Vol.2022-MBL-104 No.27
Vol.2022-UBI-75 No.27
Vol.2022-CDS-35 No.27
Vol.2022-ASD-24 No.27
2022/9/6

A preliminary study for monitoring hygiene behaviors by using multiple sensors on a wrist

Haoyu Zhuang[1,a)] Liqiang Xu[1,b)] Yuuki Nishiyama[1,2,c)] Kaoru Sezaki[1,2,d)]

Abstract: Under the epidemic of COVID-19, it is important to automatically detect epidemic protective behaviors without a user's intention. Existing studies utilized only sensor data from IMU for detecting epidemic protection behaviors. However, the performance of the classification for similar behaviors could be unsatisfactory due to the single data dimension. It is well known that washing hands and hand sterilization are essential personal hygiene behaviors. In this paper, we use multiple sensor data from an off-the-shelf smartwatch and smartphone for detecting these three behaviors. Our performance evaluation indicated that our proposed method has improved accuracy for classifying the target epidemic protective behaviors over previous methods. Furthermore, for applying our method in reality, we developed a prototype for detecting these behaviors on a wearable device, which allows us to utilize our method widely in health habits monitoring.

Keywords: Human Activity Recognition, Smart Sensors, Multimodal classification

1. Introduction

Since the pandemic of COVID-19, personal hygiene is received more attention to decrease the risk of infections. It is well known that frequent hand washing and disinfection is a crucial way to reduce the risk of infection. As shown in Fig. 1, the World Health Organization has also issued guidance[12] on hand washing practices. On other hand, hand hygiene is even more vital for specific workers in the food and healthcare industry. However, people often overlook hand washing or disinfection. Identifying these personal hygiene behaviors automatically will help people improve their hygiene practices and reduce their risk of infection.

Over the last decade, smart devices are more popular in people's daily life. Along with this rise are the increasing amount of data and ubiquitous computing. Existing environmental sensors are the basis for real-time data collection and identification of human activity, especially portable smart devices. The development of smartwatch technology and applications has led researchers to conduct many studies regarding human behavior detection based on smartwatches[17] because the smartwatch contains multiple sensors with low-energy consumption.

Inertial sensors on smartwatches can easily capture wrist motion characteristics. However, it is a challenge for hand behavior classification due to the frequent hand movements and the similarity of various types of movements in terms of motion characteristics.

Additionally, acoustic sensors on a smartwatch can capture human-centered environmental sounds, which are contextualized to human activities. Although in the real world, collected audio data can contain massive noise, acoustic data has two advantages over inertial data. Acoustic data is comprehensive for humans, leading to be extensively used in context recognition[7][1]. On the other hand, certain activities can produce sounds with specific features, thus acoustic features can contribute to classifying similar behaviors.

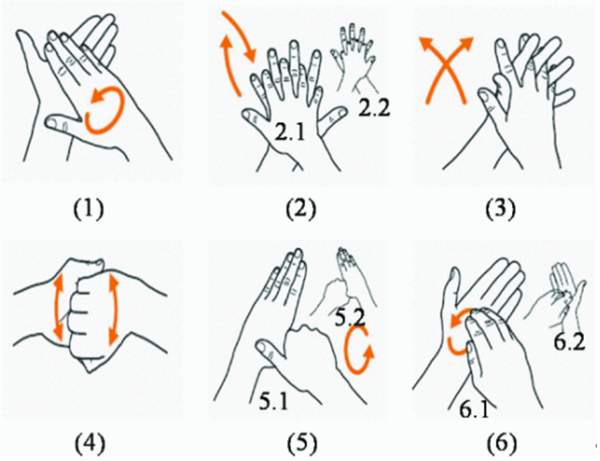


Fig. 1 Hand washing Guideline by World Health Organization (WHO).Figure from[14]

In this work, we explore to accurately classify hygiene behaviors based on inertial and acoustic data with commodity smartwatches. The hand motion data collected by the smartwatch contains a large number of hand motion features that can be used to recognize daily activities, including common personal hygiene activities such as hand washing and disinfection. By combining acoustic data, we expect that the performance of classification can become better. The specific contributions of this work are:

[1] Institute of Industrial Science, The University of Tokyo
[2] Center for Information Science, The University of Tokyo
[a)] zhuanghaoyu@mcl.iis.u-tokyo.ac.jp
[b)] xuliqiang@mcl.iis.u-tokyo.ac.jp
[c)] yuukin@iis.u-tokyo.ac.jp
[d)] sezaki@iis.u-tokyo.ac.jp

- Hygiene-related inertial and acoustic data is collected by the smartwatch in laboratary. Our dataset contains 8 participants, each of whom performed 9 activities, each lasting 1 minute.
- Processing the acoustic and inertial datasets by deep learning frameworks and the mixed data deep learning model was 8% more accurate in hand washing and sanitising behaviour when compared to the classification accuracy of the baseline model.
- In terms of overall classification accuracy, we achieved a small optimization of the classification results by combining multimodal data for classification.

