IEEE
Sensors Letters
IEEE Sensors Council
Medical, robotics and automation applications

Low-power radar-based system for real-time object recognition

Anna Coletti[1], Alessio Sanna[2], Christian Cipriani[1*] and Enzo Mastinu[1**]

[1] *Artificial Hands Area, BioRobotics Institute, Scuola Superiore Sant'Anna, Pisa, Italy*
[2] *Politecnico di Torino, Turin, Italy*
* *Senior Member, IEEE*
** *Member, IEEE*

Received in

Abstract — In recent years, radar technology has attracted a new wave of interest due to unprecedented low-power potential and its inherent low privacy concerns compared to camera systems. In particular, a radar-based object and material recognition encloses a great potential for assistive technology in healthcare scenarios, such as prosthetic hands. In this study, we aimed to explore such potential via offline and online recognition achieved by deep learning techniques. Twenty different targets were explored, including objects from daily life activity as well as biological tissue (i.e., human hand). Feasibility was confirmed by the offline and online recognition accuracies (achieving 94% and 89% correct classifications in the best case scenario respectively), and promising insights are offered in regard to the number of radars needed for such task. Remarkably, for the first time it was shown that a radar-based, real-time correct differentiation between human tissue and inanimate objects. We believe these results pave the way for an easy-to-integrate solution with wide potential benefit in industrial automation technology and novel innovative human-machine interfaces, particularly in the prosthetic field context.

Index Terms — Radar system, object recognition, Deep learning, material recognition, prosthetic hand.

I. INTRODUCTION

In recent years, radar technology has attracted a new wave of interest due to unprecedented low-power potential and its inherent limited privacy concerns compared to camera systems. Such interest is confirmed by the vast amount of radar-based research recently published, particularly for human-centered applications such as human gesture/activity recognition, vital sign monitoring, and collision avoidance in automotive applications [1], [2], [3], [4], [5], [6], [7], [8]. The use of radars for the recognition of objects and materials is also gaining interests as it opens for a wide variety of applications. Automatic radar-based object sorting was proved possible for inspection processes in industrial scenario, to detect unwanted materials in a product while preserving the sealed packages [9]. Moreover, objects, body parts and transparent materials recognition was demonstrated possible for various applications, involving human-machine interactions in wearable systems or in industrial settings, both with known and unknown target material [10]. Here, Yeo et Al. used the radar prototype proprietary system Soli by Google ATAP [11], designed for smartphone embedded human-machine interactions, particularly for miniature and precise finger gestures aimed to navigate and control graphical and virtual interfaces. Recognition accuracies as high as 97% were shown via a random forest classification algorithm.

Object and material recognition encloses a great potential also for assistive technology in healthcare scenarios, such as prosthetic hands, the ultimate focus of our research group. Indeed, an understanding of the environment via exteroceptive sensors is imperative to enable more autonomous robotic hands [12], [13], [14]. The recognition of the target object material while holding it during a grasp task could provide important information for implementing automatic grip force modulation. In this context, radars are particularly appealing for their easiness of integration in a wearable system (low-power and low computational demands), peculiar measurement features (sensitive to different materials and density, invariance to illumination conditions) and the limited privacy risks. The feasibility of using easily embedded low-power radars to recognize object shape and material in order to select the correct grasp on a robotic hand during the dynamic reaching phase was confirmed by our recent study [15] on a publicly released dataset [16]. Outcomes proved promising, albeit achieved with objects exhibiting only two types of materials and five different shapes. In this study, we aimed to further assess such radar-based object recognition capabilities with a larger target set, inspired by daily life activity and varying between different materials and shapes. Our object recognition system showed remarkable accuracies for both offline and online assessments achieving 94% and 89% correct classifications in the best-case scenario respectively, confirming once more the potential of radar technology for prosthetic applications.

