

Detecting hand washing activity among activities of daily living and classification of WHO hand washing techniques using wearable devices and machine learning algorithms

Yiyuan Zhang^{1,2}  | Tianwei Xue¹ | Zhenjie Liu¹ | Wei Chen³  | Bart Vanrumste^{1,2} 

¹ e-Media Research Lab, KU Leuven, Leuven, Belgium

² STADIUS, Department of Electrical Engineering, KU Leuven, Leuven, Belgium

³ Center for Intelligent Medical Electronics, School of Information Science and Technology, Fudan University, Shanghai, China

Correspondence

Yiyuan Zhang, e-Media Research Lab, KU Leuven, 3000 Leuven, Belgium.
Email: yiyuan.zhang@kuleuven.be

Abstract

During COVID-19, awareness of proper hand washing has increased significantly. It is critical that people learn the correct hand washing techniques and adopt good hand washing habits. Hence, this study proposes using wearable devices to detect hand washing activity among other daily living activities (ADLs) and classify steps proposed by the World Health Organization (WHO). Two experiments were conducted with 16 participants, aged from 20 to 31. The first experiment was hand washing following WHO regulation (ten participants), and the second experiment was performing eight ADLs (eight participants). All participants wore two wearable devices equipped with accelerometers and gyroscopes; one on each wrist. Four machine learning classifiers were compared in classifying hand washing steps in the leave-one-subject-out (LOSO) mode. The SVM model with Gaussian kernel achieved the best performance in classifying 11 washing hands steps, with an average F1-score of 0.8501. When detected among the other ADLs, hand washing following WHO regulation obtained the F1-score of 0.9871. The study demonstrates that wearable devices are feasible to detect hand washing activity and the hand washing techniques as well. The classification results of getting the soap and rubbing thumbs are low, which will be the main focus in the future study.

1 | INTRODUCTION

Appropriate hand washing is crucial to prevent the transmission of bacteria, viruses and parasites [1, 2] and reducing the rate of infections [3, 4]. In particular, during the period of the COVID-19 pandemic, the importance of hand washing has been emphasised [5]. There are two important aspects of hand washing: the habit (when to wash hands) and the techniques (how to wash hands). Atee et al. mentioned ten occasions when hand washing is necessary [6]. However, there are still cultural differences in the hand washing habit. For example, in an international survey among 63 countries in 2015, it appeared that around 77% of people did not wash their hands after using the toilet, and even in the developed countries, the rate was high, such as 50% in Netherlands and 43% in Italy [7]. Regarding the hand washing quality, the World Health Organization (WHO) proposes the regulation of hand washing techniques with six major steps

(steps 2–7) as shown in Figure 1 whereby hand washing lasts for 40–60 s [8]. Thus, the proper hand washing techniques must be learned, along with forming a good hand washing habit. To monitor these, sensor technology and machine learning algorithms can be used to identify different classes of the hand washing techniques proposed by the WHO and detect hand washing among other daily living activities (ADLs).

Multiple types of sensors have been investigated previously, such as internet-of-things (IoT) [9, 10], audio sensor [11, 12], and camera [13–15]. Nevertheless, these sensors still have limitations since they need to be set up close to the activity area. People wash their hands in various places at home, such as the bathroom or kitchen. Hence, IoT sensors need to be installed in all these places, which increases the expense and the complexity of setup. Additionally, it is advisable to provide individuals with feedback, especially while helping them form the hand washing habit. However, these sensors cannot identify different

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2021 The Authors. *Healthcare Technology Letters* published by John Wiley & Sons Ltd on behalf of The Institution of Engineering and Technology



FIGURE 1 WHO recommended hand washing steps [8]

users unless assisted with other sensors, for example, Beacon or motion sensors [16]. Despite being able to recognise people, cameras cause privacy concerns, especially when installed in the bathroom. In addition, they cannot detect hand washing steps (techniques).

As an alternative method, wearable sensors, including an accelerometer and gyroscope, can be applied. In the study of Mondol et al. [17], the Harmony system was developed to detect the hand washing quality for food workers. In this system, a smartwatch was attached to the participant's wrist to collect the

movement signals, a Bluetooth transmitter was integrated into the dispenser to detect whether the participant used the soap, and a Beacon sensor was used to send alerts when the participant was in the area that required hand hygiene or when the participant did not properly wash hands. The accelerometer and gyroscope signals were processed to extract the time-domain features, and a decision tree (DT) was utilised to detect classes of hand washing. The classifier was trained in the user-independent mode. In this model, all the participants are split into two groups, the data of one group is used for training,

and the other group's data is used for testing. As a result, the DT classifier received an average accuracy of 85%. In the later study [18], researchers explored the impact of the sensor position, and they tried to detect hand washing activity among other ADLs using the Feed Forward Neural Network (FFN). With the F1-scores of 0.72 and 0.74 for the left and right wrist, respectively, there was no significant difference between which sides the wearable device was worn. Because each participant only wore the smartwatch on one side, the result may be impacted by the different performances of participants. The WISDM dataset [19] was added without including hand washing to test whether other ADLs are misclassified as hand washing. The authors achieved the result by reducing the false positive rate by 77% compared to the baseline method [20]. However, they did not discuss which ADLs are more likely to be classified as hand washing.

