# User Estimation with Touch Panel Buttons Toward In-Home Activity Recognition

Kyohei Suda*, Shigemi Ishida*, Hiroshi Inamura*
* Graduate School / School of Systems Information Science, Future University Hakodate, Japan
Email: {g2122037, ish, inamura}@fun.ac.jp

***Abstract*—Smart homes that improve people's lives are attention. To realize a smart home, it is necessary to recognize user activity. Existing studies have collected user activity data by installing many sensors in the home, however, the collected activity data does not include user names, and it is not possible to recognize the activity of multiple users while distinguishing between them. In this study, we propose an in-home activity recognition method that identifies users by performing user estimation using data collected from home appliance operations. In this paper, we propose a user estimation method using button operations among various home appliance operations. We evaluated the user estimation performance of the proposed method using button operation data collected from 8 subjects and found that accuracy was 86.4% or higher when the amount of training data was 10 trials and only features obtained from the pressure were used. This result is comparable to existing studies that require multiple sensors, indicating the feasibility of user estimation with a single sensor.**

***Index Terms*—In-home activity recognition, user estimation, home appliance operation, machine learning.**

## I. Introduction

In recent years, smart homes are being developed to make people's lives and daily activities safer, more secure, and more comfortable by utilizing the latest technologies such as AI and IoT. To realize such a smart home, it is necessary to recognize the user's actions and to accumulate their histories.

As a method to recognize user activities in the home, various "sensor" methods have been reported, such as combined with ultrasonic sensors, electricity meters, and current transformer sensors [1], pressure mats, and float sensors [2]. However, the action information recognized by these methods does not include user information, and it is not possible to recognize the actions of multiple users while distinguishing between them.

In contrast, this research proposes a method of recognizing the actions of multiple users by incorporating "sensors" into home appliance "controllers" such as remote controls and operation panels, while identifying each user. By estimating users based on the "habits" of daily operations of home appliance controllers, we can identify "when" and "who" used the home appliance.

There are various types of controllers used in home appliances. Our survey of home appliances revealed that most home appliance controllers can be classified into 4 types: button switches, rotary switches, slide switches, and motion-reading controllers. If a user can be estimated for each of these types of controllers, it can estimate the user in many home appliance operations.

In [3], a user estimation method using motion-reading controllers have been presented. The results of the evaluation using data from 24 subjects confirmed that users could be estimated with a 93.5% accuracy, indicating the feasibility of user estimation using user operation habits.

In this paper, we examine the possibility of user estimation using button switches, which account for a large proportion of home appliance controllers. Using a touch panel capable of simultaneously collecting press-position and press-pressure data, we collected press-position and press-pressure data from eight subjects when they pressed a button, and evaluated the user estimation performance. As a result, we confirmed that user estimation was possible with an 82.3% accuracy when the amount of training data was 10 trials and only features obtained from the press pressure were used. This is comparable to the estimation results of an existing study [4], which required multiple sensors.

## II. Related Work

### A. In-Home Activity Recognition

A study on in-home activity recognition has reported methods of installing sensors in the home or wearing wearable devices.

Methods for installing sensors in the home have already been proposed. Nakagawa [1] proposed an activity recognition method using ultrasonic sensors, electricity meters, and current transformer sensors, but it requires the installation of more than 20 sensors in total. Kasteren [2] proposes a method using sensors such as pressure mats, float sensors, and reed switches, but at least 14 sensors need to be installed.

Thus, in-home activity recognition, in which sensors are installed, requires the installation of multiple types and many sensors, which is costly and time-consuming to install. Furthermore, it is impossible to identify the user who is active in an environment where two or more people live. Although there have been reports on camera-based activity recognition [5] and identifying users using cameras [6], it is not realistic because cameras are continually filming inside the home, so violating the privacy of the user.

