Golipoor, Sahar; Sigg, Stephan

RFID-based Human Activity Recognition Using Multimodal Convolutional Neural Networks

Published in:
2024 IEEE 29th International Conference on Emerging Technologies and Factory Automation, ETFA 2024

DOI:
10.1109/ETFA61755.2024.10710698

Published: 01/01/2024

Document Version
Peer-reviewed accepted author manuscript, also known as Final accepted manuscript or Post-print

Please cite the original version:
Golipoor, S., & Sigg, S. (2024). RFID-based Human Activity Recognition Using Multimodal Convolutional Neural Networks. In T. Facchinetti, A. Cenedese, L. L. Bello, S. Vitturi, T. Sauter, & F. Tramarin (Eds.), *2024 IEEE 29th International Conference on Emerging Technologies and Factory Automation, ETFA 2024* (IEEE International Conference on Emerging Technologies and Factory Automation, ETFA). IEEE. https://doi.org/10.1109/ETFA61755.2024.10710698

RFID-based Human Activity Recognition Using Multimodal Convolutional Neural Networks

Sahar Golipoor, Stephan Sigg
Department of Information and Communications Engineering, Aalto University, Finland
e-mails: {sahar.golipoor, stephan.sigg}@aalto.fi

***Abstract*—Recognition of human activities is crucial for enhancing safety, efficiency, and productivity within industrial and factory automation settings. This paper introduces a model for human activity recognition that leverages battery-less body-worn reflective antenna components. We perform preprocessing on both the backscattered phase and Received Signal Strength (RSS) signals. Independently and simultaneously, we extract features from phase and RSS signals using a feature extractor implementing a convolutional neural network (CNN). These features are then concatenated and fed into a fully connected (FC) layer employing the rectified linear unit (ReLU) activation function, followed by another FC layer utilizing a softmax function. This model, which merges extracted features from both phase and RSS, is termed *late fusion* model. We show that *late fusion* yields better performance than combining phase and RSS signals before feeding them into the neural network. By employing battery-free body-worn Radio frequency identification (RFID) tags, we surpass existing models, achieving an accuracy of 97.5% in recognizing five activities.**

***Index Terms*—RFID, activity recognition, human-sensing, multimodal learning**

I. Introduction

Human activity recognition (HAR) involves identifying the actions performed by an individual through the collection and analysis of data from various sources such as wearable [1] or environmental sensors [2]. Applications include smart home, surveillance, healthcare and assisted living [3]. Utilizing wearable sensors presents certain challenges [4]. Wearable sensors require frequent charging or replacement, since their operation and continuous data processing consume significant power. Additionally, for convenience, devices are lightweight. Therefore, it becomes increasingly difficult to maintain functionality given the limited available space.

We propose to integrate passive Radio frequency identification (RFID) tags with clothing as sensors for human activities. Passive RFID tags are energy efficient as they do not require an own power supply and do not process data but backscatter impinging signals emitted by environment-integrated reader devices. In the literature, RFID has been exploited for human [5] and respiration sensing [6], however, the proposed solutions expect RFID devices in the environment. RFID tags can offer enhanced security and authentication capabilities [7]. Compared to environment-based instrumentations, sensing via on-body worn RFID devices inherits several benefits, such as increased privacy control (opt-in solutions, as people need to consciously wear sensors in order to be sensed.) and distinction of various body parts [8].

In this work, we demonstrate activity recognition from body-worn RFID tags. Particularly, we propose activity recognition utilizing a multimodal learning. The system is evaluated based on data collected in an indoor office environment.

This work is part of the Horizon Europe EIC project SustAIn. In SustAIn, we develop a sustainable low energy building monitoring system.

II. Related work

Fuelled by advances in machine learning, wearable sensor-based HAR accuracy has significantly improved [9]–[11]. To capture long term dependencies in time series data, a CNN architecture that incorporates a multihead attention mechanism was introduced in [12] for HAR task. Multilevel network based on spatiotemporal attention and multiscale temporal embedding was studied in [13]. Addressing incomplete feature extraction, a multifeature extraction framework was introduce with combination of bidirectional gated recurrent unit, attention mechanism, and residual network-18 [14]. While the attention mechanism can enhance the significance of important features and reduce the impact of irrelevant ones, it also introduces architectural complexity, increases computational expenses, and complicates implementation. In this regard, A light residual neural network was introduced in [15], aiming to decrease the number of parameters and simplify the training process. Contrary to the above works that used wearable sensor data, such as inertial sensors, we use RFID technology and offer our dataset for activity recognition.

