

# A Convolutional Neural Network Approach to Classification of Human's Behaviors in a Restroom Using Doppler Radars

Mutsuki Tsuyama<sup>1</sup>, Sora Hayashi<sup>1</sup>, Kenshi Saho<sup>2,1</sup>, Masao Masugi<sup>1</sup>

<sup>1</sup>Ritsumeikan University, Shiga, Japan

<sup>2</sup>Toyama Prefectural University, Toyama, Japan

ri0082ip@ed.ritsumei.ac.jp, sora8840817@gmail.com,  
saho@pu-toyama.ac.jp, masugi@fc.ritsumei.ac.jp

## Abstract

To develop a monitoring system for the early detection of incidents, such as falling of a person in a restroom, this study introduces a classification method for the behaviors in restrooms using Doppler radars and a convolutional neural network. The proposed system used data from two Doppler radars installed at the ceiling and wall of the restroom for the classification of the behaviors, such as putting off/on pants, sitting down, standing up, and falling. As a result, an accurate classification of the assumed behaviors with over 90% accuracy was achieved.

## 1 Introduction

Population aging has made early detection of falls and syncope in elderly adults significant (World Population Prospects, 2019). In particular, monitoring places with privacy concerns, such as restrooms and bathrooms, is required. For this purpose, camera- (Gao et al., 2011 IEEE International Symposium of Circuits and Systems, 2011) or depth sensor-based approaches (Meng et al., 2016 International Conference on Advanced Mechatronic Systems, 2016) have been studied. However, their measurement accuracy is based on the lighting conditions and subject clothes. In addition, installing cameras in restrooms is difficult in most cases because of privacy issues.

The radar technique is a promising candidate to resolve these problems because it can detect the subject's behaviors while avoiding privacy concerns and the problems associate to low-light and clothes conditions (Ding et al., IEEE J. Emerging and Selected Topics in Circuits and Systems, 2018). Several studies have achieved accurate motion classification using radar images and deep learning (Zhang and Cao, IEE Sensors Letters, 2019). For the monitoring of restrooms, a detection technique for abnormal behaviors based on range information acquired using frequency-modulated radars have been proposed (Takabatake et al., 2019 IEEE Global Communications Conference, 2019). However, this study did not



© 2022 Copyright for this paper by its authors.  
Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).  
CEUR Workshop Proceedings (CEUR-WS.org)

achieve sufficient accuracy for practical use (approximately 60% accuracy for the classification of seven classes, such as falling, sitting down, and standing). Furthermore, classes with relatively larger differences in movements were considered, and the detailed classification of behaviors in restrooms was not considered.

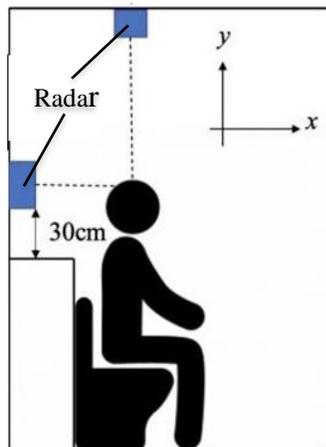
In this study, a radar-based classification of the human's behaviors in a restroom, including falling and other various behaviors, such as putting off pants, taking a toilet paper, and opening the toilet lid is proposed. This system uses two 24 GHz continuous wave Doppler radars placed at the ceiling and a wall behind the subject. The experimental results show an accurate classification of the behaviors using a convolutional neural network (CNN) and combined data of the two radars. The contributions of this study are as follows:

- We demonstrated the practicality of the radar technology, which does not have privacy issues, for the classification of humans' behaviors in the restroom.
- The effectiveness of the two radars installed on the wall and ceiling of a restroom was verified for the classification of basic behaviors and falls.
- The accuracy of 100 % for the classification of falling and other behaviors was achieved.
- The presented CNN-based method based on the fusion of two radar images achieved 95.6 % classification accuracy.

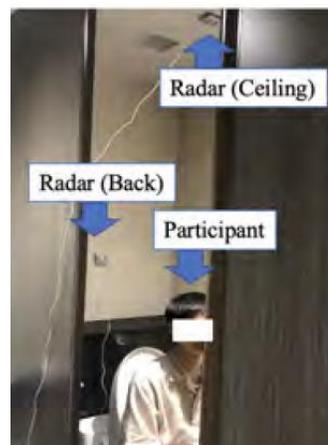
## 2 Experimental Setup

The setup and site of the proposed radar system used to detect the participant's behaviors in a restroom are shown in Figures 1 and 2. We used 24 GHz continuous wave radars (ILT office, BSS-110) with  $\pm 14^\circ$  directivity for the plane shown in Figure 1. That is, the radars can measure behaviors within  $\pm 14^\circ$  forward range. The radars were installed above (ceiling radar) and behind (wall radar) the subject. The radar signals were received by demodulation using a quadrature detector and an analog-to-digital converter with a sampling frequency of 600 Hz.

