Radar-based Human Activities Classification with Complex-valued Neural Networks

Yang, Ximei ; Guendel, Ronny G.; Yarovoy, Alexander; Fioranelli, Francesco

DOI
10.1109/RadarConf2248738.2022.9763903

Publication date
2022

Document Version
Final published version

Published in
2022 IEEE Radar Conference (RadarConf22) Proceedings

Radar-based Human Activities Classification with Complex-valued Neural Networks

Ximei Yang, Ronny G. Guendel, Alexander Yarovoy, Francesco Fioranelli
Microwave Sensing, Signals and Systems (MS3), Delft University of Technology, Delft, The Netherlands

***Abstract*—Human activities classification in assisted living is one of the emerging applications of radar. The conventional analysis considers micro-Doppler signatures as the chosen input for feature extraction or deep learning classification algorithms, or, less frequently, other radar data formats such as the range-time, the range-Doppler, or the Cadence Velocity Diagram. However, these data are typically used as real-valued images, whereas they are actually complex-valued data structures. In this paper, neural networks processing radar data as complex data structures are investigated, with a focus on spectrograms, range-time, and range-Doppler plots as the data formats of choice. Different network architectures are explored both in terms of complex numbers' representations and the depth/complexity of the architecture itself. Experimental data with 9 activities and 15 volunteers collected using an UWB radar are used to test the networks' performances. It is shown that for certain data formats and network architectures, there is an advantage in using complex-valued networks compared to their real-valued counterparts.**

***Index Terms*—Micro-Doppler Classification, Deep learning, Human Activity Recognition, Complex-valued Networks**

I. Introduction

Human activities recognition (HAR) in the context of assisted living is an important and timely need in our aging societies. This includes prompt detection of critical and life-threatening activities such as falls, but also the recording and interpretation of patterns of activities of daily living, with the objective to identify anomalies or changes that can be linked to worsening physical or cognitive conditions. In this context, a number of different sensing technologies have been proposed and validated in recent years, with radar-based technologies increasingly investigated thanks to their contactless sensing capabilities and the potential advantages in terms of perceived privacy (e.g., compared to optical sensors) [1], [2].

The conventional signal processing pipeline for HAR with radar is based on the generation of micro-Doppler signatures, typically spectrograms resulting from the application of Short-Time Fourier Transform (STFT), followed by feature extraction and the application of classifiers such as Support Vector Machine (SVM), decision tree, and K-Nearest Neighbours (kNN), amongst others [2], [3]. In recent years, deep learning approaches for classification driven by the advances in image and audio processing have also been applied to radar data in HAR. These include convolutional neural networks, recurrent neural networks, autoencoders, and their combinations [1], [4]–[6]. A variety of radar data representation formats, e.g., range-time, range-Doppler, Cadence Velocity Diagrams (CVD), radar data cubes, and point clouds, have also been explored in both the traditional machine learning pipeline and with deep learning [7], [8].

However, in most cases, radar data in HAR are processed as real-valued matrices by taking their absolute value and discarding the phase. To the best of our knowledge, only few studied reported in the literature have explored phase information and phase plots as the direct input of classifiers for radar-based HAR, or in conjunction with complex-valued representations of the data, for example [9]–[11], where the phase of spectrograms and range-time plots was used to extract features or directly as input to neural networks. Other studies include [12]–[14] in the wider context of Synthetic Aperture Radar (SAR) images classification and Ground Moving Target Indication (GMTI), and [15] on drones classification using the equivalence of complex numbers and 2D vectors to redefine complex-valued operators and blocks.

In this paper, initial results of complex-valued neural networks (CVNNs) implemented for radar-based HAR are presented. A simple multi-channel architecture to process separately real and imaginary parts or absolute value and phase is compared with actual implementations of complex-valued blocks for convolutional neural networks inspired by [16]. Different depth/complexity of the network architectures are also investigated. The networks are tested with experimental data with 9 activities and 15 volunteers collected with an Ultra Wide Band (UWB) pulsed radar.

The rest of the paper is organized as follows. Section II describes the proposed implementation of the complex-valued networks. Section III presents the experimental setup and the data used to test the networks. Results are discussed in section IV, with conclusions provided in section V.

