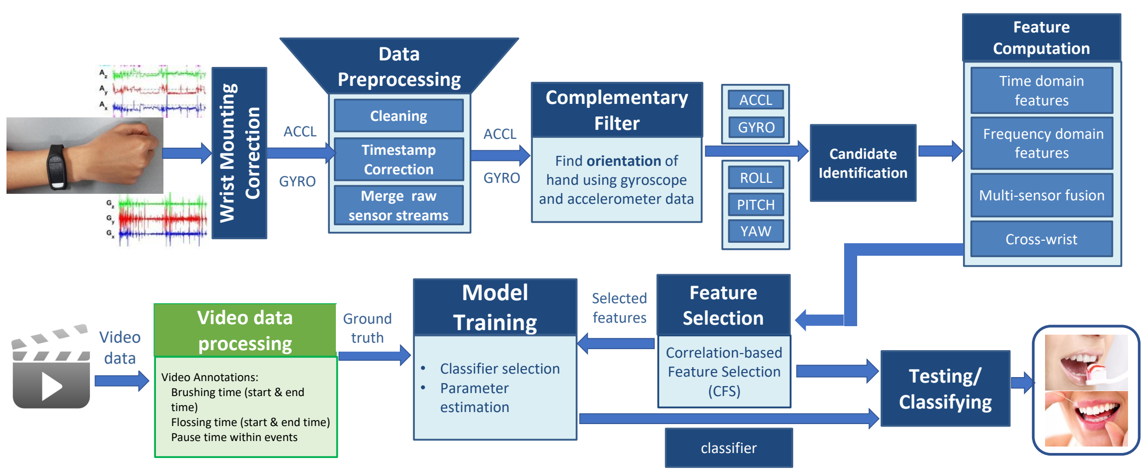


Fig. 4.9: Stages of data processing for training and testing models for brushing and flossing detection.

Figure 4.9 provides a comprehensive overview of the various intricate steps involved in the data processing pipeline for detecting oral health behaviors. This process encompasses several crucial stages, including screening, data cleaning, window size determination, candidate selection, feature computation, and machine learning model training. The following details the comprehensive workflow:

- **Sensor Mounting:** As a crucial first step, we meticulously determine the optimal placement and mounting of the sensor on the wrist to ensure accurate data collection.
- **Data Segmentation:** To ensure the integrity and relevance of the data, we meticulously exclude any segments where the sensor was not worn by the participant. This step helps maintain consistency in the collected data.
- **Data Imputation:** In situations where intermittent data gaps occur, typically of a duration less than 0.25 seconds, we employ advanced imputation techniques, as outlined in [16, 59]. By filling in these gaps, we ensure a continuous and uninterrupted flow of data.
- **Complementary Filter:** To obtain reliable and precise measurements of the wrist's orientation, we employ a sophisticated technique called the complementary filter. This technique combines data from both the accelerometer and gyroscope, leveraging their respective strengths. By filtering the accelerometer signals with a low-pass filter and the gyroscope signals with a high-pass filter, we obtain a final rate function that accurately represents the wrist's orientation with respect to gravity. This method ensures robust and accurate orientation measurements, even in the presence of noise and sensor drift.
- **Candidate Window Identification:** The next crucial step involves identifying and marking candidate windows that potentially contain oral health behaviors. These windows are identified based on specific criteria and temporal patterns observed in the sensor data.

- **Feature Computation:** Once the candidate windows are identified, we proceed to compute a diverse set of features from the sensor data within these windows. These features encompass various domains, including time-domain, frequency-domain, multi-sensor fusion, and cross-wrist features. The features capture essential characteristics and patterns related to oral health behaviors.

- **Model Training:** To classify and detect oral health behaviors effectively, we leverage machine learning techniques. Various classification models are considered, including Naive Bayes classifier (NB), Random Forest classifier (RF), Ensemble method (Ens), and Ada-Boosting (AB). Grid search is employed to optimize the hyper-parameters of each model, ensuring the best possible performance and accuracy.

In summary, the data processing pipeline for detecting oral health behaviors involves a meticulous sequence of steps, ranging from data screening and cleaning to candidate selection and feature computation. The integration of advanced techniques and machine learning models enables accurate and reliable identification of oral health behaviors, empowering individuals to monitor and improve their oral hygiene practices effectively.

4.10 Model Evaluation

We now present results from our evaluation. We start by comparing the performance of distribution-based and PCA-based methods for identifying the correct configuration of the wrist sensors. Then, we observe the performance of the selected feature set for detecting OHB events. We use the selected feature set for further experiments. We compare the performance of different models choices for detecting brushing and flossing events. We evaluate the impact of using window-based or event-based approaches to candidate identification, as they impact which segments of data the machine learning models are applied to. Finally, we analyze the error in duration and start/end times of the detected brushing and flossing events.

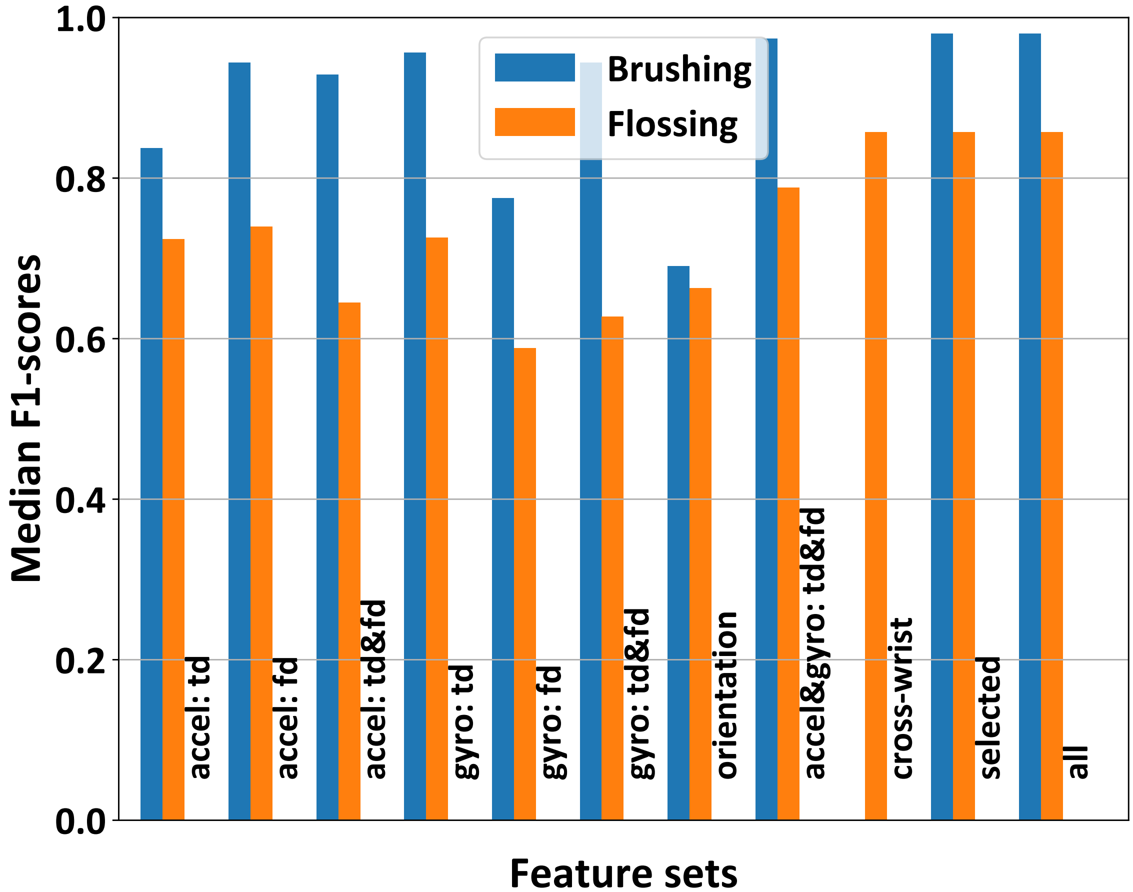


Fig. 4.10: F1-scores for brushing and flossing detection using different feature sets using LOSOCV evaluation

4.10.1 Metric for Evaluating Performance of Classifiers

For classifaction performance evaluation, we use a leave-one-subject-out cross-validation (LOSOCV) experiment where we exclude a single user's data for testing and use the remaining data for training. We repeat this procedure for all the users. We present results in box plots [60] and report median values. We note that in our dataset (similar to real-life usage), only 0.1% of the data are positive instances and rest 99.9% of the data are negative instances. If accuracy alone is used to measure the classification performance, then a simple model that classifies all testing samples into the negative class, will produce an excellent accuracy of 99.9%. But, its recall and precision will be zero and F1-score (for detecting the positive class) will be undefined. Since our interest is in reliably detecting

Table 4.4: Mean recall (R), precision (P), and F1-scores (F1) for brushing and flossing detection using leave-one-subject-out cross validation with Gaussian Naive Bayes (NB), Random Forest with 100 trees (RF-100), Random Forest with 1000 trees (RF-1000), Ensemble method (Ens), and Ada-Boosting (AB).

Models	**Brusing**						**Flossing**					
	Window based			**Event based**			**Window based**			**Event based**		
	R	*P*	*F1*	*R*	*P*	*F1*	*R*	*P*	*F1*	*R*	*P*	*F1*
NB	**0.94**	0.02	0.05	**0.95**	0.51	0.62	**0.99**	0.19	0.33	**0.97**	0.58	0.71
RF-100	0.59	0.94	0.70	0.73	0.87	0.77	0.31	0.62	0.39	0.63	0.89	0.72
RF-1000	0.61	0.94	0.71	0.73	0.87	0.77	0.31	**0.66**	0.41	0.61	0.89	0.70
Ens	0.57	**0.95**	0.68	0.75	0.88	0.78	0.28	0.63	0.35	0.61	0.89	0.72
AB	0.69	0.88	**0.76**	0.83	**0.97**	**0.86**	0.49	0.59	**0.50**	0.79	**0.93**	**0.85**

the positive class, we use recall, precision, and F1-score to measure the performance of our classifiers. In addition, we report false positives detected per day to provide a sense of the model's performance in daylong wearing.

4.10.2 Feature Selection

We evaluate the discriminatory power of different feature sets. Figure 5 shows different feature sets and their discriminatory power. We create a set of time-domain features, a set of frequency-domain features, and finally a set of both time-domain and frequency-domain features from the accelerometer data. We perform the same selection from the gyroscope data. Next, we create a set of orientation related features. We create another set by combining both accelerometer and gyroscope features. We also create another set using cross-wrist features that is used only for flossing detection. Finally, one feature set is generated using the feature selection algorithm described in 4.7. To detect OHBs, the selected feature set (consisting of 25 features) performs almost the same as using all the features. To understand the performance of the proposed model, independent of the error in detecting correct orientation, this evaluation and the rest of the experiments are done after virtual orientation of the wristband.

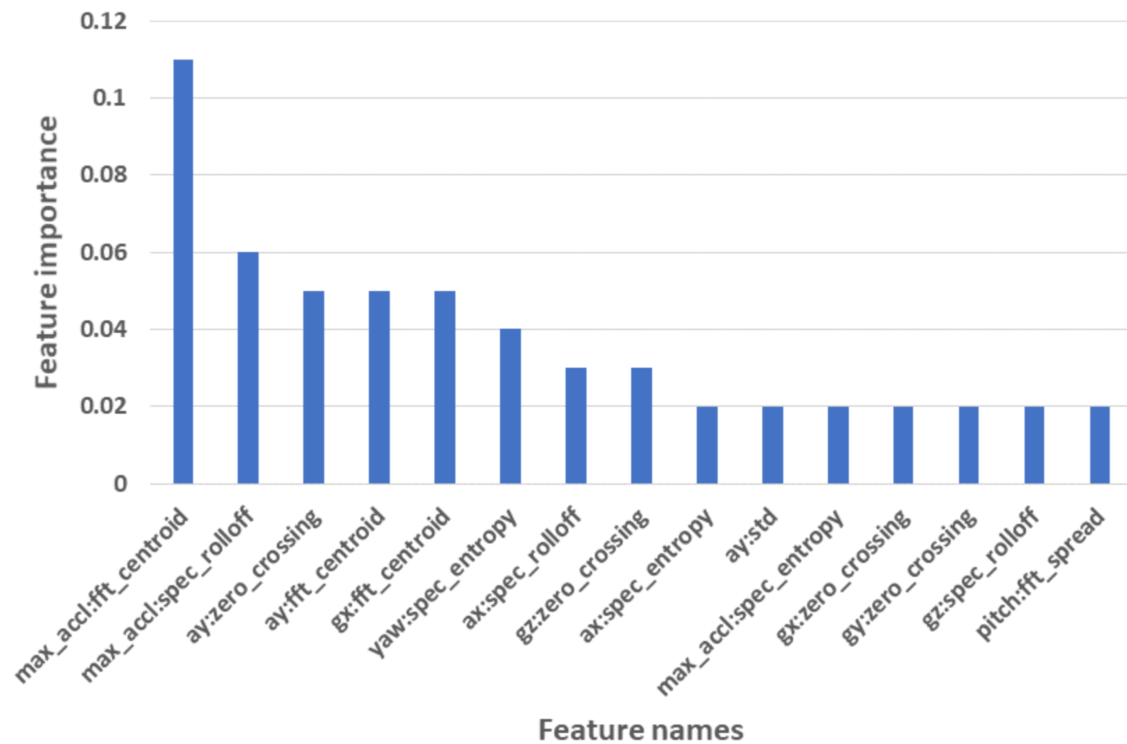


Fig. 4.11: Accuracy of brushing detection for different feature sets using LOSOCV evaluation

4.10.3 Performance of Brushing Detection

We evaluate both window-based and event-based approaches to candidate identification (see Section 4.6) for their impact on classification performance. For both experiments, Gaussian Naive Bayes, Random Forest with 100 trees and 1000 trees, Ensemble methods, and Ada-boost classifiers are applied to detect brushing behaviors. For this experiment, we include only those participants for whom we have usable sensor data for at least three manual brushing episodes (i.e., 142 manual brushing events from 21 participants over 157 days).

Window-based Approach:

As described in Section 4.6, we use a window size of 15 seconds. Recall, precision, and F1 scores appear in Figure 4.13. For the Naive Bayes model, median precision rates are lower than other models, but the recall rate is the highest. Ensemble method, Random

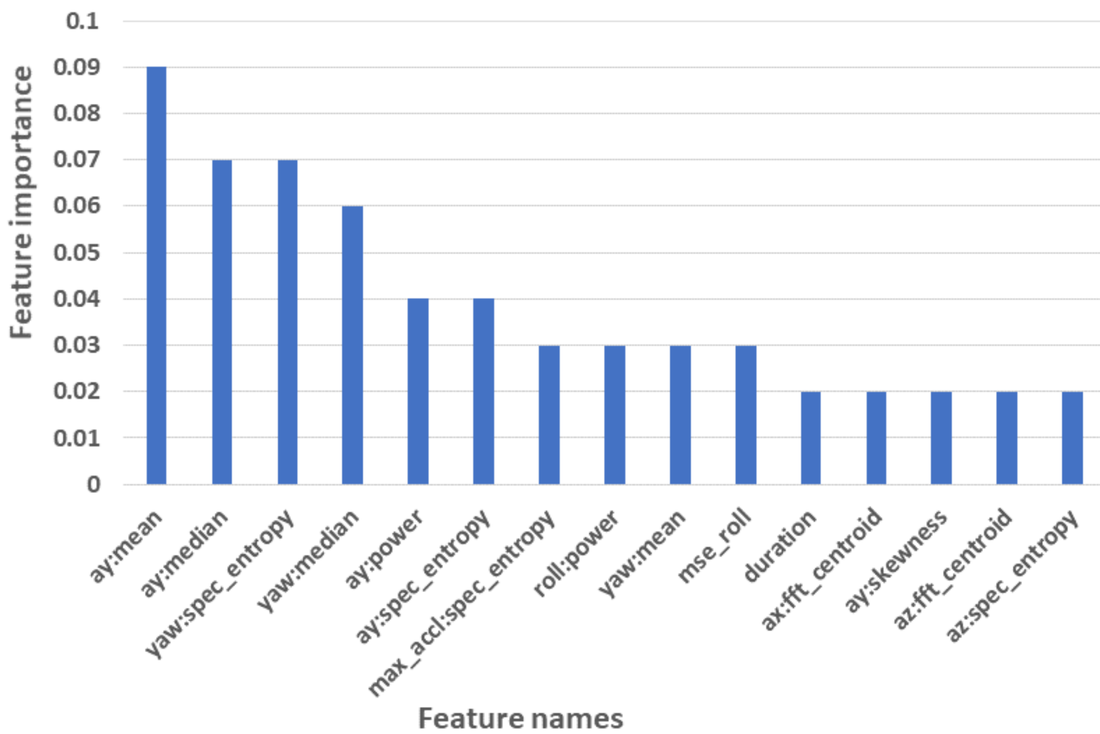


Fig. 4.12: Accuracy of flossing detection for different feature sets using LOSOCV evaluation

Forest with 100 trees and 1000 trees produce a high median precision rate of 100%, but recall rates of only 64.74%, 67.15%, and 68.27% respectively. Ada-boost's median recall rate is 70.44%, and a precision rate of 85.15%. The Ensemble method and Random forest with 1000 trees produce a median F1 scores of respectively 71.43% and 71.86%, but the highest median F1-score of 75.58% is produced by Ada-boost.

Event-based Approach:

Recall, precision, and F1 scores when using event-based approach for candidate generation are shown in Figure 4.14. The performance of all the models are higher than that for a window-based approach. The Ada-boost model provides the best recall and precision. It has a median recall of 100%, a precision of 100%, and an F1 score of 95%, which is 19.42% higher than that for the window-based approach. With the Ada-boost model, we

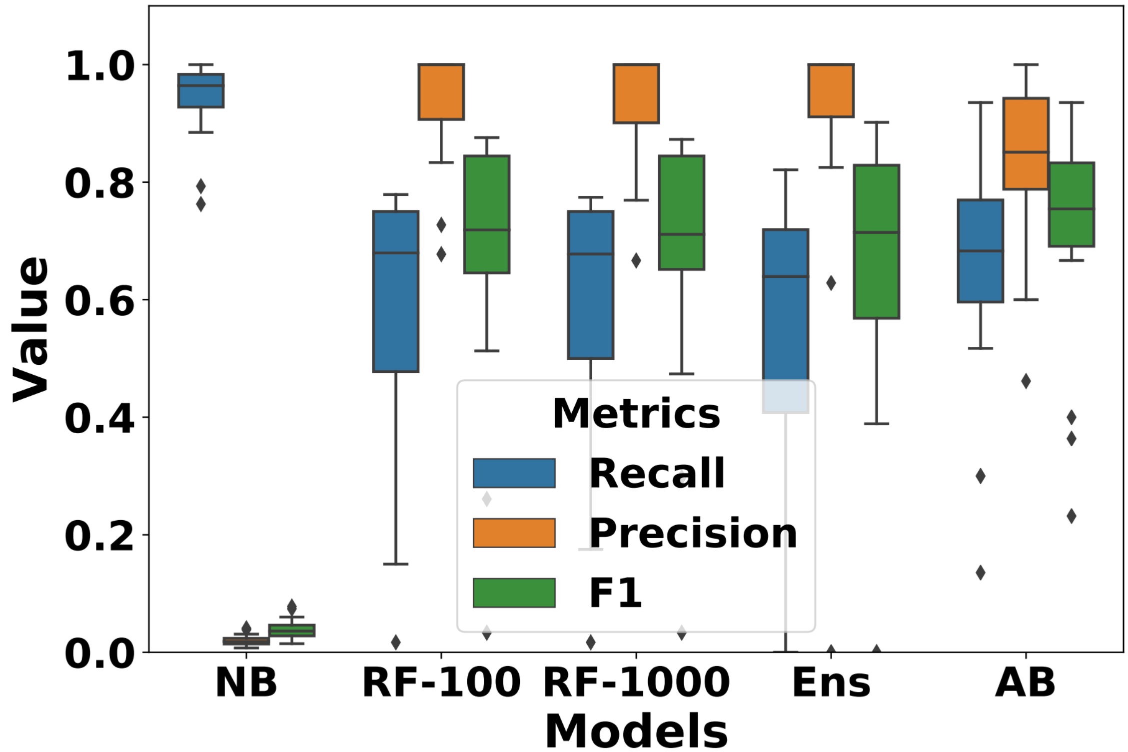


Fig. 4.13: Box-plots of leave-one-subject-out cross validation performance of brushing detection with a (15 seconds) window-based approach for candidate identification.

get about one false positive every ten days (i.e., 16 false positives on 157 person days of data).

4.10.4 Performance of Flossing Detection

We follow a similar approach for evaluating the detection of flossing as for brushing detection. For this experiment, we include data from only those participants for whom we have usable sensor data for at-least three string flossing episodes (i.e., 95 string flossing events from 16 participants over 125 days).

Window-based approach:

Recall, precision, and F1 scores appear in Figure 4.15. We obtain the best median F1 score of 55.02% for Ada-boost. It has a median precision of 65.30%, and a median recall rate of 44.50%.

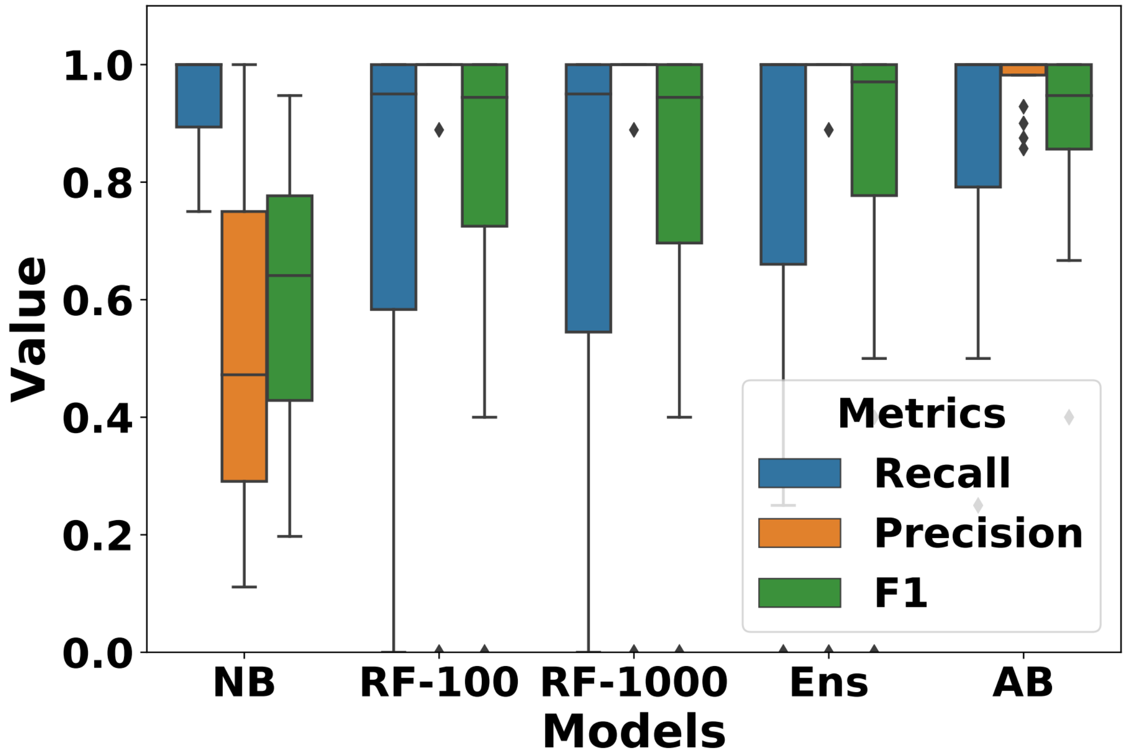


Fig. 4.14: Box-plots of leave-one-subject-out cross validation performance of brushing detection with an event-based approach for candidate identification.

Event-based Approach:

Recall, precision, and F1 scores are shown in Figure 4.16. Similar to brushing, we obtain the best median F1 score with the Ada-boost model. Its median recall is 75%, precision is 100%, and the F1 score is 82.65%, which is 27.63% higher than that for a window-based approach. With the Ada-boost model, we obtain about one false positive every twenty-five days (i.e., 5 false positives on 125 person days of data).

We observe that our detection accuracy for flossing is lower than that for brushing. One reason for a lower recall rate in detecting flossing is due to low and infrequent movement of hands (captured in inertial sensor data) during flossing. Improving the recall rate for flossing detection remains a future work.

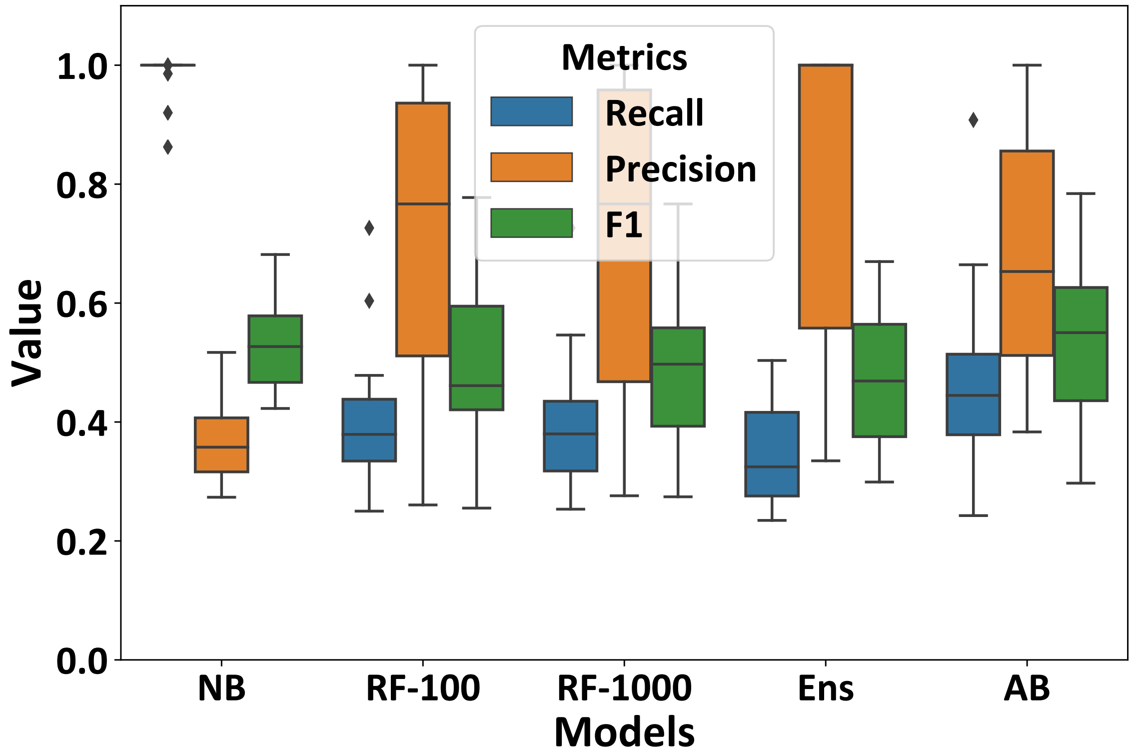


Fig. 4.15: Boxplot of leave-one-subject-out cross validation results of flossing detection with a (15 seconds) window-based approach for candidate identification.

4.10.5 Performance on Detecting Duration and Start/End Times

In addition to reliably detecting brushing and flossing events, it is also desirable to accurately detect the duration of these events (due to their predictive nature in dental disease outcomes). Further, as these events are detected from a continuous time series of sensor data, a single event can be split into multiple windows due to pauses. Portions of the actual event can also be missed or overestimated. We would like the detected event to closely match the actual event in both the start and end times. For this experiment, we use the data for brushing and flossing as described in Sections 6.4 and 6.5. We first develop some metrics to measure the accuracy of detecting start/end times and duration before reporting the performance of our models.

Event: An event consists of a start time and an end time. We define a pause event as a tuple of start time and end time ($E_p = (t_s, t_e)$). Pause events can be either inter-event

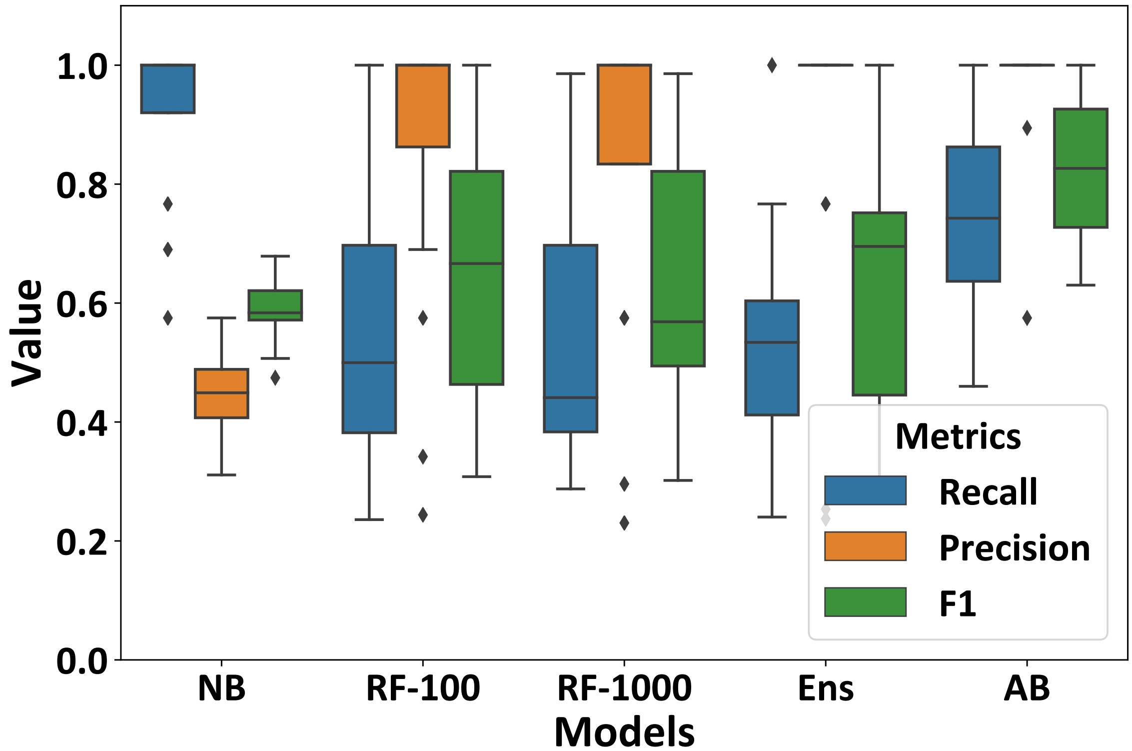


Fig. 4.16: Boxplot of leave-one-subject-out cross validation results of flossing detection with an event-based approach for candidate identification.

pause or intra-event pause. Since we are not dealing with inter-event pauses, we use pauses to refer to intra-event pauses. Each oral hygiene event consists of a start time, an end time, and a list of pause events. For example, ($E_b = (t_s, t_e, E_p)$) is a brushing event with a list of pause events, where $t_s < e.t_s$ and $e.t_e < t_e$, for all $e \in E_p$, and ($E_f = (t_s, t_e, E_p)$) is a similarly defined flossing event.

Duration of an Event: The duration of an event $E = (t_s, t_e, E_p)$ (after removing pauses) is defined as follows.

$$d(E) = (E.t_e - E.t_s) - \sum_{E_p \in E.E_p} (E_p.t_e - E_p.t_s)$$

Error in Event Localization: We divide the error by the duration of the actual event for normalization. Let E_{actual} be the actual event. For a detected event, $E_{detected}$, we use two error measurements:

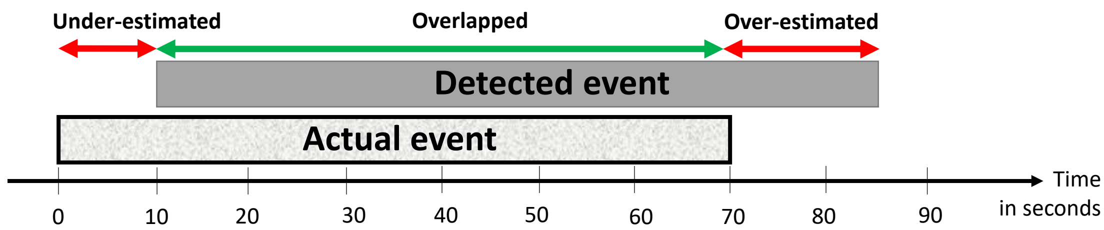


Fig. 4.17: An example of duration error ($err_{dur} = \frac{|70-75|}{70} = \frac{5}{70}$) and Localization error ($err_{boundary} = \frac{|10+15|}{70} = \frac{25}{70}$)

- *Duration Error* is defined as the difference between actual duration and detected duration, i.e.,

$$err_{dur} = \frac{|d(E_{actual}) - d(E_{detected})|}{d(E_{actual})}$$

- *Average Localization Error* of the event measures error in locating the boundary. Error can happen in both boundaries due to either over-estimation or under-estimation. We define average localization error as

$$err_{boundary} = \frac{AVG\{|E_{actual}.t_s - E_{detected}.t_s| + |E_{actual}.t_e - E_{detected}.t_e|\}}{d(E_{actual})}$$

Figure 4.17 shows a scenario where duration error is low but localization error is high.

For each brushing and flossing event detected by the best model, we evaluate the error in the duration of the detected event with that of the corresponding actual event (from video data). If a detected event does not have any temporal overlap with the actual event, it is not considered to be a true recall and is excluded from this analysis. We use regression analyses to analyze the errors in duration.

Figure 4.18 depicts regression analyses between the actual duration and the corresponding duration of the detected events (for both brushing and flossing). The Pearson correlation coefficient value r is 0.95 for brushing and 0.98 for flossing. Mean square error for brushing is 11.07 seconds and for flossing, it is 8.9 seconds. As a percentage of the actual event duration, these represent duration errors of 7.2% and 6.5% respectively.

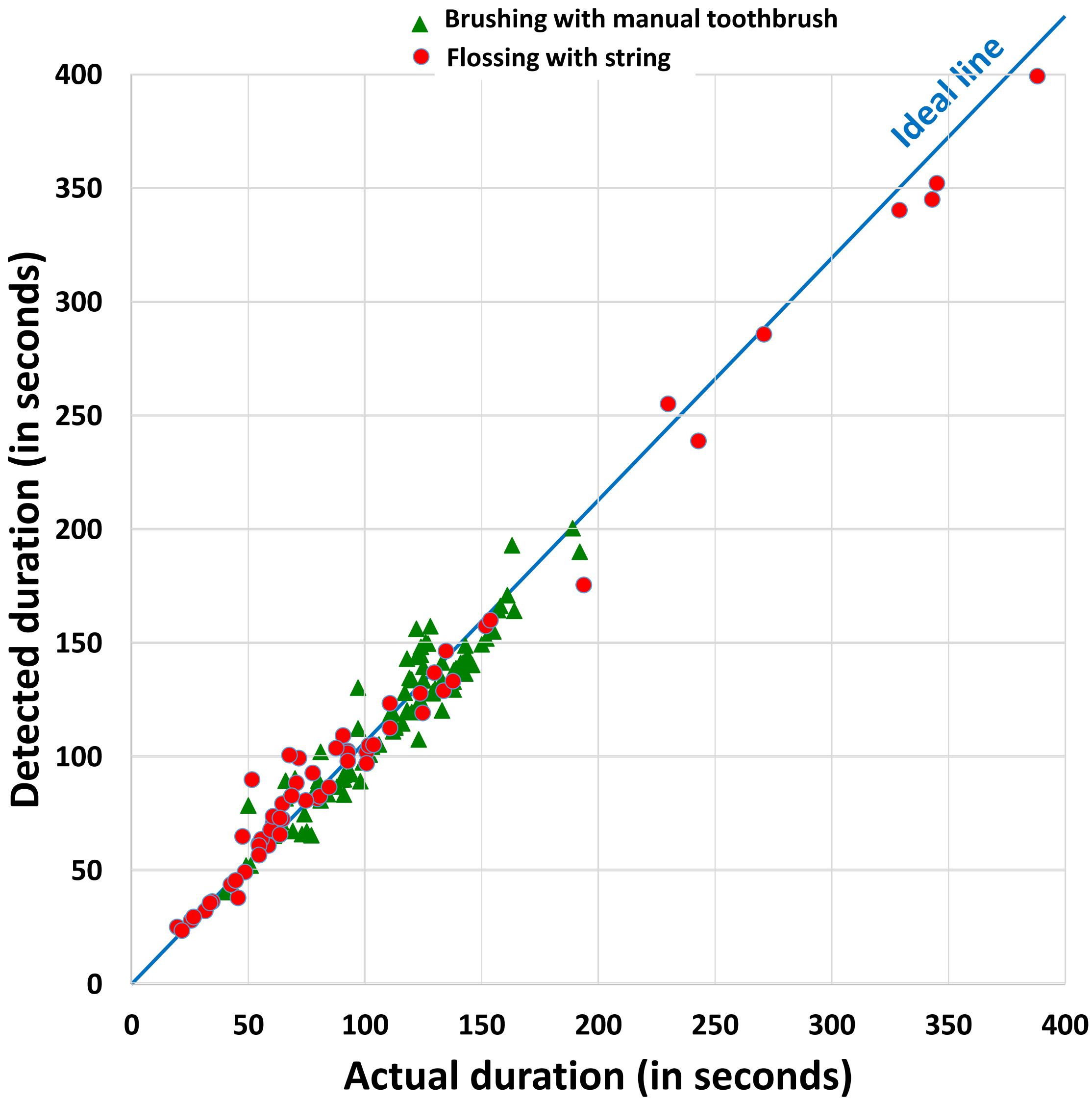


Fig. 4.18: Scatter plot of actual duration vs. detected duration, $err_{dur} = 7.2\%$ for brushing and $err_{dur} = 6.2\%$ for flossing.

Next, we analyze the temporal alignment of the detected event with the actual event. Figure 4.19 shows the results for brushing and flossing respectively. We observe that the error in alignment is small for both events. The error in accurately locating the start and end times are 4.1% for brushing and 3.5% for flossing.

4.11 Limitations and Future Works

The presented mORAL model achieves a median recall rate of 100% with one false positive every nine days. For flossing, the median recall rate is lower at 75%, but the

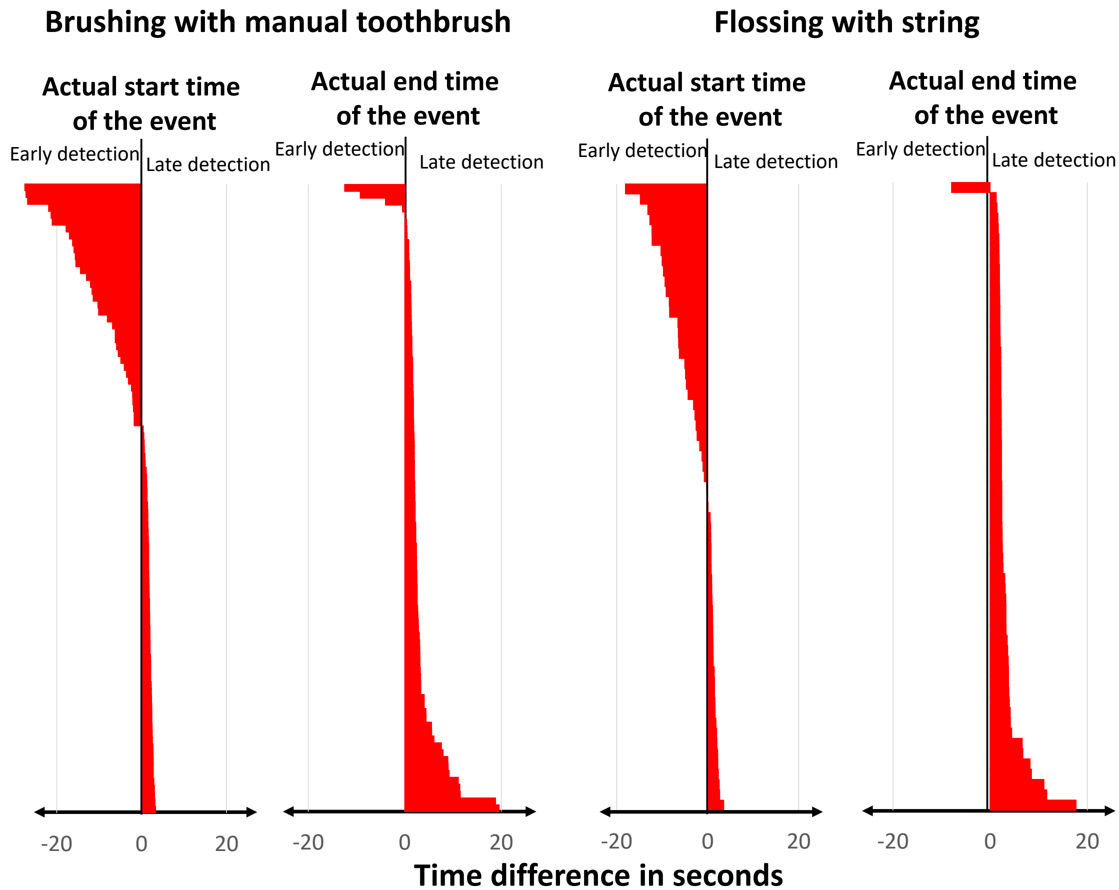


Fig. 4.19: Error in start/end times, $err_{boundary} = 4.1\%$ for brushing and $err_{boundary} = 3.5\%$ for flossing.

false positive rate is also lower at one every 37 days. In addition to low recall rate for flossing, this work has several other limitations that open up numerous opportunities for future works.

First, our model does not include oral rinsing behavior detection, which is an important oral health behavior. Especially when combined with a model for detecting eating, it can present interesting intervention opportunities. Modeling the transition in the sequence of brushing, rinsing, and flossing behavior can potentially be used to improve the detection of each of these activities. Second, our model can detect flossing with string, but not with picks. Using deep learning models, especially with a larger labeled dataset, can potentially improve detection performance further.