In activity recognition using wearable devices, each user wears a separate wearable device so that activities can be recognized while identifying users. Shahmohammadi [7] and Nandy [8] reported methods to estimate user states such as walking, running, and sitting using a smartwatch. Paraschiakos [9] report methods that use wearable devices to recognize activities in the home, such as washing dishes or vacuuming, but require the user to wear different wearable devices on different parts of the body. Since these methods require the wearable device at all times, they may cause discomfort to users who do not wear wearable devices regularly. In addition, wearable devices are expensive, so having wearable devices for all family members is expensive.

In this study, in-home activity recognition is performed by a user estimation method that installs sensors in-home appliance controllers. Although the proposed method does not recognize a variety of activities, it does not require many sensors to be installed at home or worn and can achieve in-home activity recognition with identified users.

### B. *User Estimation Using Body Movements*

A study in the field of user authentication has been conducted to estimate using data collected from the user's body movements. Gesture authentication is an authentication method that uses body movements, has been studied among user authentication methods. Mare [10] reported a method for user authentication by lifting a smartphone while wearing a wristband. Zhao [11] reported on a method that uses a hand-tracking device called Leap Motion to authenticate users with hand gestures. However, these methods require the user to repeat the same action in more than 20 trials or collect the necessary data for 30 to 40 minutes. Because user authentication studies require high user estimation performance, it is necessary to increase the amount of training data for operations and to use complex actions that are not performed daily for authentication operations.

Some studies have reported user estimation based on simple operations. These studies assume that the requirement of the estimation performance is set low, assuming that it will be used for different situations from authentication. Pohl [4] proposed a method for estimating the user using home appliances using button operations. However, this method requires the installation of multiple sensors, such as distance and pressure sensors, on the home appliance. Although touch-panel-based authentication methods have also been reported [12], multiple-time operation is required for authentication. We aim to estimate users from a single operation.

In this study, user estimation is performed for the purpose of in-home activity recognition, so high estimation performance is not required. User estimation is performed with a button press once, considering that some estimation errors can be permissible. By installing the sensor in the controller of the home appliance, the user is estimated without the need for an additional device.

## III. User Identifier Based on Touch Screen Operation

### A. *Key Idea*

The key idea of this proposal is to estimate users using their daily habits of operating home appliances. It is expected that the strength and duration of pressing a button on a home appliance will differ among users, even if the action is only for a short period. Therefore, we estimate the user by supervised learning using features that differ among users.

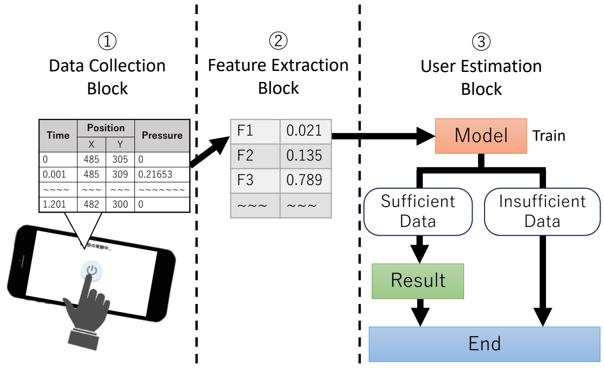


Fig. 1. System overview

### B. *Assumed Environment*

The proposed system is assumed to be used in a home, so it is sufficient to estimate a few users. According to the statistical handbook by the Statistics Bureau of Japan [13], the average size of households was 2.21 in 2020. To cover most households, the number of users in our system should be larger than the average size of households. The maximum number of users in this system is therefore set to 4.

### C. *Sysetm Overview*

Figure 1 is the proposed system overview of user estimation. The user estimation system consists of data collection, feature extraction, and user estimation blocks. When a user presses a touch panel button, the data collection block collects the position and pressure as time series data. The feature extraction block extracts proposed features from the collected data. In the user estimation block, the extracted features are input into the user estimation model and the estimation results are output. Then, the features are used to retrain the model. If sufficient data for estimation has not been collected, the estimation results are not output. In this case, only model retraining is performed.

The label data is necessary for training supervised learning models. However, this paper only confirms the possibility of user estimation and does not consider how to collect label data. In this paper, label data was used in all of the model training and evaluation. After the possibility of user estimation is demonstrated, it is necessary to devise a way to minimize the collection of label data by utilizing supervised and semi-supervised learning.