II. Proposed complex-valued networks

A. Complex-valued neural networks

Complex numbers may have a richer representational capacity than real numbers, but their usage in deep learning architectures has been limited due to the absence of complex-valued building blocks, and the fact that conventional image processing and pattern recognition algorithms typically work with real-valued data. In this paper, two approaches are proposed for the implementation of CVNNs: Multi-Channel Architecture, and Deep Complex Network (DCN) whose implementation is based on [16].

1) Multi-Channel Architecture: Inspired by the conventional processing of RGB images into three separate channels, the multi-channel approach for complex-valued networks splits

2022 IEEE Radar Conference (RadarConf22) | 978-1-7281-5368-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/RADARCONF2248738.2022.9763903

978-1-7281-5368-1/22/$31.00 ©2022 IEEE

either the magnitude and phase or the real and imaginary parts of the input data into two channels. This approach is simple, as it essentially treats the complex number as two independent real numbers, but does not capture the intrinsic relations and dependencies between the two components.

2) Deep Complex Networks: The implementation of the constituent blocks of the DCN is inspired by the work in [16], which presents a general formulation for complex-valued blocks such as complex convolutional layers, complex-valued activation functions, complex batch normalization, and max-pooling. The name "deep" should not deceive, as it is also possible to have architectures with few blocks and layers to make relatively shallow networks, especially compared to the architectures used in image processing. This is an important aspect for radar data processing to avoid overfitting in case of a limited amount of training data.

As some of the scripts in [16] appeared to be outdated, all the source code for the proposed building blocks has been implemented again in PyTorch[1] [17]. With this, complex-valued networks can be realized in a very similar manner to their real-valued counterparts, and the implementation is not tight to specific formats of radar data. The specific operations of each type of block/layer are presented in this section below.

- Complex convolutional layer (*ComplexConv*): A convolution is performed to compute the output of neurons that are connected to local regions of the input. A complex-valued filter matrix $W = A + iB$ is convolved with a complex-valued feature cube $h = x + iy$, so that the output neuron is defined as:

$$W * h = (A * x - B * y) + i(B * x + A * y) \tag{1}$$

Figure 1 illustrates the complex convolutional process.

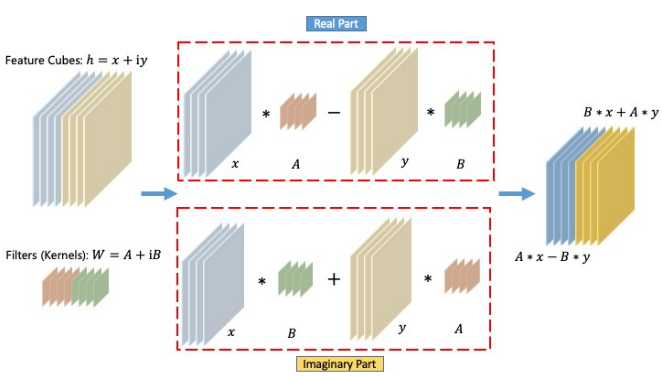


Fig. 1: An illustration of the complex convolution operator

- *Complex-valued activation function*: In [16] three possible complex-valued activation functions were explored. In this paper, we choose to use their proposed CReLU activation function, as their experiments on various image tasks proved that the CReLU function was the most effective. The CReLU is defined as:

$$\mathrm{CReLU}(z) = \mathrm{ReLU}(\Re(z)) + i\,\mathrm{ReLU}(\Im(z)) \tag{2}$$

where z is the complex input neuron.

- Complex batch normalization (*ComplexBN*): Considering that the input has real and imaginary components, whitening 2D vectors $z = (x, y)^T$ can be defined for this building block. This technique was firstly used in the multi-channel signal processing area. The complete expression is shown below:

$$\tilde{\boldsymbol{z}} = (\boldsymbol{V})^{-\frac{1}{2}}(\boldsymbol{z} - \mathbb{E}[\boldsymbol{z}]) \tag{3}$$

where V is the covariance matrix defined as

$$\begin{aligned} \boldsymbol{V} &= \begin{pmatrix} V_{rr} & V_{ri} \\ V_{ir} & V_{ii} \end{pmatrix} \\ &= \begin{pmatrix} \mathrm{Cov}(\Re\{\boldsymbol{z}\}, \Re\{\boldsymbol{z}\}) & \mathrm{Cov}(\Re\{\boldsymbol{z}\}, \Im\{\boldsymbol{z}\}) \\ \mathrm{Cov}(\Im\{\boldsymbol{z}\}, \Re\{\boldsymbol{z}\}) & \mathrm{Cov}(\Im\{\boldsymbol{z}\}, \Im\{\boldsymbol{z}\}) \end{pmatrix} \end{aligned} \tag{4}$$

