# Deep Learning-based identification of human gait by radar micro-Doppler measurements

Papanastasiou, V. S.; Trommel, R. P.; Harmanny, R. I.A.; Yarovoy, A.

**DOI**
10.1109/EuRAD48048.2021.00024

**Publication date**
2021

**Document Version**
Final published version

**Published in**
EuRAD 2020 - 2020 17th European Radar Conference

# Deep Learning-based identification of human gait by radar micro-Doppler measurements

V. S. Papanastasiou[1], R. P. Trommel[2], R. I. A. Harmanny[2], A. Yarovoy[1]
[1]Microwave Sensing, Signals and Systems
Delft University of Technology
Delft, The Netherlands
[2]Thales Nederland B.V.
Delft, The Netherlands
vasilispapanasta@gmail.com, {roeland.trommel, ronny.harmanny}@nl.thalesgroup.com, a.yarovoy@tudelft.nl

***Abstract*— For the first time identification of human individuals using micro-Doppler (m-D) features measured at X-band has been demonstrated. Deep Convolutional Neural Networks (DCNNs) have been used to perform classification. Inspection and visualization of the classification results were performed using Uniform Manifold Approximation and Projection (UMAP). Classification accuracy of above 93.5% is obtained for a population of 22 subjects. The results show that human identification on a specific population based on X-band m-D measurements can be performed reliably using a DCNN.**

***Keywords*— micro-Doppler, deep learning, radar, identification, classification.**

## I. Introduction

Due to security and surveillance reasons, the interest in sensing applications to human identification and classification is increasing more and more. The use of radar for this purpose has become popular because of its tolerance to environmental conditions such as low light, precipitation and its potentially long operation range. Human identification can be performed based on the micro-Doppler (m-D) features of the radar signal caused by the locomotion of a human body. One way to observe and exploit m-D features is by the spectrogram representation of a time-frequency response obtained by several consecutive Doppler measurements. The spectrogram is computed by applying Short Time Fourier Transform (STFT) to coherent radar measurements with sufficient sampling rate.

Human gait classification based on micro-Doppler has attracted the interest of many researchers. Non-machine learning approaches were used in [1] and [2] proposing a particle filter method and handcrafted features respectively. In [3] and [4] machine learning approaches were used to solve gait classification problems such as activity and walking style classification. Due to their capability to learn features by themselves, deep learning based methods have gained popularity in several fields. In [5]-[6] Deep Convolutional Neural Networks (DCNNs) were used to deal with more challenging problems such as personnel recognition based on multistatic m-D and multi-target human gait classification revealing the potential of such networks. In [7] and [8] DCNNs were used successfully for person identification based on human gait K-band m-D measurements revealing the potentials of DCNNs for human identification. However, identification accuracy of above 89% was achieved for less than 10 subjects in [7] and 98% in [8] considering that the subjects were walking on a treadmill.

This study aims to prove that the human walking gait differs between individuals and that it can be used as signature for personnel identification using X-band radar measurements, which is a widely used frequency band in operational (military) radar systems with a much coarser Doppler resolution than those at K-band. Also, this study employs a more realistic and challenging human identification scenario with subjects walking inbound and outbound while observed by a radar. Section II describes the data acquisition and the Deep Convolutional Neural Network (DCNN) used. The experimental results are described and analyzed in Section III. Section IV presents the conclusions.

## II. Input Data and DCNN Classifier

### A. Radar System and Experimental Setup

An X-band monostatic Continuous Wave (CW) radar at 10 GHz was used to measure responses of 22 subjects, 16 male and 6 female, varying in height between 1.67m and 2.07m and age between 21 and 55 years old. Table 1 includes the gender and height information per subject. The subjects were walking in full inbound/outbound target trajectory along the Line of Sight (LoS) path of the radar. The distance of each walk was approximately 25m and each subject was starting its walk at 3m distance from the radar resulting in measurements of various SNR levels. Each subject made three inbound and outbound walks. During each measurement only one subject participated. The experimental setup is depicted in Figure 1.