RFID-based activity recognition is gaining interest because it is lightweight and energy-efficient [16]. For instance, the authors in [17] exploit Received Signal Strength (RSS) data from RFID readings to feed a CNN and a LSTM, to sense teamwork activities in a trauma resuscitation [17] and daily routine activities [18]. Activity classification was also studied in [19] via multivariate Gaussian using maximum likelihood estimation algorithm embedding RFID array in an indoor environment. In [20], a method for generating an RFID skeleton feature matrix has been proposed. In [8], we have further demostrated that it is possible to distinguish different body parts by exploiting variable phase profiles for body-worn RFID tag groups which allows high recognition rates for complex movements of individual body parts. RFID-based sensing studies primarily focus on the phase of the backscattered

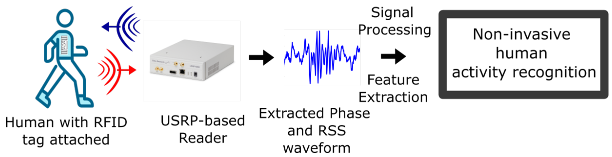


Fig. 1: RFID hardware setting for activity recognition

signal. To increase data diversity, this paper incorporates both phase and RSS, extracting time-domain statistical features and frequency-domain features as recommended in [21].

RFID utilization was also studied for occupancy detection. Using a carpet embedded with RFID tags, a random forest was trained to predict population density [22]. In [23], both RSS and phase were calibrated to mitigate the interference from the line-of-sight (LoS) and multi-path components and CNN was applied to present an occupant counting system. In [24], people counting problem turned into a clustering problem and tags' feature vectors were build. Then, the distance between two feature vectors by combining the Hausdorff distance and the Euclidean distance was used.

Although existing literature predominantly focuses on embedded fixed environmental RFID arrays, our research delves into the utilization of RFID tags affixed to clothing worn by individuals for the purpose of activity recognition. In this context, the reflection of signals is influenced by the anatomy and tissues of the human body [25]. Varied body compositions and shapes can lead to differences in how RF signals are reflected, potentially resulting in various patterns for the same activities and posing challenges to recognition.

III. Backscatter technology for HAR

Several authors have considered the use of backscatter devices for human activity recognition [26], [27]. Also commercial RFID technology has been extensively employed in wireless sensing, demonstrating high reliability and robust performance in tracking and monitoring applications across various industries [28], [29].

When an RFID tag enters a reader's electromagnetic field, it reflects the signal emitted by the reader [30]. The tag modulates and backscatters this RF signal according to the EPCglobal Gen2 (ISO/IEC 18000-63) communication protocol to ensure reliable connectivity. By analyzing the reflected signal, both the stored information and the unique tag ID can be retrieved. Additionally, the reader device provides RSS and signal phase data, enabling the detection of tagged object movements through appropriate algorithms (cf. Fig 1) [20].

The data collected from RFID devices must be preprocessed. The goal of this pre-processing is to transform raw data into a consistent and standardized format, enhancing its readability. This crucial step involves organizing the data to enhance its accessibility and intelligibility, which, in turn, simplifies the analysis and interpretation processes. The formatted tag data includes three types of information:

- **Tag ID**: Unique Identification (UID) code of each tag. The size of the UID is 96 bits.

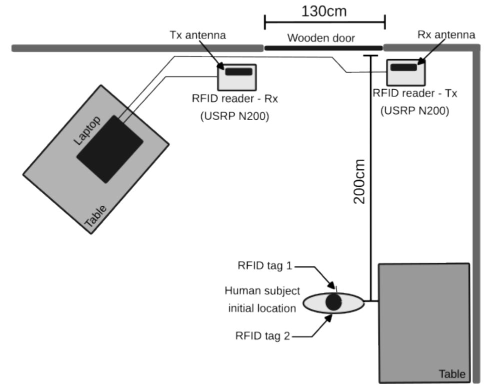


Fig. 2: Schematic of the sensing area where the experiment has been conducted.