- Complex-valued max-pooling layer (*ComxPooling*): For a complex data structure $z = x + iy$, x and y are processed separately. The equation for complex-valued max-pooling is defined as:

$$\begin{aligned} &\mathrm{ComxPooling}(z) = \\ &\mathrm{MaxPooling}(\Re(z)) + i\,\mathrm{MaxPooling}(\Im(z)) \end{aligned} \tag{5}$$

A very simple complex-valued CNN is chosen as an example to show the processing pipeline using a combination of the aforementioned DCN blocks. When real numbers are used as inputs, the operations are performed with the sequence of building blocks as follows:

$$Conv \rightarrow BN \rightarrow ReLU \rightarrow MaxPooling \tag{6}$$

When complex numbers are used as inputs, then the following pipeline of complex-valued blocks is used:

$$\begin{aligned} &ComplexConv \rightarrow ComplexBN \rightarrow CReLU \\ &\rightarrow ComxPooling \end{aligned} \tag{7}$$

It should be noted that a DCN network has twice as many parameters as its real-valued counterpart, and that complex arithmetic leads to more operations (e.g., a complex multiplication implies four real-valued multiplications). This potentially increased computational load should be carefully considered as part of the performance evaluation together when comparing real and complex networks.

B. CNN architectures

Three CNN architectures are proposed and investigated to compare real-valued and complex-valued implementations. These are specifically: a shallow ConvNet, a deep ConvNet, and ResNet-18 [18], with details provided in this section.

1) Shallow ConvNet: A basic, simple CNN architecture is demonstrated in Figure 2, consisting of a convolutional (Conv) layer, batch normalization (BN), ReLU layer, pooling (Pool) layer, flattening, fully-connected (FC) layer, ReLU layer, and sigmoid activation function. This shallow network is explicitly designed for the situation of limited data available, with only one CONV layer and is referred to as "Shallow ConvNet" in this paper. The first four layers (CONV, BN, ReLU, Pool) can be regarded as one basic building block.

[1]https://github.com/SherryYang1122/Radar-based-Human-Activities-Classification-with-Complex-valued-Neural-Networks

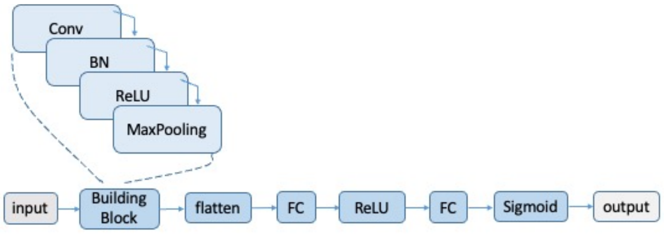


Fig. 2: Shallow ConvNet with only one CONV layer in the network

2) Deep ConvNet: Five building blocks are stacked in this proposed architecture before flattening, in order to make the network deeper. This is referred to as "Deep ConvNet" in this paper, and is illustrated in Figure 3.

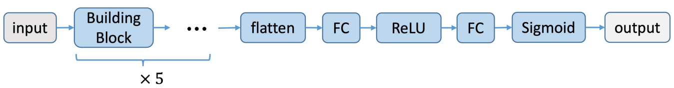


Fig. 3: Deeper ConvNet with 5 building blocks stacked (vs only 1 used in the shallow network)

3) ResNet: Deep ConvNet may lose the benefit and effectiveness of the additional layers due to the vanishing gradient problem. A modern CNN architecture, ResNet, can solve this degradation problem caused by activation functions with residual units to perform "identity shortcut connections". In this case, we mainly referred to the structure and parameters of *ResNet 18-layer*, the simplest model proposed in [18]. This is important for the small size of the radar dataset considered in this paper. A simple sketch of this network is shown in Figure 4.

