Multioccupancy Activity Recognition Based on Deep Learning Models Fusing UWB Localization Heatmaps and Nearby-Sensor Interaction

Miguel Ángel Anguita-Molina, Pedro J. S. Cardoso, João M. F. Rodrigues, Javier Medina-Quero, and Aurora Polo-Rodríguez

***Abstract*—Human activity recognition (HAR) focuses on developing systems and techniques to recognize and categorize human actions automatically based on sensor data. This study combines ultrawideband (UWB) technology and binary sensors to describe and recognize daily activities in real-world environments with multiple occupants, ensuring accurate user localization through noninvasive and privacy-respecting methods. A novel method that combines wearables with UWB technology, which allows the generation of heatmaps for accurate positioning, and binary sensors, which collect nearby interaction with daily activities in naturalistic conditions, is presented. A dataset composed of real-world data collected from three individuals in a real-life environment (house) was compiled. Advanced deep learning models are implemented to effectively fuse spatiotemporal information, leading to an encouraging performance in recognition of daily activities. The promising results suggest that this approach could be viable for large-scale deployments in future smart environments.**

***Index Terms*—Human activity recognition (HAR), localization heatmap, multioccupancy, ultrawideband (UWB).**

I. Introduction

HUMAN activity recognition (HAR) relies on the integration of various multimodal devices and sensors to detect and interpret human movements and behaviors with precision [1]. This technology is crucial in healthcare, assisted living, and smart home systems, where detailed monitoring of indoor activities is essential [2]. The ability to accurately analyze human activities can significantly improve safety, health outcomes, and overall quality of life. Recent advances in sensor technology, including binary and vision sensors and ultrawideband (UWB) technology, have played a key role in improving the HAR capabilities for indoor environments [3]. The integration of advanced machine learning techniques and the combination of data from various sensors further support these developments by minimizing errors and reducing signal interference, thus improving the reliability of the system [4], [5]. This technological synergy is crucial for a more precise interpretation of human behavior. Binary sensors are frequently used in HAR systems in this context due to their affordability and privacy-preserving feature of only detecting the presence or absence of people. However, their drawbacks become apparent in multioccupant situations, making accurate identification and tracking of activities difficult [6]. This underscores the need for more advanced or complementary technologies to improve the precision of monitoring in complex environments.

One of the primary challenges in multioccupant environments lies in the inability of environmental sensors to discriminate between individuals or accurately identify which person is performing a specific activity [7]. This limitation reduces the effectiveness of activity tracking and analysis. Furthermore, the nature of human behavior, together with the limitations of low-cost sensors and privacy concerns, complicates the deployment of HAR systems in these settings. UWB technology has become a promising approach to precise indoor location tracking, thanks to techniques, such as the Time Difference of Arrival (TDoA) and the received signal strength indicator (RSSI) [8], [9], [10]. This technology proves particularly efficient in multioccupant settings, enhancing real-time location systems (RTLS) for consistent and accurate monitoring. In healthcare, UWB-based RTLS are widely used for patient monitoring, fall detection, and emergency response [9], [11]. Furthermore, integration of nearby interaction with sensors improves the accuracy of activity detection in settings with multiple occupants by providing environmental and behavioral context [12], [13].

Building on the context of sensorization and activity recognition, this research introduces an intelligent system that leverages UWB technology to describe and recognize activities in environments with multiple occupants, ensuring accurate

Received 13 December 2024; accepted 13 January 2025. Date of publication 17 January 2025; date of current version 23 May 2025. This work was supported in part by the Spanish Institute of Health ISCIII under Project DTS21-00047; in part by the UID/04516/NOVA Laboratory for Computer Science and Informatics (NOVA LINCS) with the financial support of FCT.IP; in part by the University of Granada through the Project "Detection of Sports Activities Using IoT Devices" (6324) from the Office of Technology Transfer (OTRI); and in part by the Open Access funding provided by the Universidad de Granada/CBUA. *(Corresponding author: Javier Medina-Quero.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Regional Ethics Committee under Application No. DTS21-00047.

Miguel Ángel Anguita-Molina, Javier Medina-Quero, and Aurora Polo-Rodríguez are with the Department of Computer Engineering, Automation and Robotics, University of Granada, 18071 Granada, Spain (e-mail: javiermq@ugr.es).

Pedro J. S. Cardoso and João M. F. Rodrigues are with the NOVA LINCS and the Institute of Engineering, University of Algarve, 8005-139 Faro, Portugal.

Digital Object Identifier 10.1109/JIOT.2025.3531316

user localization through noninvasive and privacy-preserving methods. Our contributions are summarized as follows.