Third, a significant limitation in this work is the use of two wrist sensors. We used two wrist sensors to understand the value of having the wrist sensors in both hands. We note that detection of brushing uses only the data from dominant hand. Therefore, detection of brushing should work as long as the user is wearing the sensor in their dominant hand. But, the model needs to be adapted to detect flossing from sensor data collected on only one wrist.

Fourth, detecting brushing with the SmartBrush from wrist sensors is another future research direction. Even though smart toothbrushes detect the brushing event due to the user pressing a button to start/stop the device, a wrist-based model can be used to assign the brushing event to a specific user, e.g., when a brush handle is shared in a family.

Fifth, even though significant work has been done in detecting the brushing pressure and detecting the tooth surface or quadrant being brushed, using smart or instrumented toothbrushes, our work opens up opportunities to develop these capabilities for brushing with manual toothbrushes, especially due to low boundary location error with our model. Doing so will benefit a large population that still uses manual toothbrushes.

Finally, additional work is needed to adapt our model for detecting the brushing and flossing events in real-time. Although our model is computationally efficient, the virtual orientation process may require a couple hours of data. Developing a quicker method for virtual orientation can help not only with our model, but also in models developed for detecting other behaviors such as eating and smoking that involve hand-to-mouth gestures.

4.12 Chapter Summary

mHealth research is rapidly growing to incorporate detection, prediction, and just-in-time intervention for a diversity of markers of health and well-being. Our work extends this paradigm into the oral health domain, an oft-underserved realm of biomedicine and public health. It opens doors for extending several benefits only available to those with access to smart toothbrushes to a broader population still using manual toothbrushes. It also opens up new opportunities for designing sensor-triggered interventions for improving oral health behaviors. For example, incorporating the passive detection of OHB's into a context-based approach can enhance our understanding of how OHB biomarkers are associated with other biomarkers such as stress, semantic location, tobacco, alcohol, or other substance use, and may provide substantial future insight to these behaviors and lead to new supporting protective interventions.

In addition to providing a solution for monitoring oral health behaviors, our work also provides a direction for the computing community when developing models to reliably detect rare daily events from dense time series of data collected in everyday life. For example, it shows the promise of the event-based approach for identifying candidate windows before applying the machine learning models. It further indicates that identifying discerning characteristics of the target event and translating them into efficient data models results into a more accurate overall model.

Chapter 5

Characterizing of the Detected Brief Daily Behavior: Identifying Brushing Teeth Surfaces Using Wrist-Worn Sensors

5.1 Introduction

Dental disease (caries and gum disease) is very prevalent globally, affecting 53 million people in USA alone. A primary reason for continued prevalence of dental diseases despite regular brushing is that people may not be brushing each tooth surface adequately, missing some surfaces completely, while spending disproportionate time on other surfaces. When saliva combines with particles from food and drinks we consume, a colorless, sticky biofilm containing bacteria known as *dental plaque* forms on our teeth. Unmindful or poor brushing habits allow plaque to accumulate over time, leading to gum disease, tooth decay (and cavities), and tooth loss. Beyond the pain and suffering, oral health problems affect the ability to eat and swallow, speak and socialize. Importantly, because the mouth is the main portal for entry to the body, poor oral health can contribute to a range of conditions and diseases including respiratory diseases, endocarditis, cardiovascular diseases, and pregnancy and birth complications. What makes matters even worse is the accompanying steep cost of dental and health care, that many without insurance struggle to bear. However, the good news is that people can still prevent much of the complications arising from poor brushing habits through technology-driven awareness.

Smart toothbrushes [23] equipped with Bluetooth connectivity, gyroscopes, and accelerometers are beginning to address some key aspects of oral hygiene. They use beeps, vibrations, and visualizations on smartphones to reinforce a recommended routine of spending 30 seconds on each quadrant — upper right, upper left, lower right, and lower left — for adequate brushing, a key component of proper dental care [61, 62]. Identification and evaluation of toothbrushing activities coupled with a feedback system to encourage proper brushing has been a focus of several works on understanding and improving human oral health behavior. They include assistive technologies to promote brushing habits among children through playful experiences [37, 28]; supporting users in learning a complex brushing technique with realtime feedback [40, 63, 64]; encouraging regular toothbrushing using virtual aquarium or mirror [65, 25]; enabling self-examination and creating awareness about common oral health conditions [66]; creating plaque awareness [19]; and helping handicapped people without arms to brush their teeth correctly [20]. But, these works use either smart toothbrushes, electric toothbrushes, or toothbrushes fitted with sensors [24, 63, 25, 44]. As such toothbrushes are still used by a small minority, these advances do not benefit most people who still use manual toothbrushes.

Wrist-worn inertial sensors in smartwatches and activity trackers are increasingly being used to detect various activities like eating [15], smoking [17, 16], drinking [67], and hand washing [68]. A recent work [29] presented the mORAL model to detect the start and end times of brushing and flossing activities. Although this work enables monitoring of brushing and flossing events for a large population of users still brushing with a regular toothbrush, the capability of monitoring which surfaces are not adequately being brushed is still lacking.

In this paper, we present a new *mTeeth* model to detect which tooth surface is being brushed using a regular uninstrumented toothbrush from data collected by inertial sensors in wrist-worn activity trackers and smartwatches. We successfully address several challenges in detecting brushing on specific tooth surfaces, which receive only a few seconds of brushing before a user transitions to another surface.

First, we enhance the utility of a publicly available wrist-worn inertial sensor dataset collected from daily life of participants by annotating it with fine-grained labels of which surface is being brushed on and moments of transition. We develop a hierarchical categorization of teeth surfaces in nine types that is suited to detection by sensors. We analyze the labeled data to quantify between-person variability in brushing patterns, within-person between-episode variability, and within-episode between-surface variability in brushing duration.

Second, we find that there are time synchronization errors of several seconds between sensor data and associated video, even though both are collected on the same smartphone. As transition among teeth surfaces last only milliseconds, we propose an algorithm to tightly synchronize the two data sources that does not have any explicit anchor event. We find that this improves the F1 score for surface classification by 13%.

Third, we observe that time spent on a brushing surface can be as low as a few milliseconds and as high as few tens of seconds. This prevents unambiguous label assignment in fixed-length window-based approach to data segmentation. We observe that an anchor micro-event called *brushing stroke* occurs during all surface transitions. We propose a computationally lightweight method to identify brushing strokes using wrist-worn inertial sensors.

Fourth, we identify and compute several features from each brushing stroke. To leverage the hierarchical organization of teeth surfaces and sequence of transitions among them, we select and train a Dynamic Bayesian Ensemble model. We train and test on one week of brushing data from 19 participants to analyze the impact of wide between-person and within-person variability on the performance of machine learning models for dynamic brushing surface identification using wrist-worn inertial sensors.

5.2 Related Works and Key Contributions

Our proposed mTeeth model assumes that the start and end of a brushing event can be identified from wrist-worn inertial sensors. Development of such models has progressed from detecting brushing from hand gestures in the context of detecting a vast amount of activities of daily living (ADL) [7, 13] in scripted settings to recently proposed models for automatically detecting toothbrushing in the wild using wrist-worn inertial sensors [29, 69]. In the following, we discuss prior works that aim to detect the specific teeth surface being brushed on.

5.2.1 Toothbrushing Surface Detection from Smart or Instrumented Toothbrushes

In [21] and [22], the authors designed a smart toothbrush fitted with a 3-axis accelerometer and magnetometers to trace which group of teeth the user was brushing at a particular moment. This work divided the teeth into several brushing regions before developing a k-means clustering-based model to detect them [21] and determine if brushing in each of those areas was done appropriately or not. Their smart toothbrush based approach achieved an overall accuracy of 97.1% for a total of 15 brushing regions. Smart toothbrush based solutions are now commercially available that guide users on the correct way of brushing. For example, [23] includes a brushing head capable of giving real-time feedback to the user based on brushing pressure. A paired smartphone provides visual display to determine which surface is being brushed.

To provide an alternative to smart toothbrushes, [24] proposed a smartwatch based recognition system to evaluate the brushing quality. They attached magnets to a normal toothbrush to build an arm motion model with inertial data collected from wrist-worn sensors for real-time detection of brushing gestures. The system was able to detect brushing surfaces with an average precision of 85.6% by dividing the teeth set into 16 different surfaces following the Bass technique. In [25], a 3D colored ball was attached at the tail of a toothbrush to estimate which dental side was being brushed by analyzing the spatial position and orientation of the ball. As these methods rely on sensors in a toothbrush, they are not applicable to detecting brushing surfaces with regular toothbrushes.

5.2.2 Toothbrushing Surface Detection from Audio and Video

An initial work [26] evaluated brushing from acoustic signals captured by a smartphone placed next to the sink. It recorded audio signals from which 12-order Mel-Frequency Cepstral Coefficient (MFCC) features were extracted to train a Hidden Markov Model (HMM) for recognizing toothbrushing activities. It achieved a classification accuracy of 78.3%. Similarly, [27] proposed a tooth brushing monitoring system based on acoustic inputs. They deployed an asymmetrical sound-field detector which had a Bluetooth earphone and a throat microphone to capture acoustic inputs from the air and human body, respectively. The two different sources of inputs carried a rich set of characteristics from the environment and a living entity. To reduce computational complexity, a series of statistical inferences from time and frequency domains were extracted for training different models.

Some works use image analysis to detect teeth surfaces. A computer based web-cam is used in [28] to identify the position of the smart toothbrush. It has a visual feedback system, equipped with a physical avatar whose teeth are made of LEDs for tracking children's toothbrushing activities in real time. Another work [25] detects toothbrush and the face of its user with the help of a smartphone's front camera. The smartphone's display works as a "virtual mirror" to locate a person's face with a toothbrush through a face tracker and replaces the captured image with that of an avatar. The avatar is able to completely mimic the user's gestures and expressions, and points out any wrong movement. As these works rely on some instrumentation of the environment to detect teeth surfaces, their methods are not directly applicable to address the technical challenges faced in detecting teeth surfaces being brushed from wrist-worn inertial sensors alone.

5.2.3 Toothbrushing Surface Detection from Wrist-Worn Inertial Sensors

Even though [24] used an instrumented toothbrush, they also trained a model to detect brushing surfaces using only wrist-worn inertial sensors, that was further improved by [69]. Sensor data is divided in 1 second segments in [24] and 1.2 second segments in [69]. They used either a video or an observer to decide the surface labels of each data segment. For labels, [24] used 16 Bass technique surfaces, while [69] used 13 teeth surfaces, tongue brushing, and raised hand state. The number of episodes or the amount of data collected is not reported in either work. Also, details of how second-level precision was achieved in labeling either from video or observer are missing. Both trained their participants to brush using the Bass technique. Any brushing sequence not following the Bass technique are excluded from surface classification in [24]. Naive Bayes classifier together with a Hidden Markov Model is trained in [24], while an attention-based LSTM is used in [69].

A precision of 75.9% with wrist-sensor only model is reported in [24], while an accuracy of 97% is reported in [69], both using 10-fold cross-validation.

Our work differs from [24, 69] in several respects and presents an alternative approach to surface classification. First, the goal of both prior works was to achieve homogeneity in the brushing pattern of participants by training them. Our goal instead is to observe the natural brushing habits of participants and still aim to detect the surface being brushed and transitions among them, despite natural variability. Second, a fixed window of 1 or 1.2 seconds can include brushing on two different surfaces and the transition time, which creates ambiguity in which labels to assign to these windows. Leaving them unlabelled can exclude 30-40 seconds of data from a 120-second session, as an average of 30 transitions occur in a 90-second brushing session. Therefore, we use a new anchor micro-event (i.e., brushing strokes) that naturally occurs between all transitions among surfaces, and thus separates the data from different surfaces. Third, we find that the number of samples in our data segments (i.e., in a brushing stroke) consists of only 4-5 data points (at a sampling frequency of 16 Hz). They are sufficient to detect peaks and valleys, but are not suitable to train a deep learning or other models that identify complex features automatically. But, the Dynamic Bayesian Ensemble method we present is still able to achieve similar high accuracy (with median F1 scores of 94% to 100%) for distinguishing among in/out, left/center/right, and up/down surfaces.

5.2.4 Summary of Key Contributions

In summary, the presented work makes the following novel contributions over prior works.

1. Unambiguous Labeling: Ours is the first work to use the micro-event of *brushing strokes* to assign clean labels to each sensor data segment. Prior works on brushing surface detection used fixed-length windows [24, 69] that may make unambiguous label assignment difficult.

2. Brushing Stroke Detection: A method to detect brushing strokes using acoustics was presented in [24], with an average error rate of 10.3%. They posed the task of detecting brushing strokes from inertial sensors as an open problem. We successfully solve this open problem with less than 4.2% error.

3. Tight Time-Synchronization We observe that as brushing strokes and transitions usually last < 300 milliseconds, the sensor data and video (that provides a way to obtain precise labels) needs to be tightly time synchronized. Even though prior works [24, 69] used video to obtain surface labels, ours is the first work to illustrate the label alignment challenge and presents an algorithm to achieve tight time synchronization.

4. Within-Person Variability: Although between-person variability in brushing patterns have been reported previously in dentistry [70, 71], ours is the first work to report wide within-person between-episode variability, highlighting that the usual approach of single event observation from each participant may not suffice to analyze prevalent brushing patterns.

5. Between-Person Model Generalizability: We quantify between-person variability in brushing patterns from video data and analyze the challenges it poses in achieving between-person generalizability of machine learning models for brushing surface detection.

6. Challenges for Personalized Models: Although personalized models require person-specific training, they usually perform better than general models. We show that wide within-person between-episode variability impacts the performance of even personalized models for brushing surface detection.

7. Brushing Duration Estimation: Ours is the first work to present estimation of the total duration of brushing on each surface in a brushing episode, and report a median absolute error of less than 5%.

5.3 Data Description, Labeling, and Key Findings from Labeled Data

In this chapter, we explore the essential elements of data preparation, labeling methods, and the extraction of valuable insights from labeled datasets. The chapter initiates by highlighting the significance of data description in comprehending the characteristics and attributes of a dataset. Furthermore, it delves into the fundamental process of data labeling, which plays a vital role in supervised machine learning tasks. Data labeling involves assigning appropriate tags, categories, or annotations to data instances, simplifying the learning of patterns and accurate predictions by machine learning algorithms. Lastly, the chapter concentrates on extracting key findings from labeled datasets. Then we employ a variety of data analysis techniques to reveal hidden insights and practical knowledge.

5.3.1 Dataset Selection

A wrist-worn inertial sensor data set consisting of labels of start/end of brushing and flossing episodes used in our mORAL [29] study is available publicly. This study recruited participants willing to brush at least twice — once with a manual toothbrush and once with a SmartBrush and floss at least once a day. Each participant wore a MotionSense wristband on each wrist during waking hours for seven days that included a 3-axis accelerometer and a 3-axis gyroscope sampled at 16 and 32 Hz, respectively. A study provided smartphone connected via Bluetooth technology continuously timestamped and logged incoming sensor data. Besides, participants used the phone's front camera to video record themselves (in their homes) during brushing, flossing and/or oral rinsing. The mORAL dataset currently consists of data from 30 participants (15 males, 15 females; mean age 28.5±10.6 years, 2 left handed) who have contributed 197 brushing episodes with a manual toothbrush.

In the public dataset, the start and end times of brushing episodes are annotated from self-recorded videos. But, the original annotations in the mORAL dataset [1] are insufficient for our modeling because it does not include any teeth surface annotations within a brushing episode. We used the original videos from this study to label precise times for when each teeth surface (i.e., groups of teeth portions) was being brushed, including marking of transitions among the surfaces. See Section 5.3.3 for details of surface definitions proposed.

5.3.2 Dataset Curation

Out of 197 brushing episodes, videos for some episodes were not usable for stroke-level annotation of surface transitions. First, some participants moved sideways, getting outside the camera range, during brushing. Second, some participants leaned forward to spit out the excess foam and did not revert to an upright posture. Third, some participants leaned the phone against the back wall or against the sidewall. Because the camera was tilted, it was pointing diagonally at the mouth, with their hand blocking a clear view of their mouth. Therefore, it was not possible to unambiguously determine from the video which surfaces participants were actually brushing.

For the above reasons, 83 brushing sessions had to be excluded from this modeling work. We annotated the remaining 114 episodes from 19 participants. For comparison, prior works on analyzing brushing patterns via video used 96 [70] and 101 brushing episodes [71] and prior works on the detection of brushing surfaces used data from 12 [24] and 10 participants [69].

5.3.3 Organizing and Naming of Teeth Surfaces for Labeling

Various works [72, 26, 42, 37, 44, 73] organize teeth surfaces between 4 and 16 categories. For teaching brushing, the Bass technique [21, 74, 24, 69, 75] uses 16 surfaces. When observing brushing habits from self-recorded videos [70, 71], surfaces are grouped into fewer broad categories due to ambiguity and frequent transitions among teeth sur-

[1] https://mhealth.md2k.org/resources/datasets.html#mORAL

faces. We adopt a similar hierarchical organization to obtain nine categories that is suited to sensor-based detection.

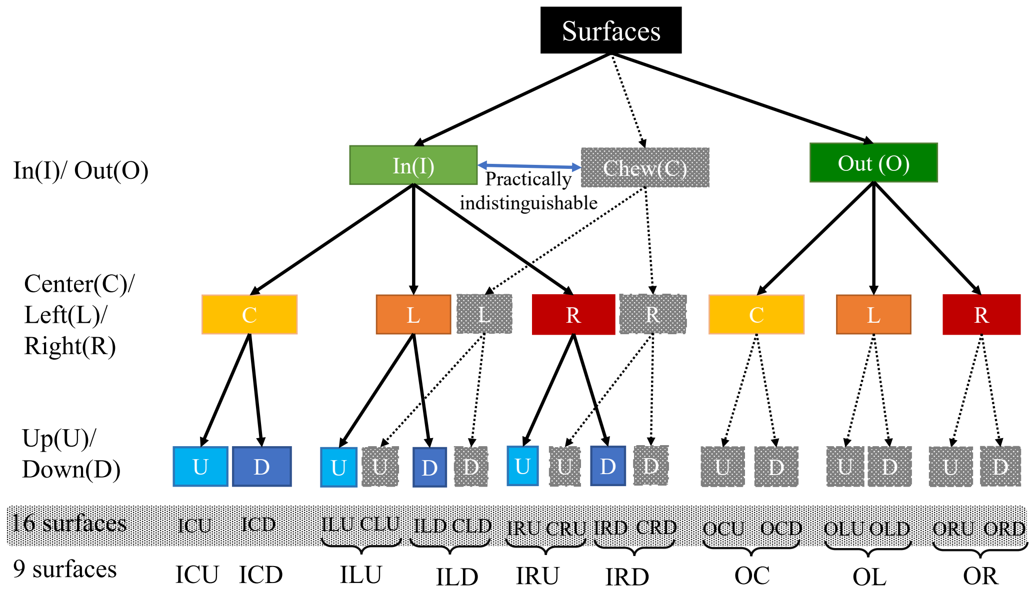


Fig. 5.1: Three layer teeth surface naming

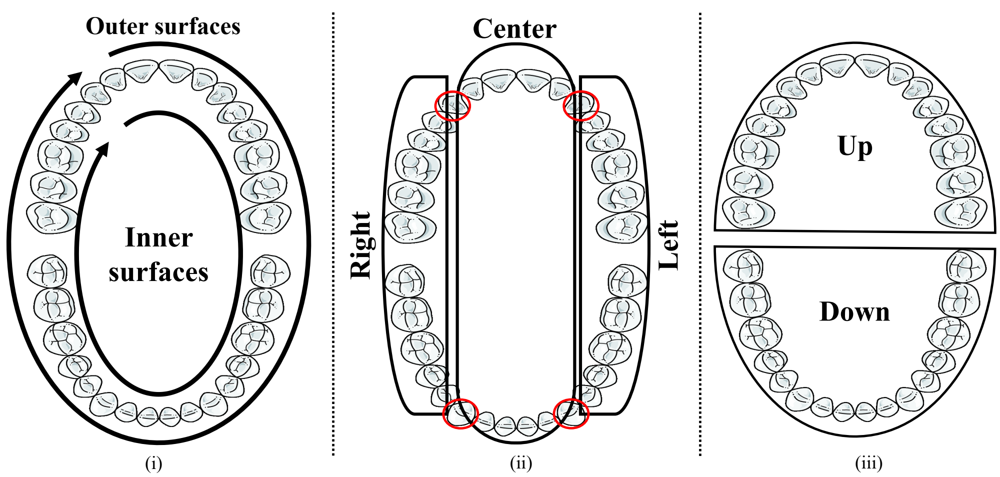


Fig. 5.2: Three broader categories of brushing surfaces

We organize teeth surfaces into three layers, as shown in Figure 5.1.

Layer 1: *In*.

The inner (tongue facing, i.e., lingual) surfaces of all the teeth and the occlusal (chewing) surfaces of the posterior teeth (premolars and molars) are labelled as the 'In' surface (see the inner arrow in Figure 5.2(i)).

Layer 1: *Out*.

The outer surfaces of the teeth (abutting lips and insides of cheeks, i.e., vestibular) are labelled as the 'Out' surface (see the outer arrow in Figure 5.2(i)).

Layer 2: *Center*.

The 'Center' surface encompasses all the anterior incisor teeth (as shown in Figure 5.2(ii)).

Layer 2: *Left/Right*.

The posterior region incorporating the premolars/molars on the left and right sides are labelled as 'Left' or 'Right', respectively (see Figure 5.2(ii)).

Layer 2: *Undecidable*.

Finally, we place the canines (red) in the 'Undecidable' class. Depending on the brushing pattern, these teeth are dynamically assigned to one of the center/left/right surfaces rather than being apriori assigned all the time. In Figure 5.2(ii), red colored teeth are considered as 'Undecidable'.

Layer 3: *Up/Down*.

We define surfaces of teeth in the upper jaw as the 'Up' surface and surfaces from the lower jaw as the 'Down' surface (as shown in Figure 5.2(iii)).

Of the 16 surfaces used in the Bass technique, we are unable to disambiguate brushing on chewing surface from inner surfaces due to frequent overlap, resulting in merging of 8 surfaces (chewing and inner) into 4 (inner) surfaces. Additionally, when brushing on the outer surfaces, we are unable to disambiguate brushing on upper and lower surfaces, due to frequent switching and overlap, resulting in merging of 6 (outer up and down) surfaces into 3 (outer) surfaces. Therefore, we end up with nine surface categories.

For naming of these nine surfaces, as we descend from Layer 1 to Layer 2 (in Figure 5.1), and then to the leaf nodes in Layer 3, we concatenate the respective categories in each layer to derive the name for each leaf surface. For example, if we traverse the nodes in the order In->Center->Up from Layer 1 to Layer 3, we get the *In-Center-Up (ICU)* surface. Names of the other eight surfaces are: In-left-up (ILU), In-right-up (IRU), In-center-down (ICD), In-left-down (ILD), In-right-down (IRD), Out-center (OC), Out-left (OL), and Out-right (OR).

5.3.4 Determining the Timings of Teeth Surface Being Brushed On and Transitions from Video

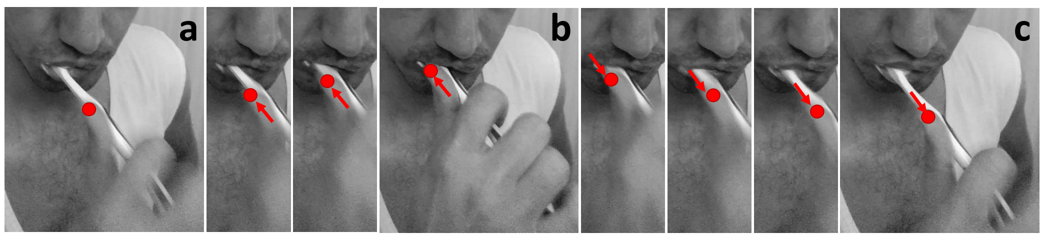


Fig. 5.3: Frame-by-frame annotation of left-right-left brushing stroke. Frames (a) in both stroke types mark the start of the stroke, Frames (b) mark the end of half stroke, and Frames (c) mark the end of the full stroke.

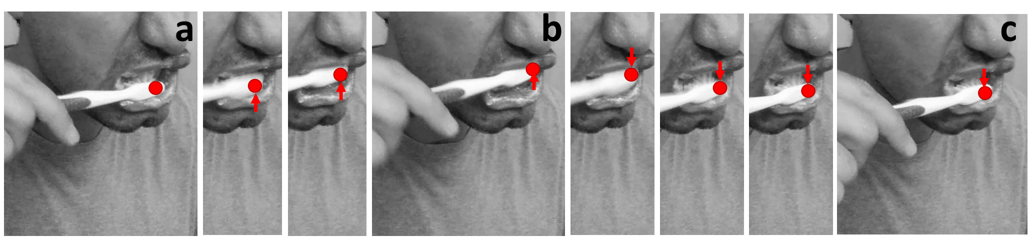


Fig. 5.4: Frame-by-frame annotation of up-down-up brushing stroke. Frames (a) in both stroke types mark the start of the stroke, Frames (b) mark the end of half stroke, and Frames (c) mark the end of the full stroke.

We analyze the video recordings to annotate the start/end times of each teeth surface being brushed on. We precisely mark the transition among surfaces so that data labeled for a surface is not contaminated by any transition data, resulting in unambiguous and clean labels for model training and testing. This is arduous and time-consuming because the duration of brushing on any surface is quite short (less than 5 seconds) and transitions are rapid (lasting few hundred milliseconds) and frequent (tens of transitions in a brushing episode).

Since achieving precision at such granularity is harder for human eyes, we used ELAN [76], a freely available software for assistance in labeling the surface and transition times. We developed the following coding definitions for this labeling to correspond to hierarchical naming of surfaces. We annotated each of the three layers in our naming hierarchy — Pass-I (Inner and Outer), Pass-II (Center, Left, and Right), and Pass-III (Up and Down). For each of these, we decide the start and end time as described below. Figures 5.3 and 5.4 show two frame-by-frame examples.

Begin time: We assign the start time to the moment whenever a participant touches and starts to go back and forth or up and down with the brush in a periodic motion in any one of the In/Out/Center/Left/Right/Up/Down surfaces for the first time or every time after a transition from the previous surface.

End time: We assign end time to the moment whenever the participant stops the periodic back and forth or up and down motion with the brush at the current surface and begins to leave the surface by changing the motion.

To distinguish a surface from transition, we annotate a surface only if it receives at least three brushing strokes.

Switching interval: We use the following criteria for declaring a transition.

1. When the participant, in a bid to move to the next surface, starts rotating the brush holding wrist to till the rotation stops, and the wrist is in a position from where it can start brushing the next surface.

2. When the brush holding wrist enters the junction of any two surfaces to when it leaves.

3. When the wrist holding the brush discontinues the periodic back and forth or up and down motion and slowly takes either a single forward or backward motion.

4. When the wrist holding the brush suddenly stops brushing the current surface.

Two independent coders labeled all videos to annotate the start and end time of brushing on each surface. The duration of surface transitions is usually < 300 milliseconds, and our goal was to annotate the timings at the stroke-level precision. Therefore, instead of using 0.96 seconds [70], we consider annotations from two coders to match only if the discrepancy for any surface is less than a half-stroke, i.e., 150 milliseconds. We observe 342 discrepancies out of 10,230 surface annotations (3.34%). Discrepancies were resolved via joint viewing of the segment in doubt, and a consensus was reached regarding the labeling of the event in consideration.

5.4 Key Observations from Labeled Data

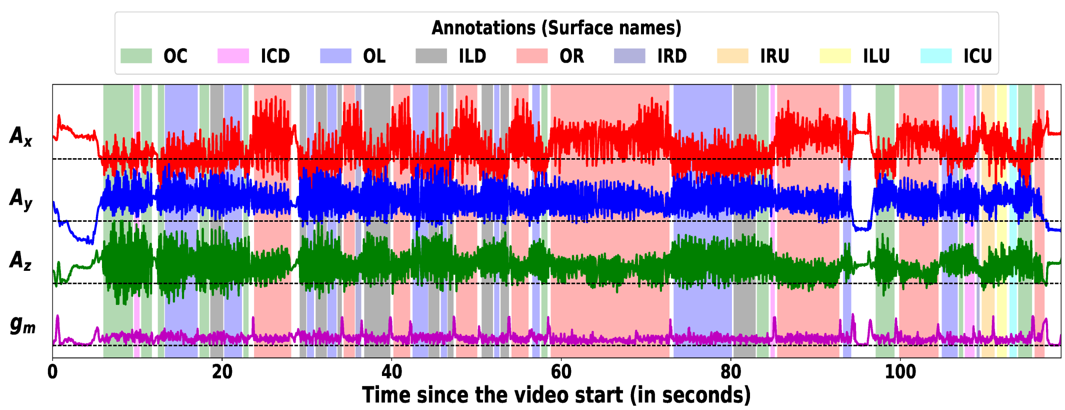


Fig. 5.5: Three axes of accelerometer and gyroscope magnitude during brushing at different annotated surfaces.

Brushing patterns from videos have been analyzed in dentistry during habitual brushing [70] and best-effort brushing [71] to assess shortcomings and to find ways to further improve brushing habits. As these works invited participants to the study site and recorded one episode from each participant, their observations largely focused on between-person variability. In contrast, we ask participants to video-record themselves in their homes without any explicit instructions, providing us repeated measurements from the same person in their natural environment. This data allows us to analyze within-person and within-episode variability during habitual brushing.

5.4.1 Between-Surface Variability in the Time Spent on Brushing Different Surfaces

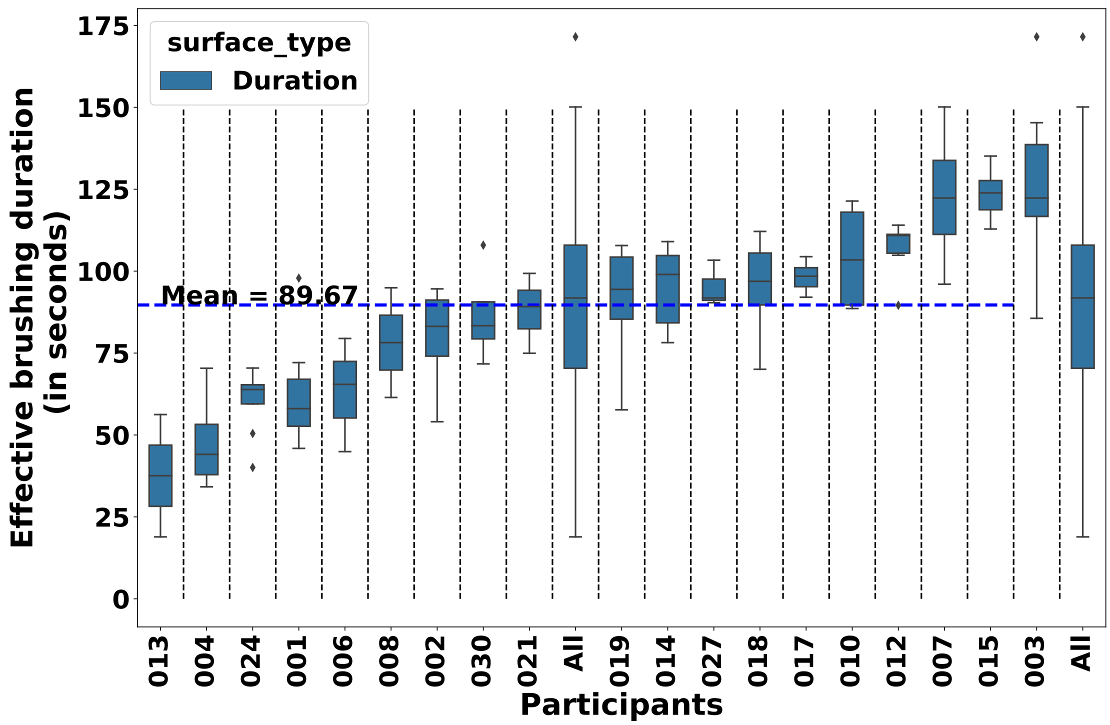


Fig. 5.6: Difference in effective brushing duration

Table 5.1: Average duration (in seconds) of brushing on each of the nine surfaces

	Surfaces									EBD
	ICU	**ICD**	**IRU**	**ILU**	**ILD**	**IRD**	**OC**	**OL**	**OR**	
DM (± SD)	2.50 (±0.9)	3 (±1.1)	5.8 (±1.7)	5.9 (±1.4)	9 (±2.2)	9.2 (±1.9)	13.3 (±2.3)	20 (±2.5)	23 (±2.7)	91.7 (±3.4)

Table 5.1 shows the mean and standard deviation duration of brushing on each of the nine surfaces. Figures 5.6, 5.7, 5.8, 5.9 show detailed distribution for each episode from each participant. Similar to [70], we find that the duration of brushing on left and right sides (across both inner and outer surfaces) are similar. But, we find that the total effective duration of brushing in our dataset is significantly lower at 92 seconds, compared with 155

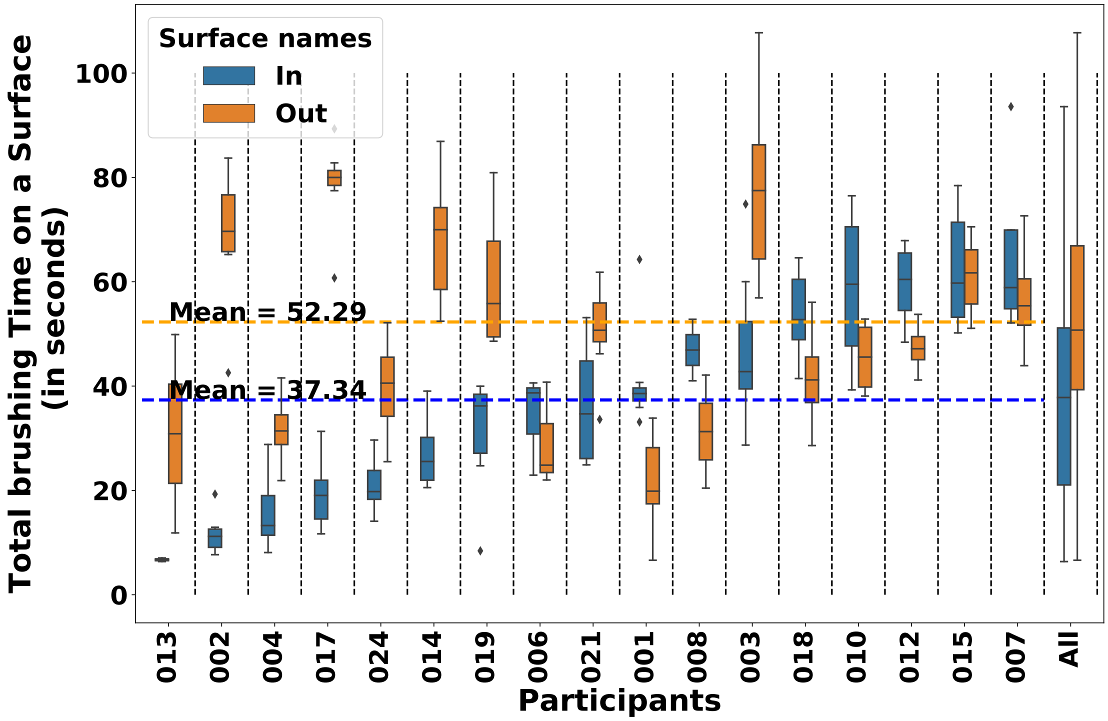


Fig. 5.7: Difference in total duration of brushing on *In* and *Out* surfaces.

seconds in [70] and 207 seconds in [71]. We make several new observations regarding between-surface variability.

To test statistical significance, we take pairwise percentage difference in duration between two brushed surfaces, i.e., ratio of brushed surfaces for all the brushing episodes across all the participants. We want to find a value a such that mean of the percentage difference is significantly greater than the value a. Without loss of generality, we assume the mean of percentage differences is positive (otherwise we switch two duration lists). To find the value of a, we perform a left tailed t-test where the alternative hypothesis is mean $\mu < a$. We want to find the maximum a such that using a t-test we can reject the null hypothesis that the mean of the list is a, i.e., $H_0 : \mu = a$.

Participants spend 40% more time brushing their outer (i.e., buccal or labial) teeth surfaces than their inner (i.e., lingual and occlusal) teeth surfaces across both upper and lower jaw. (p-value <0.008)

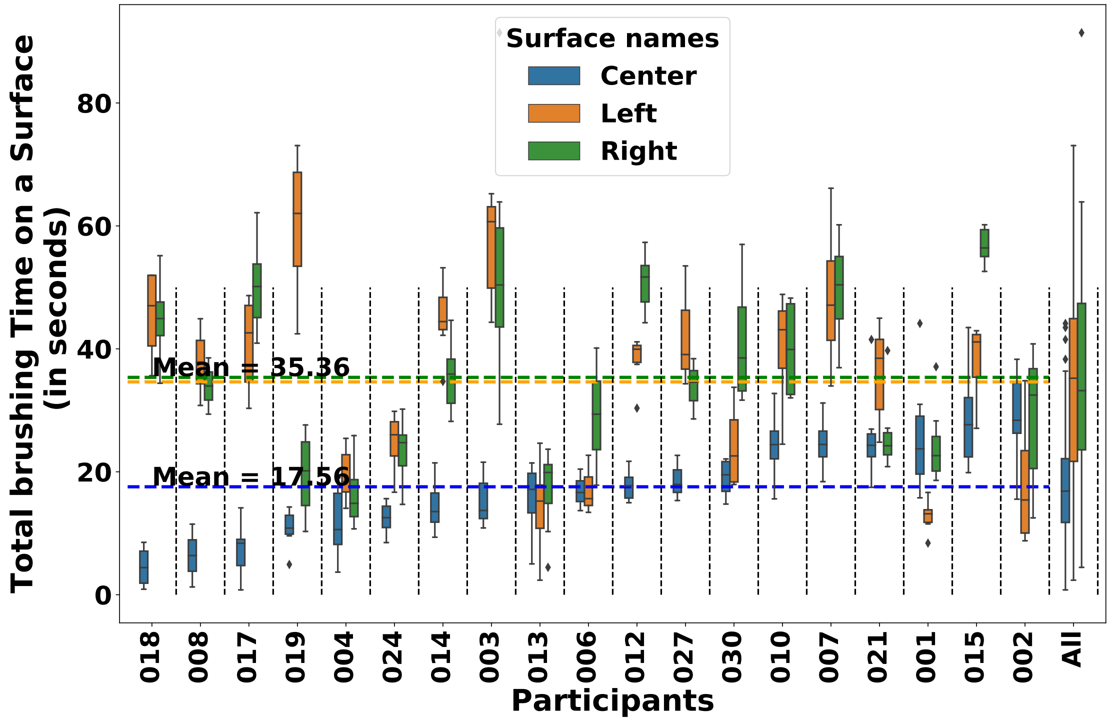


Fig. 5.8: Difference in total duration of brushing on *Center*, *Left*, and *Right* surfaces.

Participants spend 75% less time brushing their center (i.e., anterior) teeth surfaces as compared to their left or right surfaces (i.e., posterior). (p-value <0.009)

When brushing on inner (i.e., lingual) teeth surfaces, participants spend 75% more time in brushing down surfaces vs. up surfaces. (p-value <0.009)

The duration of the most brushed surface within an episode is 11 to 19 times the duration of the least brushed surface. But, the least- and most-brushed surfaces are not the same in all episodes. (p-value <0.0003)

5.4.2 Between-Person Variability in the Brushing Time on Each Surface

Between-person variability in brushing patterns have been reported previously [70, 71]. As Figures 5.6, 5.7, 5.8, 5.9 and 5.10 show, we also observe substantial within-person between-episode variability in amount of time spend on brushing surfaces. Our goal here is to quantify these differences to assess the feasibility of developing a common machine learning (ML) model that can work for all users.

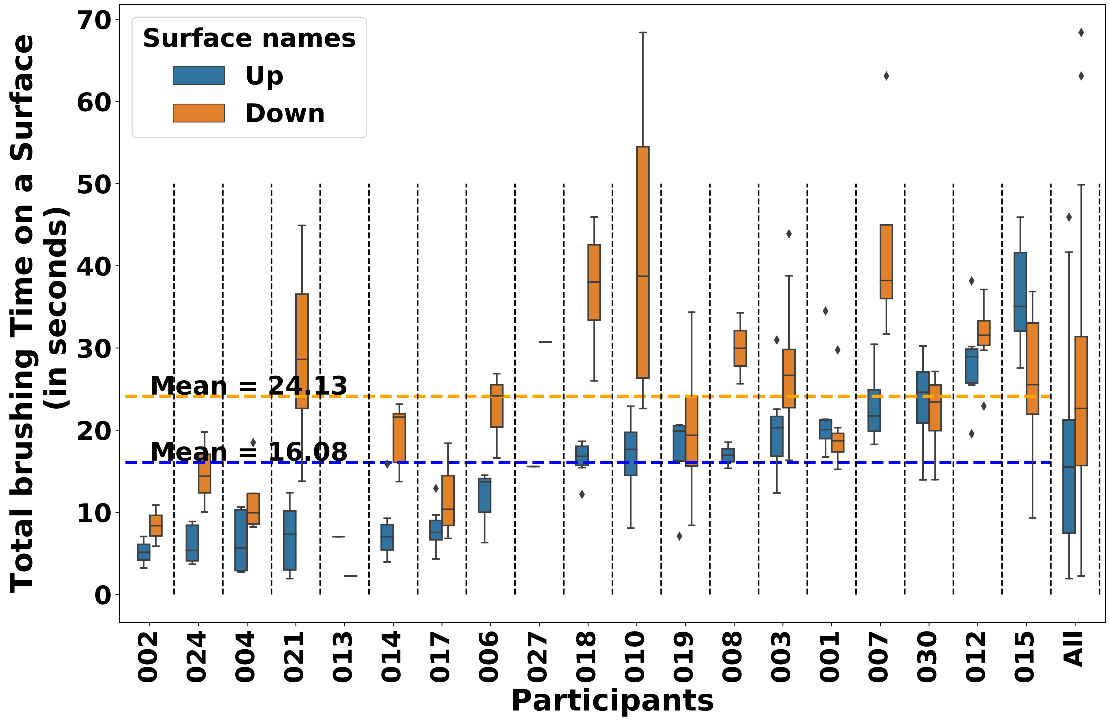


Fig. 5.9: Difference in total duration of brushing on *Up* and *Down* surfaces.

