

Article

MM-Wave Radar-Based Recognition of Multiple Hand Gestures Using Long Short-Term Memory (LSTM) Neural Network

Piotr Grobelny¹  and Adam Narbudowicz^{1,2,*} 

¹ Department of Telecommunications and Teleinformatics, Wrocław University of Science and Technology, 50-370 Wrocław, Poland; 241654@student.pwr.edu.pl

² CONNECT Centre, Trinity College Dublin, The University of Dublin, D02 PN40 Dublin, Ireland

* Correspondence: adam.narbudowicz@pwr.edu.pl

Abstract: The paper proposes a simple machine learning solution for hand-gesture classification, based on processed MM-wave radar signal. It investigates the classification up to 12 different intuitive and ergonomic gestures, which are intended to serve as a contactless user interface. The system is based on AWR1642 boost Frequency-Modulated Continuous-Wave (FMCW) radar, which allows capturing standardized data to support the scalability of the proposed solution. More than 4000 samples were collected from 4 different people, with all signatures extracted from the radar hardware available in open-access database accompanying the publication. Collected data were processed and used to train Long short-term memory (LSTM) and artificial recurrent neural network (RNN) architecture. The work studies the impact of different input parameters, the number of hidden layers, and the number of neurons in those layers. The proposed LSTM network allows for classification of different gestures, with the total accuracy ranging from 94.4% to 100% depending on use-case scenario, with a relatively small architecture of only 2 hidden layers with 32 neurons in each. The solution is also tested with additional data recorded from subjects not involved in the original training set, resulting in an accuracy drop of no more than 2.24%. This demonstrates that the proposed solution is robust and scalable, allowing quick and reliable creation of larger databases of gestures to expand the use of machine learning with radar technologies.

Keywords: radar; MM-wave radar; FMCW radar; hand-gesture recognition; machine learning; neural network; recurrent neural network



Citation: Grobelny, P.; Narbudowicz, A. MM-Wave Radar-Based Recognition of Multiple Hand Gestures Using Long Short-Term Memory (LSTM) Neural Network. *Electronics* **2022**, *11*, 787. <https://doi.org/10.3390/electronics11050787>

Academic Editor: Han Lim Lee

Received: 29 January 2022

Accepted: 27 February 2022

Published: 3 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Development of gesture recognition technology has found widespread applications, ranging from virtual reality, gaming, automotive, wearable devices, to smartphones and medical [1–5]. Recently, the COVID-19 pandemic highlighted further benefits of the contactless interface as a preventive measure against the spread of the virus. Contactless hand-gesture recognition can be implemented within various devices, i.e., optical cameras [6] or gesture-based controllers and radars [7]. Currently, the optical camera-based approach is the most widespread. The camera sensor captures the hand gestures and feeds the classification algorithm through image/video processing [8]. However, optical-based gesture recognition has many limitations in both hardware and software: The camera sensor is sensitive to lighting conditions, while background noise from dirt, weather, dim and bright light or scratching of the lens affects the collected data. On top of that, camera-based machine learning may be subjected to adversary attacks [9], with malicious visual input being easily injected into the system.

Radar-based gesture recognition is recently gaining popularity, as it is resistant to light and weather conditions. Moreover, a single sensor allows for seamless collection of data within 2D and 3D spaces, as well as velocity data from Doppler measurements. Those additional parameters can create more sophisticated classification and increase the overall accuracy.

However, machine learning algorithms used in radar signal processing are still a nascent field. Current literature reports a classification of up to 10 motion gestures with >90% accuracy [7,10–19]. In [10], it is discussed that the accuracy of such algorithms may drop by up to 40% when the classification is executed on samples from a subject not included in the training set. This emphasizes the need for large-scale open and portable data-sets, that are compatible with a number of standardized radar hardware. In [20], a classification within a set of 5 gestures demonstrated a good performance, achieving 97.6% accuracy. The work gathered 300 samples, however, with no information about the number of subjects involved. Another work [7] successfully demonstrates the classification of 6 basic hand gestures. The training set also includes some random movements, which allowed explicit classification of ‘invalid’ gestures.

In most of the reported approaches, the raw data captured by the radar are converted into spectrogram images and fed into a Convolution Neural Network (CNN) optimized for the image classification problem. While using CNN for image classification is a well studied subject, it creates large amounts of stored and processed data, which require feature extraction in the post-processing step [21].

This publication proposes a classification algorithm of up to 12 basic gestures with classification accuracy above 94%. The motivation is to create an ergonomic and intuitive contactless interface for vending machines, ATMs, and similar devices. The work studies the impact of different parameters on the neural network proposed, as well as a trade-off between accuracy and the number of gestures involved. Overall accuracies are within 94.4% to 100%, depending on the number of gestures.

Distinguishable to previous contributions, the work uses standardized off-the-shelf radar hardware, the AWR164 BOOST from Texas Instrument [22]. This allows for machine learning and data-sets that are transferable across a number of off-the-shelf radar hardware. This ensures repeatability, easy benchmarking between different solutions, and the possibility to gradually build large training data-sets from multiple sources. All recorded samples reported in the study are openly available in [23]. The proposed algorithm is based on Long short-term memory (LSTM) neural network, which is better tailored for sequential datasets. The proposed work offers 94.4% accuracy for 12 gestures and 97.3% for 10 gestures.

