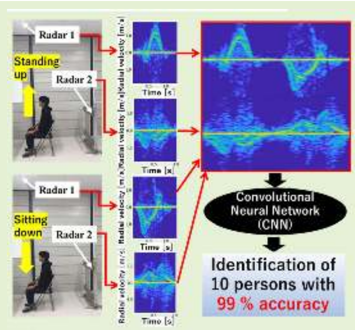
IEEE SENSORS JOURNAL, VOL. 21, NO. 4, FEBRUARY 15, 2021

Accurate Person Identification Based on Combined Sit-to-Stand and Stand-to-Sit Movements Measured Using Doppler Radars

Kenshi Saho, *Member, IEEE*, Keitaro Shioiri, and Keisuke Inuzuka

***Abstract*—This article demonstrates the identification of 10 persons with 99% accuracy achieved by combining micro-Doppler signatures of sit-to-stand and stand-to-sit movements. Data from these movements are measured using two radars installed above and behind the person. Images of Doppler spectrograms generated using the measured data are combined and input to a convolutional neural network. Experimental results show the significantly better accuracy of the proposed method compared with conventional methods that do not perform data combination. The accuracy of identifying 10 participants having similar ages and physical features was 96–99%, despite the relatively small training set (number of training samples: 50–90 Doppler radar images per person). These results suggest that combining sit-to-stand and stand-to-sit movements provides sufficient information for accurate person identification and such information can be remotely acquired using Doppler radar measurements.**



***Index Terms*— Doppler radar, identification of persons, machine learning, motion measurement.**

I. Introduction

Automatic person identification enables monitoring and authentication in security applications. Face recognition using cameras has been the most widely used method for identification [1], [2], and other devices such as smart speakers have become ubiquitous and can be used for identification [3]. However, these devices present privacy issues. Hence, to achieve person identification without privacy concerns, various biomarkers, such as fingerprint, iris, vein, brainwave, and heartbeat characteristics, have been studied [4], [5]. Nevertheless, acquiring these biomarkers often demands physical contact with the user, complicated setups, or sensor attachment.

Alternatively, person identification using behavioral characteristics is being actively developed. Behavioral characteristics include gait features that can be sensed remotely and used for continuous identification [6]. Likewise, movements such as sit-to-stand and stand-to-sit have also been considered as features to extend the applicability of identification systems [7] (e.g., user authentication by simply sitting down on a chair [8] and identification of suspicious persons based on measurement of simple standing up or sitting down movements in criminal investigations, similar to gait authentication systems [9]). Again, however, gait authentication systems relying on cameras [10], [11] present privacy issues. Although other methods for person identification consider accelerometers to characterize gait and other movements [7], [12], [13], they require the user to carry or wear the sensor device.

To achieve remote unobtrusive measurements for person identification, micro-Doppler signatures are being recently used as time–velocity features that characterize human movements from the Doppler effects reflected in acquired radio or ultrasound waves [14], [15]. Sensing based on these waves allows us to remotely obtain detailed velocity information without requiring sensors in the users. Although ultrasound micro-Doppler signatures of gait are effective for person identification [15], [16], ultrasound waves are strongly attenuated in air, thus undermining their measurement robustness for security systems at relatively long distances.

In contrast, Doppler radars using microwaves can robustly measure relatively large areas. Although many existing studies

Manuscript received September 4, 2020; revised October 15, 2020; accepted October 18, 2020. Date of publication October 22, 2020; date of current version January 15, 2021. The associate editor coordinating the review of this article and approving it for publication was Dr. Michail Antoniou. *(Kenshi Saho and Keitaro Shioiri contributed equally to this work.) (Corresponding author: Kenshi Saho.)*

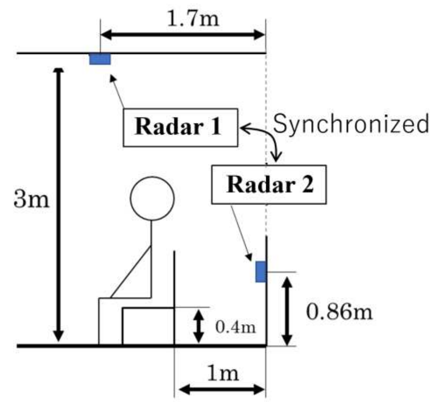
The authors are with the Faculty of Engineering, Toyama Prefectural University, Imizu 939-0398, Japan (e-mail: saho@pu-toyama.ac.jp).

Digital Object Identifier 10.1109/JSEN.2020.3032960

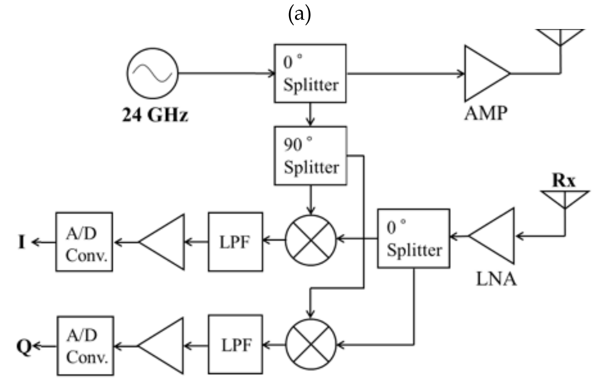
on Doppler radar measurements from persons have not been focused on person identification but on recognition of motion types (e.g., sitting down, standing up, walking, and falling) [17]–[19], the feasibility of person identification based on machine learning has been demonstrated using micro-Doppler signatures of gait [19]–[23] and other motions such as jumping and running [22]. These methods apply machine learning to images of micro-Doppler signatures and can achieve high accuracy, such as 2-person identification with 99.9% accuracy [19] and 15-person identification with 94.4% accuracy [22]. Likewise, we demonstrated the identification of 10 persons based on micro-Doppler signatures of sit-to-stand and stand-to-sit movements with accuracies of 93.9% and 94.9%, respectively [24]. To the best of our knowledge, [24] was the first report of Doppler radar-based person identification using standing-up and sitting-down movements. Similarly, Sakamoto [25] recently identified 6 persons using walk-to-sit movements (each participant was walking and then sat down in a chair) with 93.3% accuracy.

