Communication

Machine Learning-Based Classification of Human Behaviors and Falls in Restroom via Dual Doppler Radar Measurements [†]

Kenshi Saho [1,2,*,‡] [ID], Sora Hayashi [2,‡], Mutsuki Tsuyama [2,‡], Lin Meng [2] [ID] and Masao Masugi [2]

[1] Department of Intelligent Robotics, Toyama Prefectural University, Imizu 939-0398, Japan

[2] Department of Electronic and Computer Engineering, Ritsumeikan University, Kusatsu 525-8577, Japan; sora8840817@gmail.com (S.H.); ri0082ip@ed.ritsumei.ac.jp (M.T.); menglin@fc.ritsumei.ac.jp (L.M.); masugi@fc.ritsumei.ac.jp (M.M.)

[*] Correspondence: saho@pu-toyama.ac.jp

[†] This paper is an extended version of our paper published in Tsuyama, M.; Hayashi, S.; Saho, K.; Masugi, M. A Convolutional Neural Network Approach to Classification of Human's Behaviors in a Restroom Using Doppler Radars. Proc. of ATAIT 2021 (2021 International Symposium on Advanced Technologies and Applications in the Internet of Things), Kusatsu, Japan, 23–24 August 2021; Paper No. 6.

[‡] These authors contributed equally to this work.

Abstract: This study presents a radar-based remote measurement system for classification of human behaviors and falls in restrooms without privacy invasion. Our system uses a dual Doppler radar mounted onto a restroom ceiling and wall. Machine learning methods, including the convolutional neural network (CNN), long short-term memory, support vector machine, and random forest methods, are applied to the Doppler radar data to verify the model's efficiency and features. Experimental results from 21 participants demonstrated the accurate classification of eight realistic behaviors, including falling. Using the Doppler spectrograms (time–velocity distribution) as the inputs, CNN showed the best results with an overall classification accuracy of 95.6% and 100% fall classification accuracy. We confirmed that these accuracies were better than those achieved by conventional restroom monitoring techniques using thermal sensors and radars. Furthermore, the comparison results of various machine learning methods and cases using each radar's data show that the higher-order derivative parameters of acceleration and jerk, and the motion information in the horizontal direction are the efficient features for behavior classification in a restroom. These findings indicate that daily restroom monitoring using the proposed radar system accurately recognizes human behaviors and allows early detection of fall accidents.

Keywords: Doppler radar application; restroom; fall detection; human behavior classification; remote monitoring

1. Introduction

Aging is a global phenomenon responsible for various problems related to shortened health expectancy and the sudden death of elderly people [1]. Therefore, monitoring systems for early detection of accidents and abnormal behaviors in elderly adults, such as falling, have recently been developed based on sensors and Internet-of-Things technologies [2]. However, such systems are not used inside restrooms due to privacy concerns. The early detection of falls and abnormal behaviors in restrooms is important because it is one of the dangerous spaces in the home where elderly people are likely to fall [3]. Even though accelerometry-based approaches for fall detection in restrooms have been proposed [4,5], they require the subjects to wear the sensor devices.

Few studies have investigated camera-based approaches for the remote monitoring of restrooms [6] based on the image-processing-based fall detection technique [7,8]. However, their measurement accuracy depends on lighting conditions and the subjects' clothing. In addition, installing cameras in restrooms is challenging because of privacy issues. Hence,

investigations of restroom monitoring have been limited. Even though infrared, thermal, sensor-based posture estimation in restrooms has been proposed [9,10], its measurement accuracy depends on the temperature of the toilet and of the subject's clothing.

Radar technology is a promising candidate method for solving the above-described problems because it does not invade privacy during measurement and is not affected by temperature conditions or clothing. Therefore, radar-based motion recognition and fall detection are active research areas. Various approaches based on machine learning of radar images, such as time–Doppler, time–range, or range–Doppler images, have been proposed and demonstrated in realistic situations [11–13]. However, research on the use of radar in restrooms is quite limited. Though the detection of falls in restrooms using radar systems or the classification of normal and abnormal behaviors by range (distance) information [14,15] have been studied based on classical signal detection or discriminant analysis methods, it is difficult to determine what is happening in a restroom because only abnormal behaviors, such as falls, are detected. Moreover, their accuracy and practicality were both insufficient. To solve these problems, our recent paper reported the long short-term memory (LSTM)-based classification of eight types of behaviors and falls in a restroom with approximately 80% accuracy using the Doppler radar-measured velocity time series [16].

As a significant extension of our previous study [16], this study presents a more accurate classification using the convolutional neural network (CNN)-based approach and investigates the accurate classification of human behavior in restrooms via various machine-learning methods. The contributions of this study are as follows:

- The efficient implementation of Doppler radars and experimental examples for privacy-protected restroom monitoring were provided for the realistic environment; this is a significant contribution because there are only several limited reports on realistic restroom monitoring.
- Efficient classification models and types of their input data for the radar-based restroom-monitoring system were clarified via the thorough comparison of the radar-based motion recognition approaches.
- The classification of human behaviors and falls in a restroom with above 95% was demonstrated. This result shows significant improvement over other conventional studies on radar-based restroom monitoring [14–16].

This paper is an extended version of our conference paper [17] that simply presented the results for the CNN-based approach. In this study, we added the results for the comparison with other various machine learning-based classification methods, the details of the implementation of the classification methods, comparison with other conventional studies, and investigation to elucidate the efficient features.

2. Related Work

The machine learning-based human motion classification has been widely investigated in various sensing technologies such as cameras, depth sensors, and accelerometers [18–21]. For various applications, the practicality and versatility of various machine-learning models [22–24] have been demonstrated.

For the field of radar technology, the machine learning methods that have been established for the abovementioned studies have also been applied to human motion recognition using the Doppler and/or range information obtained via radar sensing [11,25,26]. For example, radar-based fall detection has been widely studied [27,28] and various efficient methods using machine learning techniques have been proposed [29,30]. In recent years, accurate fall detections using radars have been achieved with machine learning methods such as CNN [12,31–34], LSTM [13,34], random forest (RF) [35], and support vector machine (SVM) [36]. These classification techniques are properly selected based on the features of the problem and the objectives of the radar data analysis. The CNN-based classification technique has achieved the accurate classification of human motion using Doppler radar spectrograms [11,12,35,37]. Additionally, the LSTM-based technique has achieved better accuracy for the continuous classification problem [13,37,38]. Although these techniques

can achieve relatively accurate classifications, the mechanisms and/or factors of the classifications are generally unclear. Thus, the classical motion parameter-based approaches are still important in contemporary radar technology for developing the motion recognition system whose mechanism and performance are guaranteed [29,37,39,40]. The methodology for the design of the radar-based motion recognition system using these various machine learning approaches is being established for various types of experimental data.

However, as stated in the Introduction, most studies on motion recognition and fall detection studies using radar have not considered their application to the restroom. Thus, there are only several studies on radar-based restroom monitoring [14–16]. For example, in [15], the classification accuracy of seven types of behaviors and falls was approximately 60%. The low accuracy of conventional radar systems may be due either to the use of range information or to the use of classical detection/classification methods. Thus, the efficient machine learning methods that are suitable for radar-based restroom monitoring and their appropriate input data have not been investigated at all. For this purpose, our previous study [16] achieved accurate classification of eight human behaviors and falls in the restroom. This study aimed to extend this previous study with respect to the classification accuracy and to clarify the efficient models and features for behavior classification in radar-based restroom monitoring.

