

Research Article

Through-Wall Radar Classification of Human Posture Using Convolutional Neural Networks

Alper Kılıç ¹, İsmail Babaoğlu ¹, Ahmet Babalık ¹ and Ahmet Arslan²

¹Department of Computer Engineering, Konya Technical University, Konya, Turkey

²Department of Computer Engineering, Konya Food and Agriculture University, Konya, Turkey

Correspondence should be addressed to İsmail Babaoğlu; ibabaoglu@selcuk.edu.tr

Received 19 November 2018; Revised 22 January 2019; Accepted 20 February 2019; Published 31 March 2019

Academic Editor: Angelo Liseno

Copyright © 2019 Alper Kılıç et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Through-wall detection and classification are highly desirable for surveillance, security, and military applications in areas that cannot be sensed using conventional measures. In the domain of these applications, a key challenge is an ability not only to sense the presence of individuals behind the wall but also to classify their actions and postures. Researchers have applied ultrawideband (UWB) radars to penetrate wall materials and make intelligent decisions about the contents of rooms and buildings. As a form of UWB radar, stepped frequency continuous wave (SFCW) radars have been preferred due to their advantages. On the other hand, the success of classification with deep learning methods in different problems is remarkable. Since the radar signals contain valuable information about the objects behind the wall, the use of deep learning techniques for classification purposes will give a different direction to the research. This paper focuses on the classification of the human posture behind the wall using through-wall radar signals and a convolutional neural network (CNN). The SFCW radar is used to collect radar signals reflected from the human target behind the wall. These signals are employed to classify the presence of the human and the human posture whether he/she is standing or sitting by using CNN. The proposed approach achieves remarkable and successful results without the need for detailed preprocessing operations and long-term data used in the traditional approaches.

1. Introduction

The ability to image targets behind building walls or to detect people under debris also including the classification of the human body has been drawing attention since the last decade. For this reason, unlike image processing, ultrawideband (UWB) radars as radio frequency sources more precisely achieve this kind of purpose applicable to real-world problems. UWB radars are used for different applications such as the detection and classification of aircrafts, collision avoidance, detection of a target, or the heart and respiration rate of a human. This kind of radar has several key advantages over narrowband continuous wave radars like having a very high downrange resolution of a target, allowing better separation between targets and clutter due to the large bandwidth; multiple target detection capability; good immunity against multipath interference; and detection of both an object and its position [1]. The concept underlying

through-wall human detection using UWB radars lies on a similar approach with that of radar imaging. A fraction of the transmitted RF signals is traversed from a nonmetal wall, reflected from the objects—even humans, and returned to the receiver imprinted by passing the nonmetal wall again having some signature of the objects within the room. By using this received signal, imaging of the objects is possible [2].

As a form of UWB radar, stepped frequency continuous wave (SFCW) radar approaches are commonly used in many practical applications including through-wall radar imaging and target ranging [3–7], medical imaging [8], and many applications utilizing ground penetrating radar (GPR), a kind of SFCW radar, for civil engineering [9, 10], structural static testing [11], quality estimation of the road surface layer [12], detection of pipes and cables buried in the ground [13, 14], archeological purposes [15], and unexploded ordnance disposal [16, 17]. These studies rely on using SFCW radar signals since they make the spectrum accessible directly to the

user. SFCW radar techniques also have benefits including high mean transmitter power and high receiver sensitivity [2, 4]. Not only do they provide the ability to detect targets or events but SFCW radars also enhance the range accuracy, enable clutter rejection, and help reduce the multipath.

Convolutional neural networks (CNNs) have been used for solving many different artificial intelligence problems, providing significant advantages over other machine learning approaches in solving complex learning tasks. In conventional classification approaches, features were manually extracted and designed and then followed by a traditional classifier such as a support vector machine (SVM). Due to having several feature extraction and signification layers, CNNs are capable of performing automatic preprocessing along with their neural network characteristics [18–21].

In the literature, studies that classify targets by processing radar signals can be divided into two categories: studies based on creating images from radar signals and studies based on extracting different features of the target traditionally. The first category suggests either operating at very high frequencies (e.g., millimeter wave or terahertz) [22] which do not penetrate walls or using the SAR (synthetic aperture radar) algorithm [23]. The second category proposes to extract the specific features of the target such as vital signs [24–26] or movement characteristics [27–29] of the target. Micro-Doppler signatures [30] are commonly used to detect vital signs or to classify the target's specific activities such as running, walking, or even falling. However, the most important feature and disadvantage of this approach is the need for continuous data for a certain period (5-30 secs) for preprocessing and classification. This approach may not be practical for security, counterterrorism, or mission-critical operations where immediate decisions are important. Although the SAR algorithm provides detailed information for the absolute position and shape of the target, there are difficulties in implementing it in practice. Therefore, there is a need to develop methods which have the simplest configuration and can make instant classification without the need for long-term data.

This paper focuses on the classification of the human posture behind the wall using through-wall radar signals and a convolutional neural network (CNN). Thus, due to the advantages of CNN, detailed preprocessing is not required for classification. The SFCW radar is used to collect radar signals reflected from the human target, and these signals are employed to classify the human target whether he/she is standing or sitting by using CNN.

The paper is organized as follows: the SFCW radar concept as a form of UWB radar and CNNs are briefly introduced. After presenting the detection and classification approach of the study, the experimental setup and results are given and discussed. The study is concluded in Section 6, also proposing some future works.