In study [21], also one wristwatch was used, including the accelerometer and gyroscope sensors. The data were collected both in a controlled and uncontrolled environment. A hidden Markov model (HMM) classifier was used with an accuracy of 85% with firstly determining the order of hand washing classes in the user-independent mode. Compared to study [17, 18], the before (rubbing hands) and after (turning off the faucet) hand washing activities were included. These two classes were detected using an extra binary classifier, with an accuracy ranging from 35% to 54%. Rather than attaching sensors to the wrist, Wang et al. used four armbands which also included Electromyography (EMG) sensors [22]. Two armbands were placed on the forearms and two on the upper arms. This study investigated two types of hand hygiene, alcohol-based hand rubbing and hand washing with soap and water. During data acquisition, each type was performed 15 times/participant. The after hand washing classes were categorised into three classes: rinsing hands, drying hands with a towel, and turning off the faucet. EXtreme Gradient Boosting (XGBoost) combined with E.Divisive [23] was applied. Furthermore, Wang et al. discussed the impact of different sensors and the attached position. The highest accuracy was achieved using the signals from all three sensors (accelerometer, gyroscope, and EMG) from both armbands attached to the forearm for 91% in the user-independent mode. The classifier was also trained in the user-dependent mode: one part of the dataset of a participant used for training and the remaining part of the same participant used for testing. The number of sessions used for training the model was examined in the user-dependent mode, which indicated that after using the data of 16 sessions, the classification accuracy could reach 90%. In study [16], an online monitoring system including a smartwatch was developed with the Convolutional (CNN) combined with Long Short Term Memory (LSTM) neural network as the classifier. The accuracy of each step was over 90% in the real-life dataset.

Previous studies have proven the feasibility of using wearable sensors, especially accelerometers and gyroscopes, to monitor the hand washing techniques suggested by the WHO and determine the duration of hand washing. However, they did not detect all the steps covered in the hand washing techniques, for example, getting soap. In addition, few studies [18] looked at

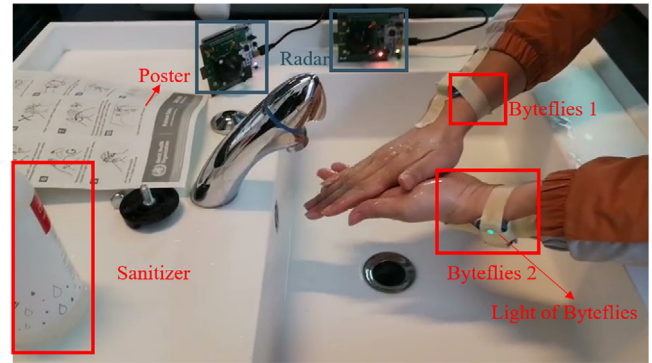


FIGURE 2 Setting up of the experiment

detecting hand washing among other ADLs. At last, the impact of sensors and the wearing side was examined based on the average classification performance rather than for each hand washing step. Therefore, this study proposes a wireless system with wearable devices to monitor hand washing. The contributions of this study are:

- Classifying hand washing steps in the WHO regulation.
- Classifying hand washing activity among other ADLs
- Differentiating hand washing activities between not following (untrained) and following (trained) the WHO regulation.
- Investigating the impact of the combination of sensors on the classification of each hand washing step.

To accomplish this, two wearable sensors composed of three-axis accelerometers and gyroscopes were attached to each wrist. Features are extracted from both the time and frequency domain, and multiple machine learning algorithms are explored in the leave-one-subject-out (LOSO) mode.

2 | METHODOLOGY

2.1 | The wearable device

As shown in Figure 2, two Byteflies sensors [24] were used, one on each wrist. The sensor includes a three-axis accelerometer and a three-axis gyroscope, both running at 100 Hz. The placement direction on each wrist was fixed. The sensors were automatically synchronised. Two radar sensors were also applied to detect the hand washing movement during the experiment, but their received signals will not be discussed in this paper.

2.2 | Dataset preparation

2.2.1 | Hand washing data collection

This study involved ten individuals (seven males, three females, aged 20–23). Participants washed their hands with the WHO regulation poster placed in front of them and instructed by a researcher. Each participant performed hand washing three

TABLE 1 Annotated hand washing classes

Class	Detail (Related steps proposed by the WHO as shown in Figure 1)
a	0, 8, 10
opening the faucet, wetting/ rinsing hands and turning off the faucet	
b	1, getting the soap
c	2, rubbing hands palm to palm
d	3, right palm over left dorsum with interlaced fingers
e	3, left palm over right dorsum with interlaced fingers
f	4, palm to palm with interlaced fingers
g	5, backs of fingers to opposing palms with interlaced fingers
h	6, rubbing left thumb
i	6, rubbing right thumb
j	7, clasped fingers of right hand in left palm
k	7, clasped fingers of left hand in right palm

times, each time spending 60–80 s. One camera was placed aside for recording as the ground-truth annotation. Eleven activities (classes a–k) were labelled, as listed in Table 1. The classes were categorised into three groups: before hand washing (steps 0, 1, and 2), hand washing (steps 3–11), after hand washing (steps 0 and 1). Class a covers three hand washing steps: opening the faucet, wetting/ rinsing hands, and turning off the faucet. These activities were performed both before and after hand washing. Therefore, they were categorised into one class. Step 9, drying hands, was not included. Considering real application, people usually first turn off the faucet and then get the towel to dry their hands. Thus, hand drying was not included in this study.

2.2.2 | ADLs data collection

To test the system's ability to classify hand washing among other ADLs, we also conducted another experiment that included eight ADLs: sitting, standing, walking, computer typing, going upstairs, and going downstairs, brushing teeth, and untrained hand washing. These activities were included because:

- These activities are basic ADLs and are normally performed every day.
- Brushing teeth is also one type of bathroom activity and it also includes high-frequency movements of the wrists.