2. Related works

2.1 Activity recognition based on inertial data

A variety of behavior detection models based on smartwatches have been proposed by researchers. Behavioral detection of sports-related human activities[18] such as running, cycling, stair climbing, and squatting has been widely used in several applications and was beneficial to personal health.

Regarding hand hygiene, For example, the detection of social distance based on hands motion data[16] has been explored with promising results. Recently, some studies on face-touching behavior recognition have been proposed[2][6][13], and the accuracy of this model is quite impressive. In addition, exploring the detection of wearing masks[10] based on smartwatches during epidemics is also a popular theme.

Innovative classification methods for hand washing relied on smartwatches are raised[11]. In this study, researchers utilized output from the penultimate layer of the neural network for behavior recognition. This approach saved computational resources but only yielded an F1-Score of about 0.78. This accuracy is not sufficient due to the real-world complicate situations where hand-washing behavior needs to be detected, especially now with the epidemic. In some articles, researchers have used traditional machine learning methods to classify hand washing with other behaviors, such as SVM with KNN[9], however, the team only divided the dataset into three categories, hand washing, hand rubbing, and other activities, which is not effective enough in the utilization of the data, and we expect that we can make full use of the collected data to identify most of the daily hand activities successfully.

2.2 Activity recognition based on Multi-sensor

Research on the use of acoustic data for human activity recognition is not a recent development. Stork *et al.* identified 22 activities using acoustic data[15]. Furthermore, a more accurate and realistic approach was recently proposed by researchers[4]. By exploring the application of inertial sensors and acoustic sensors on off-the-shelf commodity smartwatches for daily activity recognition, a model using multi-sensor data to classify daily activities performs well on the in-the-wild dataset.

Compared to previous studies, our study differs in several aspects. First, we focused on analyzing personal hygiene behaviors, especially hand hygiene behaviors. Second, we relied solely on smartwatches to collect inertial and acoustic data streams, with no externally placed sensors on the environment or the human body. Finally, we tried to combine acoustic data with inertial data and expected to enhance the classification accuracy of personal hygiene behaviors.

3. Motivation

Currently, although there are many studies on personal hygiene behaviour recognition, most of them only used inertial data as input, neglecting the application of acoustic data. However, research has demonstrated that the input of acoustic data can enhance the classification accuracy of deep learning models. Therefore, we attempted to apply both inertial and acoustic data to improve the accuracy of personal hygiene behaviour recognition.

Our objective is to evaluate and analyze inertial-acoustic data collected from a smartwatch for personal hygiene behaviors recognition. Using the data collected, we aimed to explore the following questions:

- How do the individual models perform? Is a model that incorporates acoustic data more accurate than a purely inertial data model?
- Which activities show greater variation in performance across models?
- What is the impact of different Frame sizes on model classification?

4. Dataset

The section presents the dataset founding via a supervised study with 8 participants implementing various activities in laboratory. Firstly we introduce the activities set and then presents hardware, experimental procedure, and data processing.

4.1 Types of activities

To ensure that the classification results are informative for most everyday scenarios, our main objective is to select hand behaviors that are frequent and relatively similar in terms of movement patterns. Ultimately, we selected nine categories of typical hand behaviors shown in Fig. 2, including hand washing, eating snacks, drinking, sanitizing, washing cups, writing, clapping, viewing a mobile phone screen, and strict hand washing based on WHO guidelines.

