Data in Brief 37 (2021) 107248

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Dataset for toothbrushing activity using brush-attached and wearable sensors

Zawar Hussain[a,*], David Waterworth[a], Adnan Mahmood[a], Quan Z. Sheng[a], Wei Emma Zhang[b]

[a] Department of Computing, Macquarie University, Sydney, Australia
[b] School of Computer Science, The University of Adelaide, Australia

ARTICLE INFO

Article history:
Received 7 April 2021
Accepted 24 June 2021
Available online 2 July 2021

Keywords:
Toothbrushing
Smart toothbrush
Activity recognition
Machine learning
Sensor

ABSTRACT

Maintaining oral hygiene is very important for a healthy life. Poor toothbrushing is one of the leading causes of tooth decay and other gum problems. Many people do not brush their teeth properly. There is very limited technology available to help in assessing the quality of toothbrushing. Human Activity Recognition (HAR) applications have seen a tremendous growth in recent years. In this work, we treat the adherence to standard toothbrushing practice as an activity recognition problem. We investigate this problem and collect experimental data using a brush-attached and a wearable sensor when the users brush their teeth. In this paper, we extend our previous dataset [1] for toothbrushing activity by including more experiments and adding a new sensor. We discuss and analyse the collection of the dataset. We use an Inertial Measurement Unit (IMU) sensor to collect the time-series data for toothbrushing activity. We recruited 22 healthy participants and collected the data in two different settings when they brushed their teeth in five different locations using both electric and manual brushes. In total, we have recorded 120 toothbrushing sessions using both brush-attached sensor and the wearable sensor.

* Corresponding author.
E-mail addresses: zawar.hussain@hdr.mq.edu.au (Z. Hussain), david.waterworth@hdr.mq.edu.au (D. Waterworth), adnan.mahmood@mq.edu.au (A. Mahmood), michael.sheng@mq.edu.au (Q.Z. Sheng), wei.e.zhang@adelaide.edu.au (W.E. Zhang).
Social media: (Z. Hussain)

https://doi.org/10.1016/j.dib.2021.107248

Specifications Table

Subject	Biomedical Engineering
Specific subject area	Toothbrushing analysis, human activity recognition, machine learning, classification, feature engineering
Type of data	CSV files
How data were acquired	The data was collected using two sensor devices from Mbientlab Inc [2]. One device was attached to the brush handle while the other device was used as wearable (wrist watch on the brushing hand). The participants brush their teeth for two minutes in each session following a pre-given sequence and the data was logged onto the device memory which was then transferred to laptop via Bluetooth.
Data format	Raw
Parameters for data collection	The sensor device consists of accelerometer, gyroscope and magnetometer which can capture the brush movements. We record 3-axis accelerometer, 3-axis gyroscope and 3-axis magnetometer data for the brush movements when the participants brush different regions of the mouth.
Description of data collection	The participants were given instructions (e.g., which side to be brushed first and then second etc.,) of how to brush their teeth. The experiments are performed in different washrooms. The participants brushed their teeth in natural way using toothpaste. Different types of electric and manual brushes were included in the experiments. The average duration for each experiment was two minutes. The participant repeated the same experiment next day. We recorded minimum two sessions and maximum five sessions for each participant. After each experiment, the data was transferred to laptop via Bluetooth.
Data source location	Institution: Department of Computing, Macquarie University City: Sydney Country: Australia
Data accessibility	https://data.mendeley.com/datasets/hx5kkkbr3j/1

Value of the Data

- This dataset set is very useful for research in toothbrushing analysis. This dataset provides real toothbrushing data using different brush types. It can help the researchers in the designs of smart toothbrushes as well as for developing automated machine learning algorithms to check the compliance with standard toothbrushing techniques.
- This dataset can help biomedical engineers in designing smart toothbrushes. It can be very useful for researchers in the field of artificial intelligence to design and test machine learning models to detect the compliance to standard toothbruhing practices. It can also be used to evaluate the performance of classification algorithms and feature engineering techniques.
- This dataset can be used in multiple ways. It can be used to evaluate the noise removal techniques as this dataset is in raw form. For example, the dataset from setting-1 consists of back-to-back regions with gaps between each region (a region corresponds to the surface of the jaw being brushed). It can be used for validating segmentation techniques. Also, new machine learning techniques for supervised and un-supervised classification and feature selection can be evaluated using this dataset.

1. Data Description

Pervasive computing is getting very popular in the area of human activity recognition especially for health related application due to its low-cost, easy deployment and ubiquitous nature [3]. Different types of low cost sensor are used for recognizing various activities such as sleep [4], falls, [5] and many other activities of daily life [6,7]. We consider the toothbrushing monitoring as an activity recognition problem and divide the toothbrushing activity into 16 sub-activities corresponding to the 16 regions of the teeth [8]. These regions are shown in Table 1.