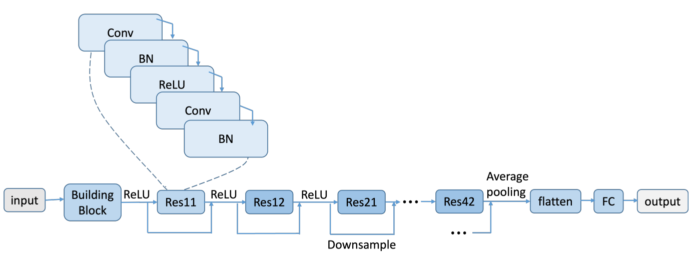


Fig. 4: ResNet 18-layer architecture stacked with one basic building block in Figure 2 and 8 res blocks

III. Experimental Setup and Radar Data Formats

The radar data was collected in the MS3 research laboratory at TU Delft. While data were acquired simultaneously with 5 distributed monostatic radars, only data from one radar are used for the initial results in this paper, as shown in Figure 5. The radar was a monostatic pulsed UWB radar, monitoring the measuring space with a circumference diameter of $4.39\,\text{m}$, highlighted by the blue lines in Figure 5. The radar has a carrier frequency of $4.3\,\text{GHz}$ and bandwidth of $2.2\,\text{GHz}$, and can detect objects in the range from $1\,\text{m}$ to $5.39\,\text{m}$. Its pulse repetition interval (PRI) was $8.2\,\text{ms}$.

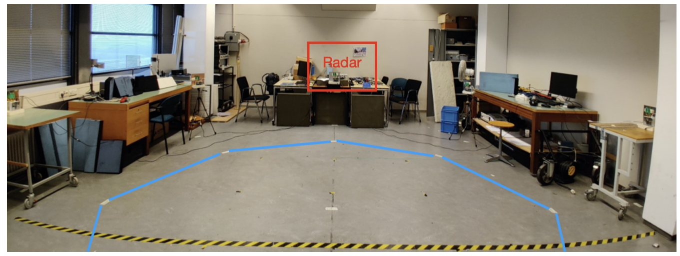


Fig. 5: Location of the experimental setup for data collection

A. Dataset description

Fifteen participants (of which 11 male and 4 female subjects) joined in the experiment to perform nine activity classes, including (1) *walking*, (2) *stationary*, (3) *sitting down*, (4) *standing up from sitting*, (5) *bending from sitting*, (6) *bending from standing*, (7) *falling from walking*, (8) *standing up from ground* and (9) *falling from standing*. It is important to note that the activities were performed as continuous sequences of actions, with random transitions between pairs of activities. Furthermore, the movements' trajectories and the angles faced by the different volunteers were also random [19].

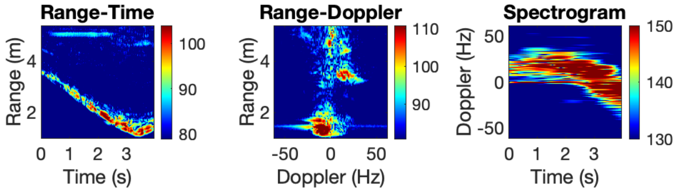
28 measurements were carried out for each participant, where the measurement recorded one person's continuous motion consisting of one (walking) or multiple activities (walking or standing plus other activities) with a total duration of about 2 minutes for each sequence (corresponding to 14634 pulses). Some activities (class 3, 4, 5, 7, 8, 9) were performed at predefined locations with an uncertainty error of approximately $1\,\text{m}$, and other activities were performed freely at random locations within the surveillance area.

B. Radar data formats

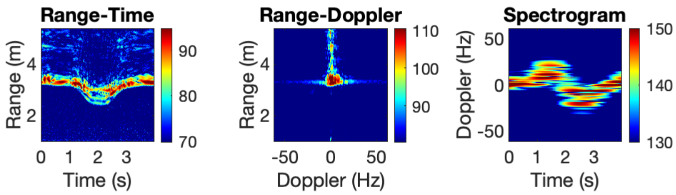
Data from radar with a single transmitter and receiver can span across the 3 domains of range, Doppler and time. The range is a measure of how far the target is located from the radar, the Doppler represents the target's radial speed, and the time dimension represents the evolution of the signals over the sequence of consecutive radar pulses. The information from these 3 domains can be mapped into different complex-valued 2D matrices, i.e., images to be used as inputs to neural networks for classification. A few examples are shown in Figure 6.