2. SFCW Radar

The SFCW radar is a UWB radar form with advanced features having considerable capabilities for a variety of applications. The main advantage of the SFCW radar is the high

dynamic range and low noise floor. Furthermore, with the ability to avoid certain frequencies for transmission, the SFCW is preferred for certain restricted applications. Depending on these advantages of the SFCW radar, these radar systems are better choices on through-wall imaging due to the range, resolution, and propagation characteristics of UWB signals through a dielectric wall [31, 32]. Detailed information about the SFCW radar can be found in [33].

Assuming that the total period of repeated pulses of the SFCW signal is T , the initial (minimum) frequency is f_0 , the frequency step increment is Δf , and the stepped frequency is N ; then, the stepped frequency $S_t(t)$ may be written as follows [3]:

$$S_t(t) = \sum_{n=0}^{N-1} \text{rect}\left(\frac{t - (n + 1/2) * T}{T}\right) \exp\{j(2\pi(f_0 + n\Delta f)t + \varphi_0)\}, \quad (1)$$

where φ_0 refers to the initial phase of the transmitted signal and

$$\text{rect}\left(\frac{t}{T}\right) = \begin{cases} 1, & -\frac{T}{2} \leq t < \frac{T}{2}, \\ 0, & \text{else.} \end{cases} \quad n = 0, 1, 2, \dots, N - 1, \quad (2)$$

The transmitted signal is reflected and echoed back from the target at the radial distance. The target echo signal $S_r(t)$ may be described as follows:

$$S_r(t) = \sum_{n=0}^{N-1} \text{rect}\left(\frac{t - (n + 1/2) * T - \tau(t)}{T}\right) \rho_n \exp\{j(2\pi(f_0 + n\Delta f)(t - \tau(t)) + \varphi_0)\}, \quad (3)$$

where $\tau(t) = (2R(t))/c$ refers to the time delay of the echoed signal considering a two-way distance of the object at $R(t)$ and c refers to the speed of light. ρ_n in (3) refers to the back-scattering coefficient of the objective. ρ_n is assumed constant and set as uniformly within the observation period of the radar signal.

The maximum unambiguous range that the radar can detect is decided by the step size Δf . The resolution which is the ability to distinguish the two closely spaced targets is determined by the bandwidth ($N\Delta f$).

In radar systems, the signal collected in any point of measuring is called an A-scan (1D data). The received signal obtained in the frequency domain from the entire bandwidth, namely, the A-scan data, is converted into the time domain by performing the Inverse Fast Fourier Transform (IFFT). As the time delay between the transmitted and received signals is related directly to the radial distance of objects in the radar's range, the spatial domain can easily be calculated by using IFFT.

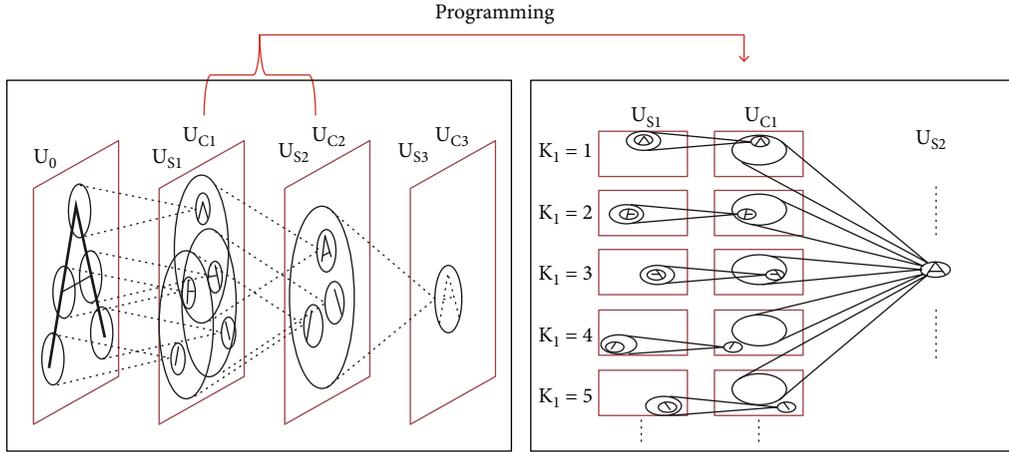


FIGURE 1: The proposed neocognitron architecture by Fukushima [52].

3. Convolutional Neural Networks

This section aims to introduce and clarify some concepts of convolutional neural networks. Detailed definitions of CNNs can be found from the literature [34–37].

A CNN is a class of deep multilayer feed-forward neural network machine learning algorithms that was inspired by the visual cortex of the brain. These network models are based on local receptive fields, shared weights, and spatial or temporal subsamplings that ensure some degree of shift, scale, and distortion invariance [38]. The CNN architecture allows the computer to “see”—recognize images by propagating raw natural images from an input layer, a feature extraction module and a classification module, to class scores in the output layer [39]. The feature extraction module (made up of convolutional and subsampling (pooling) layers) automatically gathers relevant information such as colors, blobs, corners, oriented edges, endpoints, and higher order features through a feature learning process by filtering the input for useful hierarchical abstract information [40].

In the traditional classical machine learning approaches which are used for pattern recognition, a hand-designed feature extractor such as the Histogram of Oriented Gradients (HOG), Bag of Features (BOG), scale invariant feature transform, bank of Gabor filters, Linear Binary Pattern (LBP), and Fisher vectors is used for feature extraction in a domain-specific feature-engineered process. Training such models on natural images would lead to problems such as the curse of dimensionality due to their high dimensionality and sparseness [36]. However, when training a CNN model, a filter/kernel is used to perform a cross product with the 2-dimensional input—that is, a convolution operation is performed across the input volume—to produce a 2-dimensional feature map. The convolution operation (hence, they are called CNN) is followed by an additive bias and squashing function (such as the sigmoid function, hyperbolic tangent function, and Rectifier Linear Unit (ReLU)). A cross-correlation interpretation of the kernels is seen as the input or feature map detectors for certain nonlinear features that are large on a given activation map [37].