2. Materials and Methods

The gesture recognition system is outlined in Figure 1. It is based on AWR1642 Booster Pack evaluation board from Texas Instrument [22], which is a Frequency Modulated Continuous Wave (FMCW) radar operating within 76–81 GHz band. It works in a MIMO configuration of 2 transmitters and 4 receivers. High resolution (up to 4 cm), compact size, relatively low cost, ready-to-use software, and USB interface makes it a good candidate for the proposed system. The onboard digital signal processor (DSP) processes the raw data from the radar with the Fast Fourier Transform (FFT). Processed data is transferred into a specific text data format, which contains information about the detected points, eliminating the need for external algorithm for data processing. Radar hardware classifies targets and delivers data as frames. Each frame corresponds up to 200 targets collected in a single sensing epoch. Each target contains a set of parameters: x and y coordinates (m), distance from the radar (m), velocity (m/s) and power of the reflected signal (in arbitrary units). Radar collects the data in a preconfigured speed with a maximum speed of 30 frames per second. Since the radar operates at a close range of 10–40 cm, each detected target corresponds to some area of the hand (e.g., finger), however, there is no tracking of a particular targets over multiple frames.

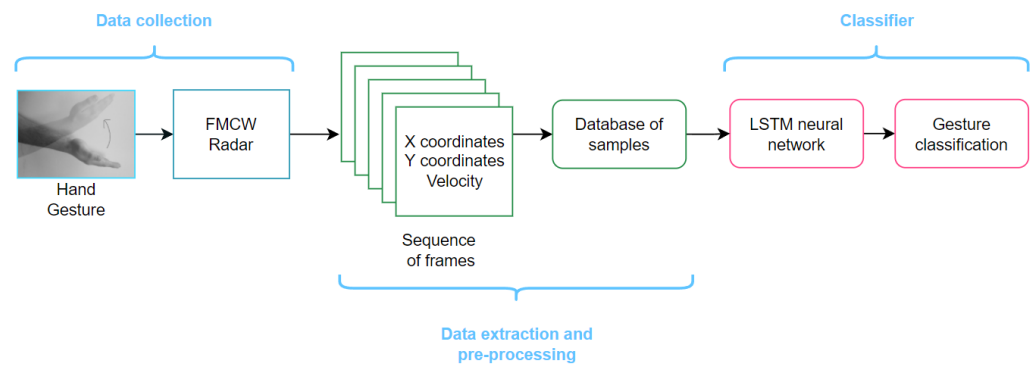


Figure 1. Pipeline of the proposed system. The hand gesture is captured by the radar and transformed by the AWR1642 hardware into samples consisting of frames that are eventually used with the LSTM neural network.

2.1. Experimental Setup

The operation of the proposed system was tested with subjects in a sitting position in front of the radar. The starting position of the hand before performing the gesture was at a distance of 10–40 cm from the antenna. AWR1642 was placed vertically with the transmit and receive antennas pointed towards the human subject (see Figure 2).

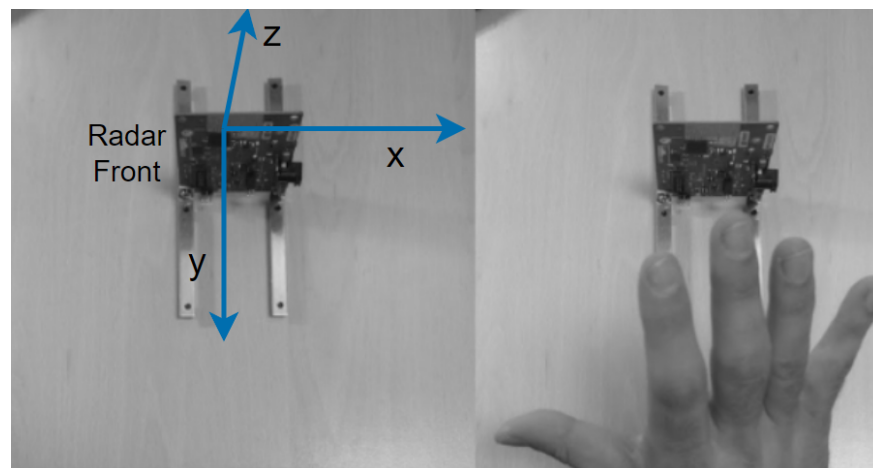


Figure 2. Position of the hand against the radar.

To prevent the collection of clutters behind the human hand, the X and Y thresholds have been established. The range of motion when capturing a gesture may vary from person to person. Therefore, we set the y -axis cut-off at 60 cm, which allows a sufficient margin of 20 cm while performing the hand movement. The software cut off all targets beyond 0.6 m in Y axis and outside the range from -0.3 m to 0.3 m in X axis. As the interface works in close range, the Best Range Resolution mode of AWR1642 was chosen. The frame rate was set up to 30 frames per second, which is considered optimal for the investigated application. A lower value can reduce the accuracy of the neural network, and a higher value creates too much data to process. Other important parameters are shown in Table 1.