### B. Measured Data and Preprocessing

By using STFT with a 128-point FFT, Hanning windowing and an overlap of 90% between continuous integration intervals, spectrograms from measurements of entire inbound or outbound walks are obtained. Final spectrogram images of dimensions 128x192 Doppler and time bins are obtained by cropping chunks from a spectrogram of an entire walk with an

2020 17th European Radar Conference (EuRAD) | 978-2-87487-061-3/21/$31.00 ©2021 EuMA | DOI: 10.23919/EuRAD48048.2021.00024

overlap of 50% between consequent chunks. Each spectrogram corresponded to 1.25s time duration. The baseband radar signal is sampled at 8 kHz. The In-phase and Quadrature-phase (I/Q) imbalance is corrected by adjusting the magnitude and the phase of time signal. Then, the I/Q corrected signal is decimated to 2 kHz, high-pass filtered to suppress static clutter and cropped in outbound and inbound parts. For each subject approximately equal amount of spectrograms were gathered. In total, a dataset of 12083 spectrograms, of which 6058 inbound and 6025 outbound, was created. Figure 3 shows the spectrograms of the inbound walking gait from a male and a female subject.

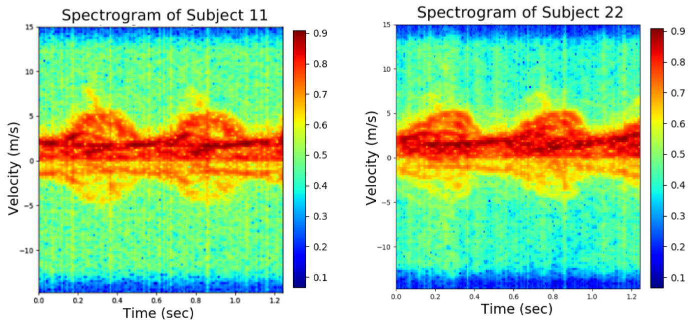
Table 1. Gender and height per subject.

| Subject | Gender | Height (m) | Subject | Gender | Height (m) |
|---|---|---|---|---|---|
| #1 | M | 1.79 | #12 | F | 1.75 |
| #2 | M | 1.73 | #13 | M | 1.85 |
| #3 | M | 1.87 | #14 | M | 1.70 |
| #4 | M | 2.07 | #15 | M | 1.80 |
| #5 | M | 1.86 | #16 | M | 1.82 |
| #6 | M | 1.82 | #17 | F | 1.70 |
| #7 | M | 1.86 | #18 | F | 1.71 |
| #8 | M | 1.85 | #19 | M | 1.83 |
| #9 | M | 1.93 | #20 | M | 1.82 |
| #10 | F | 1.74 | #21 | F | 1.67 |
| #11 | M | 1.80 | #22 | F | 1.69 |

Fig. 1. Experimental setup during measurements.

## C. DCNN Classifier

The DCNN is a form of neural network [9] specially designed to classify images and are considered the most successful network architecture for this kind of tasks to date [7][8]. Thus, a DCNN classifier architecture was selected, specifically the VGG-16 [10]. Shallower configurations lead to worse classification accuracy while a deeper one did not lead to better results increasing the number of parameters, the computational complexity and training time. Latest architectures, such as ResNet-50 [11], were also investigated resulting in slightly worse results. The VGG-16 DCNN was implemented using Keras [12] and TensorFlow [13]. As shown in Figure 2, the classifier consists of 22 layers including 13 convolutional layers with filter size 3x3, five 2x2 max pooling layers, three fully connected layers and a final softmax layer for the classification. Each layer except the softmax max pooling and the max pooling layers use the Rectified Linear Unit (ReLU) as activation function and Dropout with probability of 0.5 was applied to the first two fully connected layers.




(a) (b)

Fig. 3. Spectrograms of inbound walking gaits from male and female subjects. a) m-D signature of subject 11. b) m-D signature of subject 22.