The following sections describe each block in detail.

### D. *Data Collection Block*

The data collection block collects sensor data when the user presses a button.

The 5 types of data that can be collected from button operation are shown in Table I. If the sensor were to be included in a home appliance controller without changing the feel of operation, the data that can be collected would be limited to the pressure on the button and position. Therefore, this study collects the press position and pressure for user estimation using the touch panel button. Specifically, collect data that the press position is the x and y coordinates on the touch panel, and the press pressure is the pressure sensor value in the touch panel, along with the time.

TABLE I
DATA THAT CAN BE COLLECTED FROM BUTTON OPERATION

| Collection Data | Using sensor |
|---|---|
| Video of button presses | Camera |
| Hand movement | Wearable device |
| Wrist position | Distance sensor [4] |
| Button press position | Touch panel |
| Pressure on the button | Pressure sensor |

### E. Feature Extraction Block

In the feature extraction block, the proposed features are extracted from the data collected in the data collection block. In this paper, we extract the 28 dimensional features shown in Table II for user estimation using touch panel button operations. These features were decided with reference to related studies and Ref. [3]. The effectiveness of these features for user estimation is discussed in Section IV-D.

### F. User Estimation Block

In the user estimation block, the user is estimated as a multi-class classification problem using the features extracted in the feature extraction block. We used a SVM (Support Vector Machine) with a linear kernel referring to Ref. [3], which showed the high estimation performance of the SVM, though, the gesture used in this paper is different from that used in Ref. [3]. The impact of learning algorithms is evaluated in Section IV-G.

## IV. EVALUATION

To evaluate the user estimation performance of the proposed system, we conducted an experiment in which users pressed buttons displayed on smartphones and collected operation data. In this section, explain the details of the data collection system and the experimental environment, after that, we discuss the effective features for user estimation. Next, user estimation performance is evaluated by changing the amount of training data and the location of training data extraction to identify the training data needed for user estimation. Finally, we compare the estimation performance of various machine learning algorithms to identify appropriate algorithms for this operation.

### A. Data Collection System

A data collection system was implemented to collect data during button operations. Figure 2 is the screen transition image of this system. The data collection system is implemented as a Web page, and data is collected by accessing this page from the experimental terminal. Pressure.js[1] was used to implement this Web page. The press position refers to the JavaScript event object, and the press pressure refers to the value provided by Pressure.js. The value provided by

[1]Pressure.js, https://pressurejs.com/

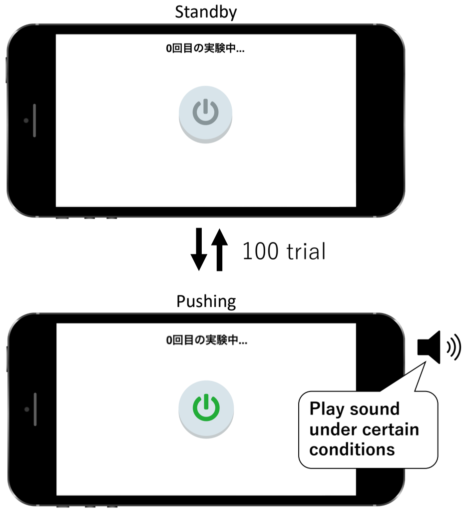


Fig. 2. Screen transition image of this system

Fig. 3. Data collection environment

Pressure.js is a min-max normalization value of the pressure value provided by the OS. Although home appliances have a variety of buttons, almost all of them have a power button, so the screen was created to imitate a power button.

There are “short press” and “long press” button operations for home appliances, and it is ideal to be able to estimate both types of presses. To evaluate the performance difference between the two presses, the data collection system was designed to collect data for the “short press” and “long press” operations.

Feedback was provided on each press so that the user could recognize that the button was pressed. Table III shows the definition of each press and feedback. When the short press data was collected, feedback was given with the once sound effect immediately after the pressure exceeded 0.5. When long press data was collected, the once sound effect was also played, but when the pressure was over 0.5 and held for 1 second, feedback was given that continued to play the sound effect until the finger was released from the button. We used the On-jin[2] wall-switch sound for the sound effect.