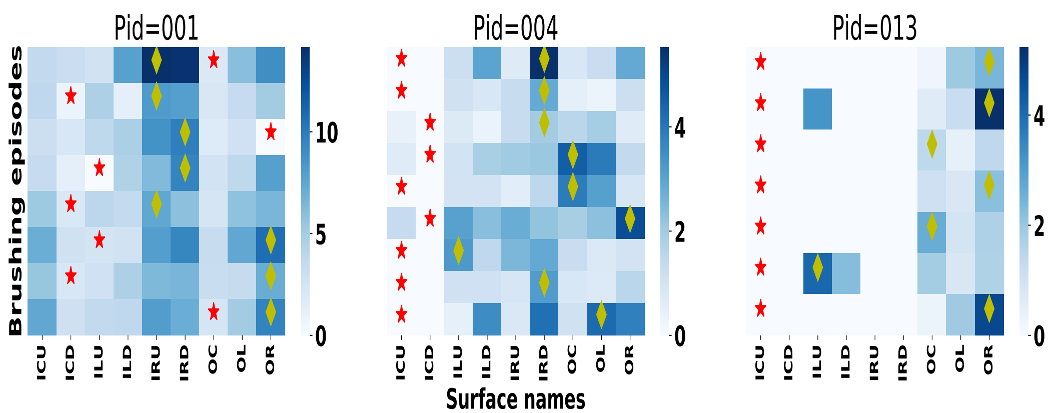


Fig. 5.10: Illustration of difference in the amount of time participants spend on each surface across different brushing episodes. Data from three representative participants are shown here. The least brushed surface and most-brushed surface within each episode (i.e., each row) are marked with a star and diamond, respectively. Numbers in legends represents surface brushing duration in seconds and darker colors represent longer duration.

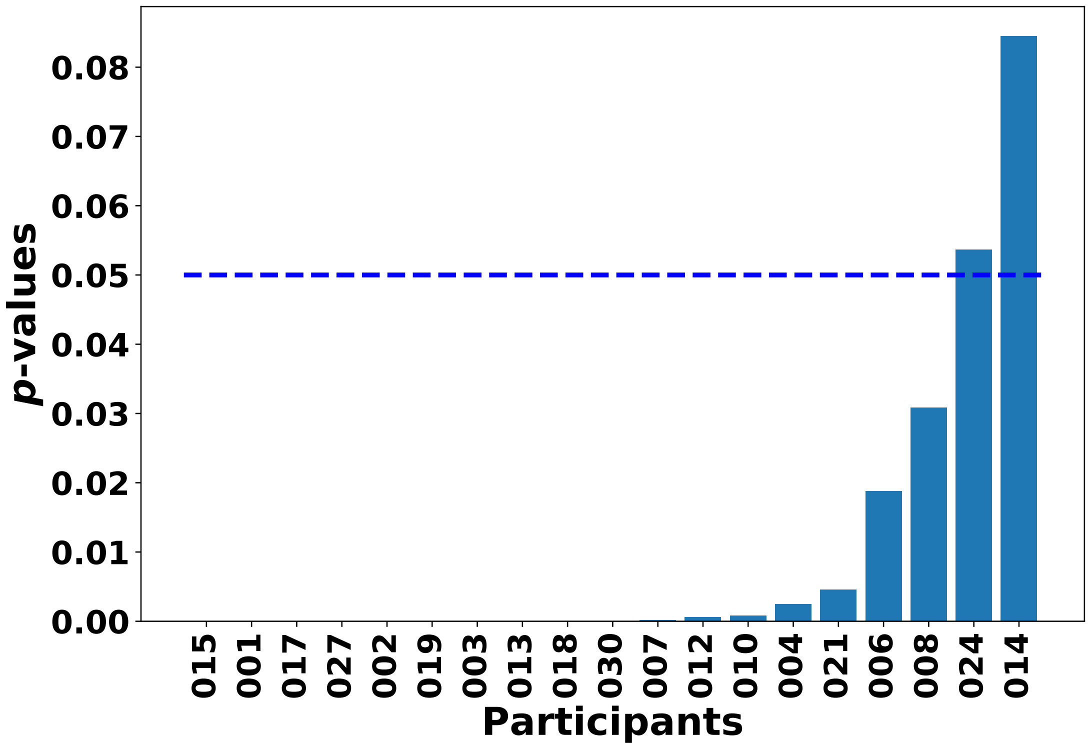


Fig. 5.11: Person-to-population similarity in the amount of time spent in each surface

Similarity of Persons with the Population Profile in Time Spent on Each Surface:

First, we quantify how many participants have a brushing duration profile that matches the population average. For that, we represent each brushing episode as a duration vector of all the brushing surfaces, i.e., a nine value vector. This way, we form a list of vectors with one participant's data and combine list of vectors from the rest of the participants to create a population profile. From these two lists of vectors, to find whether a participant's data approximates the population profile, we perform the χ^2-test (Chi-squared test). We repeat this process for all the participants, and the resulting p-values are shown in Figure 5.11. We see that only 2 out of 19 participants share profiles similar to the population one. Therefore, population-profile is not representative for most individuals.

Person-to-Person Similarity in Time Spent on Each Surface:

Next, our goal is to see if there are clusters of participants sharing a similar profile amongst themselves. We test pairwise independence for all possible pairs of participants.

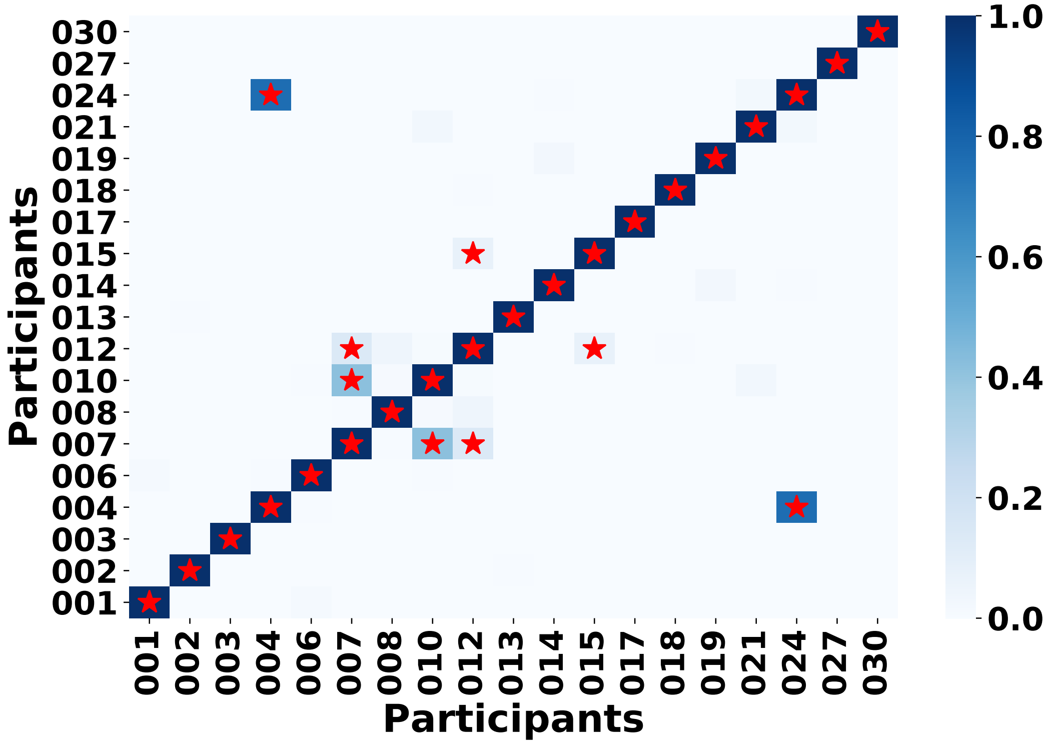


Fig. 5.12: Person-to-person similarity in the amount of time spent in each surface. Stars show statistical significance.

We form two list of vectors from two participants following a similar method mentioned in Section 5.4.2 and perform the χ^2 hypothesis test to find if they are similar to each other. We repeat this process for all possible pairs of participants and present the test results as a heatmap in Figure 5.12. Each cell (p_i, p_j) in the figure shows the p-value of the test for p_i and p_j. Only 4 out of 136 pairs show similarities in profile.

5.4.3 Within-Episode Patterns of Transitions Among Surfaces

Prior work [70] has observed preference among participants for frequent transitions among surfaces with an average of 45 transitions in brushing episodes lasting 155 seconds, on average. Figure 5.13 shows the distribution of average number of surface transitions and Figure 5.14 shows the average time spent brushing a surface between transitions in our dataset. We observe significant variability in both the frequency of transitions and the time between transitions both between-person and within-person.

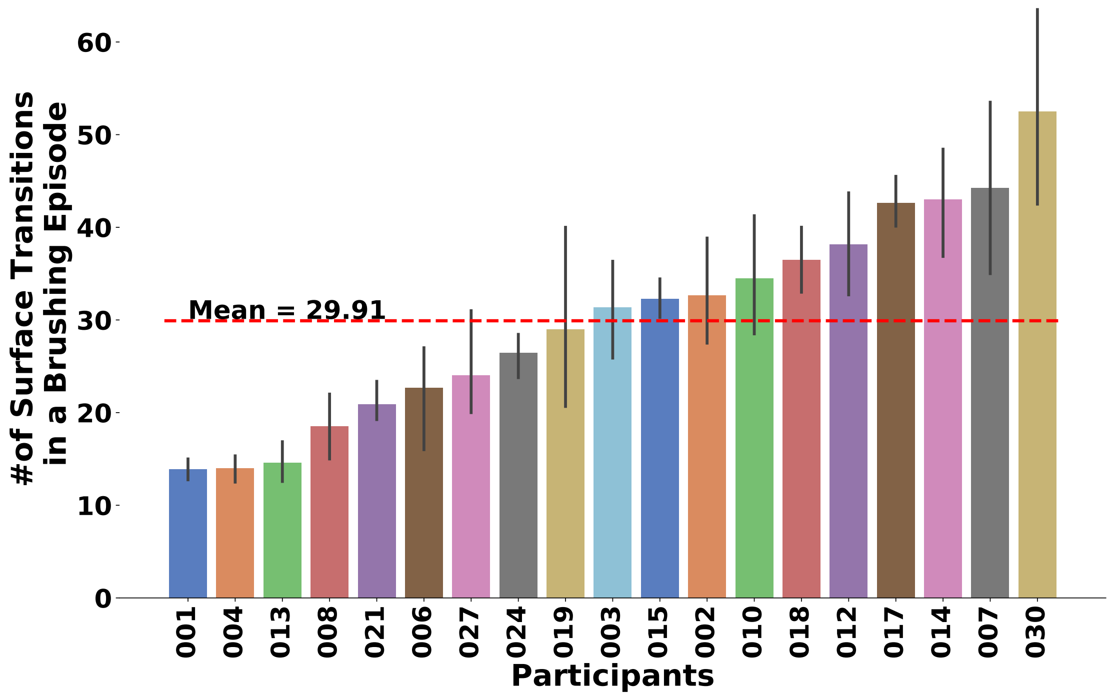


Fig. 5.13: Number of surface transitions per brushing episode.

5.5 Overview of the mTeeth Model

Figure 5.15 presents an overview of all the steps in the *mTeeth* model. The input to the model are the inertial sensor data (accelerometer and gyroscope) and the start/end of brushing episodes from a brushing detection model such as mORAL [29]. We first define an *anchor event* (detectable from sensor data) that can be used to segment the time series data cleanly so each segment can receive the unambiguous label of one surface (in Section 5.6). Subsequently, we develop a method to tightly synchronize sensor data with video so that labels of surface transitions correspond to the sensor data segments at millisecond precision (in Section 5.7). Then, we identify and compute event-based features and select distinctive features for each surface (in Section 5.8). Finally, we train a Dynamic Bayesian Ensemble model to assign each data segment to the most likely surface (in Section 5.9). This generates a sequence of brushing surfaces in each brushing episodes, with its duration and the number of strokes in it.

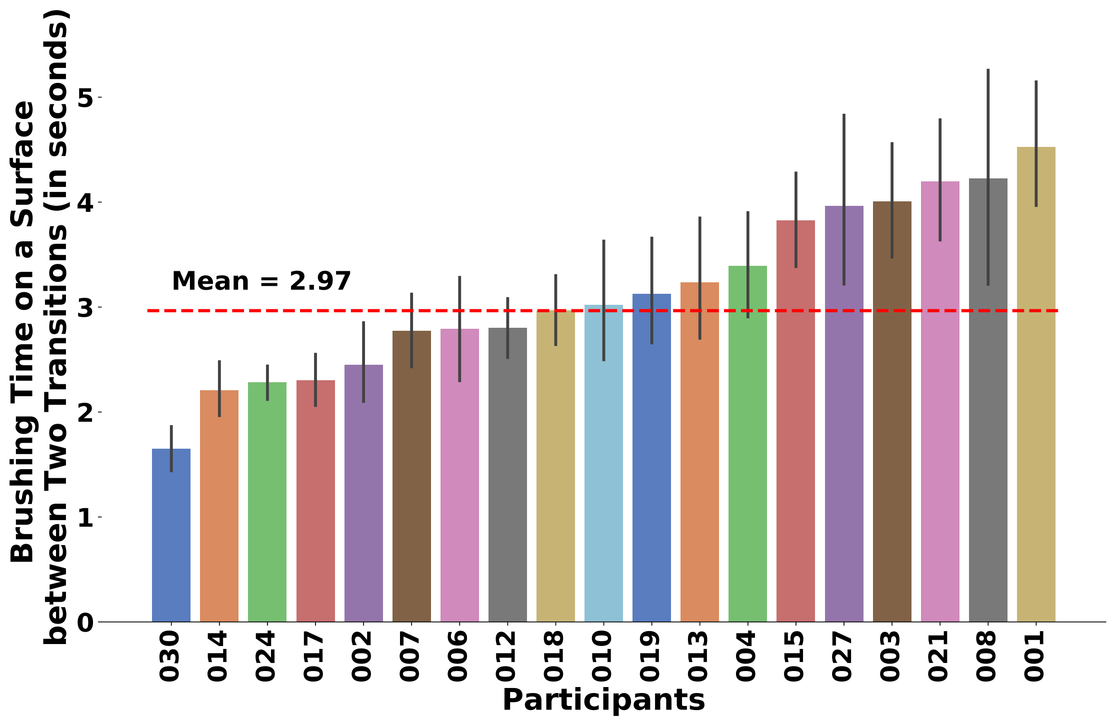


Fig. 5.14: Brushing time on a surface between two transitions.

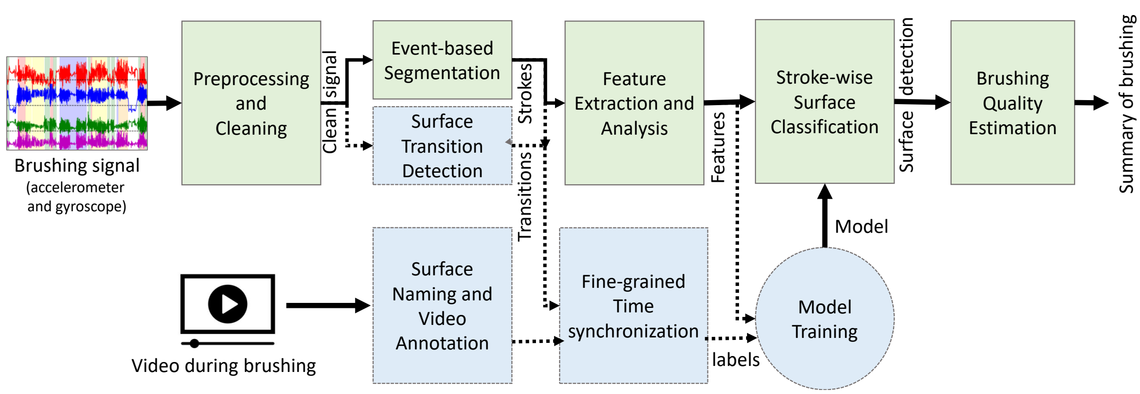


Fig. 5.15: The pipeline of data processing stages for training and testing the *mTeeth* model.

5.6 Defining and Detecting Anchor Events for Time-Series Segmentation

There is wide between-person variability in how people brush, including the pattern of back and forth or up and down motion of the brush, time spent in each surface during brushing, and transition sequence among surfaces. In our labeled data, we observe that the time spent on a brushing surface varies from a few milliseconds to as long as 10 seconds.

This poses a significant challenge for finding an optimal window of sensor data that can be treated as a single unit of assessment from which features can be extracted to train a machine learning model. The traditional approach of sliding or fixed time-based windowing is unlikely to work. If we choose a window size of few milliseconds to deal with the short duration in some surfaces, we may end up with insufficient data to find distinguishable feature(s). If the window size is too large, several transitions occurring within it may go undetected, resulting in missed transitions and mixing of surfaces in one window, creating both ambiguity and mismatch in label assignment. Therefore, we seek a dynamic-length event-based approach to data segmentation.

For an event-based approach to succeed, we need anchor events such that the likelihood of not detecting this event is low, the event should be efficiently detectable, and the event should cleanly isolate data segment belonging to different surfaces. When developing a model to detect brushing [29], flossing, eating [15], drinking [67], or smoking [17, 16], a hand-to-mouth gesture works as an anchor event. The events of hand reaching the mouth and hand coming back from the mouth isolates segments of sensor data that can be treated as a candidate for each of these hand-to-mouth gesture events and can be tested by the respective machine learning models. But, the hand-to-mouth gesture only occurs at the start and end of a brushing event and hence can't be used to segment the sensor data within a brushing event to distinguish among various teeth surfaces.

For our purpose, we need to find an anchor event that clearly demarcates when brushing surface changes, this event occurs during most surface transitions, and the event is efficiently detectable from sensor data. As our goal is to find the start and end times of brushing on each surface, the transition from one surface to another initially appears to be an obvious choice for an anchor event. But, the transition itself is so short-lived that it is improbable to detect some of the transitions from sensor data. Moreover, some of these transitions are difficult even to annotate from the video. Thus, accurate detection of all transitions is quite challenging, and failure to do so results in mixing of data from two or more surfaces. Therefore, transitions do not qualify as the anchor events.

We have observed that there is one specific activity that is both easily detectable in sensor signals and commonly performed across all surface transitions: brushing. During brushing, individuals typically engage in a back-and-forth or up-and-down periodic motion with the toothbrush. This distinctive movement pattern, referred to as a "brushing stroke," serves as our anchor event for surface identification.

Among the various types of brushing strokes described in the literature, such as circular strokes [70], we have noticed that brushing strokes predominantly follow an up-down-up or back-forth-back motion. These two primary periodic movements are consistently observed during brushing activities. Consequently, when at least one brushing stroke is performed on a surface, its trace tends to be retained in the sensor data, minimizing the likelihood of missing a stroke and reducing surface identification errors.

Additionally, we have observed that no brushing strokes occur between surface transitions. This absence of brushing strokes between transitions ensures that there is no mixing of signals from different surfaces within any given segment of sensor data used for surface identification by a machine learning model. This further enhances the accuracy of surface identification by preventing the confounding influence of brushing strokes on adjacent surfaces.

By leveraging the distinctive brushing stroke pattern and its absence between transitions, our approach effectively addresses the challenges associated with surface identification using sensor data. It provides a reliable and accurate means of differentiating surfaces based on the presence or absence of brushing strokes, thereby enhancing the overall performance and robustness of the surface identification algorithm.

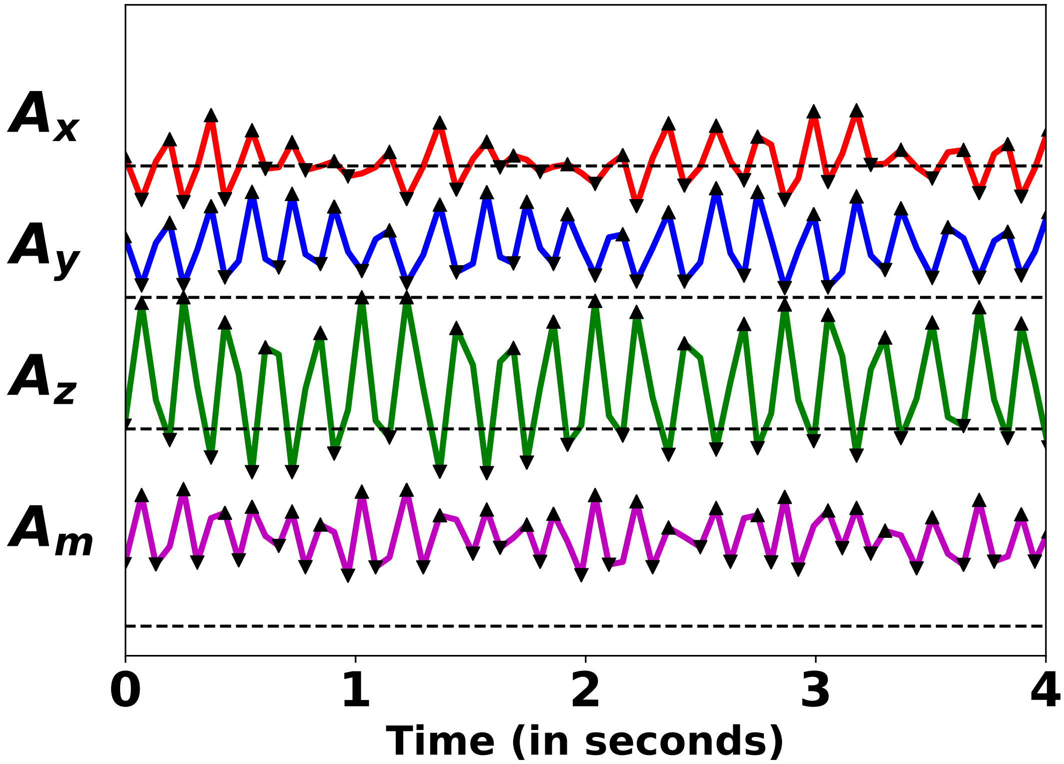


Fig. 5.16: Brushing Strokes.

5.6.1 Brushing Stroke Detection

To efficiently detect brushing stroke from sensor data, we identify the signature of periodic up-and-down or back-and-forth movement in the wrist-worn accelerometer signal in the form of a peak-valley pair.

Figure 5.16 shows plots of three axes accelerometer signal and its magnitude during one such surface brushing. In the signal time series, we define peak as the point in each cycle where the signal is at its maximum, whereas a valley as where the signal is at its minimum. We mark those peaks and valleys of a signal with black up-pointing and down-pointing triangles, respectively. Let $\mathcal{P} = \{p_1, p_2, ..., p_n\}$ be the peaks, and $\mathcal{V} = \{v_1, v_2, ..., v_n\}$ be the valleys. Once we carefully detect all these peaks and valleys using a peak-valley detection algorithm, we define brushing stroke as a cycle of valley-peak-valley combination, i.e., an i^{th} brushing stroke is $S_i = \langle v_i, p_i, v_{i+1} \rangle$.

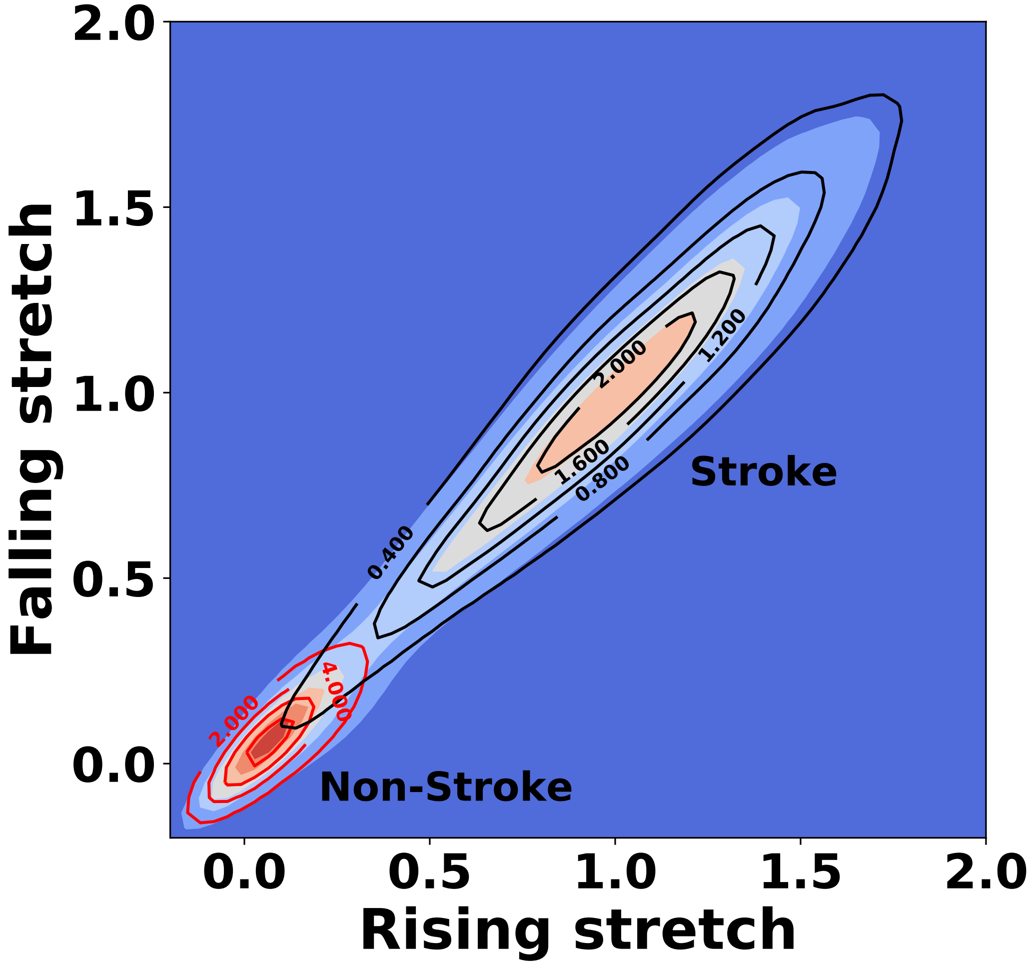


Fig. 5.17: 2D Gaussian Kernel density function of rising stretch and falling stretch.

Now, in Figure 5.16, we observe series of peaks and valleys in all the four signals, but most of them are temporally unaligned across the signals. Since we get four sequences of peak-valley cycles or therefore strokes, out of these four signals, we need to select the one that will represent the start and end times of the brushing strokes optimally. We note that even though the magnitude contains information from all the three axes, it is not a suitable choice due to lack of synchronized alignment across the three axes.

We first define the stretch of a stroke as the difference between the amplitude of its peak

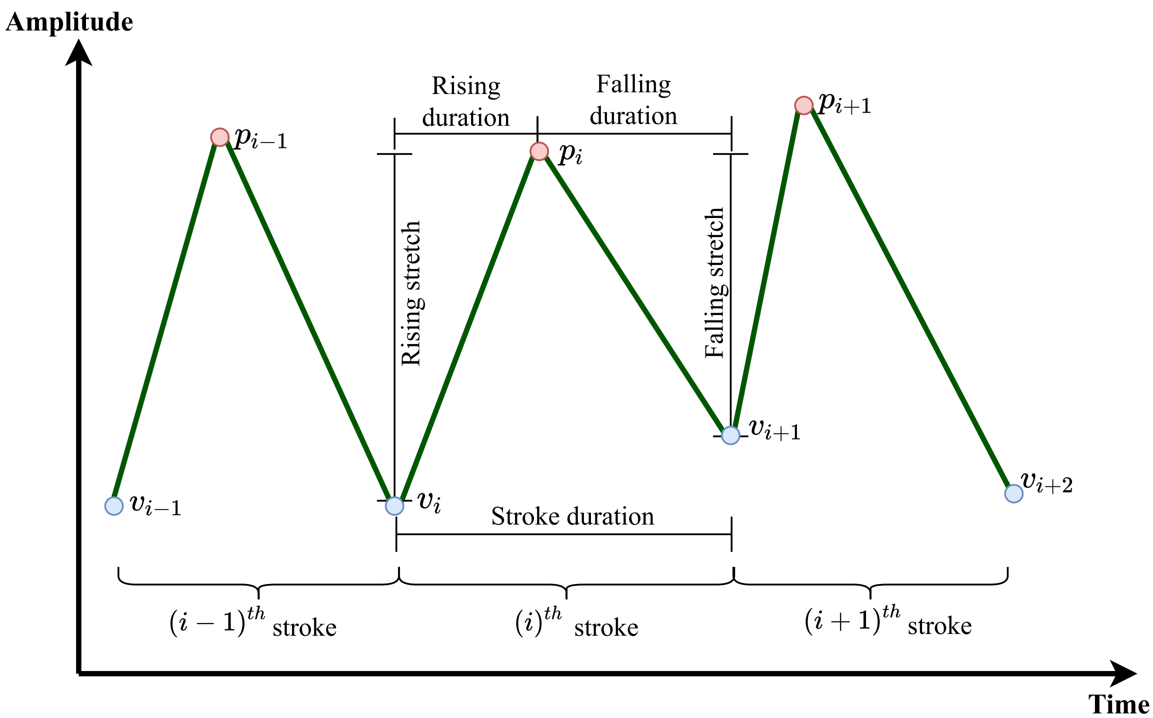


Fig. 5.18: Stroke-wise computed features from the accelerometer signal.

and valley. If the stretch of a stroke along any axis is low, that corresponds to having least wrist movement along that axis at that moment. Conversely, if the stretch of a stroke along an axis is high, that signifies a likely wrist movement along that axis during a brushing stroke, making it the dominant axis for this stroke. Hence, we select a brushing stroke along a particular axis that has the maximum stretch. To brush different surfaces, orientation of the wrist changes, so does the movement of the toothbrush along with it. Following the movement, the acceleration of the wrist along a particular axis changes the most. In addition, we observe that the dominant axis remains unchanged throughout brushing on a single surface. When the user switches to the next brushing surface, depending on the type of surface, the dominant axis may either change or continue to be the same.

To distinguish brushing from other activities (e.g., walking) which also involves periodic wrist movement, we define two thresholds, $\mathcal{T}_{dur}$ and $\mathcal{T}_{stretch}$ such that for any $\langle v_i, p_i, v_{i+1} \rangle$ peak-valley cycle to be a brushing stroke, the time difference between v_{i+1} and v_i can be at most $\mathcal{T}_{dur}$ and the stretch needs to be at least $\mathcal{T}_{stretch}$. The average duration of a stroke is $230(\pm 60)$ milliseconds, and the average stretch is $0.57(\pm 0.35)g$. We remove all peak-valley cycles that are two standard deviations away from the mean stroke duration and the mean stroke stretch. These thresholds retain all the brushing strokes in our data, i.e., achieve 100% recall.

5.7 Fine-grained Time synchronization between video and sensor data

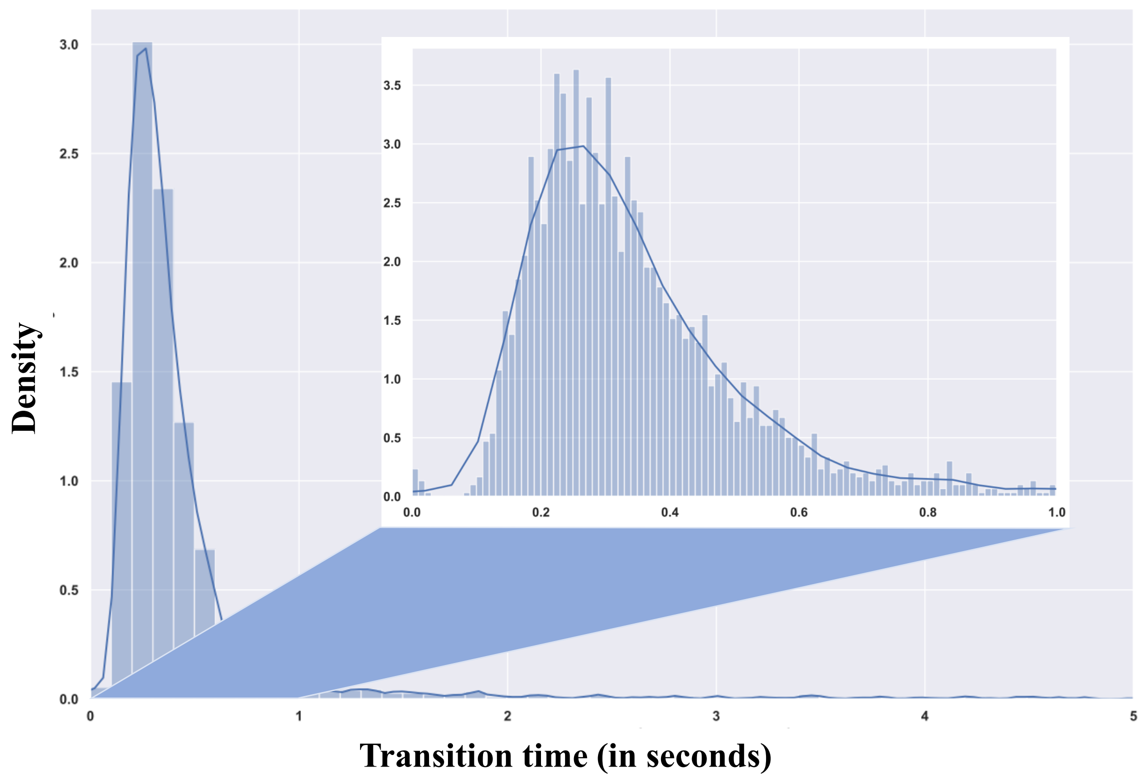


Fig. 5.19: Distribution of transition time

A key premise for our categorization of human teeth into nine surfaces is that the pattern of brushing on each of these surfaces is likely to be sufficiently unique making the corresponding sensor data distinguishable. But, to enable successful modeling for recognizing each of the nine surfaces from sensor data, a brushing episode should include at least a few seconds of brushing on each surface interspersed with milliseconds of surface switching times (Figure 5.19). However, in reality (see Figure 5.20) some users spend only milliseconds on a surface before switching to another surface. Thus, for accurate estimation of brushing surfaces from such short spans of time, precise time synchronization between the

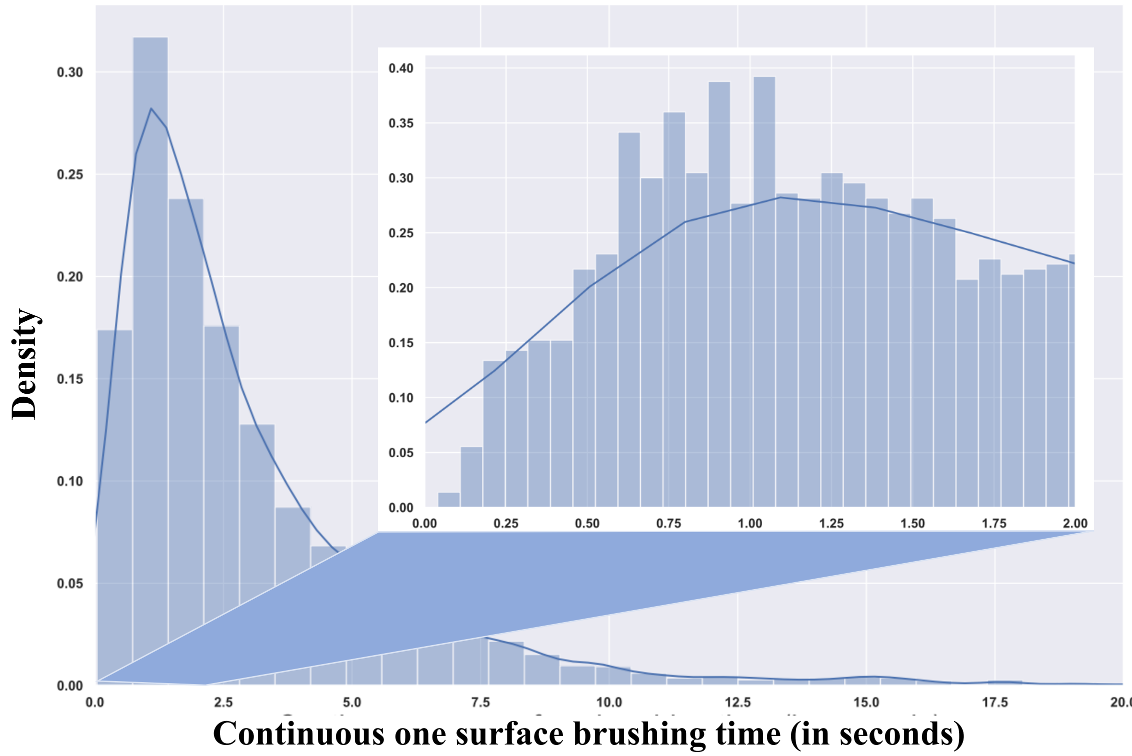


Fig. 5.20: Distribution of brushing duration on a surface

sensor and video data becomes critical. More specifically, we need to synchronize the start of a brushing session extracted from the video data with that of brushing events automatically detected from inertial sensors [29] at millisecond-level precision.

Even though it was assumed in [29] that the mORAL dataset has tight time synchronization between video and sensor data, we find that the video and sensor data have a time synchronization error of several seconds (see Figure 5.21 for an example). This may be because even though the sensor data from wrist-worn devices were streamed in real-time to the same phone recording the video, time lapse between the sensor data being received on the phone and assignment of a timestamp to them may be of the order of seconds. For the task of detecting the start and end of a brushing session that lasts 2 minutes, few seconds of time lapse may be tolerable. But, for our purposes where the brushing duration on a surface and transition times are only few milliseconds long, time synchronization errors of seconds can render the modeling process extremely challenging. If the lag between video and sensor data is not adjusted properly, part of sensor data which is actually a surface may be mistaken for a transition and vice-versa, or data from different surfaces may get mixed. The performance of a machine learning model will suffer as the quality of these labels drive the accuracy of the model.

5.7.1 Time Synchronization Problem

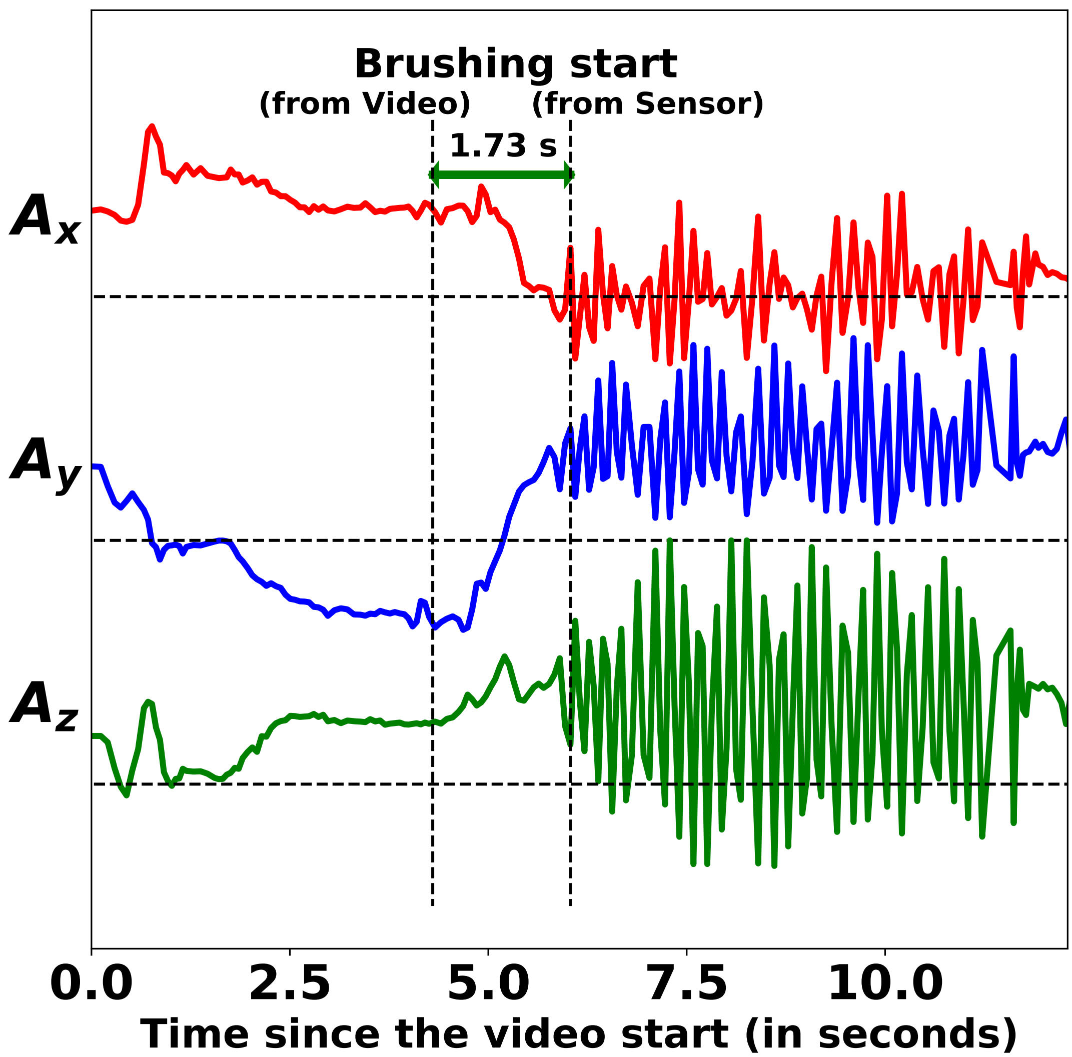


Fig. 5.21: Time difference between brushing start from video and from sensor to show the time synchronization problem.