Table 1. AWR1642 parameters.

Parameter	Value
Antenna configuration	2TX, 4RX
Azimuth resolution	15
Range resolution	0.039 m
Radial velocity resolution	0.13 m/s
Frame duration	33 ms
Range detection threshold: 30 dB	30 dB
Doppler detection threshold: 30 dB	30 dB

RX—receiver, TX—transmitter.

2.2. Implemented Gestures

The gestures used in the study were selected with an intention for intuitive control of a contactless interface that can be used by any member of the public. The following list of 12 gestures was compiled for the study, all illustrated in Figure 3: (G1) arm to left—full swipe of an arm from right to left, (G2) arm to right—full swipe of an arm from left to right, (G3) hand away—taking a hand away from radar, (G4) hand closer—taking a hand closer to the radar, (G5) arm up—an arm movement from bottom to top, (G6) arm down—an arm movement from top to bottom, (G7) palm up—rotating a palm upwards, (G8) palm down—rotating a palm downwards, (G9) hand to the left—a hand movement to the left (without an arm movement), (G10) hand to the right—a hand movement to the right, (G11) closing a fist horizontally, (G12) closing a fist vertically. All gestures are simple to perform, easily explainable to subjects without prior training, and can intuitively be assigned to different actions of the interface.

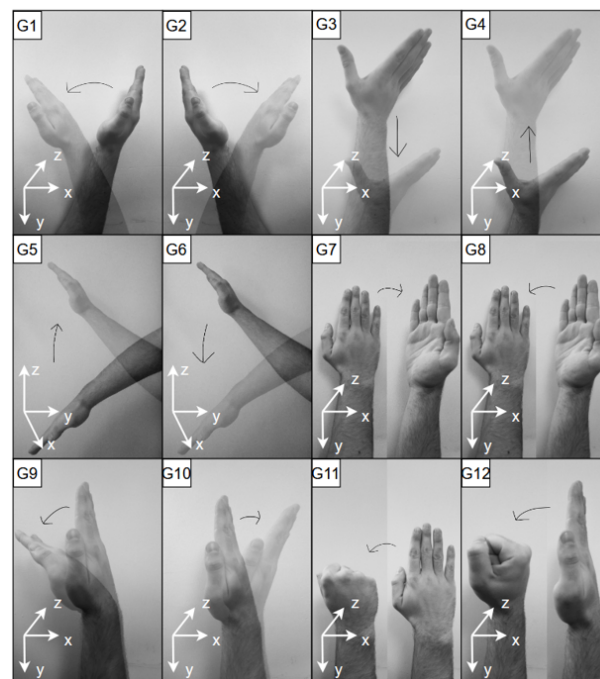


Figure 3. Gestures: (G1) arm to left, (G2) arm to right, (G3) hand away, (G4) hand closer, (G5) arm up, (G6) arm down, (G7) palm up, (G8) palm down—rotating palm down, (G9) hand to the left, (G10) hand to the right, (G11) closing a fist horizontally, (G12) closing a fist vertically.

Four subjects were involved in collection of the data samples: three females and one male. Overall, 4600 samples were collected, with the detailed distribution between subjects shown in Table 2.

Table 2. Distribution of samples per person.

Gesture Type	Person1	Person2	Person3	Person4	Sum
Arm to left	100	100	100	100	400
Arm to right	100	100	100	100	400
Closing fist horizontally	100	100	100	100	400
Close fist perpendicularly	150	50	100	100	400
Hand away	200	100	100	0	400
Hand closer	100	100	100	100	400
Hand down	100	100	100	100	400
Hand up	100	100	100	100	400
Hand rotating palm down	300	0	100	0	400
Hand rotating palm up	300	0	100	0	400
Hand to left	100	100	100	0	300
Hand to right	100	100	100	0	300

2.3. Data Processing

Configured, AWR1642 streams the processed data via USB interface to the controlling PC. Decoded information is a frame filled by targets' parameters. A single frame is a $3 \times N$ dimensional matrix, where N stands for the number of detected targets. Three columns correspond to the respective parameters: Doppler velocity (m/s), X and Y position (m). For machine learning, those values are considered intermediate parameters, i.e., one is not concerned about the error of each individual value as long as their cumulative effect yields the recognition of the correct gesture. Each gesture consists of M frames, creating a matrix of $3 \times N \times M$ size. The examples of the matrix representation of the detected targets with their respective pixel vs. frame image are shown in Figure 4. The program automatically transfers a gesture matrix into CSV file, which contains a single gesture sample per file. Each file stores 20–40 kB of data. It allows training directly on the original digitized radar-based data, without generating a spectrogram-like image and adopting the techniques used for visual image recognition. Overall, the above features make the training process more computationally efficient. The database with all gestures is freely available at IEEE Data Port [23].

