Dual-Timescale Classification of Human Activities Using Radar Point Clouds

Nicolas C. Kruse[1], Alec Daalman[2], Francesco Fioranelli[3], Alexander Yarovoy[4]
MS3 Group, EEMCS Faculty, Delft University of Technology, the Netherlands
{[1]n.c.kruse, [2]a.daalman [3]f.fioranelli, [4]a.yarovoy}@tudelft.nl

***Abstract*— The problem of radar-based, continuous Human Activity Recognition (HAR) has been studied in this work. A fixed-window segmentation method based on dual timescales has been proposed to tackle this challenge. The method is experimentally validated on a challenging publicly available dataset with 14 participants and 9 activities, and is compared to reference works from the literature. L1PO validation of the method yields a test accuracy and macro F1-score of $87.5\,\%$ and $80.1\,\%$ respectively.**

***Keywords*— Human activity recognition, machine learning, radar, point cloud processing.**

I. Introduction

Radar sensors are promising sensing modalities for a variety of healthcare applications, ranging from fall detection [1], [2], to gait analysis [3]–[7], to vital sign monitoring [8]–[10]. Radar sensors work contactless and preserve the privacy of the end user, as no visual imagery is captured that can be used to directly identify an individual. Furthermore, as an active sensor, radar does not rely on any external source of light or radiation, and functions in complete darkness or in glaring light conditions.

Classification of Activities of Daily Life (ADL) can be employed to monitor the wellbeing of vulnerable individuals. For example, the identification of wandering among patients with dementia [11], or the aforementioned fall detection. For continuous monitoring of individuals, classification approaches must account for activities of unconstrained duration. This open challenge of continuous classification can be addressed by methods based on activity sequence segmentation, where sequences are divided into either regularly spaced intervals, or adaptively changing intervals. Regular intervals [12]–[14] are simpler to implement and computationally efficient, but offer no guarantee that any interval contains only a single activity. Adaptive segmentation [15]–[18] aims to create segments of varying duration that contain only a single activity. This approach minimizes the possibility of ambiguity in classification, but is computationally more involved and also cannot guarantee single-activity segments.

In this work, a novel classification method is proposed that features a fixed-window approach, operating at two distinct time-scales. This work is an extension of the fixed-window method in [14], and aims to minimize the ambiguity issues that arise with fixed-window methods by fusing information from two classification models operating in parallel. The method is evaluated using a challenging, publicly available experimental dataset of human activities collected at TU Delft.

The rest of this paper is organized as follows: the proposed method is described in Section II, with details of the experimental dataset in Section III. The evaluation results are discussed in Section IV, with conclusions in Section V.

II. Proposed Method

The method proposed in this work is inspired by an approach in earlier work [14]. In broad terms, radar data are processed into Point Cloud (PC) representations, in two distinct timescales. The PCs from both timescales are classified by individual Point Transformer (PT) networks. The data processing for a single timescale is shown in Figure 1. The utilized parameters are based on earlier work in [14], [19] and the steps are outlined as follows:

1) Raw data in complex-valued range-time format is segmented into fixed-duration intervals in step (1). This fixed duration represents the one of the two timescales.
2) The acquired segments are subdivided into 6 (N_{sub}) shorter subsegments, and a Fast Fourier Transform (FFT) along the slow-time dimension yields 6 time-ordered range-Doppler maps.
3) With a relative threshold of $80\,\%$, the 6 range-Doppler maps are binarized. The relative threshold is computed with respect to the highest signal amplitude in each individual range-Doppler map.
4) Two operations are performed to reduce noise and clutter. A range gate of $2\,\mathrm{m}$ is first applied around the center of mass of the binarized range-Doppler map. Subsequently, the three largest connected regions in the binary image are identified and preserved. All other non-zero points in the binary image are set to zero.
5) The remaining non-zero points are converted to a Point Cloud format, using their respective coordinates in range-Doppler space. A third coordinate, signal amplitude, is retrieved from the original (non-binary) range-Doppler map. Concatenating the 6 individual point clouds yields time as a fourth coordinate.
6) Finally, the resulting 4D point cloud is upsampled or downsampled based on the required point cloud size N_{pts}, which is maintained at 512 points for this work. Downsampling is performed by removing random points from the cloud, upsampling by duplicating random points in the cloud.