1) Development of a dataset under naturalistic and domestic conditions with three participants, utilizing advanced localization technologies through UWB wearables and low-invasiveness binary sensors. This contribution represents a significant and innovative advancement in the scientific community by integrating a multioccupancy dataset with precise indoor localization, activity labeling, and recognition of basic activities.
2) Creation of an activity recognition model based on localization heatmaps. In addition, our approach reduces the complexity of 2-D spatial location introducing the degree of spatial interaction in activity zones. This transformation translates the absolute localization heatmap into a relative interaction over activities.
3) Evaluation of deep learning models gated recurrent unit (GRU), long-short-term memory (LSTM) networks, convolutional neural network (CNN) + GRU with attention (ATT) mechanism, CNN + ATT + GRU + ATT, transformer with multihead attention, transformer with positional encoding, and collection of metrics from the dataset. A study of performance and complexity is deeply included.

The remainder of this article is structured as follows. Section II presents studies related to HAR in smart multioccupancy environments. Section III details the materials and methods employed, including the integration of nearby sensors, the computation of heatmaps from UWB devices, and the combination of these data with the proposed models. Section IV describes the results obtained from the developed case study, assessing the effectiveness of the methodology in a real-world context, and providing a detailed analysis of the methods used. Finally, Section V discusses conclusions and ongoing work.

II. Related Works

HAR in multioccupant environments is a critical area of research driven by the increasing demand for intelligent systems in smart homes, assisted living facilities, and advanced healthcare settings [14]. These environments often require systems that can accurately track and differentiate activities performed by multiple individuals sharing the same space. The complexity of human interactions, coupled with the limitations of traditional HAR systems, makes it challenging to identify and monitor specific activities reliably. Conventional approaches can struggle to distinguish between individuals and accurately track their activities, leading to potential errors and inaccuracies in behavior analysis [15].

To address these limitations, researchers have explored various sensor technologies and data-fusion methods. Among these, UWB sensors have garnered considerable attention for their ability to provide precise real-time localization in indoor environments. However, UWB systems are sensitive to obstacles and high-density scenarios, which can affect their accuracy. As a result, researchers have begun using heat maps generated from UWB data to enhance spatial representation and activity recognition. Heat maps offer a flexible and intuitive visualization of spatial occupancy, helping to identify patterns and differentiate simultaneous activities in crowded environments. This approach has been validated in several studies, demonstrating its effectiveness in improving activity recognition accuracy and resolving the limitations of UWB alone [16], [17]. For instance, Naser et al. [18] showed that heat maps could effectively visualize occupancy patterns and behavior changes, while Hiley et al. [19] and Yuan et al. [20] highlighted their utility in real-time adjustments of activity recognition algorithms.

The interaction with nearby sensors has emerged as a significant advancement in HAR systems. These sensors, which include wearable devices and embedded environmental sensors, provide additional contextual information that improves the accuracy of activity detection. Integrating nearby sensor data with UWB localization allows for a more detailed and accurate representation of activities, particularly in scenarios involving multiple people performing concurrent activities in close proximity. This interaction improves both spatial and temporal resolution, addressing many of the limitations associated with single-sensor approaches [21], [22]. For example, the work of [23] demonstrated that combining UWB with other sensors could significantly enhance HAR accuracy by providing richer data.

Advanced sensor fusion techniques, including the use of fuzzy logic, have been employed to further improve HAR accuracy. Fuzzy logic offers a robust framework for managing the uncertainty and variability inherent in activity data, which is crucial when activities do not fit neatly into discrete categories. By integrating fuzzy logic with heat maps, researchers can better interpret complex data and differentiate between multiple concurrent activities [24], [25]. In addition, advanced hardware and processing techniques have been proposed to optimize data capture and analysis, leading to significant improvements in activity recognition [26].

The evolution of HAR has also been marked by the development of sophisticated models, such as those that combine CNNs with LSTM networks. CNNs are effective in extracting spatial features from sensory data, such as those obtained from UWB heat maps, while LSTMs are adept at modeling the temporal sequences of activities. This combination offers a robust framework for real-time activity recognition, particularly in dynamic and complex multioccupancy environments [27]. The integration of CNN-LSTM models improves the ability to differentiate and track concurrent activities, addressing the challenges posed by the unpredictable nature of human behavior [28], [29].

Despite these advancements, several gaps remain in the field of HAR. There is a need for further research to optimize these technologies for real-time data processing and analysis, especially in dynamic environments with high activity complexity. Future work should focus on improving the adaptability of HAR systems to real-time changes in human behavior, improving the robustness of learning algorithms, and developing more efficient solutions for data processing and analysis [30], [31]. Addressing these gaps will contribute to more accurate and flexible HAR systems, with broader

Fig. 1. Overview of ambient sensors deployed in the environment for capturing user interactions with household elements, such as door status, temperature, humidity, presence, motion, and power consumption.

applications in various fields, including healthcare, home automation, and beyond [32], [33].

