Classification of human body motions using an ultra-wideband pulse radar

Hui-Sup Cho* and Young-Jin Park
Division of Electronics and Information System, DGIST, Daegu, Korea

Abstract.
BACKGROUND: The motion or gestures of a person are primarily recognized by detecting a specific object and the change in its position from image information obtained via an image sensor. However, the use of such systems is limited due to privacy concerns.
OBJECTIVE: To overcome these concerns, this study proposes a radar-based motion recognition method.
METHODS: Detailed human body movement data were generated using ultra-wideband (UWB) radar pulses, which provide precise spatial resolution. The pulses reflected from the body were stacked to reveal the body's movements and these movements were expressed in detail in the micro-range components. The collected radar data with emphasized micro-ranges were converted into an image. Convolutional neural networks (CNN) trained on radar images for various motions were used to classify specific motions. Instead of training the CNNs from scratch, transfer learning is performed by importing pretrained CNNs and fine-tuning their parameters with the radar images. Three pretrained CNNs, Resnet18, Resnet101, and Inception-Resnet-V2, were retrained under various training conditions and their performance was experimentally verified.
RESULTS: As a result of various experiments, we conclude that detailed motions of subjects can be accurately classified by utilizing CNNs that were retrained with images obtained from the UWB pulse radar.

1. Introduction

Radar technology is becoming an indispensable element not only in the aviation and military fields but also in the automotive field for advanced driver-assistance systems (ADAS). In recent years, in addition to the traditional radar applications, attempts have been made to use radar to obtain various information from the human body [1,2], and studies have been carried out to distinguish the movements of human body using radar technology [3]. Currently, human body movements are distinguished and analyzed mainly using image sensors. However, due to the concerns about being used for monitoring individuals that could lead to infringement of privacy, radar technology is drawing attention as a means to replace the image sensors.

This study suggests radar technology as an alternative to overcome these concerns and proposes a novel motion classification method using a pulse radar. Radar can be classified into continuous wave (CW),

---

*Corresponding author: Hui-Sup Cho, Division of Electronics and Information System, 33, Techno Jungang Daero, Dalseong-Gun, Daegu, Korea. E-mail: mozart73@dgist.ac.kr.

which continuously emits waves with constant or modulated frequency, and pulse radar, which emits pulses of finite duration. CW radar has limitations in that it can only detect moving objects, as the change in the moving speed of the target is expressed by the change in Doppler shift frequencies. As a result, since CW radar does not provide the information about the distance to the target, it has fundamental limitations in describing movements and distinguishing motions. On the other hand, the pulse radar measures the distance to the target with high spatial resolution and can describe the precise movement of the target. In addition, pulse radar can detect a target regardless of whether the target is moving. In particular, this study aims to classify the motions of workers who perform tasks in a limited space, because motion classification technology that can replace image-based methods has been required for smart factory and process automation applications. Specifically, there are applications such as monitoring the movement of workers to find hazards and measure workload, and automating part supply by measuring how often an operator accesses a specific part in the assembly lines.

In this study, an ultra-wideband (UWB) pulse radar was used to maximize the superiority in terms of spatial resolution among pulse radars. In order to take advantage of the UWB radar, radar pulses reflected from the target were accumulated and converted into images to observe how the radar pulse changes when the body, head, and limbs move over time. Prior to that, a process was made to emphasize the characteristics of each movement as much as possible, and as a result, this process directly contributed to the expression of unique characteristics corresponding to specific movements in each image. Then, classifying each motion was performed using deep learning technology. Since advances in deep learning technology using the convolutional neural networks (CNN) has enabled accurate feature detection and classification [4], data sets consisting of the images corresponding to several motions and various kinds of CNNs were used to classify motions in this study.

In the experiments, radar images were created for five motions: "bending over and straightening the upper body," "swinging the arms," "sitting down and standing up," "stretching the arms up and down," and "turning the upper body to one direction." Then, these images were used to retrain the pretrained "Resnet18," "Resnet101," and "Inception-Resnet-V2" models, and verify the ability of the trained networks to distinguish the actions. As a result, classification of motions using the radar images and deep learning with Resnet18, Resnet101 [5], and Inception-Resnet-V2 [6] showed the accuracies of 99.40%, 99.20% and 96.40%, respectively.