# III. Experimental Results

## A. Training and Classification Result

The VGG-16 classifier was trained using a NVIDIA GeForce GTX TITAN X GPU. The data were split into training (88%) and test (12%) sets. A tenth of the training set was used for validation during training process and had never been seen by the classifier. Because it was crucial to ensure that training and testing sets were uncorrelated, overlapping spectrograms were strictly included either to the one or the other set. The classifier was trained over 500 epochs using Adam optimizer with initial learning rate of $10^{-5}$, learning rate decay equal

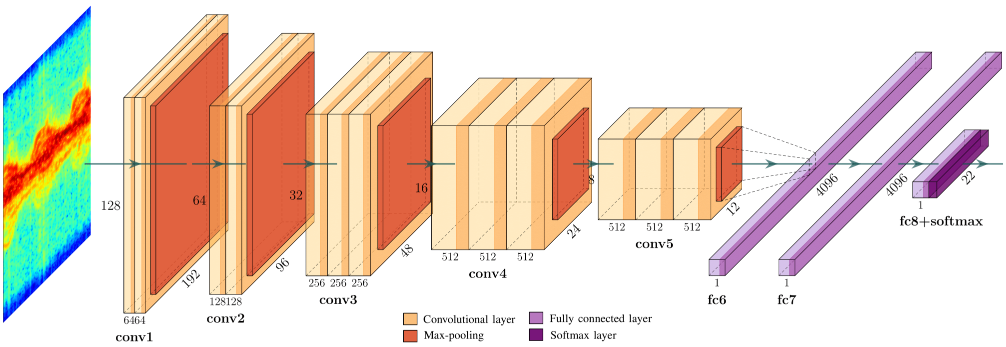



Fig. 2. The VGG-16 architecture that was implemented to perform m-D signature classification/identification.

to $10^{-6}$ and mini batches of 32 spectrograms. Training the DCNN took about 14 hours.

Figure 4 illustrates the normalized confusion matrix of the classification results. In this case, an overall classification accuracy of above 93.5% was achieved. Also, the diagonal of the normalized confusion matrix indicates the individual classification accuracies for each of the 22 subjects. All the individual accuracies are above or equal to 83% while 17 of them are above or equal to 90%. This proves that the trained VGG-16 classifier can identify the differences between the walking gaits of 22 individuals classifying them with high accuracy. Inspecting the individual classification accuracies, it can be observed that 5 out of 6 female subjects achieved individual classification accuracies of above or equal to 95%. From this observation, one can conclude that the classifier distinguishes the walking gait of female subjects from those of male very accurately. The overall high classification accuracy shows that CW X-band radar measurements are sensitive enough to capture sufficient personal traits of the walking gait for population of 22 persons. The different human walking gait traits result in discriminative m-D signatures of the human walking gait per individual enabling personnel identification with high accuracy. Also, the high classification accuracy was achieved using measurements with varying SNR levels. This is considered a more realistic scenario than the one examined in previous studies in which high classification accuracy was achieved only in case of a treadmill walk scenario using K-band radar. The novelty of this result is that human identification based on X-band m-D measurements is now proven.