- **Phase**: The relative phase difference in the signal when the tag reflects the signal emitted by the reader is captured. This phase data can be indicative of the tag's orientation and movement, offering a more nuanced understanding of the tag's spatial positioning in relation to the reader. By analyzing phase variations over time, it is possible to detect subtle changes in tag orientation, which can be used to infer the dynamics of user interactions.
- **RSS**: RSS denotes the power present in a received radio signal. This metric can be used as a proxy for estimating the distance between the RFID tag and a reader. It may hence be used to determine a tag's location relative to the reader device. The RSS data can for instance be used to track the proximity of an individual to various zones in an office, assess user engagement with specific office resources, or detect the presence of an individual within defined spatial thresholds.

IV. Problem statement

For activity recognition in an indoor space, a proper arrangement of the RFID reader is vital for robust tag reading due to its limited sensing range. RFID tags can be detected by a reader within several meters. We install an RFID reader in an office, and perform activities. The orientation of the utilized RFID antennas is configured to cover the entrances of office and open spaces as shown in Fig 2.

By detecting the tag IDs, identification of subjects is also possible. Furthermore, to improve the recognition accuracy regardless of the orientation of the subject, two RFID tags are attached to upper torso and the upper back of the study subjects' clothing. To further improve recognition, two additional tags may also be placed to the outer sides of the right and left upper arms. Activities performed by human subjects equipped with the RFID tags in this way can then be recognized. In a settings where multiple instrumented subjects reside in the same area, their tag IDs can be read separately and their movement and motion can be tracked individually.

A. System Implementation and Data Collection

We utilize the USRP-based UHF RFID reader developed in [31]. The reader software uses GNU Radio and consists

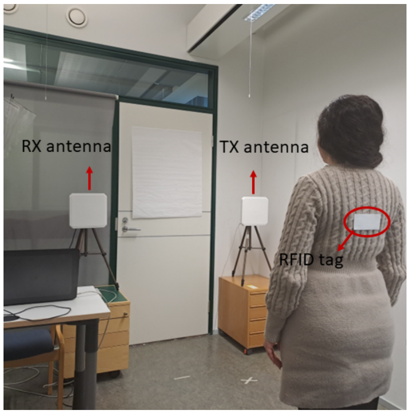


Fig. 3: Illustration of the setup

of 6 blocks (USRP source, matched filter, gate, tag decoder, Generation-2 UHF RFID logic, and USRP sink).

For the study, we attach two UHF RFID tags (Zebra Z-Select 2000T), one at the upper front torso and another in the upper back of a human subject (cf. Fig 3).

The RFID system operates in the 910MHz range. The distance between transmit and receive antennas is 130cm. RSS and phase of backscattered signals are captured for five different activities related to presence detection including Opening the door (O), Closing the door (C), Turning the light on (TL), Leaving the room (EX), and Entering the room (EN) (cf. Fig 4).

We utilized a USRP N200 with 15 dBm output power and two circularly polarized antennas with 9 dBic antenna gain as well as elevation and azimuth beamwidth of 70°. We collected data from 8 people (5 male and 3 female) with different heights and physiques in the age range 25-45. For every activity and person, we collected RSS and phase while the activities (O, C, TL, EX, EN) were performed. The total number of data points collected is 80 for every activity.

B. Data Pre-processing

The extracted RSS and phase signals are noisy and erratic and require pre-processing before feeding them to activity recognition algorithms. We pre-process both RSS and phase using signal smoothing and noise reduction as well as signal normalization.

Since high-frequency noise exists in the RSS and phase of the signal, we utilize a Savitzky-Golay (S-G) filter with polynomial order of 15, with a frame length of 500 [32] and moving average filtering, which preserves the characteristic patterns of the underlying signals.

We normalize the RSS and phase to enhance activity-relevant changes and to alleviate the impact of background signals by mapping them to a range of $[0, 1]$ and $[-1, 1]$, respectively. The normalization of the m-th RSS reading α_m is computed as

$$\tilde{\alpha}_m = \frac{\alpha_m}{\alpha_{\text{max}}} \tag{1}$$

where α_{max} denote the maximum of all RSS readings of the tag (i.e., $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \ldots, \alpha_M]$). The total number of readings is M. Besides, the normalization of the m-th phase reading φ_m is computed as

$$\tilde{\varphi}_m = \begin{cases} \frac{\varphi_m - \bar{\varphi}}{\varphi_{\text{max}}} & \varphi_m \geq 0 \\ -\frac{\varphi_m - \bar{\varphi}}{\varphi_{\text{min}}} & \varphi_m < 0 \end{cases}, \tag{2}$$

where $\bar{\varphi}$, φ_{max}, and φ_{min} denote the mean, maximum, and minimum of all phase readings of the tag (i.e., $\boldsymbol{\varphi} = [\varphi_1, \varphi_2, \ldots, \varphi_M]$), respectively.