Parameters on a convolutional layer include input/feature map size n_{in} , stride (s), zero padding (p), and filter size (k). The spatial size of the activation map is computed as follows:

$$n_{out} = \left\lfloor \frac{n_{in} + 2p - k}{s} \right\rfloor + 1. \quad (4)$$

Pooling layers handle a shift and scale invariance, thus reducing the sensitivity of the output. Moreover, they help reduce the model’s memory consumption by reducing the number of parameters from the feature maps. Nonlinear functions to implement pooling include max pooling, average pooling, and Region of Interest (ROI) pooling [37, 40, 41].

This CNN model resembles the unsupervised self-organized multilayer neocognitron model by Fukushima [42] that was inspired by experiments on the visual cortex of the cat and monkey done by Hubel and Wiesel [42, 43]. The model was made up of cells in a single cell plane (S-layer whose receptive fields are found on the input layer) as shown in Figure 1.

CNN training uses the backpropagation algorithm where a gradient descent search is performed to compute the weights that minimize the classification error [37]. In the backpropagation steps for training the CNN, the stochastic gradient descent search is performed to update the weights. This is done by evaluating gradient based on a single or a small sample of a training sample to update the approximate gradient rather than accumulating the gradients over the entire training set. At each training iteration, a parameter is updated as follows:

$$w_k \leftarrow w_k - \epsilon_k \frac{\partial E^P}{\partial w_k}. \quad (5)$$

However, in the CNN network, the partial derivative is the sum of the partial derivatives with respect to the connections that share the weight parameter.

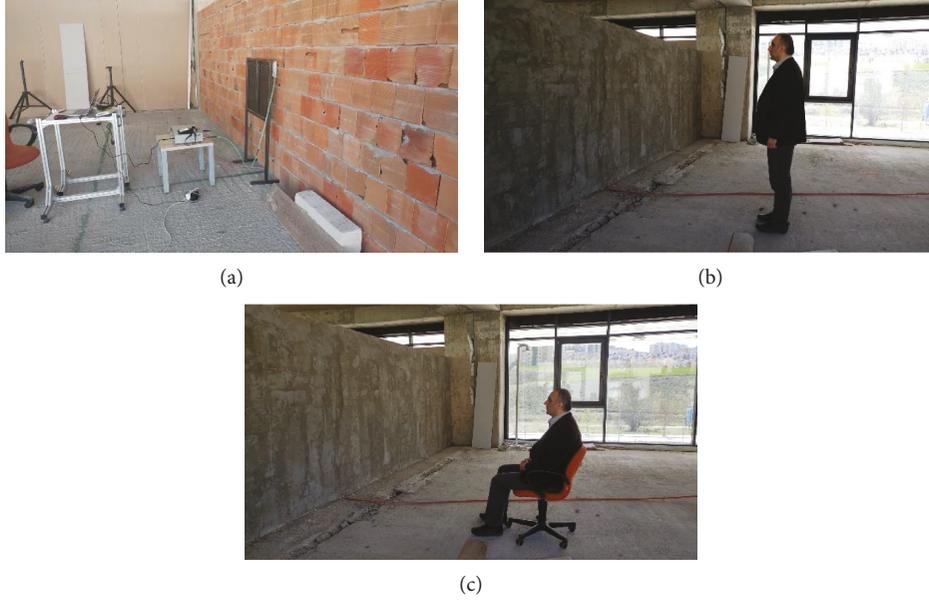


FIGURE 2: Experiment scene: (a) test setup, (b) human standing at 2.5 m, and (c) human sitting at 2.5 m.

Other training recommendations include the Gauss-Newton or Levenberg-Marquardt algorithms, such as the Broyden-Fletcher-Goldfarb-Shanno. The classification module, usually a fully connected multilayer perceptron (MLP), is a trainable classifier that categorizes the resulting feature vectors into classes in the output layer by using loss functions such as softmax [38, 41] which is given as follows:

$$\text{softmax}(x)_i = \frac{\exp(x_i)}{\sum_j \exp(x_j)}. \quad (6)$$

A recent performance of the CNN has greatly improved and surpassed humans in several tasks such as classification, segmentation, object detection, and playing games [36]. This is due to the increasing complexity of the model, increased training samples, and implementation of new training techniques [44, 45]. These new training techniques include initialization schemes [46, 47], deep rectifier networks [48], batch normalization [49], dropout [50], and softmax loss classifier networks [45, 46] and parallel programming with GPUs [40, 51].

4. Detection and Classification of the Human Posture

Using the powerful features of CNNs, it is aimed to detect and classify a human posture and activities whether he/she is standing, sitting, or absent behind a wall by using SFCW radar signals in this study. To detect the human and to classify the posture behind the plastered brick wall, the test data have been gathered by a SFCW radar system with a vector network analyzer (VNA) and two horn antennas. Since the SFCW radar signals carry valuable information about the object behind the wall, the experiment is evaluated by acquiring the SFCW radar signals for 3 different cases including an

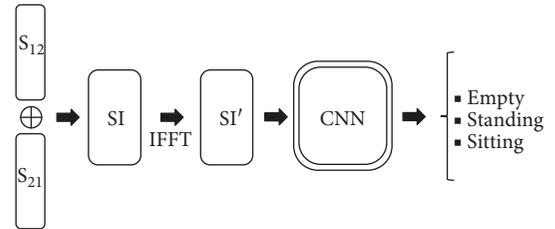


FIGURE 3: Brief demonstration of the suggested approach.

empty scene, a scene with a standing human target, and a scene with a sitting human target. The last two cases with human targets are evaluated for the presence of a human both 2.5 and 5 meters away from the wall. A brief demonstration of the experimental setup can be seen in Figure 2.