2. Comparable studies

There have been attempts to distinguish human body movements using the CW radar. Kim et al. [3] proposed a method of classifying motions such as running, walking, crawling, and boxing in front of an antenna using the CW micro-doppler radar and machine learning. In the study of Kim et al., spectrograms were generated from CW radar signals to extract features for classifying six human activities, and six kinds of features were extracted within the spectrogram. Then, after the data sets were constructed with a combination of features, support vector machine (SVM) [7], a binary classifier, was used to classify each activity. According to the result of the test, it was reported that the classification accuracy was 91.4%. This study observed the Doppler shift frequency caused by the change in the speed at which body parts move and focused on classifying large motions by observing how the Doppler shift frequency changes over time. As mentioned above, since CW radar does not provide distance information to the target, it creates the effect of compressing the space in front of the radar so that all targets seem to be on one vertical plane in front of the antenna and all the movements of the targets appear to happen on this plane. Consequently, these limitations make it impossible to express subtle movements of human body.
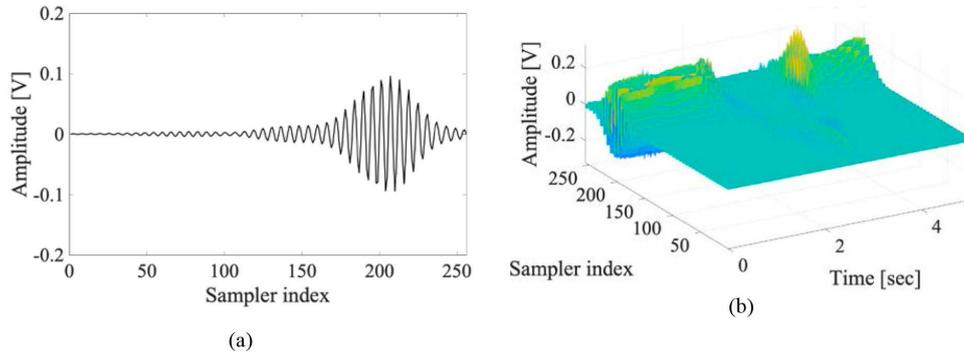
(a)

(b)

Fig. 1. Shape of (a) a frame and (b) a frame set.

In contrast, observing the collection of pulses acquired from moving subjects using the pulse radar reveals not only the range components, which are the components with a large amplitude reflected from the body, but also the micro-range components with a small amplitude reflected from the arms, legs, shoulder, head and so on. Since the micro-range component is an important factor describing the movements of the body parts, by using this feature, it is possible to classify detailed and specific movements of human body. It is also a feature that cannot be obtained with CW radar. Although it is difficult to directly compare the proposed method with the one presented in the comparable studies, it is clear that whereas CW radar is a technology that makes it possible to distinguish large motions, pulsed radars allow detailed motions to be classified.

3. Methodology

3.1. Radar system and signal

In this study, a UWB pulse radar, which utilizes extremely narrow pulses, was used. The shape of the electromagnetic wave used by the radar, in the time domain, is a sinusoidal wave with a Gaussian envelope, and the pulse width is less than 0.4 ns. An example of the pulse is shown in Fig. 1a. In terms of the characteristics in the frequency domain, the center frequency is 6.8 GHz and the bandwidth is 2.3 GHz, which occupies a fairly wide bandwidth. The effective isotropic radiated power (EIRP) of the pulse emitted by the radar is -12.6 dBm, which satisfies the radio emission safety guidelines. In the experiment for this study, the subject is placed within the range of 2 meters in consideration of the emission power and the signal-to-noise ratio (SNR) of the radar pulse. The emission frequency of the radar pulse is set to 100 MHz and the acquisition frequency is set to 50 Hz.

Radar pulses collected from the antenna are sampled by 256 samplers in the RF transceiver and converted into a set of 256 samples, which is called a "frame." Radar frames are stacked at regular time intervals (20 ms) to form a "frame set" consisting of 256 frames, which is used throughout data processing as the basic unit for further processing. Sampler index is a unit corresponding to the distance from the antenna to the target.