II. MATERIALS AND METHODS

A. Experimental setup

We aimed to investigate the use of low-power, millimeter-accurate radars for the recognition of different objects. In particular, our choice was set on the commercial pulse-coherent radar A121 (Acconeer, Sweden), which combines <0.1W power consumption, 2.5mm precision, large field-of-view of 65×53 degrees, with integrated TX/RX antennas in a 29mm^2 compact package and cost ≈10€.

This work involved human subjects. The ethical approval was provided by Ethical Committee of the Scuola Superiore Sant'Anna (ref. 12/2022).

Corresponding author: Anna Coletti (anna.coletti@santannapisa.it)
Digital Object Identifier:

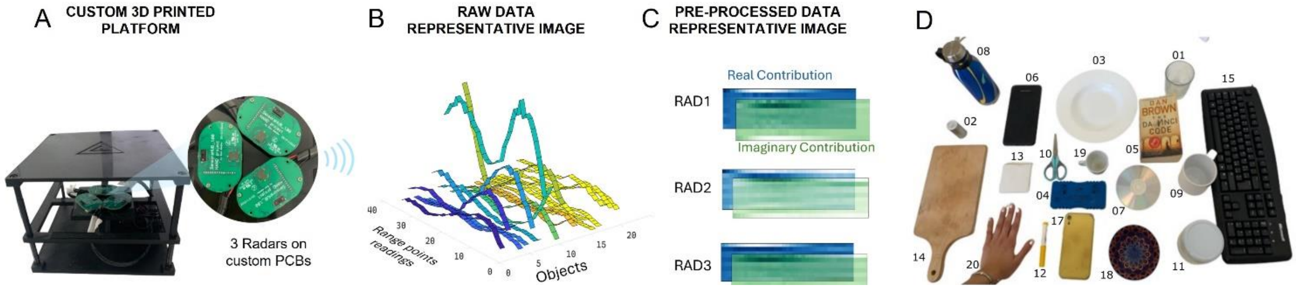

Figure 1: Representative images of the experimental setup, acquired data and processing. (A) Custom 3D printed platform with the three radars accommodated in the middle level. (B) Representative visualization of a single sweep intensity (absolute value of each range point) recorded from the twenty target objects. (C) Representative images of the water bottle data prepared for the CNN. (D) Twenty target objects evaluated in this study.

Acconeer's radar is already widely used in innovative use cases such as gesture control, obstacle detection, surface classification, parking sensors, human presence detection, alarms and access control. Here, an instrumented device was developed based on the A121 (Fig. 1 A), consisting of a custom 3D printed three-layer platform, three radars and an acquisition microcontroller (ESP32, EspressIf, China). The middle layer held the three radars and their hosting custom-made PCBs. Here, the radars were configured to emit trains of 60GHz pulses in pulsing/silent cycles alternating at 13MHz. The radar received pulses at times related to the distances of 35 range points, equally distributed from 87.5mm to 175.0mm (i.e., 35 range point, where each range point was separated by 2.5mm). Each range point measure was represented as a complex number (sparse IQ service which utilizes the phase-coherency of the A121 pulsed radar to produce stable In-phase and Quadrature components) enclosing information about amplitude and phase of the reflected pulse (Fig. 1 B). Each radar provided a complete measure (i.e., a sweep) of its range points with 15Hz frequency. Data was then transmitted via USB to a custom Matlab-based PC application.

B. Data collection protocol

The radar received intensity is influenced by the reflection and thus by the inherent dielectric properties of the target material. In order to investigate the performance of our instrumented device on a wide range of materials, 19 daily objects were selected for this experiment (Table 1), trying to maximize the variety in their dielectric constant (i.e., permittivity constant). Additionally, due to the explorative nature of this study, we decided to include also the human hand within the list of investigated targets (Fig. 1 D). All collected data is publicly accessible on Figshare [17]. The three radars were placed 65mm away from the top layer, thus from the target objects. This was chosen to avoid any artifact or distortion on the radar signals due to too short-distance from the target. Moreover, the three radars were placed in equilateral triangular configuration with ≈20mm distance between each other, although through experimental investigations it was seen that the data were not significantly influenced by different intra-radar distance, for example with respect to rotating the PCBs by 180°. This positioning was deemed as a good compromise between a small intra-radar distance and a large aggregated field-of-view, so that the objects considered were within the space for which the measurements were acquired. Data were acquired by positioning the target objects over the instrumented device, at the center of the 3D printed top layer platform, in three trials of 10s each. For each trial, the target object was repositioned in the same position with small random perturbation in orientation. This procedure was then repeated for four identical sessions. Different orientations and multiple sessions were deemed necessary to secure enough data and variability for the following deep learning part. Lastly, to avoid potential environmental influences, all data acquisition sessions were conducted within the same day, to ensure some consistency in room conditions (temperature and humidity).