Only [24], [25] have addressed person identification using Doppler radar measurements of standing-up and/or sitting-down movements, and these have not investigated its accuracy for various conditions considering such as the number of persons and training samples. Moreover, a comprehensive study on person identification using the micro-Doppler signatures of such movements and the performance improvement based on feature combinations of these movements have not been conducted including in our previous study [24]. Further, the combination of data acquired from multiple radars is expected to improve the identification accuracy, as reported in a gait classification study [19], [23]. However, the combination of multiple radar data has not been investigated for the person identification using standing-up and/or sitting-down movements.

This study proposes an accurate person identification method that combines the data from sit-to-stand and stand-to-sit movements to improve the method introduced in [24]. The data from the movements are measured using two Doppler radars installed directly above and behind the person to be identified. By combining the images of micro-Doppler signatures from the two movements and two radars, the improved method achieves accurate person identification using a convolutional neural network (CNN) for classification. We evaluated the identification accuracy of the proposed method by implementing variants of the radar/movement data and applying different CNNs. Furthermore, we compared the effectiveness of the adopted CNN with that of recognition based on the conventional bag of features [26]. Finally, we evaluated the accuracy of the proposed method according to the number of identified persons and the number of training samples. The main contribution of this study is the achievement of 99% accuracy for 10-person identification by combining movement and radar data. Further, conduction of a more in-depth analysis than [24] reveals the detailed performance and effectiveness of person identification via sit-to-stand/stand-to-sit movements. The contributions of this article as compared with our previous one [24] are as follows:



(a)



(b)

Fig. 1. Doppler radar measurement system. (a) Diagram of measurement setup. (b) Block diagram of Doppler radar setup (Tx, transmitter; Rx, receiver; AMP, amplifier; LNA, low-noise amplifier; LPF, lowpass filter; A/D Conv., analog-to-digital converter).

- Accuracy improvement via fusion of data from two radars.
- Accuracy improvement via fusion of data for both movement types (sit-to-stand and stand-to-sit).
- Further accuracy improvement via fusion of all the data mentioned above.
- Clarification of identification accuracy according to number of persons.
- Clarification of identification accuracy according to number of training samples.
- Clarification of the effectiveness of use of CNN via comparisons with classical methods of image processing and Doppler radar spectrogram processing.

II. Doppler Radar Measurement and Dataset

The Doppler radar system to measure the sit-to-stand and stand-to-sit movements is similar to the one we used for [24]. Fig. 1 illustrates the system and shows a block diagram to obtain the Doppler radar measurements. Radars 1 and 2 are above and behind the person, respectively, and the operation of both radars is synchronized. The radars transmit 24 GHz sinusoidal waves with directivities of $\pm 14°$ for the plane shown in Fig. 1(a). The received radar signal is demodulated with a quadrature detector (Fig. 1(b)), and the I/Q (in-phase/quadrature) signal is the received signal with sampling frequency of 600 Hz (being a sufficiently large for

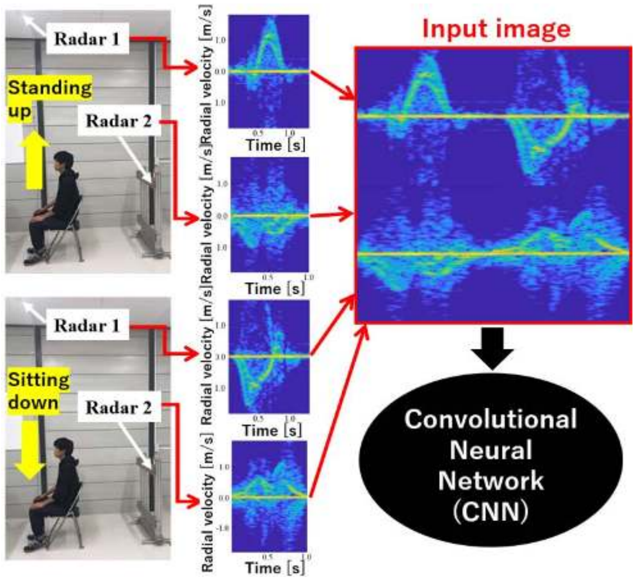


Fig. 2. Experimental setup and outline of proposed method.

measuring the sit-to-stand and stand-to-sit movements), which corresponds to velocity measurement range of ± 1.875 m/s.