This study aims to address these gaps by integrating advanced sensor fusion techniques of noninvasive devices and adaptable learning models to offer a comprehensive solution for HAR in multioccupancy environments. The proposed approach aims to improve the accuracy of activity monitoring while minimizing deployment complexities and optimizing performance in various settings.

III. Materials and Methods

This section outlines the rationale behind the selection of devices and methodologies employed to achieve flexible yet precise activity recognition. To fully understand this choice, it is essential to distinguish between two categories of activities. The first category includes activities that can be recognized only based on the location of the user and the time spent in specific areas, such as sleeping, resting, leaving the house, and staying in the kitchen. The second category encompasses more complex activities that involve interaction with various elements of the household, such as cooking, taking a shower, using a computer, or using the bathroom. All activities are defined in Section IV-B. To address these needs, we deployed an RTLS based on UWB technology throughout the environment, supplemented by various ambient sensors placed on key household elements. By combining user location heatmaps with activation data from these sensors, our system is capable of accurately discerning which users performed specific actions or activities while preserving privacy and maintaining low deployment costs and high scalability. This is achieved without the need for a large number of beacons or a high density of sensors in the home, as the devices used are commercially available, low-cost, and feature long battery life. This approach not only enhances the granularity of activity recognition, but also ensures that the system can adapt to a variety of contexts and user behaviors.

A. Minimal Ambient Devices for Nearby-Sensor Interaction

To capture user interactions with various household elements, such as the microwave, refrigerator, computer, stove, faucets, and others (detailed in Section IV-A), we used a variety of ambient sensors (Fig. 1). These sensors are connected through a central hub with Home Assistant and transmit data through message queuing telemetry transport (MQTT) [34], allowing seamless integration and communication within the system.

1) *Aqara Door and Window Sensor:* This sensor registers activations using a magnetic contact, sending a 0 when the door is closed and 1 when it is open.
2) *Aqara Temperature and Humidity Sensor:* This device records temperature variations in degrees Celsius and humidity levels as a percentage, providing continuous environmental data.
3) *Aqara Motion Sensor:* This sensor detects movement, transmitting a 0 when motion is not detected and a 1 when motion is present.
4) *Aqara Vibration Sensor:* Equipped with an inertial sensor, this device detects vibrations or movements, sending a 1 upon detection and a 0 if no movement is observed over a period.
5) *Shelly-EM Power Consumption Sensor:* Installed on the main electrical panel, this sensor monitors the total consumption of household power, allowing us to determine when the stove is used (in watts).
6) *TP-Link Smart Plug:* This plug measures the power consumption of small devices, such as a computer, in real time (in watts).

In a formal representation, the ambient sensor stream $\vec{s}$ from a given sensor s is composed of a set of measures $\vec{s} = [s_0, \ldots, s_i]$. A measure s_t is collected at a timestamp t within the timeline $T = \{t_-, t_- + \Delta, \ldots, t, \ldots t_+\}$, determined by the start (t_-) and end (t_+) moments, and a timestep Δ. In a shake of simplicity, in this work, we translate the status of ambient sensors into binary representation $s[t] \in \{0, 1\}$. For example, open-door sensors are straightforwardly related from the state *active, inactive* to 0, 1, respectively, or the power consumption sensor is thresholded at 50 W to denote the use of the computer with binary values, 0 and 1.

B. Computing Localization Heatmap From UWB Wearable Devices

Our proposal focuses on leveraging UWB technology for location-based services through the innovative use of wearable UWB devices in the form of wristwatches. These wearables function as low-invasiveness tags, providing a seamless and user-friendly experience for tracking and locating individuals within a given space. The novelty of our method lies in the use of fingerprinting techniques to achieve improved and precise location estimation in Non-Line-of-Sight, which also reduces the number of anchors required per room. This reduction leads to considerable cost savings and shorter deployment times [35]. In our deployment, we used the Ubitrack RTLS (see Fig. 2), employing only one or two anchors per room and assigning one tag to each resident of the household. The proposed wearable UWB wristwatches bring about a significant step forward in applying UWB technology for efficient and cost-effective location services.

The location estimation for each user is visually represented in the form of a localization heatmap [23], [36]. This visual approach is particularly suitable for aggregating of locations in segmentation processes, such as those proposed in this work, with a timestep of 1 min. The localization heatmap provides an intuitive and comprehensive view of spatial data, enabling

Fig. 2. Overview of UWB devices deployed for indoor location of individuals within the environment. The deployment uses anchors and wearable UWB tags (wristwatches) assigned to each resident.

Fig. 3. Localization heatmap of the three users involved in the dataset where they are located in the kitchen, livingroom, and bedroom (07/25 15:13:00).

more accurate and insightful analysis of location patterns and movements over time.