In the frame shown in Fig. 1a, the horizontal axis is the index representing each sampler of the transceiver. It is also a unit corresponding to the distance from the antenna to the target. The smaller the number, the closer the target is to the antenna. Figure 1b shows an exemplary frame set made of radar pulses reflected from a subject who bends over and straightens the upper body, and Fig. 2 shows the top
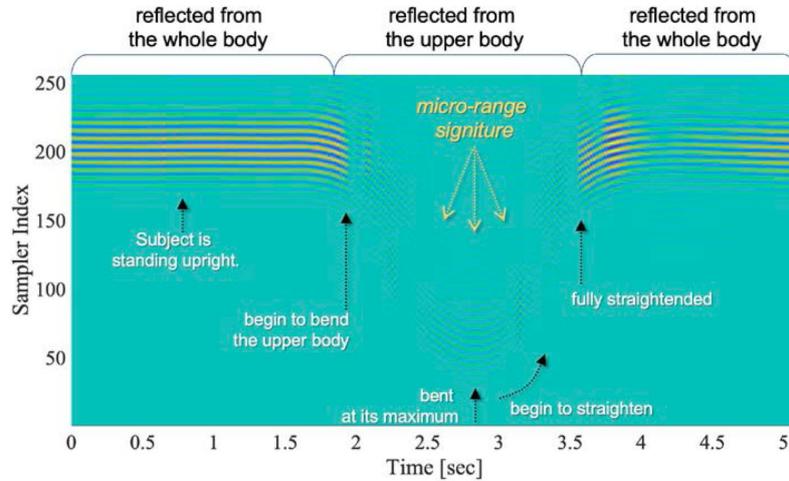
Fig. 2. Top view of the frame set obtained from a subject bending over and straightening the upper body.

view of this frame set. In Figs 1b and 2, the pulse near the 200th sampler index has a large amplitude because it was reflected from a whole body. The small-sized components found in the 2 to 4 second interval of the 20th to 170th sampler indexes are reflected from the head and shoulders while the upper body is bent, and these are the components called "micro-range signature". This study utilizes the shape of micro-range that appears differently for each body movement. As shown in detail in Fig. 2, it can be clearly seen that when the upper body is bent and straightened, the micro-range component reflected from the head and shoulders which have narrow areas for electromagnetic wave reflection is weakly displayed. In addition, it is confirmed that the entire process of bending over and straightening the upper body is well expressed through the micro-range signature.

3.2. Deep learning for motion classification

Machine learning is the ability of an artificial intelligence system to perform complicated tasks by learning through input data without pre-defined rules for reasoning; it is used in several fields [8]. Machine learning models accept a variety of input data, and they create and provide complex inferences based on the relationships between the inputs [9]. Deep learning is a sub-discipline of machine learning that consists of multiple layers to create complex and high-level representations based on input data. Upon looking at the structure of a deep learning model, the concept of a machine learning model called artificial neural network (ANN) inevitably appears, as deep learning is used to refer to a new ANN that goes beyond the limitations of the existing ANN. ANN comprises several layers of neurons, and it operates by multiplying the input by weight, adding bias, and sending the result through an activation function to the next neuron.

The most widely used deep learning model is the convolutional neural network (CNN) [4], which is used in this study. While CNN inherits the features of ANN, it is generally deep because it is composed of many layers. CNN is composed of many layers, and in general, as the number of layers increases, the meaning of the output becomes more complex and richer [10]. It basically includes a convolution layer, a pooling layer, and a fully-connected layer. The convolution layer is a layer used for convolution with the input image, which can be seen as implementation of an image filter. It plays a role of extracting features from images, for example, extracting shapes that describe detailed structures from simple edges.

Table 1
Comparison of pretrained CNNs used in the study

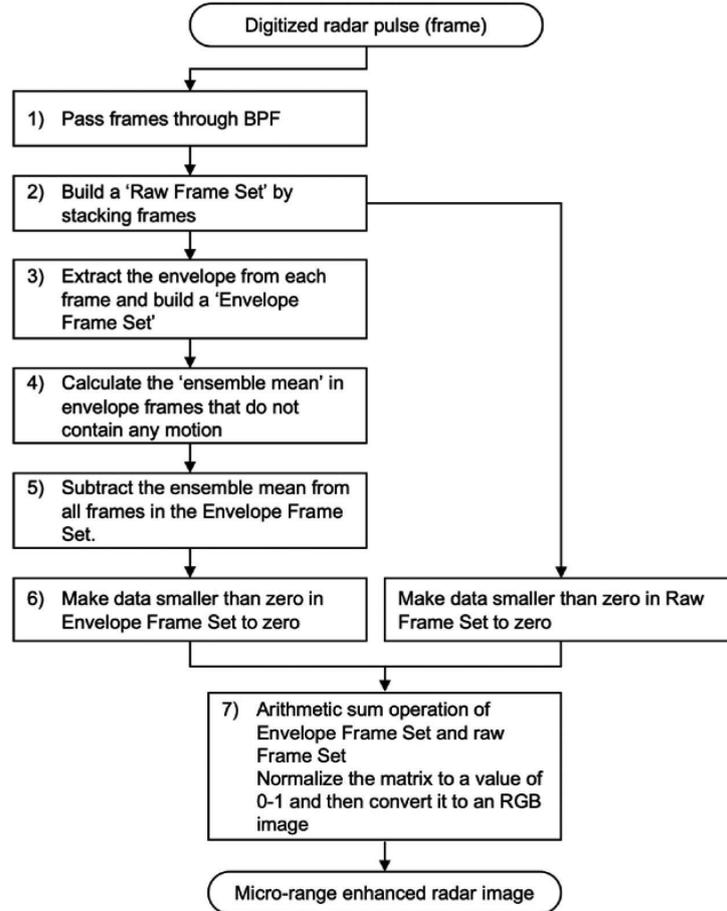| Network model	Depth	Parameters (Millions)	Input image size
Resnet18	18	11.7	224×224
Resnet101	101	44.6	224×224
Inception-Resnet-V2	164	55.9	299×299