We generated a dataset for the same 10 persons of [24], who were 10 healthy adults (8 men, 2 women; age, 21–24 years; mean height, 168.8 ± 7.87 cm; mean weight, 59.1 ± 6.15 kg) and could easily perform the standing-up and sitting-down movements. There are no publicly available datasets on micro-Doppler signatures of the sit-to-stand and stand-to-sit movements and we thus generated the dataset for this study. Upon collection of sit-to-stand and stand-to-sit data, the participants were asked to stand up or sit down at a self-selected speed. We conducted 100 measurements for both the sit-to-stand and stand-to-sit movements per participant. Thus, $100 \times 10 = 1000$ samples were collected from both movements per radar. The data were collected over some days with sufficient intervals (at least 7 days) to prevent the participants from becoming fatigued and to collect data including effects on the inter-day reproducibility of their movements. No restrictions were imposed regarding the arm motion, foot position, and types of clothes and shoes worn by the participants. The spectrograms of the received signals were calculated using the short-time Fourier transform with Hamming window of 213 ms, which is empirically optimized and corresponds to a velocity resolution of 2.93 mm/s. Note that the acquired spectrograms were not subjected to any other processing such as moving target indicators.

The first column in Fig. 2 shows the experimental setup, and the second column shows the corresponding spectrograms, whose detailed interpretation is provided in [24]. The spectrograms were converted into JPEG images of 224×224 pixels. Each image corresponds to a sample in the constructed dataset.

III. Person Identification

A. Proposed Person Identification Using Combined Doppler Radar Measurements

Fig. 2 outlines the proposed person identification method. We input the spectrograms to a CNN for person identification.

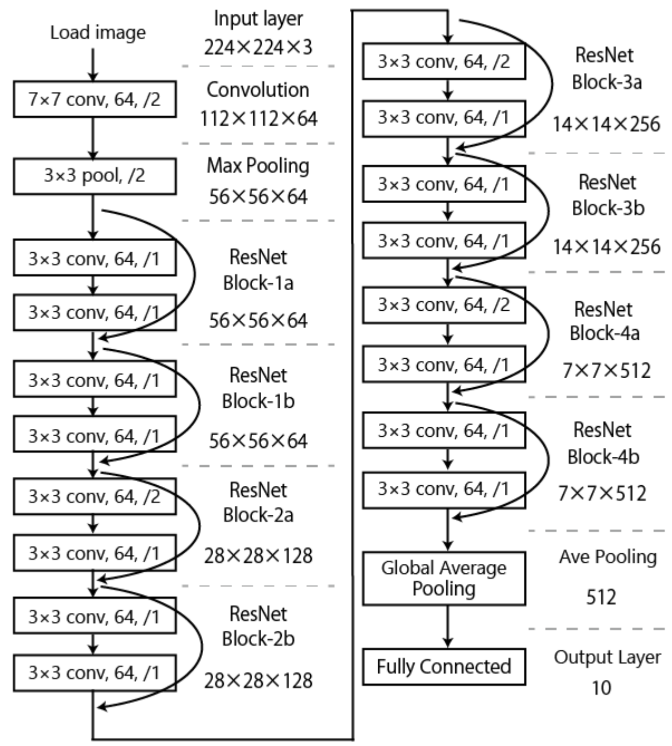


Fig. 3. Structure of the CNN used in the proposed method (ResNet-18).

Specifically, the input of the CNN is the combination of the four spectrograms, that is, the sit-to-stand and stand-to-sit spectrograms from radars 1 and 2. Image combination simply groups the four images, as shown in the input image of Fig. 2. The upper, lower, left, and right parts of the input image correspond to radar 1, radar 2, sit-to-stand, and stand-to-sit data, respectively. The CNN for person identification has the ResNet-18 architecture [27], [28], because it was the most effective network for radar-based person identification in our previous study [24] and other recognition methods based on CNNs [17], [22]. This selection of the network is validated via the comparison with other machine-learning methods as presented in the next subsection.

Fig. 3 shows the structure of the CNN used in the proposed method. We used the basic ResNet structure [28]. When input image (224×224 pixels) is loaded, convolution and pooling operations are performed before the ResNet blocks. Then, then ResNet blocks are configured as a sequence of BN (batch normalization)-ReLU (rectified linear unit)-Conv-BN-ReLU-Conv. At the end of the network, global average pooling is performed and the fully connected layer is then used for prediction, whose weights are optimized using the stochastic gradient descent with momentum optimization algorithm to minimize the difference between the training and output labels. The loss function was a cross-entropy function. We trained for 50 epochs and used a batch size of 64. The learning rate was 0.01 and was decayed by multiplying 0.5 every 10 epochs. These hyperparameters were empirically optimized (similarly, the parameters of the comparative methods presented in the next section were also empirically optimized).

B. Comparisons

We evaluated the accuracy of person identification in various cases to demonstrate the effectiveness of the proposed method. First, we compared the effectiveness of the CNN-based extracted features with classical features of the spectrogram envelopes presented in [29]. For this comparison, we evaluated the identification accuracy using a support vector machine and the extracted spectrogram features. Then, we compared the effectiveness of the proposed CNN with that of conventional image recognition based on bag of features, for which we selected the Harris and SURF features and used a support vector machine for identification [26].

We also compared the identification accuracy with other types of CNNs. We selected the following representative networks used in identification of human motion based on Doppler radar measurements for comparison: AlexNet, GoogLeNet, VGG16, and ResNet-18. These networks also provided high accuracy in [24], but we did not consider the data combination of radars/movements in that work. Thus, we verified the effectiveness of ResNet-18 against the other networks using the combined spectrograms. Furthermore, we considered deeper networks such as ResNet-50, ResNet-101 [27], and DenseNet-201 [30] to investigate the effect of increasing the number of CNN layers on person identification.