The occupancy location of the user u is represented in a spatiotemporal matrix $M^u = [m_{txy}]$, such that $m_{txy} = 1$ if the user is at position (x, y) at time $t \in T$, and 0 otherwise. In this context, the plane is divided into $H \times V$ equally sized rectangles, where being at position (x, y) means that the user is located in the xth horizontal and yth vertical rectangle. This discretization of space and time generates a 3-D data stream, whose shape is $T \times H \times V$, which captures the user's location over time.

Moreover, the occupancy location can be summarized straightforwardly in the temporal dimension, namely, through the definition of the localization heatmap $\mathrm{HM}^u = [h_{xy}]$, such that, for user u, the heatmap is computed as follows:

$$\mathrm{HM}_t^u = \frac{1}{|T_t|} \sum_{t' \in T_t} M_{t'}^u$$

where $T_t = \{t_-, t_- + \Delta, \ldots, t\} \subset T$ is a time window of $|T_t|$ timesteps and $M_{t'}^u$ is the submatrix of M^u containing only the fixed instante/index t', i.e., $M_{t'}^u = [m_{t'xy}]$. This process allows for the aggregation of the user's location over a time window, providing a more general view of their location over time.

Fig. 3 provides an example of the localization heatmap of three users from the real-life environment dataset deployed in this work, namely, where they are located in the kitchen, livingroom, and bedroom.

C. From Localization Heatmap to Interaction Degree in Activity Zones

The main issue with absolute localization is the dependency on context. For example, the location where each user sleeps or rests can vary, taking place in one or more areas within a home. This variability complicates the generalization of behaviors in different households, as room locations differ. To address this, we propose calculating the degree of interaction in the activity zones [13] over time, as relative characteristics between users and environments, to improve the recognition of activity.

As we described previously, we can compute the time window-location of inhabitants using a heatmap HM_t^u, for each user u at the timestep t. To translate the localization heatmap into relative zone paths, we introduce the use of activity zones as regions of interest $C_a^u = [c_{xy}]$, where a is an activity, (x, y) corresponds to the xth horizontal and yth vertical rectangle, as defined before, and c_{xy} equals 1 if user u performs activity a in position (x, y) and 0 otherwise. They are represented as activity location masks, which are straightforwardly defined by expert knowledge [13] or by aggregating the localization heatmap for each user and activity over time [36].

From C_a^u, we compute the degree of membership of each user u in the activity zone a [13]. The degree of interaction P_t^{au} with the activity a, at instant t, is computed through aggregation and spatial intersection, using element wise multiplication between the user location heat map HM_t^u and the activity mask C_u^a, i.e.,

$$P_t^{au} = \mathrm{HM}_t^u \otimes C_a^u$$

where $\otimes$ is the component wise multiplication of the matrix elements. If $\max P_t^{au} > 0$ then user u was within the a activity zone.

We note that the degree of interaction in activity zones P_t^{au} overcomes the limitation of dependency on the absolute location of users with relative activity paths, enhancing the applicability and robustness of location-based analyses in diverse living environments. In Fig. 4, we provide an example which translates the absolute location heat maps to the degree of interaction of the relative activity zones.

D. Ensemble of Deep Learning Model Fusing Localization Heatmaps and Nearby-Sensor Interaction in Multioccupancy

In this work, we propose using an ensemble of deep learning models to predict user activities in multioccupancy environments. The model inputs are based on the previous formulation of user interaction in activity zones, where the interaction degree is calculated through the aggregation and spatial intersection of the localization heatmap of user and activity masks, and also the nearby-sensor interaction with the environment by means of ambient sensors.

1) Interaction degree of the user in activity zones, representing where the user has been within the activity zones. It is straightforwardly introduced by the interaction degree P_t^{au} between the user u and each activity a.
2) Interaction degree of other users in activity zones. The inclusion of other users' location information as

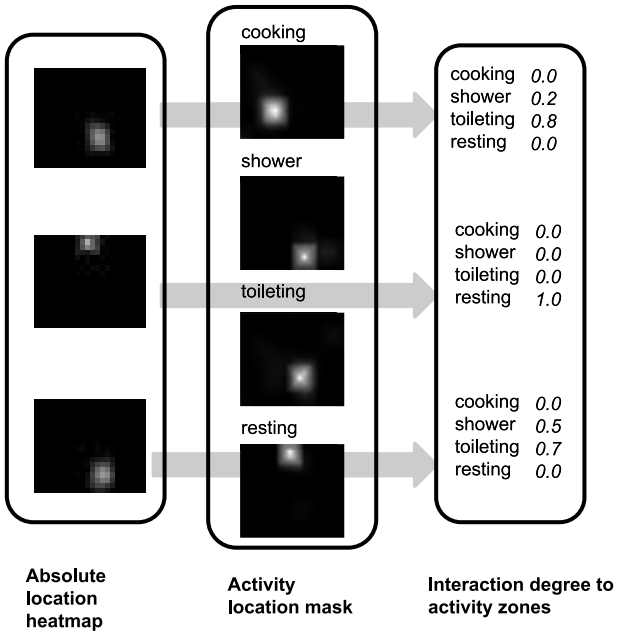

Fig. 4. Interaction degree of relative activity paths from absolute localization heatmaps with an example of four activities (cooking, toileting, showering, and resting).

