This article has been accepted for publication in IEEE Transactions on Mobile Computing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TMC.2022.3227299

WashRing: An Energy-efficient and Highly Accurate Handwashing Monitoring System via Smart Ring

Weitao Xu, Huanqi Yang, Jiongzhang Chen, Chengwen Luo, Jia Zhang, Yuliang Zhao, Wen Jung Li

Abstract—The outbreak of COVID-19 has greatly changed everyone's lifestyle all over the world. One of the best ways to prevent the spread of infections is by washing hands properly. Although a number of hand hygiene monitoring systems have been proposed, they either cannot achieve high accuracy in practice or work only in limited environments such as hospitals. Therefore, a ubiquitous, energy-efficient and highly accurate hand hygiene monitoring system is still lacking. In this paper, we present *WashRing*—the first smart ring-based handwashing monitoring system. In WashRing, we design a Partially Observable Markov Decision Process (POMDP) based adaptive sampling approach to achieve high energy efficiency. Then, we design an automatic feature extraction scheme based on wavelet scattering and a CNN-LSTM neural network to achieve fine-grained gesture recognition. Finally, we model the handwashing gesture classification as a few-shot learning problem to mitigate the burden of collecting extensive data from five fingers. We collect data from 25 subjects over 2 months and evaluate the system performance on both commercial OURA ring and customized ring. Evaluation results show that WashRing achieves 97.8% accuracy which is 10.2%–15.9% higher than state-of-the-arts. Our adaptive sampling approach reduces energy consumption by 64.2% compared to fixed duty cycle sampling strategies.

Index Terms—Hand washing, Wearable devices, Deep learning, Energy-efficiency

1 Introduction

1.1 Background and Motivation

THE coronavirus disease 2019 (COVID-19) has emerged as a pandemic, which spread in over 200 countries, infected more than 172,630,000 people and caused more than 3,718,000 deaths, according to the latest data from the World Health Organization (WHO) [1]. The pandemic changes people's lifestyles, places unprecedented demands on the world's health systems, devastates vulnerable populations, and poses an unprecedented threat to global societies [2]. While significant efforts are being made on the front lines to detect the virus, provide treatments, and research vaccinations, it is also critical to prevent infection in our daily life. Handwashing, also known as hand hygiene, is one of the best ways to prevent infection, as claimed by the WHO [3].

According to guidelines published by the WHO [4], proper handwashing consists of six stages, which are shown in Fig. 1. The six stages can be further divided into nine steps which ensure that every area of the hands is properly covered. The whole hand hygiene duration should last 20–30 seconds. Unfortunately, a report by the United States Department of Agriculture found that up to 97% of people wash their hands incorrectly [1]. Therefore, it is important to monitor and measure hand hygiene adherence.

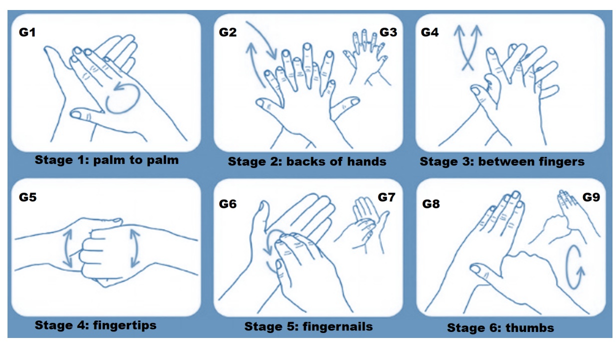


Fig. 1: WHO handwashing procedure.

Over the past decade, many automated hand hygiene monitoring systems based on various sensing technologies have been proposed, such as wall-mounted sensor-based [5], [6], camera-based [7], and radio-based [8]. However, there exist some drawbacks when they are used in real-world environments. For example, approaches based on placing or attaching devices around the sink are limited by their high installation and maintenance cost because they need to be installed at each sink. Additionally, camera-based techniques raise the privacy issue and their accuracy are highly environment-dependent. In comparison, wearable technology provides a continuous and non-obtrusive way to monitor user's handwashing activities. Several automatic handwashing monitoring systems based on wrist-worn devices have been developed recently [9],

- *Weitao Xu and Huanqi Yang are with the City University of Hong Kong Shenzhen Research Institute. They are also affiliated with the Department of Computer Science, City University of Hong Kong, Hong Kong. E-mail: weitaoxu@cityu.edu.hk, huanqi.yang@my.cityu.edu.hk*
- *Jiongzhang Chen and Chengwen Luo are with College of Computer Science and Software Engineering, Shenzhen University, China. Email: chenjiongzhang2020@email.szu.edu.cn, chengwen@szu.edu.cn.*
- *Jia Zhang is with the School of Information Science and Engineering, Shandong Normal University, China. Email: zhangjia@sdnu.edu.cn.*
- *Yuliang Zhao is with School of Control Engineering, Northeastern University at Qinhuangdao, China. Email:zhaoyuliang@neuq.edu.cn*
- *Wen Jung Li is with Department of Biomedical Engineering, City University of Hong Kong, Hong Kong. E-mail:wenjli@cityu.edu.hk*

1. https://www.usda.gov/media/press-releases/2018/06/28/study-shows-most-people-are-spreading-dangerous-bacteria-around [Online, accessed on Dec 4th, 2022].

TABLE 1: Comparison of different solutions.

	Ubiquity	Usability	Quality Assessment	Privacy
wall-mounted sensor [12], [13]	✗	✗	✗	✓
Camera-based [7]	✗	✗	✓	✗
Radio-based [8]	✗	✗	✓	✓
WashRing	✓	✓	✓	✓

[10], [11], but their practical accuracy is not satisfactory. This is because some handwashing steps are very similar to each other and they only have subtle difference in some finger movements, which cannot be fully captured by wrist-worn devices. Moreover, wearable sensors have limited battery life and continuous handwashing monitoring will drain battery quickly. Unfortunately, this problem is largely ignored in existing studies.

Since electronic components are constantly getting smaller and smaller, wearables are evolving from wrist-worn to be finger-worn. Recently, smart ring has started to enter people's life. Smart ring offers several benefits compared to wrist-worn devices such as smart watch and smart wristband. On the one hand, smart rings are more comfortable to wear both night and day because they are less obtrusive. On the other hand, smart rings can capture the minor finger motions and thus are better suited for finger-level gesture monitoring. Motivated by this, we design the first smart ring-based hand hygiene monitoring system—WashRing. As shown in Table 1, WashRing provides a ubiquitous, energy-efficient and highly accurate hand hygiene monitoring solution compared to existing systems.

1.2 Challenges and Contributions

To achieve the goals above, we need to address several non-trivial challenges.

Challenge 1: Resource constraint. Smart ring is a resource-constrained wearable device because of its small form factor. Although the duration of one-time handwashing is short, people need to wash their hands multiple times a day. Therefore, continuous sensor sampling will quickly drain the battery. To address this challenge, we model the hand washing process as a Partially Observable Markov Decision Process (POMDP). Then, we propose an adaptive sampling strategy to balance energy efficiency and recognition accuracy.

Challenge 2: Variability of handwashing action. Different people have different handwashing behaviors and even for the same person, non-trivial variations exist during different times of handwashing activities. Although deep learning has achieved great success in various recognition tasks [14], [15], [16], [17], it is often regarded as a blackbox and requires significant amount of data to train an accurate model. To address this problem, we design an automatic feature extraction scheme based on wavelet scattering to extract discriminative features.

Challenge 3: Impact of different fingers. Different people may wear the ring on different fingers, and the sensor data recorded from different fingers are slightly different even for the same gesture. Fig. 2 shows the accelerometer data when the user is performing the same handwashing action (G6 in Fig. 1). We can see that although the overall pattern is similar, there are some minor difference. However, collecting sufficient data to train a model working for all the five fingers is a non-trivial work. To address this challenge, we adopt the "learn to learn" concept from meta-learning and formulate the task as a few-shot learning problem to enable quick system adaptation in a data-efficient manner.

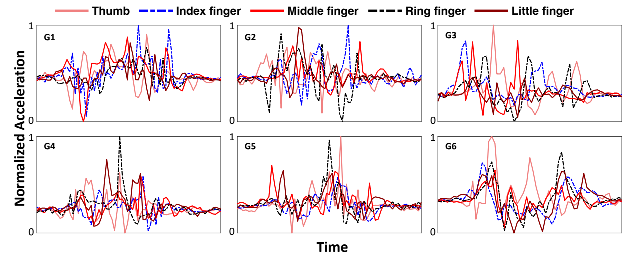


Fig. 2: Acceleration data of different fingers for six gestures.

We implement a prototype of WashRing on both off-the-shelf smart ring and customized ring, and conduct extensive experiments to evaluate the system performance. Results show that our system achieves high energy-efficiency and recognition accuracy. To our best knowledge, WashRing is the first smart ring-based hand washing monitoring system. In summary, the contributions of the paper are as follows:

- We propose WashRing, the first smart ring-based handwashing monitoring system. Compared to state-of-the-arts, WashRing achieves ubiquitous, energy-efficient and highly accurate handwashing monitoring. WashRing bring fine-grained handwashing monitoring from hospitals to daily life. We hope it can help people improve their handwashing quality and thus reduce the chance of infection.
- To achieve high energy efficiency, we model the handwashing procedure as a Partially Observable Markov Decision Process (POMDP) and propose an adaptive sampling strategy to dynamically adjust sampling actions to achieve an optimal trade-off between gesture recognition and device lifetime.
- To achieve fine-grained handwashing monitoring, we design an automatic feature extraction scheme based on wavelet scattering which can extract more discriminative features. Based on the extracted features, we design a CNN-LSTM based classification model to achieve highly accurate handwashing gesture recognition.
- To reduce the data collection effort for each finger, we model the handwashing gesture classification as a few-shot learning problem and design a few-shot learning module to rapidly adapt to new finger locations when a few training samples are available. Our strategy outperforms the popular fine-tuning method by up to 15.6%.
- We implement WashRing on both off-the-shelf smart ring and customized smart ring. We collect a dataset from 25 subjects over 2 months and conduct extensive evaluation to investigate the performance of WashRing. Evaluation results show that WashRing improves classification accuracy by 10.2%–15.9% compared to state-of-the-arts and achieves up to 64.2% energy saving with the proposed energy-efficient sampling strategy.