C. Data Augmentation

As illustrated in Fig 5, we apply time series augmentation techniques such as jittering, flipping, scaling, magnitude warping [33], and amplitude and phase perturbations (APP) in the frequency domain [34].

V. Activity recognition models

We introduce a multi-modal learning architecture in which features of RSS and phase are extracted. For evaluation, we compare it to a random forest classifier [21] and to a residual network [15].

A. Random Forest Model

We extract 13 statistical features from RSS and phase including the mode, the median, the first quartile, the third quartile, the mean, the max, the min, the range, the variance, the standard deviation, the third-order central moment, the kurtosis, and the skewness. Moreover, we use discrete wavelet transform with Daubechies wavelet and extract both low-frequency and high-frequency coefficients.

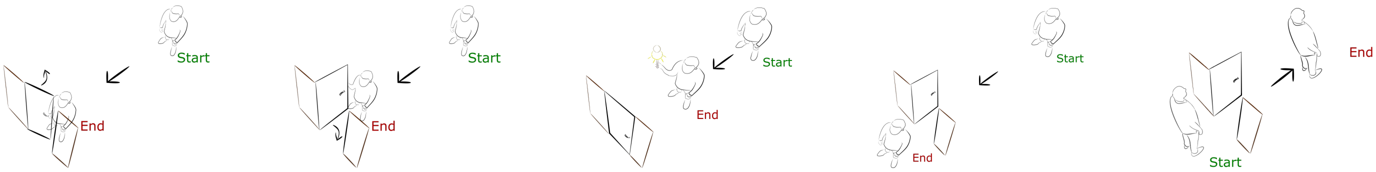
Accordingly, we create three datasets:

Dataset **SPR**:	Only statistical features from phase and RSS and their concatenation
Dataset **SWP**:	Statistical features and wavelet coefficients from phase
Dataset **SWPR**:	Statistical features and wavelet coefficients from phase and RSS and their concatenation

We then use the above datasets to train three separate classifiers based on random forest.

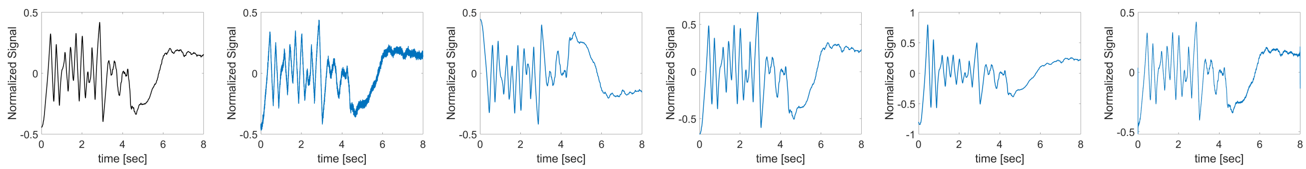
B. Residual Network (Early Fusion Model)

In [15], a residual network is proposed where the inputs are tensors, each containing nine 1-D signals captured by inertial sensors of a smartphone. These signals are triaxial acceleration, triaxial estimated body acceleration, and triaxial angular velocity. Drawing inspiration from this data-feeding method, we construct input data incorporating both phase and RSS values, with dimensions of (2, 1, 25600), where 25600 represents the samples (length) of the phase and RSS readings. We then feed this tensor data into the residual network [15]. Since we combine phase and RSS before extracting features, we refer to this approach as *early fusion*.



(a) Opening the door (O) (b) Closing the door (C) (c) Turn on the light (TL) (d) Leaving the room (EX) (e) Entering the room (EN)

Fig. 4: Schematic illustration of the activities conducted during the study.