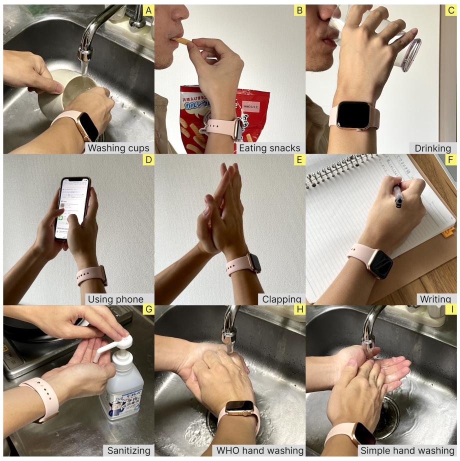
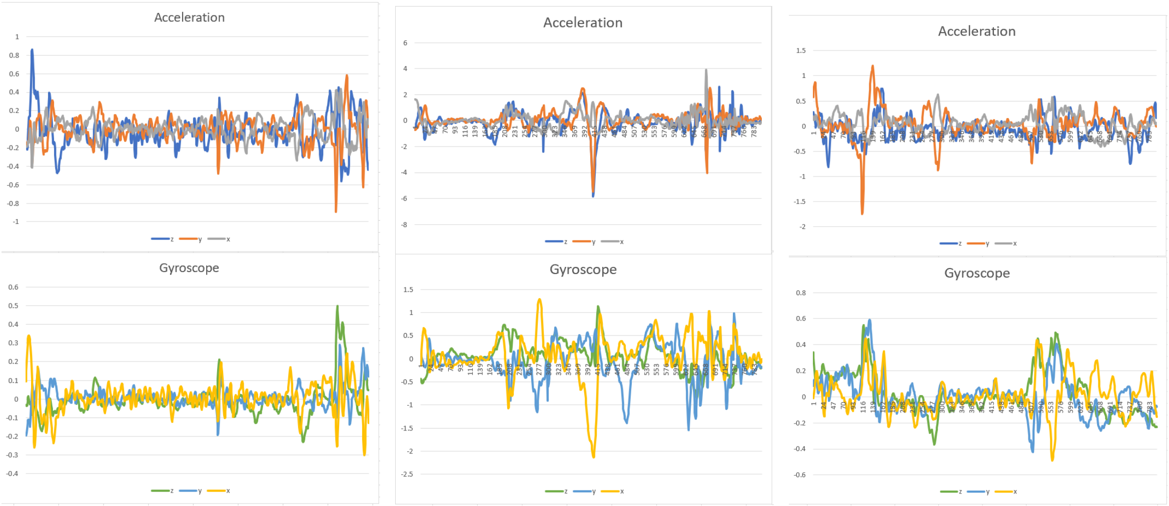


Fig. 2 Nine types of hand behaviors captured by a commodity smartwatch

The real author is the Editorial Board of JIP.



(a) Clapping (b) Simple hand washing (c) Sanitizing

Fig. 3 Example of raw inertial data of 3 activities in 8s clips

4.2 Data collection

We ensured the availability of the experimental apparatus in advance before the beginning of the experiment. We used an Apple Watch and an iPhone11 with iOS 15.0.2 on it and downloaded an app named Sensor Logger. The app supported us in collecting IMU data through the watch's motion sensor at a sampling rate of 100Hz. Considering that our target was to capture the hand motion patterns of daily activities and the participants' dominant hand is the right hand., the smartwatch was worn on the wrist of the right hand.

We ran a study with participants in our laboratory to capture experiment data. During the experiment, participants performed a variety of activities successively lasting 1 minute. Data collection was discontinuous from the beginning of the 1st activity to the ending of the last activity, in other words, which didn't involve any in-between movements. Between activities, participants were asked to familiarize themselves with the next activity and the experimenter changed the data file name of the previous activity to the label, which is a part of data annotation. At the end of all experiments, we labeled all activities with data according to the file names.

4.3 Data Processing

Sixteen-axis IMU data was sampled at 100Hz and saved as data files with different names. The raw acoustic data was saved as a file with a sampling rate of 22.05 kHz. Raw inertial data are shown in Fig. 3 During data pre-processing, we merged the inertial data sampled at the same time and transferred the acoustic data file to .wav format. The frame size was important to the performance of the final model[5]. Therefore, we chose many frame sizes to compare the classification performance and the effect of frame size on the model classification results will be discussed in detail later.

5. Hygiene Behavior Detection Methods using Hand Motion and Audio

Given the collection of the inertial and acoustic data, the crucial step of system building is to apply different models to train and classify the data. Our intent is to examine and promote personal hygiene activities recognition with the combination of inertial and acoustic sensors from smartwatches. To that end, we utilized classical machine learning and deep learning frameworks with different datasets to generate comparable results.

We explored three traditional machine learning models to establish baseline performance: (1) Support vector machine classifier; (2) K-Nearest Neighbor Classifier; (3) XGBoost Classifier as the base classifier. Simple standard machine learning models were deployed. Six frame-level statistical features, including mean, median, variance, maximum, minimum and root mean square were extracted from inertial data per axis separately and Mel Frequency Cepstral Coefficients (MFCC) were computed from acoustic data. Within the acoustic data, the segmented frames extracted 20 MFCCs and averaged across them.

In addition to traditional machine learning methods, we adopted the convolutional neural network (CNN) model as the primary model for processing inertial and acoustic data in our experiments, due to the excellent performance of CNN in acoustic processing and time-series data processing. We expect that the training results of CNN models can be significantly better than traditional machine learning methods.