Table 1: The table shows the 19 daily objects and the natural hand listed with their permittivity constants.

	Objects	Permittivity coefficients
01	Glass empty	3.7-10
02	Aluminium	3.5-5.5
03	Plate (ceramic)	4.5-7
04	Eraser (polyethylene)	2.25
05	Book	1.4
06	Power bank	3.5-5.5
07	CD	2.5
08	Water bottle (aluminium)	3.5-5.5
09	Cup (ceramic)	4.5-7
10	Scissors (steel)	2-7
11	Jar (glass)	3.7-10
12	Marker (plastic)	4
13	Wood	2.25
14	Tile	3
15	Keyboard (plastic)	4
16	Air (No object condition)	1
17	Phone case (plastic)	4
18	Bowl (ceramic)	4.5-7
19	Coffee cup (ceramic)	4.5-7
20	Human Hand (body tissue)	8

C. Deep learning models

The convolutional neural network (CNN) was used for the object recognition evaluation. Thus, similar analysis was performed in three data conditions, using respectively the data acquired from one, two and all the three radars (1RAD, 2RAD or 3RAD). Each model consisted of three convolutional input blocks and five linear output layers. Then, depending on the data condition, the CNN included one, two or three parallel input branches. Each branch consisted of three convolutional blocks including four layers each (convolution 3x3 kernel + max-pooling + batch-normalization + leaky-ReLU activation). The linear output layers consisted of five layers including batch-normalization and leaky-ReLU activation function. For 2RAD and 3RAD data conditions, the outputs of the last convolutional blocks were concatenated before passing through the linear output layers. The CNNs were implemented with TensorFlow framework, trained and tested via GPU (T4, Nvidia, USA) via the Adam stochastic optimizer for 200 epochs using a batch size of 8. The models' architectures and hyperparameters were found via preliminary testing. Before being fed to the models, the radar readings were pre-processed. Specifically, data was divided in frames of 5 sweeps, then FFT was applied to each frame along the temporal axis (i.e., doppler maps) and lastly, images were created from both real and imaginary parts of the data (Fig. 1 C). Recognition performances were evaluated with both offline and online tests. We evaluated the offline recognition performance of the proposed radar-based system by using a session-based cross-validation. Thus, three sessions of data were used for training and the last session for testing. This leave-one session-out process was then repeated for all the combinations of sessions and data conditions (1RAD vs 2RAD vs 3RAD). The most performing CNNs for each data conditions and for each session-based iteration were then exported and used for online evaluation on newly collected, unseen data. For the online tests, a custom Matlab-based PC application was developed enabling real-time data acquisition and consequent evaluation of the pre-trained exported CNNs. New data was collected for each object for 10 seconds, and the online-predicted labels issued every 5 sweeps were stored to file for accuracy calculations every ~0.3s. Additionally, we performed a power analysis to explore effects of training set size, repeating the offline analysis using only one or two training sessions.

III. RESULTS

A. Offline object recognition

The proposed system, using all the dataset, shown an average accuracy of 90.9% across the 4 testing sessions in the first case scenario (1RAD), which increased up to 91.1% of average accuracy for the 2RAD condition and finally reached 94% of average accuracy if the readings of all three radars are processed (3RAD). The system showed a considerable decrease in the median accuracy when the training set was reduced to only one or two sessions (Fig 2). Indeed, when using a single training session, the offline accuracies were 63%, 67% and 68.5%, respectively in 1RAD, 2RAD and 3RAD conditions, which then increased to 79%, 81% and 86% when using two training sessions.