1) Range-time maps: In a pulsed UWB radar, each received radar echo is digitized to generate fast-time samples representing the distance of targets. Repeating this for multiple pulses over the slow-time at every PRI generates the range-time (RT) map s(d,t), essentially a matrix where d is the distance and t is the (slow) time.

2) Range-Doppler maps: Range-Doppler (RD) maps are generated by performing a Fast Fourier Transform (FFT) across the slow-time dimension of the RT maps. This allows to simultaneously characterize the distance and the velocity of human movements via the Doppler effect.



(a) Label 1: walking



(b) Label 6: bending from standing

Fig. 6: Examples of range–time, range–Doppler, and spectrogram patterns for two human activities (Label 1 and 6)

3) Micro-Doppler spectrograms: Micro-Doppler spectrograms, essentially patterns of Doppler frequencies over time, describe how the Doppler frequency of the different body parts varies with time and reflect the unique patterns of specific movements and activities. The STFT is applied on the RT maps to obtain spectrograms. A variety of STFT window sizes and step-width between consecutive RT maps were tested. Finally, superior classification was achieved by Hann window, the STFT step-width of 4 samples (32.8ms), and a window of 244 samples (approximately 2s).

IV. Results and discussion

In the tests presented in this section, the classification accuracy and depth of real-valued CNN models are compared with their corresponding complex-valued CNN models. All models are trained using the Adam optimizer with dynamic learning rates from 1e-4 to 1e-6. Cross-entropy loss is used as a cost value or loss function to minimize. The training stage of each model lasts $40 \sim 60$ epochs with a batch size of 32, depending on the models. It should be noted that the learning rate and the number of training epochs are crucial hyper-parameters for performances. Hence they need to be tuned for every model.

A. Accuracy of various radar data formats

In this subsection, validation accuracy results are presented for the three aforementioned radar data formats and for the different proposed implementations of complex and real-valued networks. For complex networks, both DCN blocks and multi-channel architectures with real & imaginary parts and magnitude & phase are compared. For real-valued networks, only the magnitude of the data is considered. The validation accuracy is calculated on 20% of randomly selected samples from the dataset in a hold-out manner.

1) Range-time maps: The overall accuracy of the different networks on the validation set is shown in Table I. It can be seen that the best prediction accuracy is 92.6% for the case of ResNet & multi-channel network with real and imaginary parts. In this case, implementing CVNNs helped improve the accuracy only a bit (+1.5%) on the ResNet model, whereas no advantages or even significantly worse performances were seen for the deep and shallow network, respectively.

Notably, the classification accuracy obtained for this dataset with RT maps yielded the highest classification accuracy across all the data formats considered.

TABLE I: Accuracy on range-time maps

Models	Shallow ConvNet	Deep ConvNet	ResNet
Real-valued network	**78.4%**	**90.4%**	91.1%
Multi-channel: Abs&phase	60.3%	**90.4%**	91.5%
Multi-channel: Real&imaginary	67.4%	89.1%	**92.6%**
DCN (Deep complex network)	36.7%	89.3%	91.5%

2) Range-Doppler maps: Table II shows that the highest classification accuracy values for this radar data format are 63.5%, 76.6%, and 75.5%, respectively for the complex-valued shallow ConvNet, the complex-valued deep ConvNet, and the complex-valued ResNet. Compared with the corresponding real-valued CNN model, the prediction accuracy of the complex-valued model can improve by between +8% and +10%. This is an interesting result compared with what is seen with the range-time maps, where the usage of complex information appeared to improve the performance only a little compared to real-valued networks.

TABLE II: Accuracy on range-Doppler maps

Models	Shallow ConvNet	Deep ConvNet	ResNet
Real-valued network)	55.3%	66.2%	65.9%
Multi-channel: Abs&phase	60.8%	66.0%	65.8%
Multi-channel: Real&imaginary	**63.5%**	**76.6%**	**75.5%**
DCN (Deep complex network)	41.4%	76.4%	74.6%

3) Spectrograms: The classification results for spectrograms are shown in Table III. For deep ConvNet and ResNet models, the accuracy of DCN is slightly higher (86.6% and 87.5%) than for their real-valued CNN (86.2% and 87.0%). From these results, it would appear that there is no significant benefit in using complex-valued networks on spectrograms, suggesting that the phase of the spectrogram has little information for HAR, at least in this case.