For acoustic classification, we decided to extract features from log-Mel spectrograms that have been proved effective in prior work[3]. By computing the short-time Fourier transform (STFT), we extracted the spectrograms of our audio data, using a Hanning window of 1024 samples and a hop size of 512 samples. Example log Mel spectrograms of each activity class are shown in Fig. 4.

About data fusion, traditional fusion strategies include representation-level fusion[3] and score-level fusion[8]. We believe that the data fusion strategy should be able to combine the motion features represented by IMU data with the acoustic features represented by acoustic data, resulting in a more accurate output from the framework. we used two data fusion techniques: (1) a simple concatenation of feature maps. The feature maps from the penultimate layer of the inertial and acoustic data modules are concatenated and then classified by the softmax layer. The method can reduce the impact of inter-class similarity on classification accuracy in a single dimension by concatenating features. (2) Score-level fusion is to classify the inertial data module and acoustic data module separately and average the class probabilities to get the final probabilities, which is more straightforward and can combine the outputs of two convolutional models to achieve better classification results. The whole framework is shown in Fig. 5.

6. Evaluation and Results

The models were trained and classification results were obtained in the section 5. For the evaluation of the classification results, we used the traditional classification accuracy-based evaluation method, focusing on the behaviours that are easily confused and on the performance of the multi-sensor models.

6.1 Comparison between models

As shown in Table. 1, within our expectations, both CNN models using only inertial motion data and CNN models using only acoustic motion data perform far better than traditional machine learning methods for classification by 15% to 20%. And among the traditional machine learning methods, the best performer is the XGBoost model. The models based on IMU and acoustic data by concatenating fusion and score fusion methods were better than models using only acoustic or motion data. However, this difference is not significant. As far as the F1-score is concerned, the model using concatenating strategy is 2% higher than which is based on only motion data.

Table 1 Table showing average F1-score for each motion and audio single-modality classification and each combination of fusion strategies. Concatenate fusion means a combination between IMU and acoustic data by concatenating. Score fusion means combining IMU and acoustic data by averaging the class probabilities.

	Mean-F1Score	precision	Recall
SVM	0.69	0.84	0.65
KNN	0.67	0.68	0.67
XGBoost	0.74	0.76	0.74
IMU CNN	0.86	0.87	0.85
Acoustic CNN	0.83	0.83	0.82
Concatenate Fusion	0.90	0.89	0.91
Score Fusion	0.87	0.87	0.87

In the final results, a large discrepancy is in the recognition rate between different activities, as shown in Fig. 6. Among them, clapping had the lowest classification accuracy. Furthermore, the classification of accuracy in the classes drinking, writing, simple hand washing, and WHO hand washing is not over 90%. 10% of clapping and 5% of drinking were confused. 6.9% of simple hand washing and 5.6% of WHO hand washing was confused with washing cups.

Both Fusion strategies significantly improved classification accuracy in terms of key activities simple hand washing, sanitization, and WHO-hand washing. Taking Concanating Fusion as an example, the classification accuracy for hand washing was 10% higher than the CNN model using only IMU data and 8% higher than the CNN model using only acoustic data; the classification accuracy for disinfection was 3% higher than the CNN model using only IMU data and 8% higher than the CNN model using only acoustic data; the classification accuracy for WHO hand washing classification accuracy was 7% higher than the CNN model using only IMU data and 15% higher than the CNN model using only acoustic data.

6.2 Influence of Frame size

The frame size affects the performance of the model, therefore, we investigated the effect of the length of the frame size on the classification results of the model using IMU data.

The frame sizes with 1s, 3s, 5s, and 7s were set experimentally when we did not change other parameters. The results are shown in Fig. 7. Obviously, we can discover that the performance of the CNN model classification became significantly worse as the frame size became larger. The same pattern appeared for the CNN model only using acoustic data.