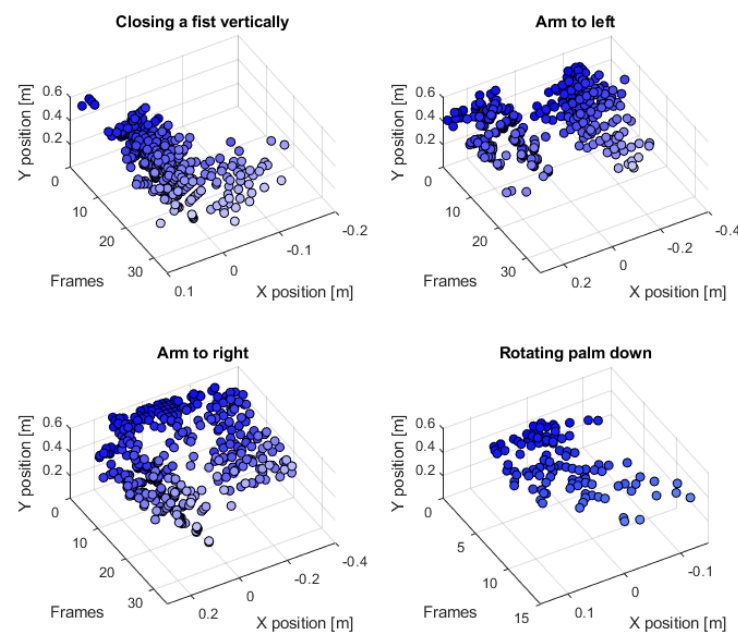


Figure 4. Exemplary gesture samples recorded by the radar with detected targets along X and Y position for each frame.

2.4. Gesture Classification Algorithm

The LSTM network was studied with two use-cases: for the ‘simplified’ case only data about X and Y position of each target and frame were included, while the ‘normal’ case included data about X and Y position, as well as the measured Doppler velocity. Each gesture sample from the database was converted from a matrix into a vector of 19,200 elements for the ‘normal’ case (i.e., including X position, Y position and Doppler velocity) or 12,800 for the ‘simplified’ case (i.e., without considering the Doppler effect). To overcome the varying number of detected targets for different time-slots, zero-padding was implemented to keep the constant number of 80 frames and 80 targets. Hence, the extracted gestures have a dimension of $80 \times 3 \times 80$ and, $80 \times 2 \times 80$, respectively.

Long short-term memory (LSTM) network for each element in input sequence, each layer computes the following function:

$$\begin{aligned} i_t &= \sigma(W_{ii}x_t + b_{ii} + W_{hi}h_{t-1} + b_{hi}) \\ f_t &= \sigma(W_{if}x_t + b_{if} + W_{hf}h_{t-1} + b_{hf}) \\ g_t &= \tanh(W_{ig}x_t + b_{ig} + W_{hg}h_{t-1} + b_{hg}) \\ o_t &= \sigma(W_{io}x_t + b_{io} + W_{ho}h_{t-1} + b_{ho}) \\ c_t &= f_t \odot c_{t-1} + i_t \odot g_t \\ h_t &= o_t \odot \tanh(c_t) \end{aligned} \quad (1)$$

where h_t is the hidden state at time t , x_t is the input at time t , h_{t-1} is the hidden state of the layer at time $t - 1$ or the initial hidden state at time o and i_t, f_t, g_t, o_t are the input, forget, cell, and output gates respectively. σ is the sigmoid function, and \odot is the Hadamard product. In a multilayer LSTM, the input $x_t^{(l)}$ of the l -th layer ($l \geq 2$) is the hidden state $h_t^{(l-1)}$ of the previous layer multiplied by dropout, $\delta_t^{(l-1)}$ where each $\delta_t^{(l-1)}$ is a Bernoulli random variable which is 0 with probability dropout [24]. Tanh activation function is defined as:

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (2)$$

The output layer provides the probability of input data against each gesture. It uses LogSoftmax activation function to normalize output o_1 of the network to a probability distribution over the predicted output classes.

$$\log \text{Softmax}(o_i) = \log \left(\frac{\exp(o_i)}{\sum_{j=0} \exp(o_j)} \right) \quad (3)$$

The largest value of the output indicates the classified gesture, accordingly the number of nodes of the output layer corresponds to the number of gestures being classified.

2.5. Validation

Each gesture from the validation set was fed into the trained neural network. The algorithm classifies each input as a gesture with the highest probability, i.e., the highest value of the respective output. The overall accuracy was calculated as:

$$\text{Accuracy} = 100\% \left(\frac{N_{\text{correct}}}{N_{\text{total}}} \right) \quad (4)$$

where N_{correct} is the number of correctly recognized gestures and N_{total} is the total number of tested gestures.

3. Results

The database was divided into a training set and a validation set (90% and 10% of the database, respectively). The order of gesture samples in both sets was randomized before

feeding to the neural network, i.e., with random sequence of gesture type and subjects. Ten epochs for the training set were used. Other training parameters are shown in the Table 3. The impact of the Doppler velocity on the LSTM network accuracy was examined by the two use-cases. To find the optimal number of layers and nodes for each layer, the network was trained with 12 types of gestures against different number of hidden layers with different number of neurons. The experiment was repeated for both 'simplified' and 'normal' use cases. The summary of this investigation is shown in Figure 5.

Table 3. Training parameters.