Fig. 3. Process to acquire the micro-range enhanced images.

The fully connected layer plays a role of generating a specific decision using features extracted in the convolution layer. In this study, it distinguishes what kind of motion the input radar image represents in the decision-making process. The pooling layer makes the representations smaller and more manageable by performing the role of downsampling the results of the convolution layer. A set of images is given as an input to the network, and training of the network is accomplished through an iterative process of calculating the error between the predicted value of the network and the desired output, and using this to adjust the weights of the network in the direction of reducing the error.

Transfer learning is widely used in deep learning applications. It is conducted by bringing in a pretrained network and fine-tuning it for the application, which is much faster and more efficient than training the network from scratch using randomly initialized weights and biases [11]. In general, there is a concern
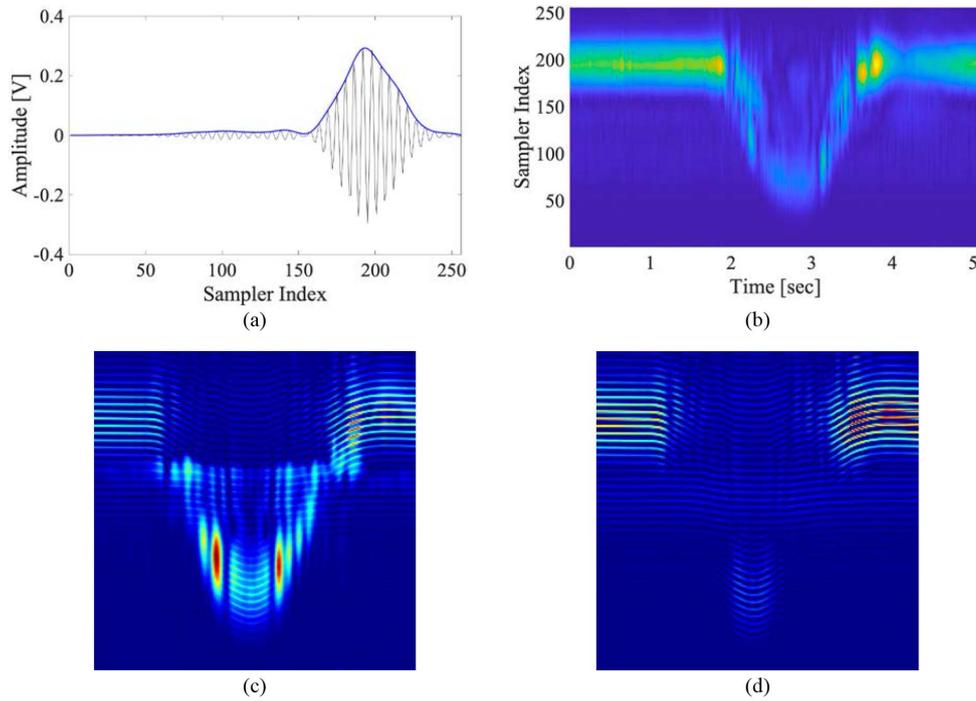
Fig. 4. Shape of (a) frame and its envelope, (b) envelope frame set, (c) micro-range enhanced radar image, and (d) raw radar image.

that overfitting may occur in the network when training is performed with a small number of data sets. Use of many pretrained layers by utilizing transfer learning reduces the risk of overfitting. In this study, the pretrained network models of CNN, Resnet18, Resnet101, and Inception-Resnet-V2 models are used. Table 1 shows the number of layers constituting the three network models used in the study, the size of the input image, and the number of parameters that require training.