The study participants who consented to this study were 24 young men (age:  $22.4 \pm 1.1$  years, height:  $173.8 \pm 5.1$  cm). Each participant performed the following eight types of behaviors three times each: (a) opening the toilet lid, (b) putting off the pants, (c) sitting down, (d) taking the toilet paper, (e) standing up, (f) putting on the pants, (g) closing the toilet lid, and (h) falling, where the falling is the motion of falling forward from a seated position.



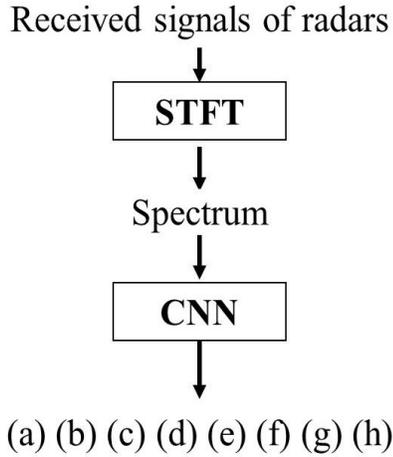
**Figure 1:** Radar system setup



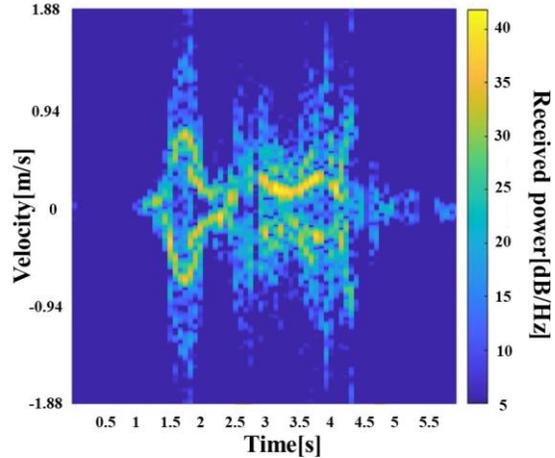
**Figure 2:** Experimental site

### 3 Classification Method

Figure 3 presents an outline of the classification method. First, the short-time Fourier transform (STFT) of the received signals generates radar spectrogram images. For the STFT, a Hamming window function with a length of 128 samples was used. The radar spectrogram showed the time-velocity distribution of the measured movements. Figure 4 shows an example of the spectrogram images of behavior (a) calculated from the received signals of the ceiling radar, where the positive velocity indicates the velocity in the direction approaching the radar and the negative velocity indicates the velocity away from the radar.



**Figure 3:** Outline of the classification method



**Figure 4:** An example of the spectrogram for behavior (a)

Then, the spectrograms were converted to PNG images of size  $164 \times 218$  with RGB color channels, and the generated images were input to the CNN for the classification of eight behaviors. Tables 1 and 2 show the structures of CNNs with one and two input images, respectively. The CNN using one input image assumes the use of the ceiling or wall radar image, while that using two input images assumes the combination of the images of the two radars. Thus, the classification accuracy for three input cases was investigated and compared: only ceiling radar, only wall radar, and the fusion of two radars. Both CNNs are composed of convolution, max pooling, batch normalization (Ioffe and Szegedy, the 32nd International Conference on Machine Learning, 2015), and fully connected layers. The CNN using two input images combines the outputs of the fully connected layer in a concatenate layer. Stochastic gradient descent with the momentum optimization algorithm was used. The loss function is a cross-entropy function: A total of 100 epochs were trained and a batch size of four was used. The learning rate was 0.01, and was decayed by multiplying by 0.5, every 10 epochs. These hyperparameters were optimized empirically.