3. Experiments for Dataset Generation

3.1. Doppler Radar Experiments

Figure 1 shows the experimental site and outline of the measurement system. Twenty-one healthy young men (age: 22.4 ± 1.1 years, height: 173.8 ± 5.1 cm) consented to participate in this study and were instructed regarding the testing procedures prior to the experiments. Informed consent was obtained from all participants. Each participant performed the following eight types of behaviors three times: (a) opening the toilet lid, (b) pulling down the pants, (c) sitting, (d) taking the toilet paper, (e) standing, (f) pulling up the pants, (g) closing the toilet lid, and (h) falling. Falling is defined as the motion of falling forward from a seated position, which is one of the realistic falling motions in restrooms [6]. For example, a person seated on the toilet falls when leaning forward slowly.

(a)   (b)
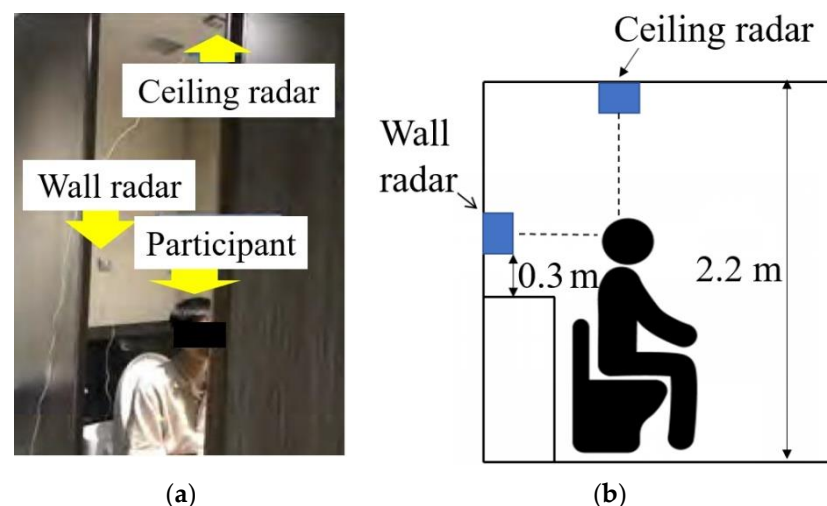
Figure 1. Doppler radar sensing system for measuring behaviors in a restroom. (**a**) Experimental site and (**b**) measurement setup.

We used 24 GHz continuous-wave radars (ILT office, BSS-110) with ±14° plane directivity mounted as shown in Figure 1. The radars were installed above (ceiling radar) and behind (wall radar) the participant. The ceiling and wall radars measured the motion along the vertical and horizontal directions, respectively. The Doppler radar transmitted a 24 GHz sinusoidal wave with an effective isotropic radiated power of 40 mW to each

participant. The demodulated in-phase/quadrature radar signal $s(t)$ was obtained using a quadrature detector and an analog-to-digital converter with a sampling frequency of 600 Hz, and a measurement velocity range of -1.875–1.875 m/s.

3.2. Generation of Spectrogram Dataset

Similar to our previous study [41], the short-time Fourier transforms (STFT) of the received signals were calculated to generate the spectrogram images (time–velocity–power distribution images) as follows. First, we removed the zero-Doppler frequency components from the received signals using a one-dimensional Butterworth high pass filter with a cutoff frequency of 30 Hz to eliminate echoes from static objects such as walls and toilet seats. The STFT of $s(t)$ was calculated as $S(t, f_d) = \int s(\tau)w(\tau - t)\exp(-j2\pi f_d \tau)d\tau$, where t is the time, f_d is the Doppler frequency, and $w(t)$ is the window function. For $w(t)$, the Hamming window function, with a length of 213.3 ms (corresponding to the frequency resolution $\Delta f_d = 1/(213.3 \times 10^{-3}) = 4.69$ Hz) and overlap length of 211.6 ms, was empirically used for the STFT process. Doppler velocity v_d was calculated with $v_d = cf_d/(2f_0)$, where f_0 is the frequency of the transmitting signal (24.0 GHz) and c is the speed of light (2.999×10^8 m/s). Using this equation, we obtained the spectrogram $|S(t, v_d)|^2$ (based on the above setting, the resolution of v_d in the spectrogram was $\Delta v_d = c\Delta f_d/(2f_0) = 0.0293$ m/s). Finally, we removed the components with a received power density of less than 0 dB/Hz, assuming that these components corresponded to random noises.

Figures 2 and 3 show examples of the generated spectrograms for all behaviors of the ceiling and wall radars, respectively. The motion toward the ceiling and back of the subject is positive for the ceiling and wall radars that measure the head and torso, respectively. For example, we confirmed the significant negative velocity components in Figures 2f and 3f that correspond to the fall forward motion for behavior (f). Similarly, the motion characteristics of each of the other behaviors could also be confirmed. For behaviors (d) and (f), no characteristic velocity components were obtained because they did not involve large motions compared with the other behaviors. Although the different features of each behavioral spectrogram were confirmed to some extent, some behaviors were difficult to classify. For example, the spectrograms of the behaviors (b) and (e) in Figure 2 have relatively similar characteristics. Therefore, this study aimed to classify the behaviors using various machine learning methods.
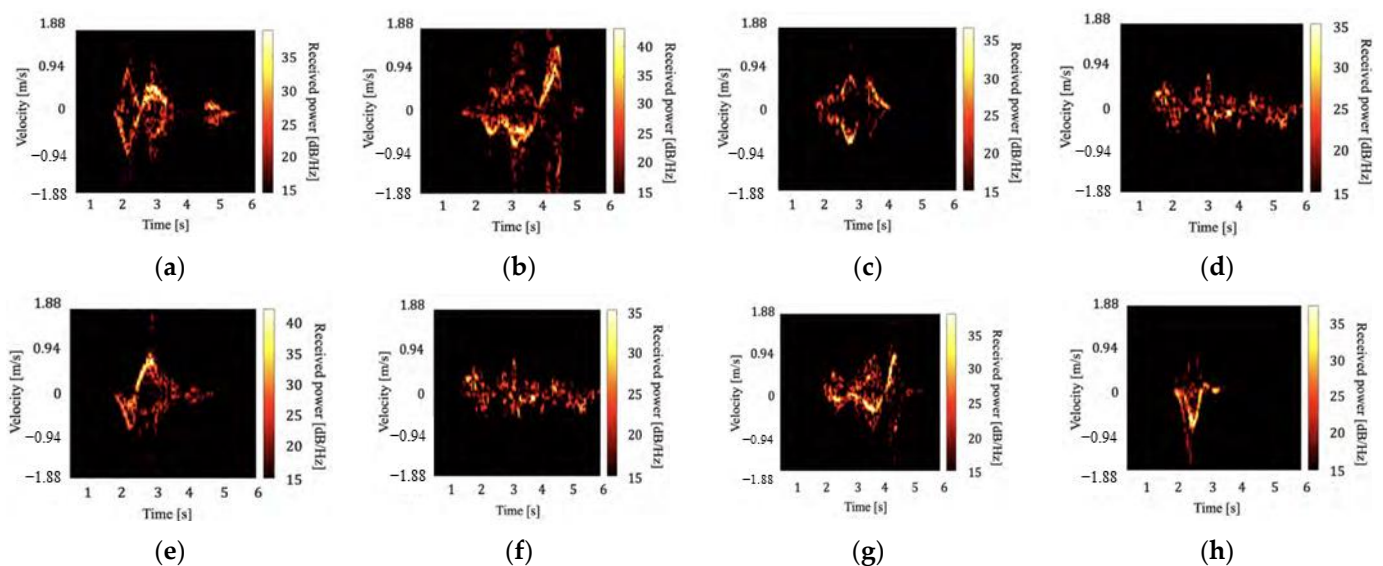
Figure 2. Examples of spectrograms measured with the ceiling radar. (**a**) Opening the toilet lid, (**b**) pulling down the pants, (**c**) sitting, (**d**) taking the toilet paper, (**e**) standing, (**f**) pulling up the pants, (**g**) closing the toilet lid, and (**h**) falling.
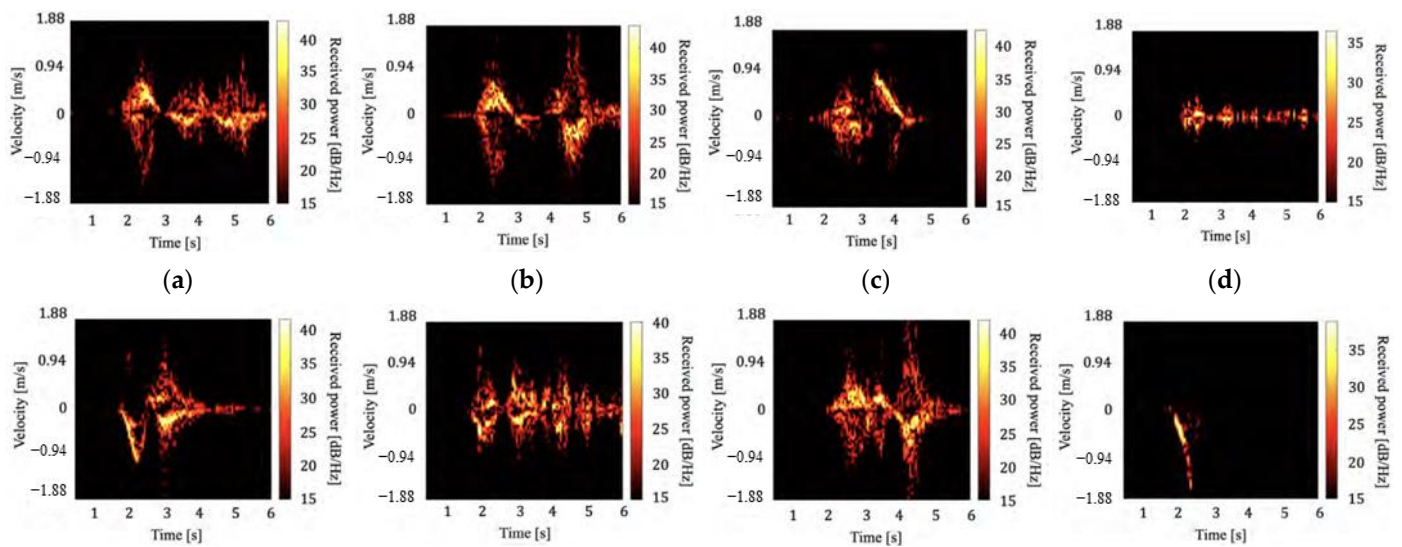
Figure 3. Examples of spectrograms measured with the wall radar. (**a**) Opening the toilet lid, (**b**) pulling down the pants, (**c**) sitting, (**d**) taking the toilet paper, (**e**) standing, (**f**) pulling up the pants, (**g**) closing the toilet lid, and (**h**) falling.