Eight participants joined the experiment (three females and five males, aged 23–31). The right wrist was the dominant wrist of all participants. Among the eight participants, two people

joined the experiment of hand washing dataset collection as well. Same as the experiment of hand washing, all participants wore two Byteflies sensors, one on each wrist. In this experiment, participants were asked to perform activities in their own way. The performed activities and procedures are as following:

1. Sitting (3 min). Participants could talk, use smartphones, or with slight arm movements.
2. Computer typing (3 min). Participants were asked to type the given text material.
3. Standing (3 min). The requirement is the same as sitting.
4. Walking (180 m). Participants were asked to walk at their normal speed.
5. Going upstairs (60 stairs \times 2 times). Participants were asked to go at their normal speed.
6. Going downstairs (60 stairs \times 2 times). Participants were asked to go at their normal speed.
7. Hand washing (3 times). The time was recorded from participants opening the faucet to rinsing hands and closing the faucet.
8. Brushing teeth. The time was recorded from participants squirting the toothpaste on the toothbrush to rinsing the mouth. All the participants used manual toothbrushes.

One researcher was by aside to annotate activities. The experiments were approved by the KU LEUVEN Social and Societal Ethics Committee (SMEC), with the assigned serial number G-2020-2705. All participants signed the informed consent form before participating in the experiment.

2.3 | Data analysis

2.3.1 | Data pre-processing

The collected accelerometer and gyroscope signals were segmented based on the annotation of the video recording. Afterwards, segmented signals were then sliced by a fixed-length window. In the previous studies, the applied windows size ranged from 0.06 s (three data points) [21] to 1 s (50 data points) [17], with the overlap rate in the range of (50%, 70%). In this study, the testing window size was varied between (0.08 s, 3.5 s), with an overlap rate of 0/50%. A Support Vector Machine (SVM) classifier was used as the classifier, trained in the user-independent mode. The result of optimal window size is shown in Section 3.

2.3.2 | Feature extraction

The features were extracted from each axis and the magnitude of the three axes. In the time domain, the extracted features were the mean, standard deviation, root mean square (RMS), minimum (min), maximum (max), range (= max–min), interquartile range, skewness, kurtosis, mean-crossing times (= the number of the time the signal value passing through the mean value), and

TABLE 2 Hyper-parameters tuning results of classifiers in the user-independent mode

Classifier	Hyper-parameter	Searching range	Tuned value
SVM	C - penalty parameter	$2^n, n \in (-5, 15)$ [28]	2^5
	γ	$2^n, n \in (-15, 5)$	2^{-7}
KNN	K	1 – 15	3
XGBoost	maximum depth of decision trees	1 – 10	2
	minimum child weight	1 – 5	4
	γ	$2^n, n \in (-5, 5)$	2^5
FFN	Number of hidden layers	1 – 5	3
	Number of neurons	$2^n, n \in (1, 8)$	128 - layer 1
			128 - layer 2
64 - layer 3			

energy. In addition, the correlation between every two axes and zero-crossing times ($=$ the number of the time the signal value passing through zero) were calculated from each axis, except for the magnitude signal. In the frequency domain, the frequency power, the frequency at the highest amplitude, 25% frequency (F25d), 75% frequency (F75d), Normalised frequency ($1 - F25d/F75d$) and spectral energy were extracted. As a result, 74 features were extracted from each sensor, with 296 features extracted from all sensors.

2.3.3 | Classification model

Four types of machine learning (ML) models were investigated as they were applied in the previous works: SVM [10, 22], K-Nearest Neighborhood (KNN) [16], XGBoost [22], and FFN [18]. The SVM model applied the Gaussian kernel and was trained in the one-versus-one mode. Hence, 55 binary classification models were trained, and the predicted class was assigned to the one with the highest score. The hyper-parameters of the models were tuned using the grid-search method with testing all the combinations [25]. The searching range and the tuning result are shown in Table 2. While developing the FFN architecture, one hidden layer was added each time, and the number of neurons in this layer was tuned. We stopped adding the layer until the classification performance dropped on the cross-validation set. Moreover, the searching range of the number of neurons of the new hidden layer did not exceed that of the previous one. All the models were trained in the LOSO mode: in each fold, one participant's data were selected for testing, one participant's data for validation, and the remaining eight participants' data were for training. In this study, the models were trained in ten folds. The F1-score was applied as the evaluation metric. The class weight balancing method [26, 27] was used to solve the imbalanced dataset problem by assigning the class weight to be the inverse of the number of windows. The model with the highest performance was used in the following cases:

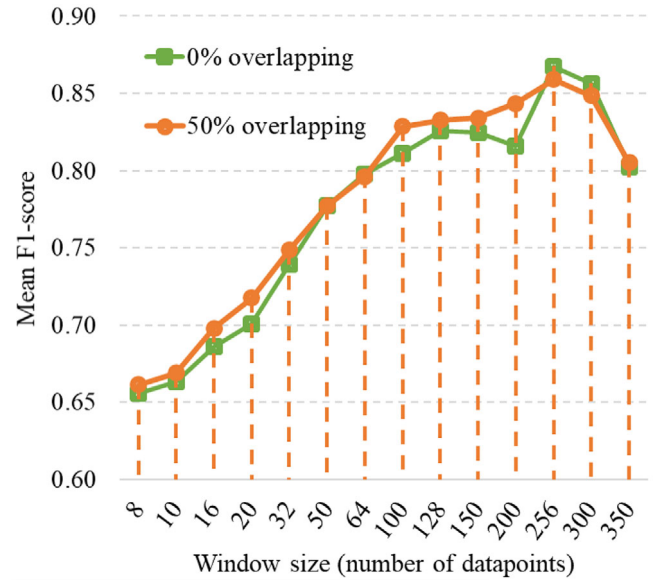


FIGURE 3 Plot of the mean F1-score of detecting handwashing steps using various sizes of sliding window, with the overlap rate of 0% and 50%, separately