We start by defining the time synchronization problem. Let the start time of the i^{th} brushing event, based on video time be t_i^v. From the time at which sensor data is captured

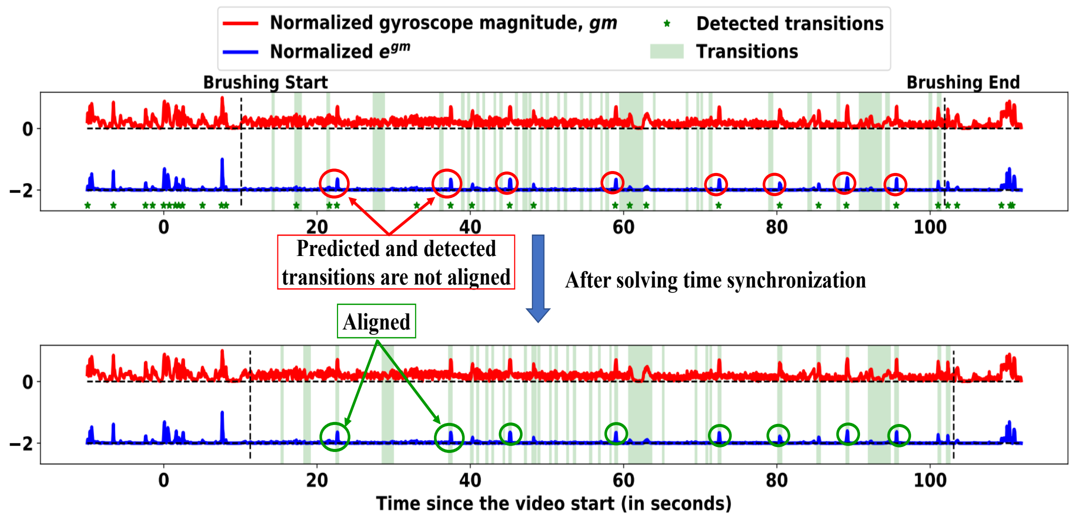


Fig. 5.22: Our approach to solve time synchronization. The top figure shows annotation before time synchronization where green start are the detected transitions from the sensor. The bottom figure shows annotation after time synchronization.

by the wrist device to when it reaches the smartphone and receives a timestamp, there is a time lag. We need to find the offset o_i such that when added to t_i^v, it corresponds to the starting time of the i^{th} brushing event from sensor data t_i^s. Therefore, $o_i = t_i^s - t_i^v$. Since each packet of sensor data contains both accelerometer and gyroscope data, offsets are the same for both.

A brushing event is composed of multiple bushing strokes. Using the brushing stroke detection method discussed in Section 6.9, we can extract the start times of all brushing strokes and use the first stroke from both video and sensor data to synchronize. Since this method is based on accurately locating the first stroke in both video and sensor data, there are at least two cases when this method may fail. First, several participants start brushing before starting the video, missing the first stroke in video. Second, when participants put toothpaste on the brush head, even one up-and-down or back-and-forth movement may create a false first brushing stroke pattern in the sensor data. Therefore, we next propose a more robust method for time synchronization.

5.7.2 Multi-point Synchronization Approach

We observe that during some transitions from one brushing surface to another, e.g., from left to right, the wrist rotation is significantly higher than that when brushing on any surface. The gyroscope can detect hand rotation, and the contrast in magnitude between brushing and some of the surface switching are clearly identifiable from the gyroscope signal in Figure 5.22. Note that for several transitions, the amount of rotation is negligible. But, if we can detect some surface transitions from the sensors, we can map these detected transitions with annotated transitions and find the offset for the synchronization. We build upon this idea to solve the time synchronization problem. Our algorithm consists of three main steps described below.

(Step 1) Rotation-based Transition Detection: During brushing, the wrist moves linearly back and forth or up and down, which are captured by accelerometers. When transitioning from one surface to another, if the wrist holding the brush changes the direction of movement, a rotational change is seen in gyroscope.

To find rotation-based transitions, we first compute the gyroscope magnitude from the 3-axes gyroscope data. Then, we normalize the gyroscope magnitude. To amplify the differences between rotation during brushing and rotation during transition, we take the exponential of each value of the normalized gyroscope magnitude sample. We find a threshold such that if the gyroscope value is higher than the threshold, we consider it as the beginning of a transition. To find the threshold, we first apply the Gaussian Mixture Model (GMM) to find two clusters: one for the lower values (during brushing or stationary) of the signal and the second for higher values (during transition). All the points in Cluster 2 are considered as transitions and time of those points are stored in $\mathcal{T}$.

(Step 2) Candidate Offset Detection: Let $\mathcal{T}_{\mathcal{G}}$ be the transitions from the video annotation. We seek to maximize the matching between the detected transitions from sensor and video. Therefore, we compute all the possible offset values to identify candidate offsets as $O = \{(t - t_g)\}_{t \in \mathcal{T}, t_g \in \mathcal{T}_{\mathcal{G}}}$.

(Step 3) Selecting the Best Offset: To find an offset that maximizes the number of matching, we find total matchings for each candidate offset value. We align the timing of the detected transitions by adding a candidate offset to the timestamp of each transition. We then find the closest distance from the marked transitions from video (i.e, ground-truth). If this distance is $< \epsilon$, we consider it a match. Then, we compute the number of detected transitions with a match. Finally, we select the offset that maximizes the number of matching to align the timestamp between video and the sensor data.

5.8 Stroke-wise Feature Extraction and Selection

After identifying brushing strokes as events within a brushing episode to segment the sensor data stream, we identify and compute several features from sensor data comprising each brushing stroke. We identify those features that are expected to vary during brushing of different surfaces, contributing to successful differentiation among each of them from the sensor data using a trained machine learning model.

5.8.1 Accelerometer Features

In Figure 5.18, we define brushing stroke i as a tuple of three points in a signal. Let ($time(v_i)$, $value(v_i)$) be the (timestamp, amplitude) of the valley v_i and ($time(p_i)$, $value(p_i)$) be the (timestamp, amplitude) of the peak p_i of the i^{th} brushing event. We identify 8 distinct features that are computed from the accelerometer 3-axes signal.

- **Peak Amplitude:** Peak amplitude corresponds to the amplitude value ($value(p_i)$) in each stroke duration, where the signal is at its maximum.

- **Valley Amplitude:** Valley amplitude corresponds to the amplitude value ($value(v_i)$) in each stroke duration, where the signal is at its minimum.

- **Rising Stretch:** Rising stretch is defined as the difference in amplitude of the peak ($value(p_i)$) and the valley immediately appearing before it ($value(v_i)$) of the i^{th} brushing cycle/stroke duration (see Figure 5.18).

- **Falling Stretch:** Falling stretch is defined as the difference in amplitude of the peak

($value(p_i)$) and the valley immediately following it ($value(v_{i+1})$) of the i^{th} stroke duration (see Figure 5.18).

- **Rise-Fall Ratio:** Rise-Fall ratio is defined as the ratio of rising stretch to the falling stretch.

- **Rising Duration:** Rising duration corresponds to the time elapsed from a valley of a stroke duration, to the subsequent peak (see Figure 5.18).

- **Falling Duration:** Falling duration corresponds to the time duration between a peak and the subsequent valley in a stroke duration (see Figure 5.18).

- **Stroke Duration:** Stroke duration is the sum of rising and falling duration.

We compute the above eight time-domain features for each of the 3-axis and magnitude signal of the accelerometer for a brushing stroke, resulting in a set of 32 features. In addition to these features, we compute **Correlation** measure that expresses the extent to which two variables are linearly related. As a result, three more correlation features among X, Y and Z axes are added, namely **corrXY**, **corrYZ**, and **corrZX**. In total, we have a set of 35 features computed for each brushing event or stroke.

5.8.2 Gyroscope Features

In addition to the accelerometer features, we also compute several features from gyroscope data. Since the gyroscope captures the amount of rotation in each axis, which is used to capture the surface switching/transition, we compute several statistical features, such as **mean** and **standard deviation**, to obtain the transition and the amount of rotation within each stroke. In total, we compute six features from three axes.

5.8.3 Orientation Features

The wrist's orientation with respect to gravity during brushing varies from surface to surface because of the position of the surface and angle of the wrist with the elbow. Recall that a brushing stroke consists of one forward movement (from the valley to peak in the signal) and one backward movement (from peak to next valley). During these two movements, the wrist has linear acceleration, but at the peak, the wrist gets stable, i.e., no linear acceleration, and prepares to move in the other direction. To capture the wrist's orientation, we compute **roll**, **pitch**, and **yaw** when the wrist is at the peak, i.e., at a stable state.

5.9 Model Selection and Training

During routine dental care, people generally initiate brushing sequence with the outer surface, followed by the inner surfaces. We observe a similar pattern among the study participants where they start and more importantly, cover all the portions of the outer surface first before moving onto the inner surface. To capture the natural layered hierarchy that is also captured in our organization of teeth surfaces (i.e., in/out, left/right/center, and up/down) as well as sequence of transition from one surface to the next, we select a hierarchical model that allows leveraging of any sequence patterns. We train a Hierarchical Bayesian Network for our model training.

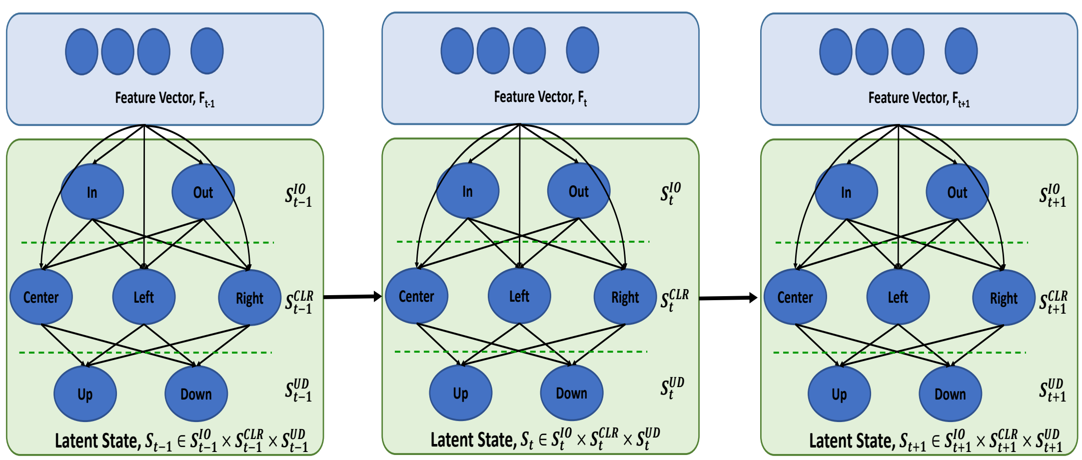


Fig. 5.23: Bayesian network with state transition

5.9.1 Bayesian Ensemble Method

A Bayesian network is a type of probabilistic graphical model that uses Bayesian inference for probability computations. Bayesian network aims to model conditional dependence, and therefore causation, by representing conditional dependence through edges in a directed graph.

Architecture of the Bayesian network for our brushing surface detection problem is shown in Figure 5.23. We organize the nine surfaces in the three surface layers as $S^{IO} = \{I, O\}, S^{CLR} = \{C, L, R\}, S^{UD} = \{U, D, *\}$, with $*$ denoting the ambiguity between up and down for the outer surfaces. Since any surface label is a combination of nodes from the tree layers, the class label set for the nine brushing surfaces is, $s \in \mathcal{S} \subset S^{IO} \times S^{CLR} \times S^{UD}$.

Now, for a given feature vector $f \in F$ of a brushing stroke, the model computes the likelihood of each surface label $s \in S$ as the predicted class using conditional probability $Pr[S = s|F = f]$. The model then outputs the class with the maximum probability as the final prediction for the brushed surface, i.e., $s = argmax_{s \in S} Pr[S = s|F = f]$.

Inference: For any feature vector f of brushing stroke, to compute probability of any surface $s \equiv (x, y, z)$, where $s \in S$, $x \in S^{IO}$, $y \in S^{CLR}$, and $z \in S^{UD}$, we use the following joint probability distribution function,

$$\begin{aligned} Pr[S = s|F = f] = Pr[S^{IO} = x, S^{CLR} = y, S^{UD} = z|F = f] \\ = \big(Pr[S^{IO} = x|F = f] \times Pr[S^{CLR} = y|F = f, S^{IO} = x] \\ \times Pr[S^{UD} = z|F = f, S^{IO} = x, S^{CLR} = y]\big) \end{aligned} \tag{5.1}$$

For example, probability of surface label $S = ICU$ is computed as,

$$\begin{aligned} Pr[S = ICU|F = f] = \big(Pr[S^{IO} = I|F = f] \\ \times Pr[S^{CLR} = C|F = f, S^{IO} = I] \\ \times Pr[S^{UD} = U|F = f, S^{IO} = I, S^{CLR} = C]\big) \end{aligned} \tag{5.2}$$

To compute these conditional probabilities, we learn a machine learning classifier–Random-forest model in each layer (a brief description of all the models is listed in Table 5.2 and Table 5.3). We then ensemble the outputs of these machine learning models using Equation 5.1 to produce the final output of the model.

Table 5.2: All the models that generate all required conditional probabilities

Models	**Tasks (Generates probabilities of surfaces)**
$M_{IO}(f)$	'in' and 'out' from feature vector f of a stroke
$M_{CLR\|I}(f)$	'center', 'left', and 'right' from f given $S^{IO} = I$
$M_{CLR\|O}(f)$	'center', 'left', and 'right' from f given $S^{IO} = O$
$M_{UD\|I,C}(f)$	'up' and 'down' from f given $S^{IO} = I$ and $S^{CLR} = C$
$M_{UD\|I,L}(f)$	'up' and 'down' from f given $S^{IO} = I$ and $S^{CLR} = L$
$M_{UD\|I,R}(f)$	'up' and 'down' from f given $S^{IO} = I$ and $S^{CLR} = R$

Table 5.3: All the models that generate outputs

Models	**Outputs**
$M_{IO}(f)$	$\langle Pr[S^{IO} = I\|f], Pr[S^{IO} = O\|f]\rangle$
$M_{CLR\|I}(f)$	$\langle Pr[S^{CLR} = C\|I, f], Pr[S^{CLR} = L\|I, f]\ , Pr[S^{CLR} = R\|I, f]\rangle$
$M_{CLR\|O}(f)$	$\langle Pr[S^{CLR} = C\|O, f], Pr[S^{CLR} = L\|O, f]\ , Pr[S^{CLR} = R\|O, f]\rangle$
$M_{UD\|I,C}(f)$	$\langle Pr[S^{UD} = U\|C, I, f], Pr[S^{UD} = D\|C, I, f]\rangle$
$M_{UD\|I,L}(f)$	$\langle Pr[S^{UD} = U\|L, I, f], Pr[S^{UD} = D\|L, I, f]\rangle$
$M_{UD\|I,R}(f)$	$\langle Pr[S^{UD} = U\|R, I, f], Pr[S^{UD} = D\|R, I, f]\rangle$

5.9.2 Dynamic Bayesian Ensemble (DBE) Method

Despite the wide variability in the brushing duration on each surface, we also observe stable patterns in surface transitions [24] for most of the participants, as shown in Figures 5.24, 5.25, 5.26 and 5.27. Dynamic Bayesian Ensemble (DBE) method uses the transitions to update the probabilities when it computes the probability of a surface that is different from the previously detected surface. Let T^* be the transition probability matrix, where each $T^*_{i,j}$ is the transition probability from surface i to surface j, where $i, j \in S$. We use the $*$ in a symbol to denote that the states can be over all nine surfaces or only over the groups of surfaces. We end up with four transition matrices, one for all nine surfaces and one each for the three layers. Note that we only consider transition probability when the current surface is changed, i.e., $T_{i,i} = 0$. Therefore, the updated probabilities are computed

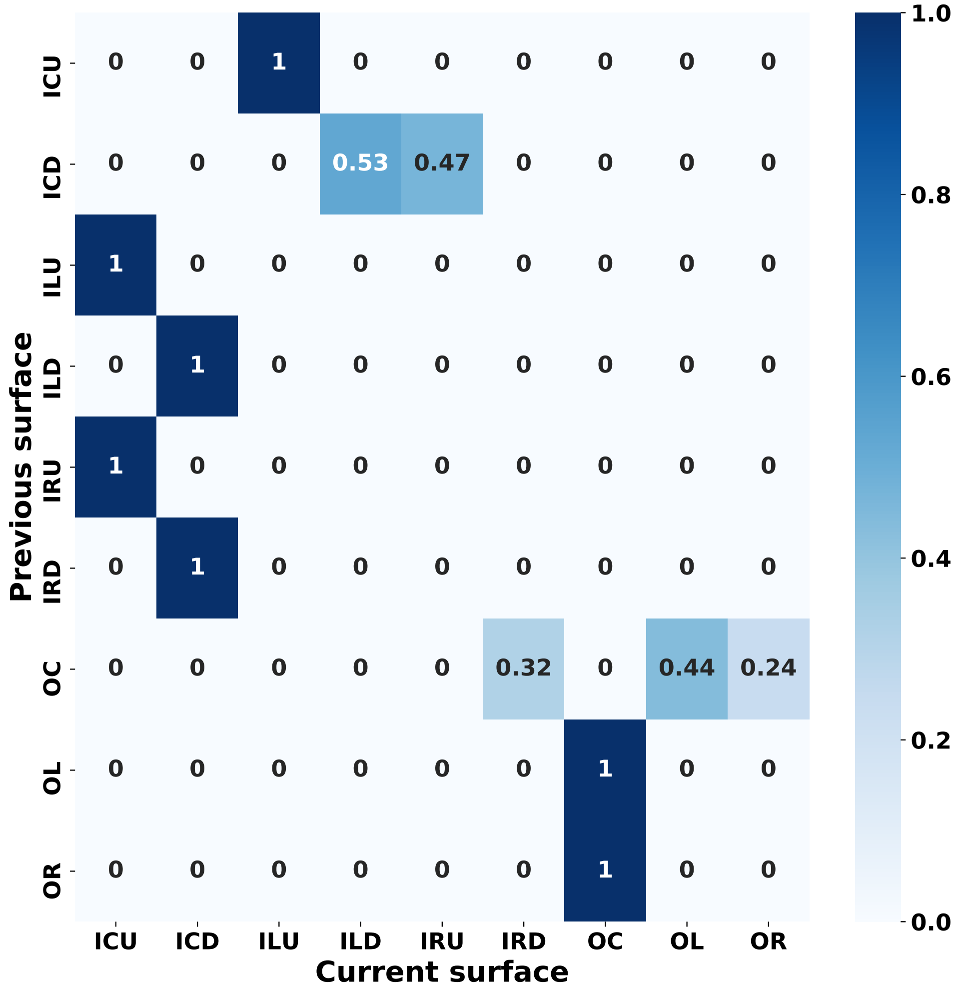


Fig. 5.24: Consistency in state-to-state transitions for participant '001'

as follows,

$$Pr'[S_t^* = x|F_t = f_t, S_{t-1}^*] = \begin{cases} Pr[S_t^* = x|F_t = f_t] & \text{, if } S_{t-1}^* == x \\ \alpha * Pr[S_t^* = x|F_t = f_t] + (1-\alpha) * T_{y,x}^* & \text{, else if } S_{t-1}^* == y, \forall y \neq x \end{cases}$$

Here, f_t is the feature vector of t^{th} brushing stroke, α is the parameter of the weighted average of two values, and $Pr[S_t^* = x|F_t = f_t]$ is computed using Equation 5.1. We use Pr' to denote the updated probability.

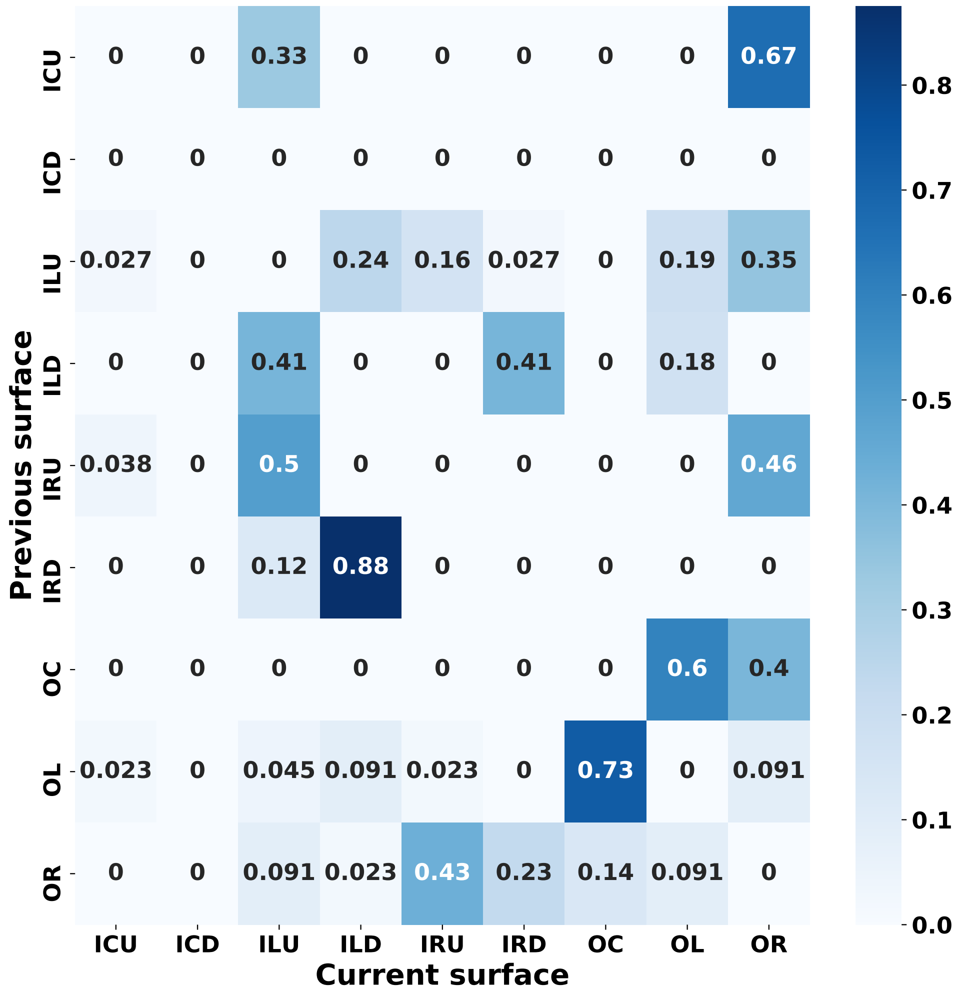


Fig. 5.25: Consistency in state-to-state transitions for participant '003'

Inference:

The selected class of t^{th} brushing stroke is given by

$$s_t =_{s \in S} Pr'[S_t = s | F_t = f_t, S_{t-1} = s_{t-1}],$$

where $\langle f_1, f_2, ..., f_m \rangle$ denotes a sequence of features. The model produces the surface sequence, i.e., $\langle s_1, s_2, ..., s_m \rangle$.

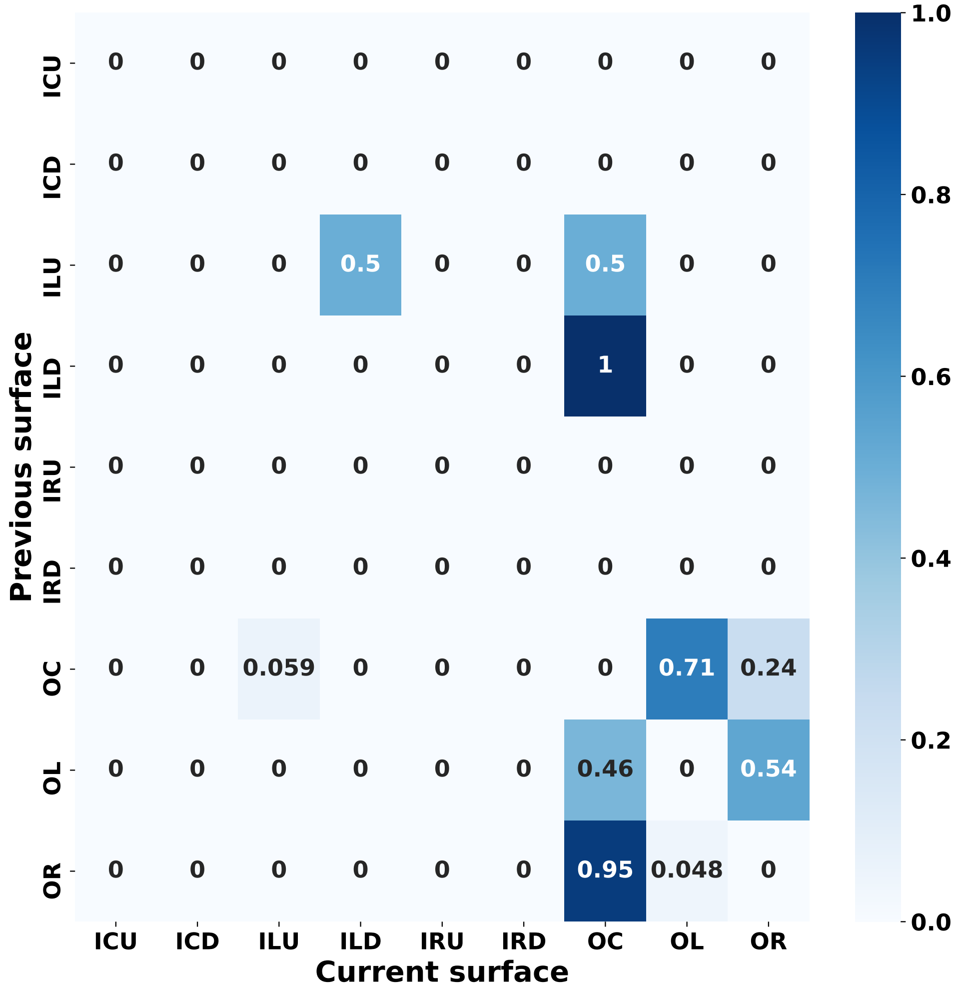


Fig. 5.26: Consistency in state-to-state transitions for participant '013'

5.10 Model Evaluation

The dataset we use confirms the wide between-person variability reported in dentistry [70, 71]. Additionally, it reveals substantial within-person between-episode variability not analyzed in prior works due to lack of such data. Recent works on detecting brushing patterns from wrist-worn inertial sensors [24, 69] collected multiple episodes from the same participants, but used 10-fold cross-validation. Hence, between-person generalizability of a machine learning model for detecting brushing surfaces has not yet been studied. The dataset we use has a larger number of episodes compared with [70, 71], more participants as compared with [24, 69], and is unique in representing natural brushing patterns in the

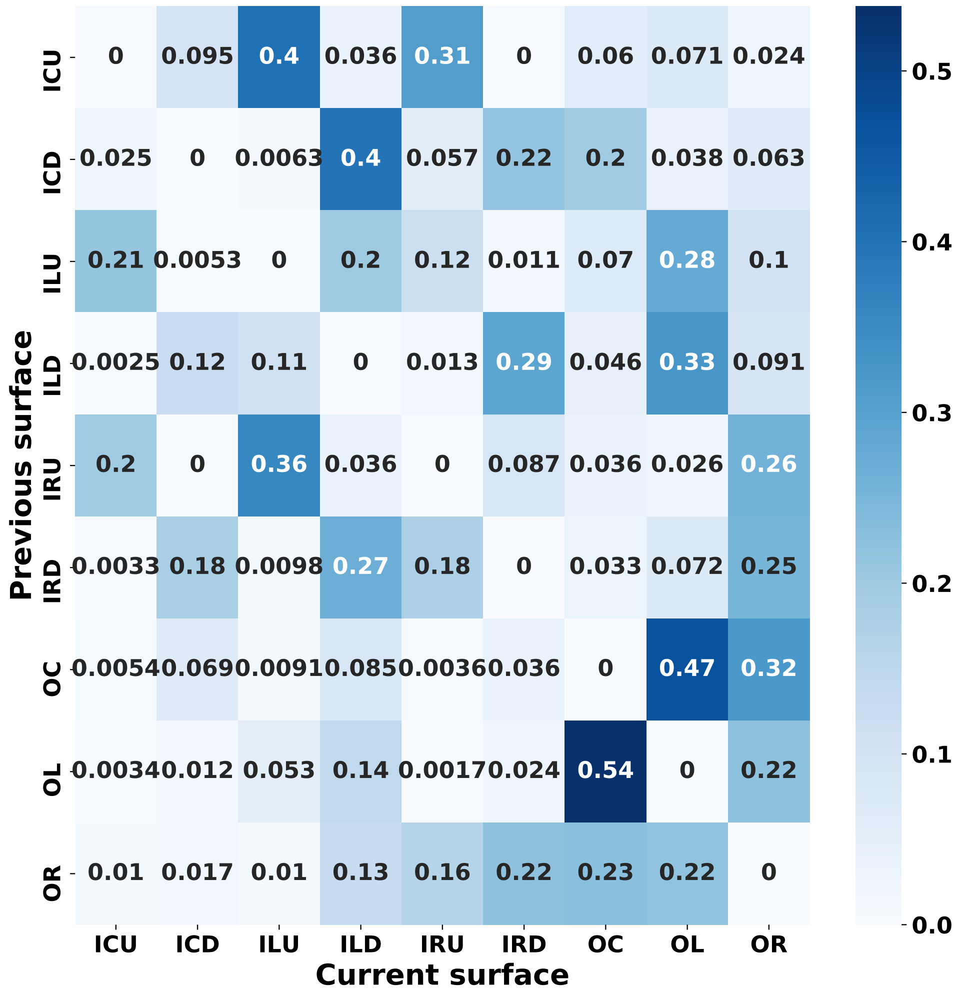


Fig. 5.27: Consistency in state-to-state transitions for all participants

users' home environment, without any specific brushing instructions (as in [24, 69]) that may reduce the natural between-person variability. To evaluate the between-person generalizability of our model, we start with Leave-One-Subject-Out-Cross-Validation (SCV), but also present 10-fold cross-validation (10CV) results to both allow a comparison with recent works on brushing surface detection and to show the impact of between-person variability in natural brushing on the model's performance. In addition, to study the impact of within-person variability among brushing episodes, we also perform Leave-One-Episode-Out-Cross-validation (ECV), where for each participant, we take one brushing episode as test data and the remaining episodes from that participant as training data, yielding a personalized model for each participant.

We start by evaluating the accuracy of detecting brushing strokes, which is a key enabling and distinguishing aspect of our model. Next, we evaluate the performance of our model for surface detection via all three validation methods. Finally, we study the impact of time synchronization on model performance, and conclude our evaluation by reporting the accuracy of estimating the total brushing duration on each surface, that can improve oral care.

5.10.1 Accuracy in Detecting Brushing Strokes

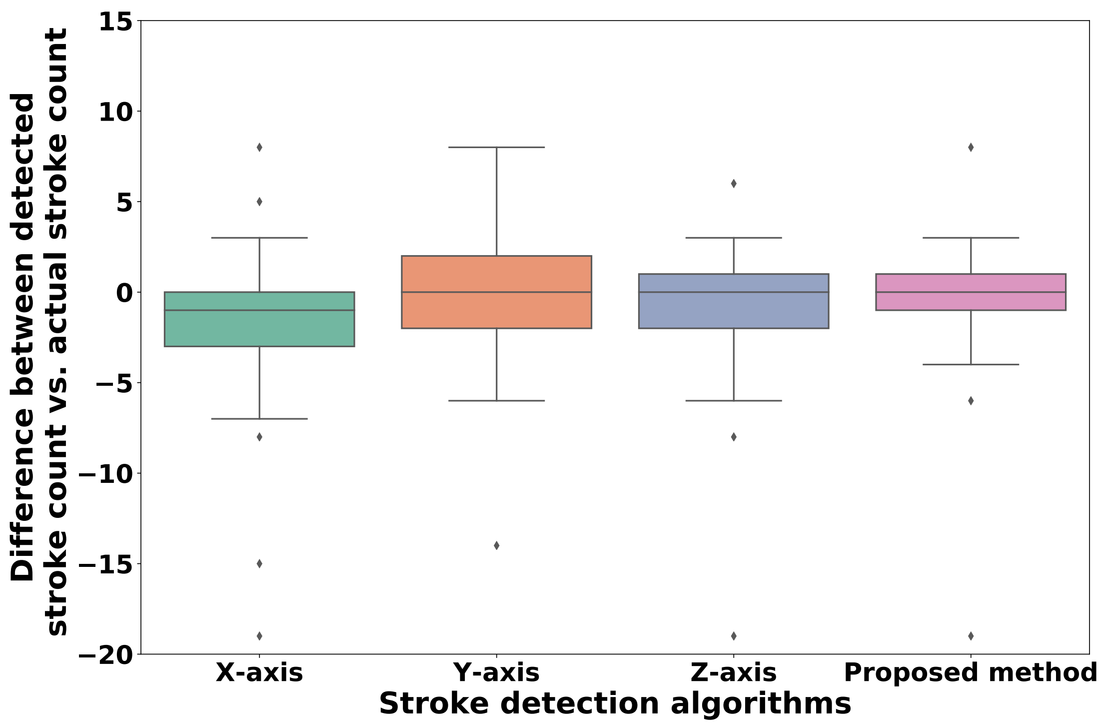


Fig. 5.28: Distribution of stroke count differences between detected stroke counts and ground-truth

To evaluate the accuracy of our brushing stroke detection method (see Section 6.9), we compare the number of brushing strokes detected in each brushed surface from sensor data with that from video annotation. As annotating each brushing stroke (lasting only milliseconds) for all the episodes is even more arduous than annotating each brushing surface, we

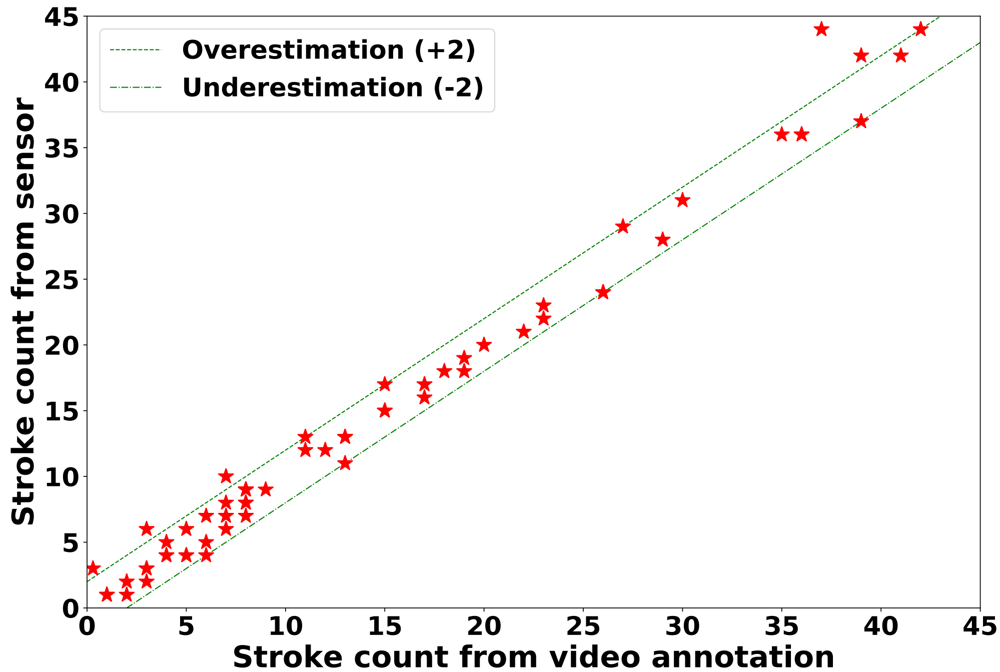


Fig. 5.29: Scatter plot of actual count vs. detected count of brushing strokes in a brushing surface

limit our annotation to 1,456 brushing strokes from 100 surfaces. For each surface, we count each and every periodic movement (valley-peak-valley) with no condition imposed on brushing strokes for each of the x, y and z-axis of the accelerometer separately and also through our brushing stroke detection model discussed in Section 6.9. We calculate the difference between the counts of brushing strokes from sensor data and that from video annotation for each annotated episode and present the results in Figure 5.28. We observe from the distributions that our proposed method results in the lowest error (mean absolute error is 1.5). In Figure 5.29, the scatter plot shows the counts of brushing strokes from the video and sensor data using our proposed method. We observe that most of the errors are limited to no more than 2 strokes, even when the number of strokes are as high as 40 (in a brushing surface). We note that [24] estimated the number of brushing strokes using acoustic sensors with an average error of 10.3%, leaving the task of stroke detection using inertial sensors open. Our stroke detection algorithm solves this open problem with less than 4.2% error.

5.10.2 Impact of Between- and Within-Person Variability on Brushing Surface Detection

We trained Naive Bayes (NB), Random Forest (RF), Bayesian Ensemble (BE), and Dynamic Bayesian Ensemble (DBE) for brushing surface detection (see Figure 5.30) and find that the DBE produces the best performance (for 10 CV). Hence, we use DBE for subsequent evaluations. As brushing recommendations are usually based on broad surface category, we begin by evaluating the performance for detecting the three surface layers. Recall that our Bayesian Ensemble model consists of six models each of which is a Random-forest classifier, as shown in Table 5.2. We evaluate the performance of each model for classification at each teeth surface layer and present the results in Figure 5.31. Note that any model in the form of $M_{x|y}$ is trained on a filtered dataset belonging to surface label y from the upper layer. For example, $M_{CLR|I}$ is trained on only the feature set from inner surface. Table 5.4 shows the results for both SCV and 10CV. The $M_{UD|I,C}$ model for inner-center surfaces achieves the best performance.

For nine-surface classification, the model obtains median recall, precision and F1-score of 65.26%, 65.30% and 63.14% for SCV, which improves to 82.14%, 82.66% and 79.50% for ECV, and further improves to 87.06%, 86.96% and 86.02% for 10CV. Low performance for SCV as compared with 10CV (used in prior works on brushing surface detection [24, 69]) can be explained by wide between-person variability. As described in Section 5.4.2, there does not exist a population profile or even clusters of participants with similar brushing patterns (see Figures 5.10, 5.11 and 5.12). Hence, a model trained on other participants' data performs poorly when tested on a different participant.

We observe that training a personalized model for ECV improves the performance substantially from SCV, but still falls short of the 10CV performance due to within-person between-episode variability exhibited in natural brushing habits of participants (see Sec-

Table 5.4: Cross Validation performance for identifying broad teeth surface categories. Median values are reported here.

	$M_{IO}(f)$		$M_{CLR\|I}(f)$		$M_{CLR\|O}(f)$		$M_{UD\|I,C}(f)$		$M_{UD\|I,L}(f)$		$M_{UD\|I,R}(f)$	
	SO	**10F**	**SO**	**10F**	**SO**	**10F**	**SO**	**10F**	**SO**	**10F**	**SO**	**10F**
R(%)	77.6	94.1	83.9	94.1	85.8	96.8	86.8	100	65.4	96.6	77.5	97.1
P(%)	79.3	94.1	84.1	94.4	88.9	96.9	100	100	81.3	96.9	78.5	97.4
F1(%)	77.4	93.9	82.2	94.2	84.9	96.8	91.7	100	60.9	96.5	72.1	97.1

tion 5.4.2). Figure 5.33 displays breakdown of the result by participant with individual precision, recall and F1-score to show participant wise between-episode variability. We compute between episode variability as follows: as discussed in Section 5.4.2, each episode is represented as a duration vector of all the brushing surfaces, i.e., a nine value vector, and we use the Euclidean distance metric to compute the distance between two such vectors. We compute Euclidean distance of all pair-wise combinations of a participant's episodes and take the mean of the distances as a representative of the between-episode variability for that participant. To show the relative variation over the participants, we plot the normalized measurements of all the participants. We observe that between-episode variation in the total time spent in the nine surfaces highly affects the performance of correctly identifying the surfaces. As discussed in Section 5.4.2, we observe that some participants completely miss some surfaces in many of the brushing episodes due to their personal brushing habits. So, when the model is trained with mostly missing data for a surface from most of the episodes and asked to detect the surface when it is present in the test episode, it fails to do so.