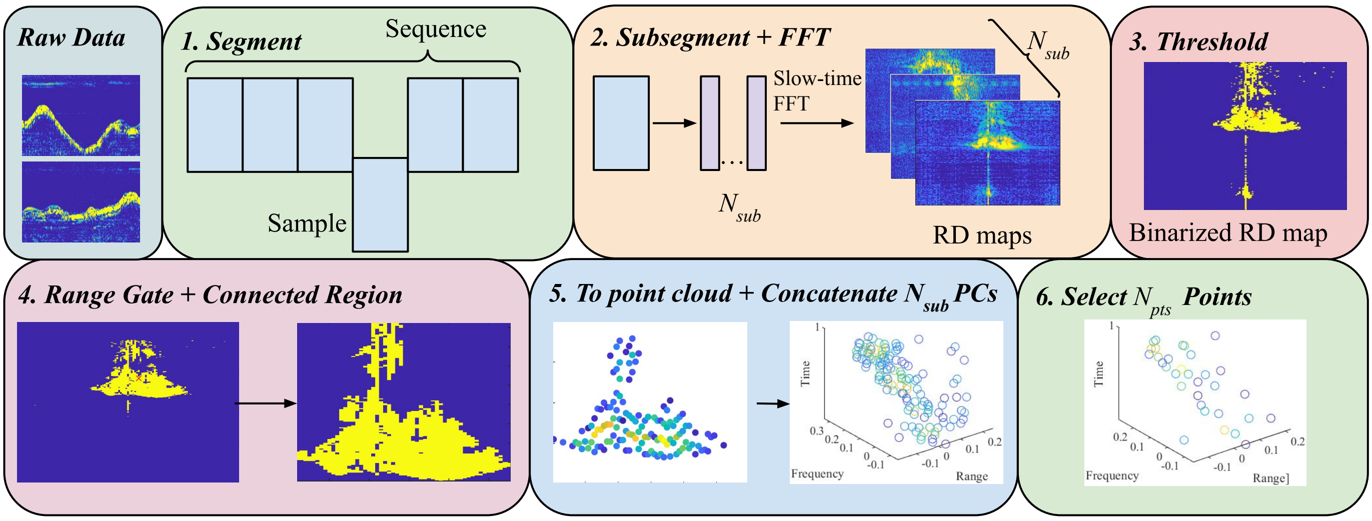

Fig. 1. Diagram of the segmentation processing pipeline for the generation of point cloud samples suitable as inputs for Point Transformer networks.

Each point cloud is labelled as an activity class, which is based on the most prevalent ground truth label within the duration of the sample. The point cloud samples for each timescale are used to train two individual PT networks [20]. PT networks fall in the family of Transformer neural networks, which utilize the so-called attention mechanism [21] to classify sequence-type data. In this PT implementation, the sequence is the unordered list of points in the cloud, with an added mechanism to account for local correlations between neighbouring points. The input to a PT model is an N_D-dimensional point cloud with N_{pts} points. In the case of this work, $N_D = 4$ and $N_{pts} = 512$. The output of the model is a logit vector of the activity classes $\vec{y}_c$ for each point cloud, which correlates to class probabilities.

To determine the predicted activity class at any given time, the corresponding segments at that time, and thus point clouds, for both the timescales is determined, and the logit vectors are summed. The class with the highest value after summation is output as the dual-timescale prediction, as:

$$\tilde{y} = \arg\max_{c} \sum_{N_C=1,2} \vec{y}_c^{N_C}. \qquad (1)$$

When multiple radar sensors are available, the logit vectors of samples originating from the different sensors can be summed in the same way in order to obtain a single prediction.

III. Experimental Setup

The proposed method is evaluated using a publicly available experimental dataset of ADL [22]. The dataset consists of sequences of activities performed by participants of varying age and gender, and is captured by a network of five simultaneously operating monostatic radars.

