Human Detection by Deep Neural Networks Recognizing Micro-Doppler Signals of Radar

Jihoon Kwon
Seoul National University
and Hanwha systems company
Seoul, Korea
jihoonkwon@snu.ac.kr

Seoungeui Lee
Seoul National University
Mipal Lab., GSCST
Seoul, Korea
dehlix@snu.ac.kr

Nojun Kwak
Seoul National University
Mipal Lab., GSCST
Seoul, Korea
nojunk@snu.ac.kr

***Abstract*— The purpose of this paper is to show the effectiveness of Deep neural networks (DNN) for recognizing the micro-Doppler radar signals generated by human walking and background noises. To show this, we collected various micro-Doppler signals considering the actual human walking motion and background noise characteristics. Unlike the previous researches that required a complex feature extraction process, we directly use the FFT result of the input signal as a feature vector without any additional pre-processing. This technique helps not to use heuristic approaches to get a meaningful feature vector. We designed two types of DNN classifier. The first is the binary classifier to classify human walking Doppler signals and background noises. The second is the multiclass classifier that is roughly able to recognize a circumstance of a place as well as human walking Doppler signals. DNN for the binary classifier showed about 97.5% classification accuracy for the test dataset and DNN(ReLU) for the multiclass classifier showed about 95.6% accuracy.**

***Keywords*— Doppler radar, micro-Doppler, human detection, Deep neural network, radar classification, radar machine learning**

I. Introduction

In the field of commercial applications, radar sensors can be widely used for LED control and alarm monitoring in indoor and outdoor environments for a smart home and a smart city. A radar sensor can be applied as an useful sensor in the area where there is restriction to the use of sound, infrared, vibration and camera sensors. For example, Acoustic and vibration sensors are very vulnerable to acoustic noise, and infrared sensors frequently generate false alarms in outdoor environments. Camera sensors are relatively expensive, require high signal processing burden, decline in performance at night, and have lens contamination problems. On the other hand, the radar sensors are robust against weather conditions and sensor performances do not decrease at night. In addition, since a signal processing for detecting a target is relatively easy and effective, it is possible to effectively utilize a radar sensor in various indoor and outdoor fields [1].

A conventional object detection method of a radar sensor is processed by comparing a received signal power with a fixed or an adaptive threshold level. This method is so simple that it has been now widely applied to indoor commercial products. However, this method has an uncontrollable problem that false alarms occur frequently due to various background noise sources. Also, RCS fluctuation of targets negatively influences on a detection performance. And it can lead to situations that a received signal is under the threshold level. So, these problems should be importantly solved and considered for commercial applications of a radar sensor.

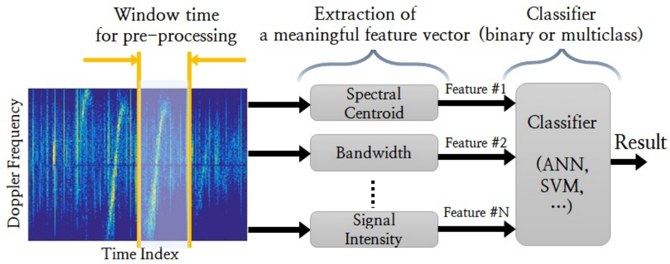

Fig. 1. The example of the conventional approach for classification of human walking. It requires complicated processes such as statistical analysis which is considered as a heuristic method in order to extract a meaningful and effective feature vector.

Doppler signals of a radar are used to resolve moving objects under environments with complicated background noises. To effectively classify real targets and various background noises, Doppler signals are regarded as one of very important data captured by a radar. In the case of air surveillance radar, since a speed of a target is fast, Doppler filtering process is relatively not complicated. However, in the case of radars for detecting human motions, classifying Doppler signals caused by human walking from various noise signals is considerably difficult, because the speed of human walking is quite slow. A human walking speed is known as approximately 1.4 m/s [2]. Therefore, in the case of radars for detecting human walking or motions, a bandwidth of Doppler signals by human motions overlaps a bandwidth of Doppler signals by background noise sources. So, it is hard to detect exactly human walking under this complicated environment. Accurate detection of human walking requires a high-level Doppler resolution for resolving the micro-Doppler signals. However, this requirement are not attractive functionally and commercially if we can not accept for increases of the Doppler processing time and the hardware cost.