5.10.3 Impact of Time Synchronization on Nine Surface Classification

We first manually check if the proposed method correctly synchronizes the sensor data to the video-obtained labels. We find that 101 out of 114 episodes (88.59%) are correctly synchronized. We carefully analyze the remaining 16 episodes and observe that wrist movement during brushing is too slow to detect the significant rotation required to identify transitions and make a cluster. We manually synchronized these 16 episodes. Next, to analyze the impact of time synchronization, we train a model without performing the time

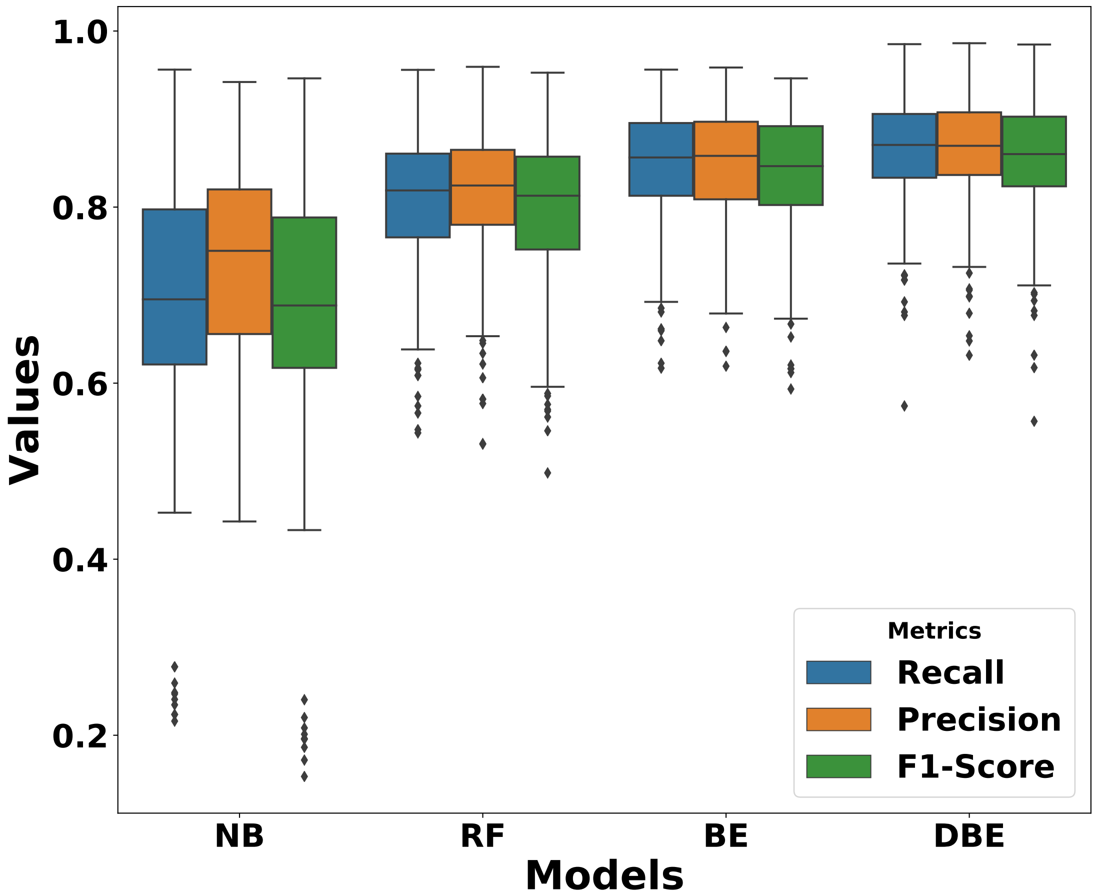


Fig. 5.30: Performance of different classification models

synchronization step. We find that the surface detection accuracy (using 10CV) drops significantly to recall, precision, and F1-score of 75.09%, 74.02%, and 73.34% respectively, showing a drop in F1-score by almost 13%.

5.10.4 Accuracy of Estimating Total Brushing Duration on Each Surface

Thus far, we have presented the accuracy of detecting when each surface is being brushed. As we present in Section 5.4, participants switch frequently between surfaces, coming back to a surface multiple times. For oral health purposes, both users and their providers may be interested in determining the total time a user spends in brushing of each surface in a brushing episode. Figure 5.34 shows the percent error (as compared with la-

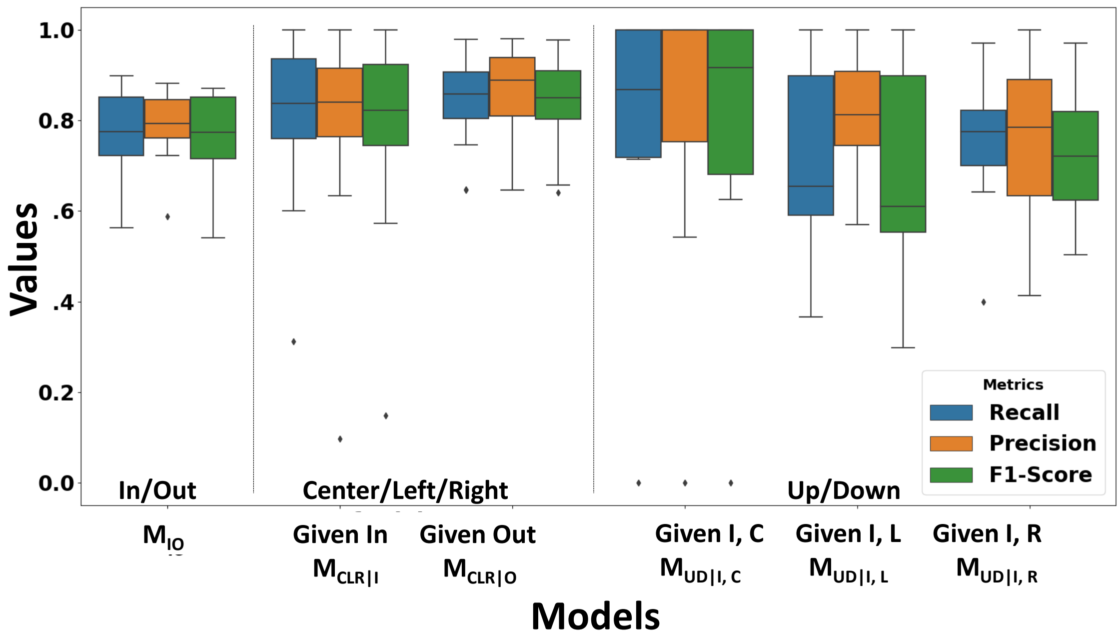


Fig. 5.31: Classification performance using Leave-one-subject-out cross-validation (SO) at each layer (for models defined in Table 5.2)

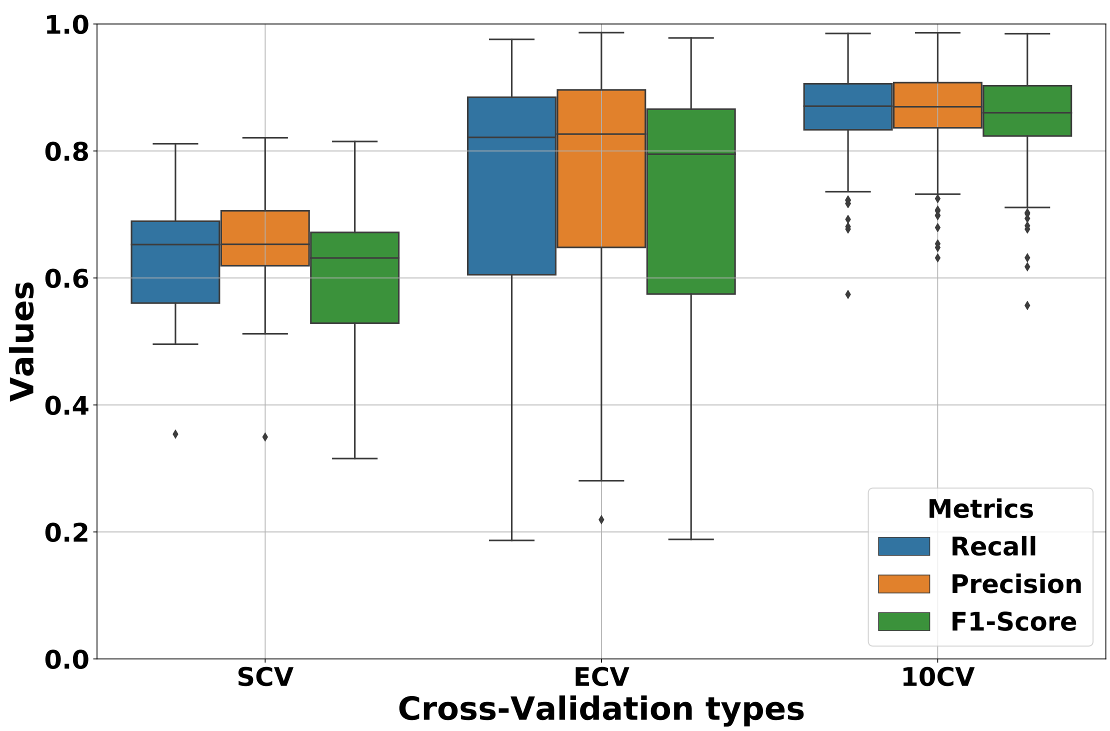


Fig. 5.32: Nine surfaces classification results for different types of cross-validation results

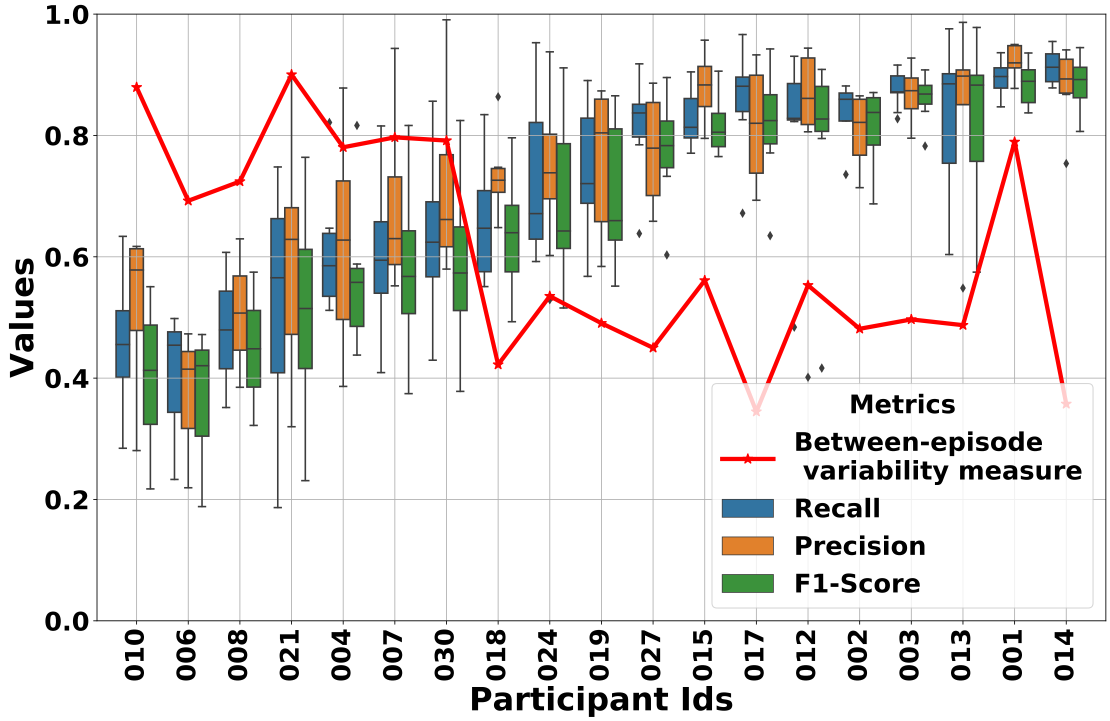


Fig. 5.33: Nine surfaces classification results for leave-one-episode-out cross validation results

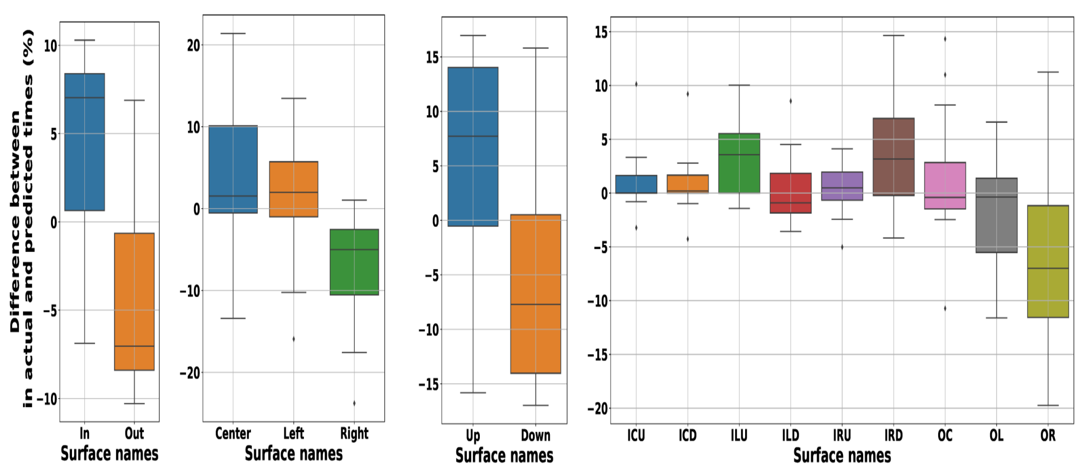


Fig. 5.34: Accuracy of estimating total duration of brushing in different surfaces

beled data from video) in estimating the total duration of brushing on surface groups and for each of the nine surfaces.

We observe that the median absolute error is $< 7.5\%$ for in vs. out, $< 2.5\%$ for center vs. left vs. right, $< 7.5\%$ for up vs. down and $< 7\%$ for all of the nine surfaces. Even the first and third quartiles are $< 5\%$ in most cases. We notice that most errors are from confusing some instances of out with in, right with left, and down with up.

Finally, we note that using mTeeth model can improve the estimates of start/end times of brushing (and total duration of brushing, a widely-used clinical variable) by models such as mORAL [29]. Our assumption is that mTeeth will be triggered upon detection of the start of a brushing event by mORAL. mORAL considers the start of the event from when the hand is in the upward direction, which includes putting the paste on the brush-head and preparing to brush; and end of the event when the hand moves to the downward position. Error in start/end times is 4.1% for the mORAL model. Once mTeeth model is activated by mORAL, by using our stroke-based approach, the error in estimating the total brushing duration will be reduced from 4.1% to $< 0.5\%$.

5.11 Limitations and Future Work

The work presented here has several limitations that open up opportunities for future work. First, our video annotation did not disambiguate between occlusal and lingual surfaces. As users are known to spend more time brushing occlusal and less time on lingual surfaces [70], future work can improve on video annotation and model training to seperately estimate the time spent on these two kinds of surfaces.

Second, this work did not estimate the pressure being applied during brushing, which is also an important component of brushing efficiency. Future work can develop methods that can leverage the stroke detection and characterization approach presented here to estimate the pressure applied during brushing of different surfaces.

Third, this work analyzed a week worth of daily brushing data from 19 participants during their natural brushing sessions. This work found significant variability between episodes of the same participant, and even greater variability among participants. As a result, although it achieved very high accuracy of classification in 10-fold cross-validation, but found the accuracy drop for leave-one-episode-out training and drop even further for leave-one-subject-out training. As between-person generalizability is important for real-life adoption of machine learning models, future work can increase the number of episodes per person and the number of participants to determine the level at which clusters emerge among both episodes and among users that exhibit sufficient similarity. These clusters can then be used to develop group specific models which can more accurately detect brushing surface for each brushing episode and for each person.

Fourth, the time synchronization method presented here missed 16 out of 122 episodes, due to very slow brushing pace of some participants. Future work can investigate better methods that can automatically synchronize video labels and sensor data for all episodes without manual intervention.

Fifth, the algorithms presented here for stroke detection and time synchronization found specific thresholds that was suitable for the current dataset. Future work can develop adaptive thresholds or other adaptive algorithms that can generalize to unseen datasets without retraining. Finally, this work only observed the natural daily brushing behavior of participants and did not attempt to teach them better brushing habits. Future work can leverage the mTeeth model to develop interventions that can help users self-reflect on their brushing habits, detect regularly missed surfaces, and present personalized behavioral nudges to help individuals optimize their oral self-care routines and proactively tackle teeth surfaces at-risk for plaque accumulation.

5.12 Chapter Summary

The orientation of a toothbrush changes noticeably when brushing different tooth surfaces, resulting in detectable changes if inertial sensors are embedded in or attached to the brush itself, as in smart or instrumented toothbrushes. However, inferring tooth surface coverage from wrist-worn sensors is much more challenging because the changes are very subtle as the general orientation of the hand does not change much when transitioning from one teeth surface to another. Give that most brushes are manual and lack sensors, we develop a model for leveraging sensor data from ubiquitous smartwatches to infer brushing coverage. This work presents several insights for detecting these subtle signatures and constructs a model to distinguish among teeth surfaces and transitions between them. By doing so, it opens up a new frontier in the detection of rare daily events such as brushing, flossing, eating, drinking, and smoking by allowing finer-grained characterization (i.e., detecting even more ephemeral embedded micro-events) of self-care activities in natural environments. This may motivate new methods for successful characterization of other rare events such as detecting smoking with e-cigarettes that only consists of one or two puffs at a time, classifying among different kinds of food or drink in an eating or drinking episode by distinguishing the subtle differences in the hand-to-mouth gesture involved, and similar other daily behaviors.

Chapter 6

Brushing Prompt: When to Intervene?

6.1 Introduction

We have detected manual tooth brushing and flossing from wrist-worn inertial sensors. After automated detection of brushing we need the capability of monitoring which surfaces are not adequately being brushed. Since a primary reason for continued prevalence of dental diseases despite regular brushing is that people may not be brushing each tooth surface adequately, missing some surfaces completely, while spending disproportionate time on other surfaces. Then we have detected brushing teeth surfaces from wrist-worn inertial sensors. By analyzing the data, we know which surface needs to be more attention. Each of these brushes and these technologies have been developed. Now with the surface detection and intervention can be provided to users to remember to brush the right surface. We can have a model in future that can leverage the mTeeth model to develop interventions that can help users self-reflect on their brushing habits, detect regularly missed surfaces. Power of the mobile sensor should be find those right time to deliver intervention, only then it has higher chance to succeed. It is required to detect when they are about to brush. Just detecting their when they are at sink that will not suffices because people can go the sink many times . When they have gone to the sink to wash their hands telling them about intervention they may not remember. So that's why detecting about to brush moment using wrist worn inertial sensors is important. It is also required to detect detect that a user has started toothbrushing so that we can give intervention quickly.

A primary reason for the continued prevalence of dental diseases despite regular brushing is that people may not be brushing each tooth surface adequately, missing some teeth surfaces completely, while spending disproportionate time on other teeth surfaces. A proper intervention can help the participants to correct their brushing habits. The success of the intervention depends on two things – *what to deliver*, i.e., content of the intervention, and *when to deliver*, i.e., timing of the intervention. That is a successful intervention depends on delivering the right content at the right time.

What to deliver: In chapter 3, we observe most of the participants spend most of the time brushing on outer surfaces. Furthermore, they just touched few teeth surfaces without proper brushing. The proposed mTeeth model is able to find the brushing time in each surface with more than 90% accuracy. This model provides which surface needs to be more attention. Detected summary of the brushing session can be very effective content of the intervention since it can help users self-reflect on their brushing habits, detect regularly missed surfaces.

When to deliver: After having the right content to deliver, the next task is to find the right time to deliver the intervention. Even if we have the right content but delivered it at the wrong time, it may not be effective. For example, if intervention is provided too early or when the user is busy, the user may forget the missing surface in this brushing scenario. The power of the mobile sensor enables the ability to deliver intervention at the right time, only then it has a higher chance to succeed. All the existing works only detect after the brushing ends. By that time, delivering the intervention to the user may be too late. The user may not be able to go back and correct their brushing. Therefore, better chance for the user to see the instruction and remember the message when they are about-to-brush, that what requires detecting when they are about-to-brush. Just detecting when they are at the sink is not sufficient because people can go to the sink many times for many reasons. When they have gone to the sink to wash their hands, telling them what surface to brush on might not be effective since they may not remember. Therefore detecting about-to-brushing moments using wrist-worn inertial sensors is important.

6.2 Problem Statement

There are several observables that happens before somebody begins to brush. One of those is walking but that is highly non-specific. Since people walk indoors more frequently, creating many false intervention moments. Also, walking to the sink is challenging only with the wrist-worn inertial sensors since indoor localization requires environmental sensors or some additional sensors. Thus, we seek for an event that is unique and close to the brushing event. About-to-brushing events pertain to actions like applying toothpaste to the toothbrush immediately before brushing, while the beginning of brushing refers to the initiation or initial strokes of the toothbrushing event. By detecting these pivotal moments, we can deliver timely interventions and tailored guidance, drawing from the aggregated data of toothbrushing activities from previous days. This approach encourages the development of consistent and precise brushing habits.

6.2.1 About-to-Brush Moment

Fig. 6.1: Five micro-events for about-to-brush moment

The moments just before performing an event, known as "about-to-event" moments, are crucial in predicting human behavior. In the context of toothbrushing, we identified five sequences of micro-activities that constitute an **about-to-brushing** event:

- **Take out paste**: This involves retrieving the toothpaste tube from a drawer, shelf, or counter, ensuring that toothpaste is readily available for brushing.
- **Open cap**: The toothpaste tube is equipped with a cap, which needs to be flipped or rotated to access the toothpaste. Opening the cap signifies the readiness to apply toothpaste.

- **Apply paste**: Applying toothpaste onto the toothbrush is a key step in the toothbrushing process. It involves squeezing the toothpaste tube with one hand while holding the toothbrush with the other.

- **Close cap**: Once the desired amount of toothpaste has been dispensed, the cap of the toothpaste tube is closed by flipping or rotating it.

- **Put paste away**: After applying toothpaste, the toothpaste tube is returned to its original location, and the toothbrush may be moistened with water before starting the brushing activity.

These sequences of micro-activities form a specific pattern that characterizes the about-to-brushing event as shown in Figure 6.1.

There are instances when people engage in similar activities, such as taking out a comb or utensils while cooking, which reduces the specificity of the action of taking out the toothpaste. Similarly, opening a cap is not commonly associated with combing hair. However, the sequence of taking out the paste, followed by opening the cap, adds a level of specificity. Nonetheless, it is important to note that other scenarios, such as applying cream to the face or painting, can involve similar actions.

To further refine the specificity, we introduced the sequence of taking out the toothpaste tube, opening the cap, and then applying the paste. This combination narrows down the context to toothbrushing-related activities. However, it is worth mentioning that the act of hand proximity to the mouth and taking a puff while smoking closely resembles the sequence of opening the cap and applying the paste.

By incorporating the additional steps of closing the cap and putting the paste away, the problem becomes significantly more specific, aligning with the distinct pattern of toothbrushing activities. These sequences provide a clearer understanding of the context and enable a more precise identification of the about-to-brushing event.

6.2.2 Brushing Initiation

Brushing initiation means we can use some initial portion of brushing. How early in the brushing episode could we detect that person is now brushing. If we can detect initially few brush stroke then it prompt the intervention during the initial portion of brushing. It should still work and increase the detection accuracy. We need reliable stroke detection though mTeeth [30] has stroke detection model but it not be applicable here. mTeeth stoke detection only applied during toothbrushing brushing time to detect toothbrushing surface where as precision is high. However, for this intervention problem we have to apply new stroke detection model to detect toothbrushing activity from whole day time series data. This new stroke detection model should Filter the periodic signal such as walking, toothbrushing. Here we have to concentrate two criteria such as

- How quickly can we detect that a user has started toothbrushing?
- What is the minimum number of strokes typically associated with random activities other than toothbrushing?

6.3 Related Works

The paper by Rahman [77] focuses on developing a predictive model for identifying "about-to-eat" moments to support just-in-time eating interventions. Using a mobile application, researchers collected data on participants' eating behaviors, including meal time, location, duration, and contextual information. The machine learning model, utilizing features such as time, location, social context, and user behavior patterns, successfully predicted these moments with high accuracy. The study suggests leveraging this predictive capability to deliver personalized interventions for healthier eating habits, such as providing suggestions or messages prior to predicted eating events. Overall, the paper highlights the potential of predictive models and mobile technology in promoting healthier dietary behaviors by intervening at crucial moments in individuals' eating routines (Rahman, 2016).

The paper by Gustafson et al [78] discusses a mobile technology-based system aimed at improving outcomes for individuals in recovery from alcohol dependence. The system integrates mobile phones, web interfaces, and interactive voice response (IVR) systems to deliver evidence-based interventions, support self-monitoring, and facilitate communication. By incorporating theories and frameworks such as the Transtheoretical Model and Motivational Interviewing, the system addresses individual needs and stages of change. Pilot studies suggest that the system has the potential to enhance treatment adherence, increase abstinence rates, and improve self-efficacy. Overall, the paper highlights the promise of mobile technology in supporting alcohol dependence recovery.

The paper by Mohr et al. [79] presents the Behavioral Intervention Technology (BIT) model—an integrated framework for eHealth and mHealth interventions. The model combines behavioral science theories with technological advancements to improve the effectiveness and reach of digital interventions. It comprises theory, content, features, and human support components. The BIT model offers benefits such as increased accessibility, scalability, and personalization of interventions. The paper emphasizes the importance of rigorous evaluation and continuous refinement of interventions based on user feedback and data analysis. Overall, the BIT model aims to optimize the impact of digital interventions on health behaviors and outcomes.

The paper by Pina et al. [80] explores the use of mobile technology to provide in situ cues for parents implementing ADHD parenting strategies. They developed a mobile application that delivers real-time prompts and reminders based on the child's behavior, location, and time of day. User studies showed positive feedback, indicating that the cues helped parents stay consistent and manage challenging behaviors. The authors suggest that mobile applications with in situ cues can enhance ADHD parenting interventions and provide valuable support to families in real-time.

The paper by Wansink and Johnson [81] explores the "clean plate" behavior, finding that approximately 92% of self-served food is eaten. The study highlights the influence of environmental factors, such as plate size, on food consumption. Larger plates tend to lead to larger portion sizes and increased food intake. The findings underscore the importance of mindful portion control to support healthy weight management.

The paper by Witkiewitz and Marlatt [82] discusses relapse prevention for alcohol and drug problems, comparing the Zen and Tao approaches. The Zen approach emphasizes acceptance, mindfulness, and non-attachment to cravings, while the Tao approach focuses on balance and addressing underlying imbalances contributing to substance use. The authors suggest that integrating elements from both approaches can lead to comprehensive and effective relapse prevention strategies. They emphasize the importance of tailored interventions, including cognitive-behavioral techniques, coping skills training, and social support. By combining acceptance, mindfulness, balance, and ongoing support, individuals can enhance their chances of maintaining long-term recovery from alcohol and drug problems.

The paper [83] explores the application of offline contextual multi-armed bandits (CMAB) for mobile health interventions, specifically focusing on emotion regulation. The study develops a recommendation system using offline data to provide personalized interventions for individuals seeking support in emotion regulation. Through a case study, the researchers evaluate the performance of the offline CMAB approach by collecting data from a mobile health app. The study demonstrates the potential of offline CMAB algorithms in delivering personalized mobile health interventions for emotion regulation, leveraging contextual information to provide tailored recommendations for effective emotion management. The findings contribute to the field of mobile health interventions, showcasing the effectiveness of CMAB algorithms in providing personalized support for individuals' emotional well-being.

The paper [84] introduces the Food Watch system. This system utilizes wearable sensors and video recordings to capture hand-to-mouth gestures during meals, enabling the detection and characterization of eating episodes. Through experiments, the authors demonstrate the accuracy of the Food Watch system in identifying eating behaviors and discuss its potential applications in dietary monitoring, weight management, and behavioral interventions for promoting healthier eating habits. Overall, the paper showcases the feasibility and effectiveness of utilizing feeding gestures to monitor and analyze eating episodes for various health-related purposes.

The paper Fingerprints [85] system, which uses mobile sensing to detect and categorize meaningful moments for health interventions. By combining data from various sensors, the system identifies contextually relevant moments in individuals' daily lives that can be targeted for interventions. The research validates the effectiveness of Fingerprints in detecting and classifying these moments, offering potential for timely and personalized mobile health interventions.

The paper [86] focuses on evaluating the availability of users to engage in just-in-time interventions in their natural environment.The study aims to understand the feasibility of delivering interventions at the right moment by assessing users' availability and receptiveness in their daily lives. The authors conducted a research study involving participants equipped with mobile devices and sensors to collect data on their availability, engagement patterns, and contextual factors. Through their analysis, the researchers examine the participants' willingness and ability to engage in just-in-time interventions based on various contextual factors, such as location, time of day, and activity. They discuss the implications of these findings for designing effective just-in-time interventions and highlight the challenges associated with timing and user availability. The study provides insights into the feasibility of delivering interventions in users' natural environment, taking into account contextual factors. The findings contribute to the design and implementation of effective just-in-time interventions that align with users' availability and enhance their engagement.

6.4 Data Overview

In the realm of data-driven decision making and analysis, understanding the characteristics and properties of data is paramount. This chapter delves into the essential aspects of data selection, exploration, description, and the crucial process of data labeling. The chapter begins by emphasizing the importance of gaining an overview of the data. A comprehensive data overview involves understanding the sources, formats, and structure of the data, as well as the context in which it was collected. This step sets the foundation for subsequent analyses and enables researchers to identify potential biases, missing values, and other data quality issues. Next, the chapter focuses on data description, which involves summarizing and characterizing the dataset. A comprehensive data overview involves understanding the sources, formats, and structure of the data, as well as the context in which it was collected. This step sets the foundation for subsequent analyses and enables researchers to identify potential biases, missing values, and other data quality issues. Finally, the chapter delves into the critical task of data labeling. Labeling can be performed manually or through automated techniques, depending on the nature of the data and the labeling task.

6.4.1 Dataset Selection

A wrist-worn inertial sensor data set consisting of labels of start/end of brushing and flossing episodes used in our mORAL [29] study is available publicly. This study recruited participants willing to brush at least twice — once with a manual toothbrush and once with a SmartBrush and floss at least once a day. Each participant wore a MotionSense wristband on each wrist during waking hours for seven days that included a 3-axis accelerometer and a 3-axis gyroscope sampled at 16 and 32 Hz, respectively. A study provided smartphone connected via Bluetooth technology continuously timestamped and logged incoming sensor data. Besides, participants used the phone's front camera to video record themselves (in their homes) during brushing, flossing and/or oral rinsing. The mORAL dataset currently consists of data from 30 participants (15 males, 15 females; mean age 28.5±10.6 years, 2 left handed).

In the public dataset, the start and end times of brushing episodes are annotated from self-recorded videos. This annotations in the mORAL dataset [1]are used for brushing ini-

[1]https://mhealth.md2k.org/resources/datasets.html#mORAL

tiation model for giving intervention. However, we have also another approach for detect about-to-brush moment for giving intervention to the people. The original annotations in the mORAL dataset are insufficient for our modeling, because it does not include any about-to-brush micro event annotations. We used the original videos from this study to label precise times for each about-to-brush micro event was performed.

6.4.2 Dataset Details

Out of 412 brushing episodes of 30 participants for manual and smartbrush videos for most of the episodes were not usable for about-to-brush micro event annotations.During ROBAS phase 1 data collection study we did not ask for participant to record the video for about-to-brush moment. Most of the participants did not record their pre brushing steps or preparation steps before tooth brushing. Most of the participants record their toothbrushing, flossing and rinsing steps. For that reasons, 83 brushing sessions had to be excluded from the about-to-brush moment modeling work. We annotated the remaining 70 episodes from 12 participants. However, for brushing initiation modeling work we use 197 manual brushing episodes from 30 participants. Since stroke detection model does not work on smart brush brushing episodes so rest of the 215 smart brush brushing episodes had to be excluded.

6.4.3 Annotation Protocol for About-to-Brush Approach

We annotated data from twelve participants' videos, focusing on four time series: *about-to-brush*, *tube holding wrist*, *toothbrush holding wrist*, and *active wrist*. Each time series consists of five sequences. The *about-to-brush* time series is labeled with the following actions: *bring*, *open*, *paste*, *close*, and *back*. *Bring* refers to taking out the toothpaste tube from the cabinet, *open* means opening the toothpaste tube cap by rotating or flipping it, *paste* denotes applying the paste to the toothbrush, *close* indicates closing the toothpaste tube cap, and *back* signifies putting the toothpaste back into the cabinet. The *tube holding wrist* is labeled as the left or right wrist, depending on which wrist is used for each micro event in the *about-to-brush* time series. Similarly, the *toothbrush holding wrist* is labeled as the left or right wrist, based on the wrist used for each micro event during the *about-to-brush* time series.

During the annotation process, we encountered several challenges that arose from the intricate coordination between the two wrists during the "about-to-brush" time series. Each micro event, such as *bring*, *open*, *paste*, *close*, and *back*, requires a specific combination of wrist movements. The complexity arises from the fact that the roles of the wrists can alternate for different actions.For example, when taking out the toothpaste tube from the cabinet, the right wrist may be primarily responsible for this task, whereas opening the toothpaste tube cap might be better suited for the left wrist. This dynamic interplay between the wrists can create a complex pattern of movements.

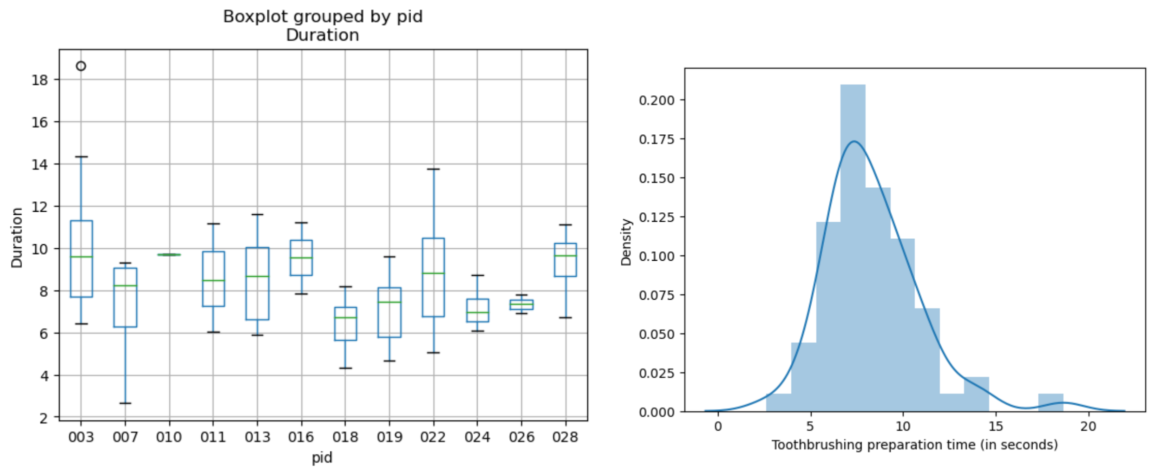
Moreover, there are instances where the roles of the wrists switch within a single micro event. Consider the scenario where the left wrist opens the toothpaste tube cap, but the right wrist takes charge of applying the paste to the toothbrush. These intricate variations in wrist involvement make the annotation process challenging but essential for accurately capturing the participants' actions. Despite the simultaneous engagement of both wrists throughout the "about-to-brush" time series, we made efforts to determine which wrist played the most significant role in each micro event. By identifying the primary active wrist, whether left or right, we aimed to provide a clearer understanding of the participant's actions and movements during the toothbrushing process.

This detailed annotation of the wrist activities not only contributes to a comprehensive analysis of the brushing behavior but also opens doors for further research and potential improvements in dental care techniques and ergonomic design of oral hygiene products.

6.5 Insights from Newly Annotated Dataset: A Descriptive Analysis

This dissertation explores various aspects of toothbrushing behaviors, aiming to uncover the individual differences and patterns that influence oral hygiene practices. Through meticulous data analysis and observations, four key sections shed light on different facets of toothbrushing, including the duration of the "about-to-brush" phase, the variability in opening and closing the toothpaste tube cap, the role of wrists during toothbrushing actions, and the accelerometer signals during the "about-to-brush" moment. The findings presented in this dissertation offer valuable insights into the intricate nature of toothbrushing, highlighting the need to consider individual variations in behavior and preferences. With this knowledge, tailored recommendations and interventions can be developed to optimize oral hygiene practices and promote overall dental health for individuals.

6.5.1 Variability of About-to-brush Timeseries Duration



((a)) Analysis of about-to-brush moment durations obtained from video annotations.

((b)) Distribution of about-to-brush moment durations.

Fig. 6.2: Variability of about-to-brush timeseries duration.

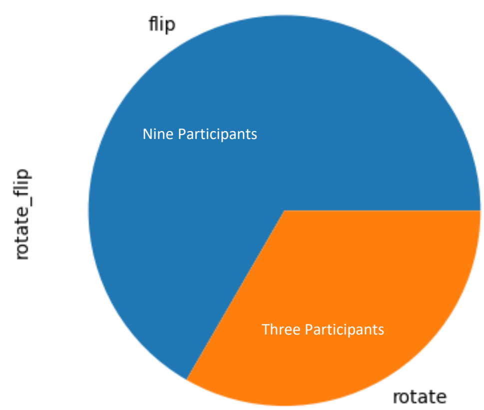
Figure 6.2(a) depicts the duration of the "about-to-brush" moments for each participant, ranging from 2.5 seconds to 14.5 seconds. The box plot showcases the person-to-person variability in these moments. Each box represents the median duration calculated from seven days' worth of data for each participant. The medians range from seven seconds to nine seconds, reflecting the varying lengths of time individuals spend in the "about-to-brush" phase.

Additionally, Figure 6.2(b) showcases the distribution plot of "about-to-brush" moments performed by all twelve participants. This plot provides an overview of the frequency and spread of these moments across the entire group. It offers insights into the collective behavior of the participants during this crucial phase of the toothbrushing process.

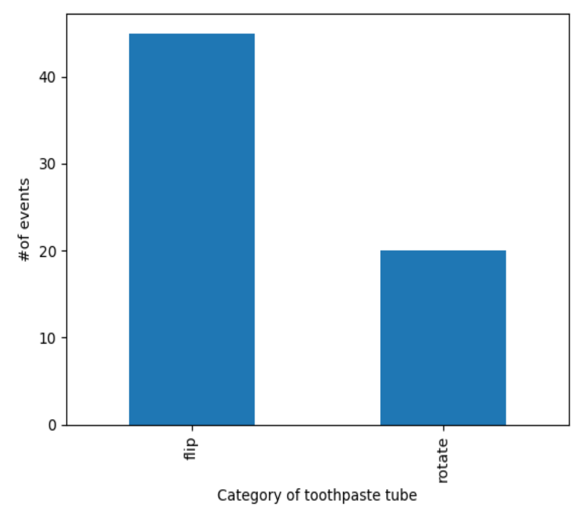
Together, these plots highlight the significant variability in the duration of the "about-to-brush" timeseries among the participants, underscoring the importance of considering individual differences in toothbrushing behavior.

Overall, these plots highlight the complex nature of the "about-to-brush" phase and the importance of recognizing and accommodating individual differences in toothbrushing behaviors. This understanding can inform personalized oral care strategies, enabling tailored recommendations and interventions to optimize oral hygiene practices for each individual.

6.5.2 Variability in Opening and Closing Toothpaste Tube Cap



((a)) Toothpaste tube open-close types by participants.



((b)) Distribution of toothpaste tube open-close type by brushing episodes.

Fig. 6.3: Variability of toothpaste tube open-close types.

We have analyzed 70 brushing episodes from 12 participants, specifically focusing on the "about-to-brush" moment. During this phase, we observed two different types of behaviors in opening and closing the toothpaste tube cap. Typically, individuals either rotate or flip the cap to accomplish this task. The distribution of these opening-closing types among the participants is shown in Figure 6.3(a).

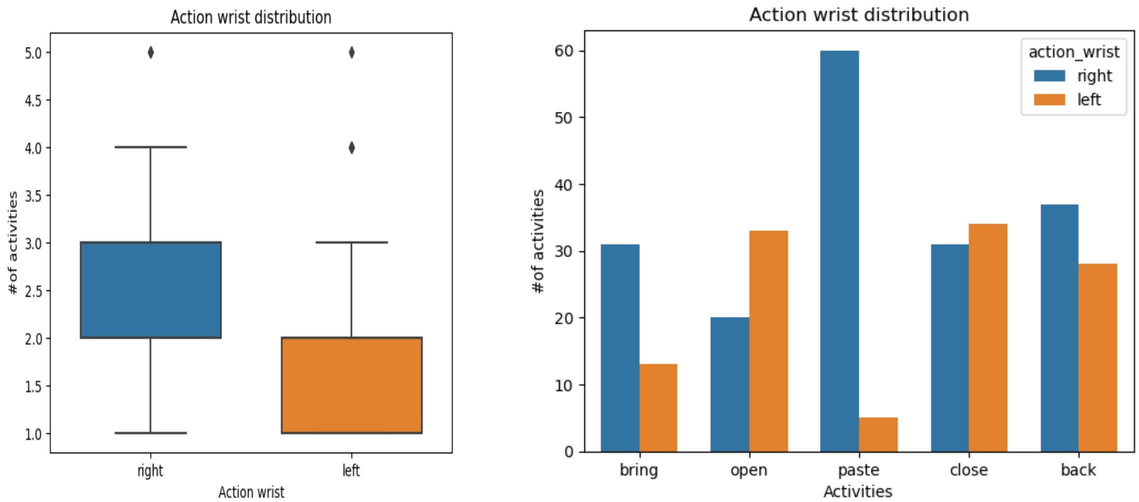
Interestingly, we found that only three out of the twelve participants preferred to rotate the toothpaste tube cap for both opening and closing. On the other hand, the remaining nine participants consistently chose to flip the cap for both actions. This finding indicates that individuals exhibit distinct preferences and habits when it comes to handling and accessing the toothpaste tube.

In addition, the distribution of opening and closing methods used during brushing episodes is shown in Figure 6.3(b). Out of the 70 recorded episodes, rotating the toothpaste tube cap was used in 20 instances, while flipping was used in the remaining 50 episodes. This distribution sheds light on the frequency of each method used in the overall brushing routine.

The observed behaviors surrounding the act of opening and closing the toothpaste tube cap just before brushing demonstrate a wide range of diversity. By comprehending these preferences and patterns, we can enhance oral care practices and create toothpaste packaging that accommodates and respects individual habits and preferences.