We conclude that the smaller the frame size, the more samples

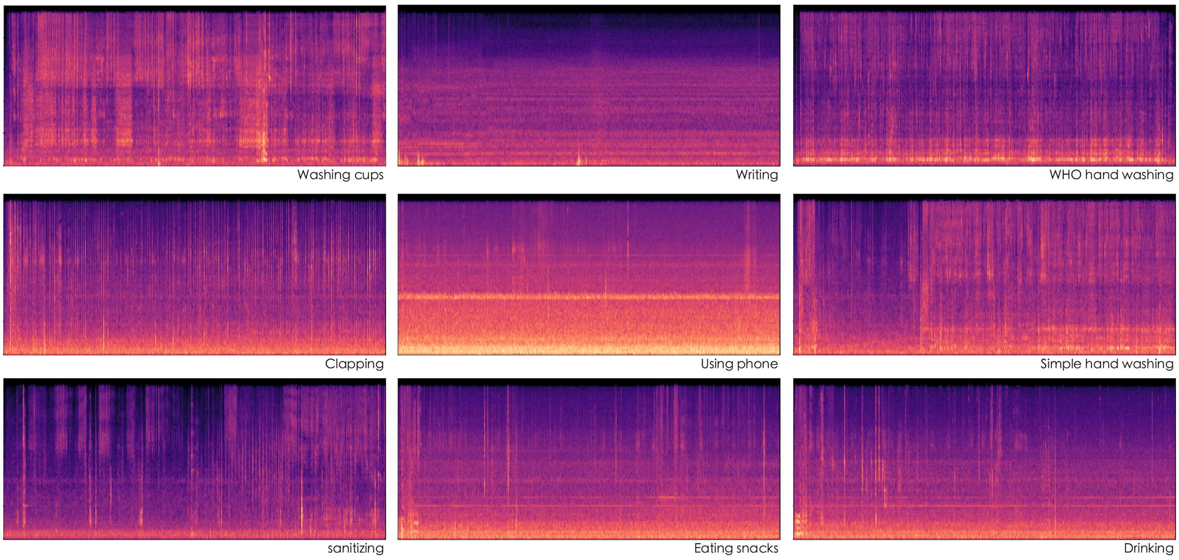
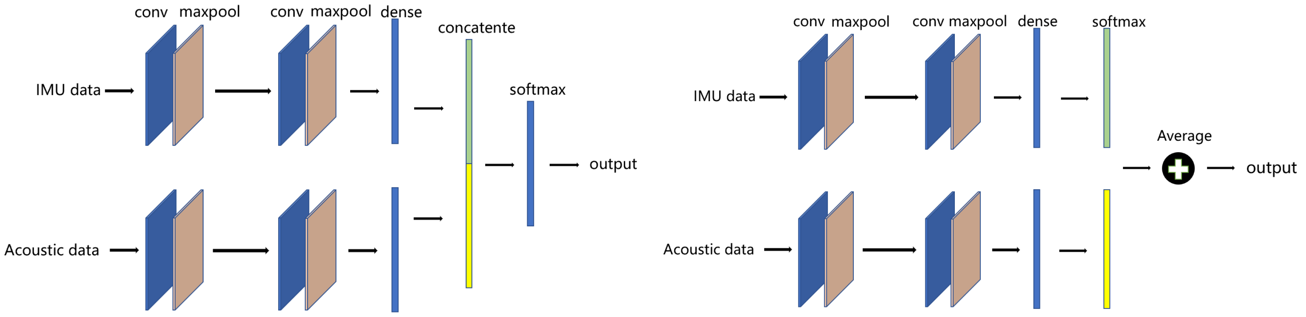


Fig. 4 Example Log Mel spectrograms for 9 activity classes



(a) Represent-level fusion(concatenate)

(b) Score-level fusion

Fig. 5 An overview of the activity recognition framework. The overall training process is shown. Represent-level fusion and score-level fusion are shown separately.

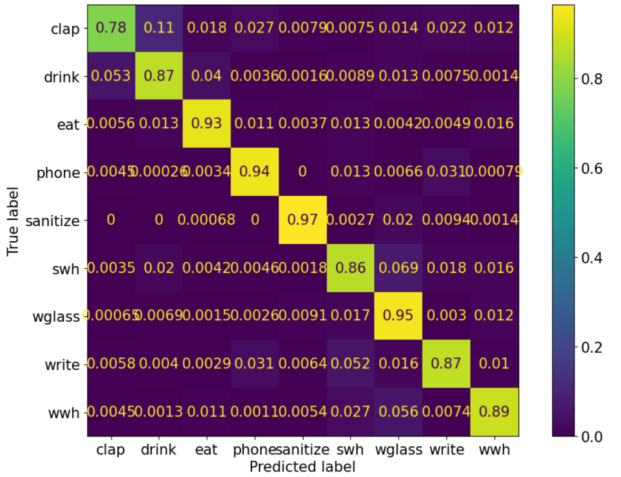


Fig. 6 Confusion matrix of the model using concatenating fusion strategy. The label ”swh” represents simple hand washing, ”wglass” represents washing cups, and ”wwh” represents WHO hand washing.

there are and the more accurate the model will be. Moreover, as there is not much noise in the experimental environment, a single frame contains more useful features than noise, so the smaller the frame size is, the better the model training will be.

7. Discussion

In this section, we will discuss the details of the classification results. Firstly, we will discuss behaviours that are easily confounded in multiple models, and secondly, the advantages of multi-sensor data models and the reasons for them will be mentioned.