input to the model is optional. Ultimately, the decision to include or exclude other users' location data should be made based on a careful assessment of the tradeoffs between improved accuracy, architectural feasibility, performance, and privacy implications within the specific application context. This input describes the locations of other users, which is key for understanding crowd dynamics and interactions while preserving privacy without identificating of other users. We compute it by aggregating the interaction degrees of other users $\overline{U} = \{u_1, u_2, \ldots, u_k\} \subset U - \{u\}$ as

$$P_t^{au\overline{U}} = P_t^{au} \oplus P_t^{au_1} \oplus P_t^{au_2} \oplus \cdots \oplus P_t^{au_k}$$

where, as before, $\oplus$ represents the element direct sum of matrices. The sum of interaction degrees of other users provides an aggregated value representing the influence of other inhabitants in a multioccupancy environment. This value is independent of the total number of users in the smart environment.

3) Activity of the environmental sensor, which is represented by the activation s_t of a given sensor s in the timestamp t. They provide the interaction between users and objects in the environment.

The inputs are represented by a sequence under a sliding-window approach [1], where for each timestamp t, we compute a set of timesteps $t^* \in [t - t^*_-, t + t^*_+]$ defined by a temporal window size $T^* = [t^*_-, t^*_+]$.

The output of the model $O_t^{au} \in \{0, 1\}$ is a binary prediction for each activity a and user u on a given time stamp t. For learning purposes, we develop an ensemble approach [37] in which each model is focused on learning a given activity. According to the recommendations of previous works [37], [38], we compose the training dataset for each activity with descriptive, but also conflicting, data based on the activity.

1) Including data where the activity is developed $O_t^{au} = 1$.
2) Including data where the activity is not developed $O_t^{au} = 0$, but prioritizing (with a stochastic selection) when the evaluated user is within the activity zone ($\max P_t^{ua} = 1$). Includes conflicting learning, increasing learning performance, and detailed capabilities [37] in HAR.

Based on this approach, we evaluated several models for predicting activities in multioccupancy environments: [GRU], [LSTM], [CNN + GRU + ATT], [CNN + ATT+ GRU + ATT], [Transformer+multihead], and [Transformer+position encoding].

1) *[GRU] and [LSTM]:* GRU and LSTM sequence models are used as the baseline due to their efficiency and robustness in learning sequence patterns. The GRU and LSTM models consist of three input branches (sensor data, the location of the evaluated user, and the location of the other users), which are reshaped and fed into individual blocks to capture temporal dependencies. The model employs dropout layers throughout to prevent overfitting and is optimized using binary cross-entropy loss and the Adam optimizer.
2) *[CNN + GRU + ATT], [CNN + ATT + GRU + ATT]:* The addition of CNN layers aims to extract local spatial and temporal features before applying the recurrent GRU layer. This convolutional block enhances the ability of the model to identify fine-grained patterns within each sequence before feeding into a GRU layer to capture the temporal dependencies. After these individual blocks, the outputs are concatenated and further processed by additional GRU layers to integrate information across all inputs. We evaluated two variants of ATT locations, incorporating ATT layers at different stages of feature processing: ATT applied to higher level features [CNN + GRU + ATT], and ATT applied to lower level features [CNN + ATT + GRU + ATT]. The ATT mechanism is based on scaled dot-product ATT, which allows the model to dynamically focus on relevant parts of the input sequence. This enhances its ability to capture complex activity patterns, especially in the presence of multiple users and contextual interactions.
3) *[Transformer + Multihead] and [Transformer + Position Encoding]:* Two Transformer-based models have been evaluated. [Transformer + Multihead] uses three parallel Transformer encoders with multihead attention to process three separate input sequences, then concatenates their outputs for binary classification. [Transformer + Position Encoding] encodes three input sequences with positional encoding, processes them with separate multihead attention encoders, and concatenates their reduced representations for binary classification.

These models were evaluated using standard HAR metrics (f1 score, precision, recall, and accuracy).

In Figs. 5 and 6, we detail the components and inputs, the outputs, and the set of proposed models.