Two (S_{12}^j, S_{21}^j) S-parameter vectors are used for each sample. The sample matrix which is utilized as the input data for the CNN structure generated by using these two vectors and the input data matrix SI has the form as $2 \times (N-1)$ which is shown as follows:

$$SI^j = \left(\left(S_{12,0}^j, S_{21,0}^j \right), \left(S_{12,1}^j, S_{21,1}^j \right), \dots, \left(S_{12,N-1}^j, S_{21,N-1}^j \right) \right). \quad (7)$$

To classify the human target into three classes (empty, standing, and sitting), the generated data matrix SI is utilized as the input of the CNN. The CNN structure is constructed sequentially including convolutional, batch normalization, ReLU and pooling layers, second convolutional, batch normalization, ReLU, and pooling layers and fully connected, softmax, and classification layers at the last. Dropout is also applied in the fully connected layer. A brief demonstration of the suggested approach is given in Figure 3.

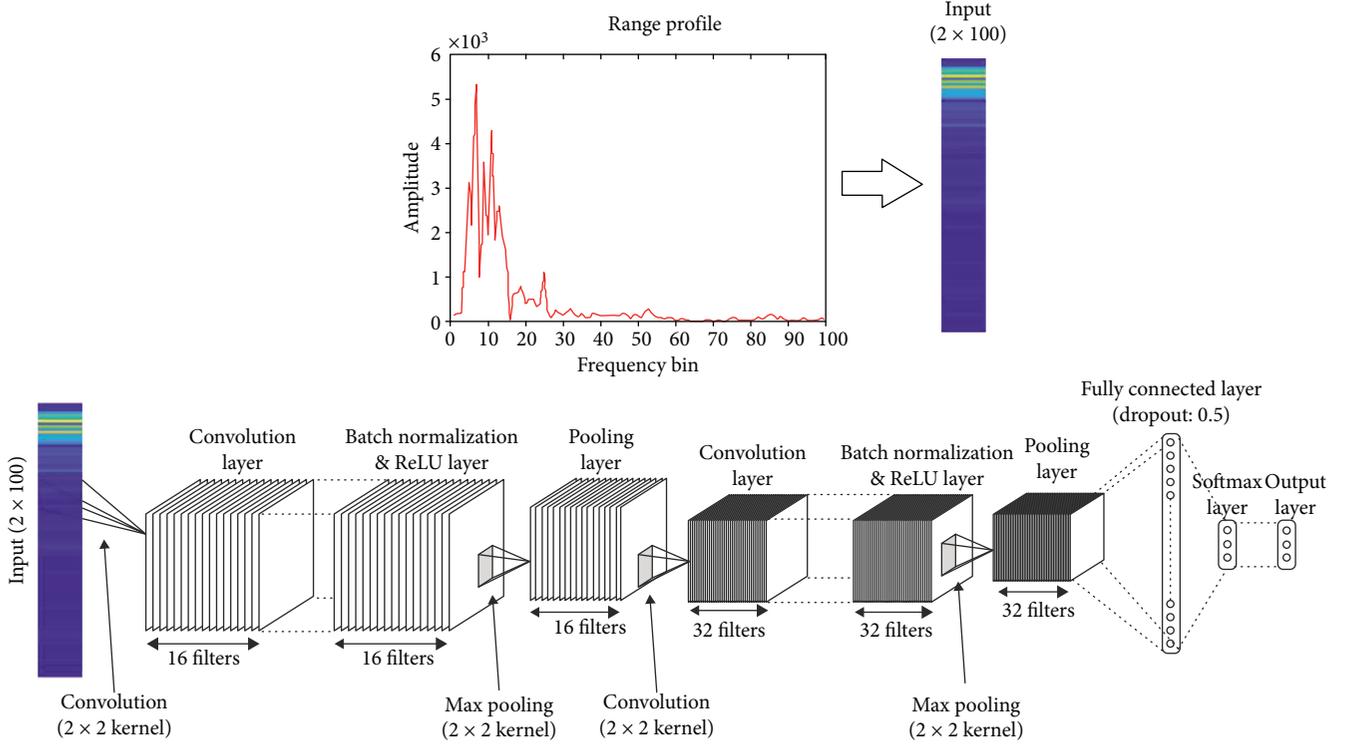


FIGURE 4: CNN structure.

5. Experimental Setup and Results

In order to collect RF data reflecting a human behind the building wall, one Copper Mountain S5065 model vector network analyzer (VNA) having the frequency range of up to 6.5 GHz is used. The stepped frequency waveforms are generated within the range of 2.0 GHz and 4.0 GHz (2 GHz bandwidth (BW)) having a step size $\Delta f = 5$ MHz. Hence, the number of frequency points $N_f = 201$. Also, 1000 readings per sample are collected regarding the given configuration. According to this setup, the maximum unambiguous radar range can be calculated as follows;

$$R_{\max} = \frac{c(N_f - 1)}{2BW} = 15.00 \text{ m}, \quad (8)$$

$$\Delta r = \frac{R_{\max}}{(N - 1)} = \frac{c}{2BW} = 7.5 \text{ cm},$$

where R_{\max} denotes the maximum range, c stands for the speed of light, and Δr refers to the radar's downrange resolution. Horn antennas for transmitting and receiving RF signals are used during the experiments, and the distance between antennas is set to 40 cm. Both transmit and receive antennas are placed about 80 cm above the ground, and the distance between the wall and the antennas is approximately zero (see Figure 2(a)). Despite the 201 points being set as the frequency steps for better downrange resolution, the first 100 data are considered for ease of calculation since the measurement distance in our test scenario is less than 7.5 m. According to the experimental setup, micro movements can be

TABLE 1: CNN structure definitions of the proposed approach.