4. Implementation of the Machine Learning-Based Classification Methods

This study implemented three types of classification methods and compared their accuracy to determine the most efficient method and investigate the features for efficiently classifying behaviors and falls in a restroom.

4.1. Spectrogram Image-Based Method Using CNN

This method (CNN method) directly uses the generated spectrogram images as the CNN input. Figure 4 shows the process and structure of the network. The spectrogram PNG images of size 168×218 generated from the two radars were input into the CNN. CNN has a structure similar to AlexNet [22]. However, to avoid overfitting, we used the batch normalization layer instead of the dropout layer. To fuse the two images obtained from the dual radars, we constructed two similar CNNs and combined their outputs using the concatenate layer that was then used in the fully-connected layer to determine the output class. A stochastic gradient descent with a momentum optimization algorithm was used for the network optimization. We trained for 100 epochs with a learning rate of 0.01 and batch size of 8. In addition, to compare the classification accuracy with the single and dual radars, the input from a single spectrogram image was fed into the CNN structure without the concatenate layer.

This classification method using spectrogram images and CNN is known as the most efficient method for various applications of radar motion classification [11]. Although the efficient input data and CNN structure depend on the applications, many studies demonstrated the best accuracy with the CNN-based method compared to the use of other classifiers. However, existing studies fail to explain the classification mechanism or to determine the efficient features. Therefore, this study compared the proposed CNN method with other methods to determine the efficient features and reasons for classification.

4.2. Spectrogram Envelope-Based Method Using LSTM

The LSTM method uses the velocity time-series spectrogram envelopes, as shown in Figure 5, for classification [16]. We extracted three types of envelopes from the spectrograms with the same process as in [41]: the upper envelope $v_{\mathrm{u}}(t)$, lower envelope $v_{\mathrm{l}}(t)$, and power-weighted mean velocity $v_{\mathrm{m}}(t)$. These extracted envelopes were then input into the LSTM as outlined in Figure 6. The data length of each envelope was 102 points, and the dimensions of the input data for the single and dual radar fusion were 102×3 and 102×6, respectively. We empirically optimized the hyperparameters using an Adam optimizer [23]. Thus, the

number of hidden layers was 100, batch size was 64, learning rate was 0.001, and number of epochs for the training was 300.
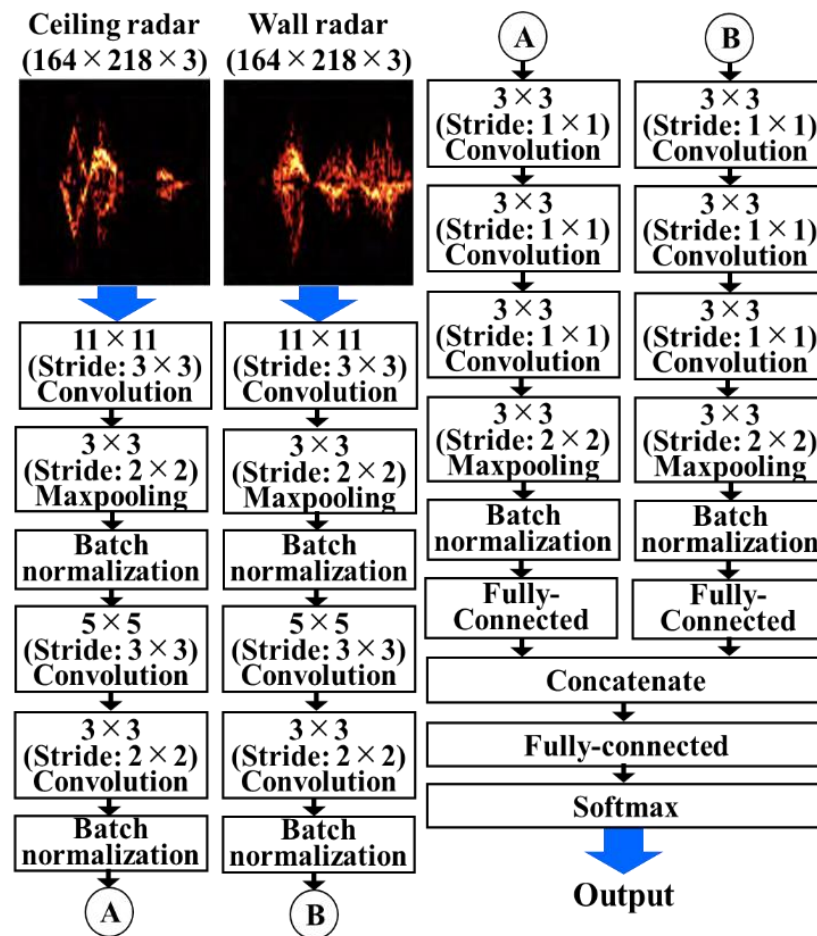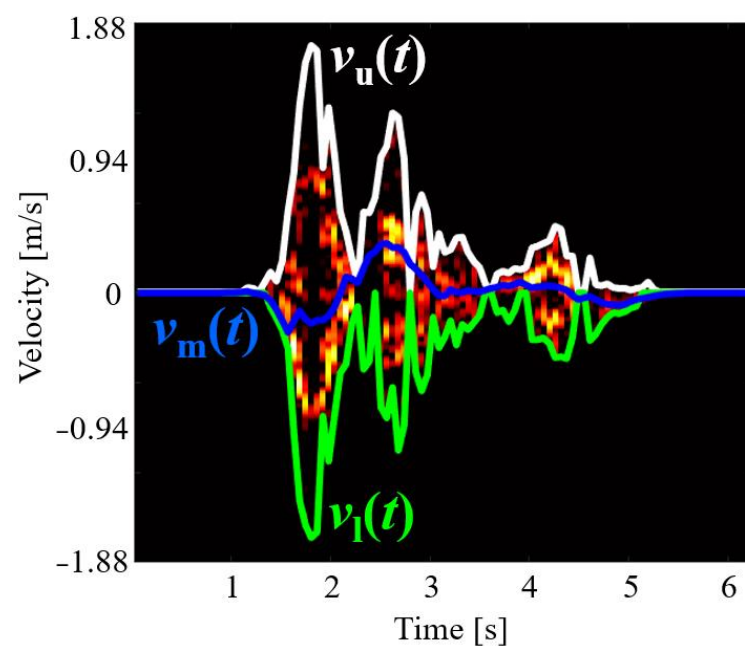