- Investigating the optimal combination of the sensors. Various combinations of sensors were tested, including only the accelerometer sensor from one side, both accelerometer and gyroscope sensors on one side, to all sensors.
- Detecting the hand washing activity among other ADLs. The hand washing dataset was combined with the ADLs dataset. As aforementioned, the ADLs dataset did not include hand washing activity that followed the WHO regulation. Hence, half of the participants' data were used for training in each dataset and the remaining for testing. In other words, the data of five participants in the hand washing dataset and four participants in the ADLs dataset were selected as the training dataset. The data of the remaining participants were selected as the testing dataset.
- Differentiate trained and untrained hand washing activities. While combining the hand washing dataset and ADLs dataset, the hand washing activities in these two datasets were categorised into two classes.

3 | RESULT

3.1 | Classification of hand washing techniques

3.1.1 | Optimal window size selection

Figure 3 displays the F1-score result of multiple sizes of windows and overlap rates. The result implied that with the increase in the window length, the classification performance rises as well. This is because more information can be extracted in a longer window, and it can reduce the impact of abrupt changes in the signal. The 0% overlap has a similar result as the 50%

a	62	10					1				1
b	6	20	1	1							
c	1	2	82							3	
d		1		30							
e			1		28					1	
f	1					80					
g	2	1					78		1		
h							25	8			
i	1						7	23			
j		1	6						23		
k	1					1					29
	a	b	c	d	e	f	g	h	i	j	k

FIGURE 4 Confusion matrix of the classification result of the SVM model using all the sensors. a: opening the faucet, wetting/rising hands, turning off the faucet; b: getting soap; c: rubbing hands palm to palm; d: right palm over left dorsum; e: left palm over right dorsum; f: palm to palm with fingers interlaced; g: backs of fingers to opposing palms with fingers interlocked; h: rubbing left thumb; i: rubbing right thumb; j: rubbing with clasped fingers of right hand; k: rubbing with clasped fingers of left hand

one, and it is even a little better than the 50% when the window length equals to 2.56 s and 3.00 s. Hence, in the following study, the 2.56 s window with no overlap is selected.

3.1.2 | Comparison among classifiers

Table 3 presents the classification result of each classifier for detecting hand washing classes. Among the 11 classes, KNN obtains the highest performance in classes b, h, and i, and SVM obtains the highest result in the remaining eight classes. As a result, SVM achieves the highest average F1-score of 0.8501 ± 0.1655 , followed by FFN with the F1-score of 0.8236 ± 0.2014 . The mean F1-score of class j is only 0.2286 using the KNN classifier, which impacts the average performance of KNN. This will be explained in detail in Section 3.1.3. The standard deviation (std) of the F1-score of all classifiers ranges from 0.1655 to 0.2799, which implies the subjective difference.

3.1.3 | Misclassification among hand washing classes

Figure 4 presents the confusion matrix of the SVM model using features extracted from all sensors. The main misclassifications exist in three groups: classes a & b, classes h & i, and classes j & c. The misclassification between classes a and b could be because different people use different sides of the hand to open and turn off the faucet and get the soap. During the experiment,

TABLE 3 The comparison of the F1-score (mean \pm std) result of hand washing classes detection among classifiers. The highest result of each class is highlighted in bold

Classifier	Classes											Mean
	a	b	c	d	e	f	g	h	i	j	k	
SVM	0.8332 \pm 0.1657	0.6555 \pm 0.2353	0.9168 \pm 0.0714	0.9657 \pm 0.0698	0.9500 \pm 0.1500	0.9856 \pm 0.0300	0.9676 \pm 0.0408	0.6883 \pm 0.3679	0.6861 \pm 0.3158	0.7564 \pm 0.2898	0.9457 \pm 0.0842	0.8501 \pm 0.1655
KNN	0.8349 \pm 0.1743	0.6589 \pm 0.2385	0.8427 \pm 0.0722	0.9314 \pm 0.0859	0.9400 \pm 0.0917	0.9404 \pm 0.1176	0.9428 \pm 0.0995	0.7336 \pm 0.2945	0.6967 \pm 0.1939	0.2286 \pm 0.3255	0.9057 \pm 0.1310	0.7869 \pm 0.1659
XGBoost	0.7507 \pm 0.1794	0.5657 \pm 0.3006	0.8129 \pm 0.2783	0.8350 \pm 0.2950	0.7536 \pm 0.2073	0.8365 \pm 0.3192	0.9105 \pm 0.1866	0.6163 \pm 0.3541	0.6557 \pm 0.2861	0.4919 \pm 0.3693	0.7833 \pm 0.3036	0.7284 \pm 0.2799
FFN	0.7767 \pm 0.1339	0.6217 \pm 0.3581	0.9227 \pm 0.0964	0.9300 \pm 0.1522	0.8900 \pm 0.1578	0.9856 \pm 0.0300	0.9531 \pm 0.0719	0.6832 \pm 0.3974	0.6824 \pm 0.3759	0.7186 \pm 0.2833	0.8957 \pm 0.1555	0.8236 \pm 0.2014