The rest of the paper is organized as follows. Sec. 2 provides an overview of WashRing. Sec. 3 discusses energy-efficient data sampling and processing. Then, Sec. 4 presents the design details of feature extraction, classification model

This article has been accepted for publication in IEEE Transactions on Mobile Computing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TMC.2022.3227299

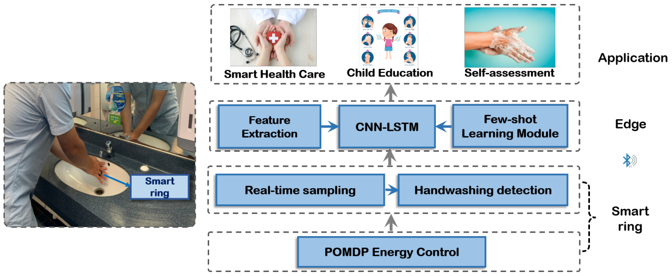


Fig. 3: An overview of WashRing.

and few-shot learning module. Sec. 5 presents the evaluation results. Sec. 7 discusses the related work. Finally, Sec. 8 concludes the paper.

2 System Overview

The architecture of WashRing is shown in Fig. 3. WashRing consists of a smart ring and an edge device (e.g., a smartphone). The smart ring is worn by the user on his/her preferred finger and is used to sample the IMU sensor during handwashing. A light-weight handwashing detector is integrated in smart ring to detect whether the user is washing hand or not. Meanwhile, a POMDP-based sampling strategy is running in background to save battery life and achieve energy efficiency. Then, the collected data is transmitted to the edge device via Bluetooth. The edge device first applies an automatic feature extraction scheme to extract discriminative features from the sensor data. Then, the extracted features are fed into a CNN-LSTM classification model to achieve fine-grained handwashing analysis. A few-shot learning module is pre-trained to make sure the classification model can achieve high performance when limited training data is available.

Users do not necessarily strictly follow the prescribed steps, namely, from G1 to G6. Instead, WashRing still works well when users perform the handwashing gestures in an arbitrary order. Meanwhile, users do not need to perform any special gestures to start. WashRing runs a handwashing activation detection module to detect the start of handwashing. The module continuously runs in the background to monitor the user's current activities (such as walking, running, typing, and eating). Once it detects the current action is handwashing, it will enable the handwashing classification module to monitor the whole handwashing process. WashRing will send the continuous sensor data to the user's smartphone for fine-grained analysis during handwashing. In the following sections, we will describe the design details of each component.

3 Data Sampling and Processing

3.1 Handwashing Activity Detection

Everyday people perform different kinds of activities, such as walking, cooking, and cleaning. Therefore, the first step is to detect handwashing activity to avoid unnecessary computation for non-handwashing activities. Since the number of non-handwashing activities is numerous, it is hard to collect a large amount of data for each activity. Therefore, we design a one-class activity classifier based on Support Vector Data Description (SVDD) [18]. SVDD is one of the most well-known support vector learning methods for the one-class problem. It only requires the data of the handwashing activity to build a prediction model whose output is positive (i.e., handwashing) or negative (i.e., non-handwashing). Moreover, based on the evaluation in Section 5.2, SVDD outperforms other one-class classifier such as one-class Support Vector Machine (SVM) and one Class Mini-max Probability Machine (OCMPM); therefore, we choose SVDD in our handwashing detection module.

Unlike smartwatch which is always worn on the wrist with fixed orientation, the smart ring can be worn with arbitrary orientations. Therefore, we first calculate the amplitude of raw accelerometer data which is orientation-independent: $Acc = \sqrt{Acc_x^2 + Acc_y^2 + Acc_z^2}$, where Acc_x^2, Acc_y^2 and Acc_z^2 represent the linear accelerometer data along three axes. Then a moving average filter of order 3 is applied on the acceleration amplitude to remove noise. Afterwards, continuous accelerometer data is segmented into $1\,\mathrm{s}$ sliding windows with 50% overlap. The window size of $1\,\mathrm{s}$ is chosen to balance between classification accuracy and latency. Wearable sensor-based activity classification is a well-studied field, and features from both time and frequency domain are widely used to achieve accurate activity recognition [19]. For computation efficiency, we calculate six light-weight time domain features: MAX, MEAN, MIN, VARIANCE, RANGE, CROSS ZERO RATE. These features are used to train the classifier. The evaluation results in Section 5.2 show that the designed SVDD classifier can achieve over 98% handwashing detection accuracy.

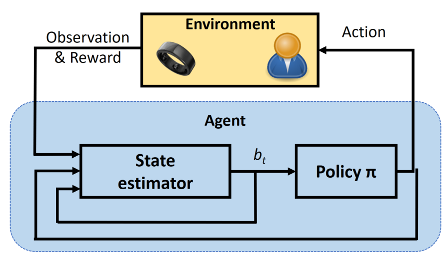


Fig. 4: POMDP Sampling strategy.

3.2 POMDP-based Dynamic Sampling Strategy

Continuous sampling sensory data will quickly drain the battery of smart ring because of its extremely small form factor. Therefore, it is crucial to improve the energy-efficiency of WashRing to maximize the device lifetime. In WashRing, we model the sampling strategy optimization as a POMDP problem. POMDP provides a formal probabilistic framework for solving tasks involving action selection and decision making under uncertainty [20]. POMDP is a common technique to model a variety of real-world sequential decision processes. Although some recent works use deep reinforcement learning to solve similar decision problems [21], [22], deep reinforcement learning is known to be computationally expensive and require extensive data for training.

The decision making process of the proposed method is illustrated in Fig. 4. We consider the smart ring and user as the environment, and the goal of the agent is to find an optimal sampling policy based on the observations from the environment. When the agent executes a sampling policy, the environment enters a new state. The agent receives an observation of the new state and a reward. In order to solve the optimization problem, the agent maintains a belief b_t which is a probability distribution over states of the environment. This belief is computed iteratively using Bayesian inference by the belief state estimator. An action for the current time step is provided by the learned policy π, which maps belief states to actions. Although the environment is changing over time and not fully predictable, the agent is able to find the optimal sampling policy based on new observations at each time step. Below we describe the details of the proposed sampling strategy.

The POMDP model used in our sampling strategy consists of six elements (S, A, O, F, Ω, R), where S, A, O, F, Ω, R are the state space, action space, observation space, state transition function, observation function, and the reward function, respectively. In WashRing, the state at time t is specified by $S_t = (e_t, \zeta_t, t)$, where e_t is the battery level of the smart ring at time t, i.e., $e_t = \{0, 1, 2, \ldots, 100\}$, and $\zeta_t = \{0, 1, 2, \ldots, 9, 10\}$ represents the percentage time that the user is washing hands during the current observation window whose default value is $10\,\mathrm{s}$. For example, 0 means the user is not washing hands and 2 means the time of handwashing occupies 20% of the current observation window. The actions are represented by $A_t = \{0, 1, 2, 3, 4, 5\}$, indicating the duty cycle of sensor sampling for energy control. For example, 0 represents no sampling, 1 represents 20% duty cycling, and 2 represents 40% duty cycling. The observation at time t is defined as $O_t = \{0, 1, 2, \ldots, 9, 10\}$, representing the observed percentage of time during the window that user is washing hand. The observation are obtained by the light-weight handwashing detector in Section 5.2.

After defining all these elements, the dynamic process of handwashing data sampling can be modeled as transitions over a finite set of states. Specifically, we define three functions: state transition function $p(S_{t+1}|S_t, A_t)$, observation function $\Omega = p(O_t|S_t, A_t)$ and the reward function $R(S_t, A_t)$. The state transition functions $p(e_{t+1}|e_t, A_t)$ and $p(\zeta_{t+1}|\zeta_t, t)$ represent the transition between different sampling strategies and the probability of handwashing activity during different periods of a day. These two functions can be learned from the past data. The observation function $p(O_t|S_t, A_t)$ can be calculated as follows:

$$p(O_t|S_t, A_t) = p(O_t|\zeta_t, A_t) = \frac{\binom{A_t}{O_t} \cdot \binom{5 - A_t}{\zeta_t - O_t}}{\binom{5}{\zeta_t}} \tag{1}$$

where the last expression $\frac{\binom{A_t}{O_t} \cdot \binom{5 - A_t}{\zeta_t - O_t}}{\binom{5}{\zeta_t}}$ means the probability of observing O_t given the sampling duty cycle A_t and the user's handwashing level ζ_t. The goal of the agent is to maximize the expected sum of future rewards: $\max E[\sum_0^\infty \gamma^t R(S_t, A_t)]$, where $\gamma \in [0, 1]$ is the discount factor which determines the weights of immediate reward and future reward. The reward function used in WashRing is defined as follows [23]:

$$R(S_t, A_t) = \begin{cases} r_a - r_p & e_t \neq 0 \\ 0 & e_t = 0 \end{cases} \tag{2}$$

where $r_a = A_t \cdot \eta_t$ represents the rewards of using high duty cycles when the user is washing hands, $r_p = \frac{c \cdot A_t^2}{e_t}$ represents the penalty of using high sampling duty cycles when the battery level is low, and c is a parameter to adjust the penalty weights. The reward function is chosen in this way so that when the battery level is sufficiently high, the system will be more aggressive to use higher duty cycles, while when the battery left is low, the sampling strategy becomes more conservative.

In WashRing, the agent can observe e_t and t but not ζ_t. Therefore, the model must choose actions based on the history of observations and actions. This information is succinctly captured by the "belief state", which is the posterior probability distribution over states at time t, given past observations and actions. The states in our system are discrete, so the belief state is a vector b_t whose size is equal to the size of states. The i-th component of b_t is the posterior probability of state i: $b_t(i) = P(s_t = i|o_t, a_t, o_{t-1}, a_{t-1}, \ldots, o_0, a_0)$. The goal of the agent then becomes maximizing the expected future reward by finding an optimal "policy" π which maps a belief state b_t to an appropriate action at $\pi(b_t) = a_t$.