7.1 Analysis of confusion-prone behaviours

The inter-class similarity is one of the reasons that affect the model's classification performance. We mentioned in Section 6 those activities that are mixed during classification. The confusion between washing cups and hand washing is understandable. They presented strong agreement on both IMU data and acoustic data, plus the hand washing behavior was not consistent for each experimenter, so those data would likely be classified as washing cups. In fact, among these activities, washing cups and hands washing were the most liberating activity, and we did not specify how participants should perform the two activities.

7.2 Multimodal Sensor Fusion

Disinfection and hand washing, which are more closely related to individual behavior, performed well in the model, with disinfection achieving a 97% classification accuracy. Therefore, it is highly feasible to identify human hygiene behavior, especially disinfection behavior, through deep learning.

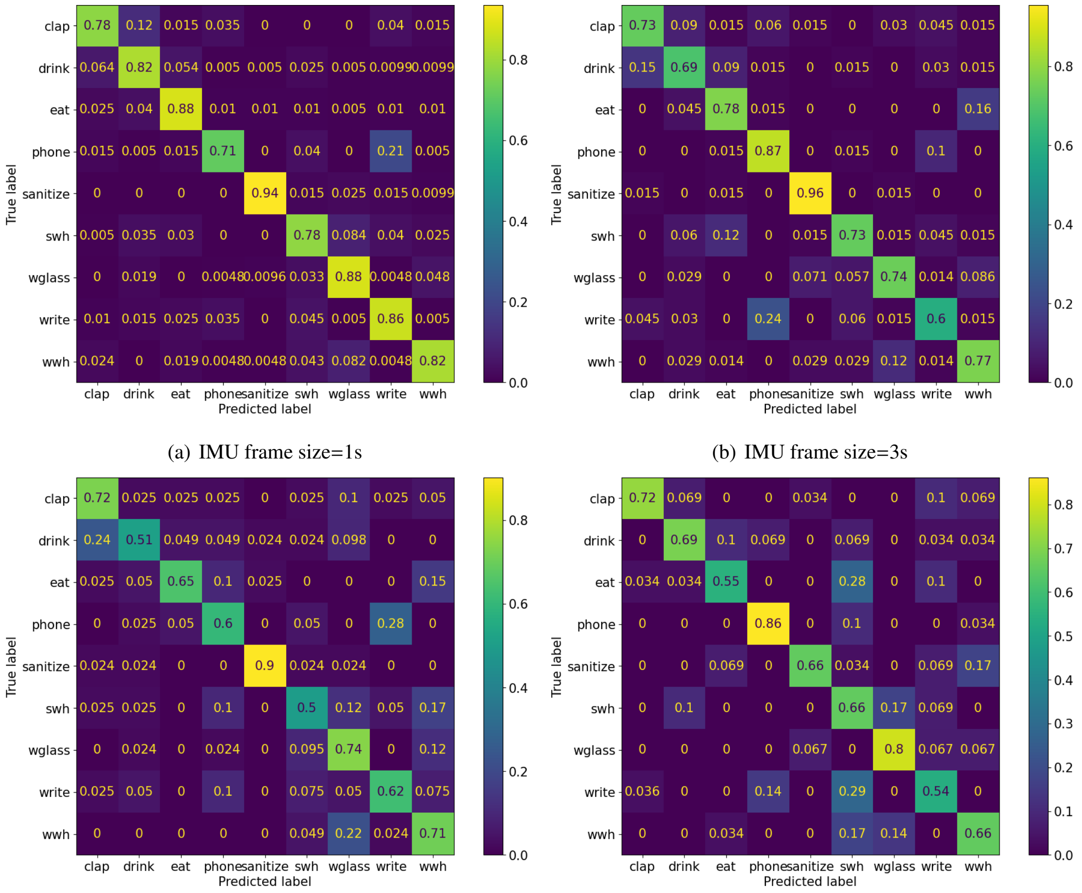
The model did not perform much better when using both IMU data and acoustic data than when using only a single data. We speculate that the reason for this may be due to the smaller sample size, resulting in better classification performance of the model using only inertial data, whereas factors affecting this model classification cannot be masked or eliminated by the acoustic data, such as errors in participant subtle individual behavior in the experiment, for example, errors that occurred when labeling the data.

By combining acoustic and inertial data, models trained under both Fusion strategies do yield higher results than those using a single data set. For more representative activities in terms of personal hygiene behaviors, such as hand washing and disinfection, there is an improvement in classification accuracy of more than 8%, indicating that the combination of multimodal data has a significant improvement in the classification accuracy of personal hygiene behaviors.