Epochs	10
Loss function	Cross Entropy Loss
Optimizer algorithm	Adam
Optimizer step value	0.001
Machine learning framework	PyTorch

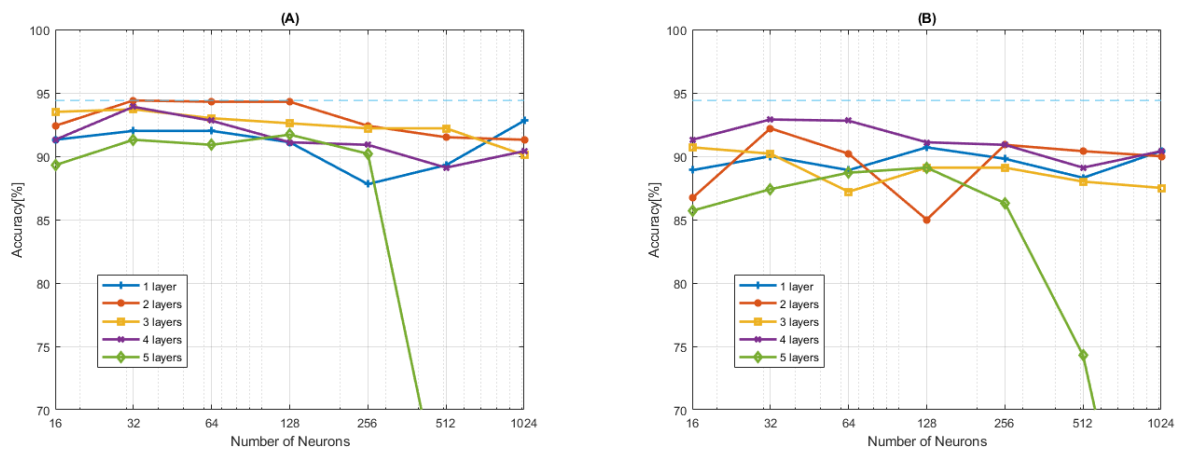


Figure 5. Accuracy vs. number of neurons for different sets of hidden layers. (A) 'Normal' use case with Doppler velocity feature included. (B) 'Simplified' use case without Doppler velocity.

The highest accuracy of 94.4% was achieved by a LSTM network consisting of 2 layers of 32 neurons each for 'normal' case with Doppler velocity. Overall, networks trained with the Doppler velocity data resulted in higher efficiency: An average accuracy decline of the 'simplified' use case in relation to the 'normal' use case ranged from 1.33% to 11.64% depending on the setup. The confusion matrix in Figure 6 shows that, for 12 gestures, a very high accuracy can be obtained for most of the gestures. Gestures G5 (hand down), G6 (hand up), and G11 (closing fist horizontally) remain below 90%. This is because the antennas of the AWR1642 device allows for distinction only in a single plane, measuring targets along X and Y axes. The three gestures in question include significant up and down movements, i.e., along the Z axis. This problem is expected to be solved by using FMCW radar with 2-dimensional antenna array. After eliminating gestures G5 and G6, the experiment was repeated for 10 remaining gestures, which resulted in 97.3% accuracy.

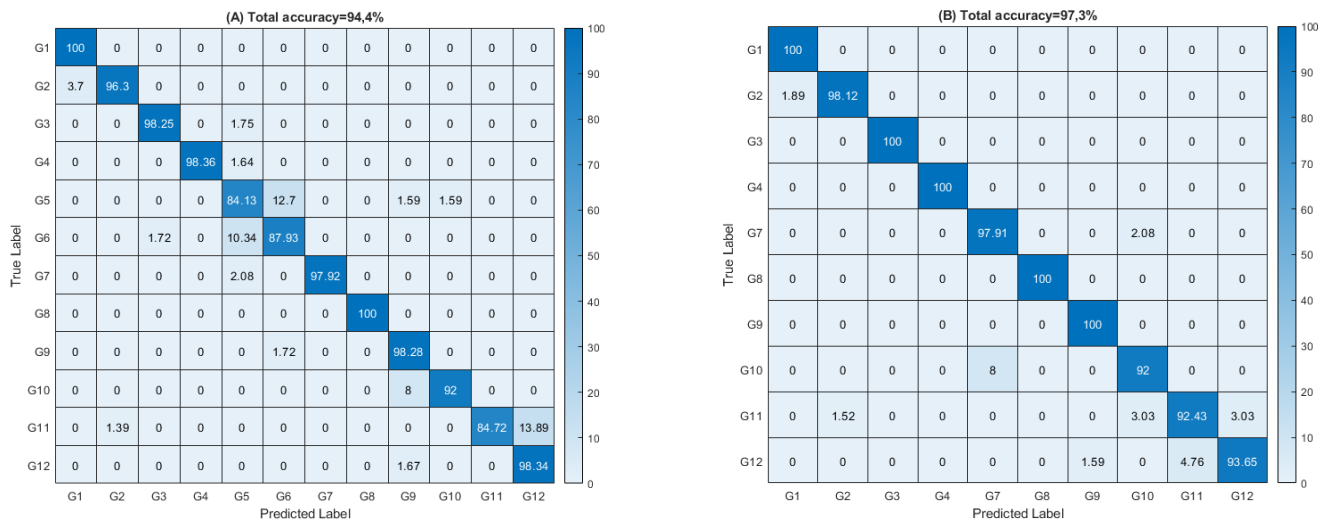


Figure 6. (A) Normalized to percentage confusion matrix for 12 gestures. (B) Normalized to percentage confusion matrix for 10 gestures, without gesture G5 (hand down) and G6 (hand up).