3.3. Method to create radar images containing the features of motions

A subject is asked to repeat five specific motions and the pulses reflected from the human body are collected. This study proposes a process of dealing with radar signals reflected from the body, extracting motion features characterized by the micro-range components, and converting the radar data into an image for being used for machine learning. The process to acquire the micro-range enhanced images is shown in Fig. 3.

1) Since the radar pulse collected from the antenna contains DC components and noises, each frame is passed through a band-pass filter (BPF) that selectively passes a signal of 2.3 GHz bandwidth around 6.8 GHz. Then, a raw frame set, which is a 2D matrix of 256×256 in size, is created by stacking the frames which have passed through the BPF at regular time intervals (20 ms).
2) Movements of the head, arms, and legs generate micro-range components, and amplitude of micro-range components is smaller than that of range component. Thus, it is necessary to compensate for their amplitude.
3) An "envelope frame set" is constructed by extracting the envelope of pulse peaks from each frame in the raw frame set. Figure 4a shows the frame from Fig. 1a together with its envelope.
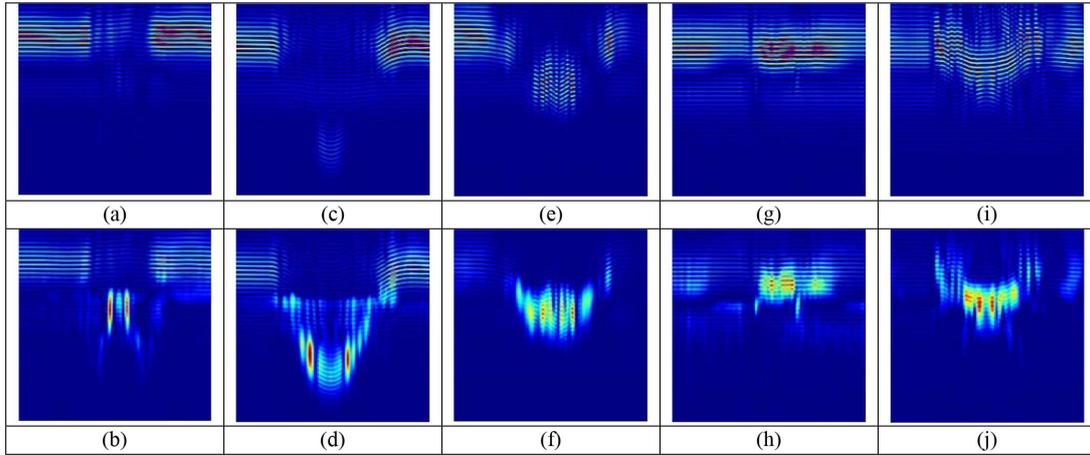
Fig. 5. (a) Raw radar image and (b) micro-range enhanced radar image correspond to "swinging arms"; (c) raw radar image and (d) micro-range enhanced radar image correspond to "bend/straighten upper body"; (e) raw radar image and (f) micro-range enhanced radar image correspond to "sit down/up"; (g) raw radar image and (h) micro-range enhanced radar image correspond to "stretch arms"; and (i) raw radar image and (j) micro-range enhanced radar image correspond to "turn upper body".

4) Separately, after extracting the envelope frame set in the same manner as 3) from the frame set collected while the subject is still without motion, their ensemble mean is calculated.
5) Next, the ensemble mean is subtracted from all envelopes in the envelope frame set. The resulting frame set shows only the area where the micro-range appears in the raw frame set, so this is used to compensate the amplitude of the micro-range component. Figure 4b shows the envelope frame set extracted from Fig. 1b in the top view.
6) As shown in Fig. 1a, a raw frame is a bipolar signal, and such characteristic is maintained even after passing through several filters. Since the positive and negative parts are in symmetry, information loss does not occur even if only a component of one polarity is used. In this study, only the positive parts of the raw frame set and envelope frame set are used. This process makes it possible to create full-scale (full dynamic range) images with half range of the data when imaging the radar data.
7) The raw frame set and envelope frame set composed of only positive values are added by arithmetic matrix operation. Then, this is normalized and converted into an RGB image, which is called a "micro-range enhanced radar image". An RGB image created with a raw frame set is called a "raw radar image." All images are labeled with corresponding actions. The micro-range enhanced radar image and raw radar image created with the frame set from Fig. 1b are shown in Fig. 4c and d, respectively. The micro-range components that are not quite visible in the middle part of Fig. 4d are clearly visible in Fig. 4c.

The micro-range enhanced radar images corresponding to the five motions are shown in Fig. 5 along with the raw radar images.