8. Limitation

As we wrote in Section 4, all our experiments were conducted in a laboratory environment, so the noise included in the acoustic data is much less compared to the real environment. For the model to be deployed in a real environment, a module or algorithm related to noise removal must be added. Our experimenters were uniform in age level, and the number of activities we collected is still not large compared to daily life, which will affect our model performance. In future research, we can use data augmentation methods to increase the amount of data in our models, in addition to collecting more samples.

In terms of data collection, another concern is the power consumption of smartwatches in continuously acquiring multiple sensor data over long periods. Although the data processing and model training parts are not deployed on the watch, prolonged data collection through the microphone and IMU data collection module will inevitably reduce the normal usage hours of the watch. Some experiments have shown[4] that the usable hours of



True label \ Predicted label	clap	drink	eat	phone	sanitize	swh	wglass	write	wwh
clap	0.78	0.12	0.015	0.035	0	0	0	0.04	0.015
drink	0.064	0.82	0.054	0.005	0.005	0.025	0.005	0.0099	0.0099
eat	0.025	0.04	0.88	0.01	0.01	0.01	0.005	0.01	0.01
phone	0.015	0.005	0.015	0.71	0	0.04	0	0.21	0.005
sanitize	0	0	0	0	0.94	0.015	0.025	0.015	0.0099
swh	0.005	0.035	0.03	0	0	0.78	0.084	0.04	0.025
wglass	0	0.019	0	0.0048	0.0096	0.033	0.88	0.0048	0.048
write	0.01	0.015	0.025	0.035	0	0.045	0.005	0.86	0.005
wwh	0.024	0	0.019	0.0048	0.0048	0.043	0.082	0.0048	0.82

(a) IMU frame size=1s

True label \ Predicted label	clap	drink	eat	phone	sanitize	swh	wglass	write	wwh
clap	0.73	0.09	0.015	0.06	0.015	0	0.03	0.045	0.015
drink	0.15	0.69	0.09	0.015	0	0.015	0	0.03	0.015
eat	0	0.045	0.78	0.015	0	0	0	0	0.16
phone	0	0	0.015	0.87	0	0.015	0	0.1	0
sanitize	0.015	0	0.015	0	0.96	0	0.015	0	0
swh	0	0.06	0.12	0	0.015	0.73	0.015	0.045	0.015
wglass	0	0.029	0	0	0.071	0.057	0.74	0.014	0.086
write	0.045	0.03	0	0.24	0	0.06	0.015	0.6	0.015
wwh	0	0.029	0.014	0	0.029	0.029	0.12	0.014	0.77

(b) IMU frame size=3s

True label \ Predicted label	clap	drink	eat	phone	sanitize	swh	wglass	write	wwh
clap	0.72	0.025	0.025	0.025	0	0.025	0.1	0.025	0.05
drink	0.24	0.51	0.049	0.049	0.024	0.024	0.098	0	0
eat	0.025	0.05	0.65	0.1	0.025	0	0	0	0.15
phone	0	0.025	0.05	0.6	0	0.05	0	0.28	0
sanitize	0.024	0.024	0	0	0.9	0.024	0.024	0	0
swh	0.025	0.025	0	0.1	0	0.5	0.12	0.05	0.17
wglass	0	0.024	0	0.024	0	0.095	0.74	0	0.12
write	0.025	0.05	0	0.1	0	0.075	0.05	0.62	0.075
wwh	0	0	0	0	0	0.049	0.22	0.024	0.71

(c) IMU frame size=5s

True label \ Predicted label	clap	drink	eat	phone	sanitize	swh	wglass	write	wwh
clap	0.72	0.069	0	0	0.034	0	0	0.1	0.069
drink	0	0.69	0.1	0.069	0	0.069	0	0.034	0.034
eat	0.034	0.034	0.55	0	0	0.28	0	0.1	0
phone	0	0	0	0.86	0	0.1	0	0	0.034
sanitize	0	0	0.069	0	0.66	0.034	0	0.069	0.17
swh	0	0.1	0	0	0	0.66	0.17	0.069	0
wglass	0	0	0	0	0.067	0	0.8	0.067	0.067
write	0.036	0	0	0.14	0	0.29	0	0.54	0
wwh	0	0	0.034	0	0	0.17	0.14	0	0.66

(d) IMU frame size=7s

Fig. 7 Confusion matrices about different frame size

a smartwatch collecting data via the microphone and IMU modules are half of the market for data collection using only the IMU module. Therefore, reducing the frequency of data collection and identifying the environment for reasonable data collection is what can be done to save the power consumption of smartwatches.