6.5.3 Variability in Using Wrist to Perform About-to-moment

Both wrist are active during about-to-moment. Naturally, people bring the toothpaste tube in one wrist but open the toothpaste tube cap by another wrist. So based on active wrist we have to detect the micro event of about-to-brush moment from that wrist signals. Figure 6.4(a) represents the distribution of two wrists during about-to-moment. We observed participants are doing much activities by using right wrist since we observed that all 12 participants are right dominant wrist. Since active wrist preserve data that which wrist is actively doing the micro event of about-to-brush moment, we plotted another figure 6.2(a) to represent the distribution of active wrist during each micro event. We observed most of the participants bring the toothpaste tube by right dominant wrist, then hold the toothpaste tube by right wrist and open the tube cap by left wrist. Most of the participants were using right wrist to put paste in the toothbrush. Similarly for closing the toothpaste tube cap, if participant hold the toothpaste tube right wrist then close the tube cap by left wrist. This above about-to-moment active wrist behaviors depicts in the figure 6.5



((a)) Distribution of active wrist ((b)) Distribution of active wrist for each micro-event

Fig. 6.4: Active wrist distribution during about-to-brush moment performance

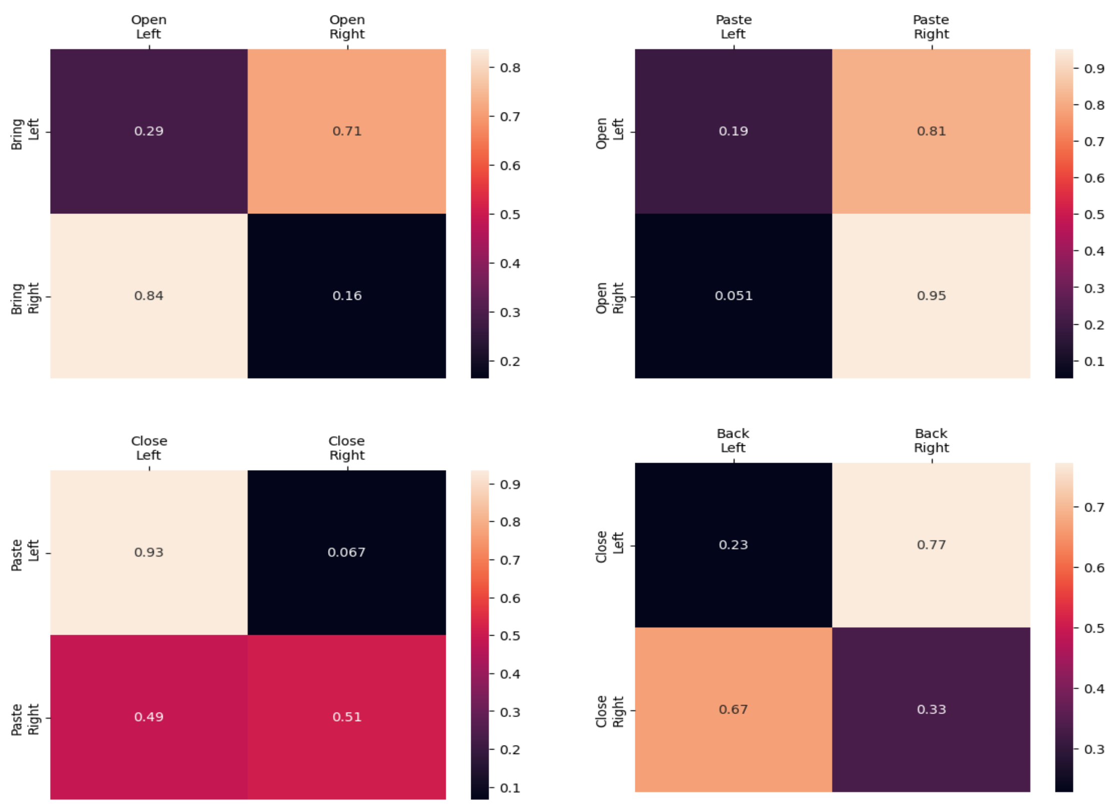


Fig. 6.5: Transition between wrist during the about-to-brush moment performance.

6.5.4 Feasibility of Detecting About-to-Brush Moment using Accelerometer Time-series

Figure 6.6 represents the signal of accelerometer during about-to-brush moment activity. During each of the activity which wrist is active that represents by bold line. For example, participant bring the toothpaste tube by left wrist, then hold the toothpaste tube by left wrist and open the tube cap by right wrist. Then using right wrist to put paste in the toothbrush. Similarly for closing the toothpaste tube cap, participant hold the toothpaste tube left wrist and close the tube cap by right wrist.

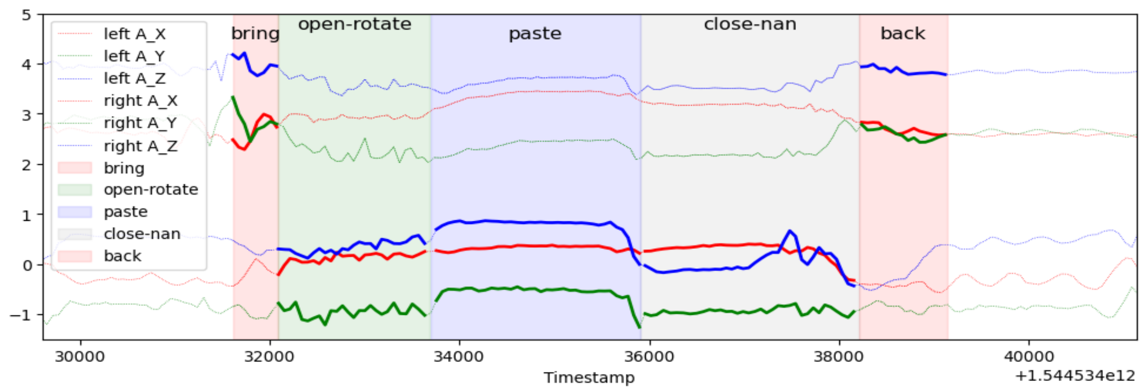


Fig. 6.6: Accelerometer signal for left and right wrist during the about-to-brush moment performance.

Overall, these findings provide valuable insights into the variability of toothbrushing behaviors, including the duration of the "about-to-brush" phase, preferences for opening and closing the toothpaste tube cap, and the use of wrists during the process. Understanding these individual differences can inform the development of personalized oral care strategies and the design of toothpaste packaging that accommodates diverse habits and preferences.

6.6 Model Development for Monitoring Brushing Behavior and Delivering Timely Interventions

We propose two approaches we design and develop to track tooth brushing behaviors and giving intervention. One approach is to detect about-to-brush moments to give intervention before starting of the tooth brushing behavior. The second approach is to detect initial portion of the tooth brushing behavior to give intervention after starting the tooth brushing behavior.

6.7 About-to-brush Moment Detection Model

In Section 6.2.1, we introduced a set of sequential micro-events that occur prior to the initiation of tooth brushing. Each micro-event is labeled and depicted in Figure 6.1. Our objective is to accurately detect and identify these micro-events using wrist-worn inertial sensors in real-world settings.

6.7.1 Challenges in Detecting Five Micro Events

Designing and detecting the five micro-events posed several challenges, which are outlined below:

- High variability in task performance: The tooth brushing task can be performed in various ways by different individuals, leading to a high level of variability in how each micro-event is executed.

- Short duration and lack of repetitions: Each micro-event occurs for a very brief duration, ranging from milliseconds to seconds, and there are no repeated instances of these micro-events within a single tooth brushing session.

- Hand switching and non-dominant hand dependence: The micro-events involve the use of different hands and frequent switching of wrists during the task. Additionally, not all micro-events are dependent on the dominant hand of the individual.

- Simultaneous actions: Some micro-events require performing two actions simultaneously, such as opening and bringing for flipping or closing and returning for flipping.

- Incomplete video recordings: The video recordings capturing these activities may be incomplete or missing certain segments. Some participants start recording only after applying the toothpaste, while others may fail to maintain a consistent timeline, forgetting to close the toothpaste tube after opening it and proceeding with brushing. Additionally, a percentage of participants may not exhibit certain micro-events in their recorded videos.

- Lack of unique patterns: In some micro-events, there is no unique visual pattern or signal to identify the event, except for the rotation of the cap.

- Limited data for certain cap types: Two types of toothpaste tube caps are rotating and flipping. Among the 12 participants, only three participants used the rotating cap to open and close the toothpaste tube. Consequently, the amount of available data for training a model is limited.

These challenges highlight the complexities involved in designing and detecting the micro-events accurately, necessitating innovative approaches and robust algorithms to overcome these hurdles.

6.7.2 Training data generation

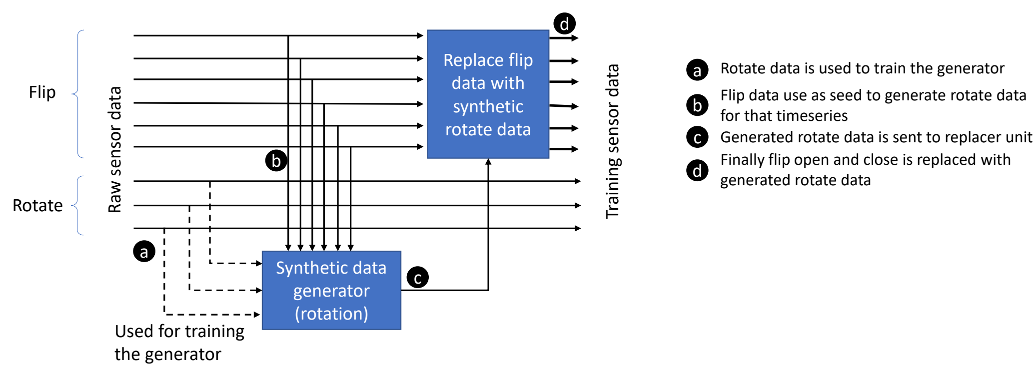


Fig. 6.7: Initially, we trained a generative model using rotation-only sensor data from toothpaste open/close actions. Subsequently, we replaced each toothpaste open/close with flip segment in the dataset with generated rotation-only sensor data. This augmented dataset was then utilized for further training.

Within our dataset, we have a limited amount of data from only three participants who performed the crucial action of rotating the toothpaste tube cap to open and close it. However, we possess a more substantial amount of data from nine participants who executed the action of flipping the cap. To overcome this discrepancy and enrich our dataset, we employed a clever technique. We utilized the data from these nine participants who performed the flipping action as a foundation and generated new data specifically for the rotating action. This was accomplished by substituting the original data with newly synthesized data from three additional participants who were tasked with rotating the cap. By incorporating these generated data points, we were able to capture the inherent variability and intricacies of the participants' movements during the rotation process. For a comprehensive overview of the entire data generation process, please refer to Figure 6.7. In the initial phase, a generative model was trained using rotation-only sensor data from toothpaste open/close actions. This model generated synthetic rotation-only sensor data that replicated the patterns observed during toothpaste open/close with flip segments. The generated rotation-only sensor data was then used to replace the toothpaste open/close with flip segments in the original dataset. This augmented dataset combined real and synthetic data, which was subsequently used for further training. By incorporating the generated data, we aimed to enhance the training process and improve the model's ability to capture the distinct characteristics and patterns of toothpaste open/close with flip actions.

6.7.3 Transformer-based Generative Model

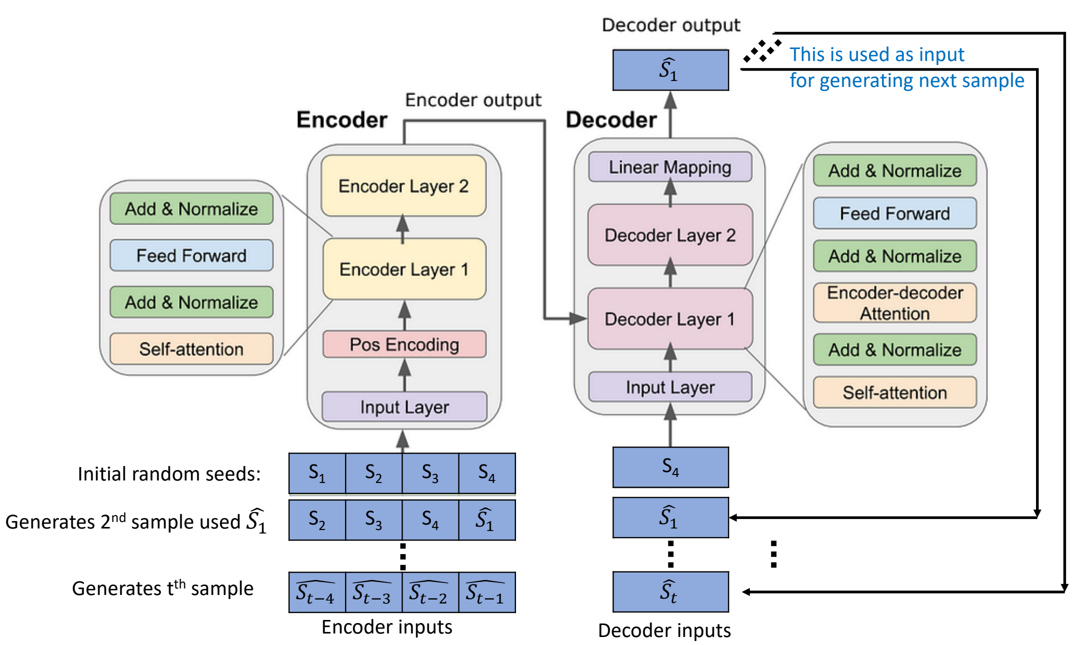


Fig. 6.8: Architecture of transformer-based synthetic sensor data generation.

The architecture Figure 6.8 of transformer-based synthetic sensor data generation encompasses a sophisticated framework designed to generate synthetic sensor data using the transformer model. This architecture leverages the power of transformers, which are highly effective in capturing intricate patterns and dependencies within sequential data. At its core, the architecture consists of multiple components working in synergy. The input to the system includes a combination of real sensor data and synthetic seed data. The real sensor data serves as a foundation, providing meaningful patterns and characteristics extracted from actual sensor measurements. The synthetic seed data acts as a starting point, injecting diverse and novel patterns into the generated data. The transformer model takes this combined input and applies its attention mechanism to learn complex relationships between the input sequences. The attention mechanism allows the model to selectively focus on different parts of the input, capturing dependencies and correlations between different time steps and sensor readings.

The learned representations from the transformer model are then fed into a generative component, which is responsible for generating the synthetic sensor data. This component utilizes the learned representations to generate data points that closely resemble real sensor measurements while incorporating novel patterns and variations.

To ensure the generated data aligns with the characteristics of real sensor data, a feedback loop is established. The generated data is compared with the real sensor data, and a loss function is employed to quantify the similarity between the two. This loss is then backpropagated through the architecture, allowing the model to iteratively refine its generation process.

The architecture is trained using a dataset comprising both real sensor data and corresponding ground truth labels. This training process enables the model to learn and generalize the underlying patterns, enabling it to generate realistic synthetic sensor data. Overall, the architecture of transformer-based synthetic sensor data generation represents a powerful approach for generating high-quality synthetic sensor data that closely mimics real-world sensor measurements. It combines the strength of transformers in capturing complex dependencies with generative components to produce data that exhibits meaningful patterns and variability.

6.7.4 Performance of Synthetic Data Generation

In this section, we evaluate the performance of our synthetic data generation approach in replicating the characteristics and patterns of sensor data during the opening/closing of a toothpaste cap with rotation. Our goal is to assess the effectiveness of the synthetic data in accurately capturing the behavior and dynamics of the real-world sensor data.

To assess the performance of our synthetic data generation approach, we conducted a comparison between the synthetic data and the ground truth real data obtained from actual sensor recordings. This evaluation involved employing several metrics, including the mean square error (MSE) and root mean square error (RMSE), to quantify the similarity and reliability between the synthetic and real data.

To ensure an unbiased evaluation, we split the data into training and testing sets, with 70% of the data used for training and the remaining 30% used for testing. The model was trained using the MSELoss loss function and optimized with the Adam optimizer.

Table 6.1: Test results

Metric	**Value**
Mean Squared Error (MSE)	0.0005
Root Mean Squared Error (RMSE)	0.0229

Our evaluation of the synthetic data generation approach yielded promising results (Table 6.1). The mean squared error (MSE) between the synthetic data and the ground truth real data was found to be 0.0005, indicating a very low level of deviation. Similarly, the root mean squared error (RMSE) was measured at 0.0229, further confirming the high similarity and accuracy of the synthetic data in replicating the real sensor data.

These results demonstrate the effectiveness of our approach in generating synthetic data that closely aligns with the characteristics and patterns observed in the actual sensor recordings. The low MSE and RMSE values signify the reliability and precision of the

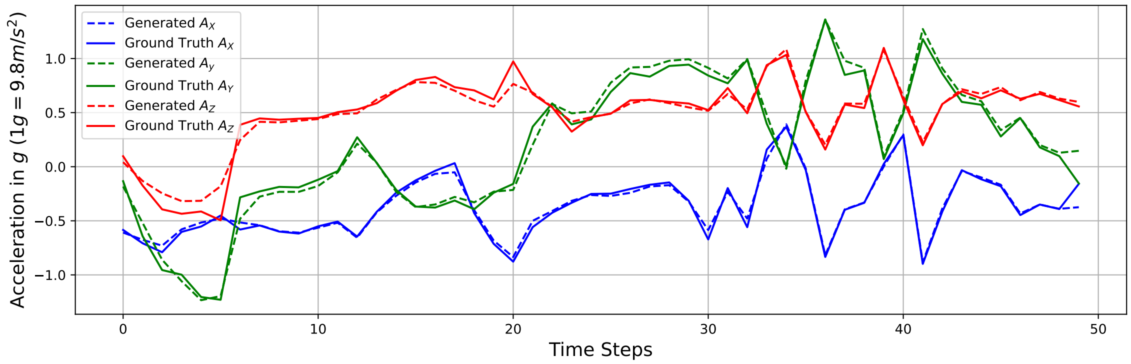


Fig. 6.9: The figure illustrates the comparison between the actual and predicted signals. Each predicted signal at time t is generated based on the model's observation of the actual signal from time 0 to $(t-1)$. The graph shows that the predicted signal closely aligns with the actual signal, demonstrating a high accuracy and similarity between the two.

generated synthetic data, making it a valuable tool for various applications and analyses in the field of sensor data processing and modeling.

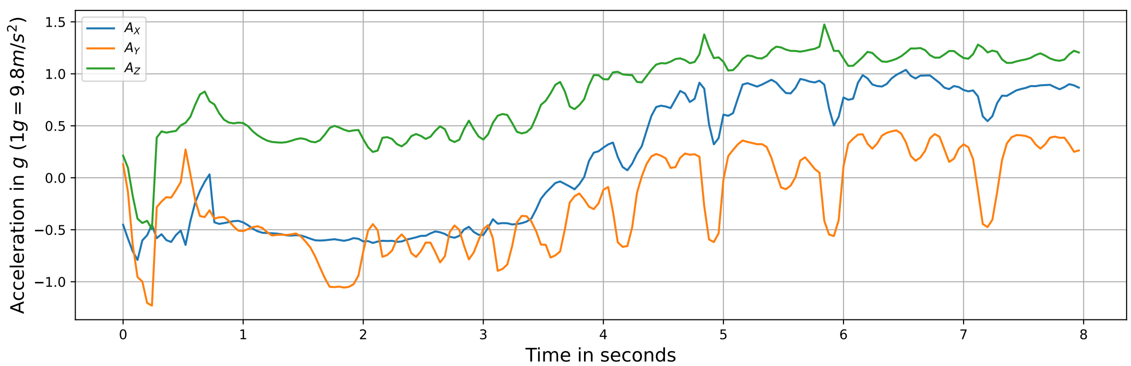


Fig. 6.10: Generated signal is shown here from random seeds.

Figure 6.10 displays the generated signal using random seeds. Each seed corresponds initial values of the rotation signal used to generate the signal. The plot provides a visual representation of the resulting signal obtained from these random seeds.

6.7.5 LSTM-based Sequence-to-sequence Model for Detecting About-to-brushing Moment

The LSTM-based Sequence-to-sequence model is designed to detect the “about-to-brushing" moment based on sensor data. This model leverages the capabilities of Long

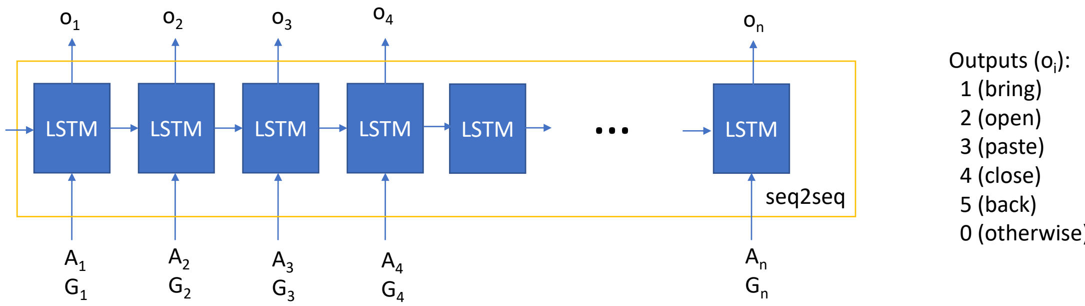


Fig. 6.11: LSTM-based sequence-to-sequence model for detecting about-to-brushing moment.

Short-Term Memory (LSTM) networks, which are well-suited for capturing sequential patterns in time series data.

The model takes as input a sequence of sensor readings and aims to predict the occurrence of the “about-to-brushing" moment, which refers to the specific moment before the initiation of the brushing action. By accurately detecting this moment, the model can provide timely prompts or notifications to users, enhancing their oral hygiene habits.

The sequence-to-sequence architecture allows the model to learn the temporal dependencies and patterns inherent in the sensor data. It consists of an encoder network that encodes the input sequence into a fixed-length representation and a decoder network that generates predictions based on the encoded information. The LSTM units within the model enable it to capture long-term dependencies and make informed predictions about the occurrence of the “about-to-brushing" moment.

The LSTM-based Sequence-to-sequence model generates outputs, denoted as o_t, at each time step. Each timestamp is assigned a specific label (an integer value), indicating the predicted action or state based on the input sensor data. The following are the labels and their corresponding interpretations:

- 1 (for ‘bring’): signifies that the model predicts the user is in the process of bringing the toothpaste container.

- 2 (for 'open'): indicates that the model predicts the user is opening the toothpaste container.
- 3 (for 'paste'): suggests that the model predicts the user is applying toothpaste to the toothbrush.
- 4 (for 'close'): indicates that the model predicts the user is closing the toothpaste container.
- 5 (for 'back'): suggests that the model predicts the user is moving the toothpaste container away.
- 0 ('otherwise'): represents any other action or state that does not fall into the above categories. It indicates that the model does not predict the occurrence of the "about-to-brushing" moment based on the given sensor data.

6.7.6 Detection Performance of About-to-Brush Moment

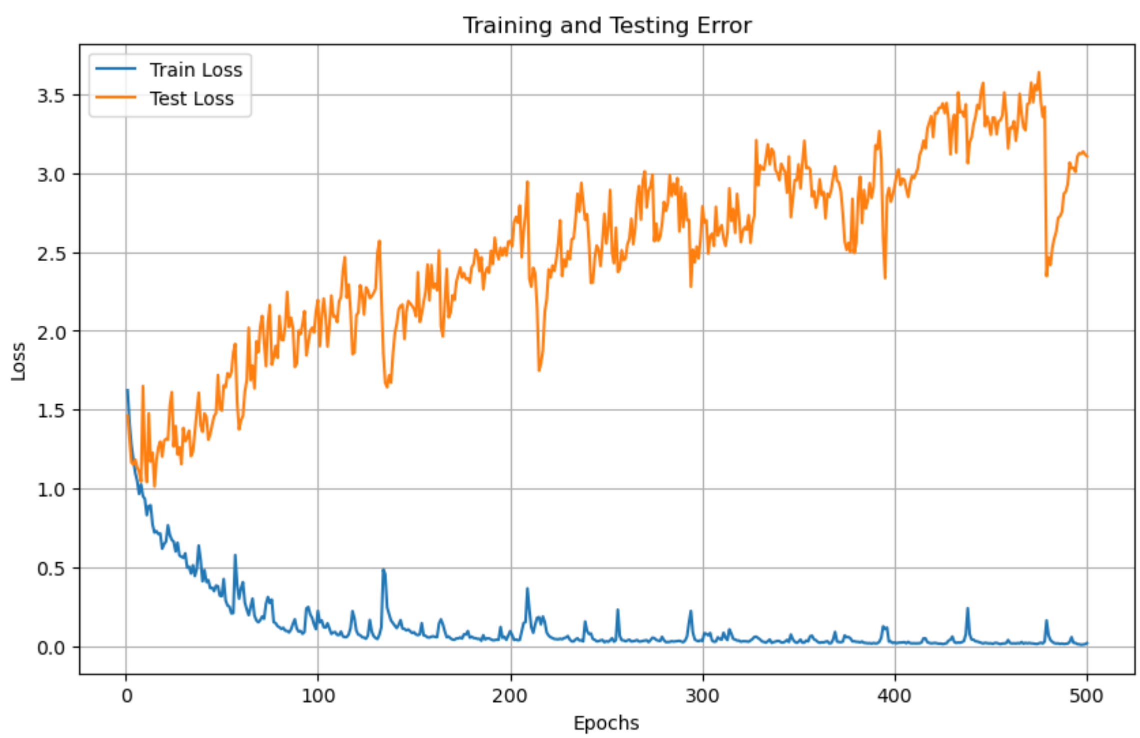


Fig. 6.12: Training and testing loss.

The observed trend of decreasing training loss over epochs but increasing testing error raises an interesting discussion about the generalization and overfitting of the model.

The decreasing training loss indicates that the model effectively learns and fits the training data. As the model iteratively adjusts its parameters during training, it becomes more adept at minimizing the error between the predicted outputs and the ground truth labels in the training set. This reduction in training loss suggests that the model is capturing the underlying patterns and relationships present in the training data.

However, the increasing testing error suggests that the model's performance is deteriorating when presented with unseen data. This phenomenon is known as overfitting, where the model becomes too specialized and fails to generalize well to new examples. Overfitting occurs when the model becomes overly complex and starts to memorize the training data instead of learning the underlying patterns that can be applied to unseen data.

Several factors can contribute to the observed overfitting. One possible explanation is that the model is becoming too complex relative to the size and diversity of the training data. When the model has a large number of parameters, it can easily adapt to noise and idiosyncrasies in the training data, resulting in poor performance on new, unseen data.

Another factor could be the lack of regularization techniques employed during training. Regularization methods, such as L1 or L2 regularization, can help prevent overfitting by adding penalties to the model's loss function, discouraging excessive parameter values and promoting simpler models.

Insufficient or unrepresentative testing data can also contribute to the increasing testing error. If the testing set does not adequately reflect the true distribution of the target population, the model may struggle to generalize its learnings to new examples.

Several strategies can be employed to address the issue of overfitting and improve generalization. One approach is collecting more diverse and representative data for training and testing. Increasing the size and variability of the dataset can help the model learn more robust and generalizable patterns.

As mentioned earlier, regularization techniques can also be employed to introduce constraints and prevent the model from overfitting. By balancing the complexity of the model and its ability to generalize, regularization methods can improve performance on unseen data.

Furthermore, fine-tuning the model architecture and hyperparameters can also help mitigate overfitting. Simplifying the model by reducing its complexity, adjusting learning rates, or exploring different optimization algorithms can lead to improved generalization performance.

In conclusion, the observed phenomenon of decreasing training loss and increasing testing error indicates the presence of overfitting in the model. Understanding the factors contributing to overfitting and employing appropriate strategies, such as regularization and data augmentation, can help address this issue and improve the model's generalization performance.

6.8 Alternate Intervention Opportunities: Brushing Initiation Moments

Initiating the toothbrushing process involves the crucial task of determining the precise moment when an individual commences brushing their teeth. This prompts the question: how early in the brushing episode can we accurately detect this initiation? Detecting the initial few brush strokes can play a significant role in triggering timely interventions during the early stages of brushing, thereby potentially enhancing the overall detection accuracy. Although an existing stroke detection model called mTeeth [30] already exists, it is not directly applicable to our specific problem. The mTeeth stroke detection model primarily focuses on identifying toothbrushing surfaces during dedicated toothbrushing time, ensuring a high level of precision. However, for our intervention problem, a different approach is required. We need a new stroke detection model capable of effectively identifying toothbrushing activity from the comprehensive time series data spanning an entire day.

6.9 Tooth Brushing Stroke

During brushing, individuals typically engage in a back-and-forth or up-and-down periodic motion with the toothbrush. This distinctive movement pattern, referred to as a *brushing stroke*, serves as our anchor event for surface identification.

Among the various types of brushing strokes described in the literature, such as circular strokes [70], we have noticed that brushing strokes predominantly follow an up-down-up or back-forth-back motion. These two primary periodic movements are consistently observed during brushing activities.

6.9.1 Event-based Segmentation for Brushing Stroke Detection

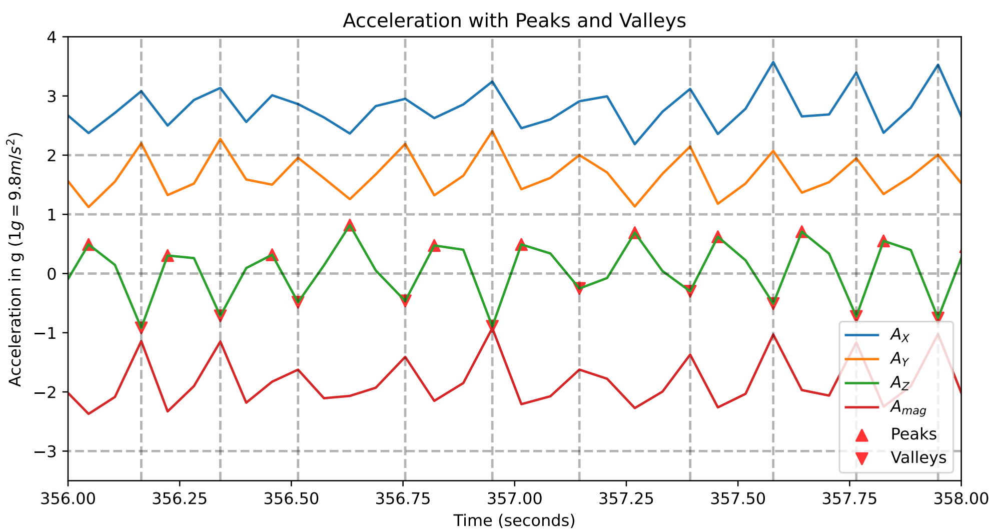


Fig. 6.13: Brushing Strokes.

To effectively identify brushing strokes from sensor data, we leverage the characteristic pattern of periodic up-and-down or back-and-forth movement captured by the wrist-worn accelerometer signal. We define a brushing stroke as a cycle of peak-valley-peak combinations, representing the characteristic motion during toothbrushing.

Figure 6.13 illustrates plots of the accelerometer signal along three axes and the accelerometer's magnitude during a small toothbrushing segment. In the time series of the signal, we identify peaks as the points where the signal reaches its maximum, and valleys as the points where the signal reaches its minimum. These peaks and valleys are marked with red up-pointing and down-pointing triangles. Let $\mathcal{P} = \{p_1, p_2, ..., p_n\}$ represent the set of peaks, and $\mathcal{V} = \{v_1, v_2, ..., v_n\}$ represent the set of valleys. By accurately detecting these peaks and valleys using a peak-valley detection algorithm, we can define each brushing stroke as a combination of a valley, peak, and the subsequent valley, denoted as $S_i = \langle v_i, p_i, v_{i+1} \rangle$.

In Figure 5.16, we can observe a series of peaks and valleys in all four signals, representing the motion patterns during toothbrushing. However, these peaks and valleys are not temporally aligned across the signals. Out of the four signals, we need to select the one that can best represent the start and end times of the brushing strokes.

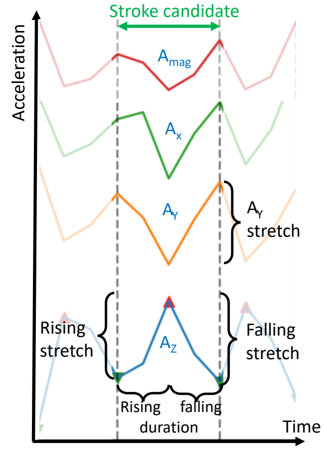
Upon analysis, we find that the magnitude signal, which incorporates information from all three axes, is not suitable for this purpose due to the lack of synchronized alignment across the axes. However, we notice that the peak-valley cycles from the Z-axis signal exhibit better alignment with the ground truth obtained from video recordings.

Based on this observation, we consider the segments created by the Z-axis signal as candidate windows for determining the optimal start and end times of the brushing strokes. By focusing on the Z-axis signal, we can effectively capture the temporal alignment of the peak-valley cycles, enhancing the accuracy of stroke detection and ensuring alignment with the ground truth information from video recordings.

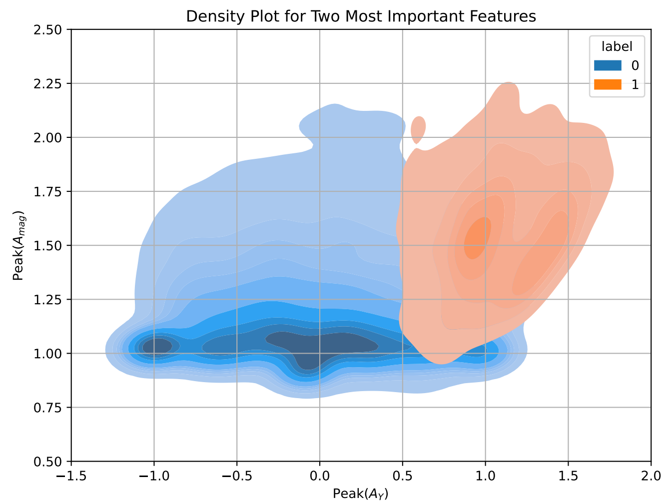
6.9.2 Feature Extraction and Selection

In our analysis, we consider several features related to the characteristics and dynamics of brushing strokes. These features provide insights into different aspects of the stroke, such as its amplitude, duration, and rate of change. Below is a description of each feature computed from the Z-axis of the accelerometer signal (Figure 6.14(a)):

- **rising stretch**: This feature represents the stretch of the rising phase of a stroke, which is calculated as the difference between the amplitude of the peak and the amplitude of the preceding valley.



((a)) Stroke-wise computed features from the accelerometer signal.



((b)) Density Plot for Two Most Important Features

Fig. 6.14: Stroke-wise computed features from the accelerometer signal.

- **falling stretch**: Similarly, this feature represents the stretch of the falling phase of a stroke, calculated as the difference between the amplitude of the peak and the amplitude of the subsequent valley.

- **stretch diff**: This feature captures the difference between the rising and falling stretches of a stroke, providing insights into the asymmetry of the stroke's amplitude profile.

- **stretch ratio**: The stretch ratio feature represents the ratio between the rising and falling stretches of a stroke, indicating the relative magnitude of the two phases.

- **rate of change**: This feature describes the rate of change of the stroke's amplitude, providing information about the speed or intensity of the brushing motion.

- **rising duration**: This feature represents the duration of the rising phase of a stroke, indicating the time taken for the amplitude to increase from the valley to the peak.

- **falling duration**: Similarly, this feature represents the duration of the falling phase of a stroke, indicating the time taken for the amplitude to decrease from the peak to the subsequent valley.

- **duration**: This feature captures the overall duration of a complete brushing stroke, including both the rising and falling phases.

We consider additional features specific to the $X-$axis, Y-axis, and magnitude signals to capture the characteristics of brushing strokes. Since there is no guarantee of having a valley, peak, and another valley for these signals, we focus on amplitude-based features. The following features are considered for $X-$axis, Y-axis, and magnitude signals:

- **valley**: This feature represents the minimum amplitude value within the segment, indicating the lowest point reached during the brushing stroke.
- **peak**: Similarly, this feature represents the maximum amplitude value within the segment, indicating the highest point reached during the brushing stroke.
- **avg**: This feature represents the average amplitude value within the segment, providing insights into the overall magnitude of the brushing motion.
- **stretch**: This feature captures the difference between the maximum and minimum amplitude values within the segment, indicating the range or extent of the brushing stroke.

By analyzing these features, we can gain a comprehensive understanding of the dynamics and characteristics of brushing strokes. These features provide valuable insights into the amplitude variations, ranges, and overall patterns of the brushing motion. By leveraging this information, we can effectively distinguish brushing strokes from other activities or gestures captured by the sensor data. This enables us to develop robust and accurate algorithms for brushing stroke detection and classification.

Important Features: Using the *random forest* algorithm, we can identify the most important features that contribute significantly to the classification of brushing strokes. As

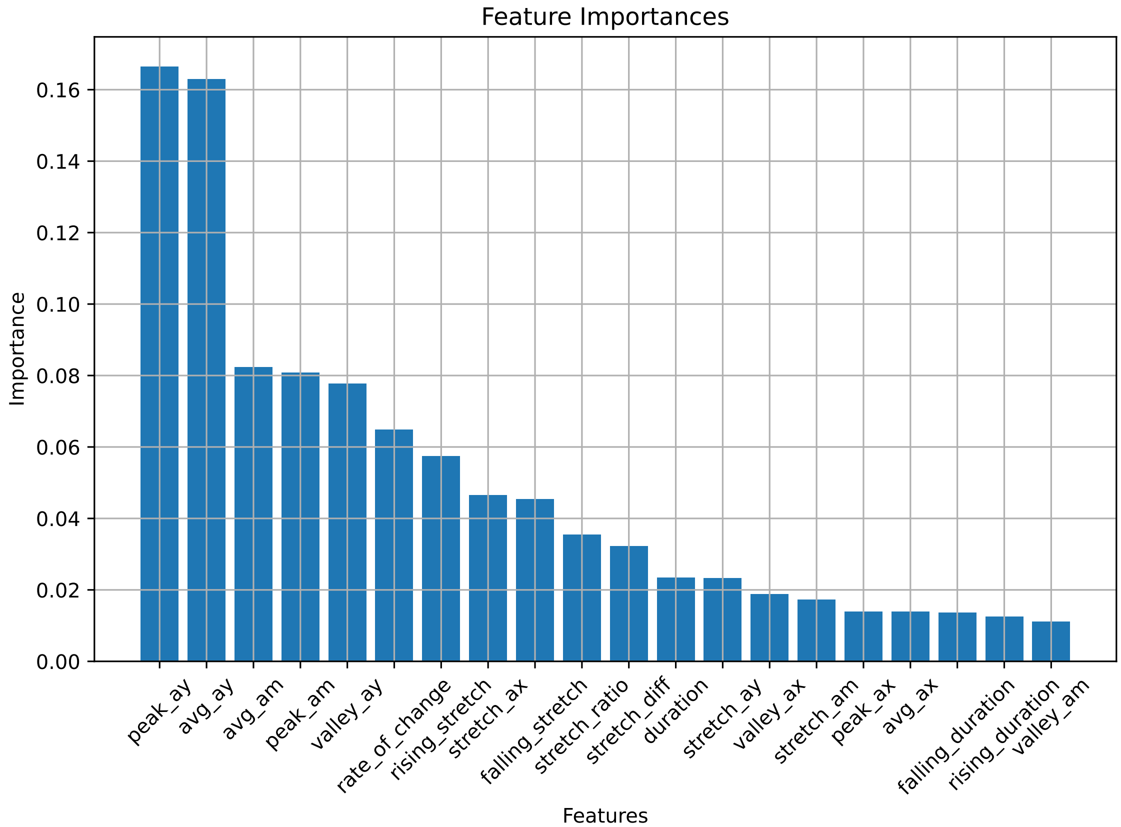


Fig. 6.15: Important features for brushing strokes classification.

shown in Figure 6.15, by analyzing the importance scores assigned to each feature by the random forest model, we can determine which features have the greatest impact on accurately distinguishing brushing strokes from other activities. This information is crucial for feature selection and model optimization, as it allows us to focus on the most informative and discriminative features for improving the overall classification performance.

6.9.3 Brushing Stroke Model Development

In the model development phase, we explored the performance of various machine learning models for the task of brushing stroke detection. The models we considered included Logistic Regression, Decision Tree Classifier, Random Forest Classifier, Gradient Boosting Classifier, and MLP Classifier.

Each model offers its own strengths and characteristics, making it important to evaluate their performance on our specific task. **Logistic Regression** is a linear model that provides interpretable results and is commonly used in binary classification problems. **Decision Tree Classifier**, on the other hand, is a non-linear model that learns hierarchical decision rules. **Random Forest Classifier** is an ensemble model that combines multiple decision trees to improve performance and reduce overfitting. **Gradient Boosting Classifier** is another ensemble model that sequentially trains weak learners to create a strong classifier. MLP Classifier, which stands for **Multi-Layer Perceptron**, is a type of neural network model that can capture complex patterns and relationships in the data.

By considering a diverse range of models, we aimed to find the one that best suits our brushing stroke detection task. Each model was trained and evaluated using a 70% training and 30% testing data split. This allowed us to assess their performance and determine which model yielded the most accurate and reliable results.

In the subsequent sections, we present the evaluation metrics, including precision, recall, and F1-score, for each of the considered models. These metrics provide insights into the models' ability to classify brushing strokes and distinguish them from other activities correctly. By analyzing these results, we can identify the model that demonstrates superior performance and select it for further analysis and application in brushing stroke detection.

6.9.4 Performance of Brushing Stroke Detection

In this section, we present the results of our experiments on training and evaluating various classification models using the computed features from the previous section. We considered several models, including Logistic Regression, Decision Tree Classifier, Random Forest Classifier, Gradient Boosting Classifier, and MLP Classifier.

To assess the performance of these models, we divided the data into a training set consisting of 70% of the samples and a testing set consisting of the remaining 30% of the samples. The models were then trained on the training data and their hyperparameters were tuned for optimal performance.