The radars are five Humatics PulsON P410 pulsed Ultra Wideband (UWB) sensors [23]. They operate at a center frequency of $4.3\,\mathrm{GHz}$, with a bandwidth of $2.2\,\mathrm{GHz}$. The resolution in range is thus approximately $6.8\,\mathrm{cm}$. The radars

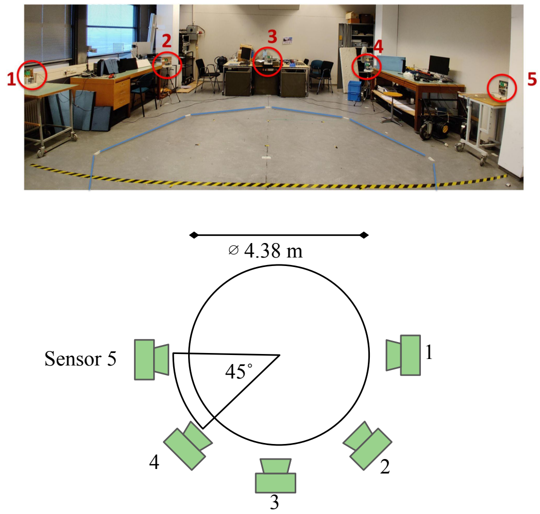

Fig. 2. Photograph (top) and diagram (bottom) of the experimental measurement area. Five simultaneously operating radar sensors are arranged in a semicircle, spaced at regular $45°$ intervals.

are operated at a Pulse Repetition Frequency (PRF) of $122\,\mathrm{Hz}$, which yields a maximum unambiguous velocity of $2.13\,\mathrm{m\,s^{-1}}$. The sensors are Single Input Single Output (SISO), and equipped with antennae that feature a pattern symmetric in azimuth. As such, no angle-of-arrival information can be obtained from an individual radar sensor.

The radar network geometry of the five radar sensors is shown in Figure 2. The nodes are arranged in a semicircle of diameter $6.38\,\mathrm{m}$, spaced at regular $45°$ intervals. The minimum and maximum ranges of the radar nodes are $1\,\mathrm{m}$ and $5.38\,\mathrm{m}$ respectively. As such, the observation area is approximately circular, with a diameter of $4.38\,\mathrm{m}$. The individual sensors operate simultaneously, they are unsynchronized and employ pseudo-random PRF variations to mitigate interference.

The collected data consist of continuous sequences of

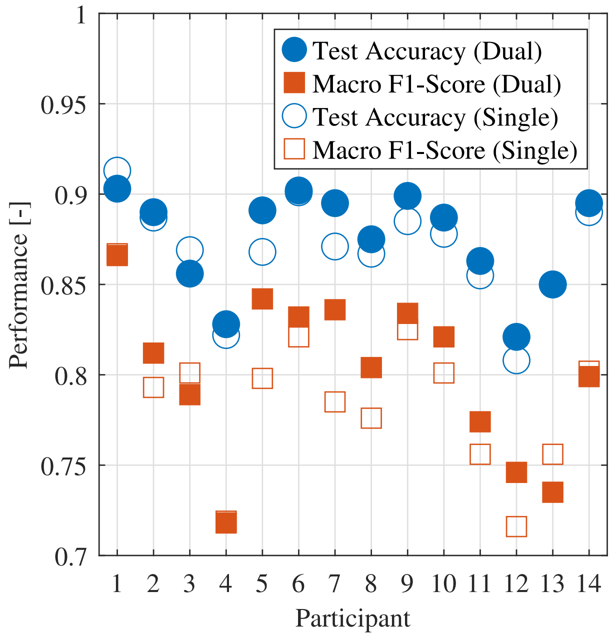

Fig. 3. Results for the proposed dual timescale classification method, as well as the single timescale reference method. Both are evaluated using a L1PO testing scheme. Test accuracy and Macro F1-score results are shown for each of the 14 participants. Solid markers represent the proposed method, whereas empty markers represent the single timescale reference.

human activities. Each sequence is 2 min in duration and features a variety of motions. All motions are performed at random locations in the measurement area, in random directions, and start and stop at random times. In total, 14 participants perform 30 sequences each, yielding 420 sequences in total. All sequences are composed of a total of nine distinct activities, and more details on the sequence types is available with the dataset documentation [22].