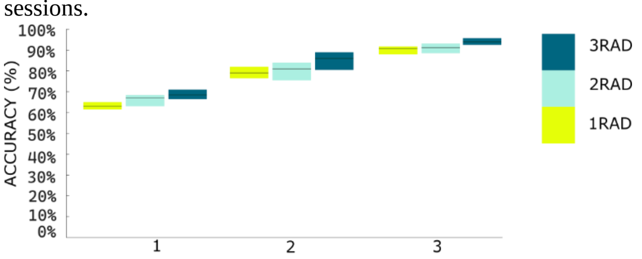

Figure 2: Offline test accuracies (median:iqr) for the three different scenarios (1RAD,2RAD,3RAD), using different training sets.

B. Online object recognition

The most performing CNNs, which they achieved respectively 91.9%, 94.2% and 96.7% testing accuracy during offline evaluation for 1RAD, 2RAD and 3RAD condition, were exported for online test. The labels predicted using these models during the real time test-session for each object were used to calculate the accuracy of the system for online classifications. Online object recognition accuracies in the three data conditions reached respectively 66.0%, 70.0% and 89.2%. Results confirmed the feasibility of the radar-based object recognition, however three radars appeared to be necessary to keep consistency between offline and online accuracy. Indeed, accuracy dropped considerably for 1RAD and 2RAD data conditions. Most of the misclassifications (>85%) regarded the eraser vs no object vs the marker (Fig. 3), pointing to a potential limitation of the system to differentiate low density materials. Interestingly, the system showed little difficulties in recognizing the human hand (83.3% of right classification, F1-scores: 0.91), opening to promising applications with organic material.

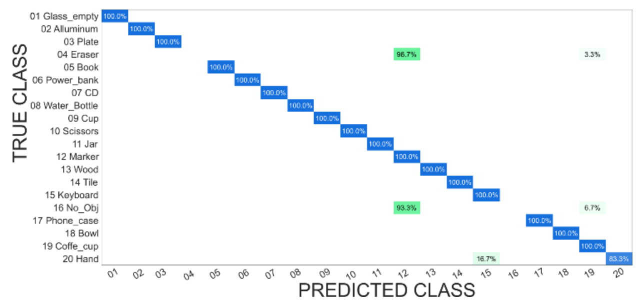

Figure 3: Confusion matrix for the online object recognition test using data from the three radars.

IV. DISCUSSION AND CONCLUSION

In this study, we explored the potential of using low-power radars for object and material recognition. Here, twenty different targets were explored, including objects from daily environment as well as the human hand. The proposed system proved promising recognition of all objects and for all data conditions, clearly showing the feasibility to differentiate also objects made of the same material (e.g., small aluminium parallelepiped vs water bottle). Arguably, the overall lower online accuracy can be attributed to the limited amount of training data, and to the real-time data normalisation which was entirely based on the training data sessions. A more extensive data collection, combined with a real-time calibration session could resolve or reduce this accuracy mismatch. Similarly, more training data could also resolve the destructive performance drop seen in online object recognition with a single radar. The power analysis also generally suggests a larger training set for a more reliable classification, for all the case scenarios. Nevertheless, the offline recognition results from the ablation study seem to point towards the feasibility of using even a single radar for such task. However, more testing are imperative to shine some light on the contradictory online results for the single radar condition. The data from such simple radars appeared highly sensitive to different materials as well as to their distribution and density in space. Indeed, it was possible to differentiate different objects despite them being made of the same material (e.g., small aluminium parallelepiped vs the water bottle). Remarkably, for the first time it was shown a correct differentiation between human tissue and inanimate objects. Interestingly, the human hand was correctly classified in 100% and 83% of the cases for offline and online testing, respectively. This result paves the way for an easy-to-integrate solution with wide potential benefit in human-robot interactions, particularly for those context where human safety is a major concern such as industrial automation and assistive robotics.