### B. Data Collection Environment

The environment for the data collection experiment is shown in Figure 3. An Apple iPhone7 was used as the

[2]Free sound effects On-Jin, https://on-jin.com/

TABLE II
FEATURES

| Dimension | Features | Position | Pressure | Features Group |
|---|---|---|---|---|
| 1 | Operation time length | ◯ | ◯ | Common |
| 1 | Number of data samples | ◯ | ◯ | Common |
| 2 | First $x, y$ coordinates obtained | ◯ | | Position |
| 2 | Width of $x, y$-axis variation of press position trajectory | ◯ | | Position |
| 6 | Mean, median, and standard deviation of the speed and acceleration of the press position trajectory | ◯ | | Position |
| 1 | Total length of the press position trajectory | ◯ | | Position |
| 2 | Time when the pressure value was 1 and the ratio of this time in the total | | ◯ | Pressure |
| 1 | Time is taken to reach the first button press decision threshold | | ◯ | Pressure |
| 1 | Final pressure obtained | | ◯ | Pressure |
| 11 | Frequency of pressure distribution (0.1 separators) | | ◯ | Pressure |

TABLE III
DEFINITION OF EACH PRESS AND FEEDBACK

| Press type | Requirements of press(pressure) | Feedback |
|---|---|---|
| Short | 0.5 or higher | Play sound effect once immediately after exceeding 0.5 |
| Long | 0.5 or higher for 1 second | Play sound effect once immediately after exceeding 0.5, and play the sound effect until the hand is released when the conditionis met. |

experimental terminal, and the data collection system shown in Section IV-A was worked on the terminal to collect data. There were 8 subjects, 7 males and 1 female in their 20s. As shown in the Figure 3, the terminal was fixed almost vertically to the desk, and the subjects were instructed to operate the terminal while sitting on a chair. After the subjects were allowed to practice until they became accustomed to the operation, data was collected for a total of 200 trials (100 short presses and 100 long presses). In this experiment, we assumed a case in which a button is operated only once in the operation of a home appliance, and subjects were instructed to put their hands down each time a button is pressed once in order to collect each trial as an independent button operation. Because the data collected in this experiment may contain personally identifiable information, the experiments were conducted under permission from the ethics committee of Future University Hakodate (permission #2021016).

### C. Evaluation Method

The user estimation performance was evaluated by taking 4 subjects from 8 subject data, performing user estimation, and evaluating accuracy. In other words, user estimation was performed for each of the 70 ( ${}_8\mathrm{C}_4 = 70$ ) possible user combinations, and the average accuracy was calculated.

### D. Effect of Accuracy by Feature Selection

To examine effective features for user estimation, we compared the user estimation performance of the following 3 patterns from the features shown in Table II.

1) Positional Features:
   Features with "◯" in the position column
2) Pressure Features:
   Features with "◯" in the pressure column
3) All Features:
   All features in the table

The evaluation was conducted using all 100 trials of data collected from each subject in a 10-fold cross-validation.

Table IV shows the accuracy for each pattern. When pressure features were used, the accuracy of 92.5% or higher. Looking at each feature type shows that the "pressure feature" and "all features" have about the same accuracy, whereas the "position feature" has about 8% lower accuracy. Using position and pressure together improves performance by about 1%, but not significantly. In a related study [4] using multiple sensors, accuracy was about 90% for 4 subject estimations. Although a simple performance comparison is difficult because the experimental environment and the amount of training data differ from related studies [4], the estimation performance in this evaluation is about equal for all press types. Based on these results, the collection of press positions is considered unnecessary for estimation.