Layer	Parameters	Value
Input layer	Input size	2×100
	Kernel size	2×2
Convolutional layer	# of filters	16
	Transfer function	ReLU
Batch normalization layer	Transfer function	ReLU
Pooling layer	Kernel size	2×2
	# of filters	16
Convolutional layer	Transfer function	ReLU
	Kernel size	2×2
Batch normalization layer	# of filters	32
	Transfer function	ReLU
Pooling layer	Kernel size	2×2
	# of filters	32
Fully connected layer	Transfer function	ReLU
	Dropout	0.5
Fully connected layer	# of neurons	100
	Classifier	Softmax
Output layer	# of outputs	3
	Classes	{empty, standing, sitting}

TABLE 2: Confusion matrices of experimental results for the free-space scenario.

(a)

		Actual										<i>Total</i>
		Empty	Standing				Sitting					
			Micro @ 2.5 m	Micro @ 5 m	Macro @ 2.5 m	Macro @ 5 m	Micro @ 2.5 m	Micro @ 5 m	Macro @ 2.5 m	Macro @ 5 m		
Predicted	Empty	200	1	3	1	1	1	3	3			
	Micro @2.5 m		198				1					
	Micro @5 m			196				1				
	Macro @2.5 m				196				2			
	Macro @5 m					196				1		
	Micro @2.5 m		1				198					
Sitting	Micro @5 m			1				196				
	Micro @2.5 m				3				198			
	MACRO @5 m					3				196		
	<i>Total</i>	200	200	200	200	200	200	200	200	200	1800	
<i>Accuracy</i>		100%	99%	98%	98%	98%	99%	98%	99%	98%	98.56%	

(b)

		Actual			<i>Total</i>
		Empty	Standing	Sitting	
Predicted	Empty	200	6	7	
	Standing	0	786	5	
	Sitting	0	8	788	
	<i>Total</i>	200	800	800	1800
	<i>Accuracy</i>	100.00%	98.25%	98.50%	98.56%

defined as the minimal movement of the human, just heart beats and aspiration are valuable, and macro movements can be defined as the movement of the head and extremities including the heart beats and the aspiration.

The experimental data is obtained for 9 different scenario measurements. These scenarios can be defined as the absence of a human (empty) scene, standing human scene in which the human is standing at 2.5 m and 5 m away from the wall, and sitting human scene in which the human is sitting on a chair at 2.5 m and 5 m away from the wall. In all scenes excluding the empty scene, the micro and macro movement variations of the scenarios are also collected. Therefore, the experiments are evaluated on 9 totally different scenario data. Each scenario has 1000 readings related with 3 classes which are empty, standing, and sitting. All experiments are evaluated without the wall, states as free space, and two different dielectric wall structures which are a brick wall and a drywall. The plastered brick wall is made of standard 135 mm thickness bricks, and the plaster thickness is approximately 10 mm and is not homogeneous. The drywall was formed by combining 5 pieces of 12.5 mm thick plates side-by-side. There is a nonhomogenous air gap of less than 10 mm between the plates.

In the classification process, a CNN structure is generated specific to the study. The demonstration of the proposed CNN structure is given in Figure 4, and the CNN structure

and the detailed layer definitions are given in Table 1. The training dataset is constructed by using 80% of the randomly selected readings, and the test dataset is constructed by using the remaining part of the readings. Therefore, 7200 readings are used in the training phase and 1800 readings are used for the test phase of the experiment. The proposed approach is run 30 times with different random seeds, and the results are presented by using the means of the 30 runs. The confusion matrices of experimental results are given in Tables 2–4 for the free-space, brick wall, and drywall scenarios, respectively, and the overall experimental results are given in Table 5.

While data are available for longer ranges, data regarding a shorter distance is used in these tests and experiments. This reduced unwanted data such as the reflection and multipath effects the most. Since the reflected RF signals contain important information about the obstacles at the radial distance, the CNN will be able to successfully differentiate situations where dimensional differences are high. Increasing the resolution of downrange by increasing bandwidth will, therefore, increase the success of discrimination. By increasing the number of SFCW frequencies, more data will be obtained, and again, classification success will be increased for longer distances.

Although researchers focus on the reduction of negative effects of the wall, this approach may reduce success in the

TABLE 3: Confusion matrices of experimental results for the brick wall scenario.

(a)

		Actual										
		Empty	Standing				Sitting					
			Micro @ 2.5 m	Micro @ 5 m	Macro @ 2.5 m	Macro @ 5 m	Micro @ 2.5 m	Micro @ 5 m	Macro @ 2.5 m	Macro @ 5 m		
Predicted	Empty	200	1	4	1	7	4	7	2	8		
	Standing	Micro @2.5 m	196				2					
		Micro @5 m		194				1				
	Sitting	Macro @2.5 m			194					4		
		Macro @5 m				190						
	Sitting	Micro @2.5 m		3			194					
		Micro @5 m			2			192				
	Macro @2.5 m				5					194		
Macro @5 m					3					192		
	<i>Total</i>	200	200	200	200	200	200	200	200	200	1800	
	<i>Accuracy</i>	100%	98%	97%	97%	95%	97%	96%	97%	96%	97%	

(b)

		Actual			
		Empty	Standing	Sitting	
Predicted	Empty	200	13	21	
	Standing	0	774	7	
	Sitting	0	13	772	<i>Total</i>
	<i>Total</i>	200	800	800	1800
	<i>Accuracy</i>	100.00%	96.75%	96.50%	97%

learning phase. The use of micro-Doppler methods would be appropriate for imaging and detection; the tests would need to be deepened for classification. Preprocessing is applied with the assumption that the wall is constant, and filters will destroy some valuable data on the environment.