All data in the dataset is processed following the proposed method. The data of all but one participant is then used to train a PT network, which is tested with the data of the remaining participant. This Leave-One-Person-Out (L1PO) validation approach is repeated until the data of all participants has been used once for testing. This validation scheme emphasizes the ability of the model to generalize to unseen participants.

The selection of the two timescales is key to the correct implementation of the proposed method. For this work, the two timescales are set to the statistical mode and median of the ground truth activity segment lengths. They are 1.31 s and 2.30 s respectively.

For the PT networks, three parameters are relevant: the number of transformer blocks, the number of neighbours considered for each point, and the size of each transformer layer. They are set to 4, 16, and 128 respectively, based on the research in [14], [19]. The code for the PT model is adapted from [24].

Table 1. Comparison of the proposed dual timescale method with the single timescale alternative, as well as two classification methods from the literature using the same experimental dataset and a L1PO validation scheme. [25] features a hybrid CNN-BiGRU method for classification, [26] features a sensor fusion approach, paired with a CNN-BiLSTM architecture.

Reference	**Test Accuracy**	**Macro F1-Score**
Proposed Method	0.875	0.801
Single Timescale	0.869	0.787
Difference	+0.69 %	+1.49 %
CNN-BiGRU [25]	0.851	-
Sensor Fusion [26]	0.874	0.819

IV. Results

As mentioned before, results are presented in a L1PO validation format. Figure 3 shows two key classification metrics for both single and dual-timescale classification for each participant. Test accuracy is a measure of the correct predictions over the entire test dataset. Macro F1-score is a performance measure that takes class imbalance into account. For the single-timescale experiment, the amount of points per point cloud has been doubled to 1024, in order for a more equal comparison in terms of computational requirements, and sample information contents. It can be seen in Figure 3 that both the L1PO test accuracy and the macro-F1 score are improved in the majority of cases.

For the sake of comparison, the proposed dual-timescale method is compared to a single timescale reference using a single Point Transformer, as well as two reference methods from the literature that have been benchmarked on the same experimental dataset. The results are summarized in Table 1. It is noted that the addition of the secondary timescale brings the proposed method to the same test accuracy as the work in [26], which reports the highest L1PO test accuracy on this dataset to date. Additionally, an increase in macro F1-score of 1.21 % is demonstrated over the single timescale reference.

The confusion matrix in Figure 4 displays the most common error types over all participants. The two most prevalent error types are a confusion between the two types of falls, and a confusion between *Walking* and *Falling from Walking*. These errors can be explained by the ambiguous boundary between a walk and subsequent fall.

V. Conclusion

In this work, classification of continuous sequences of human activities is explored using a dual-timescale approach. It is demonstrated that the proposed method performs superior to a single-timescale equivalent method, with no increase in the size of a single sample. A L1PO test accuracy and Macro F1-score of respectively 87.4 % and 80.3 % are attained, making the proposed method competitive with reference classification methods from the literature.

In future work, studies on optimal timescales will be performed for the dual-timescale approach. Additionally, adding additional timescales is a possible extension of this method, and the balance between classification performance and computational requirements will be studied.