Figure 4. Process and structure of the CNN method.

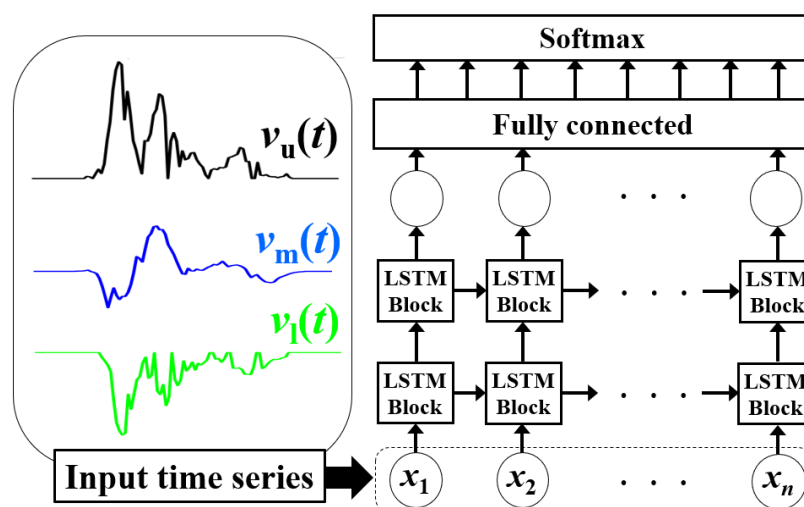**Figure 5.** Example of extraction of spectrogram envelopes.

Figure 6. Process and structure of the LSTM method.

4.3. Motion Parameter-Based Methods

These methods use kinematic parameters extracted using the spectrogram envelopes for classification and can directly obtain efficient motion-feature parameters for behavior classification, as shown in Figure 7. First, we extracted the three envelopes $v_u(t)$, $v_l(t)$, and $v_m(t)$, as with as the LSTM method. We calculated four representative values of mean, maximum, minimum, and standard deviation with respect to time for each envelope. Then, we calculated the time derivative of the envelopes to obtain the acceleration and jerk time series. For example, for $v_m(t)$, we calculated the acceleration time series $a_m(t) = \mathrm{d}v_m(t)/\mathrm{d}t$ and jerk time series $j_m(t) = \mathrm{d}a_m(t)/\mathrm{d}t$. Empirically designed, moving-average low pass filters with an average length of 0.15 s were used to remove small errors in each time series. Similar to the velocity time series, we also calculated the four representative values of $a_m(t)$ and $j_m(t)$. The same process was also applied to $v_u(t)$ and $v_l(t)$. Thus, we extracted 4 (parameters) × 3 (envelopes) × 3 (time series of velocity, acceleration, and jerk) = 36 parameters for each radar. For dual radar fusion, 36 × 2 = 72 parameters were obtained as candidate feature parameters for classification. From these parameters, efficient feature parameters for each classifier were automatically selected using the filter method [24]. The filter method determines the relevance of each parameter for classification, and the top 20% is selected. We selected the widely used random forest (RF) and support vector machine (SVM) as the classifiers in this study. Their hyperparameters were optimized using a grid search, and the Gaussian kernel was used for the SVM.
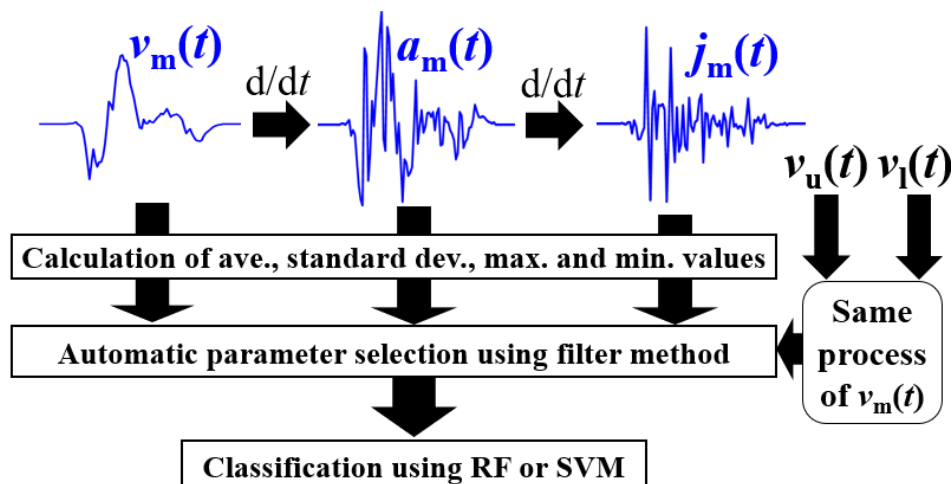

Figure 7. Outline of the RF and SVM methods using the motion parameters.

5. Evaluation and Discussion

5.1. Main Evaluation Results

We evaluated and compared the classification accuracy of the four classification methods in Section III using hold-out validation. For all of the methods, we also compared the accuracy for three cases: only ceiling radar data, only wall radar data, and fused data from the two radars. The classification model was first trained using 80% of the data for each case and then tested with the remaining 20%. We performed 30 trials of the test processes by randomly varying the training data.

Table 1 summarizes the mean and standard deviation of classification accuracies from the 30 test trials for the four classification methods. The CNN method achieved the best accuracy of 95.6%, indicating that the spectrogram images were more effective than the spectrogram envelope [16] or motion parameter-based approach for human behavior and fall classification in restrooms. However, other classification methods also achieved moderate accuracy, making it possible to obtain efficient motion information and/or parameters necessary for classification. This is discussed in the next subsection.

Table 1. Summary of Classification Results.

Method	Ceiling Radar Data	Wall Radar Data	Both Radars
RF	$41.5 \pm 4.82\%$	$55.2 \pm 3.80\%$	$63.8 \pm 3.72\%$
SVM	$60.4 \pm 5.37\%$	$62.4 \pm 4.31\%$	$63.4 \pm 4.27\%$
LSTM [16]	$72.3 \pm 4.96\%$	$82.6 \pm 4.24\%$	$83.2 \pm 3.93\%$
CNN	$90.3 \pm 2.66\%$	$91.5 \pm 3.07\%$	$95.6 \pm 2.28\%$

Furthermore, better accuracy was obtained when using dual radar data than by using the ceiling or wall radar data only. In particular, a significant improvement was obtained for the CNN and RF methods when dual radar data were used. Therefore, we conclude that the motions in both upward and horizontal directions included the differences in the assumed behaviors and falls.