Layer	Filter Size	Stride	Output Size
Input			(164,218,3)
Convolution 1	(11,11)	(3,3)	(55,73,48)
Maxpooling 1	(3,3)	(2,2)	(27,36,96)
Batch 1			(27,36,96)
Convolution 2	(5,5)	(3,3)	(9,12,128)
Maxpooling 2	(3,3)	(2,2)	(4,5,128)
Batch 2			(4,5,128)
Convolution 3	(3,3)	(1,1)	(4,5,192)
Convolution 4	(3,3)	(1,1)	(4,5,192)
Convolution 5	(3,3)	(1,1)	(4,5,128)
Maxpooling 3	(3,3)	(2,2)	(1,2,128)
Batch 3			(1,2,128)
Fully Connected 1			2048
Fully Connected 2			2048
Fully Connected 3			8

**Table 1:** The Construction of CNN for one input radar image

Layer	Filter Size	Stride	Output Size
Input			(164,218,3), (164,218,3)
Convolution 1	(11,11)	(3,3)	(55,73,48), (55,73,48)
Maxpooling 1	(3,3)	(2,2)	(27,36,96), (27,36,96)
Batch 1			(27,36,96), (27,36,96)
Convolution 2	(5,5)	(3,3)	(9,12,128), (9,12,128)
Maxpooling 2	(3,3)	(2,2)	(4,5,128), (4,5,128)
Batch 2			(4,5,128), (4,5,128)
Convolution 3	(3,3)	(1,1)	(4,5,192), (4,5,192)
Convolution 4	(3,3)	(1,1)	(4,5,192), (4,5,192)
Convolution 5	(3,3)	(1,1)	(4,5,128), (4,5,128)
Maxpooling 3	(3,3)	(2,2)	(1,2,128), (1,2,128)
Batch 3			(1,2,128), (1,2,128)
Fully Connected 1			(2048,2048)
Concatenate			4096
Fully Connected 2			8

**Table 2:** The Construction of CNN for two input radar images

## 4 Performance Evaluation

The classification accuracy of eight behaviors in the restroom were evaluated using hold-out validation for the three input cases explained in the previous section. Each CNN was trained using 80% of the generated spectrogram images, and the remaining 20% of the data were tested. Then, 30 trials of validations were performed by randomly dividing the dataset into training and test data. The mean classification accuracy of the trials was calculated.

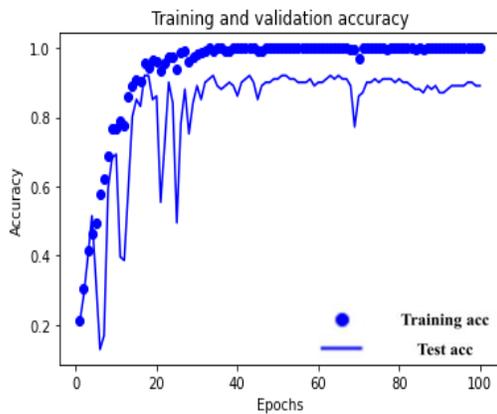
Table 3 lists the mean classification accuracies of the three input cases. All the results achieved accurate classification with an accuracy of over 90%. The CNN using the combination of the two radars achieved 95.6% accuracy and was significantly better than the CNN using only a single radar data.

Finally, the validity and details of our results were discussed using the convergence curve and confusion matrices. First, Figures 5, 6, and 7 show the convergence curves for all input cases. In all

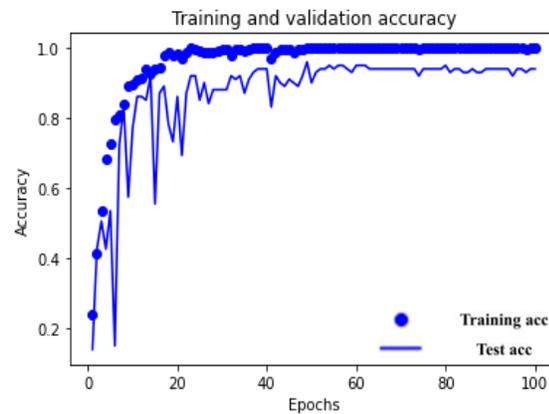
cases, we see that there were no overfittings, and the accuracies converged in less than 40 epochs. The confusion matrices of all the cases are shown in Tables 4, 5, and 6, where the values in the tables are rounded off to the two decimal places. As a result, of the ceiling radar, the classification accuracy of "(f) putting on the pants" was worse. In the wall radar, the accuracy of "(b) putting off the pants" was worse. Table 6 indicates that the worse classification accuracy of (b) is not resolved even though the ceiling radar is combined, while the worse accuracy of (f) in the ceiling radar is resolved by the combination of the radar images. The solution to this problem, which will improve classification accuracy, is an important direction for future studies. However, for all results, the classification accuracy of "(h) falling" was 100%. The most important function in practical use is fall detection, which has already been achieved in this study.