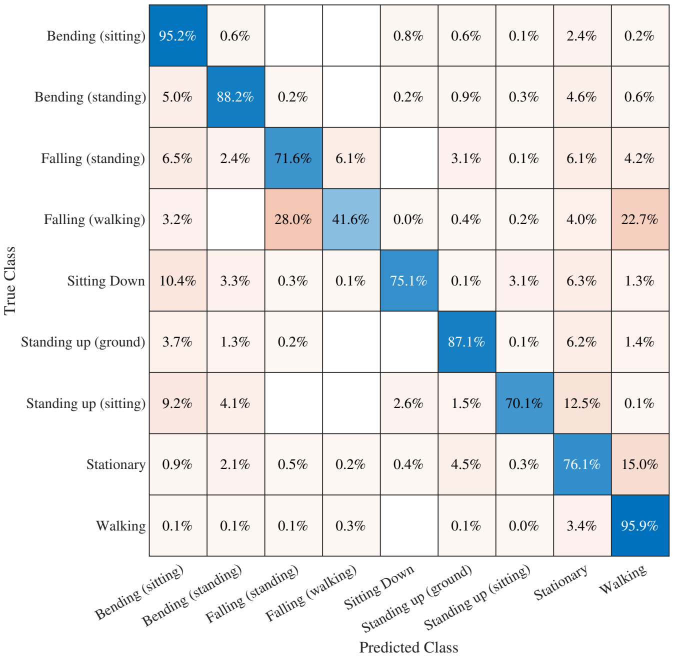

Fig. 4. Row-normalized confusion matrix for all samples processed for the L1PO validation scheme.

ACKNOWLEDGMENT

This research is funded by the Dutch Research Council (NWO) through the project *RAD-ART* (Radar-aware Activity Recognition with Innovative Temporal Networks).

REFERENCES

[1] S. A. Shah, A. Tahir, J. Le Kernec, A. Zoha, and F. Fioranelli, "Data portability for activities of daily living and fall detection in different environments using radar micro-doppler," *Neural Computing and Applications*, vol. 34, no. 10, pp. 7933–7953, 2022, ISSN: 0941-0643.

[2] Y. Yao, C. Liu, H. Zhang, *et al.*, "Fall Detection System Using Millimeter Wave Radar Based on Neural Network and Information Fusion," *IEEE Internet of Things Journal*, pp. 1–1, 2022, ISSN: 2327-4662.

[3] M. M. Rahman, D. Martelli, and S. Z. Gürbüz, "Gait Variability Analysis with Multi-Channel FMCW Radar for Fall Risk Assessment," in *2022 IEEE 12th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, IEEE, 2022, pp. 345–349, ISBN: 978-1-6654-0633-8.

[4] H. Ji, C. Hou, Y. Yang, F. Fioranelli, and Y. Lang, "A One-Class Classification Method for Human Gait Authentication Using Micro-Doppler Signatures," *IEEE Signal Processing Letters*, vol. 28, pp. 2182–2186, 2021, ISSN: 1070-9908.

[5] A.-K. Seifert, M. G. Amin, and A. M. Zoubir, "Toward Unobtrusive In-Home Gait Analysis Based on Radar Micro-Doppler Signatures," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 9, pp. 2629–2640, 2019, ISSN: 0018-9294. arXiv: 1809.06653.

[6] Z. Ni and B. Huang, "Open-Set Human Identification Based on Gait Radar Micro-Doppler Signatures," *IEEE Sensors Journal*, vol. 21, no. 6, pp. 8226–8233, 2021, ISSN: 1530-437X.

[7] H. Li, A. Mehul, J. Le Kernec, *et al.*, "Sequential Human Gait Classification with Distributed Radar Sensor Fusion," *IEEE Sensors Journal*, vol. 21, no. 6, pp. 7590–7603, 2021, ISSN: 15581748.

[8] T. Sakamoto, P. J. Aubry, S. Okumura, H. Taki, T. Sato, and A. G. Yarovoy, "Noncontact Measurement of the Instantaneous Heart Rate in a Multi-Person Scenario Using X-Band Array Radar and Adaptive Array Processing," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 8, no. 2, pp. 280–293, 2018, ISSN: 2156-3357.

[9] X. Dang, Z. Chen, and Z. Hao, "Emotion recognition method using millimetre wave radar based on deep learning," *IET Radar, Sonar & Navigation*, vol. 16, no. 11, pp. 1796–1808, 2022, ISSN: 1751-8784.

[10] K.-C. Peng, M.-C. Sung, F.-K. Wang, and T.-S. Horng, "Noncontact Vital Sign Sensing Under Nonperiodic Body Movement Using a Novel Frequency-Locked-Loop Radar," *IEEE Transactions on Microwave Theory and Techniques*, vol. 69, no. 11, pp. 4762–4773, 2021, ISSN: 0018-9480.

[11] Q. Lin, D. Zhang, L. Chen, H. Ni, and X. Zhou, "Managing elders' wandering behavior using sensors-based solutions: A survey," *International Journal of Gerontology*, vol. 8, no. 2, pp. 49–55, 2014, ISSN: 1873-9598.