After training, we applied the models to the testing data and evaluated their performance. Figure 6.16 illustrates the results obtained from the testing phase, providing insights

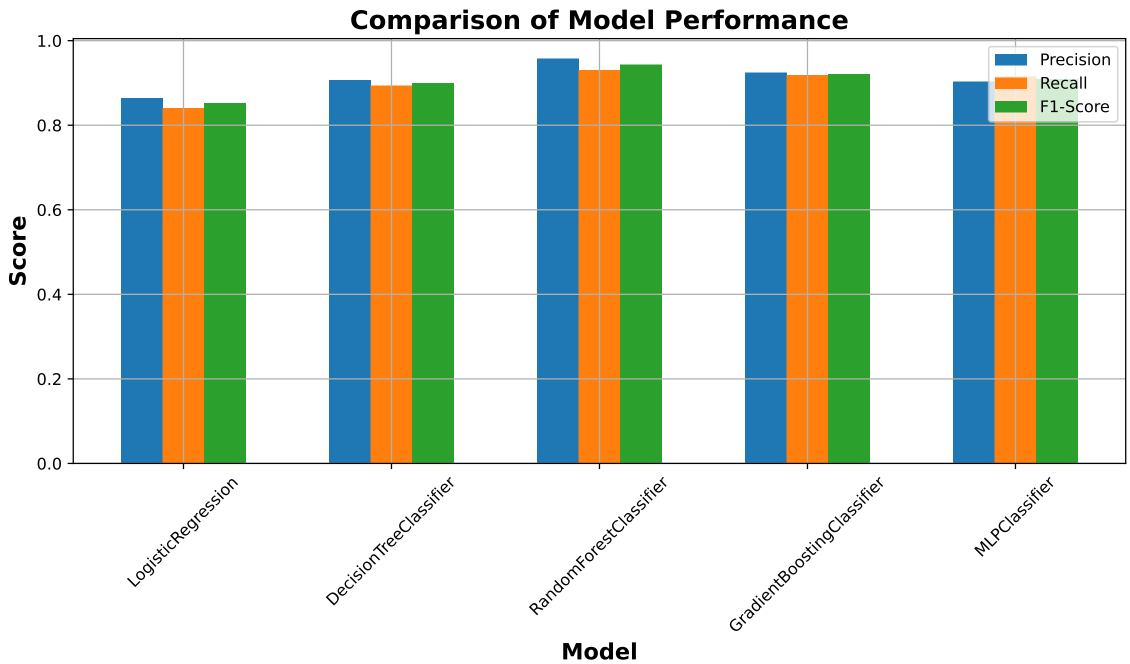


Fig. 6.16: Important features for brushing strokes classification.

into the accuracy and effectiveness of the classification models in distinguishing brushing strokes from other activities.

Looking at the precision metric, which measures the accuracy of positive predictions, we observed that the Random Forest Classifier achieved the highest precision score of 0.9570. This indicates that the model had a low rate of false positive predictions, correctly identifying a significant portion of brushing strokes. The Decision Tree Classifier also demonstrated high precision with a score of 0.9063, indicating its ability to make accurate positive predictions.

In terms of recall, which measures the ability to correctly identify positive instances, the Random Forest Classifier achieved a recall score of 0.9305, indicating its ability to capture a high proportion of actual brushing strokes. The Gradient Boosting Classifier also performed well in terms of recall, achieving a score of 0.9184.

The F1-score, which is the harmonic mean of precision and recall, provides a balanced measure of a model's overall performance. The Random Forest Classifier exhibited the highest F1-score of 0.9436, followed closely by the Gradient Boosting Classifier with a score of 0.9212. These results indicate that both models achieved a good balance between precision and recall, indicating their overall effectiveness in accurately identifying brushing strokes.

Overall, the results demonstrate the effectiveness of the trained classification models in distinguishing brushing strokes from other activities based on the computed features. The Random Forest Classifier emerged as the top-performing model, showcasing its ability to achieve a high level of precision, recall, and F1-score. These findings highlight the potential of using machine learning algorithms in automatic brushing stroke detection, which can contribute to improving oral hygiene monitoring and dental health assessment.

6.10 Tooth Brushing Initiation Detection Model

Initiating the toothbrushing process involves the crucial task of determining the precise moment when an individual commences brushing their teeth. This prompts the question: how early in the brushing episode can we accurately detect this initiation? Detecting the initial few brush strokes can play a significant role in triggering timely interventions during the early stages of brushing, thereby potentially enhancing the overall detection accuracy. Although an existing stroke detection model called mTeeth [30] already exists, it is not directly applicable to our specific problem. The mTeeth stroke detection model primarily focuses on identifying toothbrushing surfaces during dedicated toothbrushing time, ensuring a high level of precision. However, for our intervention problem, a different approach is required. We need a new stroke detection model capable of effectively identifying toothbrushing activity from the comprehensive time series data spanning an entire day.

6.10.1 Algorithm for Detecting Tooth Brushing Initiation

The Detection of Brushing Prompt Events utilizes the Stroke Detection and Clustering algorithm. This algorithm aims to identify brushing strokes from a raw accelerometer signal and group them into clusters based on their temporal proximity. By utilizing a trained stroke detection model, segments of the signal are classified as brushing strokes or non-strokes. The algorithm then applies a clustering technique to group the detected strokes into distinct brushing events, as well as other similar behaviors that produce signals similar to brushing strokes.

Additionally, the algorithm determines the minimum number of strokes that reduces the number of false alarms. It achieves this by considering clusters that have at least this minimum number of strokes. This optimization step helps to ensure that only significant clusters with a sufficient number of strokes are considered as valid brushing prompt events. By applying this threshold, the algorithm can accurately identify and distinguish true brushing prompt events from other activities captured by the accelerometer signal.

Algorithm 2 Stroke Detection and Clustering

1: Train a stroke detection model, M, to classify segments as brushing strokes or non-strokes.
2: Find peaks and valleys from the raw accelerometer signal: $V_1, P_1, V_2, P_2, V_3, P_3, \ldots$.
3: Let C be the set of detected stroke candidates, where each candidate $c_i = (V_{i-1}, P_i, V_i)$.
4: **for** each candidate c_i in C **do**
5:     Classify the candidate as a stroke or non-stroke using model M: $s_i = M(c_i)$.
6: **end for**
7: Group the detected strokes into clusters if the time difference between consecutive strokes is below a specified threshold.
8: Mark each cluster as a true brushing stroke.
9: Determine the minimum number of strokes threshold that minimizes both the detection time of true brushing events and false alarms.

The algorithm consists of the following steps:

1. **Training the Stroke Detection Model**: A stroke detection model, denoted as M, is trained using labeled data to classify segments of the accelerometer signal as brushing strokes or non-strokes. The model learns to differentiate the patterns associated with brushing strokes from other activities.

2. **Peak and Valley Extraction**: Peaks and valleys are extracted from the raw accelerometer signal. Peaks represent the maximum values, and valleys represent the

minimum values in the signal. These peaks and valleys provide essential landmarks for identifying potential brushing strokes.

3. **Stroke Candidate Detection**: Based on the extracted peaks and valleys, stroke candidates are identified. Each stroke candidate, denoted as c_i, consists of a sequence of a valley, peak, and subsequent valley, representing a potential brushing stroke.

4. **Stroke Classification**: The stroke detection model M is applied to each stroke candidate c_i to classify it as a stroke or non-stroke. The classification output, denoted as s_i, determines whether a candidate is identified as a brushing stroke or not.

5. **Stroke Clustering**: The detected strokes are grouped into clusters based on their temporal proximity. If the time difference between two consecutive strokes is below a specified threshold, they are considered part of the same brushing event cluster. Clustering helps to separate distinct brushing events and consolidate strokes that occur closely together in time.

6. **Identification of True Brushing Strokes**: Each cluster represents a true brushing stroke event. By marking the clusters as true brushing strokes, we obtain a collection of distinct brushing events from the raw accelerometer signal.

7. **Optimization of Strokes Threshold**: A minimum number of strokes threshold is determined to optimize the algorithm's performance. This threshold minimizes both the detection time of true brushing events and false alarms, ensuring accurate and reliable identification of brushing strokes.

Following these steps, the Stroke Detection and Clustering algorithm effectively identifies brushing strokes from the raw accelerometer signal, clusters them into distinct events, and provides valuable insights into brushing behavior patterns. Figure 6.17 illustrated the brushing prompt detection.

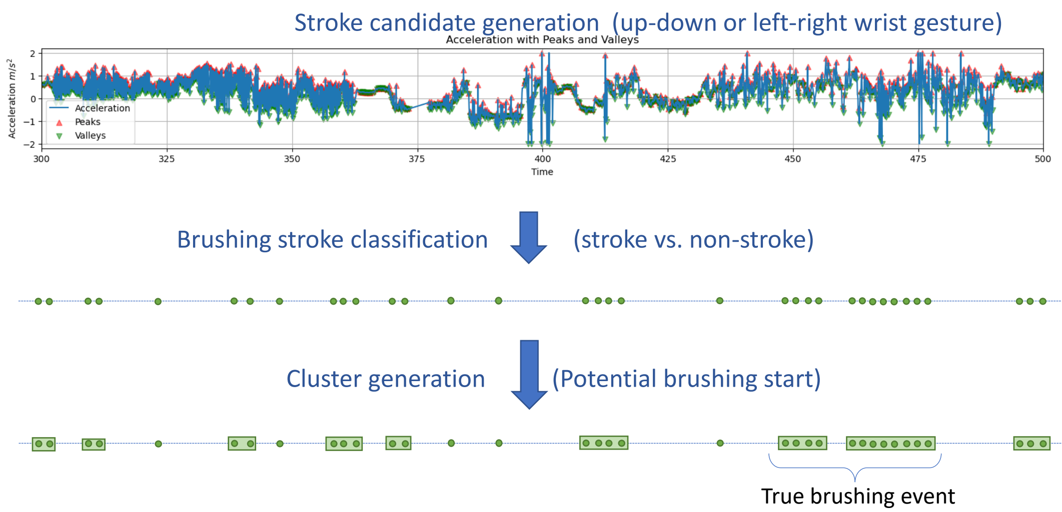


Fig. 6.17: Illustration of brushing prompt detection.

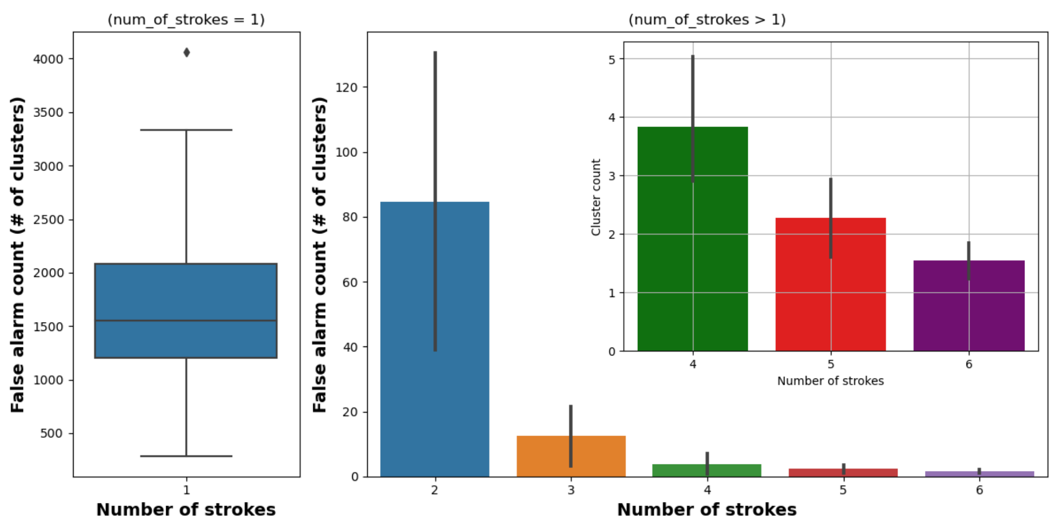


Fig. 6.18: Anticipated false positive rate for stroke count

6.10.2 Projected Count of False Alarm per Day

Figure 6.18 provides valuable insights into the occurrence of false brushing prompts per day. It presents a comprehensive overview of the relationship between the number of strokes and the corresponding number of clusters generated. When examining the scenario where the number of strokes is one, a significant number of clusters, exceeding 3000, are observed. This suggests a high frequency of false prompts in this particular case. Moving on to the case of two strokes, the number of clusters decreases to 130, indicating a notable reduction in false prompts. Further analysis reveals that when the number of strokes increases to three, the number of false prompts decreases even further to approximately 20. This downward trend continues, as the number of false prompts drops to five, three, and two when considering four, five, and six strokes, respectively.

It is important to note that individuals typically engage in toothbrushing once or twice a day, indicating that setting the threshold at six strokes is a reasonable choice. By doing so, we ensure that the number of interventions required does not exceed twice a day, providing a practical and manageable approach. This insight allows for the development of a more efficient and effective brushing prompt system, reducing unnecessary interventions while still ensuring timely reminders for best oral hygiene practices.

6.10.3 Anticipated Time of Detection Since Brushing Commencement

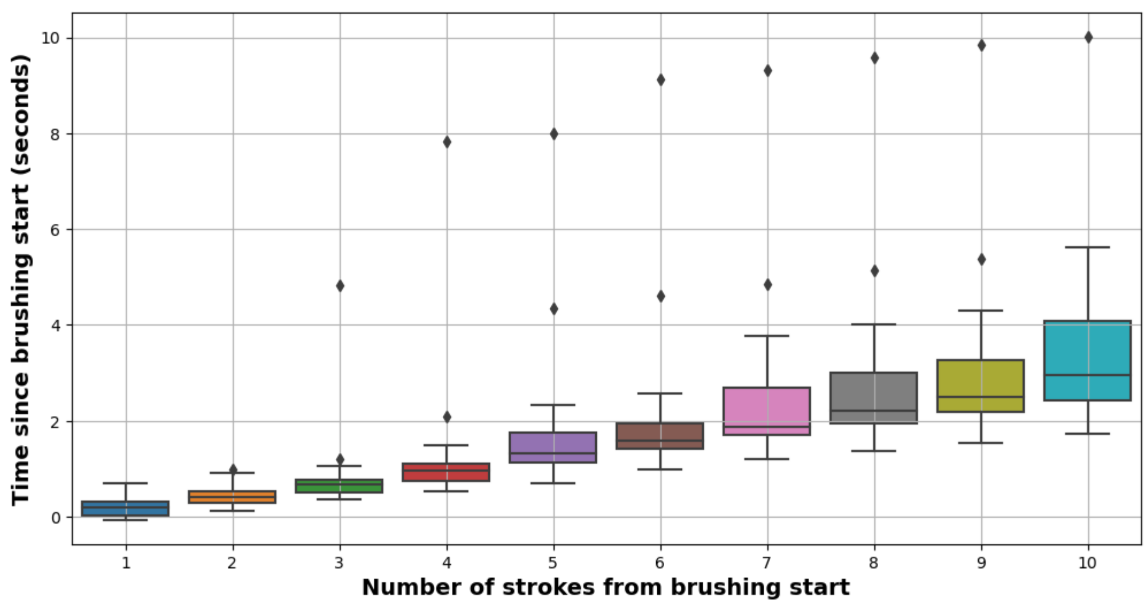


Fig. 6.19: Expected detection time based on the number of strokes

In the preceding section, we established that when there are six strokes, the number of interventions does not exceed two times in a day. Now, let us delve into the aspect of detection speed for these six strokes. The visualization in Figure 6.19 provides valuable insights into the time it takes to detect the strokes based on the number of strokes performed.

When there is only one stroke, we can expect the detection to occur in no more than one second. Similarly, for two, three, and four strokes, the detection time is estimated to be less than two seconds. As the number of strokes increases to five and six, the expected detection time remains below three seconds. This implies that the timing of interventions can be accomplished within a brief span of time, requiring less than three seconds. It is worth noting that this duration is significantly shorter when compared to the 30-second interventions typically associated with smart toothbrushes. This expedited detection time allows for more timely and efficient interventions to enhance oral care practices.

6.11 Discussion

In this chapter, we thoroughly examine the limitations associated with two distinct approaches: the about-to-brush moment detection approach and the brushing-initiation moment detection approach. Our objective is to shed light on the challenges and constraints faced when employing these methods to provide intervention to individuals regarding their toothbrushing habits.

We critically analyze the shortcomings inherent in the about-to-brush moment detection approach, which focuses on identifying the optimal time just before an individual starts brushing their teeth. We explore the potential issues, such as inaccuracies in detection or difficulties in capturing the precise moment when brushing commences. We delve into the challenges of accurately identifying this precise moment, including potential discrepancies in data collection or variations in individuals' brushing techniques. By understanding these limitations, we propose an alternative strategy which is brushing-initiation to enhance intervention accuracy. This method centers on detecting the exact instant when an individual initiates the act of toothbrushing. Similarly, we investigate the limitations of the brushing-initiation moment detection approach.

Furthermore, this chapter goes beyond discussing limitations and presents a comprehensive comparison between the two proposed approaches. By thoroughly comparing and contrasting these methods, we can gain valuable insights into their respective strengths and weaknesses. Through a thorough comparison of proposed approaches, it aims to contribute to the development of more effective interventions that provide individuals with valuable information and guidance regarding their toothbrushing routines.

6.11.1 Limitations of About-to-brush Moment Detection Approach

The accuracy of the Leave-one-subject-out cross-validation for about-to-brush moment is less than 50%. This suggests that the approach's performance is relatively poor when applied to unseen data from different subjects. Such low accuracy undermines the approach's reliability and effectiveness in real-world scenarios. In order to understand the reasons behind this low accuracy, let's delve into the detailed explanations below:

- **Overfitting Due to Limited Data:** The number of available data points for training and evaluation is significantly small. This scarcity of data increases the risk of overfitting, where the model becomes too closely tailored to the limited dataset, resulting in poor generalization to new, unseen instances. To mitigate overfitting, a larger and more diverse dataset is necessary to enhance the approach's performance.

- **Short Duration and Lack of Repetitions in About-to-brush Moments:** The about-to-brush moments tend to be brief in duration, making them challenging to capture accurately. Additionally, the lack of repetitions or consistent patterns in the about-to-brush moment signal further complicates the detection process. These characteristics hinder the development of robust algorithms that can reliably identify and classify these micro events within the toothbrushing routine.

- **Person-to-Person and Episode-wise Variability:** There is a high degree of variability observed between individuals in terms of their toothbrushing behavior and performance. Moreover, even within the same person, variations can occur across different episodes of toothbrushing. This inter- and intra-personal variability poses a significant challenge when attempting to develop a generalized approach that can effectively detect about-to-brush moments for diverse individuals and across multiple instances.

- **Person-to-Person and Episode-wise Variability:** There is a high degree of variability observed between individuals in terms of their toothbrushing behavior and performance. Moreover, even within the same person, variations can occur across different episodes of toothbrushing. This inter- and intra-personal variability poses a significant challenge when attempting to develop a generalized approach that can effectively detect about-to-brush moments for diverse individuals and across multiple instances.

- **Lack of Unique Patterns in Signal for Micro Event Detection:** The signal associated with about-to-brush moments often lacks distinct and easily identifiable patterns. This absence of clear cues makes it difficult to precisely identify and distinguish the micro events within the about-to-brush moment. Consequently, the current approach may struggle to reliably detect these subtle events, requiring further exploration and development of novel signal analysis techniques.

To address these limitations, future research and improvements in the About-to-brush Moment Detection Approach should focus on enhancing cross-validation accuracy, increasing the size and diversity of the dataset, developing robust algorithms capable of capturing short-duration and low-repetition events, accounting for individual and episode-wise variability, and exploring innovative signal analysis methods to identify micro events within the about-to-brush moment more effectively. By overcoming these challenges, we can improve the reliability and applicability of the approach, ultimately leading to more accurate and personalized interventions for individuals' toothbrushing habits.

6.11.2 Limitations of Brushing-initiation Moment Detection Approach

The Brushing-initiation Moment Detection approach achieves an exceptional accuracy of 99.9%, indicating a high level of precision in identifying the initiation of toothbrushing.

This remarkable accuracy is undoubtedly commendable. However, it is important to note that the approach may encounter limitations when implemented with smart brushes. Smart brushes generate data in formats that may vary from the assumed format employed by the Brushing-initiation Moment Detection approach. These discrepancies can include differences in data acquisition, preprocessing methods, or feature extraction techniques. As a result, the approach's algorithms may not be well-suited to analyze the specific data generated by smart brushes, potentially leading to reduced accuracy or performance degradation.

6.11.3 A Comparative Analysis of About-to-Brush and Brushing-Initiation Methods

In this comparative analysis, our aim is to explore the limitations and advantages associated with the about-to-brush and brushing-initiation methods, with the ultimate goal of gaining valuable insights for developing effective intervention strategies. By thoroughly comparing and contrasting these approaches, we can obtain a deeper understanding of their respective strengths and weaknesses. This analysis serves as a foundational framework for devising intervention strategies that provide individuals with valuable information and guidance pertaining to their toothbrushing routines.

To facilitate this comparative analysis, we have identified four key metrics that will help us assess and evaluate the about-to-brush and brushing-initiation methods. These metrics act as benchmarks for measuring the strengths and weaknesses of each approach. Let's now delve into a detailed description of these metrics:

1. **Intervention Time:** The about-to-brush Moment Detection approach focuses on detecting toothbrushing just before it begins, allowing for a timely intervention. In contrast, the brushing-initiation approach detects toothbrushing within 4 seconds of its initiation, offering a slightly delayed but still effective intervention.

2. **Accuracy:** The brushing-initiation approach exhibits significantly higher accuracy, reaching as high as 99.9%, while the about-to-brush approach lags behind with an accuracy of less than 50%. This disparity in accuracy indicates the contrasting performance levels between the two approaches.

3. **Complexity:** The brushing-initiation method of detection is notably less complex compared to the about-to-brush approach. This implies that the brushing-initiation approach is simpler to implement and execute, potentially leading to enhanced practicality and ease of use.

4. **Person Variability:** The about-to-brush approach demonstrates high variability across individuals, meaning that its effectiveness may vary significantly from person to person. Conversely, the brushing-initiation approach is more robust, displaying greater consistency and reliability in detecting toothbrushing initiation across different individuals.

By evaluating the about-to-brush and brushing-initiation methods based on these four metrics, we can gain valuable insights into their characteristics and performance. This comprehensive analysis empowers us to design more effective intervention strategies that leverage the strengths of each approach while addressing their respective limitations. Ultimately, the goal is to provide individuals with tailored and impactful guidance to optimize their toothbrushing routines and improve their overall oral health.

6.12 Chapter Summary

The chapter focuses on two approaches for tracking tooth brushing behavior and providing interventions: detecting about-to-brush moments and detecting the initial portion of tooth brushing. It addresses challenges such as limited data by employing a synthetic data generation technique using a transformer-based architecture. The effectiveness of the synthetic data generation approach is evaluated, showing low deviation from real data.

The chapter discusses the detection performance of about-to-brush moments and suggests strategies to address overfitting and improve generalization. It then presents the Stroke Detection and Clustering algorithm, which effectively identifies brushing strokes and groups them into distinct events. The algorithm optimizes the strokes threshold to reduce false alarms while ensuring timely detection of true brushing events. Highlights of the chapter are describe below:

- **Synthetic data generation technique:** The authors employ a transformer-based architecture to generate synthetic sensor data that closely resembles real-world sensor measurements during the opening/closing of a toothpaste cap with rotation. This approach captures the intricacies and patterns of participants' movements.

- **Performance evaluation of synthetic data:** The synthetic data generation approach is evaluated using metrics such as mean squared error (MSE) and root mean squared error (RMSE). The results show a low level of deviation between the synthetic data and the ground truth real data, indicating the effectiveness of the approach in replicating sensor data characteristics.

- **Detection of about-to-brush moments:** The chapter discusses the challenges of overfitting in detecting about-to-brush moments. It suggests strategies such as collecting diverse data, employing regularization techniques, and fine-tuning the model architecture and hyperparameters to address overfitting and improve generalization performance.

- **Stroke detection and clustering algorithm:** The proposed algorithm, called Stroke Detection and Clustering, effectively identifies brushing strokes from raw accelerometer data and groups them into distinct brushing events. It employs a trained stroke detection model and a clustering technique to achieve accurate detection of brushing prompt events.

- **Optimization of strokes threshold:** The algorithm determines the minimum number of strokes threshold that reduces false alarms while ensuring timely detection of true brushing events. Setting the threshold at six strokes is considered practical and manageable for interventions.

Chapter 7

Conclusion and Future Directions

This dissertation has made significant contributions to the field of oral hygiene research, addressing various aspects of behavior analysis, intervention opportunities, and dataset improvement. These contributions have paved the way for future goals focused on enhancing oral health care practices and bringing advanced solutions to end users. This chapter provides a comprehensive recapitulation of the main theme and significant contributions presented in this thesis, while also highlighting emerging directions in behavior detection advancements. Additionally, it emphasizes the unique contributions and novel approaches introduced in each chapter, underscoring their relevance and impact on advancing behavior detection methodologies. Finally, the chapter concludes by outlining potential future directions and areas of research that naturally extend from the thesis's findings, serving as a foundation for future studies and inspiring further innovation in the exciting realm of behavior detection.

7.1 Summary and Key Contributions of This Dissertation:

This chapter provides a comprehensive summary of the key contributions and findings of the dissertation, which focuses on monitoring oral hygiene behaviors. Through years of research, the dissertation successfully tackled challenges related to accurate detection, precise analysis, and timely interventions in toothbrushing activities were addressed. The research efforts led to the development of advanced models, algorithms, and datasets, providing valuable insights into oral health practices in real-world settings. This chapter highlights the major accomplishments of this dissertation, emphasizing their potential impact on personalized oral health monitoring, interventions, and the broader field of activity monitoring.

7.1.1 Collection and Fine Grained Labeling of Oral Hygiene Behaviors Data in Natural Settings

Chapter 3 of our dissertation delved into the intricate process of collecting oral hygiene behaviors in natural settings through the utilization of wrist-worn inertial sensors. Recognizing the dearth of available open datasets in this domain, we conducted a comprehensive field study in Memphis, which received ethical approval from the Institutional Review Board (IRB), with all participants providing written informed consent.

We employed a lab-made smartwatch and a smartphone as ground truth devices to capture the necessary data. In order to preserve the integrity of the sensor data and ground truth data, we upgraded the mCerebrum software to enable video recording during oral hygiene activities. Recruitment efforts involved distributing informative flyers outlining the study's objectives and requirements. The study design was carefully crafted to align with our research goals. Participants were instructed to wear the wearable sensors and carry the study smartphone for a duration of seven (7) days. Throughout this period, participants were requested to record videos of their brushing activities in comfortable settings using the provided smartphone. With a diverse sample of 30 participants, incorporating factors such as gender, age, and dominant hand, we amassed an extensive dataset comprising over three thousand hours of sensor data, along with more than 400 videos for ground truth annotation. The meticulous task of annotating the brushing levels from these videos proved to be challenging and time-consuming, with over 500 hours dedicated to meticulously annotating over 10 thousand brushing levels.

These rigorous data collection and annotation efforts laid a solid foundation for our research, ensuring the accuracy and reliability of subsequent analysis and modeling. The substantial dataset and ground truth annotation obtained from the study served as valuable resources for developing and evaluating our approach to identifying brushing teeth surfaces using wrist-worn inertial sensors.

7.1.2 Open Dataset for Advancing Activity Monitoring: Introducing the mORAL Dataset for Oral Hygiene Research

In order to foster collaboration and facilitate further research in the field, we have made the collected data available as an open dataset called mORAL (Monitoring Oral Hygiene Activities) on the Internet. This dataset grants researchers access to a longitudinal collection of time-series data spanning over seven days (> 112 hours) from each participant. The extensive duration of data collection enables in-depth analysis of activity patterns throughout the day.

The mORAL dataset has proven to be highly valuable not only for developing activity detection methods but also for research on privacy and re-identification. The availability of this dataset encourages the exploration of innovative techniques and algorithms to tackle various challenges in activity monitoring. By providing open access to this data, we aim to encourage collaboration, accelerate advancements, and contribute to the broader scientific community's knowledge in this field.

7.1.3 Invariance to Variability in Wrist-Mounting

Chapter 4 delves into the challenge of detecting and correcting device orientation and wrist position during various activities. Considering that participants may wear the device in different positions, such as the top or bottom of the wrist, it is essential to accurately account for these variations to detect wrist-based activities. To address this, we introduced quaternion-based algorithms for orientation estimation that utilize accelerometer and gyroscope readings to align sensor data with hand orientation during brushing. This ensures precise inference of brushing movements and tooth surfaces.

Through extensive experimentation and evaluation, we successfully validated the effectiveness of the virtual orientation technique. This advancement played a vital role in achieving more precise inference of daily behaviors in real-world scenarios. The virtual orientation technique not only improves the accuracy of inferring oral hygiene behaviors but also holds the potential to enhance all activity detection algorithms by addressing the challenges associated with device orientation and wrist position variations. This contribution significantly strengthens the practical utility and generalizability of our research findings.

7.1.4 Dynamic Event-Based Time Series Segmentation: Improved Segmentation Method for Brushing Behavior Analysis

One of the key aspects explored is time series segmentation which involves breaking down a continuous time series into meaningful segments or intervals. When it comes to brushing their teeth, people have different brushing patterns, including motion, duration, time spent on each surface, and transitions between them. Traditional segmentation methods are ineffective due to the variability in brushing duration and surface transitions. We use a dynamic-length event-based approach to overcome these challenges that require an anchor event for effective segmentation. Our contribution is an improved event-based segmentation method that moves away from fixed or sliding window-based methods. By using event-based segmentation, we can accurately detect both toothbrushing and surface detection, precisely detecting the start and end of the brushing event and characterizing the entire brushing session.

During brushing and flossing, the wrist position tends to be higher than the elbow. To generate candidate segments for brushing events, we employ a different approach to detecting upward and downward wrist movements. A threshold is determined for the lateral axis of the accelerometer to filter out samples below this threshold, resulting in 81% data filtering. Temporal clusters are merged if the time difference between retained samples is less than 1 second. Further refinement of candidate segments is done based on their time duration, with optimal values of 11 seconds minimum and 2.5 minutes maximum duration. This method rejects 91% of the data, generating an average of 100 candidate segments per day, compared to 1,000 segments with the 2-minute window-based approach.

To identify surface transitions accurately to characterize the brushing session precisely, we found that the brushing stroke, a distinctive back-and-forth or up-and-down motion during brushing, serves as a reliable and easily detectable anchor event. This motion pattern is performed consistently across all surface transitions, making it suitable for surface identification. By analyzing the accelerometer signal for peak-valley pairs and selecting the brushing stroke with the maximum stretch along a specific axis, our approach effectively distinguishes surfaces based on the dominant wrist movement.

7.1.5 Detecting Start and End Times of Tooth Brushing and Flossing Behaviors

Another key contribution of this research is introducing a novel approach, namely *mORAL*, that utilizes wrist-worn inertial sensors to accurately detect and analyze daily behaviors such as toothbrushing and flossing in real-world settings.

Chapter 5 emphasizes the importance of monitoring oral hygiene behaviors and highlights the limitations of previous methods in capturing these activities accurately. Traditional approaches relied on self-reporting or scripted settings, leading to unreliable data due to memory biases or the artificial nature of the environment. To overcome these challenges, mORAL leveraged wearable sensor technology to provide a more natural and unobtrusive monitoring experience. The utilization of wrist-worn inertial sensors in the mTeeth model enabled convenient and unobtrusive monitoring of brushing activities in real-world settings.

A notable aspect of the research was the innovative approach to selecting candidate events for time series segmentation, moving away from fixed or sliding window-based methods. This approach offered greater flexibility and accuracy in segmenting the data. Leveraging machine learning techniques and sophisticated algorithms, mORAL successfully inferred brushing and flossing behaviors in real-world scenarios, providing valuable insights into individuals' oral hygiene habits without relying on self-reporting or controlled settings. The model's high accuracy and reliability were demonstrated through comprehensive evaluation metrics and validation experiments, further establishing its potential impact in the field of oral health monitoring.

Overall, the development of mORAL in 2019 represented a significant advancement in activity monitoring, particularly in the context of oral hygiene behaviors. The research showcased ongoing efforts to improve the accuracy and effectiveness of monitoring technologies, ultimately benefiting individuals' oral health and well-being.

7.1.6 Characterizing Tooth Brushing Behaviors

Chapter 6 presents the mTeeth model, which was rooted in the idea that effective oral hygiene is not only about the duration of brushing but also the coverage and effectiveness of brushing on different tooth surfaces. Its primary objective was to identify and characterize the specific tooth surfaces being brushed using a manual toothbrush in natural free-living environments. The model aimed to overcome the limitations of previous methods by focusing on the identification of tooth surfaces being brushed, thus offering a more granular analysis of brushing behaviors. The utilization of wrist-worn inertial sensors in the mTeeth model allowed for convenient and unobtrusive monitoring of brushing activities in real-world settings

One of the key aspects explored is time series segmentation, as we discussed in the previous section. Brushing teeth involves a series of strokes that target different tooth surfaces. Brushing teeth involves multiple strokes that address various tooth surfaces. The mTeeth model detects and analyzes these individual strokes to accurately identify the specific tooth surfaces being brushed. The chapter presents a robust classification model capable of recognizing different tooth surfaces, including the inner, outer, center, left, right, up, and down surfaces.

The development of the mTeeth model in 2021 signifies ongoing efforts to refine activity monitoring techniques and enhance oral hygiene practices. By focusing on micro events detection and tooth surface analysis, this model provides a comprehensive understanding of brushing behaviors, paving the way for future advancements in oral health monitoring. The mTeeth model showcases the potential of leveraging wearable sensor technologies to gain insights into oral hygiene practices and promote improved dental care habits. The model is a significant achievement in oral health research, emphasizing the importance of detailed analysis to optimize oral hygiene practices.

7.1.7 New Intervention Opportunities for Oral Hygiene Behaviors

Inadequate brushing of tooth surfaces remains a key factor contributing to the persistence of dental diseases, despite regular brushing. Certain surfaces may be overlooked entirely, while excessive time may be spent on others. To combat this issue, it is crucial to provide effective interventions that help individuals correct their brushing habits. However, the timing of these interventions is equally vital for their efficacy. Even with the right content, interventions delivered too early or when users are occupied may be forgotten or not fully absorbed. Mobile sensors have the potential to deliver interventions at the opportune moment, increasing the likelihood of success. Unfortunately, existing methods in the field typically detect brushing behavior after it has concluded, missing the window for timely intervention.

Chapter 6 delves into two innovative approaches for tracking tooth brushing behavior and providing interventions: detecting about-to-brush moments and detecting the initial portion of tooth brushing. These approaches address challenges such as limited data by employing a synthetic data generation technique that utilizes a transformer-based architecture. The effectiveness of the synthetic data generation approach is thoroughly evaluated, demonstrating minimal deviation from real data.

Furthermore, the chapter discusses the performance of detecting about-to-brush moments and proposes strategies to mitigate overfitting and enhance generalization. It introduces the Stroke Detection and Clustering algorithm, which efficiently identifies brushing strokes and groups them into distinct events. This algorithm optimizes the stroke threshold to minimize false alarms while ensuring the timely detection of genuine brushing events.

In summary, Chapter 6 addresses the critical factor of intervention timing by detecting about-to-brush moments and the initial portion of tooth brushing. It showcases the utilization of synthetic data generation. It introduces the Stroke Detection and Clustering algorithm to improve the accuracy and effectiveness of detecting brushing behavior, ultimately contributing to better oral hygiene practices. These advancements hold significant promise for enhancing oral health and fostering healthier brushing habits.

7.1.8 Impact on Health and Wellness

Collectively, all the work in this dissertation contributes to the advancement of mobile health (mHealth) technologies in the field of oral health monitoring. This work demonstrates the potential of wrist-worn inertial sensors and data analysis techniques for capturing and analyzing oral hygiene behaviors. These contributions pave the way for personalized interventions, improved oral health outcomes, and enhanced oral care practices. Furthermore, integrating multi-modal data fusion techniques and fine-grained tooth surface identification demonstrates the ability to capture detailed brushing information and comprehensively monitor oral hygiene activities. By focusing on real-world application and evaluation, these papers bridge the gap between laboratory-based research and practical implementation, offering promising prospects for the widespread adoption of mHealth solutions in oral health monitoring. By providing accurate and personalized insights into individuals' oral health practices, these contributions lay the foundation for improved oral care and better oral health outcomes.

7.2 Future Goals

The Future Goals section lays the foundation for future research and development in the field of oral health monitoring and activity tracking, outlining key areas for exploration and advancement. While significant strides have been made in understanding brushing behaviors and leveraging wearable sensor technologies, several untapped opportunities and challenges remain. This section outlines the key objectives and aspirations for future advancements to enhance further activity monitoring systems' accuracy, effectiveness, and usability. The goal of exploring novel approaches, incorporating emerging technologies, and addressing critical gaps is to propel the field forward, make meaningful contributions to oral health research, and promote optimal dental care practices.

7.2.1 Improvement in the Oral Hygiene Dataset by Conducting Longitudinal Studies

Through this study, we have gained valuable insights into the missing elements within our dataset that were previously unknown. Questions such as the number of participants required, the behaviors to be captured, and the duration of data collection have been addressed. Notably, we have discovered a significant variability among participants regarding brushing surface, indicating the need to considerably expand the participant pool. While 10 participants were insufficient, the next iteration may necessitate a larger sample size, such as 100 participants.

Another important lesson learned pertains to the timing of camera activation. Initially, we instructed participants to turn on the camera when they began brushing. However, we now recognize the value of having the camera activated before the actual brushing starts, during the preparation phase. This adjustment will provide additional context and insights into the entire oral hygiene routine.

The third lesson revolves around the camera positioning. While we initially suggested keeping the camera anywhere for convenience, we have realized that placing it on the sink counter surface does not yield useful results. Hence, we need to reconsider and refine the camera placement strategy for more effective data collection.

The analysis in this dissertation emphasizes the importance of an expanded dataset to explore brushing variability and participant differences. By expanding the ROBAS Dataset, we can investigate clustering patterns and identify users with similar habits, enhancing model reliability. Longitudinal studies conducted over a longer duration are crucial for validating activity monitoring systems in real-world contexts. Population-based research on a larger scale can assess the benefits and challenges of implementing these systems more broadly. By examining the long-term impact of activity monitoring on oral health and user adherence, valuable evidence can be gathered to enhance oral hygiene practices and prevent dental diseases. This comprehensive approach of expanding the dataset and conducting longitudinal studies in real-world settings contributes to personalized interventions and improved oral hygiene monitoring techniques, ultimately improving oral health outcomes.

7.2.2 Advancing Behavior Detection using Advanced Machine Learning/AI Models

After having a large dataset from Longitudinal Study that contains data from a large population, future research should focus on advancing behavior detection using more sophisticated machine learning and AI models. This study has demonstrated the potential of deep learning techniques and the stroke detection model to distinguish toothbrushing from other activities and identify specific tooth surfaces being brushed. Expanding on these findings, researchers can further explore the application of reinforcement learning and other innovative approaches to enhance the accuracy and analysis of brushing behaviors. These advanced algorithms can provide more detailed insights into brushing habits, enabling personalized interventions for better oral hygiene.

7.2.3 Real-time Intervention Design and Optimization

The primary objective of this dissertation is to implement the entire research findings on a smartwatch platform, enabling real-time operation of the developed models and algorithms. This implementation will bridge the gap between research and practical application, allowing for the evaluation of the developed techniques in real-world scenarios.

To achieve the second objective, a study will be conducted where the system operates in real time, providing an opportunity to test and compare different intervention ideas. Adopting a micro-randomized trial design approach, various intervention strategies can be explored to identify the optimal one for enhancing oral health behavior.

The micro-randomized trial design will focus on key areas of investigation, including the comparison of haptic feedback versus audio feedback to determine their respective effectiveness. Furthermore, the optimal timing for interventions will be investigated, considering factors such as switching tooth surfaces and specific intervals. These investigations will contribute to refining the design and implementation of personalized oral hygiene interventions.

By conducting this study, valuable insights can be gained, leading to the optimization of real-time interventions for oral health behavior. The findings will shed light on the effectiveness of different feedback modalities and provide guidance on the timing and frequency of interventions, ultimately enhancing the design and implementation of personalized oral hygiene interventions.

In parallel, it is crucial to design activity monitoring systems with the user in mind to ensure their widespread acceptance and usage. Future research should prioritize understanding user preferences, motivations, and obstacles to adoption, enabling the creation of user-friendly and intuitive interfaces. Collaborative clinical trials and studies involving dental professionals will provide strong evidence for the effectiveness of these systems, facilitating their seamless integration into regular oral care routines. This integration will contribute to improved oral health outcomes for individuals and populations as a whole.

7.2.4 Taking Oral Health to the End User Dissemination

The commercialization of toothbrushing detection, characterization, and prediction models has the potential to revolutionize oral health care. However, it is essential to ensure these advancements reach the end users effectively. To achieve this, future research should concentrate on deploying these commercialized solutions in real-life settings and actively seeking feedback from users and stakeholders.

Continuous monitoring and evaluation of the solution's performance are crucial to identify areas for improvement and address any concerns raised by users. By closely monitoring the effectiveness and user experience of the oral health care solutions, necessary adjustments can be made to enhance their functionality and address any issues that arise. User feedback, technological advancements, and emerging research findings should serve as valuable inputs for regular updates and enhancements to the toothbrushing detection and prediction models.

In addition, it is vital to stay informed about the latest developments in oral health care. By keeping pace with evolving trends, industry standards, and user expectations, the product can adapt and remain competitive in the market. Collaboration with oral health care professionals, researchers, and industry experts will help to stay at the forefront of advancements and ensure that the commercialized solutions align with current practices and offer cutting-edge features.

Ultimately, the dissemination of advanced oral health care solutions relies on continuous cycle of improvement, user engagement, and staying ahead of the curve. By prioritizing user feedback, monitoring performance, and innovation, these solutions can effectively transform oral health care and improve the well-being of individuals on a larger scale.

In summary, the future goals encompass several important areas for further research and development. Expanding the current ROBAS Dataset will enable researchers to explore the variability among brushing episodes and develop adaptive algorithms for personalized interventions. Conducting longitudinal studies and real-world validation will provide insights into the long-term effectiveness and challenges of implementing activity monitoring systems on a larger scale. Advancing behavior detection using advanced machine learning/AI models will enhance the accuracy and analysis of brushing behaviors, enabling personalized interventions and improved oral hygiene practices. Implementing real-time feedback and interventions will guide users in optimal brushing techniques and promote better oral health outcomes. Adopting a user-centered design approach and conducting clinical validation and adoption studies will ensure widespread acceptance and integration of activity monitoring systems into routine oral care practices. Finally, commercializing toothbrushing detection models and continuously improving them based on user feedback and technological advancements will contribute to advanced oral health care solutions. These future goals will collectively drive advancements in oral hygiene monitoring and interventions, ultimately leading to enhanced oral health outcomes for individuals.