TABLE III: Accuracy on spectrograms

Models	Shallow ConvNet	Deep ConvNet	ResNet
Real-valued network	**76.0%**	86.2%	87.0%
Multi-channel: Abs&phase	74.6%	84.7%	86.5%
Multi-channel: Real&imaginary	74.2%	84.6%	85.6%
DCN (Deep complex network)	44.5%	**86.6%**	**87.5%**

B. Leave one person out

Every time the dataset is split randomly in a holdout evaluation, the test accuracy may differ depending on the split. A more robust methodology for testing classification approaches in human activities is that the sequences of 15 participants are split into test and training sets by excluding one participant from the training data and keeping that for testing. The procedure is well-known as *leave one person out* (L1PO). This also enables to evaluate performances for "unseen" participants in a perspective of generalisation of capabilities. After 15 tests, one per participant, the average accuracy (mean) and standard deviation (std) is calculated. In this subsection, L1PO results are reported.

1) Case 1: Range-time maps + real-valued network: The first case refers to the range-time domain with real-valued networks, which offered good results and low complexity in the previously discussed hold-out evaluation. The results are reported in Table IV. It is shown that these leave one person out results are lower than the hold-out results in the previous subsection (about 84-85% vs 90-91% obtained before). The depth of the network appears to help, with the "deep ConvNet" model achieving about +15% accuracy compared with its shallow counterpart as well as lower standard deviation. However, excessive model complexity, like in the case of ResNet, is shown not to produce significant improvements in this case.

TABLE IV: L1PO mean and std accuracy (train; test) on range-time maps and real-valued network

Models	**Shallow ConvNet**	**Deep ConvNet**	**ResNet**
Mean	82.70%; 69.56%	97.95%; 84.96%	100.0%; 85.39%
Std	8.98%; 8.18%	0.44%; 6.24%	0%; 6.50%

2) Case 2: Range-Doppler + DCN: The second case with range-Doppler maps and DCN implementations is discussed, where it was shown that the accuracy is improved with complex-valued architectures compared to the corresponding real-valued models. Table V shows that these L1PO accuracy values are still lower than the hold-out results, dropping around 5% in line with what seen in case 1. This is somewhat expected as L1PO validation is more challenging than the simpler hold-out evaluation, yet more realistic. Accuracy and model stability of the complex-valued shallow ConvNet using the DCN approach are low, while the performance of the deep ConvNet is dramatically improved, achieving about +45% accuracy, as well as some reduction of the standard deviation across participants, which is typically a challenge in L1PO implementations. ResNet suffers from the overfitting problem since this complex model with many convolutional layers does not boost performance in this case, compared to deep ConvNet (which has 5 stacked building blocks, so to some extent is shallow compared to ResNet).

TABLE V: L1PO mean and std accuracy (train; test) on range-Doppler maps and complex-valued DCN

Models	**Shallow ConvNet**	**Deep ConvNet**	**ResNet**
Mean	28.81%; 26.07%	92.18%; 70.62%	100.0%; 71.04%
Std	11.98%; 9.49%	1.16%; 6.30%	0%; 5.87%

C. Depth of the network

The effect of varying depths of the network is explored in this subsection. Each DCN and its real-valued counterpart are trained independently with seven different architectures obtained by stacking one to seven basic building blocks shown in section II. All three data formats are considered, with the RT maps and spectrograms providing the highest accuracy.

The results are displayed in Figure 7. When n = 1, the DCNs perform poorly, and the performance of real-valued networks is much better, which appears to suggest that complex-valued representations make little sense for a very shallow network. From n = 2, the gap between DCN and real-valued CNN becomes narrow, and DCN catches up or even outperforms the corresponding real-valued CNN. The highest accuracy of RT maps, range-Doppler maps, and spectrogram datasets is 90.4%, 76.4%, 86.6% respectively, with five building blocks.

It is noted that depth of more than five building blocks does not provide classification advantages. Also, a neural network with five blocks is not truly "deep" as deep CNNs generally have hundreds of layers in the image-processing domain [18]. However, with the relatively small radar datasets, shallow networks like those explored in this paper can still be suitable or even preferable in some cases.