[12] R. Yu, Y. Du, J. Li, A. Napolitano, and J. Le Kernec, "Radar-based human activity recognition using denoising techniques to enhance classification accuracy," *IET Radar, Sonar & Navigation*, vol. 18, no. 2, pp. 277–293, 2024, ISSN: 1751-8784.

[13] I. Ullmann, R. G. Guendel, N. C. Kruse, F. Fioranelli, and A. Yarovoy, "Radar-Based Continuous Human Activity Recognition with Multi-Label Classification," in *Proceedings of IEEE Sensors*, IEEE, 2023, pp. 1–4, ISBN: 9798350303872.

[14] N. C. Kruse, F. Fioranelli, and A. Yarovoy, "Radar Point Cloud Processing Methods for Human Activity Classification With Point Transformer Networks," *IEEE Transactions on Radar Systems*, vol. 2, pp. 1–12, 2023, ISSN: 2832-7357.

[15] M. J. Bocus, K. Chetty, and R. J. Piechocki, "UWB and WiFi Systems as Passive Opportunistic Activity Sensing Radars," in *2021 IEEE Radar Conference (RadarConf21)*, Atlanta, GA, USA: IEEE, 2021, ISBN: 9781728176093.

[16] A. Gorji, H.-U.-R. Khalid, A. Bourdoux, and H. Sahli, "On the Generalization and Reliability of Single Radar-Based Human Activity Recognition," *IEEE Access*, vol. 9, pp. 85 334–85 349, 2021, ISSN: 2169-3536.

[17] S. Z. Gürbüz, E. Kurtoglu, M. M. Rahman, and D. Martelli, "Gait variability analysis using continuous RF data streams of human activity," *Smart Health*, p. 100 334, 2022, ISSN: 23526483.

[18] Z. Yang, H. Wang, P. Ni, P. Wang, Q. Cao, and L. Fang, "Real-time Human Activity Classification From Radar With CNN-LSTM Network," in *2021 IEEE 16th Conference on Industrial Electronics and Applications (ICIEA)*, Chengdu, China: IEEE, 2021, pp. 50–55, ISBN: 978-1-6654-2248-2.

[19] N. C. Kruse, F. Fioranelli, and A. Yarovoy, "Continuous Human Activity Classification with Radar Point Clouds and Point Transformer Networks," in *2023 20th European Radar Conference (EuRAD)*, IEEE, Berlin, 2023, pp. 302–305, ISBN: 9782874870743.

[20] H. Zhao, L. Jiang, J. Jia, P. Torr, and V. Koltun, "Point Transformer," in *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, IEEE, 2021, pp. 16 239–16 248, ISBN: 978-1-6654-2812-5. arXiv: 2012.09164.

[21] A. Vaswani, N. Shazeer, N. Parmar, *et al.*, *Attention is all you need*, 2017. arXiv: 1706.03762 [cs.CL].

[22] R. G. Guendel, M. Unterhorst, N. Kruse, F. Fioranelli, and A. Yarovoy, *Dataset of continuous human activities performed in arbitrary directions collected with a distributed radar network of five nodes*, 2021.

[23] *Datasheet pulson410*, Time Domain, Nov. 2013.

[24] Z. Guo, R. G. Guendel, A. Yarovoy, and F. Fioranelli, "Point Transformer-Based Human Activity Recognition Using High-Dimensional Radar Point Clouds," in *Proceedings of the IEEE Radar Conference*, Delft University of Technology, vol. 2023-May, 2023, ISBN: 9781665436694.

[25] S. Zhu, R. G. Guendel, A. Yarovoy, and F. Fioranelli, "Continuous Human Activity Recognition With Distributed Radar Sensor Networks and CNN–RNN Architectures," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–15, 2022, ISSN: 0196-2892.

[26] N. C. Kruse, R. G. Guendel, F. Fioranelli, and A. G. Yarovoy, "Reconstruction of Extended Target Intensity Maps and Velocity Distribution for Human Activity Classification," *IEEE Transactions on Radar Systems*, 2025.