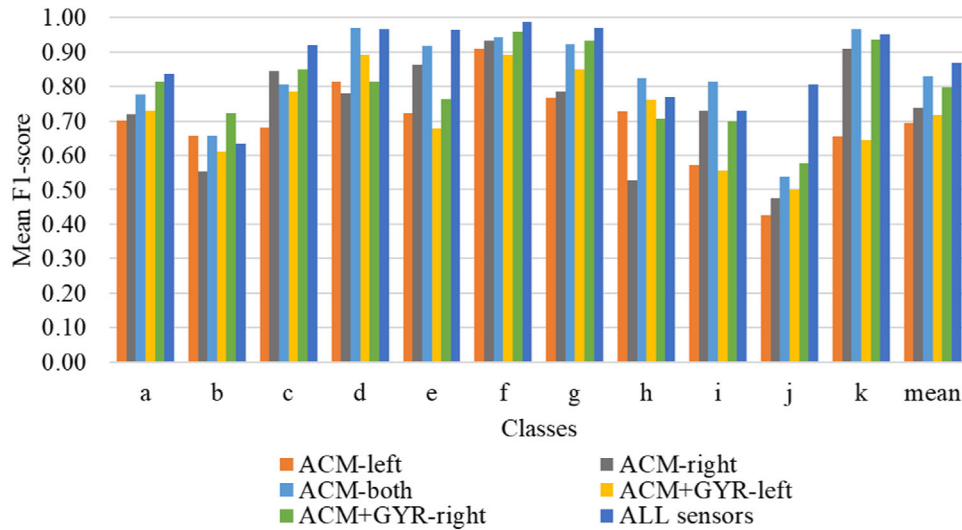


FIGURE 5 Comparison of the different combinations of the sensors in the classification of WHO hand washing steps. ACM, accelerometer; GYR, gyroscope; left, left wrist; right, right wrist

six of the participants used their right hands to open and turn off the faucet, and five of them kept using their right hand to get the soap. In addition, the major movement of classes a and b is the same, which is moving the hand from one point to another point. Classes h and i belong to the same step in the WHO regulation, but the hand side varies. While performing these two classes, participants were asked to rub their thumb, but different subjects did so differently, either rotating only one hand or both hands simultaneously. At last, the movements of classes c and j were similar, with the right hand rotating on the palm of the left hand. The only difference is that the fingers of the right hand were clasped while performing class j, which our sensors cannot detect. Thus, it is reasonable that these two classes are confused with each other.

3.1.4 | Combination of sensors

Figure 5 presents the mean F1-score results of various combinations of sensors:

- only using the accelerometer on the left/right wrist
 - only using accelerometer and gyroscope sensors on the left/right wrist
 - using all the sensors on both wrists.
1. The average results of only using the accelerometer on the left/right side are similar. The accelerometer worn on the left wrist performs better in detecting classes b, d and h. In contrast, the accelerometer on the right side achieves significantly better results in classes c, e, i, and k. The reason could be that the main wrist is different in every step, and the accelerometer is sensitive to the movement amplitude. Hence, although classes d and e are categorised as the same step in the WHO regulation, their classification results are

different while using an accelerometer on a different side. This is the same reason for the results of classes h and i. For class k, the requirement was clasping the left finger on the right palm. Since among the participants, nine of them had the right wrists as their dominant wrists, they also rotated their right hands while performing class k. Moreover, the rotation amplitude was larger than the one of class j.

2. After including the gyroscope sensors, the performance of the sensors on the right side is improved in most classes, except for class i. The average F1-score of using the accelerometer and gyroscope attached to the left wrist is a little higher than only using the accelerometer. This is because class b is getting the soap. Therefore, the gyroscope sensor can only detect rotations when the participants using their hands to catch the soap drops, but which hand to get the soap was varied in participants.
3. Among all the combinations, using all the sensors on both sides achieved the best results, except for class b. This indicates the importance of the sensors on both sides and gyroscope sensors. The reasons for the reduced performance of class b can be that: (1) the number of windows of class b is smaller than other classes; (2) the movements performed in classes a and b were similar to each other. Hence, with the increased types of features, classes a and b were misclassified, which mainly reduced the precision and recall result.

3.2 | Hand washing detection among other ADLs and with hand washing activity without following the WHO regulation

3.2.1 | Combination of sensors

In the ADLs dataset, the right wrist was the dominant wrist of all participants. In addition, while performing brushing teeth, all

TABLE 4 The F1-score of trained hand washing among other daily activities with various combinations of sensors

Sensors	ACM- right wrist	ACM- both wrists	ACM + GYR - right wrist	All sensors
F1-score	0.7625	0.8256	0.9545	0.9871

True Class	Predicted Class								
	sit	computer type	stand	walk	going upstairs	going downstairs	brush teeth	WH1	WH2
sit	204	46	29					1	
computer type	69	199	3				2	4	3
stand	34	1	236		2	1	5		1
walk				134	109	10	9		2
going upstairs	2		4	36	115	12	13	1	4
going downstairs			4	28	60	51	11		9
brush teeth	6	1	9	10	7		217	2	21
WH1	4	1	1					345	
WH2			1	15	2	1	32		149

FIGURE 6 Confusion matrix of detecting hand washing among other ADLs using all sensors. WH1, untrained hand washing; WH2, trained hand washing. The confusions with WH1 and WH2 are highlighted in the red blocks

participants held the toothbrush using their right wrists. Hence, in this part, the data of sensors attached to the right wrist were always included. The F1-score of trained hand washing of each combination is shown in Table 4. Using all the sensors achieved the highest result in detecting trained hand washing. In addition, the result of using both the accelerometer and gyroscope on the right wrist is around 0.2 higher than only using one accelerometer and 0.13 higher than using two accelerometers. This could be because hand washing is only associated with hands movements and most hand washing postures are periodic rotation movements, which is detectable by the gyroscope.