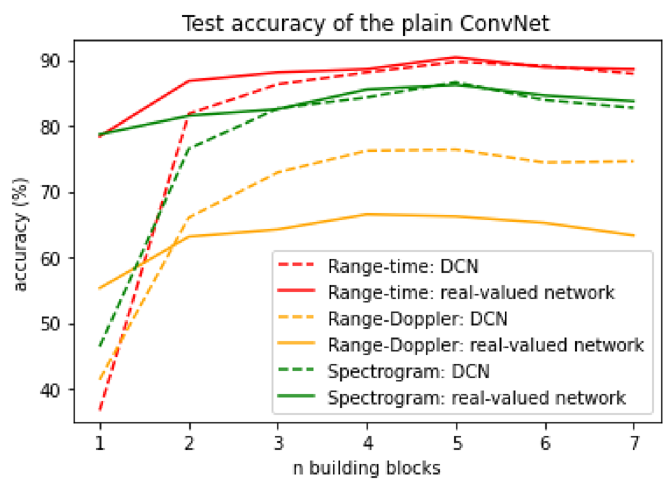


Fig. 7: Test accuracy of the ConvNet with different depths and with the three radar data formats

D. Discussion on the results

A summary of the main points from the aforementioned results is provided:

- Comparing the best results across all network architectures for the three radar data formats, RT maps and spectrograms perform best, with 92.6% and 87.5% accuracy respectively, while the performance of RD maps is lower with only 76.6% accuracy. The high accuracy obtained for the RT format (without explicit Doppler information) is notable, and it is hypothesized that the very fine discretization in range enabled by UWB radar was beneficial in this regard.
- Complex-valued networks appear to improve results for the RD maps (around +10% accuracy), but not much in

the other two formats. In general, the benefit of complex-valued networks in terms of accuracy compared to real-valued networks is visible in Deep ConvNet and ResNet models, especially for the DCN implementation and multi-channel with real and imaginary parts, compared to the multi-channel with abs and phase. For the shallow network, such a benefit is not visible.

- The analysis on the networks' depth appears to show that very deep networks are not suitable for this case study. This is likely due to overfitting problems because of the relatively small size of the radar dataset, also accounting for the larger number of parameters that complex-valued networks have, compared to their real-valued counterparts.
- Leave one person out validation proves that deeper networks improve the prediction accuracy for this more challenging validation compared to simple hold-out. However, when the networks become too deep, it is important to avoid overfitting problems caused by the small amount of available radar data.

V. Conclusions

This paper presents initial results of complex-valued neural networks implemented for radar-based HAR, with tests performed on a dataset measured with an UWB radar and containing 15 participants performing 9 activities. Specifically, two implementations of complex-valued networks are proposed: one separating real/imaginary parts or magnitude/phase into two separate channels, and one with purposely-coded complex-valued building blocks for convolutional networks.

The results show that complex-valued implementations yield an increase in classification accuracy only for certain data formats and certain architectures. Specifically, the usage of complex-valued networks showed improvement compared with real-valued networks in the case of range-Doppler input formats, but the overall best results were obtained for range-time maps as the radar input data - for which, interestingly, the complex-valued implementations did not appear to provide a noticeable improvement in accuracy. L1PO validation was also performed. With this validation approach, each participant is tested as the unseen test subject for which the classifier was not previously trained, in order to better test generalisation capabilities. While L1PO results show about $5\% \sim 15\%$ decrease in accuracy compared to the simpler hold-out validation, the results are still promising considering the relatively large number of participants and the continuous and unconstrained nature of the recorded activities. The effect of varying the depth of the networks was also explored by stacking different building blocks to go from a shallow to a deeper network with 7 stacked blocks. It is shown that there is no performance improvement, at least for these data, in increasing the depth of the network beyond 5 stacked building blocks.

Future work will explore complex-data processing in conjunction with data augmentation techniques that can potentially improve the training of the neural networks.

References

[1] S. Z. Gurbuz and M. G. Amin, "Radar-based human-motion recognition with deep learning: Promising applications for indoor monitoring," *IEEE Signal Processing Magazine*, vol. 36, no. 4, pp. 16–28, 2019.