3.1. Performance of Different Network Types

Obtained results were compared with three types of neural networks: Gated Recurrent Units (GRU), Elman Recurrent Neural Network (RNN), and Feedforward neural network. An example of differences in training of two hidden layers architecture is shown in the first plot in Figure 7. RNN obtained the worst results, with declining accuracy as the network becomes more complex. This is caused by Vanishing Gradient problem: For long time-series data, RNN tends to lower the gradient to the value where parameter updates become insignificant and have no real impact on the learning. The solution is to use LSTM and GRU architectures. LSTM obtained slightly better results than GRU. For GRU, LSTM, and Feedforward networks the accuracy is almost linear with peaks for different range of neurons. The second plot in Figure 7 shows how networks behave for a declining number of gestures. LSTM and GRU exceeds 96% accuracy for 10 gestures. RNN and Feedforward networks reach a satisfying level (above 94%) for six and eight gestures, respectively.

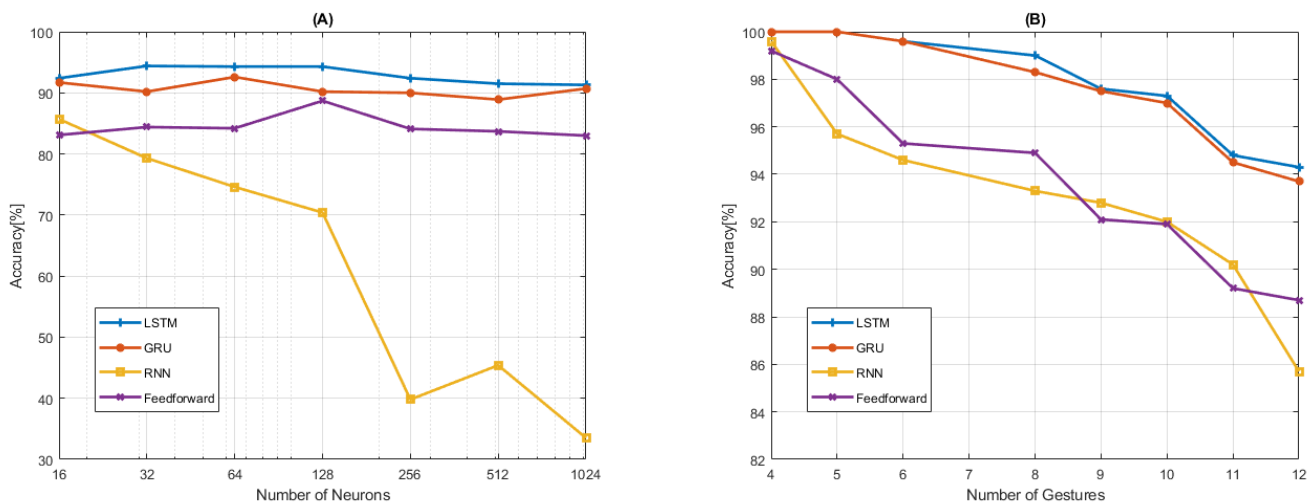


Figure 7. (A) Comparison of accuracy vs. number of neurons between LSTM, GRU, RNN, and Feedforward architecture for two hidden layers. (B) Comparison of accuracy vs. number of gestures between LSTM, GRU, RNN, and Feedforward architecture.

3.2. Performance with New Subjects

To assess how general the proposed solution is, additional samples were collected from two new subjects, who were not involved in the previous study. Each new subject recorded 20 samples per gesture, resulting in 480 new samples to be classified. The new samples were fed to the existing best resulting classification algorithm (i.e., LSTM with 2 hidden layers and 32 neurons each) without any retraining. The confusion matrices are shown in Figure 8. It can be seen that the accuracy is 92.16% and 96.09% for the recognition of 12 and 10 gestures, respectively. This correspond to a respective accuracy drop of 2.24% and 1.21 %, which is considered insignificant. By comparison, the work [10] reported up to a 40% drop when the classification is executed on samples from a subject not included in the training set.