4. Experiment and result

4.1. Setup for radar signal acquisition

Figure 6 shows the setup of the experiment for the study. The UWB pulse radar was equipped with an RF transceiver, namely, NVA6201 [12], with transmitting (Tx) and receiving (Rx) antennas. The radar

Table 2
Five motions under study

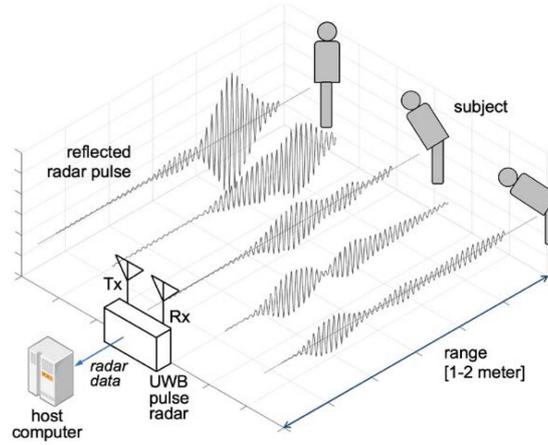| Motion label	Description
Swing arms	Swinging the arms horizontally
Bend/straighten upper body	Bending over and straightening the upper body
Sit down/up	Sitting down and standing up
Stretch arms	Stretching the arms up and down
Turn upper body	Turning the upper body to one direction

Fig. 6. Experimental setup.

system was connected to the host computer via serial communication. The radar pulses are emitted by the transmitting antenna Tx, and the pulses reflected from the human body are collected by the receiving antenna Rx. Subjects are positioned in the range of 1.0 to 2.0 meters from the antenna and asked to make five motions: "bending over and straightening the upper body," "swinging the arms," "sitting down and standing up," "stretching the arms up and down," and "turning the upper body to one direction." The labels for the five motions and their descriptions are shown in Table 2. The host computer controls the radar system, receives data from the radar system, and runs the proposed algorithm and deep learning. Since the commercial radar module used in this study complies with the safety guidelines and emits relatively low power pulses, the level of noise increases to a level where the signal cannot be used when the target is at a distance of more than 2 meters. In the actual application, the power could be adjusted to increase the range between the antenna and target.

4.2. Design of experiments to train and test convolutional neural networks

A "raw radar image set" and a "micro-range enhanced radar image set" were used for deep learning and classification. The raw radar image set was a reference data set to prove the effectiveness of the proposed method. In the data set used in this experiment, the data corresponding to five motions, which are the five classes, were evenly distributed by the same number so that "accuracy" could be used as an evaluation metric. Accuracy is defined as the ratio of the number of images correctly predicted by the network model to the total number of images used for evaluations.

Transfer learning using pretrained CNN can be categorized into two methods: using pretrained CNN as a fixed feature extractor and fine tuning the entire network. The former method replaces the last learnable

Table 3
Test result of Resnet18

Input	Ratio of frozen layers [%]	0.0	39.4	62.0	84.5	94.4
Raw radar image	Validation accuracy	1.0000	0.9867	0.9733	0.9733	0.9667
	Test accuracy	**0.9680**	0.9480	0.9560	0.9240	0.8960
Micro-range enhanced	Validation accuracy	1.0000	0.9933	0.9933	0.9667	0.9400
radar image	Test accuracy	**0.9840**	0.9680	0.9640	0.9200	0.8880

layer and the final classification layer with the new layers to make new classification, and newly trains them with the training set. The latter method finely tunes the entire layers of the network, or freezes the layers (make layers excluded when training with a new train set) close to the initial layer and finely tunes the remaining layers. In the experiment, in order to analyze the correlation between the number of frozen layers and accuracy, training and testing were repeated while gradually increasing the number of frozen layers. Since the layer structure of each CNN model was different, freezing was not performed for the same number of layers. Resnet18, Resnet101, and Inception-Resnet-V2 had 4, 6, and 3 steps to increase the number of layers to freeze according to the layer structure, respectively. In addition, the experiment was also repeated while changing the number of epochs, which is the number of training iterations. A train set consisting of 1000 images was created by selecting 200 images for each motion from the image data set. Among them, 70% of the images were allocated for training and 30% for validation. Then, a test set consisting of 500 images was made by selecting 100 images for each motion from the image data sets that are not used for training and validation. Data augmentation is used to prevent the network from remembering the details of the training images and the risk of overfitting. Augmented images are created through a random combination of resizing and translation transformations of training images. The resizing range is 80 to 120% and the translation range is (-32) to 32 pixels.