Case	Only ceiling radar	Only wall radar	Combination of two radars
Classification accuracy	90.3%	91.5%	<b>95.6%</b>

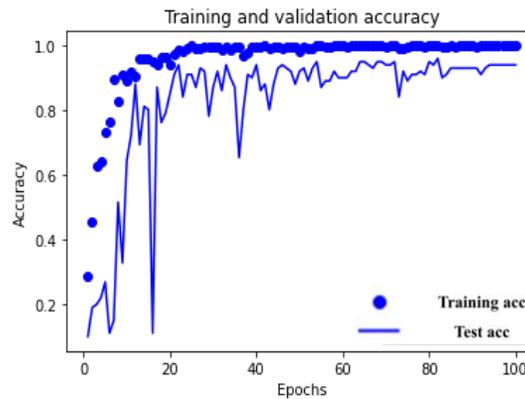
**Table 3:** Average classification rate for all cases



**Figure 5:** Learning curve of ceiling radar



**Figure 6:** Learning curve of wall radar



**Figure 7:** Learning curve of both radar

		Predict Label							
		(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
True Label	(a)	0.90	0	0	0	0	0	0.09	0
	(b)	0	0.85	0	0.15	0	0	0	0
	(c)	0.08	0	0.92	0	0	0	0	0
	(d)	0	0	0	1	0	0	0	0
	(e)	0	0	0	0	0.92	0.08	0	0
	(f)	0	0	0	0.22	0.06	0.72	0	0
	(g)	0.08	0	0	0	0	0	0.92	0
	(h)	0	0	0	0	0	0	0	1

**Table 4:** Confusion matrix of ceiling radar

		Predict Label							
		(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
True Label	(a)	0.93	0	0	0	0	0	0.07	0
	(b)	0	0.77	0	0	0	0	0.23	0
	(c)	0.	0	1	0	0	0	0	0
	(d)	0	0	0	0.9	0.1	0	0	0
	(e)	0	0	0	0	1	0	0	0
	(f)	0	0	0	0	0	0.92	0.07	0
	(g)	0	0	0	0	0	0	1	0
	(h)	0	0	0	0	0	0	0	1

**Table 5:** Confusion matrix of wall radar

		Predict Label							
		(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
True Label	(a)	1	0	0	0	0	0	0	0
	(b)	0	0.79	0	0	0	0	0.21	0
	(c)	0	0	1	0	0	0	0	0
	(d)	0	0	0	1	0	0	0	0
	(e)	0	0	0	0	1	0	0	0
	(f)	0.07	0	0.07	0	0	0.87	0	0
	(g)	0.07	0	0	0	0	0	0.92	0
	(h)	0	0	0	0	0	0	0	1

**Table 6:** Confusion matrix of both radar

## 5 Conclusion

In this study, a classification method for the participant's behaviors in the restroom using two Doppler radars is proposed. In this method, the Doppler radar spectrograms acquired with the radars installed above and behind the participants were input to the CNN. The experimental results showed that an accurate classification of 95.6% was achieved for eight types of realistic behaviors in the restroom, and the classification accuracy of falling and other behaviors was 100%. Future studies will improve the feasibility of the proposed radar system by incorporating participants of different ages and conducting additional experiments in various types of restrooms.

## References

World Population Prospects, The 2019 Revision The Key Findings , [https://esa.un.org/unpd/wpp/Publications/Files/WPP2019\\_10KeyFindings.pdf](https://esa.un.org/unpd/wpp/Publications/Files/WPP2019_10KeyFindings.pdf)., 2019.

H. Gao, W. Lin, X. Yang, H. Li, N. Xu, J. Xie and Y. Li, "A new Network-based algorithm for Multi-camera abnormal activity detection," in *2011 IEEE International Symposium of Circuits and Systems*, Rio de Janeiro, Brazil, 2011.

L. Meng, X. Kong and D. Taniguti, "Danger situations detection for the senior in toilet room using the center of gravity," in *2016 International Conference on Advanced Mechatronic Systems*, Melbourne, VIC, Australia, 2016.

C. Ding, L. Zhang, C. Gu, L. Bai, Z. Liao, H. Hong, Y. Li and X. Zhu, "Non-Contact Human Motion Recognition Based on UWB Radar," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, 2018, pp. 306-315.

R. Zhang and S. Cao, "Real-Time Human Motion Behavior Detection via CNN Using mmWave Radar," *IEEE Sensors Letters*, vol. 3, 2019.

W. Takabatake, K. Yamamoto, K. Toyoda, T. Ohtsuki, Y. Shibata and A. Nagate, "FMCW Radar-Based Anomaly Detection in Toilet by Supervised Machine Learning Classifier," in *2019 IEEE Global Communications Conference*, Waikoloa, HI, USA, 2019.

S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," *Proceedings of the 32nd International Conference on Machine Learning*, 2015, pp. 448-456.