When a user adjusts the brush to reach different regions of the mouth, the attached-device experiences rotations and accelerations shown in Fig. 1. We record these measurements from the sensor devices attached to the brush handle and the wrist because the hand orientation also changes while brushing different regions. The data collected can capture the brush movements and can be used to recognize the teeth regions being brushed.

We collected the data in two settings: setting-1 and setting-2. In setting-1, the participants brushed their teeth in a natural way without making any pauses in between different regions of the teeth (continuous brushing). In setting-2, the participants were asked to pause for a few seconds in between regions i.e., bring the brush to a reference point and wait for a few seconds

Table 1
Sub-activities in toothbrushing.

1. Left Lower Jaw Front	9. Right Lower Jaw Back
2. Left Lower Jaw Top	10. Right Upper Jaw Front
3. Left Lower Jaw Back	11. Right Upper Jaw Top
4. Left Upper Jaw Front	12. Right Upper Jaw Back
5. Left Upper Jaw Top	13. Lower Incisors Front
6. Left Upper Jaw Back	14. Lower Incisors Back
7. Right Lower Jaw Front	15. Upper Incisors Front
8. Right Lower Jaw Top	16. Upper Incisors Back

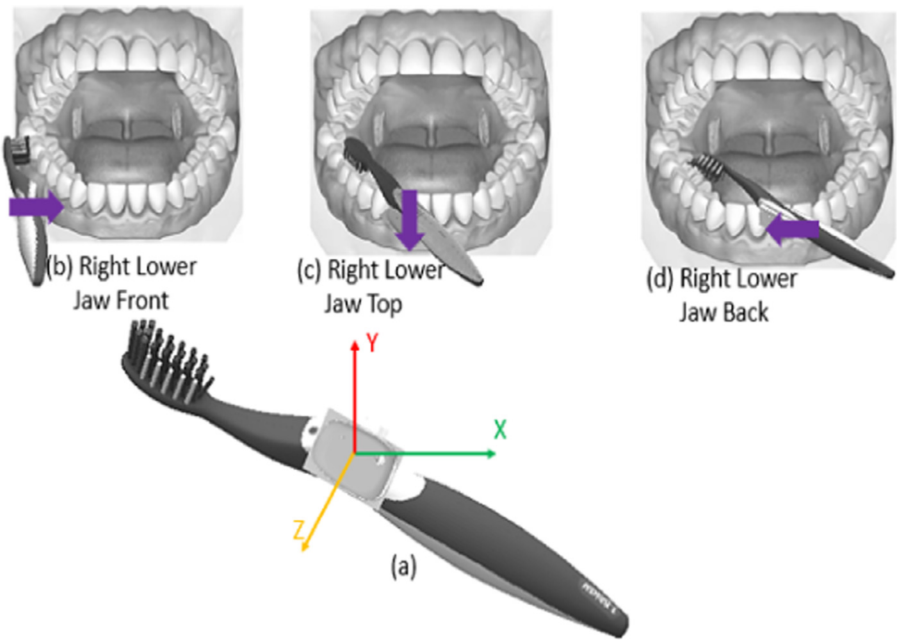


Fig. 1. Sensor's orientation changes while brushing different regions.

Table 2
File name description.

Notation	Description
S	It represents settings of the sessions as we have performed experiments in two different settings.
S	It represents the subject number.
S	It represents the session number
M/F	It represents the gender of the subject. M for male while F for female
L/R	It represents the brushing hand of the subject. R for right and L for left.
A/W	It represent the position of the sensor. A for brush-attached and W for wearable.
YY	It represents the age of the participant in years
M/E	It represents the type of the brush used. M for manual and E for electric brush.
L	It represents the location of the experiment as we have performed experiments in different locations.
A/G/M	It represents the type of sensor. A for accelerometer, G for gyroscope and M for magnetometer.

Table 3
Column description.