The influence of the permeability parameter in the tests is clearly observed. For example, it can be said that the classification success rate has been slightly reduced because there are fewer signals from especially plastered brick walls with lower permeability.

In the tests and experiments, the micro movements which are more difficult to detect by the conventional methods and the macro movements whose effects are observed on the whole frequency band are evaluated separately. Macro movements which reduce valuable information about the environment behind the wall have reduced the classification success. However, increasing the amount of data in the learning and training process will reduce this effect. Thus, by increasing the training data, similar high classification accuracy has been achieved for both micro and macro movements as seen in Table 5.

6. Conclusion

This study focuses on the assessment of the classification of the SFCW radar (as a form of UWB radar) signals in order

to detect the absence of a human and the human posture using CNNs. In the literature, there are techniques and methods using micro-Doppler signatures and the slow time data of target movements. However, in this study, long-term data are not needed, and classification is done instantaneously. Consequently, remarkable results have been obtained with this study, which is the first (to the best of the authors' knowledge) in the scope of using CNN to classify raw data obtained from through-wall radars without both detailed preprocessing and slow time data.

SFCW radar signals of the generated test scenes are obtained using a VNA, and after constructing the dataset by utilizing the readings, it is used to classify the presence of the human and the human posture whether he/she is standing or sitting by using CNN.

In another form of UWB, frequency-modulated continuous wave (FMCW) radar signals could be used as the data source for CNN. Since SFCW radar techniques have benefits including high mean transmitter power and high receiver sensitivity, the SFCW radar has been used in this study.

The proposed approach achieves remarkable and successful results without detailed preprocessing operations used in the traditional approaches. Future works include the usage of synthetic aperture radar (SAR) for dataset construction, evaluating the performance of different kinds of deep learning classifiers for the study concept.

TABLE 4: Confusion matrices of experimental results for the drywall scenario.

(a)

		Actual										<i>Total</i>
		Empty	Standing				Sitting					
			Micro @ 2.5 m	Micro @ 5 m	Macro @ 2.5 m	Macro @ 5 m	Micro @ 2.5 m	Micro @ 5 m	Macro @ 2.5 m	Macro @ 5 m		
Predicted	Empty	200	1	3	2	7	2	5	3	6		
	Standing	Micro @2.5 m	198				2					
		Micro @5 m		196				1				
	Macro	@2.5 m			194				3			
		@5 m				192					2	
	Macro	@2.5 m	1				196					
Sitting	Micro @5 m		1				194					
	Macro @2.5 m			4				194				
	Macro @5 m				1					192		
	<i>Total</i>	200	200	200	200	200	200	200	200	200	1800	
<i>Accuracy</i>		100%	99%	98%	97%	96%	98%	97%	97%	96%	97.56%	

(b)

		Actual			<i>Total</i>
		Empty	Standing	Sitting	
Predicted	Empty	200	13	16	
	Standing	0	780	8	
	Sitting	0	7	776	
	<i>Total</i>	200	800	800	1800
	<i>Accuracy</i>	100.00%	97.50%	97.00%	97.56%

TABLE 5: Experimental results (accuracy, in %).

Scene	Empty	Standing (m)				Sitting (m)			
		Micro 2.5	Micro 5.0	Macro 2.5	Macro 5.0	Micro 2.5	Micro 5.0	Macro 2.5	Macro 5.0
Free space	100	99	98	98	98	99	98	99	98
Brick wall	100	98	97	97	95	97	96	97	96
Drywall	100	99	98	97	96	98	97	97	96

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors would like to thank the STM Savunma Teknolojileri Mühendislik ve Ticaret A.Ş. company for their cooperation and support.

References

- [1] A. G. Yarovoy, X. Zhuge, T. G. Savelyev, and L. P. Ligthart, "Comparison of UWB technologies for human being detection with radar," *2007 European Microwave Conference*, vol. 1-4, pp. 1574-1577, 2007.
- [2] D. V. Kadaba, K. Bachina, S. A. Subhan, V. K. Bansal, G. Gowtham, and L. Ramakrishnan, "Real-time through-wall imaging using SFCW radar system," in *9th International Radar Symposium*, pp. 1-6, India, 2013.
- [3] F. G. Qi, F. L. Liang, H. Lv, C. T. Li, F. M. Chen, and J. Q. Wang, "Detection and classification of finer-grained human activities based on stepped-frequency continuous-wave through-wall radar," *Sensors*, vol. 16, no. 6, p. 885, 2016.
- [4] B. Y. Lu, Q. Song, Z. M. Zhou, and H. N. Wang, "A SFCW radar for through wall imaging and motion detection," in