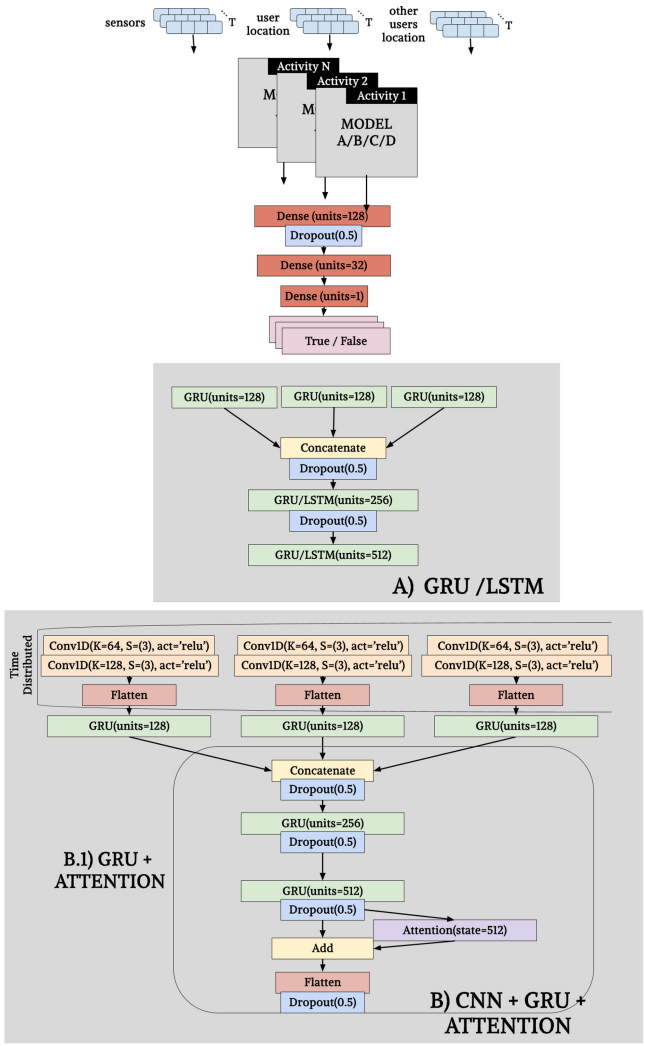

Fig. 5. Scheme of the inputs and the proposed ensemble architecture, and description of the base classification models: A) [GRU] and [LSTM], and B) [CNN + GRU + ATT].

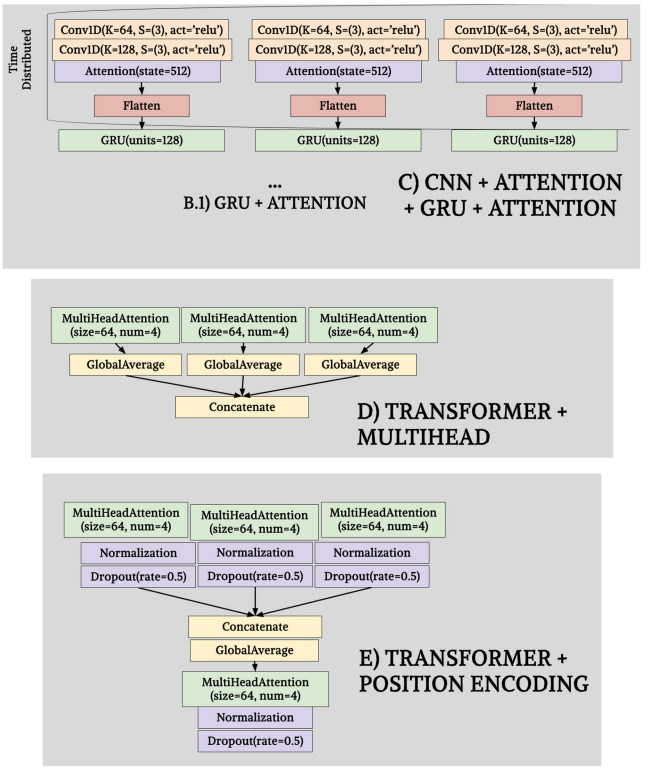

Fig. 6. Description of classification models: C) [CNN + ATT + GRU + ATT], D) [Transformer+multihead], and E) [Transformer+position encoding].

Fig. 7. Room distribution in the real environment: Living Room, Office, Bedroom 1, Kitchen, Bathroom, and Bedroom 2.

IV. Results

In this section, we outline the case study and experimental setup designed in this work to evaluate and demonstrate the application of the methodology in a real-world context.

A. Experimental Setup

The deployment was carried out in a residential apartment of 150 m^2, occupied by three individuals: a woman and two men who live in their own home. The research, carried out as part of the DTS21-00047 project, received approval from the regional ethics committee. Participants were informed about the objectives, methodologies, and legal rights of the study and provided their written informed consent prior to participating.

The apartment has six distinct areas: a living room, office, kitchen, bathroom, and two bedrooms (see Fig. 7). The distribution of sensors throughout the apartment is described in the following for each room.