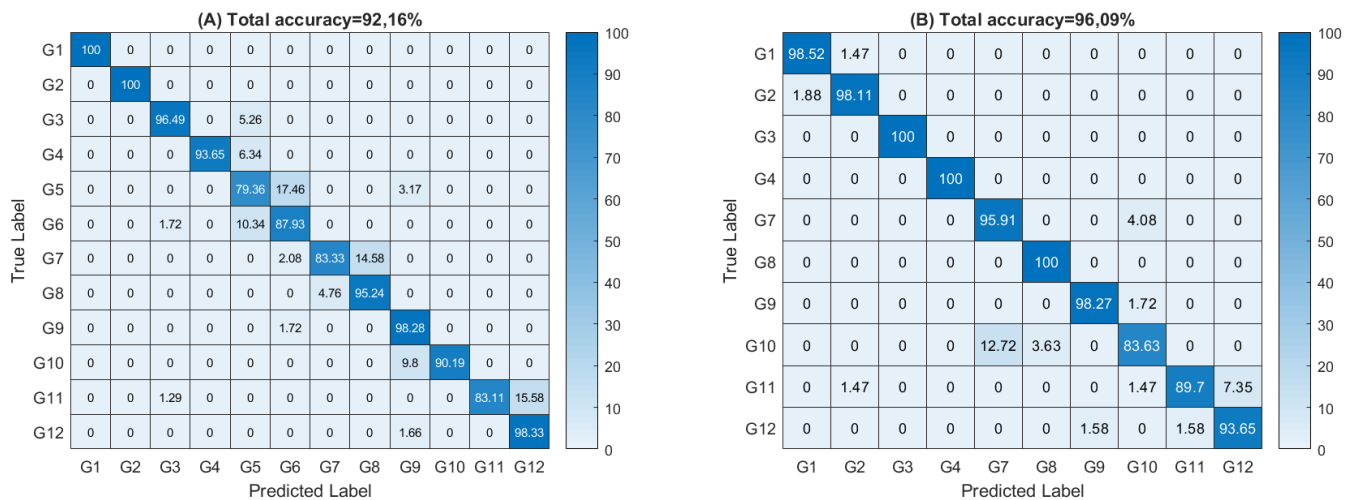


Figure 8. Normalized confusion matrices for the classification of data from subjects not involved in the network’s training. (A) Normalized to percentage confusion matrix for 12 gestures. (B) Normalized to percentage confusion matrix for 10 gestures, without gestures G5 (hand down) and G6 (hand up).

4. Conclusions

The proposed work demonstrates the LSTM network classification algorithm for radar-based signals to correctly classify a number of hand-gestures for use as a contactless interface in vending machines. The network allows the classification of up to 12 different gestures observed with a single radar with 94.4% total accuracy, outperforming comparable radar-based solutions reported in the literature (demonstrated in Table 4). The work also studies the impact of different inputs and number of hidden layers, as well as comparison with alternative neural network types. The proposed solution uses single standard off-the-shelf radar hardware, with neural networks consisting only of two hidden layers. This allows for scalability and repeatability, as well as the inclusion of additional samples into the training sets. The database of 4600 gesture samples used in the study is provided through IEEE Data Port platform [23]. The produced data requires little storage space, with a single sample CSV file of 20–40 kB.

Table 4. Comparison between the authors' method and other methods in the gesture recognition field.

Source	Year	Gestures Number	Accuracy	Classification Algorithm	Radar Type
This work	2022	12	94.3%	LSTM	FMCW
This work	2022	10	97.4%	LSTM	FMCW
[13]	2021	8	98.13%	Depthwise2D+CNN2D	FMCW
[14]	2021	4	98.91%	Spiking LSM	FMCW
[17]	2019	4	96%	KNN	FMCW
[15]	2020	6	94.21%	CNN	CW
[18]	2018	7	91%	LSTM	FMCW
[16]	2020	7	98.8%	3D-CNN	FMCW
[19]	2019	10	96.17%	TS-I3D	FMCW

Author Contributions: Conceptualization and methodology, A.N. and P.G.; software, validation, analysis, investigation, and writing—original draft preparation, P.G.; writing—review and editing, supervision, and funding acquisition, A.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by the Science Foundation Ireland, grant number 18/SIRG/5612.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Anonymized radar data used in this study is available via IEEE Data Port at dx.doi.org/10.21227/wh5w-c362.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Zhao, H.; Wang, S.; Zhou, G.; Zhang, D. Gesture-Enabled Remote Control for Healthcare. In Proceedings of the 2017 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE), Philadelphia, PA, USA, 17–19 July 2017; pp. 392–401.
- Tateno, S.; Zhu, Y.; Meng, F. Hand Gesture Recognition System for In-car Device Control Based on Infrared Array Sensor. In Proceedings of the 2019 58th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE), Hiroshima, Japan, 10–13 September 2019; pp. 701–706.
- Lee, D.H.; Hong, K.S. Game interface using hand gesture recognition. In Proceedings of the 5th International Conference on Computer Sciences and Convergence Information Technology, Seoul, Korea, 30 November–2 December 2010; pp. 1092–1097.
- Rani, S.S.; Dhruv, K.J.; Ahalyadas, M. Hand gesture control of virtual object in augmented reality. In Proceedings of the 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Udipi, India, 13–16 September 2017; pp. 1500–1505.
- Weissmann, J.; Salomon, R. Gesture recognition for virtual reality applications using data gloves and neural networks. In Proceedings of the IJCNN'99. International Joint Conference on Neural Networks. Proceedings (Cat. No.99CH36339), Washington, DC, USA, 10–16 July 1999; Volume 3, pp. 2043–2046.
- Czyszyński, K.; Rumiński, J.; Kwaśniewska, A. Gesture Recognition With the Linear Optical Sensor and Recurrent Neural Networks. *IEEE Sens. J.* **2018**, *18*, 5429–5438. [[CrossRef](#)]
- Liu, C.; Li, Y.; Ao, D.; Tian, H. Spectrum-Based Hand Gesture Recognition Using Millimeter-Wave Radar Parameter Measurements. *IEEE Access* **2019**, *7*, 79147–79158. [[CrossRef](#)]
- Chung, H.-Y.; Chung, Y.-L.; Tsai, W.-F. An Efficient Hand Gesture Recognition System Based on Deep CNN. In Proceedings of the 2019 IEEE International Conference on Industrial Technology (ICIT), Melbourne, VIC, Australia, 13–15 February 2019; pp. 853–858.
- Akhtar, N.; Mian, A. Threat of Adversarial Attacks on Deep Learning in Computer Vision: A Survey. *IEEE Access* **2018**, *6*, 14410–14430. [[CrossRef](#)]
- Skaria, S.; Al-Hourani, A.; Lech, M.; Evans, R.J. Hand-Gesture Recognition Using Two-Antenna Doppler Radar With Deep Convolutional Neural Networks. *IEEE Sens. J.* **2019**, *19*, 3041–3048. [[CrossRef](#)]
- Kim, Y.; Toomajian, B. Hand Gesture Recognition Using Micro-Doppler Signatures With Convolutional Neural Network. *IEEE Access* **2016**, *4*, 7125–7130. [[CrossRef](#)]