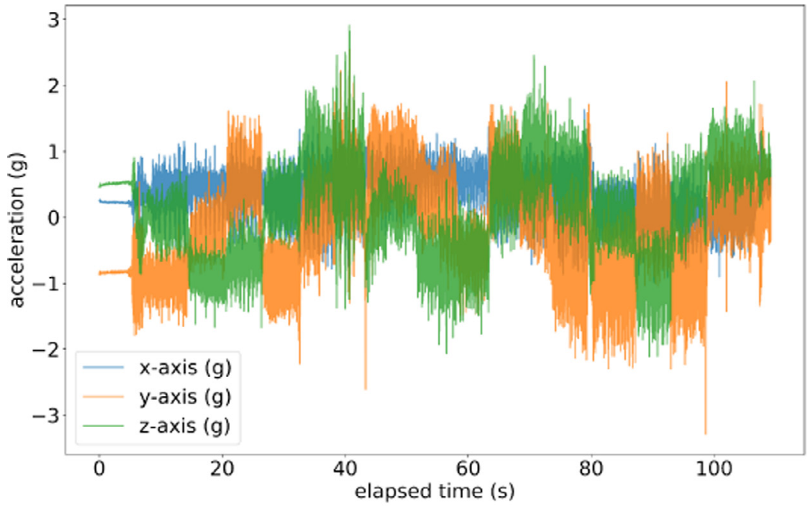
Column name	Description
epoc (ms)	unix epoch time
timestamp (+1000)	date and time in local timezone
elapsed (s)	elapsed time since sensor recording started
x-axis (g)	*x*-axis from raw accelerometer data
y-axis (g)	*y*-axis from raw accelerometer data
z-axis (g)	*z*-axis from raw accelerometer data
x-axis (deg/s)	*x*-axis from raw gyroscope data
y-axis (deg/s)	*y*-axis from raw gyroscope data
z-axis (deg/s)	*z*-axis from raw gyroscope data
x-axis (T)	*x*-axis from raw magnetometer data
y-axis (T)	*y*-axis from raw magnetometer data
z-axis (T)	*z*-axis from raw magnetometer data

before moving to the next region. The purpose of the setting-2 is to ease the process of segmentation because in setting-1, the regions are back-to-back which is very challenging for any segmentation technique [9]. A sample of the data collected for the brush-attached accelerometer in setting-1 and setting-2 are shown in Fig. 2.

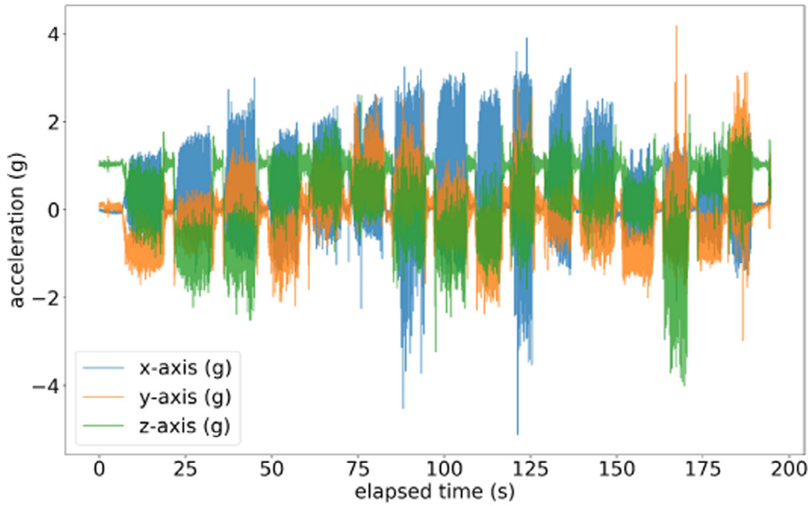
The collected dataset consists of toothbrushing data for 120 sessions performed by 22 participants (11 males, 11 females). Some of the participants are common in both settings i.e., they participated in both the settings. The age group of the participants is 22-40 years. We collected a minimum of two sessions and a maximum of five sessions (in each setting) for each participant. The average duration of each session is two minutes for setting-1 and three minutes for setting-2 (the extra one minute is because of the pauses in-between the regions). The dentists recommend to brush the teeth twice a day for two minutes [10]. The dataset contains 120 folders, one for each session. Each folder consists of four (for setting-1) or six (for setting-2) files. For setting-1, there are four files in each session: one file per sensor (accelerometer and gyroscope) for both wearable and brush attached-device. In setting-2, there are six files in each session: one file per sensor (accelerometer, gyroscope, and magnetometer) for both brush-attached and wearable device. The name of each data file consists of multiple letters. The details for the naming convention are given in Table 2. Each file consists of multiple columns. The details of the data columns are given in Table 3.

2. Experimental Design, Materials and Methods

To collect the toothbrushing data, we used an IMU called MMR developed by Mbientlab Inc [2] which is very light weight (0.2 oz) and small in size ($27 \times 27 \times 4$ mm), and can be easily



(a) Setting-1



(b) Setting-2

Fig. 2. A sample of toothbrushing data captured from brush-attached accelerometer sensor. The first example is from setting-1 with no gap between sub-activities. The second is for setting-2 where the participants paused between sub-activities.