[2] J. Le Kernec, F. Fioranelli, C. Ding, H. Zhao, L. Sun, H. Hong, J. Lorandel, and O. Romain, "Radar signal processing for sensing in assisted living: The challenges associated with real-time implementation of emerging algorithms," *IEEE Signal Processing Magazine*, vol. 36, no. 4, pp. 29–41, 2019.

[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, 2009.

[4] A. Shrestha, H. Li, J. Le Kernec, and F. Fioranelli, "Continuous human activity classification from fmcw radar with bi-lstm networks," *IEEE Sensors Journal*, vol. 20, no. 22, pp. 13 607–13 619, 2020.

[5] M. S. Seyfioglu, B. Erol, S. Z. Gurbuz, and M. G. Amin, "Dnn transfer learning from diversified micro-doppler for motion classification," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 55, no. 5, pp. 2164–2180, 2019.

[6] X. Li, Y. He, and X. Jing, "A survey of deep learning-based human activity recognition in radar," *Remote Sensing*, vol. 11, no. 9, 2019. [Online]. Available: https://www.mdpi.com/2072-4292/11/9/1068

[7] B. Erol and M. G. Amin, "Radar data cube processing for human activity recognition using multisubspace learning," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 55, no. 6, pp. 3617–3628, 2019.

[8] Y. Kim, I. Alnujaim, and D. Oh, "Human activity classification based on point clouds measured by millimeter wave mimo radar with deep recurrent neural networks," *IEEE Sensors Journal*, vol. 21, no. 12, pp. 13 522–13 529, 2021.

[9] R. G. Guendel, F. Fioranelli, and A. Yarovoy, "Phase-based classification for arm gesture and gross-motor activities using histogram of oriented gradients," *IEEE Sensors Journal*, vol. 21, no. 6, pp. 7918–7927, 2021.

[10] J. Guo, C. Shu, Y. Zhou, K. Wang, F. Fioranelli, O. Romain, and J. Le Kernec, "Complex field-based fusion network for human activities classification with radar," in *IET International Radar Conference (IET IRC 2020)*, vol. 2020, 2020, pp. 68–73.

[11] X. Wang, P. Chen, H. Xie, and G. Cui, "Through-wall human activity classification using complex-valued convolutional neural network," in *2021 IEEE Radar Conference (RadarConf21)*, 2021, pp. 1–4.

[12] T. Scarnati and B. Lewis, "Complex-valued neural networks for synthetic aperture radar image classification," in *2021 IEEE Radar Conference (RadarConf21)*, 2021, pp. 1–6.

[13] H. Mu, Y. Zhang, C. Ding, Y. Jiang, M. H. Er, and A. C. Kot, "Deepimaging: A ground moving target imaging based on cnn for sar-gmti system," *IEEE Geoscience and Remote Sensing Letters*, vol. 18, no. 1, pp. 117–121, 2021.

[14] A. H. Oveis, E. Giusti, S. Ghio, and M. Martorella, "Cnn for radial velocity and range components estimation of ground moving targets in sar," in *2021 IEEE Radar Conference (RadarConf21)*, 2021, pp. 1–6.

[15] D. A. Brooks, O. Schwander, F. Barbaresco, J.-Y. Schneider, and M. Cord, "Complex-valued neural networks for fully-temporal micro-doppler classification," in *2019 20th International Radar Symposium (IRS)*, 2019, pp. 1–10.

[16] C. Trabelsi, O. Bilaniuk, Y. Zhang, D. Serdyuk, S. Subramanian, J. F. Santos, S. Mehri, N. Rostamzadeh, Y. Bengio, and C. J. Pal, "Deep complex networks," 2018.

[17] X. Yang, *TU Delft MSc Thesis: Complex-Valued Neural Networks for Radar-based Human-Motion Classification*, 2021. [Online]. Available: http://radar.tudelft.nl/Education/bio.php?id=1243

[18] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *arXiv 1512.03385*, 2015.

[19] R. G. Guendel, M. Unterhorst, F. Fioranelli, and A. Yarovoy, "Dataset of continuous human activities performed in arbitrary directions collected with a distributed radar network of five nodes," Nov 2021. [Online]. Available: https://doi.org/10.4121/16691500.v2