TABLE IV
ACCURACY FOR EACH FEATURE[%]

| | | Features(Dimension) | | |
|---|---|---|---|---|
| | | Position (13) | Pressure (17) | All (28) |
| Press type | Short | 84.1 | 92.5 | 92.6 |
| | Long | 85.9 | 93.1 | 94.6 |

To corroborate that the collection of only pressure is sufficient, the contribution of each feature to estimation was additionally evaluated. To determine the extent to which each feature contributed the estimation, training with Random Forest using "all features" and the feature importance was output. Feature importance was calculated for all features, and the total feature importance of the 3 feature groups shown in Table II was compared. Details of the 3 groups are shown below.

1) Position Feature Group:
   Features with "◯" in the position column
2) Pressure Feature Group:
   Features with "◯" in the pressure column
3) Common Feature Group:
   Features with "◯" in both the position and pressure columns

The data used in the evaluation are the same as those used in the accuracy evaluation.

The results are shown in Table V. The total feature importance is about 0.6 for the "pressure feature group" while it is about 0.15 for the "position feature group". This indicates that the features in the "position feature group" do not contribute much to the estimation. Furthermore, the highest feature importance in the "position feature group" is 0.0333, which is about the same value as the fifth lowest in the "pressure feature group", so it is difficult to indicate that it contributes highly to estimation. The total feature importance of the "position feature group" which has 11 dimensions, is the same or lower than that of the "common feature group" which has only 2 dimensions. In the results for the accuracy

TABLE V
FEATURE IMPORTANCE

| | | Feature Group | | |
|---|---|---|---|---|
| | | Position | Pressure | Common |
| Press type | Short | 0.1367 | 0.5968 | 0.2666 |
| | Long | 0.1786 | 0.6429 | 0.1785 |

and feature importance, it is clear that there is almost no advantage to collecting the position. Furthermore, there was almost no performance difference in press type. In a real environment, only pressure should be collected for estimation in button operation, because it is desirable to have fewer types of sensors for user estimation.

In subsequent evaluations, only the features of the "pressure feature group" will be used.

### *E. Effect of the Amount of Training Data on the Accuracy*

If the amount of training data could be reduced, the effort of collecting training data in a real environment would be reduced. Therefore, user estimation performance was evaluated when the amount of training data was varied. The evaluation data was taken from the 91st to 100th trial of each subject, and the extracted training data was varied from the 1st to the 90th trial. The training data were retrieved in order from the back of the trial. For example, in the case of 10 trials, the data from the 81st to 90th trial was extracted, and in the case of 20 trials, the data from the 71st to 90th trial was extracted.

Figure 4 shows accuracy based on the amount of training data. When 90 trials were used for training, accuracy was 88.4%, whereas when only 10 trials were used, accuracy was 82.3%. This result indicates that there was almost no difference in estimation performance based on the amount of training data. In a related study [4] that required multiple sensors, the estimation performance was about the same when only 10 trials were used for training since accuracy was about 85%. Using 10 trials for training equals pressing a button 10 times in a real environment. Even for home appliances such as coffee makers, which are used only once a day, user estimation can be achieved by collecting data for 10 days. Considering the convenience of the system, it is desirable to collect less training data, so the data to be collected for training is 10 trials.

In subsequent evaluations, the amount of training data was set to 10 trials.