4.3. Experimental results

In order to retrain the pretrained CNN using the raw radar images and the micro-range enhanced radar images, it is necessary to set the hyper parameters. As part of an effort to compare performance by training three CNNs under the same conditions as possible, the hyper parameters, except the epoch number and learning rate, were set to the same values prior to retraining each CNN. In setting the hyper parameter for network learning, the mini batch size was set to 15, L2 Regularization with a factor of 1e-4 was used, and stochastic gradient descent with momentum was used as the optimizer.

4.4. Training and testing of Resnet18

Since Resnet18 has 18 layers, it took the shortest time to train. The learning rate is set to 3e-4, and it stays unchanged during training. The results obtained at 40 iterations showed the highest test accuracy and these were listed in Table 3. In the table, the 'ratio of frozen layers' represents the ratio of the number of frozen layers to the total number of layers in the network, and layers within the range from the initial layer to the position corresponding to a specific ratio were frozen. For example, since 39.4% of 18 layers corresponds to the 7th layer, it means that the layers from the initial layer to the 7th layer were frozen, and parameters belonging to those layers were not updated during the training process. According to Table 3, the network retrained using the micro-range enhanced images showed the highest test accuracy of 98.40% when the entire layers were retrained.

Table 4
Test result of Resnet101

Input	Ratio of frozen layers [%]	0.0	19.6	40.4	60.5	80.7	98.8
Raw radar image	Validation accuracy	0.9733	0.9933	0.9533	0.9533	0.9333	0.9133
	Test accuracy	0.9440	**0.9160**	0.8400	0.8360	0.8400	0.8520
Micro-range enhanced	Validation accuracy	1.0000	0.9933	0.9867	0.9867	0.9800	0.9600
radar image	Test accuracy	0.9800	**0.9920**	0.9800	0.9840	0.9760	0.9600

Table 5
Test result of Inception-Resnet-V2

Input	Ratio of frozen layers [%]	0.0	34.6	64.8
Raw radar image	Validation accuracy	99.33	94.00	86.00
	Test accuracy	**92.80**	84.40	78.00
Micro-range enhanced	Validation accuracy	98.67	98.00	94.00
radar image	Test accuracy	**96.40**	95.20	88.00

4.5. Training and testing of Resnet101

The learning rate is set to 1e-4 and other setting of hyper parameters for retraining Resnet101 is the same as that of Resnet18, except for the number of epochs. The results obtained at 30 iterations showed the highest test accuracy and these were listed in Table 4. According to the result, the network retrained using the micro-range enhanced images showed the highest test accuracy of 99.20% when the first 19.6% of the layers were frozen.

4.6. Training and testing of Inception-Resnet-V2

The setting of hyper parameters for retraining the network is the same as that of Resnet101, except for the number of epochs. Inception-Resnet-V2 is a deep network consisting of 164 layers with 55.9 million parameters to train. Therefore, more training iterations were required to train than the previous two CNNs. The results obtained at 150 iterations showed the highest test accuracy and these were listed in Table 5. The accuracy was saturated to a constant value even if the number of iterations was increased to above 150. According to the result, the network retrained using the micro-range enhanced images showed the highest test accuracy of 96.40% when the entire layers were retrained.

The classification accuracy was higher when the micro-range enhanced radar images were used than when the raw radar images were used, which was a common result for all three CNN models. This indicates that the micro-range components described the characteristics of each motion and could be used to classify human motions. Resnet18 and Inception-Resnet-V2 showed the highest accuracy when the entire layers were finely retrained. However, Resnet101 showed the highest accuracy when 19% of the front layers were frozen. This is due to the fact that the radar images have unique characteristics that are different from the natural images used in the pretraining, so the number of layers that can extract unique features of radar images decreased as the number of frozen layers increased. Therefore, it can be seen that the use of radar images for transfer learning requires a different approach than that for normal images. The difference between validation and test accuracy is a factor that determines how closely the network fits into the input image set. The difference with the raw radar images was larger than that with the micro-range enhanced radar images in all three networks. This means that when the AI model was trained with the micro-range enhanced radar images, the possibility of overfitting could be lowered.

In addition, although the direct comparison between the comparable studies using the CW radar and the proposed method does not seem to be reasonable, it can be seen that more specific motions can be

classified with higher accuracy by the proposed method. Furthermore, even the classification accuracy seen in the test results using only raw radar images appears to be superior to that of the comparable study, which proves that the pulsed radar is a more suitable method for expressing the minute movements of the human body.