We use the SARSOP algorithm [24] to solve this optimization problem because of its fast computation speed. At any time t, the sampling policy is determined by considering different factors such as current battery level, user's action, and past observations. Therefore, it can achieve a good balance between fine-grained handwashing monitoring and battery lifetime.

4 Handwashing Activity Classification

4.1 Automatic Feature Extraction Using Wavelet Scattering

Since the handwashing motions are different from person to person and even the behaviors of the same person may change over time, we need to extract robust and low-variance features from IMU data. In WashRing, we use wavelet analysis to obtain the time-frequency representation of handwashing motions. It is based on the observation that handwashing activity involves many minor and transient finger motions.

After receiving the data from the smart ring, the edge device first applies a Butterworth low-pass filter with cut-off frequency $20\,\mathrm{Hz}$ to remove noise. Then, suppose the amplitude of denoised accelerometer data is $Acc(t)$, we calculate the continuous wavelet transform (CWT) of $Acc(t)$ as follows:

$$W(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} Acc(t) \psi^* \left(\frac{t - b}{a} \right) \mathrm{d}t \tag{3}$$

where ψ is a mother wavelet function and the symbol $*$ represents the complex conjugate operation, a and b are the scale and translation parameters, respectively. In WashRing, we use the Morlet wavelet because it exhibits superior time localization. After CWT, we can get the CWT coefficients. To better visualize the time-frequency representation of each handwashing action, the scalogram which is the absolute value of the CWT coefficients, is plotted in the second row of Fig. 5. As we can see from Fig. 5, CWT can provide a time-frequency representation for each gesture. However, we notice that the wavelet transform is spread out over a region. This is due to the fact that any time-frequency transform that employs filters, such as wavelets, blurs the signal's

This article has been accepted for publication in IEEE Transactions on Mobile Computing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TMC.2022.3227299

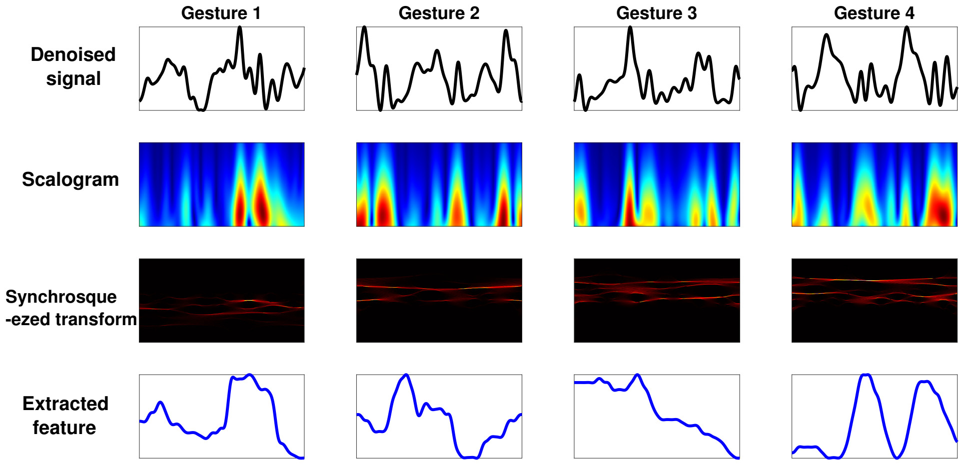


Fig. 5: Illustration of feature extraction. (1) The first row shows the denoised signal of four gestures. (2) The second row plots the scalograms. (3) The third row plots the wavelet synchrosqueezed transform. (4) The fourth row plots one feature vector for each gesture after wavelet scattering.

image in time and frequency. Therefore, we further sharpen the time-frequency representations using synchrosqueezing technique [25]. Synchrosqueezing is a time-frequency signal processing procedure that "squeezes" the transform along the frequency axis to compensate for the smearing. The third row of Fig. 5 plots the synchrosqueezed transform of CWT for each gesture.

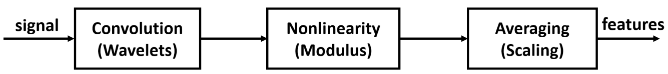


Fig. 6: Workflow of automatic feature extraction scheme.

After obtaining time-frequency representations of hand gestures, a common way is to directly feed these 2-dimensional images into a deep convolutional neural network (CNN). However, there are a few challenges to use CNN directly in the wearable device based handwashing recognition. First, CNN models typically require large datasets and significant computing resources for training and evaluation. Secondly, even a well-tuned network can achieve high accuracy, it is often treated as a blackbox. Thus, it can be difficult to understand and interpret the features that are extracted. To address these problems, it is desirable to first extract discriminative features and then feed these features into a CNN model. However, traditional feature engineering requires a large amount of empirical experience and work. To this end, we design an automatic feature extraction scheme based on a wavelet scattering framework.

A wavelet scattering transform builds translation invariant, stable, and informative representations from the input signal. As shown in Fig. 6, our wavelet scattering framework consists of three stages: wavelet transform, nonlinear modulus, and averaging. First, we perform a traditional convolution to generate a locally translation invariant feature of input signal; however, this step loses high-frequency information. In step two, the lost high-frequency information is recovered by a wavelet modulus transform. Finally, we average each of the moduli with the scaling filter to obtain the first-order scattering coefficients. By repeating this process, we can obtain a feature matrix which aggregates scattering coefficients of all orders to describe the features of input signal. For the technical details of wavelet scattering, the readers are encouraged to referred to [26], [27].

Our designed wavelet scattering framework includes two layers whose wavelets per octave are 8 and 1, respectively. We use Morlet wavelet and set the invariance scale to 20. For segmented handwashing gestures, we further divide them into $0.3\,\mathrm{s}$ time windows with 50% overlap. Then for each window, we calculate its wavelet synchrosqueezed transform and feed the transform into the wavelet scattering network, which will output a tensor feature matrix with the size of $55 \times 51 \times 45$. The forth row of Fig. 5 plots a single feature vector extracted from each gesture, it is evident that they are more discriminative compared to the original scalogram in the second row.

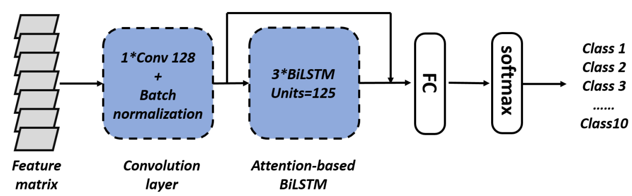


Fig. 7: Classification model.

4.2 Classification

The framework of the classification model is shown in Fig. 7. Concretely, the feature matrix extracted in the last section first passes through a convolution layer to extract the local features. The filter size and kernel size are 128 and 5×5, respectively. Then, the results are fed into three BiLSTM layers to mine the overall variation trends of the features. We employ BiLSTM because it has a high degree of adaptability to time series data, and the bidirectional learning process not only learns the characteristics of the forward time flow, but also learns the reverse time flow. However, the standard BiLSTM cannot detect which part is necessary to identify the features of the fine-grained handwashing gestures. In order to solve this problem, we add an attention scheme that

This article has been accepted for publication in IEEE Transactions on Mobile Computing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TMC.2022.3227299

can capture the key information of the input handwashing gesture characteristics. The attention scheme is used to simulate the way of thinking of the human brain, that is, giving high attention to important information and low attention to unimportant information. Finally, the output is mapped into ten gesture classes via a fully connected layer with ten units and a softmax function. The reason we have ten classes is that we notice people have their own handwashing habits which are unofficial. Therefore, in addition to the nine official gestures defined by the WHO, we add the tenth class "Other gestures".

4.3 Few-shot learning module

As mentioned earlier, the IMU signals collected from different fingers are slightly different and our evaluation results in Sec. 5.4 also demonstrate that the model trained from one finger has poor performance for other fingers. To reduce the data collection effort for each finger, we design a few-shot learning module to enable fast system adaptation for new finger locations.

Problem Formulation. Few-shot learning aims to build a robust deep learning model when the dataset contains limited information. It adopts the "learn to learn" concept and uses prior knowledge to rapidly generalize to new tasks containing only a few samples. In WashRing, we formulate the problem as follows. We have two datasets: a source dataset $\mathcal{D}_s$ and a target dataset $\mathcal{D}_t$. The target dataset can be further divided into two sets: a support set and a query set. Note that the support set and query set share the same label space but are disjoint with each other. In our problem, the source dataset $\mathcal{D}_s$ contains sufficient data from a few fingers (e.g., index finger), the support dataset contains limited data from other fingers (e.g, all the fingers except the index finger). Therefore, the source dataset and target dataset share the same label space. We adopt this setting because the meta learning process can effectively leverage the common knowledge in the same label space [28]. With the above definitions, our goal is to use source dataset to train a classifer f that can quickly adapt to the recognition task in the query set based on the limited samples in the support set. If the support set contains K labelled examples for each of C unique classes, it is called C-way K-shot problem.