REFERENCES

[1] R. M. Benjamin, "Oral health: the silent epidemic," *Public health reports*, vol. 125, no. 2, pp. 158–159, 2010.

[2] CDC *et al.*, "Oral health: Preventing cavities, gum disease, tooth loss, and oral cancers. at a glance 2011," *Atlanta: Division of Oral Health, National Center for Chronic Disease Prevention and Health Promotion*, 2010.

[3] B. L. Chadwick, D. White, D. Lader, and N. Pitts, "Preventive behaviour and risks to oral health-a report from the adult dental health survey 2009," 2011.

[4] *Wearable device forecast*, Accessed January, 2019. [Online]. Available: https://www.statista.com/chart/3370/wearable-device-forecast/

[5] S. Kumar, G. D. Abowd, W. T. Abraham, M. al'Absi, J. Gayle Beck, D. H. Chau, T. Condie, D. E. Conroy, E. Ertin, D. Estrin, *et al.*, "Center of excellence for mobile sensor data-to-knowledge (md2k)," *Journal of the American Medical Informatics Association*, vol. 22, no. 6, pp. 1137–1142, 2015.

[6] *Usage of manual toothbrushes*, Accessed June, 2021. [Online]. Available: https://www.statista.com/statistics/278116/us-households-usage-of-manual-toothbrushes/

[7] B. Bruno, F. Mastrogiovanni, A. Sgorbissa, T. Vernazza, and R. Zaccaria, "Human motion modelling and recognition: A computational approach," in *2012 IEEE International Conference on Automation Science and Engineering (CASE 2012)*. IEEE, 2012, pp. 156–161.

[8] B. Bruno, F. Mastrogiovanni, and A. Sgorbissa, "A public domain dataset for adl recognition using wrist-placed accelerometers." in *RO-MAN*, 2014, pp. 738–743.

[9] B. Bruno, F. Mastrogiovanni, A. Sgorbissa, T. Vernazza, and R. Zaccaria, "Analysis of human behavior recognition algorithms based on acceleration data," in *Robotics and Automation (ICRA), 2013 IEEE International Conference on*. IEEE, 2013, pp. 1602–1607.

[10] S. Zhang, A. V. Rowlands, P. Murray, T. L. Hurst, *et al.*, "Physical activity classification using the genea wrist-worn accelerometer," 2012.

[11] K. Ellis, J. Kerr, S. Godbole, G. Lanckriet, D. Wing, and S. Marshall, "A random forest classifier for the prediction of energy expenditure and type of physical activity from wrist and hip accelerometers," *Physiological measurement*, vol. 35, no. 11, p. 2191, 2014.

[12] T. G. Pavey, N. D. Gilson, S. R. Gomersall, B. Clark, and S. G. Trost, "Field evaluation of a random forest activity classifier for wrist-worn accelerometer data," *Journal of science and medicine in sport*, vol. 20, no. 1, pp. 75–80, 2017.

[13] P.-J. Hsieh, Y.-L. Lin, Y.-H. Chen, and W. Hsu, "Egocentric activity recognition by leveraging multiple mid-level representations," in *Multimedia and Expo (ICME), 2016 IEEE International Conference on*. IEEE, 2016, pp. 1–6.

[14] A. Mannini, S. S. Intille, M. Rosenberger, A. M. Sabatini, and W. Haskell, "Activity recognition using a single accelerometer placed at the wrist or ankle," *Medicine and science in sports and exercise*, vol. 45, no. 11, p. 2193, 2013.

[15] E. Thomaz, I. Essa, and G. D. Abowd, “A practical approach for recognizing eating moments with wrist-mounted inertial sensing,” in *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 2015, pp. 1029–1040.

[16] N. Saleheen, A. A. Ali, S. M. Hossain, H. Sarker, S. Chatterjee, B. Marlin, E. Ertin, M. Al’Absi, and S. Kumar, “puffmarker: a multi-sensor approach for pinpointing the timing of first lapse in smoking cessation,” in *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 2015, pp. 999–1010.

[17] A. Parate, M.-C. Chiu, C. Chadowitz, D. Ganesan, and E. Kalogerakis, “Risq: Recognizing smoking gestures with inertial sensors on a wristband,” in *Proceedings of the 12th annual international conference on Mobile systems, applications, and services*. ACM, 2014, pp. 149–161.

[18] E. Grossman, M. Cronin, W. Dembling, and H. Proskin, “A comparative clinical study of extrinsic tooth stain removal with two electric toothbrushes [braun d7 and d9] and a manual brush.” *American journal of dentistry*, vol. 9, pp. S25–9, 1996.

[19] T. Yoshitani, M. Ogata, and K. Yatani, “Lumio: a plaque-aware toothbrush,” in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 2016, pp. 605–615.

[20] “Sonic-powered automatic toothbrush,” https://bfbe9e.kckb.st.

[21] Y.-J. Lee, P.-J. Lee, K.-S. Kim, W. Park, K.-D. Kim, D. Hwang, and J.-W. Lee, “Toothbrushing region detection using three-axis accelerometer and magnetic sensor,” *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 3, pp. 872–881, 2011.

[22] Y.-J. Lee, P.-J. Lee, J.-W. Lee, K.-S. Kim, J.-S. Jeong, W. Park, and K.-D. Kim, “Quantitative assessment of toothbrushing education efficacy using smart toothbrush,” in *2011 4th International Conference on Biomedical Engineering and Informatics (BMEI)*, vol. 2. IEEE, 2011, pp. 1160–1164.

[23] “Oral B Genius brush,” Accessed: August 2020, https://oralb.com/en-us/products/electric-toothbrushes/genius-8000-rechargeable-electric-toothbrush/.

[24] H. Huang and S. Lin, “Toothbrushing monitoring using wrist watch,” in *Proceedings of the 14th ACM Conference on Embedded Network Sensor Systems CD-ROM*. ACM, 2016, pp. 202–215.

[25] M. Marcon, A. Sarti, and S. Tubaro, “Toothbrush motion analysis to help children learn proper tooth brushing,” *Computer Vision and Image Understanding*, vol. 148, pp. 34–45, 2016.

[26] J. Korpela, R. Miyaji, T. Maekawa, K. Nozaki, and H. Tamagawa, “Evaluating tooth brushing performance with smartphone sound data,” in *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 2015, pp. 109–120.

[27] Z. Ouyang, J. Hu, J. Niu, and Z. Qi, “An asymmetrical acoustic field detection system for daily tooth brushing monitoring,” in *GLOBECOM 2017-2017 IEEE Global Communications Conference*. IEEE, 2017, pp. 1–6.

[28] N. Lee, D. Jang, Y. Kim, B.-C. Bae, and J.-D. Cho, "Denteach: A device for fostering children's good tooth-brushing habits," in *Proceedings of the The 15th International Conference on Interaction Design and Children*. ACM, 2016, pp. 619–624.

[29] S. Akther, N. Saleheen, S. A. Samiei, V. Shetty, E. Ertin, and S. Kumar, "moral: An mhealth model for inferring oral hygiene behaviors in-the-wild using wrist-worn inertial sensors," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 3, no. 1, pp. 1–25, 2019.

[30] S. Akther, N. Saleheen, M. Saha, V. Shetty, and S. Kumar, "mteeth: Identifying brushing teeth surfaces using wrist-worn inertial sensors," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 5, no. 2, pp. 1–25, 2021.

[31] S. M. Hossain, T. Hnat, N. Saleheen, N. J. Nasrin, J. Noor, B.-J. Ho, T. Condie, M. Srivastava, and S. Kumar, "mcerebrum: A mobile sensing software platform for development and validation of digital biomarkers and interventions," in *Proceedings of the 15th ACM Conference on Embedded Network Sensor Systems*. ACM, 2017, p. 7.

[32] V. Shetty, J. Yamamoto, and K. Yale, "Re-architecting oral healthcare for the 21st century," *Journal of dentistry*, vol. 74, pp. S10–S14, 2018.

[33] A. Yap, "Oral health equals total health: A brief review," *Journal of Dentistry Indonesia*, vol. 24, no. 2, pp. 59–62, 2017.

[34] C.-Y. Li, Y.-C. Chen, W.-J. Chen, P. Huang, and H.-h. Chu, "Sensor-embedded teeth for oral activity recognition," in *Proceedings of the 2013 international symposium on wearable computers*. ACM, 2013, pp. 41–44.

[35] "Disney Magic Timer by Oral-B," https://oralb.com/en-us/product-collections.

[36] "ToothTunes by Arm & Hammer," https://www.armandhammer.com/spinbrush-pro-clean.

[37] Y.-C. Chang, J.-L. Lo, C.-J. Huang, N.-Y. Hsu, H.-H. Chu, H.-Y. Wang, P.-Y. Chi, and Y.-L. Hsieh, "Playful toothbrush: ubicomp technology for teaching tooth brushing to kindergarten children," in *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 2008, pp. 363–372.

[38] T. Nakajima, V. Lehdonvirta, E. Tokunaga, M. Ayabe, H. Kimura, and Y. Okuda, "Lifestyle ubiquitous gaming: making daily lives more plesurable," in *Embedded and Real-Time Computing Systems and Applications, 2007. RTCSA 2007. 13th IEEE International Conference on*. IEEE, 2007, pp. 257–266.

[39] Y.-J. Lee, P.-J. Lee, K.-S. Kim, W. Park, K.-D. Kim, D. Hwang, and J.-W. Lee, "Toothbrushing region detection using three-axis accelerometer and magnetic sensor," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 3, pp. 872–881, 2012.

[40] K.-S. Kim, T.-H. Yoon, J.-W. Lee, and D.-J. Kim, "Interactive toothbrushing education by a smart toothbrush system via 3d visualization," *Computer methods and programs in biomedicine*, vol. 96, no. 2, pp. 125–132, 2009.

[41] L. G. Bozgeyikli, E. C. Bozgeyikli, and A. Raij, "Keep brushing! developing healthy oral hygiene habits in young children with an interactive toothbrush."

[42] J. Korpela, R. Miyaji, T. Maekawa, K. Nozaki, and H. Tamagawa, "Toothbrushing performance evaluation using smartphone audio based on hybrid hmm-recognition/svm-regression model," *Journal of Information Processing*, vol. 24, no. 2, pp. 302–313, 2016.

[43] T. Hachisu and H. Kajimoto, "Augmentation of toothbrush by modulating sounds resulting from brushing," in *Advances in Computer Entertainment*. Springer, 2012, pp. 31–43.

[44] ——, "Modulating tooth brushing sounds to affect user impressions," *International Journal of Arts and Technology*, vol. 8, no. 4, pp. 307–324, 2015.

[45] W. Dargie, "Analysis of time and frequency domain features of accelerometer measurements." in *ICCCN*, vol. 9, 2009, pp. 1–6.

[46] H. G. Min and E. T. Jeung, "Complementary filter design for angle estimation using mems accelerometer and gyroscope," *Department of Control and Instrumentation, Changwon National University, Changwon, Korea*, pp. 641–773, 2015.

[47] S. Jain, C. Borgiattino, Y. Ren, M. Gruteser, Y. Chen, and C. F. Chiasserini, "Lookup: Enabling pedestrian safety services via shoe sensing," in *Proceedings of the 13th Annual International Conference on Mobile Systems, Applications, and Services*. ACM, 2015, pp. 257–271.

[48] M. A. Hall, "Correlation-based feature selection for machine learning," 1999.

[49] T. Rahman, M. Czerwinski, R. Gilad-Bachrach, and P. Johns, "Predicting about-to-eat moments for just-in-time eating intervention," in *Proceedings of the 6th International Conference on Digital Health Conference*. ACM, 2016, pp. 141–150.

[50] H. Sarker, M. Sharmin, A. A. Ali, M. M. Rahman, R. Bari, S. M. Hossain, and S. Kumar, "Assessing the availability of users to engage in just-in-time intervention in the natural environment," in *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 2014, pp. 909–920.

[51] S. B. Kotsiantis, I. Zaharakis, and P. Pintelas, "Supervised machine learning: A review of classification techniques," *Emerging artificial intelligence applications in computer engineering*, vol. 160, pp. 3–24, 2007.

[52] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[53] T. G. Dietterich, "Ensemble methods in machine learning," in *International workshop on multiple classifier systems*. Springer, 2000, pp. 1–15.

[54] L. Rokach, "Ensemble-based classifiers," *Artificial Intelligence Review*, vol. 33, no. 1-2, pp. 1–39, 2010.

[55] H. He and E. A. Garcia, "Learning from imbalanced data," *IEEE Transactions on Knowledge & Data Engineering*, no. 9, pp. 1263–1284, 2008.

[56] A. Estabrooks, T. Jo, and N. Japkowicz, "A multiple resampling method for learning from imbalanced data sets," *Computational intelligence*, vol. 20, no. 1, pp. 18–36, 2004.

[57] R. Caruana, A. Niculescu-Mizil, G. Crew, and A. Ksikes, "Ensemble selection from libraries of models," in *Proceedings of the twenty-first international conference on Machine learning*. ACM, 2004, p. 18.

[58] Y. Freund, R. E. Schapire, *et al.*, "Experiments with a new boosting algorithm," in *Icml*, vol. 96. Citeseer, 1996, pp. 148–156.

[59] M. M. Rahman, R. Bari, A. A. Ali, M. Sharmin, A. Raij, K. Hovsepian, S. M. Hossain, E. Ertin, A. Kennedy, D. H. Epstein, *et al.*, "Are we there yet?: Feasibility of continuous stress assessment via wireless physiological sensors," in *Proceedings of the 5th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics*. ACM, 2014, pp. 479–488.

[60] A. Stisen, H. Blunck, S. Bhattacharya, T. S. Prentow, M. B. Kjærgaard, A. Dey, T. Sonne, and M. M. Jensen, "Smart devices are different: Assessing and mitigatingmobile sensing heterogeneities for activity recognition," in *Proceedings of the 13th ACM Conference on Embedded Networked Sensor Systems*. ACM, 2015, pp. 127–140.

[61] T. Attin and E. Hornecker, "Tooth brushing and oral health: how frequently and when should tooth brushing be performed?" *Oral health & preventive dentistry*, vol. 3, no. 3, 2005.

[62] H. Löe, "Oral hygiene in the prevention of caries and periodontal disease," *International dental journal*, vol. 50, no. 3, pp. 129–139, 2000.

[63] C. Graetz, J. Bielfeldt, L. Wolff, C. Springer, K. M. F. El-Sayed, S. Sälzer, S. Badri-Höher, and C. E. Dörfer, "Toothbrushing education via a smart software visualization system," *Journal of periodontology*, vol. 84, no. 2, pp. 186–195, 2013.

[64] K.-D. Kim, J.-S. Jeong, H. N. Lee, Y. Gu, K.-S. Kim, J.-W. Lee, and W. Park, "Efficacy of computer-assisted, 3d motion-capture toothbrushing instruction," *Clinical oral investigations*, vol. 19, no. 6, pp. 1389–1394, 2015.

[65] T. Nakajima, V. Lehdonvirta, E. Tokunaga, and H. Kimura, "Reflecting human behavior to motivate desirable lifestyle," in *Proceedings of the 7th ACM conference on Designing interactive systems*. ACM, 2008, pp. 405–414.

[66] Y. Liang, H. W. Fan, Z. Fang, L. Miao, W. Li, X. Zhang, W. Sun, K. Wang, L. He, and X. Chen, "Oralcam: Enabling self-examination and awareness of oral health using a smartphone camera," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020, pp. 1–13.

[67] T. Hamatani, M. Elhamshary, A. Uchiyama, and T. Higashino, "Fluidmeter: Gauging the human daily fluid intake using smartwatches," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 2, no. 3, pp. 1–25, 2018.

[68] H. Li, S. Chawla, R. Li, S. Jain, G. D. Abowd, T. Starner, C. Zhang, and T. Plötz, "Wristwash: towards automatic handwashing assessment using a wrist-worn device," in *Proceedings of the 2018 ACM International Symposium on Wearable Computers*. ACM, 2018, pp. 132–139.

[69] C. Luo, X. Feng, J. Chen, J. Li, W. Xu, W. Li, L. Zhang, Z. Tari, and A. Y. Zomaya, "Brush like a dentist: Accurate monitoring of toothbrushing via wrist-worn gesture sensing," in *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*. IEEE, 2019, pp. 1234–1242.

[70] T. Winterfeld, N. Schlueter, D. Harnacke, J. Illig, J. Margraf-Stiksrud, R. Deinzer, and C. Ganss, "Toothbrushing and flossing behaviour in young adults—a video observation," *Clinical Oral Investigations*, vol. 19, no. 4, pp. 851–858, 2015.

[71] R. Deinzer, S. Ebel, H. Blättermann, U. Weik, and J. Margraf-Stiksrud, “Toothbrushing: to the best of one’s abilities is possibly not good enough,” *BMC Oral Health*, vol. 18, no. 1, pp. 1–7, 2018.

[72] K.-H. Lee, J.-W. Lee, K.-S. Kim, D.-J. Kim, K. Kim, H.-K. Yang, K. Jeong, and B. Lee, “Tooth brushing pattern classification using three-axis accelerometer and magnetic sensor for smart toothbrush,” in *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2007, pp. 4211–4214.

[73] L. N. Wijayasingha and B. Lo, “A wearable sensing framework for improving personal and oral hygiene for people with developmental disabilities,” in *2016 IEEE Wireless Health (WH)*. IEEE, 2016, pp. 1–7.

[74] M. B. Alonso, J. Stienstra, and R. Dijkstra, “Brush and learn: transforming tooth brushing behavior through interactive materiality, a design exploration,” in *Proceedings of the 8th International conference on Tangible, embedded and embodied interaction*. ACM, 2014, pp. 113–120.

[75] “Bass tooth brushing technique,” Accessed: August 2020, https://tube.medchrome.com/2013/04/brushing-technique-bass-and-modified.html.

[76] “ELAN software,” Accessed: August 2020, https://archive.mpi.nl/tla/elan.

[77] T. Rahman, M. Czerwinski, R. Gilad-Bachrach, and P. Johns, “Predicting" about-to-eat" moments for just-in-time eating intervention,” in *Proceedings of the 6th International Conference on Digital Health Conference*, 2016, pp. 141–150.

[78] D. H. Gustafson, B. R. Shaw, A. Isham, T. Baker, M. G. Boyle, and M. Levy, “Explicating an evidence-based, theoretically informed, mobile technology-based system to improve outcomes for people in recovery for alcohol dependence,” *Substance use & misuse*, vol. 46, no. 1, pp. 96–111, 2011.

[79] D. C. Mohr, S. M. Schueller, E. Montague, M. N. Burns, and P. Rashidi, “The behavioral intervention technology model: an integrated conceptual and technological framework for ehealth and mhealth interventions,” *Journal of medical Internet research*, vol. 16, no. 6, p. e3077, 2014.

[80] L. Pina, K. Rowan, A. Roseway, P. Johns, G. R. Hayes, and M. Czerwinski, “In situ cues for adhd parenting strategies using mobile technology,” in *Proceedings of the 8th international conference on pervasive computing technologies for healthcare*, 2014, pp. 17–24.

[81] B. Wansink and K. A. Johnson, “The clean plate club: About 92% of self-served food is eaten,” *International Journal of Obesity*, vol. 39, no. 2, pp. 371–374, 2015.

[82] K. Witkiewitz and G. A. Marlatt, “Relapse prevention for alcohol and drug problems: that was zen, this is tao.” 2009.

[83] M. K. Ameko, M. L. Beltzer, L. Cai, M. Boukhechba, B. A. Teachman, and L. E. Barnes, “Offline contextual multi-armed bandits for mobile health interventions: A case study on emotion regulation,” in *Proceedings of the 14th ACM Conference on Recommender Systems*, 2020, pp. 249–258.

[84] S. Zhang, R. Alharbi, W. Stogin, M. Pourhomayun, B. Spring, and N. Alshurafa, "Food watch: Detecting and characterizing eating episodes through feeding gestures," in *Proceedings of the 11th EAI International Conference on Body Area Networks*, 2016, pp. 91–96.

[85] Y. Wang, L. Duan, S. Butscher, J. Mueller, and H. Reiterer, "Fingerprints: detecting meaningful moments for mobile health intervention," in *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*, 2016, pp. 1085–1088.

[86] H. Sarker, M. Sharmin, A. A. Ali, M. M. Rahman, R. Bari, S. M. Hossain, and S. Kumar, "Assessing the availability of users to engage in just-in-time intervention in the natural environment," in *Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing*, 2014, pp. 909–920.

Appendix A

Appendix

Demographics Questionnaire

Participant ID ________________ **Date and Time____________________**

1. How did you learn about the study? ____

2. Age: ___________

3. Gender: Male Female

4. Race:

   - ☐ Black – Non Hispanic
   - ☐ American Indian/Alaskan Native
   - ☐ Hispanic
   - ☐ Asian/Pacific Islander
   - ☐ White – Non Hispanic
   - ☐ Other

5. Employment Status (check all that apply):

   - ☐ Employed Full-Time
   - ☐ Employed Part-Time
   - ☐ Self Employed
   - ☐ Not Employed
   - ☐ Student Full-Time
   - ☐ Student Part-Time
   - ☐ Other ____________________

6. Yearly Income:

   - ☐ Less than $10,000
   - ☐ $10,000 to $14,999
   - ☐ $15,000 to $24,999
   - ☐ $25,000 to $34,999
   - ☐ $35,000 to $49,999
   - ☐ $50,000 to $74,999
   - ☐ $75,000 to $99,999
   - ☐ $100,000 or more

7. Do you smoke cigarettes? YES NO

   If YES, approximately how many cigarettes per day? ________________________________

   If NO have you ever smoked? ____

8. How often do you brush your teeth on a typical day? __________ times

   Do you typically brush your teeth when you wake up? YES NO

   Do you typically brush your teeth before going to bed? YES NO

9. How often do you floss your teeth on a typical day? ___________ times

Do you typically floss your teeth when you wake up? YES NO

Do you typically floss your teeth before going to bed? YES NO

10. How often do you use an oral rinse (e.g., Listerine) on a typical day? ___________ times

Do you typically use oral rinse when you wake up? YES NO

Do you typically use oral rinse before going to bed? YES NO

Institutional Review Board

315 Administration Bldg.
Memphis, TN 38152-3370
Office: 901.678.2705
Fax: 901.678.2199

Consent to Participate in a Research Study

The utility of inferring health-related behaviors from smartwatches & other mobile sensors

WHY ARE YOU BEING INVITED TO TAKE PART IN THIS RESEARCH?

You are being invited to take part in a research study about the utility of mobile sensor devices like smartwatches and smart toothbrushes. If you volunteer to take part in this study, you will be one of about 20 people to do so.

WHO IS DOING THE STUDY?

The person in charge of this study is Dr. Santosh Kumar of the Department of Computer Science at the University of Memphis. You are invited to ask him or any of the study staff questions about this study or your participation at any time.

WHAT IS THE PURPOSE OF THIS STUDY?

The primary aim of this study is to collect information on mobile health (mHealth) platforms that can detect health-related behaviors. This information will help us to plan future studies that look at how mHealth technologies can improve health.

We (the research team) also hope to learn how using mobile sensors and “smart” technologies like smart watches and smart toothbrushes can identify, health status, stress, and other health-related behaviors such as oral health behaviors.

ARE THERE REASONS WHY YOU SHOULD NOT TAKE PART IN THIS STUDY?

- We are recruiting healthy individuals between the ages of 18-64.
- We are recruiting individuals who brush their teeth at least twice every day and floss their teeth at least once every day.
- We are recruiting individuals who are comfortable wearing and using a set of mobile sensors (wearing wrist sensors, a study smart phone, and using a sensor-enabled toothbrush) for the next seven (7) days.
- We are recruiting individuals who are comfortable taking videos of themselves whenever they perform an oral health behavior (brushing or flossing).
- We are recruiting individuals who can attend a total of two (2) lab sessions at the beginning and the end of the next seven (7) days.

If, for any reason, you feel uncomfortable with the study procedures, you should not take part in this study. If you decline to take part in this study, it will not be held against you.

WHERE IS THE STUDY GOING TO TAKE PLACE AND HOW LONG WILL IT LAST?

THE UNIVERSITY OF MEMPHIS
Dreamers. Thinkers. Doers.

Institutional Review Board

315 Administration Bldg.
Memphis, TN 38152-3370
Office: 901.678.2705
Fax: 901.678.2199

The research procedures will be conducted at the mobile sensors lab at Dunn Hall Room 216A on the University of Memphis main campus. You will need to come to the mobile sensors lab two (2) times: once at the beginning and once at the end of the next seven (7) days. Each visit should each take about one (1) hour.

WHAT WILL YOU BE ASKED TO DO?

- If you agree to be in this study, we will provide you with two wrist-worn sensors (smart watches), a sensor-enabled toothbrush, a smartphone, and a mounting device for the smartphone at an in-person visit. You will use these sensors and smartphone during the study. The smartphone will only be for study use, and we will expect you to return the smartphone and all sensors at the end of the study. If you complete the study, you can keep the sensor-enabled toothbrush.
- The two in-person visits will each take about 1 hour. During the first visit, we will show you how to wear the sensors and how to use the smartphone device.
- We will ask you to wear the sensors and the smartphone device during all waking hours (removing them only when bathing or swimming) as you engage in your typical day.
- While wearing all of the sensors, we will ask you to use the sensor-enabled ("smart") toothbrush once daily every time you brush your teeth in the morning. We will ask you to use a manual toothbrush every time you brush your teeth in the evening.
- We will provide you tooth brushing and flossing instructions to ensure that we can record a complete set of tooth brushing and flossing motions.
- We will ask you to use the mounting device to position the study smartphone to take videos of yourself whenever your brush or floss your teeth while wearing the wrist sensors.
    - You will:
        1. Mount the study smartphone to a suitable location (e.g., bathroom mirror, sink countertop, or bathroom tile),
        2. Align yourself in the video frame,
        3. Push a button on the phone to begin taking a video before you start brushing or flossing, and then
        4. Push a button to stop taking video when you are done with your oral health routine.
    - If you use an oral rinse at the same time you brush or floss, we ask that you continue taking video as you use the oral rinse until after you are done.
    - We will collect this video data from the study smartphones after you return the phone to the lab.
    - We will use this information to make detailed notes about the data coming from the wrist sensors. We will use these notes to build models that can tell us when people are brushing or flossing just from the wrist sensor data alone.
    - We ask you to take this video every time that your brush and/or floss your teeth. However, we also ask that you take this video when and where you are comfortable (e.g., at home). If you are uncomfortable taking the video at a given

Institutional Review Board

315 Administration Bldg.
Memphis, TN 38152-3370
Office: 901.678.2705
Fax: 901.678.2199

time, we ask that you document the time you start and stop brushing and flossing, using the smartphone device.
    - You will have the opportunity to review your video and delete it if you don't want to share it with the research team. You will be able to do this after you have taken the video, or at the end of the study when you return the device.
- We understand that it may not be practical to wear the sensors and/or smartphone device all the time. When you cannot wear them or you do not want to wear them, you may take them off. We will also show you how to temporarily pause the phone and sensors from collecting data, if removing the sensors isn't practical.
- You will need to charge the sensors and the smartphone every night.
- During the study, the sensors will be collecting data about your physical movement. This information will be collected by the study smartphone. This information by itself won't be able to identify you individually, but we will be able to use this information about your heart and wrist movements to determine certain behaviors and health states: For example, we will be able to know if you are stressed, smoking, or eating.
- The phone will send this information to a secure computer periodically throughout the day via a secure cellular connection.
- During the study, the smartphone will be collecting your location (where you are) via GPS. This information will stay on the phone, on an encrypted card and will not be uploaded in real-time. After the study is over and you've returned the device, this information will be uploaded, via a secure cellular connection, to the secure computer.
    - This means we won't be tracking where you are during the study (in real-time). After the study, we will know every place that you have visited while the study was going on. We will use this information to see if your location has anything to do with your oral health behaviors.
- We will collect the sensors and devices at an in-person visit at the end of the study. If you complete the study, you can keep the sensor-enabled toothbrush.
- At the final in-person visit, we will ask you to answer some questions about your experiences with and attitudes toward the devices.
- The final in-person visit will take approximately 1 hour.

WHAT ARE THE POSSIBLE RISKS AND DISCOMFORTS?

- To the best of our knowledge, the things you will be doing have no more risk of harm than you would experience in everyday life.
- Wearing the wrist sensors may be uncomfortable, but most people get used to them.
- You will need to carry around an extra phone with you while you are wearing the sensors.
- You will need to charge the study devices at night. There are four (4) devices, total. This may be burdensome, but we will provide you with a device that makes it easier to charge all of the devices at once.

Institutional Review Board

315 Administration Bldg.
Memphis, TN 38152-3370
Office: 901.678.2705
Fax: 901.678.2199

- The phone will use Global Positioning System (GPS) to collect location data. This location data could identify you by showing where you are. Research staff will protect this information by restricting access only to authorized personnel. All researchers and research staff have been trained in human subjects research (research that involves living people). All researchers and research staff will agree to not use the location or the video data to identify you as an individual. Neither your location nor video data will be made available to the public.
- We will make every reasonable effort to ensure that your information is kept secure, but in the event of a data breach, your location data (all of the places you've been during the study) and/or your video data could be exposed. We will minimize this risk by following best practices in information security, and will delete information that can identify you (your location and video data) after the study is over.
- We will keep all of your information confidential within the limits allowed by law. However, we cannot guarantee complete secrecy. For example, we are required by law to report evidence of child abuse or neglect.

WILL YOU BENEFIT FROM TAKING PART IN THIS STUDY?

There is no guarantee that you will get any benefit from taking part in this study. However, you may find this study helpful to improve your oral health. The toothbrush may provide you with feedback that may help you improve your own brushing techniques. Your willingness to take part may help society as a whole. Researchers may learn valuable information from this study to understand how using mHealth technologies like wearable sensors and "smart" devices can detect certain behaviors and improve health.

DO YOU HAVE TO TAKE PART IN THE STUDY?

- Whether you choose to take part is up to you.
- You can choose to say "no" and not take part.
- You can agree to take part in this research and change your mind later.
- Your decision to say "no", whether now or if you change your mind later, will not be held against you.
- You can ask any and all questions that you want before deciding to take part.

IF YOU DON'T WANT TO TAKE PART IN THE STUDY, ARE THERE OTHER CHOICES?

If you do not want to be in the study, there are no other choices except not to take part in the study.

WHAT WILL IT COST YOU TO PARTICIPATE?

Institutional Review Board

315 Administration Bldg.
Memphis, TN 38152-3370
Office: 901.678.2705
Fax: 901.678.2199

You will have to take the time to come to the lab twice, at the beginning and end of the next seven (7) days. If you are not a current student, faculty, or staff person, you will have to pay for your own travel costs (like parking costs) to come to the lab.

WILL YOU RECEIVE ANY REWARDS FOR TAKING PART IN THIS STUDY?

You will be able to keep the sensor-enabled ("smart") toothbrush if you complete the study. This toothbrush is worth about $129.

WHO WILL SEE THE INFORMATION THAT YOU GIVE?

This research study is a part of a research collaboration that aims to develop new ways of improving health by using mobile sensor technology. Because of this, we want to save your research data and share it with other researchers who study mobile health. The rest of this section describes how we will use your data and share it with other researchers.

Data security & confidentiality: We will use best practices to prevent unauthorized access to your information. For example, we will make sure that all paper study records are kept in locked areas not accessible to the general public. All electronic study records and data (including video data and GPS) will be kept on computers that are secured with appropriate technical safeguards such as password protection and encryption. This means it will be very difficult for someone to access your information if they do not have permission. We will protect your identity to the extent required by law. However, we cannot guarantee complete secrecy. For example, we are required by law to report evidence of child abuse or neglect. All of your records will be open to inspection by the research study staff, the IRB, other representatives of this institution, and the sponsor of this study: the U.S. Department of Health and Human Services, National Institutes of Health.

How we will use your data: We will link your contact information and your research data (the data collected by all of the questionnaires, study sensors, and the phone) with a code number. A master key that links your name, your contact information, and your code number will be maintained in a separate and secure location from your research data. We will only use your contact information for the purposes of contacting you about this research study, and future research studies if you choose. We will use your research data for scientific progress and for publication of study results. Information that we make public will only be in the form of summaries that make it impossible to tell who the individual participants were.

How we will share your data: The research data that we collect about you in this study may also be shared with other researchers. Some of these researchers may be at other universities/institutions. We will only share information that can't be used to identify you.

GPS data will contain all of the places you have visited during the study. Because of this, GPS data could be used to identify you. We will not share any data that includes your raw location data (GPS) with other researchers. We will use your GPS data to make a code for certain

Institutional Review Board

315 Administration Bldg.
Memphis, TN 38152-3370
Office: 901.678.2705
Fax: 901.678.2199

points of interest (POI, also called “clusters”). These POI might say “home,” work,” “school,” “car,” or some other generic code for where you have been at a given time. We will use this code in datasets that we share with other researchers, so that they couldn’t know your exact location while you were in the study.

We will share your research data with these other researchers only after:

1. They describe, in writing, how they will use the data.
2. They agree that they will keep your research data secure. They will agree that only people working on their study will be able to see your research data.

We will share your research data from a special database. This database will be saved on a secure computer that is operated at the University of Memphis. People who want to use this data will need permission from the person in charge of this study, Santosh Kumar. Anyone who uses the data from this study will sign a confidentiality agreement, meaning that they must get permission to share the data with anyone else.

After we finish this study, we will also make a copy of the data that will be stripped of all information that could identify you (including raw location data/GPS and your code). This dataset will not have the code that identifies you. We may share this dataset with any researcher who has a use for it.

If you agree, we will also make a dataset that only has the non-identifiable sensor data from the devices (the phone, the wrist sensors, and the toothbrush). This dataset will have no identifiable information and no codes that could link back to you. This dataset can help the wider research community, and we would share it with anyone who could find it useful. This is completely optional. You can agree to take part in the study but not include your data in this open data project.

<u>How long we will keep your data:</u> We expect to complete this study after 3 years. After the study is over, we will we will make a copy of the data that will be stripped of all information that could identify you (including raw location data/GPS, your video, and your code). We will save this copy of your research data as long as we think it is still useful. We expect that the data will be useful for 10 years after the study is over, but we may keep it much longer.

<u>Canceling your permission:</u> If you change your mind later and you don’t want your research data shared with other researchers, you can cancel your permission. To cancel your permission, you have to write a letter to Santosh Kumar. When you write us a letter and cancel your permission, we will delete your information from the database. No new researchers will be able to get a copy of the data. We will not be able to take back the research data from researchers who already have the data.

After we finish this study, we will also make a copy of the data that will be stripped of all information that could identify you (including raw location data/GPS and your code). This

Institutional Review Board

315 Administration Bldg.
Memphis, TN 38152-3370
Office: 901.678.2705
Fax: 901.678.2199

dataset will not have the code that identifies you. If you decide to cancel your permission to share your data with other researchers after we finish this study, we will not be able to find or delete your individual research data from this copy of the dataset.

CAN YOUR TAKING PART IN THE STUDY END EARLY?
You can leave this research study at any time and it will not be held against you. There is no penalty for deciding to leave the study now, or in the future.

If you want to leave the study at any time, contact the study team. We will arrange for the study devices to be returned. If you don't want your data to be used in the study, tell the study team. They will permanently delete any data that has been collected from you as a part of the study.

The individuals conducting the study may need to withdraw you from the study. This may occur if you are not able to follow the directions they give you, if they find that your being in the study is more risk than benefit to you, or if the agency funding the study decides to stop the study early for a variety of scientific reasons.

WHAT IF YOU HAVE QUESTIONS, SUGGESTIONS, CONCERNS, OR COMPLAINTS?

Before you decide whether to accept this invitation to take part in the study, please ask any questions that might come to mind now. If you have questions, concerns, complaints, or think that you've been hurt by the research, call us as soon as possible. You can call Santosh Kumar at telephone number 901-678-2487. You can also call the study coordinator, Shahin Samiei, at 901-678-3369 with any questions about this study.

This research has been reviewed and approved by an Institutional Review Board ("IRB"). The role of the IRB is to protect the rights and welfare of people who take part in research studies. You may contact the IRB at 901-678-2705 or irb@memphis.edu if you have any questions about your rights as a volunteer in this research. Some reasons you might want to contact the IRB are if:

- Your questions, concerns, or complaints are not being answered by the research team.
- You cannot reach the research team.
- You want to talk to someone who is not on the research team.
- You want to get more information or provide input about this research study.
- You have any questions about your rights as a research participant.

WHAT IF NEW INFORMATION IS LEARNED DURING THE STUDY THAT MIGHT AFFECT YOUR DECISION TO PARTICIPATE?

If the researcher learns of new information in regards to this study, and it might change your willingness to stay in this study, the information will be provided to you. You may be asked to sign a new informed consent form if the information is provided to you after you have joined the study.

Institutional Review Board

315 Administration Bldg.
Memphis, TN 38152-3370
Office: 901.678.2705
Fax: 901.678.2199

WHAT ELSE DO YOU NEED TO KNOW?

This study is supported by a grant from the National Institutes of Health (Grant #R01DE025244). This study is designed to research how using data from mobile sensors can be used to inform and improve health outcomes. This study is a part of collaborative research between the University of Memphis, the University of California Los Angeles, and the Ohio State University.

Institutional Review Board

315 Administration Bldg.
Memphis, TN 38152-3370
Office: 901.678.2705
Fax: 901.678.2199

Your signature documents your permission to take part in this research. By signing this consent form, you agree that we can share your coded research data with other researchers who are working on this study. You also agree that we can share your de-identified research data with any other researchers.

______________________________________ ___________
Signature of person agreeing to take part in the study Date

____
Printed name of person agreeing to take part in the study

______________________________________ ___________
Name of [authorized] person obtaining informed consent Date

If you agree, we will also make a dataset that only has sensor data from the devices (the phone, the wrist sensors, and the toothbrush) that cannot identify you as an individual. This dataset will have no identifiable information and no codes that could link back to you. This dataset can help the wider mobile health (mHealth) community to learn how mobile sensor data can improve health. We would post this dataset on a publicly accessible website for anyone to download. This is completely optional. You can agree to take part in the study but not include your data in this open data project. By signing next to “Yes,” you agree that we can keep and share the de-identified sensor signals that we collect from you on a publicly accessible website.

Will you allow us to share the sensor data we collect for this study on a publicly accessible website?

☐ YES ______________________

Signature of Participant

☐ NO ______________________

Signature of Participant

______________________________________ ___________
Name of [authorized] person obtaining informed consent Date

Office use

Data entry date & initials.____________
Verification date & initial.____________

Date: ______________________

Subject ID: ________________

ROBAS Equipment and Experience Study Exit Questionnaire

Thank you very much for your participation in this study! Please think about your experience with the mobile phone, smart toothbrush, and the wearable sensors over the last seven days and please indicate to what extent you agree or disagree with the following statements. Please circle one answer for each statement below.

Over, please →

Experience with Study Smartphone

1. The phone interfered with my daily activities.
   a) Strongly agree
   b) Agree
   c) Disagree
   d) Strongly disagree

2. The phone interfered with my social interactions.
   a) Strongly agree
   b) Agree
   c) Disagree
   d) Strongly disagree

3. I felt self-conscious using the phone in public.
   a) Strongly agree
   b) Agree
   c) Disagree
   d) Strongly disagree

4. I felt self-conscious taking videos with the phone.
   a) Strongly agree
   b) Agree
   c) Disagree
   d) Strongly disagree

5. Overall, the phone was easy to use.
   a) Strongly agree
   b) Agree
   c) Disagree
   d) Strongly disagree

6. Overall, the phone was a nuisance.
   a) Strongly agree
   b) Agree
   c) Disagree
   d) Strongly disagree

7. Overall, the phone was enjoyable to use.
   a) Strongly agree
   b) Agree
   c) Disagree
   d) Strongly disagree

Experience with Smart Toothbrush

8. Using the smart toothbrush interfered with my daily activities.
   a) Strongly agree
   b) Agree
   c) Disagree
   d) Strongly disagree

9. I felt self-conscious using the smart toothbrush.
   a) Strongly agree
   b) Agree
   c) Disagree
   d) Strongly disagree

10. The smart toothbrush caused physical discomfort.
    a) Strongly agree
    b) Agree
    c) Disagree
    d) Strongly disagree

11. Overall, the smart toothbrush was easy to use.
    a) Strongly agree
    b) Agree
    c) Disagree
    d) Strongly disagree

12. Overall, the smart toothbrush was a nuisance.
    a) Strongly agree
    b) Agree
    c) Disagree
    d) Strongly disagree

13. Overall, the smart toothbrush was enjoyable to use.
    a) Strongly agree
    b) Agree
    c) Disagree
    d) Strongly disagree

14. I felt more engaged in the study knowing that I could keep the smart toothbrush for my participation in the study.
    a) Strongly agree
    b) Agree
    c) Disagree
    d) Strongly disagree

Over, please →

Experience with Wrist Band Usage

15. The wrist band interfered with my daily activities.
    a) Strongly agree
    b) Agree
    c) Disagree
    d) Strongly disagree

16. The wrist band interfered with my social interactions.
    a) Strongly agree
    b) Agree
    c) Disagree
    d) Strongly disagree

17. I felt self-conscious using the wrist band in public.
    a) Strongly agree
    b) Agree
    c) Disagree
    d) Strongly disagree

18. The wrist band caused physical discomfort.
    a) Strongly agree
    b) Agree
    c) Disagree
    d) Strongly disagree

19. Overall, the wrist band was easy to use.
    a) Strongly agree
    b) Agree
    c) Disagree
    d) Strongly disagree

20. Overall, the wrist band was a nuisance.
    a) Strongly agree
    b) Agree
    c) Disagree
    d) Strongly disagree

21. Overall, the wrist band was enjoyable to use.
    a) Strongly agree
    b) Agree
    c) Disagree
    d) Strongly disagree