12. Zhang, Z.; Tian, Z.; Zhou, M. Latern: Dynamic Continuous Hand Gesture Recognition Using FMCW Radar Sensor. *IEEE Sens. J.* **2018**, *18*, 3278–3289. [[CrossRef](#)]
13. Chmurski, M.; Mauro, G.; Santra, A.; Zubert, M.; Dagan, G. Highly-Optimized Radar-Based Gesture Recognition System with Depthwise Expansion Module. *Sensors* **2021**, *21*, 7298. [[CrossRef](#)] [[PubMed](#)]
14. Tsang, I.J.; Corradi, F.; Sifalakis, M.; Van Leekwijck, W.; Latré, S. Radar-Based Hand Gesture Recognition Using Spiking Neural Networks. *Electronics* **2021**, *10*, 1405. [[CrossRef](#)]
15. Yu, M.; Kim, N.; Jung, Y.; Lee, S. A Frame Detection Method for Real-Time Hand Gesture Recognition Systems Using CW-Radar. *Sensors* **2020**, *20*, 2321. [[CrossRef](#)] [[PubMed](#)]
16. Lee, H.R.; Park, J.; Suh, Y.-J. Improving Classification Accuracy of Hand Gesture Recognition Based on 60 GHz FMCW Radar with Deep Learning Domain Adaptation. *Electronics* **2020**, *9*, 2140. [[CrossRef](#)]
17. Ritchie, M.; Jones, A.M. Micro-Doppler Gesture Recognition using Doppler, Time and Range Based Features. In Proceedings of the 2019 IEEE Radar Conference (RadarConf), Boston, MA, USA, 22–26 April 2019; pp. 1–6.
18. Suh, J.S.; Ryu, S.; Han, B.; Choi, J.; Kim, J.-H.; Hong, S. 24 GHz FMCW Radar System for Real-Time Hand Gesture Recognition Using LSTM. In Proceedings of the 2018 Asia-Pacific Microwave Conference (APMC), Kyoto, Japan, 6–9 November 2018; pp. 860–862.
19. Wang, Y.; Wang, S.; Zhou, M.; Jiang, Q.; Tian, Z. TS-I3D Based Hand Gesture Recognition Method With Radar Sensor. *IEEE Access* **2019**, *7*, 22902–22913. [[CrossRef](#)]
20. Wang, P.; Lin, J.; Wang, F.; Xiu, J.; Lin, Y.; Yan, N.; Xu, H. A Gesture Air-Writing Tracking Method that Uses 24 GHz SIMO Radar SoC. *IEEE Access* **2020**, *8*, 152728–152741. [[CrossRef](#)]
21. Franceschini, S.; Ambrosiano, M.; Vitale, S.; Baselice, F.; Gifuni, A.; Grassini, G.; Pascazio, V. Hand Gesture Recognition via Radar Sensors and Convolutional Neural Networks. In Proceedings of the 2020 IEEE Radar Conference (RadarConf20), Florence, Italy, 21–25 September 2020; pp. 1–5.
22. Texas Instruments. AWR1642 Evaluation Module (AWR1642BOOST) Single-Chip mmWave Sensing Solution. 2017. Available online: <https://www.ti.com/lit/ug/swru508c/swru508c.pdf?ts=1643402432641> (accessed on 1 January 2022).
23. Grobelny, P.; Narbudowicz, A. Hand Gestures Recorded with mm-Wave FMCW Radar (AWR1642). IEEE Dataport. 31 May 2021. Available online: <https://iee-dataport.org/open-access/hand-gestures-recorded-mm-wave-fmcw-radar-awr1642> (accessed on 1 January 2022).
24. pytorch.org. Available online: <https://pytorch.org/docs/1.9.1/generated/torch.nn.LSTM.html> (accessed on 28 January 2022).