The results from the CNN method based on the convergence curve are shown in Figure 8, and the confusion matrix is further discussed to validate its performance. No overfitting was observed in either test or training processes. The accuracy in the test process converged in less than 50 epochs. Table 2 shows the confusion matrices for the data from the ceiling, wall, and dual radars. The classification accuracies of "(f) pulling up the pants" and "(b) pulling down the pants" are worse for the ceiling and wall radar data, respectively. However, the classification accuracy of (f) improves when the fused data are considered whereas that of (b) is not improved. The classification accuracy of "(h) falling" is 100% in all cases, and is the most important function for the practical use of fall detection.
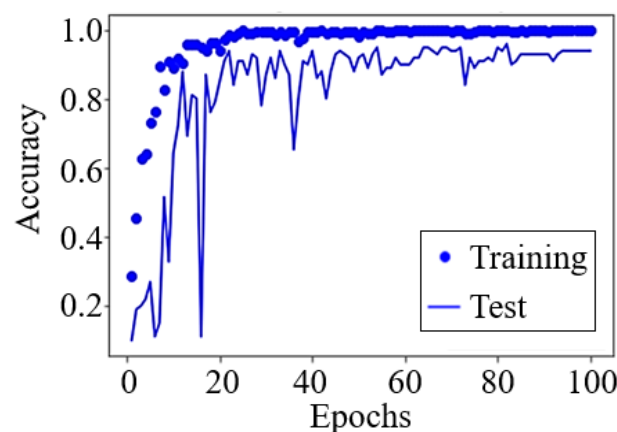
Figure 8. Sample learning curve of the CNN method using the dual radar data.

Table 2. Confusion matrix of the CNN method.

		Predicted Label							
		(a)	**(b)**	**(c)**	**(d)**	**(e)**	**(f)**	**(g)**	**(h)**
	(a)	0.90/ 0.93/ 1	0/ 0/ 0	0/ 0/ 0	0/ 0/ 0	0/ 0/ 0	0/ 0/ 0	0.10/ 0.07/ 0	0/ 0/ 0
	(b)	0/ 0/ 0	0.85/ 0.77/ 0.79	0/ 0/ 0	0.15/ 0/ 0	0/ 0/ 0	0/ 0/ 0	0/ 0.23/ 0.21	0/ 0/ 0
	(c)	0.08/ 0/ 0	0/ 0/ 0	0.92/ 1/ 1	0/ 0/ 0	0/ 0/ 0	0/ 0/ 0	0/ 0/ 0	0/ 0/ 0
True Label	(d)	0/ 0/ 0	0/ 0/ 0	0/ 0/ 0	1/ 0.9/ 1	0/ 0.1/ 0	0/ 0/ 0	0/ 0/ 0	0/ 0/ 0
	(e)	0/ 0/ 0	0/ 0/ 0	0/ 0/ 0	0/ 0/ 0	0.92/ 1/ 1	0.08/ 0/ 0	0/ 0/ 0	0/ 0/ 0
	(f)	0/ 0/ 0.07	0/ 0/ 0	0/ 0/ 0.07	0.22/ 0/ 0	0.06/ 0/ 0	0.72/ 0.92/ 0.86	0/ 0.08/ 0	0/ 0/ 0
	(g)	0.08/ 0/ 0.08	0/ 0/ 0	0/ 0/ 0	0/ 0/ 0	0/ 0/ 0	0/ 0/ 0	0.92/ 1/ 0.92	0/ 0/ 0
	(h)	0/ 0/ 0	0/ 0/ 0	0/ 0/ 0	0/ 0/ 0	0/ 0/ 0	0/ 0/ 0	0/ 0/ 0	1/ 1/ 1

Each cell represents the results for ceiling/wall/dual radars.

5.2. Discussion on Efficient Features

This section discusses the efficient features measured with each radar when classifying human behaviors in restrooms. First, we discuss the effectiveness of the data from each radar and the fused data. Tables 3–5 show the confusion matrices for the LSTM, RF, and SVM methods, respectively. As indicated in the confusion matrices of the CNN (Table 2) and LSTM methods, all behaviors and falls are accurately classified using deep learning methods. However, we can see the different classification accuracies for some classes. For example, as indicated in Table 2, the classification accuracy of the classification of behaviors (b) and (g) was worse in the results of the CNN method with the ceiling radar. In contrast, these were accurately classified with the LSTM method with the ceiling radar data as shown in Table 3. These results indicate that some of the behaviors accurately classified by these methods varied because of the differences in the included features in the spectrogram images and envelopes. In addition, for the motion parameter-based methods (the RF and SVM methods), behaviors (b) and (g) were classified with better accuracy, as indicated in Tables 4 and 5, even though the overall accuracies were significantly worse than the CNN method. Because the motion parameters were extracted from the envelopes that were also used in the LSTM method, the efficient features for the classification might be included in the spectrogram envelopes extracted from dual radars. In the following, we discuss the efficient features and factors of our results.

Table 3. Confusion matrix of the LSTM method.

		Predicted Label							
		(a)	**(b)**	**(c)**	**(d)**	**(e)**	**(f)**	**(g)**	**(h)**
True Label	(a)	0.71/ 0.90/ 0.62	0/ 0/ 0	0/ 0/ 0	0/ 0/ 0	0/ 0/ 0	0.29/ 0/ 0.31	0/ 0.10/ 0.077	0/ 0/ 0
	(b)	0.091/ 0/ 0	0.27/ 0.6/ 0.83	0.091/ 0.2/ 0.083	0.18/ 0/ 0.083	0/ 0.067/ 0	0/ 0.067/ 0	0.36/ 0/ 0	0/ 0.067/ 0
	(c)	0/ 0.11/ 0.07	0.083/ 0/ 0	0.83/ 0.89/ 0.93	0/ 0/ 0	0/ 0/ 0	0/ 0/ 0	0/ 0/ 0	0.083/ 0/ 0
	(d)	0/ 0.083/ 0	0.071/ 0/ 0	0/ 0/ 0	0.86/ 0.92/ 1	0/ 0/ 0	0/ 0/ 0	0.071/ 0/ 0	0/ 0/ 0
	(e)	0/ 0/ 0	0/ 0/ 0	0.11/ 0/ 0	0/ 0/ 0	0.78/ 0.94/ 0.78	0.056/ 0.06/ 0.11	0/ 0/ 0.11	0.056/ 0/ 0
	(f)	0.25/ 0.11/ 0.08	0.13/ 0/ 0	0/ 0.11/ 0.07	0/ 0/ 0	0/ 0/ 0.17	0.38/ 0.78/ 0.75	0.25/ 0/ 0	0/ 0/ 0
	(g)	0/ 0/ 0	0.16/ 0.17/ 0.14	0/ 0/ 0.07	0.077/ 0/ 0	0/ 0/ 0	0.077/ 0.083/ 0	0.62/ 0.75/ 0.79	0.077/ 0/ 0
	(h)	0.059/ 0/ 0	0/ 0/ 0	0/ 0/ 0.08	0/ 0/ 0	0/ 0/ 0	0/ 0/ 0	0/ 0/ 0	0.94/ 1/ 0.92

Each cell represents the results for ceiling/wall/dual radars.

Table 4. Confusion matrix of the RF method.