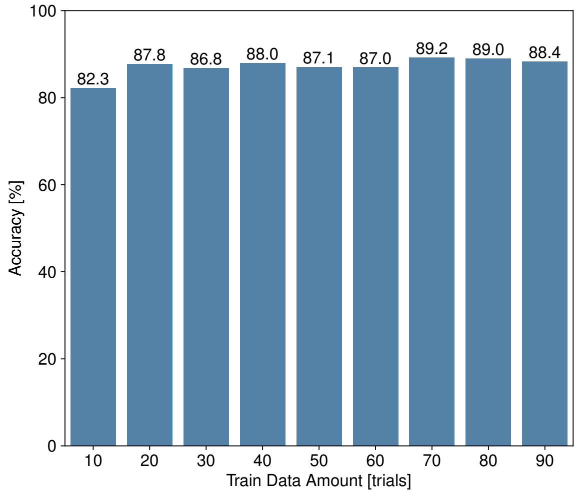


Fig. 4. Accuracy based on the amount of training data

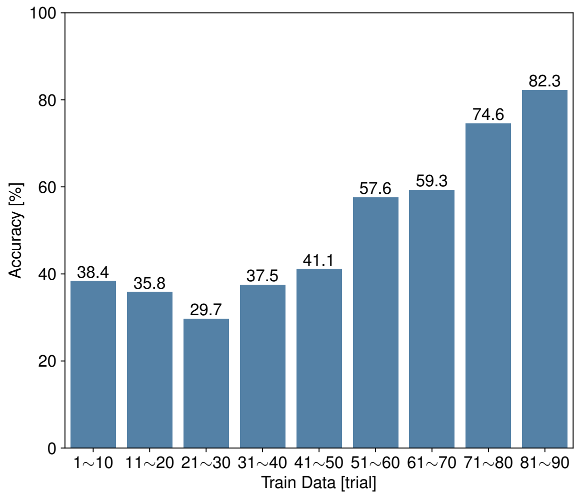


Fig. 5. Accuracy based on the location of the training data retrieval

### *F. Effect of Familiarity with Operation*

Since the collected data in this paper is based on a continuous collection of button operations, it can be possible that familiarity with the operation will affect user estimation. Therefore, we evaluated how the extraction position of the training data affects accuracy. The evaluation data were from the 91st to 100th trials for each subject, and the training data were retrieved by shifting the retrieval position by 10 trials out of the 1st to 90th trials.

Figure 5 shows the accuracy based on the location of the training data retrieval. Accuracy was 82.3% when the data was learned in the 81-90 trials immediately before the evaluation data, while accuracy was 38.4% when the data was learned in the 1-10 trials farthest from the evaluation data. The lowest accuracy was 21 to 30 trials, which was considered due to the lack of practice by each subject. Most of the subjects practiced only 1 to 3 times, and even the most practiced about 10 times. We consider that about 20 trials were necessary for the subjects to understand the operation method and stabilize the operation. The data from the 21st to the 90th trial shows a gradual increase in accuracy, indicating that familiarity with the operation has a significant impact on accuracy, even when the data was collected over a short period. In other words, it was necessary to keep collecting new training data, indicating that data collected on different days may increase the difficulty of estimation.

### *G. Estimation Performance by Learning Algorithm*

We used the SVM (Linear Kernel) in our proposed system referring to Ref. [3]. However, since the operations to be recognized in this paper are different from Ref. [3], a different learning algorithm may provide higher performance. Therefore, we studied the appropriate learning algorithm by training with various learning algorithms corresponding to multi-class classification problems and evaluating the performance of each. The algorithms compared were SVM (Linear Kernel), Random Forest, logistic regression, k-nearest neighbor method (with $k = 1$), Gaussian naive Bayes, stochastic gradient descent, decision tree, and LightGBM. In SVM, logistic regression, k-nearest neighbor method, and stochastic gradient descent method, the features used for training are standardized.

Table VI shows the accuracy when each learning algorithm was used. The results show that the highest accuracy was obtained when Random Forest was used. We guess that this is due to the higher dimensionality of the features used in this paper, which resulted in higher accuracy than SVM. The other learning algorithms were overlearning when using the data collected in this experiment, resulting in a lower accuracy compared to Random Forest. Therefore, Random Forest was shown to be appropriate for user estimation by button press operations.

TABLE VI
ACCURACY BASED ON THE LEARNING ALGORITHM

| Learning algorithm | Accuracy[%] (Short press) | Accuracy[%] (Long press) |
|---|---|---|
| Random Forest | 89.8 | 86.4 |
| SVM (Linear Kernel) | 84.7 | 83.6 |
| Logistic Regression | 84.0 | 84.2 |
| k-nearest neighbor ($k = 1$) | 81.4 | 80.4 |
| Gaussian Naive Bayes | 81.4 | 78.9 |
| Stochastic Gradient Descent | 78.8 | 78.9 |
| Decision Tree | 73.2 | 70.7 |
| LightGBM | 53.2 | 57.7 |

## V. Conclusion

To realize in-home activity recognition identified user information, we proposed a method to recognize multiple user activities with user identification by installing sensors into controllers installed in home appliances. Among various controller operations, this paper proposes a user estimation method using button switches and evaluated its estimation performance.