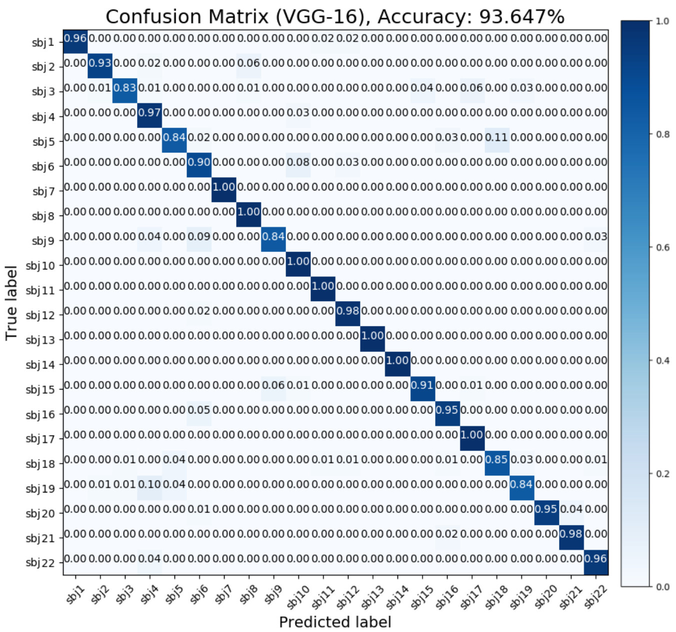
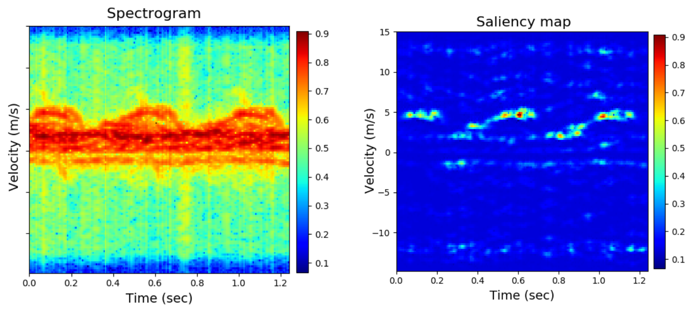



Fig. 4. Normalized confusion matrix and classification accuracy of the m-D signature VGG-16 classifier.

## B. Saliency Map

Saliency maps based on guided back propagation [14] were used to gain insight into the behavior of the trained VGG-16 classifier. Given an input spectrogram a saliency map determines which pixels of the spectrogram need to be changed the least to affect the classifier's prediction scores the most. Figure 5 depicts the saliency map of the m-D signature spectrograms of an inbound walking gait. The saliency map shows that the VGG-16 classification results are determined by the high frequency regions of the m-D signature, while the region corresponding to the torso contribution is not relevant. The high frequency components of the m-D signatures are induced by the motion of legs and feet, which is linked to gait features such as the walking speed and stride length. The classifier learns to use regions of the m-D signature which are correlated with the theoretically two strongest features [15] of the walking gait.




(a) (b)

Fig. 5. Spectrogram and saliency map of inbound human gait. a) Spectrogram. b) Saliency map.

After inspection, it was figured out that some of the misclassified spectrograms were saturated belonging to subjects 3, 5, 9, 18 and 19. These are classes with the lowest classification accuracies. The cause of that is that the used radar is less robust in terms of saturation, in particular for objects at close range. More thorough inspection of the dataset and removal of the saturated spectrograms beforehand is expected to improve the individual classification accuracies and as a result the overall classification accuracy.

## C. UMAP Scatter Plot

To visualise and inspect the behaviour of the classifier, the Uniform Manifold Approximation and Projection (UMAP) [16] dimensionality reduction method was used. Figure 6 depicts the 2-dimensional space representation of the testing data after applying the UMAP method on the output of the last fully connected layer of the classifier. UMAP tries to learn the manifold structure of the high-dimensional data (spectrograms) and find a low-dimensional embedding that preserves the topological structure of that manifold. In this case, dimensionality reduction from 4096-dimensional to 2-dimensional space, which is described by Feature 1 and Feature 2, was performed. It is observed that measurements are mapped into the 2-dimensional space forming clusters per subject. The formed clusters are colored by gender and labeled by each subject's number. However, the distance between the clusters in the 2-dimensional space should not be interpreted as a similarity metric. Misclassified data points from opposite gender are distinguishable as well. In clusters corresponding to subjects with lower individual classification accuracy, more misclassified data points are observed. Also, the clusters associated to the male and the female subjects (blue group and pink clusters respectively) seem to be separated between them creating two groups. This shows that the VGG-16 classifier extracted features which incorporate the gender information. This agrees with the classification results in Figure 4 which shows that the classifier distinguishes the walking gait of female subjects from those of male with high accuracy.


Fig. 6. UMAP plot of the 2-dimensional space representation of the measured testing data using the softmax output of the trained VGG-16 classifier. The 22 clusters are labeled by subject number.