		Predicted Label							
		(a)	**(b)**	**(c)**	**(d)**	**(e)**	**(f)**	**(g)**	**(h)**
True Label	(a)	0.64/ 0.45/ 0.62	0.091/ 0.27/ 0.23	0/ 0.091/ 0	0.18/ 0/ 0	0.091/ 0/ 0	0/ 0.11/ 0.15	0/ 0/ 0	0/ 0/ 0
	(b)	0/ 0/ 0	0.21/ 0.78/ 0.71	0.14/ 0.11/ 0	0.21/ 0/ 0.14	0.21/ 0/ 0	0.071/ 0.11/ 0.071	0.071/ 0/ 0	0.071/ 0/ 0.071
	(c)	0/ 0.059/ 0.1	0/ 0.059/ 0	0.45/ 0.59/ 0.65	0/ 0.18/ 0	0.45/ 0.059/ 0.25	0/ 0.059/ 0	0/ 0/ 0	0.091/ 0/ 0
	(d)	0.083/ 0/ 0	0/ 0/ 0	0/ 0/ 0.2	0.5/ 0.78/ 0.6	0.083/ 0.22/ 0.067	0.17/ 0/ 0.067	0.083/ 0/ 0.067	0.083/ 0/ 0
	(e)	0/ 0/ 0	0/ 0/ 0	0.1/ 0.091/ 0	0.1/ 0.45/ 0	0.6/ 0.27/ 1	0.1/ 0/ 0	0.1/ 0.091/ 0	0/ 0.091/ 0
	(f)	0.31/ 0.043/ 0.091	0/ 0.26/ 0	0/ 0/ 0.091	0.15/ 0.043/ 0.091	0.077/ 0.26/ 0.18	0.23/ 0.35/ 0.27	0.15/ 0.043/ 0.27	0.077/ 0/ 0

Table 4. *Cont.*

		(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
					Predicted Label				
	(g)	0.15/ 0/ 0	0.15/ 0.36/ 0.29	0.15/ 0.21/ 0.071	0/ 0.071/ 0	0.15/ 0.071/ 0.071	0/ 0/ 0.071	0.15/ 0.29/ 0.43	0.077/ 0/ 0.071
	(h)	0.24/ 0/ 0	0/ 0/ 0	0.29/ 0/ 0	0.059/ 0/ 0	0.12/ 0/ 0	0.059/ 0/ 0	0/ 0/ 0	0.24/ 1/ 1

Each cell represents the results for ceiling/wall/dual radars.

Table 5. Confusion matrix of the SVM method.

		Predicted Label							
		(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
	(a)	0.6/ 0.57/ 0.7	0/ 0.071/ 0	0/ 0/ 0.085	0.1/ 0.14/ 0	0/ 0/ 0	0.2/ 0.14/ 0.085	0/ 0/ 0.13	0.1/ 0.071/ 0
	(b)	0.17/ 0.077/ 0.071	0.33/ 0.38/ 0.5	0.083/ 0/ 0	0/ 0/ 0	0.17/ 0.077/ 0.071	0/ 0.15/ 0.21	0.25/ 0.23/ 0.071	0/ 0.077/ 0.071
	(c)	0.067/ 0.077/ 0.18	0/ 0/ 0	0.33/ 0.077/ 0.46	0/ 0.46/ 0	0.27/ 0.15/ 0.36	0.067/ 0/ 0	0/ 0.077/ 0	0.27/ 0.15/ 0
True Label	(d)	0.23/ 0/ 0	0.23/ 0/ 0	0.077/ 0.23/ 0.1	0.38/ 0.69/ 0.9	0/ 0/ 0	0.077/ 0/ 0	0/ 0/ 0	0/ 0.08/ 0
	(e)	0/ 0/ 0	0/ 0/ 0	0.4/ 0/ 0.3	0/ 0.55/ 0	0.33/ 0.091/ 0.4	0/ 0/ 0.2	0/ 0.18/ 0	0.27/ 0.18/ 0.1
	(f)	0.067/ 0.28/ 0.14	0.13/ 0.11/ 0.06	0.13/ 0/ 0	0.2/ 0.06/ 0	0/ 0.17/ 0	0.2/ 0.28/ 0.6	0.13/ 0.11/ 0.2	0.13/ 0/ 0
	(g)	0.11/ 0.12/ 0.077	0.11/ 0.12/ 0	0/ 0/ 0.15	0/ 0/ 0.077	0.11/ 0.12/ 0.15	0.033/ 0.12/ 0.077	0.22/ 0.5/ 0.38	0.11/ 0/ 0.077
	(h)	0.17/ 0/ 0	0/ 0/ 0	0.083/ 0/ 0	0/ 0/ 0	0.083/ 0/ 0	0/ 0/ 0	0.083/ 0/ 0	0.58/ 1/ 1

Each cell represents the results for ceiling/wall/dual radars.

Next, we discuss the effectiveness of using dual radar data. Similar to the results for the CNN method, better performance was observed with the wall radar data than with the ceiling radar data. These results indicate that the motion information in the horizontal direction obtained with the wall radar includes significant information for classifying the assumed human behaviors. Another reason is that the wall radar received the data for the whole body, whereas the ceiling radar mainly obtained the motion of the head. The confusion matrices further confirm the differences between the two radars' results for the classified behaviors. In particular, the confusion matrices of the RF and SVM methods in Tables 4 and 5 indicate that the combination of the data from the two radars significantly improves the classification accuracy because the data from the two radars complement each other. Similar accuracy improvements based on the fusion of dual radar data also can be confirmed from the confusion matrices of other methods, further verifying the effectiveness of the dual radar data.

We now discuss the efficient features included in the spectrograms. Because the classification accuracy of the eight behaviors with RF and SVM methods was above 60%, we consider the selected feature parameters for these methods. Table 6 shows the selected features for the RF and SVM methods using the filter method. The acceleration and jerk parameters were selected for all radar cases. These results indicate that the detailed motion parameters of acceleration and jerk were more effective than the velocity parameters obtained directly from Doppler radar measurements. However, the LSTM and CNN methods performed better than the RF and SVM methods using the motion parameters.

Table 6. Selected feature parameters for the RF and SVM methods.

Radar	Selected Parameters
Ceiling radar	$v_{\text{c-u-std}}$, $a_{\text{c-u-max}}$, $a_{\text{c-u-std}}$, $a_{\text{c-m-max}}$, $j_{\text{c-u-mean}}$, $j_{\text{c-m-mean}}$, $j_{\text{c-u-std}}$, $j_{\text{c-m-max}}$
Wall radar	$v_{\text{w-m-max}}$, $a_{\text{w-u-min}}$, $a_{\text{w-u-mean}}$, $a_{\text{w-u-std}}$, $a_{\text{w-m-max}}$, $j_{\text{w-m-max}}$, $j_{\text{w-m-min}}$, $j_{\text{w-l-max}}$
Dual radar	$v_{\text{w-u-min}}$, $v_{\text{w-u-mean}}$, $v_{\text{w-u-std}}$, $a_{\text{c-u-max}}$, $a_{\text{w-u-min}}$, $a_{\text{w-u-mean}}$, $a_{\text{w-u-std}}$, $a_{\text{w-m-max}}$, $a_{\text{w-m-min}}$, $a_{\text{w-m-mean}}$, $a_{\text{w-l-min}}$, $j_{\text{c-u-std}}$, $j_{\text{w-m-max}}$, $j_{\text{w-l-max}}$, $j_{\text{c-l-std}}$

v, a, and j denote velocity, acceleration, and jerk, respectively. The subscript "X-Y-Z" indicates radar type–envelope operation, X can be ceiling (c) or wall (w) radars. Y can be u, l, or m, indicating upper, lower, or power-weighted mean velocity, respectively; the parameter was extracted from $v_{\text{u}}(t)$, $v_{\text{m}}(t)$, and $v_{\text{l}}(t)$. Z indicates the calculation for the envelopes (std is standard deviation).

We conclude that deep learning can grasp the detailed information in the spectrograms corresponding to higher-order derivative parameters. In addition, because the CNN method had better accuracy than the LSTM method, detailed motion information was obtained from the main components extracted as the spectrogram envelopes and from other components corresponding to the micromotion of the various body parts.

The findings regarding the efficient features for the classification of human behaviors and falls in restrooms are summarized as follows:

- The wall radar that measured motion in the horizontal direction was more effective than the ceiling radar that measured motion in the vertical direction.
- The accurately classified classes for the two radars were different. Hence, a fusion of the two radars was effective.
- The proposed method effectively used the detailed higher-order derivative parameters of acceleration and jerk.
- Detailed motion information was diffused across the whole of the spectrograms and was not limited to the main components, and was efficiently extracted via the CNN.