Meta-training. During meta-training, we generate a set of tasks $\mathcal{T}$ from $\mathcal{D}_s$. Each individual subtask $\mathcal{T}_i \in \mathcal{T}$ is formulated by randomly picking C classes with K labelled samples to act as the support set $\mathcal{S}_{\mathcal{T}_i} = \{(X, Y)\}^m (m = K \times C)$ and the other samples serve as the query set $\mathcal{Q}_{\mathcal{T}_i} = \{(X, Y)\}$. This support/query set split in $\mathcal{D}_s$ is designed to simulate the support/query set that will be countered in $\mathcal{D}_t$. The classifier f is randomly initialized with parameters θ_0 and then trained by all the subtasks $\mathcal{T}_i \in \mathcal{T}$. The parameters are updated by the method proposed in MAML [29]. For each subtask $\mathcal{T}_i$, a new task-specific parameters $\theta'_{\mathcal{T}_i}$ is learned via gradient descent:

$$\theta'_{\mathcal{T}_i} = \theta_0 - \alpha \nabla_\theta \mathcal{L}_{\mathcal{T}_i}\left(f_{\theta_0}, S_{\mathcal{T}_i}\right)$$

where α is a hyperparameter and $\mathcal{L}_{\mathcal{T}_i}(f_{\theta_0}, S_{\mathcal{T}_i})$ is the $\mathcal{T}_i$-specific cross-entropy loss which is defined as

$$\mathcal{L}_{\mathcal{T}_i}\left(f_\theta, S_{\mathcal{T}_i}\right) = \sum_{(\mathrm{x}_j, \mathrm{y}_j) \in S_{\mathcal{T}_i}} \mathrm{y}_j \log f_\theta\left(\mathrm{x}_j\right) + \left(1 - \mathrm{y}_j\right) \log f_\theta\left(1 - \mathrm{x}_j\right)$$

After obtaining all the task-specific parameters $\theta'_{\mathcal{T}_i}$, we define a meta-objective function as follows:

$$\arg\min_\theta \sum_{\mathcal{T}_i \in \mathcal{T}} \mathcal{L}_{\mathcal{T}_i}\left(f_{\theta'_{\mathcal{T}_i}}\right)$$

This objective function is designed to find parameters θ that can minimize the sum of all the task losses. The parameters θ can be obtained by minimizing the objective function via stochastic gradient descent (SGD):

$$\theta \leftarrow \theta - \beta \nabla_\theta \sum_{\mathcal{T}_i \in \mathcal{T}} \mathcal{L}_{\mathcal{T}_i}\left(f_{\theta'_{\mathcal{T}_i}}\right)$$

where β denotes the meta learning rate.

Adaptation. After the above meta-learning stage, a classifier $f_{\theta'}$ with good parameter initialization θ' is trained from source dataset $\mathcal{D}_s$. When deploying $f_{\theta'}$ on target dataset $\mathcal{D}_t$, the new parameters θ_t can be learned after a few gradient descent steps as: $\theta_t \leftarrow \theta' - \alpha \nabla_\theta \mathcal{L}_{\mathcal{D}_t}\left(f_{\theta'}\right)$. Finally, the output of the few-shot learning module is a classifier f_{θ_t} that can achieve good performance on the target dataset $\mathcal{D}_t$ where each class only contains a few samples.

5 Evaluations

5.1 Goals, Metrics, and Methodologies

In this section, we conduct extensive evaluation to evaluate the performance of WashRing. The goals of evaluation are fourfold: 1) to evaluate the individual components of WashRing; 2) to evaluate the overall classification accuracy of WashRing in different conditions; 3) to compare WashRing with state-of-the-art handwashing monitoring systems; 4) to evaluate the resource consumption of WashRing.

Prototype. The WashRing system consists of two components: a smart ring and an edge device. As shown in Fig. 8 we implement WashRing on both commercial OURA smart ring and a custom-built smart ring. OURA smart ring is equipped with Bluetooth Low Energy (BLE), 3-axis accelerometer and gyroscope. OURA weighs in at only $6\,\mathrm{g}$ with a width of $8\,\mathrm{mm}$ and a thickness of $2.55\,\mathrm{mm}$. Moreover, OURA smart ring is water resistant, so users can wear it during handwashing. Since OURA smart ring does not provide any method to measure energy consumption, we build a customized smart ring. The customized ring is based on DA14583 chipset which includes a BLE chip, a 3-axis accelerometer and gyroscope chip, and a 3-axis magnetometer chip. The size of the ring is $19\,\mathrm{mm} \times 15\,\mathrm{mm} \times 25\,\mathrm{mm}$ and the weight is approximately $7\,\mathrm{g}$ in total, including a $1.7\,\mathrm{g}$ sensor board and a $5.3\,\mathrm{g}$ ring structure. Because the customized ring is not water-proof, the data collection is conducted by OURA ring while the energy profiling is performed on the customized ring. In our future work, we will design a water-proof customized ring to ensure all the experiments are conducted on the same hardware platform. For the edge device, we use Samsung S10 smartphone which is equipped with a 2.84 GHz Snapdragon CPU. The deep

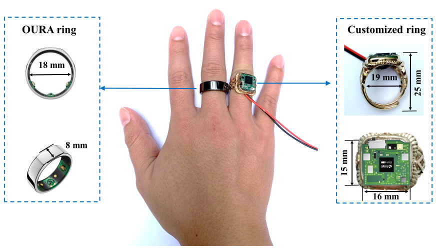


Fig. 8: Prototypes.

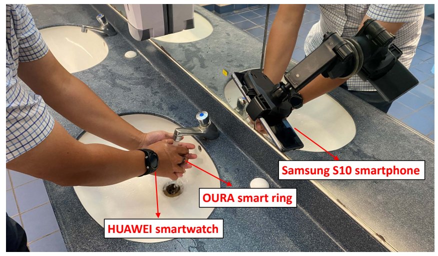


Fig. 9: Data Collection.

Property	Value
Subjects	25 (14 males, 11 females)
Age	18-46
Height	155-182 cm
Weight	58-93 kg
Sessions	2
Hand locations	6 left+9 right
Finger locations	5
Sample frequency	100 Hz

TABLE 2: Dataset information.

learning model is implemented based on Pytorch Mobile framework.

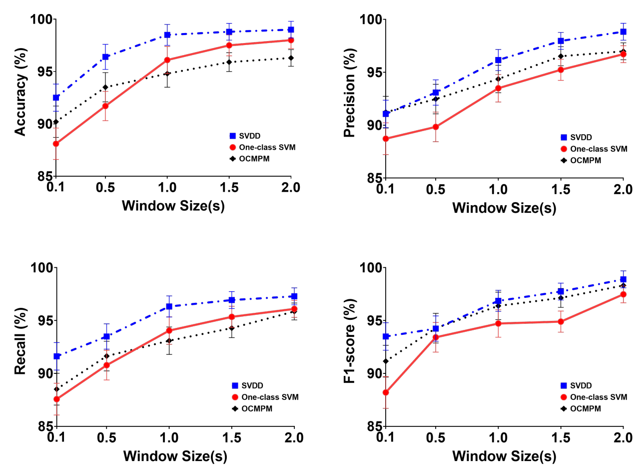
Data Collection. The process of data collection is illustrated in Fig. 9. During data collection[2], we first show the volunteers a short video of correct handwashing techniques provided by WHO [3]. Then, we ask them to wear the smart ring on one of their preferred hands. Six of them choose to wear the ring on their left hand while the others wear it on their right hand. Meanwhile, different people may wear the ring on different fingers, thus we ask the volunteers to repeat the whole handwashing process on each of the five fingers so that we have the data for all the locations. In order to compare the wrist-worn devices based solutions, we also ask the volunteers to wear a HUAWEI smartwatch on the same hand so that we can obtain the data from both fingers and wrist. Each participant was asked to perform the whole handwashing process multiple times so that we have enough data for evaluation. We also ask the volunteers to perform their own handwashing actions which are not the same as the nine official gestures. These data will be regarded as the tenth class "O" (others).

To obtain the ground-truth, we set up a smartphone which is used to record the handwashing process. The data is collected by smart ring and then transmitted to the smartphone. By synchronizing the timestamps of the sensor data and the recorded video, we can obtain the ground-truth of each gesture. Because the handwashing behavior of the same person may change slightly overtime, we record two independent sessions for each person. The time intervals between these two sessions vary from 1 to 2 months. The original sampling rate of the data is 100 Hz. In total, we collected approximately 67,500 hand gestures which is 37× larger than the private dataset used in RFWash [8]. More details of the dataset are summarized in Tab. 2.

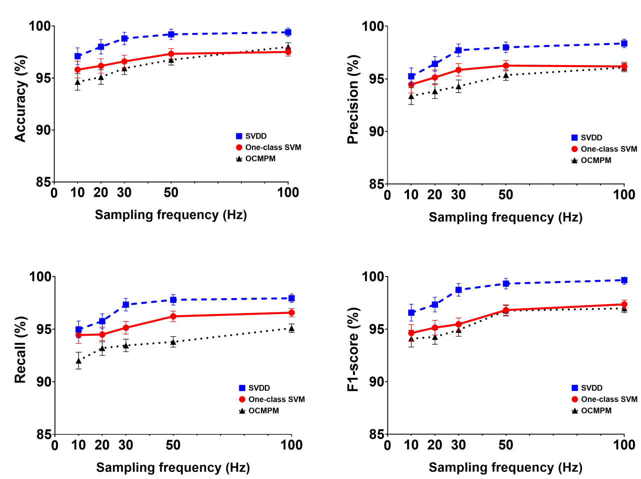
Metrics and Methodology. Unless otherwise stated, the default training and testing process is as follows. We randomly divide the whole dataset into three parts: training set (70%), validation set (15%), and test set (15%). We report the average result from 10 runs. We use the following four commonly used metrics: accuracy, precision, recall, and F1-score, which are defined as follows.

- **Accuracy.** Accuracy is the number of correct classifications divided by the number of total classifications. It is a useful metric only when the dataset has an equal distribution of classes.
- **Precision.** Precision is the number of true positives divided by the number of total positive predictions. It measures the model's accuracy in classifying a sample as positive.
- **Recall.** Recall is the number of positive samples correctly classified as positive divided by the total number of positive samples. It measures the model's ability to detect positive samples.
- **F1-score.** F1-score is a metric that combines the precision and recall metrics into a single metric, which is defined as $F1 = \frac{2 \times precision \times recall}{precision + recall}$. It is a useful metric when evaluating on imbalanced data.

The benchmarks used in the evaluation are three state-of-the-art handwashing monitoring systems: Harmony [9], Wristwash [10], and AWash [11].