Finally, we evaluated the input spectrograms by considering the following combinations:

- All spectrograms (method proposed in this article)
- Only sit-to-stand data from radars 1 and 2
- Only stand-to-sit data from radars 1 and 2
- Sit-to-stand and stand-to-sit data from radar 1
- Sit-to-stand and stand-to-sit data from radar 2
- No combination (method proposed in [24])

C. Evaluation

We evaluated the identification accuracy on the data from the 10 persons for each method and variant described in Section 3.2. For the evaluations, we used hold-out validation as follows. For each method, the corresponding CNN was trained using 700 spectrograms (70% of data, 70 images per person), and the remaining 300 spectrograms (30 images per person) were used as test data. Then, 30 trials of hold-out validations were conducted by randomly varying the training data selection. The mean and standard deviation of the accuracy across the 30 trials were calculated. The significance of the differences in the mean accuracy across trials between the proposed and comparison methods was determined from the p-value of a Welch's t-test at a significance level of 0.05.

We also investigated the performance of the proposed method according to the number of participants and the number of training samples. Again, 30 trials of evaluations using hold-out validation were conducted by randomly varying the training data selection.

Moreover, we investigated the efficiency of areas in the input images using a Grad-CAM (gradient-weighted class activation mapping) visualization [31]. The Grad-CAM visualizes critical regions in input images by producing a heatmap showing the importance of the decision. We can see efficient movement (sit-to-stand and/or stand-to-sit) and radar positioning (radar 1 and/or 2) for the identification of each participant and discuss the reasons for performance changes occurring owing to the image combination of the proposed method.

TABLE I
MEAN AND STANDARD DEVIATION (SD) OF 10-PERSON IDENTIFICATION ACCURACY USING COMBINED SPECTROGRAMS AND DIFFERENT CLASSIFICATION METHODS (TRAINING DATA: 70%)

Method		Mean accuracy (%)	SD (%)
Spectrogram features + SVM		64.2 [a]	2.2
BoF+SVM	Harris	16.2 [a]	2.0
	SURF	54.5 [a]	2.1
CNN	AlexNet	97.7 [b]	1.0
	GoogLeNet	94.7 [a]	1.1
	VGG16	95.2 [a]	1.5
	ResNet-50	93.6 [a]	1.3
	ResNet-101	95.8 [a]	1.2
	DenseNet-201	95.1 [a]	1.3
	ResNet-18	**98.3**	**0.84**

[a] p of t-test for differences between proposed method and each method below 0.001; [b] $p = 0.0125$.

SVM: Support vector machine.

BoF+SVM: Recognition based on bag of features with identification performed by SVM.

IV. RESULTS AND DISCUSSION

A. Identification Accuracy for Different Methods

We first compared the accuracy to identify the 10 participants using the proposed method and the variants and methods described in Section 3-B. Table I lists the results for the evaluated methods. ResNet-18, adopted in the proposed method, achieves the highest accuracy, being significantly superior to the other methods ($p < 0.05$). Classic spectrogram features fail to achieve accurate identification because sufficient information to identify the individuals was not extracted. In addition, conventional recognition based on bag of features also fails to achieve accurate identification and its accuracy is worse than the recognition based on spectrogram features. Because bag of features is not designed for radar spectrograms, the identification accuracy using spectrogram features was relatively better. However, these results highlight the effectiveness of CNNs. It appears that CNNs can efficiently extract the features of individuals through their deep learning-based process. Similar to our previous study, ResNet-18 is the most effective network for person identification. Although deeper networks were also examined, their accuracy was below that of ResNet-18.

Table II lists the results of the variants of the proposed method according to radars and movements using ResNet-18 for identification. The proposed method significantly outperforms the other variants, including the method proposed in our previous work [24], which did not combine the spectrograms. These results demonstrate the effectiveness of combining the movement and radar data.

TABLE II

MEAN ACCURACY OF DATA VARIANTS OF PROPOSED METHOD FOR 10-PERSON IDENTIFICATION USING RESNET-18 (TRAINING DATA: 70%)(COMPARISON WITH OUR PREVIOUS STUDY [24])

Movement	Radar 1	Radar 2	**Both radars**
Sit-to-stand	93.6% [a] [24]	92.9% [a] [24]	**96.3%** [a]
Stand-to-sit	94.9% [a] [24]	93.7% [a] [24]	**96.9%** [a]
Both movements	**96.6%** [a]	**96.4 %** [a]	**98.3%**

[a] p-value of t-test for differences between proposed method and variants is below 0.001.

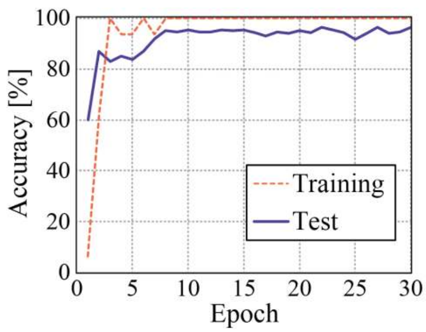


Fig. 4. Examples of convergence curves of proposed method for training and test sets.

Fig. 4 shows a convergence curve of the proposed method to confirm that no overfitting occurs. Furthermore, the accuracy converges in less than 10 epochs for the test set. The average training time for ResNet-18 was 11.2 seconds per epoch using a computer with an NVIDIA Quadro P5000 GPU (16 GB RAM) and a 1.7 GHz Intel Xeon bronze 3106 CPU. This computation time is adequate to use the proposed method in practice.

B. Identification Accuracy According to Number of Persons

We then investigated the accuracy of the proposed method according to the number of persons to be identified using 70 training samples per person. Fig. 5 shows the mean and standard deviation values of the identification accuracy across the 30 trials. The mean accuracy decreases as the number of persons to be identified increases. However, an accuracy above 97% was maintained until the identification of 10 participants across trials. For six persons, the mean accuracy was 98.9%, with 100% accuracy being achieved for several trials. Thus, the proposed method can suitably identify 10 or fewer persons.