However, as the limitation of this study, the concrete clarification of the efficient parameters and/or factors for our classification problem was difficult. To achieve this, we have to find the features that indicate clear divergence of the assumed behaviors in restrooms based on other various approaches (e.g., using principal component analysis, application of other classification algorithms and comprehensive comparison with the results of this study, and data acquisition from a larger number of participants).

5.3. Comparison with Conventional Studies

In this section, we compare our method with the conventional remote sensor-based monitoring methods for restrooms. Table 7 outlines the comparison of the experimental studies aimed at detecting abnormal, dangerous behaviors in restrooms. The proposed method achieved the best performance in terms of the classification accuracy, number of classified behaviors, number of participants, and detection accuracy of the human fall.

Table 7. Comparison of studies on restroom monitoring.

Study	Sensor	Problem	No. of Participants	Performance
[6]	Camera	Detection of dangerous situation	10	N. A. (Secure detection of dangerous situation continues for 60 s)
[9]	Thermal Sensor	Classification of normal/fall data	8	Accuracy over 95% (2-class classification)
[10]	Thermal Sensor	Classification of normal use/fall patterns	10	Accuracy: 97.8% (2-class classification)
[14]	Radar	Detection of dangerous state (such as falls)	3	Detection rate: 95%
[15]	Radar	Classification of normal/abnormal behaviors (including falls)	10	- Accuracy: 62.5% (7-class classification) - Fall classification rate: 83.3%
Our previous study [16]	Radar	Classification of eight behaviors (including falls)	21	- Accuracy: 83.2% (8-class classification) - Fall classification rate: 92.0%
This study	Radar	Classification of eight behaviors (including falls)	21	- Accuracy: 95.6% (8-class classification) - Fall classification rate: 100%

Due to privacy issues, the number of studies on restroom monitoring using cameras is quite limited. Reference [6] is one of the few studies that report on camera-based monitoring of restrooms to detect dangerous situations to protect the elderly. However, because sensors without privacy issues are more suitable for restroom monitoring, approaches using infrared-based thermal sensors and radars have been recently studied. However, most studies classify situations as normal or dangerous behaviors [9,10]. Although thermal sensors show a sufficiently accurate classification, detailed behaviors were not classified because the sensors cannot detect the motion information directly. By contrast, radar techniques can acquire motion velocity information and classify it into multiple behaviors as carried out in [14,15]. However, the accuracy achieved in [15] was insufficient because only the simple feature parameters related to distance and signal information and the RF method were used. Therefore, our previous study [16] proposed the LSTM method that used the rich velocity information obtained via spectrogram envelopes. While both our previous research and the present study can classify the behaviors into eight categories, the proposed method was carried out using CNN and showed higher classification accuracy, including 100% fall detection. In addition, the present study used a relatively large dataset generated from a larger number of participants, and the spectrogram images utilized the rich velocity information included in the Doppler radar signals.

6. Conclusions

This study used Doppler radar technology to classify the behaviors and falls in a restroom based on machine learning approaches. The CNN, LSTM, SVM, and RF methods were applied and compared to determine the most efficient method and features for restroom monitoring. Furthermore, dual radars mounted on the ceiling and wall of a restroom collected motion information in horizontal and vertical directions. The experimental results revealed that the CNN method using the spectrogram images as input achieved the best accuracy of 95.6% when classifying the eight behaviors of 21 participants. In addition, the classification rate of the fall and other behaviors was 100%. These results indicate that the proposed Doppler radar system can accurately recognize behaviors and detect falls in a restroom without any privacy concerns. In addition, we identified efficient features in the motions by comparing the four machine learning methods using single and dual radar data. The motion information corresponding to the higher-order derivative parameters of acceleration and jerk in the horizontal direction was efficient, and the corresponding features were extracted via the CNN method.

However, this study had the following limitations. The motion features of the efficient classification were insufficiently revealed in this study, as discussed in Section 5.2. In addition, only young people participated and eight limited behaviors were assumed.

Thus, further experiments are needed to address the above limitations of this study. In the future, more studies are needed with elderly participants. Other behaviors and falls, such as using a smartphone or falling sideways, should be considered. In addition, combining multiple models, such as LSTM and CNN, may improve classification accuracy because some of the behaviors accurately classified by the LSTM and CNN methods varied. The model combination also may lead to the clarification of the efficient features for behavior classification in restrooms. Furthermore, the use of multiple radars (more than two) is an important future study area.

References

1. World Population Prospects. The 2019 Revision The Key Findings. Available online: https://esa.un.org/unpd/wpp/Publications/Files/WPP2019_10KeyFindings.pdf (accessed on 17 January 2022).
2. Wang, X.; Ellul, J.; Azzopardi, G. Elderly fall detection systems: A literature survey. *Front. Robot. AI* **2020**, *7*, 71. [CrossRef]
3. Neslihan, L.Ö.K.; Belgin, A.K.I.N. Domestic environmental risk factors associated with falling in elderly. *Iran. J. Public Health* **2013**, *42*, 120.
4. Zhang, Y.; Wullems, J.; D'Haeseleer, I.; Abeele, V.V.; Vanrumste, B. Bathroom activity monitoring for older adults via wearable device. In Proceedings of the 2020 IEEE International Conference on Healthcare Informatics (ICHI), Oldenburg, Germany, 30 November–3 December 2020.
5. Vineeth, C.; Anudeep, J.; Kowshik, G.; Vasudevan, S.K. An enhanced, efficient and affordable wearable elderly monitoring system with fall detection and indoor localisation. *Int. J. Med. Eng. Inform.* **2021**, *13*, 254–268. [CrossRef]
6. Meng, L.; Kong, X.; Taniguchi, D. Dangerous Situation Detection for Elderly Persons in Restrooms Using Center of Gravity and Ellipse Detection. *J. Robot. Mechatron.* **2017**, *29*, 1057–1064. [CrossRef]
7. Harrou, F.; Zerrouki, N.; Sun, Y.; Houacine, A. Vision-based fall detection system for improving safety of elderly people. *IEEE Instrum. Meas. Mag.* **2017**, *20*, 49–55. [CrossRef]
8. Rougier, C.; Meunier, J.; St-Arnaud, A.; Rousseau, J. Robust Video Surveillance for Fall Detection Based on Human Shape Deformation. *IEEE Trans. Circuits Syst. Video Technol.* **2011**, *21*, 611–622. [CrossRef]
9. Kido, S.; Miyasaka, T.; Tanaka, T.; Shimizu, T.; Saga, T. Fall Detection in Toilet Rooms Using Thermal Imaging Sensors. In Proceedings of the 2009 IEEE/SICE International Symposium on System Integration (SII), Tokyo, Japan, 29 January 2009; pp. 83–88.
10. Shirogane, S.; Takahashi, H.; Murata, K.; Kido, S.; Miyasaka, T.; Saga, T.; Sakurai, S.; Hamaguchi, T.; Tanaka, T. Use of Thermal Sensors for Fall Detection in a Simulated Toilet Environment. *Int. J. New Technol. Res.* **2019**, *5*, 21–25. [CrossRef]
11. Gurbuz, S.Z.; Amin, M.G. Radar-based human-motion recognition with deep learning: Promising applications for indoor monitoring. *IEEE Signal Process. Mag.* **2019**, *36*, 16–28. [CrossRef]
12. Bhattacharya, A.; Vaughan, R. Deep learning radar design for breathing and fall detection. *IEEE Sens. J.* **2020**, *20*, 5072–5085. [CrossRef]
13. Ma, L.; Liu, M.; Wang, N.; Wang, L.; Yang, Y.; Wang, H. Room-level fall detection based on ultra-wideband (UWB) monostatic radar and convolutional long short-term memory (LSTM). *Sensors* **2020**, *20*, 1105. [CrossRef]
14. Tsuchiyama, K.; Kajiwara, A. Accident detection and health-monitoring UWB sensor in toilet. In Proceedings of the 2019 IEEE Topical Conference on Wireless Sensors and Sensor Networks (WiSNet), Orlando, FL, USA, 20–23 January 2019. [CrossRef]