3.2.2 | Difference between trained and untrained hand washing activities

Figure 6 shows the confusion matrix of detecting hand washing activity among other ADLs using all types of sensors. The trained hand washing activity is rarely confused with other ADLs activities, implying that using wearable sensors can recognise the hand washing activities among the other seven ADLs. In addition, the trained hand washing class is not confused with the untrained one. The could be due to the types of movement, as the hand washing movement following the WHO regulation is more periodic than the untrained one. One example can be found in Figure 7.

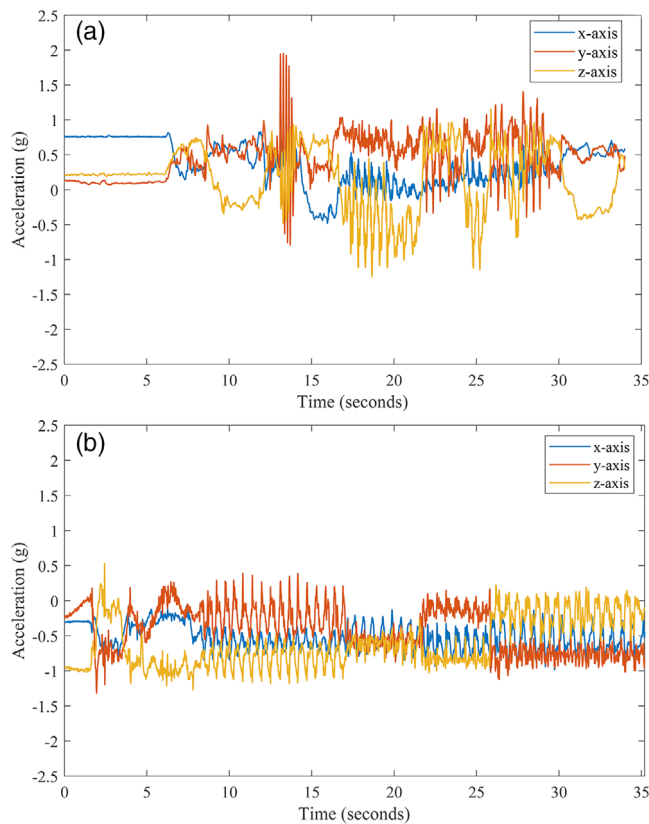


FIGURE 7 Examples of acceleration signals of hand washing; $g = 9.8 \text{ m/s}^2$

4 | DISCUSSION

4.1 | Compared with previous studies

Table 5 summarises the accuracy results of every hand washing step of the proposed method and previous studies. As some previous studies [17, 21] only presented their results in figures, their results are not listed here. Study [22] and study [21] included hand washing steps 8–10: rinsing hands, drying hands, and turning off the faucet. However, we categorised these three steps into one class (class a). Hence, to comparison, the results of steps 8 and 9 of study [21, 22] are listed together in Table 5. Study [21] and study [20] separated class f into two classes, left and right, so there are two values in that column. As a result, the proposed method is comparable to the other studies, especially for classifying classes d, e, f, and k. However, the results of classes h, i, and j for the proposed method are lower than those of other studies. The reason could be that the number of sensors used in the proposed method is less than study [22], or the dataset size is smaller.

In our study, the optimal window size is 256 data points with no overlap, which is not in the range proposed by other studies that were 3–50 data points with 50–75% overlap. This shows that the optimal window size depends on the types and number of activity classes. For instance, in study [22], the average duration of turning off the faucet was around 0.5 s, so it was reasonable to select the 0.2 s window with 75% overlap, and it can

TABLE 5 The comparison of the accuracy result of the proposed method with previous study

Study	Sensor	Classifier	Activities														
			Rinsing hands			Closing the faucet			f			g			Drying hands		
			a	b	c	d	e	f	Left	Right	Left	Right	Left	Right	Left	Right	Left
Proposed method	accelerometer+gyroscope	SVM	84%	71%	93%	97%	93%	99%	93%	99%	95%	76%	74%	77%	94%	-	-
[22]	accelerometer+gyroscope+EMG	XGBoost+E.Divisive	100%	-	83%	83%	99%	87%	99%	86%	84%	84%	84%	94%	94%	100%	100%
[21]	accelerometer+gyroscope	HMM	80%	-	97%	82%	85%	93%	93%	77%	77%	86%	75%	87%	89%	87%	87%
[20]	accelerometer+gyroscope	KNIN	-	-	93%	91%	84%	93%	98%	94%	87%	88%	84%	84%	84%	-	-

increase the detection precision in real-life application. However turning off the faucet is incorporated with turning on the faucet and rinsing hands as in one class in our study. For the overlap rate, although a higher overlap rate can enlarge dataset size, it can also increase the biased result with similar windows tested.