C. Identification Accuracy According to Number of Training Samples

The results reported thus far considered training with 70 samples per person. We also investigated the identification accuracy according to the number of training images, obtaining the results shown in Fig. 6. Although training using only 10 samples per person did not achieve a very high accuracy, using 30 samples per person substantially improved accuracy above samples per person did not achieve a very high accuracy, using 30 samples per person substantially improved accuracy above 90%. Therefore, the proposed method does not

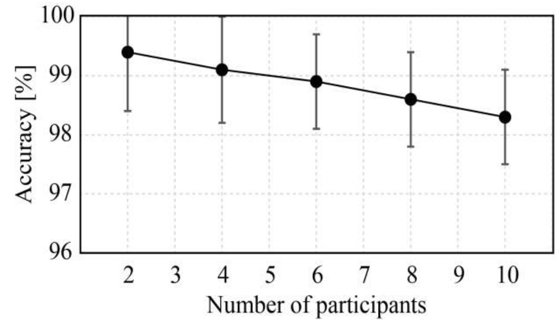


Fig. 5. Identification accuracy of proposed method according to number of persons to be identified using 70 training samples per person.

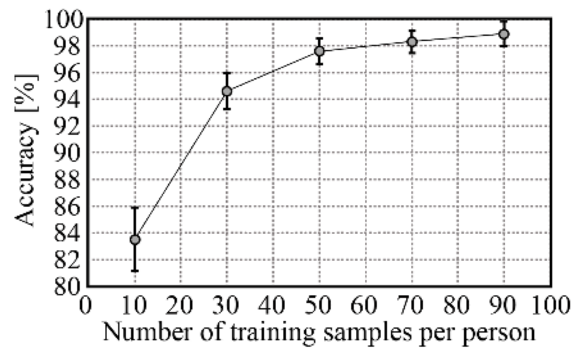


Fig. 6. Identification accuracy of proposed method for 10 persons according to number of training samples.

require a large training set to provide high accuracy. Accuracy reaches its highest of 99.0% for 90 training samples per person, being a suitable accuracy for a variety of authentication applications.

D. Discussion on Grad-CAM Visualization Results

This section investigates the efficiency of movement/radar positioning for person identification and the reasons why the proposed method achieved better accuracy. Fig. 7 shows examples of the Grad-CAM visualization results for all participants, where each panel corresponds to a different participant; the Grad-CAM visualization results of other images of each participant exhibited a similar tendency.

The Grad-CAM results shown in Fig. 7 indicate that the stand-to-sit images were particularly useful for identifying all participants, especially participants (a) and (b). Although sit-to-stand images were also used for participants (c)–(g), stand-to-sit images were more important than sit-to-stand images. Furthermore, for participants (h)–(j), although sit-to-stand is efficiently used, stand-to-sit is also used as the important features. These results indicate that stand-to-sit movement is more efficient in identifying individuals compared with sit-to-stand movement. This is consistent with the results presented in Table II.

If only the stand-to-sit data are used, identification of individuals with a characteristic stand-to-sit movement, such as participant (j), may become challenging. In fact, the features obtained from the sit-to-stand images of other participants, particularly participants (e)–(g), are efficiently combined with stand-to-sit images. The learning of movement combinations by the CNN leads to a performance improvement.