15. Takabatake, W.; Yamamoto, K.; Toyoda, K.; Ohtsuki, T.; Shibata, Y.; Nagate, A. FMCW Radar-Based Anomaly Detection in Toilet by Supervised Machine Learning Classifier. In Proceedings of the 2019 IEEE Global Communications Conference (GLOBECOM), Waikoloa, HI, USA, 9–13 December 2019. [CrossRef]

16. Hayashi, S.; Saho, K.; Tsuyama, M.; Masugi, M. Motion Recognition in the Toilet by Dual Doppler Radars Measurement. *IEICE Trans. Commun. (Jpn. Ed.)* **2021**, *J104-B*, 390–393. (In Japanese)

17. Tsuyama, M.; Hayashi, S.; Saho, K.; Masugi, M. A Convolutional Neural Network Approach to Classification of Human's Behaviors in a Restroom Using Doppler Radars. In Proceedings of the ATAIT 2021 (2021 International Symposium on Advanced Technologies and Applications in the Internet of Things), Kusatsu, Japan, 23–24 August 2021.

18. Suk, M.; Ramadass, A.; Jin, Y.; Prabhakaran, B. Video human motion recognition using a knowledge-based hybrid method based on a hidden Markov model. *ACM Trans. Intell. Syst. Technol.* **2012**, *3*, 1–29. [CrossRef]

19. Zhang, Y.; Zhang, Z.; Zhang, Y.; Bao, J.; Zhang, Y.; Deng, H. Human activity recognition based on motion sensor using u-net. *IEEE Access* **2019**, *7*, 75213–75226. [CrossRef]

20. Bilen, H.; Fernando, B.; Gavves, E.; Vedaldi, A. Action recognition with dynamic image networks. *IEEE Trans. Pattern Anal. Mac. Intell.* **2017**, *40*, 2799–2813. [CrossRef] [PubMed]

21. Ijjina, E.P.; Chalavadi, K.M. Human action recognition in RGB-D videos using motion sequence information and deep learning. *Pattern Recognit.* **2017**, *72*, 504–516. [CrossRef]

22. Krizhevsky, L.A.; Sutskever, I.; Hinton, G.E. ImageNet Classification with Deep ConvolutionalNeural Networks. *Adv. Neural Inf. Process. Syst.* **2012**, *25*, 1097–1105.

23. Kingma, D.P.; Ba, J.L. Adam: A Method for Stochastic Optimization. In Proceedings of the International Conference on Learning Representations (ICLR), San Diego, CA, USA, 7–9 May 2015.

24. Albon, C. *Machine Learning with Python Cookbook*; O'Reilly Media: Sevastopol, CA, USA, 2018.

25. Tekeli, B.; Gurbuz, S.Z.; Yuksel, M. Information-theoretic feature selection for human micro-Doppler signature classification. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 2749–2762. [CrossRef]

26. Jiang, X.; Zhang, Y.; Yang, Q.; Deng, B.; Wang, H. Millimeter-wave array radar-based human gait recognition using multi-channel three-dimensional convolutional neural network. *Sensors* **2020**, *20*, 5466. [CrossRef] [PubMed]

27. Garripoli, C.; Mercuri, M.; Karsmakers, P.; Soh, P.J.; Crupi, G.; Vandenbosch, G.A.; Pace, C.; Leroux, P.; Schreurs, D. Embedded DSP-based telehealth radar system for remote in-door fall detection. *IEEE J. Biomed. Health Inform.* **2014**, *19*, 92–101. [CrossRef] [PubMed]

28. Su, B.Y.; Ho, K.C.; Rantz, M.J.; Skubic, M. Doppler radar fall activity detection using the wavelet transform. *IEEE Trans. Biomed. Eng.* **2014**, *62*, 865–875. [CrossRef] [PubMed]

29. Cippitelli, E.; Fioranelli, F.; Gambi, E.; Spinsante, S. Radar and RGB-depth sensors for fall detection: A review. *IEEE Sens. J.* **2017**, *17*, 3585–3604. [CrossRef]

30. Amin, M.G.; Zhang, Y.D.; Ahmad, F.; Ho, K.D. Radar signal processing for elderly fall detection: The future for in-home monitoring. *IEEE Sign. Process. Mag.* **2016**, *33*, 71–80. [CrossRef]

31. Jokanović, B.; Amin, M. Fall detection using deep learning in range-Doppler radars. *IEEE Trans. Aero. Electron. Syst.* **2017**, *54*, 180–189. [CrossRef]

32. Han, T.; Kang, W.; Choi, G. IR-UWB sensor based fall detection method using CNN algorithm. *Sensors* **2020**, *20*, 5948. [CrossRef] [PubMed]

33. Saeed, U.; Shah, S.Y.; Shah, S.A.; Ahmad, J.; Alotaibi, A.A.; Althobaiti, T.; Ramzan, N.; Alomainy, A.; Qammer, H.A. Discrete human activity recognition and fall detection by combining FMCW RADAR data of heterogeneous environments for independent assistive living. *Electronics* **2021**, *10*, 2237. [CrossRef]

34. Gorji, A.; Bourdoux, A.; Pollin, S.; Sahli, H. Multi-View CNN-LSTM Architecture for Radar-Based Human Activity Recognition. *IEEE Access (Early Access)* **2022**. Available online: https://ieeexplore.ieee.org/abstract/document/9709793 (accessed on 17 February 2022).

35. Seyfioğlu, M.S.; Özbayoğlu, A.M.; Gürbüz, S.Z. Deep convolutional autoencoder for radar-based classification of similar aided and unaided human activities. *IEEE Trans. Aerosp. Electron. Syst.* **2018**, *54*, 1709–1723. [CrossRef]

36. Yacchirema, D.; de Puga, J.S.; Palau, C.; Esteve, M. Fall detection system for elderly people using IoT and ensemble machine learning algorithm. *Pers. Ubiquitous Comput.* **2019**, *23*, 801–817. [CrossRef]

37. Taylor, W.; Dashtipour, K.; Shah, S.A.; Hussain, A.; Abbasi, Q.H.; Imran, M.A. Radar sensing for activity classification in elderly people exploiting micro-doppler signatures using machine learning. *Sensors* **2021**, *21*, 3881. [CrossRef]

38. Shrestha, A.; Li, H.; Le Kernec, J.; Fioranelli, F. Continuous human activity classification from FMCW radar with Bi-LSTM networks. *IEEE Sens. J.* **2020**, *20*, 13607–13619. [CrossRef]

39. Singh, V.; Bhattacharyya, S.; Jain, P.K. Micro-Doppler classification of human movements using spectrogram spatial features and support vector machine. *Int. J. RF Microw. Comput.-Aided Eng.* **2020**, *30*, e22264. [CrossRef]

40. Saho, K.; Shioiri, K.; Fujimoto, M.; Kobayashi, Y. Micro-Doppler Radar Gait Measurement to Detect Age-and Fall Risk-Related Differences in Gait: A Simulation Study on Comparison of Deep Learning and Gait Parameter-Based Approaches. *IEEE Access* **2021**, *9*, 18518–18526. [CrossRef]

41. Saho, K.; Uemura, K.; Sugano, K.; Matsumoto, M. Using micro-Doppler radar to measure gait features associated with cognitive functions in elderly adults. *IEEE Access* **2019**, *7*, 24122–24131. [CrossRef]