9. Conclusion

In conclusion, this paper explores monitoring hygiene behaviors, including hand washing and disinfection based on acoustic and inertial sensing from smartwatches. We collected a dataset of 8 individuals with nine behaviors during 1 minute, including IMU and acoustic data. This paper adapted the CNN model to classify different behaviors by IMU and acoustic data separately. Combining acoustic data and inertial data, we also explore the concatenate fusion and score fusion methods to account for a subtle improvement in model results compared to deep learning frameworks based on a single dataset. The results also show that our method has improved prediction accuracy significantly compared with traditional machine learning methods, like SVM, KNN, and XGBoost.

Although we have made some progress in the classification results with the deep learning framework by combining acoustic and inertial data, the final accuracy is still not very high considering that the application scenario will have more noise. Thus, we will carry forward this work to construct a framework to achieve higher accuracy for classifying hygiene behaviors by other methods to combine IMU and acoustic data. Other superior deep learning will also be experimented with, like LSTM and RNN. Furthermore, we will design an algorithm to detect these behaviors timely to make it could be applied in a realistic production environment in the future.

Acknowledgment

This work was partly supported by NICT, Japan (222C01).

References

[1] Rebecca Adaimi, Howard Yong, and Edison Thomaz. “Ok Google, What Am I Doing? Acoustic Activity Recognition Bounded by Conversational Assistant Interactions”. In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5.1 (2021), pp. 1–24.

[2] Chen Bai et al. “Using smartwatches to detect face touching”. In: *Sensors* 21.19 (2021), p. 6528.

[3] Vincent Becker, Linus Fessler, and Gábor Sörös. “GestEar: combining audio and motion sensing for gesture recognition on smartwatches”. In: *Proceedings of the 23rd International Symposium on Wearable Computers*. 2019, pp. 10–19.

[4] Sarnab Bhattacharya, Rebecca Adaimi, and Edison Thomaz. "Leveraging Sound and Wrist Motion to Detect Activities of Daily Living with Commodity Smartwatches". In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6.2 (2022), pp. 1–28.

[5] Kaixuan Chen et al. "Deep learning for sensor-based human activity recognition: Overview, challenges, and opportunities". In: *ACM Computing Surveys (CSUR)* 54.4 (2021), pp. 1–40.

[6] Xiang'Anthony' Chen. "FaceOff: Detecting face touching with a wrist-worn accelerometer". In: *arXiv preprint arXiv:2008.01769* (2020).

[7] Brian Clarkson and Alex Pentland. "Extracting context from environmental audio". In: *Digest of Papers. Second International Symposium on Wearable Computers (Cat. No. 98EX215)*. IEEE. 1998, pp. 154–155.

[8] Yu Guan and Thomas Plötz. "Ensembles of deep lstm learners for activity recognition using wearables". In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1.2 (2017), pp. 1–28.

[9] Emanuele Lattanzi, Lorenzo Calisti, and Valerio Freschi. "Automatic unstructured handwashing recognition using smartwatch to reduce contact transmission of pathogens". In: *arXiv preprint arXiv:2107.13405* (2021).

[10] Huina Meng et al. "Mask Wearing Status Estimation with Smartwatches". In: *arXiv preprint arXiv:2205.06113* (2022).

[11] Md Abu Sayeed Mondol and John A Stankovic. "HAWAD: Hand washing detection using wrist wearable inertial sensors". In: *2020 16th International Conference on Distributed Computing in Sensor Systems (DCOSS)*. IEEE. 2020, pp. 11–18.

[12] Didier Pittet et al. "The World Health Organization guidelines on hand hygiene in health care and their consensus recommendations". In: *Infection Control & Hospital Epidemiology* 30.7 (2009), pp. 611–622.

[13] Hamada Rizk et al. "Smartwatch-based face-touch prediction using deep representational learning". In: *International Conference on Mobile and Ubiquitous Systems: Computing, Networking, and Services*. Springer. 2021, pp. 493–499.

[14] Sirat Samyoun et al. "iWash: A smartwatch handwashing quality assessment and reminder system with real-time feedback in the context of infectious disease". In: *Smart Health* 19 (2021), p. 100171.

[15] Johannes A Stork et al. "Audio-based human activity recognition using non-markovian ensemble voting". In: *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*. IEEE. 2012, pp. 509–514.

[16] Xin Wang et al. "Social Distancing Alert with Smartwatches". In: *arXiv preprint arXiv:2205.06110* (2022).

[17] Gary M. Weiss et al. "Smartwatch-based activity recognition: A machine learning approach". In: *2016 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI)*. 2016, pp. 426–429. DOI: 10.1109/BHI.2016.7455925.

[18] Zhendong Zhuang and Yang Xue. "Sport-related human activity detection and recognition using a smartwatch". In: *Sensors* 19.22 (2019), p. 5001.