(a) Original signal (b) Jittering (c) Flipping (d) Scaling (e) Magnitude warping (f) APP

Fig. 5: Time series augmentation operations applied to the RSS and phase sequences

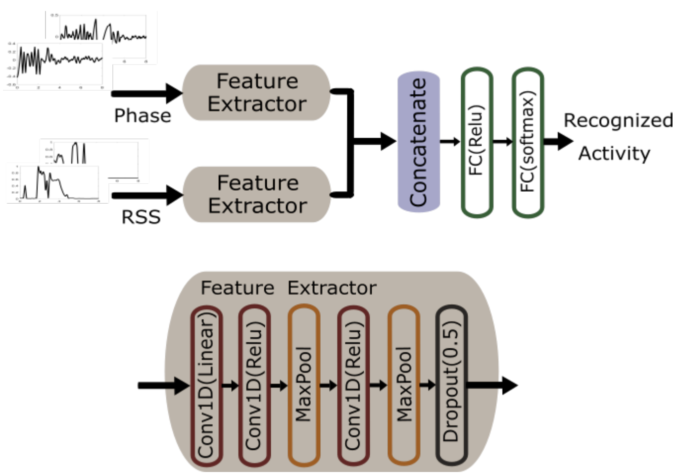


Fig. 6: In late fusion, RSS and phase features are extracted individually before combining them

C. Multimodal Learning (Late Fusion Model)

We extract features of phase and RSS independently and simultaneously by the feature extractor (cf. Fig 6). Subsequently, the concatenated features are fed into a fully connected (FC) layer utilizing the rectified linear unit (ReLU) activation function, which is then followed by another FC layer employing a softmax function. Since we merge phase and RSS only after feature extraction, we refer to this model as *late fusion*.

VI. Results and Discussion

The confusion matrices are presented in TABLE II to TABLE VI. For the random forest classifier, the datasets **SPR**, **SWP**, **SWPR** are divided into 90% and 10% for training and testing, respectively. By extracting only statistical features from both phase and RSS, i.e., dataset **SPR**, an overall test accuracy of 85.66% is achieved. Beside, extracting both statistical features and wavelet coefficients from the phase, i.e. dataset **SWP**, yields an overall test accuracy of 92.33%. Applying a random forest classifier to dataset **SWPR**, which includes statistical features and wavelet coefficients from both phase and RSS, results in superior performance compared to the two previous datasets, with an accuracy of 95%. The activity of entering the room (EN) is classified with 100% accuracy in this case and the accuracy for the remaining classes exceeds 93%. This confirms that considering both RSS and phase signals enhances recognition accuracy, and extracting wavelet coefficients as attributes offers a more sophisticated interpretation [35].

TABLE I: Accuracy (Acc.), Precision (Pre.), Recall (Rec.), and F1-score (F1) of Methods

Method \ Metric	Acc.	Pre.	Rec.	F1
Random forest classifier with **SPR**	85.66	85.82	85.61	85.71
Random forest classifier with **SWP**	92.33	92.94	92.25	92.59
Random forest classifier with **SWPR**	95	95	94.97	94.98
Early Fusion Model	95.16	95.22	95.12	95.16
Late Fusion Model	97.5	97.47	97.46	97.46

Utilizing phase and RSS, multimodal learning and the residual network outperform the random forest classifier. The datasets are split into 80%, 10%, and 10% for training, validation, and testing, respectively. Employing a residual network, the Early Fusion Model [15], achieves an overall test accuracy of 95.16%. The confusion matrix for this model is illustrated in TABLE V. Applying this model, the activity of turning on the light (TL) is identified with 100% accuracy, while the accuracy for the other classes surpasses 91%. The late Fusion Model surpasses other models and attains an overall test accuracy of 97.5% (TABLE VI). Entering the room (EN) is classified with 100% and accuracy of other classes are all above 95%. An overall comparison for methods is shown in TABLE I.