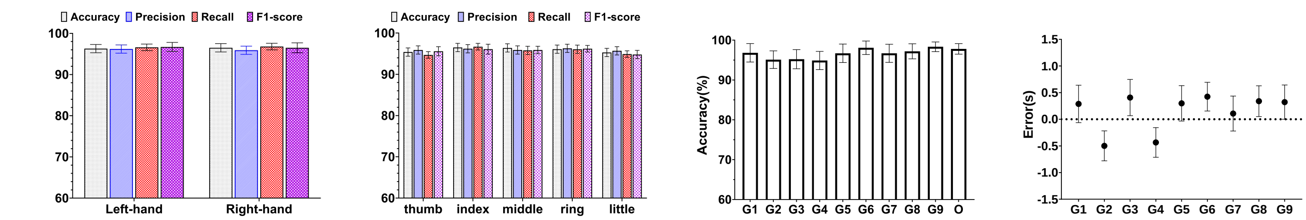
(a) Impact of window size



(b) Impact of sampling rate

Fig. 10: Results of handwashing detection.

2. Ethical clearance has been obtained to carry out this experiment.
3. https://www.youtube.com/watch?v=IisgnbMfKvI [Online, accessed on Dec 4th, 2022]



(a) Different hands (b) Different fingers (c) Different gestures (d) Timing error

Fig. 11: Accuracy of WashRing under different settings.

5.2 Handwashing Detection

To evaluate the performance of handwashing detection, we use OURA smart ring to collect over 2 hours data from non-handwashing activities, including walking, running, cleaning, toothbrushing, and cooking. We also compare SVDD with other commonly used one-class classifiers, including one-class SVM and one Class Mini-max Probability Machine (OCMPM). Fig. 10(a) and Fig. 10(b) show the performance of different methods under different window sizes and sampling frequencies, respectively. Fig. 10(a) shows we can achieve higher detection accuracy with longer window size because we can obtain more information. However, longer window size makes the detection module less responsive. The window size is set to $1\,\mathrm{s}$ in WashRing to balance the accuracy and execution time. Fig. 10(a) shows the accuracy of detection increases with higher sampling frequency, and it reaches 98.7% when over $50\,\mathrm{Hz}$ is used. However, higher sampling frequency consumes more energy. In practice, the sampling rate is dynamically adjusted by the sampling strategy which will be evaluated in Sec. 5.6. Both Fig. 10(a) and Fig. 10(b) show SVDD achieves higher accuracy compared to one-class SVM and OCMPM; therefore, we choose SVDD as the classifier in handwashing detection.

5.3 Overall Performance

In this subsection, we evaluate the performance of WashRing under different conditions. Specifically, we analyze the accuracy of handwashing gestures when the user wears the ring on different fingers and different days.

5.3.1 Performance of different locations

In this experiment, we perform an in-depth analysis of different locations of smart ring. First, we analyze the impact of hand, i.e., we aim to examine whether the accuracy is different if users wear the ring on different hands. To this end, we train the model with data from both hands but divide the test set into two parts: left-hand dataset and right-hand dataset. From the results in Fig. 11(a), we can see that the accuracy of different hands are very close to each other. Therefore, although different people have different preferences of wearing the ring on left hand or right hand, it has little impact on classification accuracy.

Next, we investigate the impact of different fingers. Same as above, we train the model with data from five fingers and divide the test set into five parts: thumb, index finger, middle finger, ring finger and little finger. Then, we calculate the accuracy of each test set. As shown in Fig. 11(b), WashRing can accurately recognize handwashing actions regardless of the locations of the smart ring.

We now examine the accuracy of recognizing different gestures. Fig. 11(c) shows the accuracy of all the gestures. We notice that the accuracy of G2, G3, and G4 is relatively low. After investigating the results, we find out that G4 (rub between fingers) is mostly misclassified as G2 (rub the back of left hand) or G3 (rub the back of right hand) because these steps are very similar[4]. Nevertheless, the overall average accuracy is up to 97.8%.

Based on the recognition results, we further evaluate the accuracy of estimating the duration of each gesture sequence. For example, if the user performs G1 for $6.3\,\mathrm{s}$ but the total time of G1 obtained by WashRing is $6.6\,\mathrm{s}$, then the error is $0.3\,\mathrm{s}$. As the results in Fig. 11(d), WashRing can accurately estimate the duration of each gesture. Specifically, the average timing error of WashRing is $0.35\,\mathrm{s}$. In comparison, the timing error of a state-of-the-art wireless sensing based hand washing system RFWash varies from 0.49s to 1.88s depending on the sequence length [8]. Therefore, WashRing outperforms RFWash in estimating the duration of gestures by 1.4–5.4$\times$.

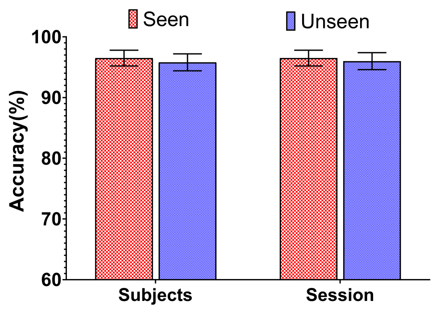
5.3.2 Unseen Scenarios

In this subsection, we evaluate the performance of WashRing in unseen scenarios, which include unseen subjects and unseen sessions. We leave the discussion of new finger locations in the next subsection.

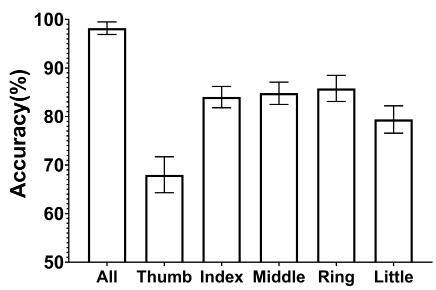
Unseen Subjects. In order to understand how the system performs for unseen subjects, we conducted a standard leave-one subject-out cross validation. Specifically, the data of one subject is excluded from the training process. Then we use this subject's data as testing set to evaluate the accuracy of the trained model. We repeat this process for all the 25 subjects and plot the average results in Fig. 12(a). We can see that the accuracy for unseen subjects is comparable to that of seen scenarios (the same subject's data are used in both training and testing), indicating WashRing has good generalization ability for new users.

Unseen Sessions. People's handwashing behaviors may change slightly over time. To understand how the system performs for data collected at different time, we use the Session 1 dataset for training and use the Session 2 dataset for testing (note Session 2 is collected 1–2 months after Session 1). From the results in Fig. 12(a), we can see there is no significant difference between seen sessions and unseen

4. If the ring is on right hand, then G4 is similar to G2. Otherwise, G4 is similar to G3.



(a) Unseen subject and time



(b) Unseen Fingers

Fig. 12: Unseen scenarios.

sessions, indicating WashRing is robust to the variations of handwashing behaviors over time.

5.4 Evaluation of few-shot learning module

In order to understand whether the model trained on one finger can be used to recognize gestures on another finger, we conduct a leave-one finger-out cross validation and compare it with the case where all the fingers' data are used in training. Fig. 12(b) shows that WashRing can achieve high accuracy when data from all the five fingers are used; however, the accuracy degrades significantly when the data of one specific finger is unavailable, especially the thumb. This can be explained by the fact that the location of thumb is far way from the other fingers and thus the motions are more different.

We now evaluate if the designed few-shot learning module can improve accuracy when we only have a few samples for some fingers. As discussed in Sec. 4.3, we formulate the recognition task as C-way K-shot classification problems, where C is the number of classes (i.e., 10 in WashRing) and K is the number of shots (i.e., number of samples available for each gesture). To investigate how many fingers are required to achieve satisfactory accuracy, we conduct a leave-M finger-out experiment, where M can be 1, 2, 3, 4. For each M, we assume there are C' instances available which make $C = C' \times M$. To demonstrate the superiority of our method, we also compare it with a common transfer learning strategy, namely fine-tuning. In this experiment, we set the hyperparameters α and β to be 0.01 and 0.001, and set C' to 1, 5, and 10, respectively. We calculate the results of different methods after 10 iterations.

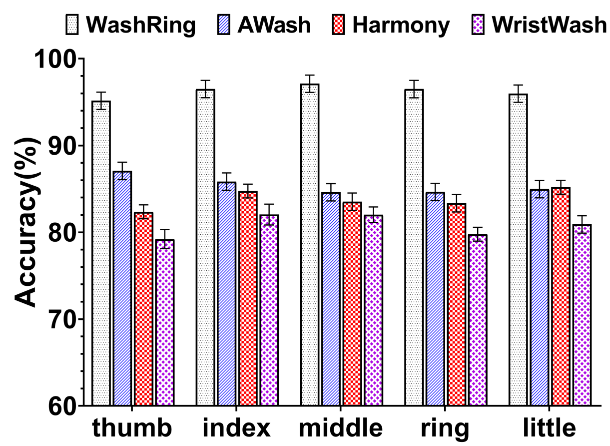
From the results in Tab. 3, we can see that our method outperforms the fine-tuning method in all scenarios. If only

TABLE 3: Performance of few-shot learning module.

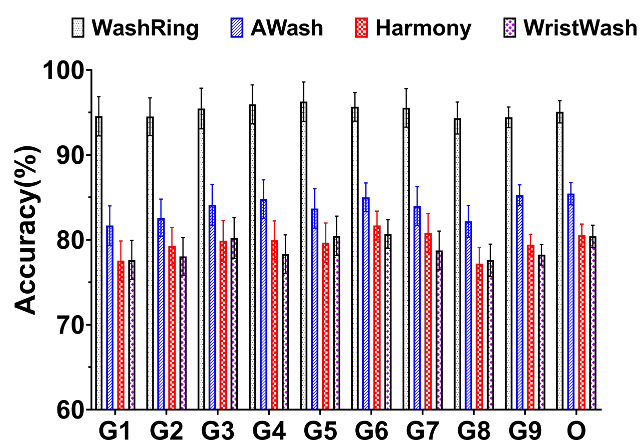
	C′=1		C′=5		C′=10	
	WashRing	Fine-tuning	WashRing	Fine-tuning	WashRing	Fine-tuning
M=1	86.3	78.2	89.4	81.3	92.3	86.6
M=2	72.1	65.7	83.2	77.6	90.7	81.5
M=3	65.7	58.4	74.6	65.2	88.4	76.3
M=4	55.6	38.1	68.7	60.8	86.5	70.9

one finger's data is available (i.e., $M = 4$), by collecting 10 samples from each of the other four unseen fingers, we can achieve 86.5% accuracy which is 15.6% higher than fine-tuning strategy. It is worth mentioning that collecting 10 samples only requires 3 s because the window size of each sample is 0.3 s. If we have data from four fingers (i.e., M=1), the accuracy can be improved to 92.3% by collecting 3 second's data from the unseen finger. Therefore, by collecting data from a few fingers only, our few-shot learning module can enable quick system adaptation and achieve reasonably good performance by collecting small-scale samples from unseen fingers.