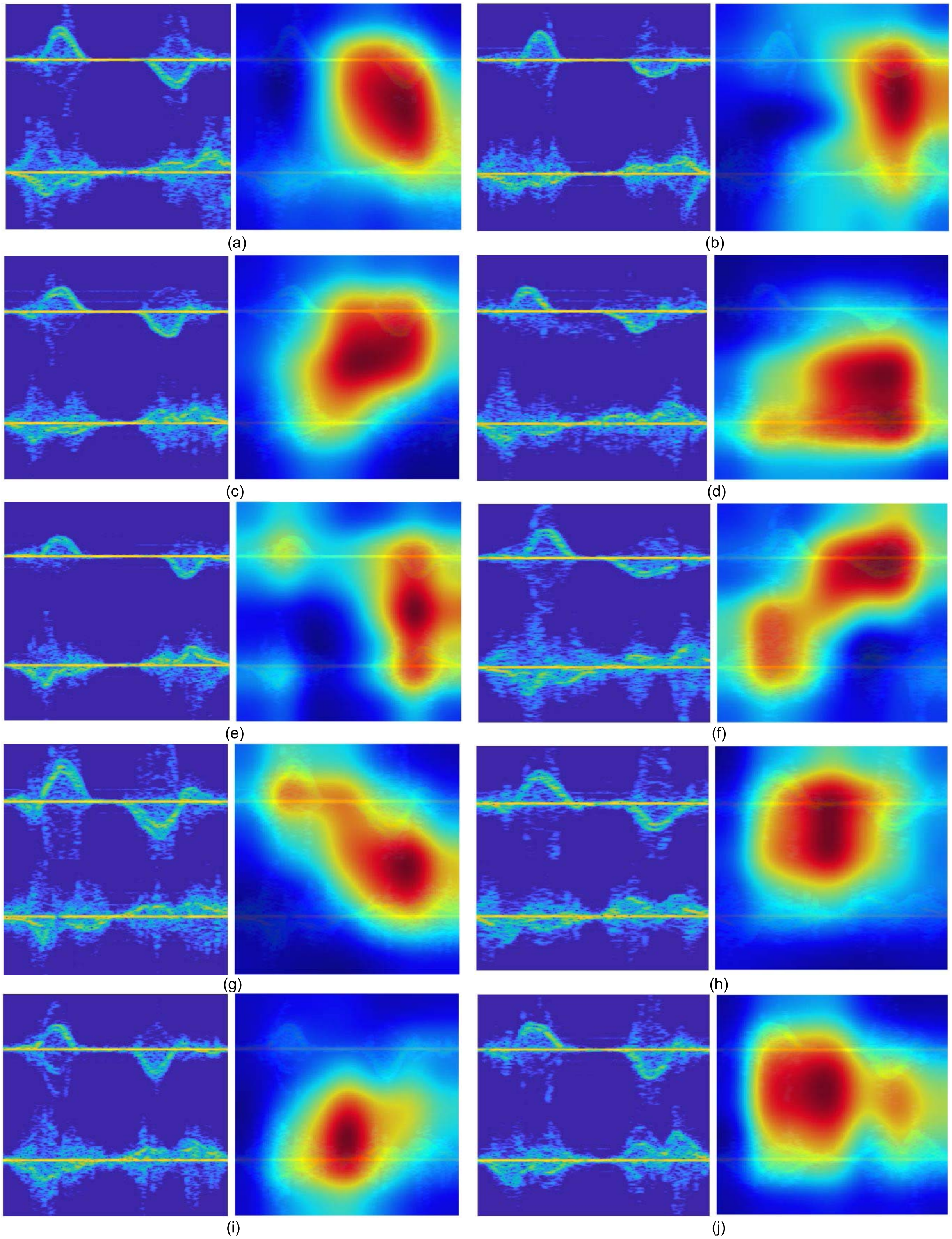


Fig. 7. Examples of Grad-CAM visualization results. Each panel corresponds to a different participant (e.g. subfigure (a) is the result of participant (a)) and shows the original input image (left) and Grad-CAM heatmap image (right). The red color in the heatmap indicates a high score of importance.

With respect to the effectiveness of using two radars, i.e., deploying motion information on the horizontal/vertical directions, radar selection was found to depend on the participant (see Fig. 7). For example, the identification of participant (h) mainly used radar 1 data, whereas that of participant (i) mainly used radar 2 data. These results indicate an important factor of the performance improvement achieved by the proposed method is the appropriate selection and combination of data collected by two radars by the CNN.

TABLE III
COMPARISON OF DIFFERENT STUDIES ON PERSON IDENTIFICATION USING MICRO-DOPPLER SIGNATURES

Study	Movement	No. persons	Accuracy (%)	No. training samples per person	No. radars
[22]	Walk	15	94.4	1500	1
	Run	15	95.2		
	Jump	15	89.2		
[19]	Walk	2	99.9	100	3
[20]	Walk	10	85.6	3200	1
[25]	Walk-to-sit	6	93.3	90	1
[24]	Sit-to-stand	10	93.6	70	1
	Stand-to-sit	10	94.9	70	
This study	Sit-to-stand	10	96.3	70	2
	Stand-to-sit	10	96.9	70	
	Sit-to-stand and stand-to-sit	**10**	**96.6**	**70**	**1**
		10	**94.5**	**30**	**2**
		6	**98.9**	**70**	
		10	**98.3**	**70**	
		10	**99.0**	**90**	

In summary, accurate person identification is mainly based on the stand-to-sit movement. Nevertheless, the combination of stand-to-sit and sit-to-stand images and the deployment of two radars significantly improved the identification accuracy.

E. Discussion on Data Fusion Scheme

In our proposed method, the simple concatenation of four images is used for data fusion (image fusion scheme). However, many other studies considering data fusion for the CNN input have used a network fusion scheme, which uses an independent convolution layer for each image and then combines the decision of the networks [32], [33]. Therefore, this subsection compares these two data fusion schemes. In the network-based fusion scheme, each image is inputted to the network shown in Fig. 3 independently, and the marge of the decisions of four networks was conducted based on the majority vote.

As a result, the identification accuracy and standard deviation for the network fusion scheme (with 70% training samples and 30 tests assuming 10-person identification) was 97.1 ± 1.2%. Although this is a good result, it was worse than that of the image fusion scheme of the proposed method (98.3 ± 0.84%, presented in Table I). Please note that although other methods of data fusion can be considered (e.g., classifier-level fusion [33]), their identification accuracies are worse than the proposed method. Thus, we selected the image fusion scheme for the proposed method.

F. Comparison With Conventional Studies

Table III lists the main characteristics of studies on person identification using micro-Doppler signatures. The method proposed in this study outperforms various person identification methods regarding both accuracy and required training samples. Compared with the conventional identification methods using standing-up or sitting-down movements, our proposal achieves significantly higher accuracy by combining data from two radars and/or two types of movements.

The specific reasons for the improved performance of the proposed method compared with the similar ones cannot be determined because the detailed mechanism of person identification based on micro-Doppler signatures and CNNs remains unclear. Nevertheless, we can consider that the reproducibility of the sit-to-stand and stand-to-sit movements increases accuracy. In fact, the inter-day reproducibility of gait is not always high [34], and the reproducibility of some participants during walking could not be confirmed in [25]. In contrast, sit-to-stand features in a timed up and go test have shown relatively high reproducibility in healthy subjects [35]. Likewise, we confirmed small inter-day variability in the spectrograms of the sit-to-stand and stand-to-sit movements in all the participants of our study. Thus, the higher accuracy with fewer training samples achieved by the proposed method may be due to the high reproducibility of these movements.