Compared to previous studies, our study has the advantages that:

- We include getting the soap as one class, even though the accuracy is 71%. Since getting soap was detected by the Bluetooth sensors integrated into the dispenser in the previous study [17], our result from wearable sensors suggests that the IoT sensors in the dispenser can be removed, which would simplify the installation and reduce the cost. In addition, this study categorises opening the faucet, rinsing hands, and turning off the faucet into one class. This is because the period of each action is short, and they are typically performed sequentially.
- We prove the feasibility of recognising WHO recommended hand washing activity among other ADLs, using two wearable devices. Hence, the next step is to detect occasions in which hand hygiene is necessary, which can help people form good hand washing habits via sending alarm signals when they forget to wash their hands, such as after going to the toilet.
- We analyse the difference between trained hand washing activity from the untrained one. This analysis result can be used to develop the training system to help people learn how to wash their hands properly.
- In our study, we do not consider the order of the hand washing steps. Thus, we used the SVM model. Moreover, the classification results of six classes are higher than the ones of study [21] using HMM model, which implies that the order of hand washing steps is not essential to classification.

4.2 | Limitations and future work

1. Although SVM achieved 0.8501 for the average F1-score in classifying hand washing steps, the results of getting the soap (class b) and rubbing thumbs (classes h and i) are much lower than the average score, less than 0.7. Thus, the future study will focus on increasing the classification performance of these steps. The possible solution can be including other sensors, for instance, the pressure sensor/motion sensors for detecting getting the soap and EMG sensors for detecting the movement of fingers.
2. In this study, we only considered one method of performing brushing teeth, which is using a manual toothbrush. The future work can also include samples of using an electric toothbrush.
3. In this study, sensors attached to both wrists achieved the highest results for both classification of hand washing techniques and detecting hand washing activity among other ADLs. However, we mainly focused on the participants with the right wrist as the dominant side. The future study will also take participants with the left wrist as the dominant side into consideration.

4. Byteflies sensors were used in our study, but they may not be convenient for users. Because the sensors were attached to the body via patches, which was not reusable. In the future, this sensor can be replaced with other similar devices, such as the smartwatch that also includes both the accelerometer and gyroscope sensors.

5 | CONCLUSION

This study demonstrates the feasibility of a system using two Byteflies sensors and an SVM classifier to classify 11 hand washing steps. The result suggests that the optimal combination of the devices is two devices integrated with an accelerometer and a gyroscope, one device on each wrist. The hand washing activity can also be classified among other ADLs and untrained hand washing activity. The system will be improved to serve as a training system for individuals who need to form a good hand washing habit.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ETHICAL APPROVAL

The experiments were approved by the KU LEUVEN Social and Societal Ethics Committee (SMEC), with the assigned serial number G-2020-2705. All participants signed the informed consent form before participating in the experiment.

DATA AVAILABILITY STATEMENT

The datasets analysed in this study are not shared due to privacy or ethical restrictions.