On the other hand, processing techniques that recognize a pattern of a received signal can be good alternatives to overcome these limitations. Many researches on the Doppler radar detection using pattern recognition have been reported [1,3-5]. According to the reported studies, many researchers have described or proposed a complex mathematical equation or approach for finding a meaningful feature vector that affects the classifier's estimation accuracy [3-5]. However, finding this equation or approach is based on heuristic methods that require a lot of effort and time [3-5]. Also, this approach needs a relatively long pre-processing time (window time) to extract a feature vector.

Fig. 1 shows an example of the extraction process of the micro-Doppler spectrum by human walking. The problem is that it can require more than a few seconds to extract these features at least in the previous researches. To overcome this problem, we proposed the human detection method by the neural network using Multilayer Perceptron in the previous research [1]. This algorithm had the simple neural network structure for applying to a low cost embedded process, and the processing speed was suggested as 0.5 second. And the paper [1] demonstrated that MLP of a binary classifier without the pre-processing can have a high classification accuracy. However, since we had a restriction condition about an computational complexity, the multi-layer perceptron proposed in the paper [1] used two hidden layers between the input and output layers. This structure can be applied as a typical form of a binary classifier but it can have a limitation on the extension to a multi-class classifier.

Therefore, in this paper, we propose the classification method using deep neural networks to classify micro Doppler signals caused by human walking and background noises. In order to eliminate a pre-processing, we follow the basic idea suggested by the paper [1]. However, in order to overcome the limitation of the paper [1], we design the multi-class classifier with high performances by using Deep Neural Networks and demonstrate it's performance by using the real micro-Doppler signal dataset.

II. DEEP NEURAL NETWORKS [6,7]

The DNN is a kind of Artificial Neural Network (ANN) with several hidden layers between the input and output layers. In the proposed DNN, all frequency components converted by the Fourier Transform are directly delivered to the input layer without any effort for frequency filtering and so on. The lower layer of the DNN receives information from the upper layer, and a feature vector can be gradually generated and rearranged by a process through passing several layers. Unlike a conventional feature extraction that required a heuristic approach and a pre-processing, this approach that we can naturally design an optimal feature extraction model using DNN that have many layers and nodes.

There are various types of the DNNs such as CNNs (Convolutional Neural Networks), RNNs (Recurrent Neural Networks) and so on. CNN has been successfully applied in the field of computer vision. In the case of radar applications, it can be applied to a image data processing related to Doppler spectrum and SAR images. However, since making a Doppler spectral image requires considerably a long pre-processing time, it is not suitable for this research purpose, that's because we consider a fast processing time to detect a moving human walking. On the other hand, RNN can be applied for speech or language classifications. Since the micro-Doppler signals by a radar are a time-series form, RNN can be an useful tool for a radar signal classification. However, at the initial time that a radar sensor begin to detect an object, a radar sensor needs a sequential input data during several time steps to obtain the first result. This problem can make a classification accuracy worse. In other words, while a radar start to detect a target, the radar's performance temporarily get worse and it can make a false result. According to these reasons and limitations, we selected the feed-forward deep neural network (DNN) which is suitable for our research purpose.

A. ReLU and SoftMax

As the number of layers in the architecture increases, DNN can suffer from a problem of vanishing gradient. This problem is related to the gradient based learning methods using certain activation functions. In other words, the traditional activation functions of neural networks such as the sigmoid or hyperbolic tangent function have gradients of the range (0, 1). However, since the back-propagation algorithm to train models computes gradients by the chain rule, the gradient gets very close to zero according to increase of the number of layers. And it is not possible to properly train a model in this situation due to this gradient vanishing problem.

ReLU (Rectified Linear Unit), which is a kind of activation function, avoids the problem of vanishing gradient and it propagates the efficient gradients to the networks. Recently, DNN mainly uses ReLU as an activation function of an input layer and hidden layers. If the input value of ReLU is less than 0, ReLU outputs 0. If the input value of ReLU is higher than 0, the output of ReLU is equal to the input value. There is no a saturation concept in the positive area unlike a sigmoid or hyper tangent function.