V. CONCLUSION

We propose a CNN-based person identification method using combined spectrograms of sit-to-stand and stand-to-sit movements measured from two Doppler radars. The effectiveness of the CNN and combined spectrograms of the radars and movements is verified and compared with other methods. The proposed method achieves 99.0% accuracy with 90 training samples per person and 98.3% accuracy with 70 training samples per person for the identification of 10 participants. The Grad-CAM visualization results indicate that the stand-to-sit movement is mainly used for the proposed identification method. However, these results also indicate that, depending on the participant, efficient movement images and/or appropriate radar data are selected, and their features are learned by the CNN to accurately identify each individual. Thus, the sit-to-stand and stand-to-sit movements and their combination reflect information to identify individuals. In future work, we will investigate the effectiveness of the proposed method for the identification of more persons based on additional experiments considering diverse population and/or open datasets such as [36] and aim toward the development of practical authentication systems.

ACKNOWLEDGMENT

The authors would like to thank Editage (www.editage.com) for English language editing.

REFERENCES

[1] I. Masi, Y. Wu, T. Hassner, and P. Natarajan, "Deep face recognition: A survey," in *Proc. 31st SIBGRAPI Conf. Graph. Patterns Images (SIBGRAPI)*, 2018, pp. 471–478.

[2] J. Zheng, R. Ranjan, C.-H. Chen, J.-C. Chen, C. D. Castillo, and R. Chellappa, "An automatic system for unconstrained video-based face recognition," *IEEE Trans. Biometrics, Behav., Identity Sci.*, vol. 2, no. 3, pp. 194–209, Jul. 2020.

[3] B. Sudharsan, P. Corcoran, and M. I. Ali, "Smart speaker design and implementation with biometric authentication and advanced voice interaction capability," in *Proc. AICS*, 2019, pp. 305–316.

[4] D. Menotti *et al.*, "Deep representations for iris, face, and fingerprint spoofing detection," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 4, pp. 864–879, Apr. 2015.