VII. Conclusion

We explored the application of body-mounted RFID tags to recognize human activities. By utilizing a USRP-based reader along with two UHF RFID tags affixed to the upper torso and upper back of subjects' clothing, we successfully identified five activities with an accuracy of 97.5% using multimodal

TABLE II: Confusion matrix resulted from random forest classifier [21] with dataset **SPR**

	Predicted label					
True label	C	EN	EX	O	TL	Recall
C	110 80.2%	2 1.4%	2 1.4%	19 13.8%	4 2.9%	80.2%
EN	4 3.3%	108 90.7%	5 4.2%	2 1.6%	0 0%	90.7%
EX	2 1.7%	4 3.4%	98 85.2%	0 0%	11 9.5%	85.2%
O	6 5.1%	3 2.5%	3 2.5%	101 87%	3 2.5%	87%
TL	1 0.8%	1 0.8%	13 11.5%	1 0.8%	97 85.8%	85.8%
Pre.	89.4%	91.5%	80.9%	82.1%	84.3%	

TABLE III: Confusion matrix resulted from random forest classifier [21] with dataset **SWP**

	Predicted label					
True label	C	EN	EX	O	TL	Recall
C	126 91.9%	0 0%	1 0.7%	3 2.1%	7 5.1%	91.9%
EN	0 0%	113 94.9%	2 1.6%	0 0%	4 3.3%	94.9%
EX	1 0.8%	1 0.8%	96 83.4%	0 0%	17 14.7%	83.4%
O	1 0.8%	0 0%	3 2.5%	111 95.6%	1 0.8%	95.6%
TL	0 0%	0 0%	3 2.6%	2 1.7%	108 95.5%	95.5%
Pre.	98.4%	99.1%	91.4%	95.6%	78.8%	

TABLE IV: Confusion matrix resulted from random forest classifier [21] with dataset **SWPR**

	Predicted label					
True label	C	EN	EX	O	TL	Recall
C	128 93.4%	1 0.7%	1 0.7%	7 5.1%	0 0%	93.4%
EN	0 0%	119 100%	0 0%	0 0%	0 0%	100%
EX	2 1.7%	0 0%	108 93.9%	1 0.8%	4 3.4%	93.9%
O	5 4.3%	1 0.8%	1 0.8%	109 93.9%	0 0%	93.9%
TL	1 0.8%	0 0%	5 4.4%	1 0.8%	106 93.8%	93.8%
Pre.	94.1%	98.3%	93.9%	92.3%	96.3%	

TABLE V: Confusion matrix resulted from Early Fusion Model [15]

	Predicted label					
True label	C	EN	EX	O	TL	Recall
C	129 92.8%	1 0.7%	1 0.7%	7 5%	1 0.7%	92.8%
EN	0 0%	110 96.4%	0 0%	0 0%	4 3.5%	96.4%
EX	2 1.8%	1 0.9%	99 91.6%	0 0%	6 5.5%	91.6%
O	2 1.6%	1 0.8%	0 0%	115 95%	3 2.4%	95%
TL	0 0%	0 0%	0 0%	0 0%	118 100%	100%
Pre.	96.9%	97.3%	99%	94.2%	89.3%	

TABLE VI: Confusion matrix resulted from Late Fusion Model

	Predicted label					
True label	C	EN	EX	O	TL	Recall
C	135 97.1%	1 0.7%	0 0%	3 2.1%	0 0%	97.1%
EN	0 0%	114 100%	0 0%	0 0%	0 0%	100%
EX	0 0%	0 0%	105 97.2%	0 0%	3 2.7%	97.2%
O	1 0.8%	0 0%	1 0.8%	116 95.8%	3 2.4%	95.8%
TL	1 0.8%	0 0%	2 1.6%	0 0%	115 97.4%	97.4%
Pre.	98.5%	99.1%	97.2%	97.4%	95%	

CNNs. The networks were trained individually with preprocessed phase and RSS signals. Our findings showed that combining both phase and RSS improves the accuracy of activity recognition. In our model, features of phase and RSS are first extracted through a CNN and then fused. We demonstrated that this method of fusion surpasses the approach of combining phase and RSS signals prior to feeding them into the neural network.

VIII. Acknowledgements

We acknowledge funding by the European Union through the Horizon Europe EIC Pathfinder project SUSTAIN (project no. 101071179). Views and opinions expressed are those of the authors and do not necessarily reflect those of the European Union.

References

[1] S. Zhang, Y. Li, S. Zhang, F. Shahabi, S. Xia, Y. Deng, and N. Alshurafa, "Deep learning in human activity recognition with wearable sensors: A review on advances," *Sensors*, vol. 22, no. 4, p. 1476, 2022.

[2] I. Ullmann, R. G. Guendel, N. C. Kruse, F. Fioranelli, and A. Yarovoy, "A survey on radar-based continuous human activity recognition," *IEEE Journal of Microwaves*, 2023.