1) *Living Room:* Two UWB anchors, a energy consumption sensor, two presence sensors located under the television and on the corner table near the sofas.
2) *Kitchen:* Two UWB anchors, six open/closed sensors distributed in refrigerator, microwave, medicine cabinet, cutlery drawer, pan drawer, and plate cabinet; a humidity and temperature sensor in the hood; two vibration sensors on the sink faucet and pantry door; two energy consumption sensors for the cooking robot and the washer/dryer and two presence sensors on the kitchen counter and the wall opposite the stove.
3) *Bathroom:* One UWB anchor, three vibration sensors located on the toilet lid, the sink faucet, and the shower screen; an open/close sensor on the shower screen; a humidity and temperature sensor inside the shower and a presence sensor on the wall opposite the mirror.
4) *Office:* One UWB anchor, a energy consumption sensor for the computer, a presence sensor positioned under the window, and an open/close sensor in the closet.

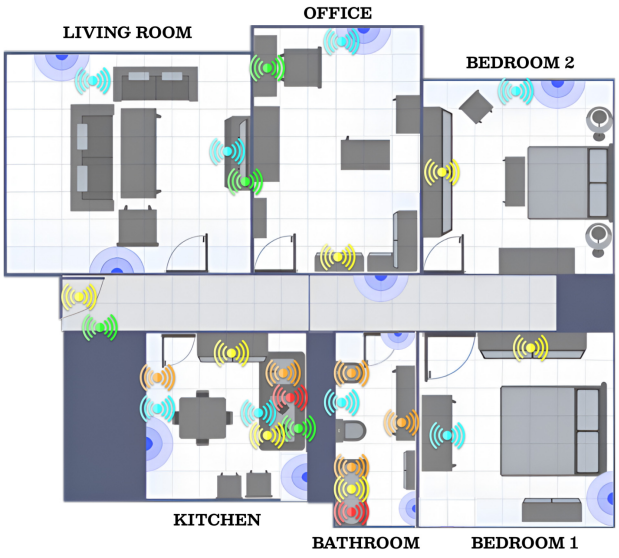

Fig. 8. Map with the distribution of the anchors per room in purple, yellow for open/close sensors, green for energy consumption sensors, blue for presence sensors, red for humidity and temperature sensors, and orange for vibration sensors.

5) *Bedroom 1:* One UWB anchor, a presence sensor placed under the table and shelf and an open/close sensor on the closet.
6) *Bedroom 2:* One UWB anchor, a presence sensor installed on the closet wall and an open/close sensor on the closet.
7) *Main Entrance:* An open/close sensor on the apartment door and a energy consumption sensor on the main panel of the house.
8) *Hallway:* One UWB anchor placed in the middle of the hallway.

Fig. 8 provides a comprehensive layout of the apartment, illustrating the precise location of each sensor and UWB anchor within the different rooms. This map serves to clarify the spatial arrangement of the sensors and anchors, highlighting their strategic placement to maximize the effectiveness of the activity recognition system. In total, 9 UWB anchors, 5 energy consumption sensors, 8 presence sensors, 11 open-close sensors, 2 humidity and temperature sensors, and 5 vibration sensors were deployed throughout the apartment to ensure comprehensive coverage and accurate detection of user activities in all areas. Fig. 9 illustrates the deployment of various sensors in the real environment, demonstrating their minimal invasiveness and seamless integration into everyday settings.

B. Daily Activity Labeling

During the data collection period, which spanned 15 days, users labeled activities using a mobile application and NFC tags (see Fig. 10) placed near the areas where activities typically occurred, facilitating the annotation process. The activities identified and labeled in this study are as follows.

1) *Toileting:* Engaging in activities within the bathroom, such as using the toilet (WC) or interacting with the faucet.

Fig. 9. Deployment of ambient sensors in the real environment, including motion sensors, door/window sensors, vibration sensors, and temperature/humidity sensors. The images demonstrate the unobtrusive placement of the sensors on everyday household items, such as faucets, doors, and appliances, highlighting their minimal invasiveness and seamless integration.

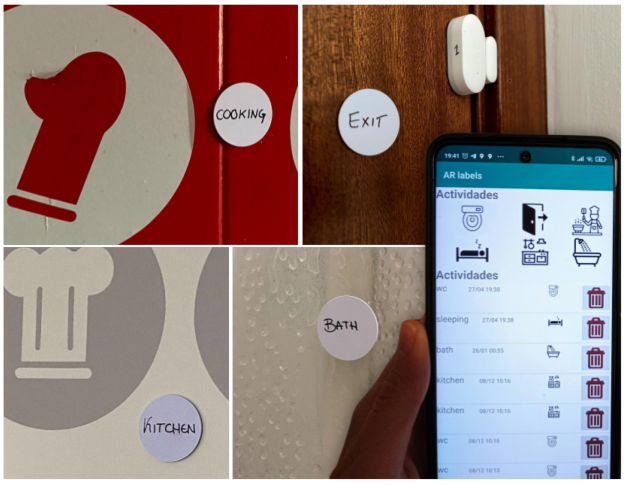

Fig. 10. NFC tagging system used during the data collection period to label daily activities. NFC tags were strategically placed near specific areas where activities commonly occurred, allowing users to annotate their actions using a mobile application quickly.