5. Conclusion

Recently, attempts have been made to distinguish human body motions using radar. Such radar-based applications can be an alternative to the existing applications based on image sensors that are raising concerns about infringement of privacy and being used for control purposes. In particular, the observation and analysis of the behavior of workers performing repeated motions in a designated space were considered as suitable fields of application of the proposed method, and classification of five representative motions was performed to explore the possibility of such applications in this study. There have been already similar attempts using CW radars, but CW radar shows a fundamental limitation in being used to classify detailed motions because it only provides information about the speed of the target's movement. In contrast, this study proposes a method of classifying human motions using a UWB pulse radar. Since the UWB pulse radar provides information on the distance to the target with a spatial resolution of several millimeters, it can be used for discriminating subtle differences in specific motions of people. This is enabled by building a frame set through accumulating radar pulses reflected from the human body at regular intervals and then observing them in the top-view. In the frame set, the change in the shape of the pulse due to the Doppler shift caused by the movement of each part of the body can be observed, therefore, not only the components of the large amplitude reflected from the torso, but also the micro-range components due to the movements of the arms, legs, and head are revealed. Especially, the radar frame set is processed to enhance the micro-range components that provide detailed information necessary for classifying various motions. Then, this is converted into a micro-range enhanced radar image, which is an RGB image, and is used in the construction of a data set for deep learning. Deep learning was used to classify motions such as "bending over and straightening the upper body," "swinging the arms," "sitting down and standing up," "stretching the arms up and down," and "turning the upper body to one direction" with radar images. In particular, this was implemented by transfer learning for a CNN model that was pretrained with a large number of natural images. The CNNs used in this study included Resnet18, Resnet101 and Inception-Resnet-V2 pretrained with a vast data set from ImageNet. As a result of the experiments, in all three CNNs, the test accuracy measured when the network was trained and tested with the micro-range enhanced radar images was higher than that measured when the raw radar images were used. Such results indicate that the micro-range components are the features that can be used for classification of each motion, and the method of creating a radar image with enhanced micro-range components is useful and effective. In Resnet18, Resnet101 and Inception-Resnet-V2, the accuracy of classifying the five motions was 98.40%, 99.20% and 96.40%, respectively. Resnet101 showed excellent accuracy even with a short training time and a small number of iterations. Moreover, the difference between the validation accuracy and test accuracy was also the smallest, which means it was robust against overfitting. The results confirm that human body motions can be classified with high accuracy by utilizing the images with enhanced micro-range features created from the radar signals and the CNNs retrained with those images.

Acknowledgments

Conflict of interest

None to report.

References

[1]   Cho HS, Park YJ. Measurement of pulse transit time using ultra-wideband radar. Technology and Health Care. 2020; Pre-press: 1-10. doi: 10.3233/THC-202626.
[2]   Cho HS, Choi BD, Park YJ. Monitoring heart activity using ultra-wideband radar. Electronics Letters. 2019; 55: 878-881.
[3]   Kim Y, Moon T. Human detection and activity classification based on micro-doppler signatures using deep convolutional neural networks. IEEE Geoscience and Remote Sensing Letters. 2016; 13: 8-12.
[4]   Valueva MV, Nagornov NN, Lyakhov PA, et al. Application of the residue number system to reduce hardware costs of the convolutional neural network implementation. Mathematics and Computers in Simulation. 2020; 177: 232-243.
[5]   He K, Zhang X, Ren X, Sun J. Deep residual learning for image recognition, Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016; 770-778.
[6]   Szegedy C, Ioffe S, Vanhoucke V, Alemi AA. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. Proceedings of the 31st AAAI Conference on Artificial Intelligence. 2017; 4278-4284.
[7]   Cristianini N, Shawe-Taylor J. An Introduction to Support Vector Machines and Other Kernel-based Learning Methods. Cambridge Uinv. Press, 2000.
[8]   Goodfellow I, Bengio Y, Courville A. Deep Learning. The MIT Press; 2016. (ISBN: 978-0262035613)
[9]   McBee MP, Awan OA, Colucci AT, et al. Deep learning in radiology. Academic Radiology. 2018; 25(11): 1472-1480.
[10] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Communications of the ACM. 2017; doi: 10.1145/3065386.
[11] Rawat W, Wang Z. Deep convolutional neural networks for image classification: A comprehensive review. Neural Computation. 2017; 29: 2352-2449.
[12] NOVELDA. NVA620x preliminary data sheet. 2013; http://www.novelda.no.