Our results found close match with the ones presented by Yeo et Al. [10]. Our set of target objects was heavily inspired by this prior study and a strong coherence was found for the correct recognition of most of the shared targets (i.e., air, aluminum, glass, plate, keyboard, wood, book, phone case and CD). Arguably, results can be considered comparable even though our study was based on a radar sensor much simpler than the prototype developed by Google ATAP. Indeed, the A121 radar used here presents several advantages such as it is commercially available, highly customizable, low-power and with single transmitting and receiving antennas. These advantages can provide an easier yet robust starting point for researchers and developers aiming to explore radars for novel human-machine interfaces. Several study limitations can be discussed. At first, the set of target objects was limited and cannot possibly represent the vast variety of objects that can be encountered during daily life activities. Our intention was merely to provide further evidence of the underlying potential of using simple low-power radar sensors for object differentiation, rather than making any statement on the absolute recognition capability of every single object. Secondly, upon closer data examination some dependence was found between the radars measurements and the environmental conditions (room humidity and temperature). While most of this dependence can be easily explained by the different water concentration in the air, it could be considered as another possible reason for the lower online recognition performances. However, future investigations are needed to formalize such dependence and thus elaborate mitigation strategies to overcome this limitation, for example adding an initial calibration session to increase the generalisation capability of the system. Thirdly, the machine learning part was not thoroughly explored here. The decision of images processing and CNNs was directly inherited from the positive results of our recent study [15], thus only a limited network optimisation was considered here. There are certainly many aspects that should be further investigated, e.g. exploring classical machine learning algorithms and manual features extraction, as well as improving models performance via explainable AI methods. Lastly, even though the presented results responded validly to the underlying technological challenge of integrating such radar-based solution in a wearable device (i.e., a prosthetic hand), aspects related to low-power optimizations of hardware and software and their embedded validation were not included here. These goals remain of interest for our group and will be explored in future work. We believe the presented results provide interesting insights for the advancement of human-robot interfaces and robots automation. The compact size and low power requirements of modern radars sensors make them friendly to integration into wearable systems. Moreover, the nature of the data processed ensures compliance with regulations concerning privacy towards third parties in possible everyday applications. Despite the experimental and preliminary nature of this study, it represents a necessary building block of the technological advancement towards more intelligent assistive technology such as prosthetic hands, the ultimate focus of our research group.

ACKNOWLEDGMENT

The authors would like to express their gratitude to Acconeer AB for supporting and freely sharing samples of the most recent radar A121, as well as to Robinson Gabriel Guachi Guachi for supporting the design and 3D printing of the enclosure.

This work was supported by European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 101029946, and the Italian National PhD program in Artificial Intelligence, XXXVIII cycle, course on Health and Life Sciences, organized by Università Campus Bio-Medico di Roma.

REFERENCES

[1] Y. Fu, C. Li, F. R. Yu, T. H. Luan, and Y. Zhang, "A Survey of Driving Safety With Sensing, Vehicular Communications, and Artificial Intelligence-Based Collision Avoidance," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 6142–6163, Jul. 2022, doi: 10.1109/TITS.2021.3083927.

[2] A. Venon, Y. Dupuis, P. Vasseur, and P. Merriaux, "Millimeter Wave FMCW RADARs for Perception, Recognition and Localization in Automotive Applications: A Survey," *IEEE Transactions on Intelligent Vehicles*, vol. 7, no. 3, pp. 533–555, Sep. 2022, doi: 10.1109/TIV.2022.3167733.

[3] M. Giordano, G. Islamoglu, V. Potocnik, C. Vogt, and M. Magno, "Survey, Analysis and Comparison of Radar Technologies for Embedded Vital Sign Monitoring," in *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, IEEE, Jul. 2022, pp. 854–860. doi: 10.1109/EMBC48229.2022.9871847.