[5] B. Choudhury, P. Then, B. Issac, V. Raman, and M. K. Haldar, "A survey on biometrics and cancelable biometrics systems," *Int. J. Image Graph.*, vol. 18, no. 01, Jan. 2018, Art. no. 1850006.
[6] V. M. Patel, R. Chellappa, D. Chandra, and B. Barbello, "Continuous user authentication on mobile devices: Recent progress and remaining challenges," *IEEE Signal Process. Mag.*, vol. 33, no. 4, pp. 49–61, Jul. 2016.
[7] M. N. Malik, M. A. Azam, M. Ehatisham-Ul-Haq, W. Ejaz, and A. Khalid, "ADLAuth: Passive authentication based on activity of daily living using heterogeneous sensing in smart cities," *Sensors*, vol. 19, no. 11, p. 2466, May 2019.
[8] M. Yamada, K. Kamiya, M. Kudo, H. Nonaka, and J. Toyama, "Soft authentication and behavior analysis using a chair with sensors attached: Hipprint authentication," *Pattern Anal. Appl.*, vol. 12, no. 3, pp. 251–260, Sep. 2009.
[9] I. Bouchrika, M. Goffredo, J. Carter, and M. Nixon, "On using gait in forensic biometrics," *J. Forensic Sci.*, vol. 56, no. 4, pp. 882–889, Jul. 2011.
[10] E. R. H. P. Isaac, S. Elias, S. Rajagopalan, and K. S. Easwarakumar, "Template-based gait authentication through Bayesian thresholding," *IEEE/CAA J. Autom. Sinica*, vol. 6, no. 1, pp. 209–219, Jan. 2019.
[11] J. P. Singh, S. Jain, S. Arora, and U. P. Singh, "Vision-based gait recognition: A survey," *IEEE Access*, vol. 6, pp. 70497–70527, 2018.
[12] S. Sprager and M. B. Juric, "An efficient HOS-based gait authentication of accelerometer data," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 7, pp. 1486–1498, Jul. 2015.
[13] D. Gafurov, E. Snekkenes, and P. Bours, "Gait authentication and identification using wearable accelerometer sensor," in *Proc. IEEE Workshop Automat. Identificat. Adv. Technol.*, Jun. 2007, pp. 220–225.
[14] F. Luo, S. Poslad, and E. Bodanese, "Human activity detection and coarse localization outdoors using micro-Doppler signatures," *IEEE Sensors J.*, vol. 19, no. 18, pp. 8079–8094, Sep. 2019.
[15] A. Balleri, K. Chetty, and K. Woodbridge, "Classification of personnel targets by acoustic micro-Doppler signatures," *IET Radar, Sonar Navigat.*, vol. 5, no. 9, pp. 943–951, Dec. 2011.
[16] G. Garreau *et al.*, "Gait-based person and gender recognition using micro-Doppler signatures," in *Proc. IEEE Biomed. Circuits Syst. Conf. (BioCAS)*, Nov. 2011, pp. 444–447.
[17] S. Z. Gurbuz and M. G. Amin, "Radar-based human-motion recognition with deep learning: Promising applications for indoor monitoring," *IEEE Signal Process. Mag.*, vol. 36, no. 4, pp. 16–28, Jul. 2019.
[18] W. Ye and H. Chen, "Human activity classification based on micro-Doppler signatures by multiscale and multitask Fourier convolutional neural network," *IEEE Sensors J.*, vol. 20, no. 10, pp. 5473–5479, May 2020.
[19] Z. Chen, G. Li, F. Fioranelli, and H. Griffiths, "Personnel recognition and gait classification based on multistatic micro-Doppler signatures using deep convolutional neural networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 5, pp. 669–673, May 2018.
[20] P. Cao, W. Xia, M. Ye, J. Zhang, and J. Zhou, "Radar-ID: Human identification based on radar micro-Doppler signatures using deep convolutional neural networks," *IET Radar, Sonar Navigat.*, vol. 12, no. 7, pp. 729–734, Jul. 2018.
[21] B. Vandersmissen *et al.*, "Indoor person identification using a low-power FMCW radar," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 7, pp. 3941–3952, Jul. 2018.
[22] Y. Yang, C. Hou, Y. Lang, G. Yue, Y. He, and W. Xiang, "Person identification using micro-Doppler signatures of human motions and UWB radar," *IEEE Microw. Wireless Compon. Lett.*, vol. 29, no. 5, pp. 366–368, May 2019.
[23] Z. Meng *et al.*, "Gait recognition for co-existing multiple people using millimeter wave sensing," in *Proc. AAAI Conf. Artif. Intell.*, vol. 34, 2020, pp. 849–856.
[24] K. Saho, K. Inuzuka, and K. Shioiri, "Person identification based on micro-Doppler signatures of sit-to-stand and stand-to-sit movements using a convolutional neural network," *IEEE Sensors Lett.*, vol. 4, no. 3, Mar. 2020, Art. no. 3500304.
[25] T. Sakamoto, "Personal identification using a convolutional neural network and ultra-wideband radar measurement of walking and sitting motion," (in Japanese), *IEICE Trans. Electron.*, vol. J103-C, no. 7, pp. 321–330, Jul. 2020. [Online]. Available: https://arxiv.org/abs/2008.02182
[26] S. Lahrache, A. El Qadi, and R. El Ouazzani, "Bag-of-features for image memorability evaluation," *IET Comput. Vis.*, vol. 10, no. 6, pp. 577–584, Sep. 2016.
[27] R. U. Khan, X. Zhang, R. Kumar, and E. O. Aboagye, "Evaluating the performance of ResNet model based on image recognition," in *Proc. Int. Conf. Comput. Artif. Intell. (ICCAI)*, 2018, pp. 86–90.
[28] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
[29] C. Karabacak, S. Z. Gurbuz, A. C. Gurbuz, M. B. Guldogan, G. Hendeby, and F. Gustafsson, "Knowledge exploitation for human micro-Doppler classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 10, pp. 2125–2129, Oct. 2015.
[30] S.-H. Wang and Y.-D. Zhang, "DenseNet-201-based deep neural network with composite learning factor and precomputation for multiple sclerosis classification," *ACM Trans. Multimedia Comput., Commun., Appl.*, vol. 16, no. 2s, pp. 1–19, Jul. 2020.
[31] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 618–626.
[32] Y. Sun, L. Zhu, G. Wang, and F. Zhao, "Multi-input convolutional neural network for flower grading," *J. Electr. Comput. Eng.*, vol. 2017, Jun. 2017, Art. no. 9240407.
[33] Z. Guo, X. Li, H. Huang, N. Guo, and Q. Li, "Medical image segmentation based on multi-modal convolutional neural network: Study on image fusion schemes," in *Proc. IEEE 15th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2018, pp. 903–907.
[34] J. F. Item-Glatthorn, N. C. Casartelli, and N. A. Maffiuletti, "Reproducibility of gait parameters at different surface inclinations and speeds using an instrumented treadmill system," *Gait Posture*, vol. 44, pp. 259–264, Feb. 2016.
[35] S. Viteckova *et al.*, "The repeatability of the instrumented timed up & go test: The performance of older adults and Parkinson's disease patients under different conditions," *Biocybern. Biomed. Eng.*, vol. 40, no. 1, pp. 363–377, 2020.
[36] F. Fioranelli, S. A. Shah, H. Li, A. Shrestha, S. Yang, and J. Le Kernec. (2019). Radar signatures of human activities. Enlighten: Research Data. [Online]. Available: http://researchdata.gla.ac.uk/848/

Kenshi Saho (Member, IEEE) received the B.E. degree in electrical and electronic engineering and the M.I. and Ph.D. degrees in communications and computer engineering from Kyoto University, Kyoto, Japan, in 2008, 2010, and 2013, respectively.

From 2014 to 2017, he was an Assistant Professor with the Department of Electronic and Computer Engineering, Ritsumeikan University, Japan. Since 2017, he has been a Junior Associate Professor with the Department of Intelligent Robotics, Toyama Prefectural University, Japan. His research interests include signal processing techniques for radar imagers, sensor fusion systems for positioning and tracking, and micro-Doppler sensors.

Dr. Saho received the Young Scientist Award from the Funai Foundation for Information Technology in 2017, the Researcher's Award from the IEICE Intelligent Transportation Systems Technical Committee in 2016, the Excellent Oral Presentation Award from the ICIAS2015, the Young Researcher's Award from the Radiation Science Society of Japan in 2013, and the Best Paper Award from the ISAP2012.

Keitaro Shioiri received the B.E. degree in intelligent systems design engineering from Toyama Prefectural University, Imizu, Japan, in 2019, where he is currently pursuing the M.E. degree.

His research interests include personal identification and biomechanical analysis of human movements using micro-Doppler radars based on machine learning approach.

Keisuke Inuzuka received the B.E. degree in intelligent robotics from Toyama Prefectural University, Imizu, Japan, in 2020.

When he was in undergraduate school, his research interest was in microwave radar data processing for personal identification. He received the Excellent Presentation Award from the Electronics, Information and Systems Society of the Institute of Electrical Engineers of Japan (IEEJ) in 2020.