ORCID

Yiyuan Zhang  <https://orcid.org/0000-0002-4522-8354>

Wei Chen  <https://orcid.org/0000-0003-3720-718X>

Bart Vanrumste  <https://orcid.org/0000-0002-9409-935X>

REFERENCES

- Burton, M., Cobb, E., Donachie, P., Judah, G., Curtis, V., Schmidt, W.-P.: The effect of handwashing with water or soap on bacterial contamination of hands. *Int. J. Environ. Res. Public Health* 8(1), 97–104 (2011). <https://doi.org/10.3390/ijerph8010097>
- Strohbehn, C., Sneed, J., Paez, P., Meyer, J.: Hand washing frequencies and procedures used in retail food services. *J. Food Prot.* 71(8), 1641–1650 (2008). Available: <https://doi.org/10.4315/0362-028X-71.8.1641>
- World Health Organization (WHO): Hand hygiene. Available: <https://www.who.int/teams/integrated-health-services/infection-prevention-control/hand-hygiene>. Accessed 4 June 2021
- Rabie, T., Curtis, V.: Handwashing and risk of respiratory infections: A quantitative systematic review. *Trop. Med. Int. Health* 11(3), 258–267 (2006). <https://doi.org/10.1111/j.1365-3156.2006.01568.x>
- Alzyood, M., Jackson, D., Aveyard, H., Brooke, J.: COVID-19 reinforces the importance of handwashing. *J. Clin. Nursing* 29(15-16), 2760–2761 (2020). <https://doi.org/10.1111/jocn.15313>
- Ataee, R.A., Ataee, M.H., Tavana, A.M., Salehi, M.: Bacteriological aspects of hand washing: A key for health promotion and infections control. *Int. J. Preventive Med.* 8(1), 16 (2017). <https://doi.org/10.4103/2008-7802.201923>
- BreakingWeb: Les Français et le savonnage des mains après être allé aux toilettes. Available: <https://www.bva-group.com/sondages/les-francais-et-le-savonnage-des-mains-apres-etre-alle-aux-toilettes/>. Accessed 4 June 2021
- World Health Organization (WHO): How to handwash and handrub. Available: <https://www.who.int/multi-media/details/how-to-handwash-and-handrub>. Accessed 4 June 2021
- Thyroff, L.: System and method for monitoring handwashing compliance including soap dispenser with integral hand-washing monitor and smart button system. US Patent US10,332,382B2 (2019). Available: <https://patents.google.com/patent/US10332382B2/en>
- Fagert, J., Mirshekari, M., Pan, S., Zhang, P., Noh, H.Y.: Monitoring hand-washing practices using structural vibrations. In: *Structural Health Monitoring 2017*. DEStech Publications, Inc., Lancaster (2017). Available: <http://www.dpi-proceedings.com/index.php/shm2017/article/view/14133>
- Chen, J., Kam, A.H., Zhang, J., Liu, N., Shue, L.: Bathroom activity monitoring based on sound. In: Hutchison, D., Kanade, T., Kittler, J., Kleinberg, J.M., Mattern, F., Mitchell, J.C., Naor, M., Nierstrasz, O., PanduRangan, C., Steffen, B., Sudan, M., Terzopoulos, D., Tygar, D., Vardi, M.Y., Weikum, G., Gellersen, H.W., Want, R., Schmidt, A. (eds.) *Pervasive Computing. Lecture Notes in Computer Science*, vol. 3468, pp. 47–61. Springer, Berlin Heidelberg (2005). [Online]. Available: http://link.springer.com/10.1007/11428572_4
- Siantikos, G., Giannakopoulos, T., Konstantopoulos, S.: A low-cost approach for detecting activities of daily living using audio information: A use case on bathroom activity monitoring. In: *Proceedings of the International Conference on Information and Communication Technologies for Ageing Well and e-Health*, pp. 26–32. SCITEPRESS - Science and Technology Publications, Rome, Italy (2016). Available: <http://www.scitepress.org/DigitalLibrary/Link.aspx?doi=10.5220/0005803700260032>
- Ivanovs, M., Kadikis, R., Lulla, M., Rutkovskis, A., Elsts, A.: Automated quality assessment of hand washing using deep learning. arXiv:2011.11383 (2020). Available: <http://arxiv.org/abs/2011.11383>
- Pirsiavash, H., Ramanan, D.: Detecting activities of daily living in first-person camera views. In: *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2847–2854. IEEE, Piscataway (2012). <http://doi.org/10.1109/CVPR.2012.6248010>
- Llorca, D.F., Parra, I., Sotelo, M.Á., Lacey, G.: A vision-based system for automatic hand washing quality assessment. *Mach. Vision Appl.* 22(2), 219–234 (2011). <https://doi.org/10.1007/s00138-009-0234-7>
- Samyoun, S., Shubha, S.S., Mondol, M.A.S., Stankovic, J.A.: iWash: A smartwatch handwashing quality assessment and reminder system with real-time feedback in the context of infectious disease. arXiv:2009.10317 (2020). Available: <http://arxiv.org/abs/2009.10317>
- Mondol, M., Stankovic, J.: Harmony: A hand wash monitoring and reminder system using smart watches. *EAI Endorsed Trans. Ambient Syst.* 2, (2015). <https://doi.org/10.4108/eai.22-7-2015.2260042>
- SayedMondol, M.A., Stankovic, J.A.: HAWAD: Hand washing detection using wrist wearable inertial sensors. In: *2020 16th International Conference on Distributed Computing in Sensor Systems (DCOSS)*, pp. 11–18. IEEE, Piscataway (2020). <http://doi.org/10.1109/DCOSS49796.2020.00016>
- Kwapisz, J.R., Weiss, G.M., Moore, S.A.: Activity recognition using cell phone accelerometers. *ACM SIGKDD Explor. Newslett.* 12(2), 74–82 (2011). Available: <https://dl.acm.org/doi/10.1145/1964897.1964918>
- Galluzzi, V., Herman, T., Polgreen, P.: Hand hygiene duration and technique recognition using wrist-worn sensors. In: *Proceedings of the 14th International Conference on Information Processing in Sensor Networks*, pp. 106–117. Association for Computing Machinery, New York (2015). <https://doi.org/10.1145/2737095.2737106>
- Li, H., Chawla, S., Li, R., Jain, S., Abowd, G.D., Starner, T., Zhang, C., Plötz, T.: Wristwash: Towards automatic handwashing assessment using a wrist-worn device. In: *Proceedings of the 2018 ACM International Symposium on Wearable Computers - ISWC '18*, pp. 132–139. ACM Press, Singapore (2018). <https://doi.org/10.1145/3267242.3267247>

22. Wang, C., Sarsenbayeva, Z., Chen, X., Dingler, T., Goncalves, J., Kostakos, V.: Accurate measurement of handwash quality using sensor armbands: Instrument validation study. *JMIR mHealth uHealth* 8(3), e17001 (2020). <https://doi.org/10.2196/17001>
23. Matteson, D.S., James, N.A.: A nonparametric approach for multiple change point analysis of multivariate data. arXiv:1306.4933 (2013). Available: <http://arxiv.org/abs/1306.4933>
24. BVBA, B.: Byteflies: We enable wearable health. <https://www.byteflies.com/>. Accessed 1 March 2019
25. Syarif, I., Prugel-Bennett, A., Wills, G.: SVM parameter optimization using grid search and genetic algorithm to improve classification performance. *TELKOMNIKA (Telecommunication Comput. Electron. Control)* 14(4), 1502 (2016). <https://doi.org/10.12928/telkomnika.v14i4.3956>
26. Kotsiantis, S., Kanellopoulos, D., Pintelas, P.: Handling imbalanced datasets: A review. *GESTS Int. Trans. Comput. Sci. Eng.* 30(1), 25–36 (2006). <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.96.9248&rep=rep1&type=pdf>
27. Yin, P.Y.: *Pattern Recognition*. IntechOpen, 2009. <https://books.google.be/books?id=enSfDwAAQBAJ>
28. Hsu, C.-w., Chang, C.-c., Lin, C.-J.: A practical guide to support vector classification (2016). Available: <https://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf>

How to cite this article: Zhang, Y., Xue, T., Liu, Z., Chen, W., Vanrumste, B.: Detecting hand washing activity among activities of daily living and classification of WHO hand washing techniques using wearable devices and machine learning algorithms. *Healthc. Technol. Lett.* 8:148–158 (2021). <https://doi.org/10.1049/htl2.12018>