[4] M. Scherer, M. Magno, J. Erb, P. Mayer, M. Eggimann, and L. Benini, "TinyRadarNN: Combining Spatial and Temporal Convolutional Neural Networks for Embedded Gesture Recognition With Short Range Radars," *IEEE Internet Things J*, vol. 8, no. 13, pp. 10336–10346, Jul. 2021, doi: 10.1109/JIOT.2021.3067382.

[5] M. Mercuri, I. R. Lorato, Y.-H. Liu, F. Wieringa, C. Van Hoof, and T. Torfs, "Vital-sign monitoring and spatial tracking of multiple people using a contactless radar-based sensor," *Nat Electron*, vol. 2, no. 6, pp. 252–262, Jun. 2019, doi: 10.1038/s41928-019-0258-6.

[6] J. Cho, H. R. Hussen, S. Yang, and J. Kim, "Radar-Based Road Surface Classification System for Personal Mobility Devices," *IEEE Sens J*, vol. 23, no. 14, pp. 16343–16350, Jul. 2023, doi: 10.1109/JSEN.2023.3279785.

[7] Z. Zhang, Z. Tian, and M. Zhou, "Latern: Dynamic Continuous Hand Gesture Recognition Using FMCW Radar Sensor," *IEEE Sens J*, vol. 18, no. 8, pp. 3278–3289, Apr. 2018, doi: 10.1109/JSEN.2018.2808688.

[8] S. Hazra and A. Santra, "Robust Gesture Recognition Using Millimetric-Wave Radar System," *IEEE Sens Lett*, vol. 2, no. 4, pp. 1–4, Dec. 2018, doi: 10.1109/LSENS.2018.2882642.

[9] J. Cho, S. Yang, and J. Kim, "PCR Radar-Based Counting System for Packaged Objects," in *2023 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*, IEEE, Feb. 2023, pp. 613–616. doi: 10.1109/ICAIIC57133.2023.10066968.

[10] H.-S. Yeo, G. Flamich, P. Schrempf, D. Harris-Birtill, and A. Quigley, "RadarCat," in *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, New York, NY, USA: ACM, Oct. 2016, pp. 833–841. doi: 10.1145/2984511.2984515.

[11] J. Lien *et al.*, "Soli," *ACM Trans Graph*, vol. 35, no. 4, pp. 1–19, Jul. 2016, doi: 10.1145/2897824.2925953.

[12] S. Došen, C. Cipriani, M. Kostić, M. Controzzi, M. C. Carrozza, and D. B. Popović, "Cognitive vision system for control of dexterous prosthetic hands: Experimental evaluation," *J Neuroeng Rehabil*, vol. 7, no. 1, p. 42, Dec. 2010, doi: 10.1186/1743-0003-7-42.

[13] P. Weiner, J. Starke, F. Hundhausen, J. Beil, and T. Asfour, "The KIT Prosthetic Hand: Design and Control," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, IEEE, Oct. 2018, pp. 3328–3334. doi: 10.1109/IROS.2018.8593851.

[14] S.-H. Heo and H.-S. Park, "Proximity Perception-Based Grasping Intelligence: Toward the Seamless Control of a Dexterous Prosthetic Hand," *IEEE/ASME Transactions on Mechatronics*, pp. 1–12, 2024, doi: 10.1109/TMECH.2023.3324051.

[15] E. Mastinu, A. Coletti, J. van den Berg, and C. Cipriani, "Explorations of Autonomous Prosthetic Grasping Via Proximity Vision and Deep Learning," *IEEE Trans Med Robot Bionics*, pp. 1–1, 2024, doi: 10.1109/TMRB.2024.3377530.

[16] E. Mastinu, A. Coletti, S. H. A. Mohammad, J. van den Berg, and C. Cipriani, "HANDdata – first-person dataset including proximity and kinematics measurements from reach-to-grasp actions," *Sci Data*, vol. 10, no. 1, p. 405, Jun. 2023, doi: 10.1038/s41597-023-02313-w.

[17] A. Coletti, A. Sanna, C. Cipriani, and E. Mastinu, "10.6084/m9.figshare.25680120."