As a result of the evaluation using a touch panel, it was confirmed that the system was able to estimate the user with an 82.3% accuracy when the amount of training data was 10 trials and only features obtained from the pressure were used. Furthermore, selecting an appropriate learning algorithm improved the accuracy by about 3%. This is about the same accuracy as existing studies [4] that require multiple sensors, so this shows the feasibility of user estimation by using only 1 type of sensor.

This paper shows the feasibility of a user estimation method for touch panel button operations among the controllers installed in home appliances. However, the data used for the evaluation was collected in a sitting, but in a real environment, it has home appliance operation in standing. Furthermore, despite the continuously collected data, the estimation was negatively affected by familiarity with the operation, and the estimation performance may be further degraded with data collected another day. In the future, it is necessary to investigate the impact on estimation performance by evaluating the data collected over a long period in a real environment.

## Acknowledgment

This work was supported in part by JSPS KAKENHI Grant Number JP21K11847 as well as the Cooperative Research Project of RIEC, Tohoku University.

## References

[1] E. Nakagawa, K. Moriya, H. Suwa *et al.*, "Toward real-time in-home activity recognition using indoor positioning sensor and power meters," *IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, pp. 539–544, 2017.

[2] T. L. M. van Kasteren, G. Englebienne, and B. J. A. Kröse, "An activity monitoring system for elderly care using generative and discriminative models," *Personal and Ubiquitous Computing*, vol. 14, no. 6, pp. 489–498, 2010.

[3] K. Suda, S. Ishida, and H. Inamura, "User identification based on mouse operation toward automatic home appliance configuration," *IEEE GCCE*, pp. 714–715, 2022.

[4] H. Pohl, M. Krause, and M. Rohs, "One-button recognizer: exploiting button pressing behavior for user differentiation," *ACM UbiComp*, pp. 403–407, 2015.

[5] L. Fiore, D. Fehr, R. Bodor *et al.*, "Multi-Camera Human Activity Monitoring," *J. Intelligent and Robotic Systems*, vol. 52, no. 1, pp. 5–43, 2008.

[6] H. Pirsiavash and D. Ramanan, "Detecting activities of daily living in first-person camera views," *IEEE CVPR*, pp. 2847–2854, 2012.

[7] F. Shahmohammadi, A. Hosseini, C. E. King *et al.*, "Smartwatch Based Activity Recognition Using Active Learning," *IEEE/ACM CHASE*, pp. 321–329, 2017.

[8] A. Nandy, J. Saha, C. Chowdhury *et al.*, "Detailed Human Activity Recognition using Wearable Sensor and Smartphones," *Int. Conf. Opto-Electronics and Applied Optics*, pp. 1–6, 2019.

[9] S. Paraschiakos, R. Cachucho, M. Moed *et al.*, "Activity recognition using wearable sensors for tracking the elderly," *User Modeling and User-Adapted Interaction*, vol. 30, no. 3, pp. 567–605, 2020.

[10] S. Mare, R. Rawassizadeh, R. Peterson *et al.*, "Continuous Smartphone Authentication using Wristbands," *Workshop on Usable Security*. Internet Society, 2019.

[11] J. Zhao and J. Tanaka, "Hand Gesture Authentication Using Depth Camera," in *Advances in Information and Communication Networks*, ser. Advances in Intelligent Systems and Computing, K. Arai, S. Kapoor, and R. Bhatia, Eds. Springer, pp. 641–654, 2019.

[12] Y. Song and I. Oakley, "PushPIN: A Pressure-Based Behavioral Biometric Authentication System for Smartwatches," *Int. J. Human–Computer Interaction*, vol. 39, no. 4, pp. 893–909, 2023.

[13] "Statistics Bureau Home Page/Statistical Handbook of Japan 2022." https://www.stat.go.jp/english/data/handbook/c0117.html