When it comes to the activation function for the output layer, when DNN is used as a binary classifier, we use the sigmoid as an active function of the output stage. On the other hand, in the case of the multi-class classifier, the probabilities of the given classes are calculated by Softmax. Softmax function returns the normalized probabilities for the classes respectively. Eq. (1) shows Softmax function. The predicted probability for the jth class given a sample vector x and a weighting vector w is:

$$P(y=j|x)_j = e^{x^T w_j} / \sum_{k=1}^{K} e^{x^T w_j},\ j = 1, ..., K \tag{1}$$

B. Loss function and Optimizer

The cross entropy (CE) is used as a cost function instead of a mean squared error (MSE), because a learning speed of CE is known to be faster than MSE. On the other hand,

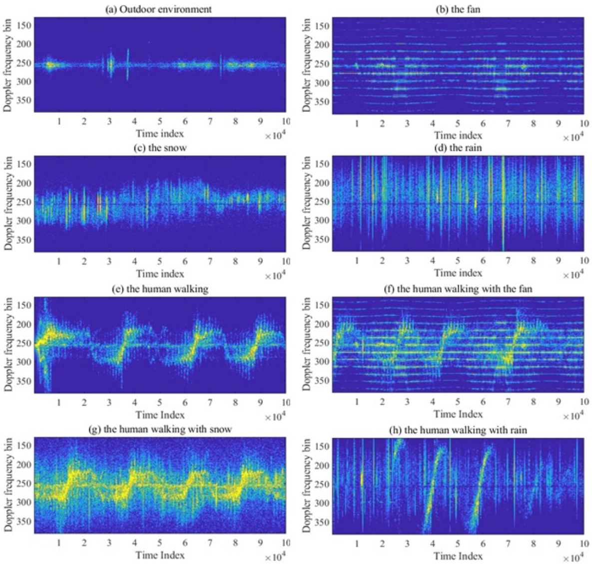

Fig. 2. The micro-Doppler Spectrograms of 8 cases: (a) outdoor environment (LOS), (b) fan, (c) snow, (d) rain, (e) human walking, (f) human walking with fan, (g) human walking with snow and (h) human walking with rain.

a optimization is a process of finding optimal variables of networks though a learning process to minimize the loss function. In this research, we use Adam algorithm as an optimizer.

III. THE MICRO- DOPPLER SIGNALS

The Doppler radar was installed in the selected test sites and the raw data was gathered for several days. Especially, in order to analyse rain and snow effects, we collected the rain and snow background noise datasets. To measure the signals by human motions, the experimenter moved only within the radar detection area. Since this research is related to a LED control purpose, human motions of a normal walking and a fast walking were considered as a general case. Strange human movements such as crawling and irregular movements were excluded. To collect Doppler signals caused by background noises, we considered the limited environments so that the various noise sources do not mix with each other. The micro-Doppler signals were collected by the following eight cases. The eight cases include (a) outdoor environment (LOS), (b) fan, (c) snow, (d) rain, (e) human walking, (f) human walking with fan, (g) human walking with snow and (h) human walking with rain. And these signals were collected three times in the three different places. Fig. 2 shows the result processed by STFT (Short Time Fourier Transform) of the micro-Doppler signals that we collected.

We selected 320,000 data from the raw data of the micro-Doppler signals. We removed the severe outliers in the given dataset and we did not use any additional pre-processing for feature selections and noise filtering. This dataset has 160,000 data of background noise class and 160,000 data of human walking class. The background noise class includes (a) 40,000 outdoor environment data, (b) 40,000 fan data, (c) 40,000 snow data and (d) 40,000 rain data. Also, the human walking class includes (e) 40,000 human walking data, (f) 40,000 human walking data with fan, and (g) 40,000 human walking data with snow, (h) 40,000 human walking data with rain. Among them, 240,000 data were randomly selected for training data. Also, 40,000 data and 40,000 data were randomly selected for test set and validation set respectively.