- 2011 8th European Radar Conference, pp. 325–328, Manchester, UK, 2011.
- [5] Y. M. Liu, H. D. Meng, H. Zhang, and X. Q. Wang, “Motion compensation of moving targets for high range resolution stepped-frequency radar,” *Sensors*, vol. 8, no. 5, pp. 3429–3437, 2008.
 - [6] B. W. Zhang and G. F. Zhu, “High frame rate stepped frequency through-wall imaging radar,” in *2015 IEEE International Wireless Symposium (IWS 2015)*, pp. 1–4, Shenzhen, China, March 2015.
 - [7] M. Mercuri, D. Schreurs, and P. Leroux, “SFCW microwave radar for in-door fall detection,” in *2012 IEEE Topical Conference on Biomedical Wireless Technologies, Networks, and Sensing Systems (BioWireless)*, pp. 53–56, Santa Clara, CA, USA, January 2012.
 - [8] J. Marimuthu, K. S. Bialkowski, and A. M. Abbosh, “Stepped frequency continuous wave software defined radar for medical imaging,” in *2014 IEEE Antennas and Propagation Society International Symposium (APSURSI)*, pp. 1909–1910, Memphis, TN, USA, July 2014.
 - [9] H. Liu and M. Sato, “In situ measurement of pavement thickness and dielectric permittivity by GPR using an antenna array,” *Ndt & E International*, vol. 64, pp. 65–71, 2014.
 - [10] X. Wei and Y. Zhang, “Autofocusing techniques for GPR data from RC bridge decks,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 12, pp. 4860–4868, 2014.
 - [11] M. Pieraccini, D. Tarchi, H. Rudolf et al., “Structural static testing by interferometric synthetic radar,” *NDT & E International*, vol. 33, no. 8, pp. 565–570, 2000.
 - [12] A. Hartikainen, T. Pellinen, E. Huuskonen-Snicker, and P. Eskelinen, “Algorithm to process the stepped frequency radar signal for a thin road surface application,” *Construction and Building Materials*, vol. 158, pp. 1090–1098, 2018.
 - [13] D. Seyfried, A. Busche, R. Janning, L. Schmidt-Thieme, and J. Schoebel, “Information extraction from ultrawideband ground penetrating radar data: a machine learning approach,” in *2012 The 7th German Microwave Conference*, pp. 1–4, German, 2012.
 - [14] D. Seyfried, R. Jansen, and J. Schoebel, “Shielded loaded bowtie antenna incorporating the presence of paving structure for improved GPR pipe detection,” *Journal of Applied Geophysics*, vol. 111, pp. 289–298, 2014.
 - [15] T. M. Urban, C. Vella, E. Bocancea, C. A. Tuttle, and S. E. Alcock, “A geophysical investigation of a newly discovered Early Bronze Age site near Petra, Jordan,” *Journal of Archaeological Science*, vol. 42, pp. 260–272, 2014.
 - [16] M. A. Mohana, A. M. Abbas, M. L. Gomaa, and S. M. Ebrahim, “Discrimination between landmine and mine-like targets using wavelets and spectral analysis,” *NRIAG Journal of Astronomy and Geophysics*, vol. 2, no. 1, pp. 54–66, 2013.
 - [17] A. Yarovoy, V. Kovalenko, F. Roth, E. Ligthart, A. Fogar, and L. Ligthart, “Landmine detection and discrimination based on GPR data,” in *Proceedings of the Tenth International Conference on Ground Penetrating Radar*, vol. 1-2, pp. 673–676, Delft, The Netherlands, 2004.
 - [18] X.-X. Niu and C. Y. Suen, “A novel hybrid CNN–SVM classifier for recognizing handwritten digits,” *Pattern Recognition*, vol. 45, no. 4, pp. 1318–1325, 2012.
 - [19] M. D. Zeiler and R. Fergus, “Visualizing and understanding convolutional networks,” in *Computer Vision – ECCV 2014*, ECCV 2014, D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds., vol. 8689 of Lecture Notes in Computer Science, pp. 818–833, Springer International Publishing, Cham, 2014.
 - [20] A. Baldominos, Y. Saez, and P. Isasi, “Evolutionary convolutional neural networks: an application to handwriting recognition,” *Neurocomputing*, vol. 283, pp. 38–52, 2018.
 - [21] J. X. Gu, Z. Wang, J. Kuen et al., “Recent advances in convolutional neural networks,” *Pattern Recognition*, vol. 77, pp. 354–377, 2018.
 - [22] Y. Zhu, Y. Zhu, B. Y. Zhao, and H. Zheng, “Reusing 60 GHz radios for mobile radar imaging,” in *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking - MobiCom '15*, pp. 103–116, Paris, France, September 2015.
 - [23] F. Adib, C. Y. Hsu, H. Z. Mao, D. Katabi, and F. Durand, “Capturing the human figure through a wall,” *ACM Transactions on Graphics*, vol. 34, no. 6, pp. 1–13, 2015.
 - [24] S. Y. Wu, K. Tan, Z. H. Xia, J. Chen, S. W. Meng, and G. Y. Fang, “Improved human respiration detection method via ultra-wideband radar in through-wall or other similar conditions,” *IET Radar Sonar & Navigation*, vol. 10, no. 3, pp. 468–476, 2016.
 - [25] S. Bakhtiari, T. W. Elmer, N. M. Cox et al., “Compact millimeter-wave sensor for remote monitoring of vital signs,” *IEEE Transactions on Instrumentation and Measurement*, vol. 61, no. 3, pp. 830–841, 2012.
 - [26] L. Anishchenko and E. Gaysina, “Comparison of 4 GHz and 14 GHz SFCW radars in measuring of small laboratory animals vital signs,” in *2015 IEEE International Conference on Microwaves, Communications, Antennas and Electronic Systems (COMCAS)*, pp. 1–3, Tel Aviv, Israel, November 2015.
 - [27] B. Tekeli, S. Z. Gurbuz, and M. Yuksel, “Information-theoretic feature selection for human micro-Doppler signature classification,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 5, pp. 2749–2762, 2016.
 - [28] C. Karabacak, S. Z. Gurbuz, A. C. Gurbuz, M. B. Guldogan, G. Hendeby, and F. Gustafsson, “Knowledge exploitation for human micro-Doppler classification,” *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 10, pp. 2125–2129, 2015.
 - [29] H. Griffiths, M. Ritchie, and F. Fioranelli, “Bistatic radar configuration for human body and limb motion detection and classification,” in *Radar for Indoor Monitoring: Detection, Classification and Assessment*, M. G. Amin, Ed., CRC Press, Boca Raton, FL, USA, 2017.
 - [30] Y. Kim and H. Ling, “Human activity classification based on micro-Doppler signatures using a support vector machine,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 5, pp. 1328–1337, 2009.
 - [31] R. J. Fontana, “Recent system applications of short-pulse ultrawideband (UWB) technology,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 52, no. 9, pp. 2087–2104, 2004.
 - [32] A. Qamar and U. Faruq, “Modelling and simulation of UWB radar system for through the wall imaging and Doppler detection,” *International Journal of Engineering Trends and Technology*, vol. 17, no. 7, pp. 325–330, 2014.
 - [33] Y. S. Yoon and M. G. Amin, “Spatial filtering for wall-clutter mitigation in through-the-wall radar imaging,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 9, pp. 3192–3208, 2009.
 - [34] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” in