5.5 Comparison with state-of-the-arts



(a) Accuracy of different fingers



(b) Accuracy of different gestures

Fig. 13: Comparison with state-of-the-arts.

In this subsection, we compare WashRing with four state-of-the-art wearable-based handwashing systems, namely, Harmony [9], Wristwash [10], and AWash [11]. The

This article has been accepted for publication in IEEE Transactions on Mobile Computing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TMC.2022.3227299

TABLE 4: System overhead (per gesture frame).

		Computation time (ms)	Energy consumption (mJ)
Smart ring	Sampling	-	0.45
	Handwashing detection	11.3	1.2
	POMDP sampling control	35	22
	Data transmission	2.4	0.013
	Total	48.7	≈ 23.6
Edge device	Feature extraction	31	32
	Classification	146	126.3
	Data transmission	2.1	0.024
	Total	179.1	≈ 158

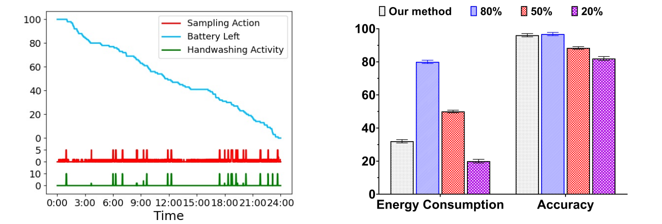
performance of each system is fine-tuned to achieve the best performance on our dataset. Specifically, for Harmony, the decision tree classier is used and the window size is set to $0.2\,\mathrm{s}$. For Wristwash, a Hidden Markov Model (HMM) classifier with 10 states is used and the window size is set to $0.12\,\mathrm{s}$. For AWash, the hybrid model is set to have 128 cells and three LSTM layers.

Fig. 13 shows the accuracy of different methods on different fingers and gestures. We can see that WashRing continuously achieves the best performance among all the methods. On average, the accuracy of WashRing is 10.2% higher than AWash, 13.8% higher than Harmony, and 15.9% higher than WristWash. For different fingers, our method improves accuracy by 9.1–20.2%. For different gestures, our method improves accuracy by 8.9–18.2%. The accuracy of Harmony and Wristwash is significantly lower because they are based on smartwatch and traditional machine learning classifiers. AWash achieves better accuracy than Harmony and Wristwash because of the use of deep learning model to extract features, but its accuracy is still lower than WashRing. These smart-watch based methods cannot achieve fine-grained gesture sensing because the wrist movements of some handwashing gestures are very similar, such as G1 and G8, G2 and G4. In contrast, WashRing can distinguish these minor difference by monitoring the fine-grained movement of fingers. Moreover, WashRing can be used by most people while AWash is specially designed for the elderly people with dementia. Therefore, WashRing provides a highly accurate and ubiquitous handwashing monitoring solution.

5.6 Evaluation of Sampling Strategy

In this subsection, we analyze the energy consumption of the propose system. WashRing includes two parts: a smart ring and an edge device. We implement the data sampling, handwashing detection, and POMDP sampling strategy on the customized ring. For the edge device, we use Samsung S10 smartphone which is equipped with a 2.84 GHz Snapdragon CPU and the Android 9.0 OS. The deep learning model is implemented based on Pytorch Mobile framework. As the results in Tab. 4, the total processing time on the ring and edge device is $49.7\,\mathrm{ms}$ and $87.1\,\mathrm{ms}$, respectively. As the window size of each gesture frame is $0.3\,\mathrm{s}$, our system can achieve real-time handwashing monitoring.

Tab. 4 shows the processing time and energy consummation of the smart ring and edge device. We can see the total processing time on the ring and edge device is $48.7\,\mathrm{ms}$ and $179.1\,\mathrm{ms}$, respectively. As the window size of each gesture frame is $0.3\,\mathrm{s}$, our system can achieve real-time handwashing monitoring. Next, we analyze the energy saving ability of the proposed sampling strategy. We use the POMDP simulation tool [24] to simulate the handwashing activities during a 24 hours period and compare the proposed adaptive sampling strategy with fixed sampling strategy. As shown in Fig. 14(a), our system keeps running at low duty cycles when the probability of handwashing is low. Once a handwashing activity is detected, our system starts to runs at higher duty cycle and turns into low duty cycle mode when the handwashing activity is over. The results of different sampling strategies are shown in Fig. 14(b). We can see that although using 20% duty cycle can achieve low energy consumption, its gesture recognition accuracy decreases to 83%. Our method can achieve comparable recognition accuracy with 80% duty cycle but reduces energy consumption by 62.4%. Overall, our sampling strategy can efficiently balance the energy consumption and handwashing activity monitoring. The amount of energy savings depends on the frequency of handwashing activities. It is recommended by WHO that people wash their hands 6–10 times per day. In this case, our sampling strategy can save energy by 34-56% compared to 80% duty cycling and 17%-28% compared to 50% duty cycling.



(a) Sampling strategy execution (b) Comparison with other strategies

Fig. 14: Evaluation of sampling strategy.

We now analyze the impact of energy consumption on smart ring. Suppose our customized smart ring uses the same battery as the OURA smart ring, whose battery capacity is $22\,\mathrm{mAh}$ ($285.12\,\mathrm{J}$). Then, the energy cost of WashRing amounts to $0.83e^{-4}$ of the total energy supply. With only 5% of the battery budget ($14.256\,\mathrm{J}$), WashRing is able to classify approximately 600 handwashing gestures per day. These results demonstrate that WashRing incurs a low system overhead.

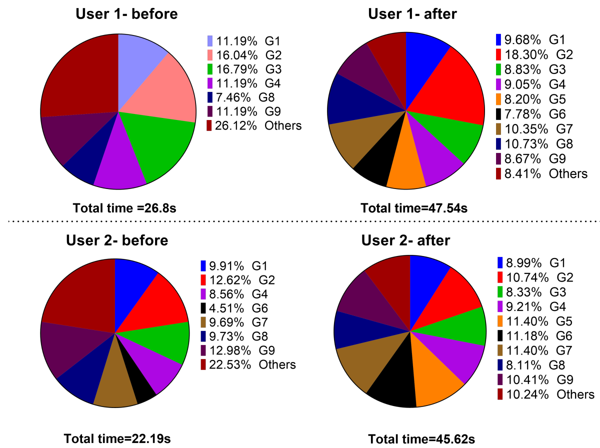
5.7 User Study

To further evaluate the effectiveness of WashRing, we recruited two participants (one male and one female) and conducted a user study in real-world environments. We asked them to wear a smart ring on their preferred fingers and install our app on their smartphones. After pairing the smart ring with their smartphones, we asked them to use our app for one week. Every time after they wash their hands, our app will provide them a summary report, notifying them which steps are missing, and the time spent on each step. The two participants did not attend the previous data collection, so we can treat them as new users.

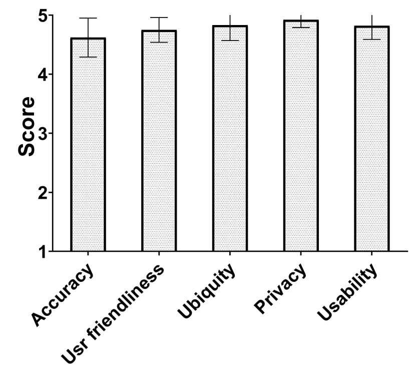
We compare their handwashing behaviors before and after using our system. Fig. 15(a) shows the changes of users' handwashing statistics. It is evident that both users established good handwashing behaviors after using our

APP for one week. For example, the handwashing process of user 1 does not include G4, G5 and G6 before testing, but after one week, the user starts to conduct G4, G5 and G6. Meanwhile, incorrect handwashing actions (i.e., others) also decrease significantly. Moreover, the total time of handwashing increases from $26.8\,\mathrm{s}$ to $47.54\,\mathrm{s}$. We can see the similar changes on user 2 as well. To be specific, the proportion of incorrect handwashing actions decreases from 22.53% to 10.24%. The total time of handwashing increases from $22.19\,\mathrm{s}$ to $45.62\,\mathrm{s}$. It is evident that WashRing helps users establish good handwashing behaviors in practice.

To further evaluate the acceptance of WashRing in real-world applications, we conducted a user survey. We invited 42 volunteers to participate in the survey, of which 24 are male and 18 are female, aged between 20 and 53 years old. The volunteers were first asked to watch a demo of WashRing and the WHO handwashing guidelines. Then, they wore an OURA smart ring on their preferred fingers and washed their hands for five times at different times of a day. Finally, our smartphone app will display the results to them after which they complete the survey based on their experience. The survey consists of five questions which ask the users to rate WashRing in terms of accuracy, user friendliness, ubiquity, privacy, and usability. As shown in Fig. 15(b), the scores of each question are 4.62 ± 0.33, 4.75 ± 0.21, 4.83 ± 0.26, 4.92 ± 0.13, and 4.82 ± 0.23, respectively. The score of accuracy is relatively lower than others probably because the volunteers do not adapt to the WHO handwashing guidelines in the short survey time. We believe, however, after using WashRing for a longer time, their hand hygiene can be improved significantly as demonstrated in Fig. 15(a). Overall, the results demonstrate the high effectiveness and acceptance of WashRing in real-world applications.



(a) Handwashing activities before/after using WashRing



(b) Results of user survey: 1–very low, 2–low, 3–medium, 4–high, 5–very high.

Fig. 15: User study.

6 Discussion

In this section, we discuss two issues related to the practical use of the proposed system.