[3] S. K. Yadav, K. Tiwari, H. M. Pandey, and S. A. Akbar, "A review of multimodal human activity recognition with special emphasis on classification, applications, challenges and future directions," *Knowledge-Based Systems*, vol. 223, p. 106970, 2021.

[4] L. Yang, O. Amin, and B. Shihada, "Intelligent wearable systems: Opportunities and challenges in health and sports," *ACM Computing Surveys*, vol. 56, no. 7, pp. 1–42, 2024.

[5] H. Landaluce, L. Arjona, A. Perallos, F. Falcone, I. Angulo, and F. Muralter, "A review of iot sensing applications and challenges using rfid and wireless sensor networks," *Sensors*, vol. 20, no. 9, p. 2495, 2020.

[6] J. L. L. Araujo, G. K. d. F. Serres, S. Genovesi, D. F. d. S. Santos, and A. J. R. Serres, "A flexible uhf rfid respiratory sensor," *Available at SSRN 4681700*.

[7] F. Costa and et al., "A review of rfid sensors, the new frontier of internet of things," *Sensors*, vol. 21, no. 9, p. 3138, 2021.

[8] S. Golipoor and S. Sigg, "Accurate rf-sensing of complex gestures using rfid with variable phase-profiles," in *2023 IEEE 32nd International Symposium on Industrial Electronics (ISIE)*. IEEE, 2023, pp. 1–4.

[9] I. Nirmal, A. Khamis, M. Hassan, W. Hu, and X. Zhu, "Deep learning for radio-based human sensing: Recent advances and future directions," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 2, pp. 995–1019, 2021.

[10] N. Y. Hammerla and T. Plötz, "Let's (not) stick together: pairwise similarity biases cross-validation in activity recognition," in *Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing*, 2015, pp. 1041–1051.

[11] P. Kumar, S. Chauhan, and L. K. Awasthi, "Human activity recognition (har) using deep learning: Review, methodologies, progress and future research directions," *Archives of Computational Methods in Engineering*, vol. 31, no. 1, pp. 179–219, 2024.

[12] T.-H. Tan, Y.-L. Chang, J.-R. Wu, Y.-F. Chen, and M. Alkhaleefah, "Convolutional neural network with multi-head attention for human activity recognition," *IEEE Internet of Things Journal*, 2023.

[13] H. Zhang and L. Xu, "Multi-stmt: Multi-level network for human activity recognition based on wearable sensors," *IEEE Transactions on Instrumentation and Measurement*, 2024.

[14] Y. Wang, H. Xu, Y. Liu, M. Wang, Y. Wang, Y. Yang, S. Zhou, J. Zeng, J. Xu, S. Li *et al.*, "A novel deep multifeature extraction framework based on attention mechanism using wearable sensor data for human activity recognition," *IEEE Sensors Journal*, vol. 23, no. 7, pp. 7188–7198, 2023.

[15] F. M. Calatrava-Nicolás and O. M. Mozos, "Light residual network for human activity recognition using wearable sensor data," *IEEE Sensors Letters*, 2023.

[16] J. Xiao, H. Li, M. Wu, H. Jin, M. J. Deen, and J. Cao, "A survey on wireless device-free human sensing: Application scenarios, current solutions, and open issues," *ACM Computing Surveys*, vol. 55, no. 5, pp. 1–35, 2022.

[17] X. Li, Y. Zhang, I. Marsic, A. Sarcevic, and R. S. Burd, "Deep learning for rfid-based activity recognition," in *Proceedings of the 14th ACM Conference on Embedded Network Sensor Systems CD-ROM*, 2016, pp. 164–175.

[18] G. A. Oguntala, Y. F. Hu, A. A. Alabdullah, R. A. Abd-Alhameed, M. Ali, and D. K. Luong, "Passive rfid module with lstm recurrent neural network activity classification algorithm for ambient-assisted living," *IEEE Internet of Things Journal*, vol. 8, no. 13, pp. 10 953–10 962, 2021.

[19] G. A. Oguntala, R. A. Abd-Alhameed, N. T. Ali, Y.-F. Hu, J. M. Noras, N. N. Eya, I. Elfergani, and J. Rodriguez, "Smartwall: Novel rfid-enabled ambient human activity recognition using machine learning for unobtrusive health monitoring," *IEEE Access*, vol. 7, pp. 68 022–68 033, 2019.