## IV. Conclusion

In this paper, m-D signatures of walking humans measured by X-band monostatic CW radar have been employed to perform human gait identification. Twenty two subjects walking in full inbound/outbound target trajectory have been measured. A VGG-16 DCNN classifier has been applied to spectrograms of m-D signatures in order to classify them. Classification accuracy of above 93.5% was achieved demonstrating for the first time that human identification based on X-band m-D measurements can be performed using a DCNN. The UMAP dimensionality reduction method was used to project and visualize the measured data on a 2-dimensional space validating the classification result. For future work, lower frequency radars such as S-band could be used. Also, investigation on the effect of increasing the group size (number of persons) and its composition (variety of age groups) could be performed.

## Acknowledgement

The work presented in this paper was conducted as part of a MSc thesis facilitated by Thales Nederland B.V. and the Microwave Sensing, Signals and Systems group of Delft University of Technology. The authors would like to acknowledge dr. D.M.J Tax of Delft University of Technology for his valuable comments and suggestions. The designing tool for neural network diagrams from H. Iqbal [17] was used to design Figure 2.

## References

[1] S. Groot, R. Harmanny, H. Driessen, and A. Yarovoy, "Human motion classification using a particle filter approach: multiple model particle filtering applied to the micro-doppler spectrum," *International Journal of Microwave and Wireless Technologies*, vol. 5, no. 3, pp. 391–399, 2013.

[2] F. Fioranelli, M. Ritchie, S. Z. Gürbüz, and H. Griffiths, "Feature diversity for optimized human micro-doppler classification using multistatic radar," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 53, no. 2, pp. 640–654, April 2017.

[3] 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, May 2009.

[4] A.-K. Seifert, L. Schäfer, A. Zoubir, and M. Amin, "Subspace classification of human gait using radar micro-doppler signatures," 08 2018.

[5] Y. Kim and T. Moon, "Human detection and activity classification based on micro-doppler signatures using deep convolutional neural networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 1, pp. 8–12, Jan 2016.

[6] R. P. Trommel, R. I. A. Harmanny, L. Cifola, and J. N. Driessen, "Multi-target human gait classification using deep convolutional neural networks on micro-doppler spectrograms," in *2016 European Radar Conference (EuRAD)*, Oct 2016, pp. 81–84.

[7] P. Cao, W. Xia, M. Ye, J. Zhang, and J. Zhou, "Radar-id: human identification based on radar micro-doppler signatures using deep convolutional neural networks," *IET Radar, Sonar Navigation*, vol. 12, no. 7, pp. 729–734, 2018.

[8] S. Abdulatif, F. Aziz, K. Armanious, B. Kleiner, B. Yang, and U. Schneider, "A study of human body characteristics effect on micro-doppler-based person identification using deep learning," *CoRR*, vol. abs/1811.07173, 2018.

[9] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–44, 05 2015.

[10] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *CoRR*, vol. abs/1409.1556, 2014.

[11] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *CoRR*, vol. abs/1512.03385, 2015.

[12] F. Chollet *et al.*, "Keras," https://keras.io, 2015.

[13] M. Abadi *et al.*, "Tensorflow: A system for large-scale machine learning," in *Proceedings of the 12th USENIX Conference on Operating Systems Design and Implementation*, ser. OSDI'16. Berkeley, CA, USA: USENIX Association, 2016, pp. 265–283.

[14] J. Tobias Springenberg, A. Dosovitskiy, T. Brox, and M. Riedmiller, "Striving for simplicity: The all convolutional net," 12 2014.

[15] C. Moler, "The Eigenwalker Model of the Human Gait," Apr. 2016. [Online]. Available: https://blogs.mathworks.com/cleve/2016/04/11/the-eigenwalker-model-of-the-human-gait/

[16] L. McInnes, J. Healy, and J. Melville, "UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction," *ArXiv e-prints*, Feb. 2018.

[17] H. Iqbal, " Latex code for making neural networks diagrams," Jul. 2019. [Online]. Available: https://github.com/HarisIqbal88/PlotNeuralNet