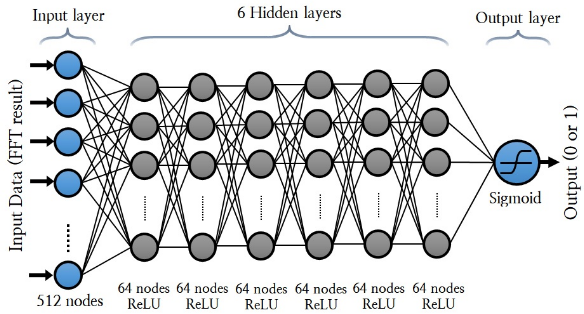

Fig. 3. The designed DNN architecture for a binary classifier : It has 1 input layer, 6 hidden layers and 1 output layer. ReLUs are used for the hidden layers. The activation function of the output layer uses Sigmoid.

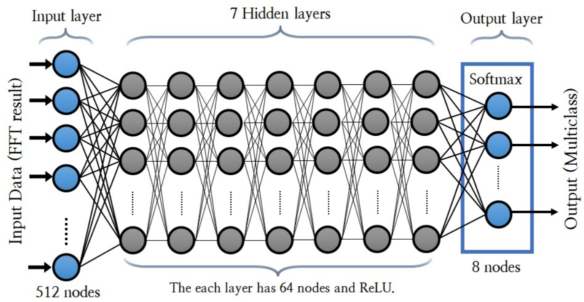

Fig. 4. The designed DNN (ReLU) architecture for a multiclass classifier : It has 1 input layer, 7 hidden layers and 1 output layer. ReLU and Softmax are used for the activation function.

IV. DESIGN AND EXPERIMENT

A. *Binary classifier*

The first type of the proposed DNN is designed as a binary classifiers, which classifies human micro-Doppler signals (e, f, g and h) and background noises (a, b, c and d). Fig. 3 shows the structure of DNN proposed in this research. The designed DNN architecture has 8 layers. The input layer consists of 512 input nodes. The hidden layer has 6 layers and the each layer has 64 nodes and ReLUs. The output layer has 1 node and Sigmoid activation function. If the number of nodes in the hidden layer increases excessively, the performances of the classifier tend to be degraded. We selected 64 nodes and 6 hidden layers optimally after repeated experiments. We also designed MLP using three hidden layers (64 nodes and Sigmoid activation function) for performance comparison. If the number of hidden layers with Sigmoid increases in MLP, the performance of the network tends to degrade because of a gradient vanishing problem.

The DNN performance as the binary classifier is better than MLP. The classification accuracy of MLP for the test dataset is about 91.2%. And the classification accuracy of DNN for

Table 1. Normalized confusion matrix of the multiclass classifier by DNN(ReLu): This result is based on the experimental dataset obtained in the limited environmental conditions. In order to get more generalized results, various additional experimental datasets can be required. Nevertheless, this result show that the designed DNN(ReLU) has a high ability to effectively classify similar pattern types of Doppler signals.

Actual class \ Estimated class	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
(a)	95.2%	0.1%	2.9%	0.8%	0.7%	0.1%	0.3%	0.0%
(b)	0.0%	95.7%	0.0%	0.0%	0.1%	0.0%	0.4%	0.0%
(c)	0.8%	0.0%	98.2%	0.4%	0.1%	0.0%	0.4%	0.0%
(d)	0.4%	0.0%	1.2%	93.2%	0.8%	0.1%	0.3%	4.0%
(e)	0.9%	0.0%	0.7%	0.6%	91.9%	0.1%	4.5%	1.2%
(f)	0.0%	1.1%	0.0%	0.1%	0.0%	97.7%	0.5%	0.3%
(g)	0.0%	0.0%	0.3%	0.1%	1.4%	0.2%	97.5%	0.5%
(h)	0.0%	0.0%	0.1%	1.5%	1.8%	0.3%	0.8%	95.5%

the test dataset is about 97.5%. Also, we tested DNN that only Sigmoid is applied as an activation function instead of ReLU and this model's classification accuracy was measured as about 89.4%.