Ring on or off? There is a debate in the medical community regarding whether one should take off the ring when they are washing hands. Several studies [30] have shown that skin underneath rings is more heavily colonized than comparable areas of skin on fingers without rings. However, several other studies also reported that there is no direct relationship between wearing rings and wearing no rings. For example, among 60 volunteers from perioperative personnel and medical students, a group of researchers [31] found no significant difference in bacterial counts on hands with or without rings when an alcohol product was used. Two other studies [32], [33] also found that mean bacterial colony counts on hands after handwashing were similar among individuals wearing rings and those not wearing rings. Therefore, whether the wearing of rings results in greater cross-transmission of pathogens remains unknown [34]. Despite different opinions, it is still strongly encouraged to take off rings when people wash their hands. However, it is hard to practically remove our rings every time we wash our hands, especially in public places such as workplaces and shopping centers. To sum up, although it is encouraged to take off rings during handwashing, considering the fact that many people still wear rings in many public places, our system has great potential to be used in people's daily life to improve their hand hygiene practices.

Contact or contactless? As discussed in Section 1, compared to contactless solutions (e.g., camera-based and radio-based), the proposed system is more accurate, cost-effective, and has no privacy issues. Despite these advantages, one concern is the requirement to wear rings on the user's fingers because the smart ring is not widely used nowadays. However, we believe that as its counterparts, such as smartphones and smartwatches, smart rings will become increasingly popular, and many people may wear a smart ring in their daily life soon. Therefore, compared to contactless solutions, our solution is more accurate, inexpensive, and ubiquitous, providing an alternative method to improve personal hygiene.

7 Related Work

This section discusses two aspects of related work: handwashing monitoring techniques and energy-efficient wearable sensing.

7.1 Handwashing monitoring

Hand hygiene is a critical activity in preventing the spread of virus infections and hence attracts considerable attention in the past decade. Existing automated hand hygiene monitoring systems can be classified into the following four catogories: wall-mounted sensor-based, camera-based, radio-based, and wearable-based.

Wall-mounted sensor-based. Many hand washing monitoring systems based on various wall-mounted sensors have been proposed. Marra et al. [12] installed an electronic handwash counter on dispensers of alcohol-based hand rub to monitor hand hygiene compliance. Kinsella et al. [13] developed a recording system that is equipped with pressure sensor to record depressions of soap and alcohol gel dispensers. Rabeek et al. [35] proposed an automated hand hygiene documentation and reminder system based a variety of sensors including vibration sensor, pressure sensor, and infrared sensor. However, these solutions can count the number of handwashing activities or the usage of washing dispenser only.

Camera-based. Since the wall-mounted sensors cannot provide fine-grained information about handwashing, researchers started to use camera to monitor hand hygiene compliance. Llorca et al. [7] proposed a camera-based hand hygiene monitoring system. Each sink is equipped with a camera which is used to record the handwashing activitity. Then, both appearance and motion features are extracted from the images and fed into a SVM classifier to obtain the gesture label. However, vision-based solutions have a number of limitations, such as privacy violation, environment-dependent accuracy, and high deployment cost. Although the privacy issue can be mitigated by using depth camera [36], [37], the captured images can still reveal a person's precise visual appearance, which may be used to breach his/her privacy.

Radio-based. Pineles et al. [38] proposed a handwashing monitoring system based on RFID technology. Khamis et al. [8] proposed a weakly supervised hand hygiene assessment system based on mmWave radar. These solutions provide contact-free handwashing monitoring and does not reveal user's privacy, but they still need to be installed in each sink which incur high deployment and maintenance expenses. Additionally, they cannot achieve ubiquitous handwashing monitoring in our daily life.

Wearable-based. Wearable devices provide a convenient way to monitor people's activities in a ubiquitous environment. Our work is closely related to Harmony [9], WristWash [10], and AWash [11], all of which are based on wrist-worn devices. Specifically, Harmony [9] uses data from a smartwatch's accelerometer and gyroscope as input and employs the decision tree classifier to recognize different handwashing actions. WristWash [10] uses 6-axis Inertial Measurement Unit (IMU) data as input and employs the hidden Markov model-based method to monitor handwashing process. AWash [11] is a smartwatch-based handwashing monitoring system specially designed for the Elderly with Dementia. AWash uses a hybrid deep learning model to extract user-independent features and designs a state machine to provide customize assistance for those with diverse cognitive impairments.

However, wrist-worn devices cannot capture the minor difference between different fingers when the user is washing hand. In comparison, WashRing can achieve higher classification accuracy based on finger-worn wearable device—smart ring. The evaluation results show that WashRing improves classification accuracy significantly compared to these methods. Moreover, these works do not consider the energy constraint issue of wearable devices. WashRing addresses this issue by proposing a POMDP-based adaptive sampling strategy.

7.2 Energy-efficient Wearable Sensing

Continuous sensor sampling is a power-intensive process for resource-limited wearable devices. As a result, lowering sensing power consumption without reducing sensing accuracy is a key factor to consider. In this subsection, we discuss the related studies that aims to reduce energy consumption in wearable sensing applications.

Some energy-efficient sampling approaches are based on dynamic sensor selection. For instance, Wang et al. [39] proposed a hierarchical sensor selection strategy that can dynamically determine a minimum set of sensors to reduce power usage. Their method not only recognizes user activities but also detects state transitions. Another line of research focuses on reducing sensing energy consumption by adaptive sampling. Krause et al. [40] studied the trade-off between decreasing the sampling frequency of accelerometer and activity classification accuracy. Their findings show that by optimizing duty cycles, the lifetime of a wearable device can be increased by four times without sacrificing prediction accuracy.

Through the analysis above, we found that although a variety of energy-efficient sampling methods have been proposed for human activity recognition, there is no related study on smart ring-based handwashing monitoring. Compared to conventional activities such as walking and running, different handwashing actions are hard to distinguish, which make energy-efficient handwashing monitoring challenging. To solve this problem, we model the sampling strategy optimization as a POMDP problem. Evaluation demonstrates the proposed method can save energy significantly while maintaining high recognition accuracy.

8 Conclusions

In this paper, we propose an accurate and energy-efficient handwashing monitoring system based on smart ring—WashRing. WashRing uses an adaptive sampling strategy to achieve energy-efficiency. Then, by using an automatic feature extraction scheme and a CNN-LSTM classification model, WashRing achieves fine-grained handwashing gesture recognition. Moreover, WashRing can achieve high recognition accuracy for different fingers based on a few-shot learning model. Extensive evaluation results show that WashRing outperforms state-of-the-art handwashing monitoring systems and can save energy consumption significantly. Overall, WashRing provides an ubiquitous, highly accurate and energy-efficient handwashing monitoring solution. It has great potential to be used in our daily life to improve our handwashing quality and prevent the spread of virus infections.

Acknowledgments

The work described in this paper was substantially sponsored by the project 62101471 supported by NSFC and was partially supported by the Shenzhen Research Institute, City University of Hong Kong. The work described in this paper was partially supported by the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. CityU 21201420 and CityU 11201422). The work described in this paper was partially supported by Shenzhen Science and Technology Funding Fundamental Research Program (Project No. 2021Szvup126), NSF of Shandong Province (Project No. ZR2021LZH010), and a grant from Chow Sang Sang Group Research Fund sponsored by Chow Sang Sang Holdings International Limited (Project No. 9229062). The work was also partially supported by CityU MFPRC grant 9680333, CityU SIRG grant 7020057, CityU APRC grant 9610485, CityU ARG grant 9667225 and CityU SRG-Fd grant 7005666.

References

[1] Coronavirus (covid-19). [Online]. Available: https://covid19.who.int/

[2] X. Ding, D. Clifton, N. Ji, N. H. Lovell, P. Bonato, W. Chen, X. Yu, Z. Xue, T. Xiang, X. Long, K. Xu, X. Jiang, Q. Wang, B. Yin, G. Feng, and Y.-T. Zhang, "Wearable sensing and telehealth technology with potential applications in the coronavirus pandemic," *IEEE Reviews in Biomedical Engineering*, vol. 14, pp. 48–70, 2021.

[3] Coronavirus disease (covid-19) advice for the public. [Online]. Available: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public

[4] Who: How to handwash? with soap and water. [Online]. Available: https://www.youtube.com/watch?v=3PmVJQUCm4E

[5] G. Kinsella, A. Thomas, and R. Taylor, "Electronic surveillance of wall-mounted soap and alcohol gel dispensers in an intensive care unit," *Journal of Hospital Infection*, vol. 66, no. 1, pp. 34–39, 2007.

[6] S. Mohamed Rabeek, A. Norman, M. Je, M. Kumarasamy Raja, R. F. Peh, and M. K. Dempsey, "A reliable handwash detector for automated hand hygiene documentation and reminder system in hospitals," in *Proceedings of the 15th International Conference on Biomedical Engineering*. Springer, 2014, pp. 892–895.

[7] D. F. Llorca, I. Parra, M. Á. Sotelo, and G. Lacey, "A vision-based system for automatic hand washing quality assessment," *Machine Vision and Applications*, vol. 22, no. 2, pp. 219–234, 2011.

[8] A. Khamis, B. Kusy, C. T. Chou, M.-L. McLaws, and W. Hu, "Rfwash: a weakly supervised tracking of hand hygiene technique," in *Proceedings of the 18th Conference on Embedded Networked Sensor Systems (Sensys)*, 2020, pp. 572–584.

[9] M. A. S. Mondol and J. A. Stankovic, "Harmony: A hand wash monitoring and reminder system using smart watches," in *Proceedings of the 12th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (Mobiquitous)*, 2015, pp. 11–20.

[10] 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 (ISWC)*, 2018, pp. 132–139.

[11] Y. Cao, H. Chen, F. Li, S. Yang, and Y. Wang, "Awash: Handwashing assistance for the elderly with dementia via wearables," in *Proceedings of the IEEE International Conference on Computer Communication (INFOCOM)*, 2021, pp. 1–10.

[12] A. Marra and M. Edmond, "New technologies to monitor healthcare worker hand hygiene," *Clinical Microbiology and Infection*, vol. 20, no. 1, pp. 29–33, 2014.