[20] Z. Wang, Y. Chen, H. Zheng, M. Liu, and P. Huang, "Body rfid skeleton-based human activity recognition using graph convolution neural network," *IEEE Transactions on Mobile Computing*, 2023.

[21] S. Zhang, Z. Ma, C. Yang, X. Kui, X. Liu, W. Wang, J. Wang, and S. Guo, "Real-time and accurate gesture recognition with commercial rfid devices," *IEEE Transactions on Mobile Computing*, 2022.

[22] C. Y. Jiang, B. Y. Wang, T. O. Tai, K. W. Tam, L. Chen, C. H. Chio, C. Teng, and N. Kong, "An intelligent occupancy detection system for smart tourism based on rfid passive tag antenna array and random forest," in *2023 IEEE 13th International Conference on RFID Technology and Applications (RFID-TA)*. IEEE, 2023, pp. 83–86.

[23] G. Xu and E. C. Kan, "Device-free occupant counting using ambient rfid and deep learning," in *2024 IEEE Topical Conference on Wireless Sensors and Sensor Networks (WiSNeT)*. IEEE, 2024, pp. 49–52.

[24] X. Wang, J. Liu, X. Yu, X. Chen, Y. Yu, and Z. Zhao, "People counting with carry-on rfid tags," in *2023 IEEE/ACM 31st International Symposium on Quality of Service (IWQoS)*. IEEE, 2023, pp. 1–10.

[25] D. Vasisht, G. Zhang, O. Abari, H.-M. Lu, J. Flanz, and D. Katabi, "In-body backscatter communication and localization," in *Proceedings of the 2018 Conference of the ACM Special Interest Group on Data Communication*, 2018, pp. 132–146.

[26] J. Ryoo, Y. Karimi, A. Athalye, M. Stanaćević, S. R. Das, and P. Djurić, "Barnet: Towards activity recognition using passive backscattering tag-to-tag network," in *Proceedings of the 16th annual international conference on mobile systems, applications, and services*, 2018, pp. 414–427.

[27] A. Saha, S. Rajak, J. Saha, and C. Chowdhury, "A survey of machine learning and meta-heuristics approaches for sensor-based human activity recognition systems," *Journal of ambient intelligence and humanized computing*, vol. 15, no. 1, pp. 29–56, 2024.

[28] M. J. C. Samonte, D. A. Medel, J. M. N. Odicta, and M. Z. L. T. Santos, "Using iot-enabled rfid smart cards in an indoor people-movement tracking system with risk assessment," *Journal of Advances in Information Technology*, vol. 14, no. 2, 2023.

[29] V. Kalaiselvi, K. Swathi, S. Rakshana, H. Shanmugasundaram, V. Kukreja *et al.*, "Implementation of iot enabled rfid driven human life rescuer in railways," in *2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC)*. IEEE, 2023, pp. 1431–1435.

[30] Y. Wang, "Contactless human activity and sport action identification based on uhf rfid tag wall and ai," *Physical Communication*, vol. 64, p. 102347, 2024.

[31] N. Kargas, F. Mavromatis, and A. Bletsas, "Fully-coherent reader with commodity sdr for gen2 fm0 and computational rfid," *IEEE Wireless Communications Letters*, vol. 4, no. 6, pp. 617–620, 2015.

[32] R. W. Schafer, "What is a savitzky-golay filter?[lecture notes]," *IEEE Signal processing magazine*, vol. 28, no. 4, pp. 111–117, 2011.

[33] B. K. Iwana and S. Uchida, "An empirical survey of data augmentation for time series classification with neural networks," *Plos one*, vol. 16, no. 7, p. e0254841, 2021.

[34] J. Gao, X. Song, Q. Wen, P. Wang, L. Sun, and H. Xu, "Robusttad: Robust time series anomaly detection via decomposition and convolutional neural networks," *arXiv preprint arXiv:2002.09545*, 2020.

[35] H. Abdelnasser, M. Youssef, and K. A. Harras, "Wigest: A ubiquitous wifi-based gesture recognition system," in *2015 IEEE conference on computer communications (INFOCOM)*. IEEE, 2015, pp. 1472–1480.