- Presented at the Proceedings of the 25th International Conference on Neural Information Processing Systems*, vol. 1, pp. 1097–1105, Lake Tahoe, Nevada, 2012.
- [35] Y. B. I. Goodfellow and A. Courville, *Deep Learning*, MIT Press, 2017.
- [36] H. H. Aghdam and E. J. Heravi, *Guide to Convolutional Neural Networks: A Practical Application to Traffic-Sign Detection and Classification*, Springer, 2017.
- [37] T. Dettmers, *Deep Learning in a Nutshell: Core Concepts*, NVidia, 2015, <https://devblogs.nvidia.com/deep-learning-nutshell-core-concepts/>.
- [38] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [39] B. Sahiner, Heang-Ping Chan, N. Petrick et al., “Classification of mass and normal breast tissue: a convolution neural network classifier with spatial domain and texture images,” *IEEE Transactions on Medical Imaging*, vol. 15, no. 5, pp. 598–610, 1996.
- [40] C. Hazirbas, L. Ma, C. Domokos, and D. Cremers, “FuseNet: incorporating depth into semantic segmentation via fusion-based CNN architecture,” in *Computer Vision – ACCV 2016. ACCV 2016*, S. H. Lai, V. Lepetit, K. Nishino, and Y. Sato, Eds., vol. 10111 of Lecture Notes in Computer Science, pp. 213–228, Springer International Publishing, Cham, 2017.
- [41] P. Pinheiro and R. Collobert, “Recurrent convolutional neural networks for scene labeling,” in *Proceedings of the 31st International Conference on Machine Learning*, pp. 82–90, Beijing, China, 2014.
- [42] K. Fukushima, “Cognitron: a self-organizing multilayered neural network,” *Biological Cybernetics*, vol. 20, no. 3, pp. 121–136, 1975.
- [43] D. H. Hubel and T. N. Wiesel, “Ferrier lecture - functional architecture of macaque monkey visual cortex,” *Proceedings of the Royal Society of London. Series B. Biological Sciences*, vol. 198, no. 1130, pp. 1–59, 1977.
- [44] W. Shi, Y. Gong, X. Tao, J. Wang, and N. Zheng, “Improving CNN performance accuracies with min-max objective,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 7, pp. 1–14, 2017.
- [45] C. Szegedy, W. Liu, Y. Jia et al., “Going deeper with convolutions,” in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1–9, Boston, MA, USA, 2015.
- [46] X. Glorot and Y. Bengio, “Understanding the difficulty of training deep feedforward neural networks,” in *Presented at the Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, Proceedings of Machine Learning Research*, pp. 249–256, Sardinia, Italy, 2010, <http://proceedings.mlr.press>.
- [47] I. Sutskever, J. Martens, G. Dahl, and G. Hinton, “On the importance of initialization and momentum in deep learning,” in *Presented at the Proceedings of the 30th International Conference on Machine Learning, Proceedings of Machine Learning Research*, pp. 1139–1147, Atlanta, GA, USA, 2013, <https://dl.acm.org/citation.cfm?id=3043064>.
- [48] X. Glorot, A. Bordes, and Y. Bengio, “Deep sparse rectifier neural networks,” in *Proceedings of the 14th International Conference on Artificial Intelligence and Statistics (AISTATS)*, pp. 315–323, Ft. Lauderdale, FL, USA, 2011, <http://proceedings.mlr.press>.
- [49] S. Ioffe and C. Szegedy, “Batch normalization: accelerating deep network training by reducing internal covariate shift,” in *Presented at the Proceedings of the 32nd International Conference on Machine Learning, Proceedings of Machine Learning Research*, pp. 448–456, Fort Lauderdale, FL, USA, 2015.
- [50] N. Srivastava, G. E. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, “Dropout: a simple way to prevent neural networks from overfitting,” *Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929–1958, 2014.
- [51] D. C. Cireşan, U. Meier, J. Masci, L. M. Gambardella, J. Schmidhuber et al., “Flexible, high performance convolutional neural networks for image classification,” in *Presented at the Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence*, vol. 2, Barcelona, Catalonia, Spain, 2011.
- [52] K. Fukushima, “Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position,” *Biological Cybernetics*, vol. 36, no. 4, pp. 193–202, 1980.