[13] G. Kinsella, A. Thomas, and R. Taylor, "Electronic surveillance of wall-mounted soap and alcohol gel dispensers in an intensive care unit," *Journal of Hospital Infection*, vol. 66, no. 1, pp. 34–39, 2007.

[14] L. Lu, J. Yu, Y. Chen, H. Liu, Y. Zhu, Y. Liu, and M. Li, "Lippass: Lip reading-based user authentication on smartphones leveraging acoustic signals," in *Proceedings of the IEEE International Conference on Computer Communication (INFOCOM)*. IEEE, 2018, pp. 1466–1474.

[15] 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 *Proceedings of the IEEE International Conference on Computer Communication (INFOCOM)*. IEEE, 2019, pp. 1234–1242.

[16] D. Ma, G. Lan, W. Xu, M. Hassan, and W. Hu, "Simultaneous energy harvesting and gait recognition using piezoelectric energy harvester," *IEEE Transactions on Mobile Computing*, 2020.

[17] R. Ayyalasomayajula, A. Arun, C. Wu, S. Sharma, A. R. Sethi, D. Vasisht, and D. Bharadia, "Deep learning based wireless localization for indoor navigation," in *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking (Mobicom)*, 2020, pp. 1–14.

[18] D. M. Tax and R. P. Duin, "Support vector domain description," *Pattern Recognition Letters*, vol. 20, no. 11-13, pp. 1191–1199, 1999.

[19] M. Shoaib, S. Bosch, O. D. Incel, H. Scholten, and P. J. Havinga, "A survey of online activity recognition using mobile phones," *Sensors*, vol. 15, no. 1, pp. 2059–2085, 2015.

[20] L. P. Kaelbling, M. L. Littman, and A. R. Cassandra, "Planning and acting in partially observable stochastic domains," *Artificial intelligence*, vol. 101, no. 1-2, pp. 99–134, 1998.

[21] F. Fraternali, B. Balaji, D. Sengupta, D. Hong, and R. K. Gupta, "Ember: energy management of batteryless event detection sensors with deep reinforcement learning," in *Proceedings of the 18th Conference on Embedded Networked Sensor Systems (Sensys)*, 2020, pp. 503–516.

[22] Z. Shen, K. Yang, Z. Xi, J. Zou, and W. Du, "Deepapp: a deep reinforcement learning framework for mobile application usage prediction," *IEEE Transactions on Mobile Computing*, 2021.

[23] C. Luo and M. C. Chan, "Socialweaver: Collaborative inference of human conversation networks using smartphones," in *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems (Sensys)*. ACM, 2013, pp. 1–14.

[24] H. Kurniawati, D. Hsu, and W. S. Lee, "Sarsop: Efficient point-based pomdp planning by approximating optimally reachable belief spaces." in *Robotics: Science and systems*. MIT Press, 2008.

[25] I. Daubechies, J. Lu, and H.-T. Wu, "Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool," *Applied and Computational Harmonic Analysis*, vol. 30, no. 2, pp. 243–261, 2011.

[26] S. Mallat, "Group invariant scattering," *Communications on Pure and Applied Mathematics*, vol. 65, no. 10, pp. 1331–1398, 2012.

[27] J. Andén and S. Mallat, "Deep scattering spectrum," *IEEE Transactions on Signal Processing*, vol. 62, no. 16, pp. 4114–4128, 2014.

[28] T. Gong, Y. Kim, J. Shin, and S.-J. Lee, "Metasense: few-shot adaptation to untrained conditions in deep mobile sensing," in *Proceedings of the 17th Conference on Embedded Networked Sensor Systems (Sensys)*, 2019, pp. 110–123.

[29] C. Finn, P. Abbeel, and S. Levine, "Model-agnostic meta-learning for fast adaptation of deep networks," in *Proceedings of International Conference on Machine Learning (ICML)*. PMLR, 2017, pp. 1126–1135.

[30] W. E. Trick, M. O. Vernon, R. A. Hayes, C. Nathan, T. W. Rice, B. J. Peterson, J. Segreti, S. F. Welbel, S. L. Solomon, and R. A. Weinstein, "Impact of ring wearing on hand contamination and comparison of hand hygiene agents in a hospital," *Clinical Infectious Diseases*, vol. 36, no. 11, pp. 1383–1390, 2003.

[31] M. D. Wongworawat and S. G. Jones, "Influence of rings on the efficacy of hand sanitization and residual bacterial contamination," *Infection Control & Hospital Epidemiology*, vol. 28, no. 3, pp. 351–353, 2007.

[32] G. JACOBSON, J. E. THIELE, J. H. MCCUNE, and L. D. FARRELL, "Handwashing: ring-wearing and number of microorganisms," *Nursing Research*, vol. 34, no. 3, pp. 186–188, 1985.

[33] D. M. Salisbury, P. Hutfilz, L. M. Treen, G. E. Bollin, and S. Gautam, "The effect of rings on microbial load of health care workers' hands," *American Journal of Infection Control*, vol. 25, no. 1, pp. 24–27, 1997.

[34] D. Pittet, B. Allegranzi, J. Boyce, W. H. O. W. A. for Patient Safety First Global Patient Safety Challenge Core Group of Experts *et al.*, "The world health organization guidelines on hand hygiene in health care and their consensus recommendations," *Infection Control & Hospital Epidemiology*, vol. 30, no. 7, pp. 611–622, 2009.

[35] S. M. Rabeek, A. Norman, M. Je, M. K. Raja, R. F. Peh, and M. K. Dempsey, "A reliable handwash detector for automated hand hygiene documentation and reminder system in hospitals," in *The*

15th International Conference on Biomedical Engineering. Springer, 2014, pp. 892–895.

[36] A. Haque, M. Guo, A. Alahi, S. Yeung, Z. Luo, A. Rege, J. Jopling, L. Downing, W. Beninati, A. Singh *et al.*, "Towards vision-based smart hospitals: a system for tracking and monitoring hand hygiene compliance," in *Machine Learning for Healthcare Conference*. PMLR, 2017, pp. 75–87.

[37] H. Zhong, S. S. Kanhere, and C. T. Chou, "Washindepth: Lightweight hand wash monitor using depth sensor," in *Proceedings of the 13th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (Mobiquitous)*, 2016, pp. 28–37.

[38] L. L. Pineles, D. J. Morgan, H. M. Limper, S. G. Weber, K. A. Thom, E. N. Perencevich, A. D. Harris, and E. Landon, "Accuracy of a radiofrequency identification (rfid) badge system to monitor hand hygiene behavior during routine clinical activities," *American journal of infection control*, vol. 42, no. 2, pp. 144–147, 2014.

[39] Y. Wang, J. Lin, M. Annavaram, Q. A. Jacobson, J. Hong, B. Krishnamachari, and N. Sadeh, "A framework of energy efficient mobile sensing for automatic user state recognition," in *Proceedings of the 7th international conference on Mobile systems, applications, and services (Mobisys)*, 2009, pp. 179–192.

[40] A. Krause, M. Ihmig, E. Rankin, D. Leong, S. Gupta, D. Siewiorek, A. Smailagic, M. Deisher, and U. Sengupta, "Trading off prediction accuracy and power consumption for context-aware wearable computing," in *Proceedings of the 9th IEEE International Symposium on Wearable Computers (ISWC)*. IEEE, 2005, pp. 20–26.

Weitao Xu is an Assistant Professor at the Department of Computer Science at City University of Hong Kong. Before that, he was a Postdoctoral Research Associate at the School of Computer Science and Engineering (CSE) at UNSW from June 2017 to August 2019. He obtained his PhD degree from the University of Queensland in 2017 (advised by Prof. Neil Bergmann and Dr. Wen Hu). His research areas include mobile computing, sensor network and IoT. He is a member of IEEE.

Huanqi Yang is currently a Ph.D. student (since 2021) at Department of Computer Science, City University of Hong Kong. He is supervised by Dr. Weitao Xu. Huanqi Yang received his Bachelor degree (in 2021) from University of Electronic Science and Technology of China. His research interests lay in smart sensing, IoT security, IoT+AI, wireless network.

Jiongzhang Chen received the B.S. degree (in 2020) from Shenzhen University (SZU), China. He is currently working toward the master's degree (since 2020) at School of Computer Science and Software Engineering, Shenzhen University (SZU), China. He is supervised by Prof. Chengwen Luo. His current research intersets include edge AI, deep leaining, and split learning.

Chengwen Luo received the PhD degree from the School of Computing, National University of Singapore (NUS), Singapore. He is currently an associate professor in the College of Computer Science and Software Engineering, Shenzhen University (SZU), China. Before joining SZU, he was a postdoctoral researcher in CSE, The University of New South Wales (UNSW), Australia. His research interests include mobile and pervasive computing and security aspects of Internet of Things.

Jia Zhang (S'10, M'14) received the Ph.D. degree in Information and Communication Engineering from Shandong University, Jinan, China, in 2013. Before that, she received the B.E. and M.E. degrees from Harbin Engineering University in 2006 and 2009, respectively. Her research interests are in the area of wireless communications, including non-orthogonal multiple access for 6G and beyond, secure and private wireless transmission, and RIS aided communications.

Yuliang Zhao received a B. S. degree in Mechanical Engineering from the Hubei University of Automotive Technology, an M. S. degree in Mechanical Engineering from the Northeastern University, and a Ph. D. degree in Mechanical and Biomedical Engineering from the City University of Hong Kong in 2016. He is an associate professor with the Northeastern University at Qinhuangdao, Qinhuangdao, China. His research interests include intelligent sensors, machine learning, motion analytics.

Wen Jung Li received the B.S. and M.S. degrees in aerospace engineering from the University of Southern California (USC), in 1987 and 1989, respectively, and the Ph.D. degree in aerospace engineering from the University of California at Los Angeles (UCLA), in 1997. He is currently a Chair Professor with the Department of Mechanical Engineering, City University of Hong Kong. His current research interests include intelligent cyber physical sensors, and nanoscale sensing and manipulation.