attached to the handle of the brush. MMR consists of multiple sensors but we only use the accelerometer, gyroscope and magnetometer in our experiments. One MMR device was attached to the brush handle as shown in Fig. 1(a) while the other was worn by the users on their brushing hand as wristwatch during the experiments. We collected the data in two separate settings (setting-1 and setting-2). In both the settings, the participants were given instructions before the experiments about the sequence of the regions to be brushed i.e., which region to start with and which region next and so on. All sessions contains one example of each of the 16 sub-activities from Table 1 in the specific order outlined in the table. In setting-1, the participants brushed their teeth in natural way (i.e., without pausing in between regions) and we collected a total of 62 sessions for 17 participants over one week in 5 different locations. In setting-1, we collected the data from two sensors (a 3-axis accelerometer and a 3-axis gyroscope) embedded in both brush-attached device and the wearable device. In setting-2, the participants brushed their teeth with having small gapes in between regions and we collected a total of 58 sessions for 17 participants over one week in 5 different locations. In setting-2, we collected the data for three sensors (a 3-axis accelerometer, a 3-axis gyroscope, and a 3-axis magnetometer) from both brush-attached device and the wearable device. For all the experiments, the sampling rate for accelerometer and gyroscope is 200 Hz while for magnetometer it is 20 Hz.

In both settings, five participants used electric brushes while the rest of the participants used different models of manual brushes. Most of the participants brushed their teeth once every day for 4–5 days while some of them brushed their teeth with gap of couple days (due to their busy schedule). All the experiments lasted for 120 s on average and the data from both brush-attached and wearable sensors were logged on to the on-chip memory of the devices. After each experiment, the data was transferred to a smart phone using Bluetooth device and stored on cloud (One Drive).

Ethics Statement

All the participants of the study were informed before the experiments and their written consent was obtained using a consent form. They also allowed their data to be shared publicly. This study was approved by the faculty ethics subcommittee at Macquarie university (reference number: 52020898720470)

CRediT Author Statement

Zawar Hussain: Conceptualization, Methodology, Data curation, Writing - original draft; **David Waterworth:** Data curation, Writing - review & editing; **Adnan Mahmood:** Writing - review & editing; **Quan Z. Sheng:** Supervision, Writing - review & editing; **Wei Emma Zhang:** Supervision, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships which have, or could be perceived to have, influenced the work reported in this article.

Acknowledgments

The work of Zawar Hussain is funded by the International Research Training Program (iRTP) of the Government of the Commonwealth of Australia (Allocation No. 2017587). The work of

David Waterworth is funded by the RoZetta Institute, the CSIRO, and CIM Pty Ltd. Adnan Mahmood's work is supported by the Government of the Commonwealth of Australia via its International Research Training Program (Allocation No. 2017560) and Macquarie University, Australia's Postdoctoral Research Fellowship. Quan Z. Sheng's work is partially supported by Australian Research Council (ARC) Discovery Project DP210101723, ARC Linkage Project LP190100140, ARC Linkage Infrastructure, Equipment and Facilities Project LE180100158.

References

[1] Z. Hussain, D. Waterworth, M. Aldeer, W.E. Zhang, Q.Z. Sheng, Toothbrushing data and analysis of its potential use in human activity recognition applications: dataset, in: Proceedings of the Third Workshop on Data: Acquisition To Analysis, 2020, pp. 31–34.
[2] MbientLab, Mmr - metamotionr, 2020. URL: https://mbientlab.com/metamotionr/. [Online; accessed 9-Aug-2020].
[3] U. Varshney, Pervasive healthcare and wireless health monitoring, Mob. Netw. Appl. 12 (2) (2007) 113–127.
[4] Z. Hussain, S. Sagar, W.E. Zhang, Q.Z. Sheng, A cost-effective and non-invasive system for sleep and vital signs monitoring using passive RFID tags, in: Proceedings of the 16th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services, 2019, pp. 153–161.
[5] W. Banno, N. Shinomiya, Monitoring system for the elderly on staircase using passive RFID sensor tags, in: 2019 IEEE 8th Global Conference on Consumer Electronics (GCCE), IEEE, 2019, pp. 816–817.
[6] J. Lu, X. Zheng, Q.Z. Sheng, Z. Hussain, J. Wang, W. Zhou, MFE-HAR: multiscale feature engineering for human activity recognition using wearable sensors, in: Proceedings of the 16th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services, 2019, pp. 180–189.
[7] Z. Hussain, Q.Z. Sheng, W.E. Zhang, A review and categorization of techniques on device-free human activity recognition, J. Netw. Comput. Appl. 167 (2020) 102738.
[8] H. Huang, S. Lin, Toothbrushing monitoring using wrist watch, in: Proceedings of the 14th ACM Conference on Embedded Network Sensor Systems CD-ROM, 2016, pp. 202–215.
[9] A. Khamis, B. Kusy, C.T. Chou, M.-L. McLaws, W. Hu, RFWash: a weakly supervised tracking of hand hygiene technique, in: Proceedings of the 18th Conference on Embedded Networked Sensor Systems, 2020, pp. 572–584.
[10] M. Poyato-Ferrera, J. Segura-Egea, P. Bullón-Fernández, Comparison of modified bass technique with normal toothbrushing practices for efficacy in supragingival plaque removal, Int. J. Dental Hyg. 1 (2) (2003) 110–114.