B. *Multi-class classifier*

The second type of the proposed DNN is designed as a multi-class classifier that is for not only recognizing human walking motions but also classifying background noises sources. The designed DNN architecture is similar to the architecture of the binary classifier described above. However, there is a big difference in the output layer. We used Softmax as an activation function in the output layer for the multi-class classifier. In the case of MLP with four or less hidden layers, the classification accuracy performance was measured to be less than 80%. And we added a hidden layer to improve the performance of MLP and we selected the optimal MLP, which has 6 hidden layers, having Sigmoid activation function. In the case of MLP, when it has 6 hidden layers, it showed the best performance. If additional hidden layers are used, it can degrades the performance of the classifier. This architecture is understood as a kind of DNN because it has the deep layer structure. The average classification accuracy of DNN with Sigmoid for test dataset is as about 85.0%

On the other hand, Fig. 4 shows the structure of DNN with ReLU. Likewise, we determined this optimal architecture after several repeated experiments. The average classification accuracy of DNN with ReLU for test dataset is about 95.6%. The performance of the multiclass classifier using DNN with ReLU is better than DNN with Sigmoid. Table 1 shows the confusion matrix of the multiclass classifier using DNN(ReLU). The average classification accuracy is defined as the average of the diagonal values in the normalized confusion matrix (Table 1) of a multiclass classifier. The each diagonal value in this table shows the classification accuracy of each class. The values in the table were rounded at the second decimal place.

What we have to be the most careful is the case that the background noises (a, b, c and d) are incorrectly estimated as the human walking cases (e, f, g and h). These errors generate false alarms and very harmful to the system. In Table 1, some of these errors were measured to be higher than 0.5%. Conversely, there are errors that do not have a serious impact. For example, if (e, f, g and h) related to human walking are estimated as one of (a, b, c and d), there is going not to be harmful to a system. In the table 1, most of these error results are less than 1% and these are an acceptable degree.

V. CONCLUSION

The purpose of this paper is to show the effectiveness of DNN for recognizing the micro-Doppler radar signals generated by human walking and background noises. In this paper, we directly use the FFT result of the input signal as feature vectors for DNN in order to remove or simplfy complex and heuristic feature extractions (pre-processing) that many previous researches have used. In this research, we designed the binary and multiclass classifiers by using MLP and DNN. As a result, DNN for the binary classifier and DNN(ReLU) for the multiclass classifier showed the excellent classification accuracy performance for the test dataset. This result showed that the designed DNN can have an excellent and effective capability for recognizing micro-Doppler signals of a radar.

Of course, we need new datasets considering various situations and places for overcoming a generalization problem. Nevertheless, even if new datasets are added, we are sure that DNN classifier can be optimized easily than conventional methods. When it comes to a future work, an ensemble learning method that combines two or three machine learning techniques can be a good solution to improve a classification accuracy performance.

REFERENCES

[1] Jihoon Kwon and Nojun Kwak, " Human detection by neural networks using a low-cost short-range Doppler radar sensor", 2017 IEEE Radar Conference, Seattle, WA, May 2017.

[2] Browning, R. C., Baker, E. A., Herron, J. A. and Kram, R, "Effects of obesity and sex on the energetic cost and preferred speed of walking", Journal of Applied Physiology, 100 (2), pp. 390398, 2006

[3] Y. Kim, S. Ha and J. Kwon, Human detection using Doppler radar based on physical characteristics of targets, IEEE Geoscience and Remote Sensing Letters, vol.12, pp. 289 293, Feb. 2015.

[4] Y. Kim and H. Ling, "Human activity classification based on micro-Doppler signatures using an artificial neural network." 2008 IEEE Antennas and Propagation Society International Symposium, 2008.

[5] Van Dorp, Ph, and F. C. A. Groen, "Feature-based human motion parameter estimation with radar." IET Radar, Sonar and Navigation, Vol. 2 no. 2, pp. 135-145, 2008.

[6] Goodfellow, I., Bengio, Y., Courville, A., and Bengio, Y, Deep learning (Vol. 1), Cambridge: MIT press, 2016.

[7] Schmidhuber, Jrgen. "Deep learning in neural networks: An overview", Neural networks 61 pp. 85-117, 2015.