2) *Resting:* Remaining stationary for an extended period in the living room, particularly in the area near the sofa.
3) *Exit:* Being outside the home.
4) *Cooking:* Preparing food or interacting with key kitchen elements, such as the microwave, stove, pans, plates, fridge, and tap, is detected through sensors placed in each drawer where these items are stored and directly on the microwave, fridge, and tap.
5) *Shower:* Using the shower, which includes interacting with the shower door, being in the shower area, and observing an increase in humidity or temperature levels.
6) *PC:* Using the computer, identified by the power consumption of the device and presence in the designated workspace.
7) *Sleep:* Sleeping in an extended period in the bedroom, specifically in the area near the bed.
8) *Kitchen:* Remaining in the kitchen and interacting with kitchen appliances, such as the microwave, stove, and fridge.

C. Dataset Description

The dataset is designed for the analysis of HAR and interaction patterns within a multioccupancy environment. It combines raw sensor data and processed information that captures the interactions of users with various household elements and their movements throughout the space. Data are collected from a variety of ambient sensors and wearable devices from the UWB, providing comprehensive insights into user behavior, activity patterns, and sensor interactions. The dataset is publicly accessible on the Kaggle platform [39][1] and comprises a wider range of ambient sensors than those used in this study. The full dataset is accessible to the scientific community for research purposes. Additionally, to promote scientific replicability and encourage advances in future work, we have included the source code of models, preprocessing data, and results publicly accessible on GitHub.[2]

The dataset comprises a total of 15 days recorded in a real environment (see Section IV-A) and is structured into two main folders: 1) *raw-har* and 2) *data-har*.

1) Description of the "Raw-Har" Folder: The "raw-har" folder contains three primary subfolders that organize the raw sensor data collected from various devices deployed in the environment. In this dataset, we have indeed included the raw, continuous signals collected from the sensors. Each subfolder contains data corresponding to different types of sensors, which capture various aspects of user interactions, environmental conditions, and positioning data. Below is a detailed description of each subfolder and its contents.

1) *Binary:* This subfolder contains raw data from all binary sensors deployed in the environment, such as door/window sensors, motion sensors, and presence sensors.
   a) *Data Format:* The files are in tab-separated values (TSV) format, where each row contains:
      i) *Timestamp:* Time in seconds since the epoch (UNIX timestamp).
      ii) *Formatted Date:* Human-readable date and time (e.g., `2024-05-03 21:04:10`).
      iii) *Sensor Value (`day.*_presence`, `day.*_apertura`):* The status of the sensor, usually `0` or `1`, indicating the sensor state (e.g., on/off and detected/not detected).
2) *Sensor:* This subfolder contains data from nonbinary environmental sensors, such as temperature, humidity, and power consumption sensors.
   a) *Data Format:* The files are also in TSV format, capturing various measurements.
      i) *Timestamp:* Time in seconds since the epoch.
      ii) *Formatted Date:* Date and time in human-readable format.
      iii) *Sensor Value:* Depending on the sensor type, the value could represent:
         A) *Temperature (`day.temp_*`):* Measured in degrees Celsius.
         B) *Humidity (`day.hum_*`):* Measured as a percentage (%).
         C) *Power Consumption (`day.*_current_consumption`):* Measured in watts (W).
3) *"Ubitrack-uwb":* This subfolder holds data from the UWB localization system, which provides detailed location information of users within the environment.
   a) *Data Format:* The files are also in TSV format:
      i) *Timestamp:* Time in seconds since the epoch.
      ii) *Formatted Date:* Date and time in human-readable format.
      iii) *UWB Values:* The value could represent:
         A) *Source Data (`day.user.sourcedata`):* Files with distance readings between UWB tags (worn by users) and anchors (fixed reference points), measured in centimeters. This data helps determine the exact location of the user within the space.
         B) *RSSI Data (`day.user.rssi`):* Files containing RSSI values indicating signal strength between the user's UWB device and each anchor, providing additional context on the relative positioning of the user.

2) Description of the "Data-Har" Folder: The "data-har" folder contains processed data organized by days, which integrates sensor readings into more interpretable formats, such as heatmaps, interval logs, and aggregated sensor data. This processed data is essential for understanding daily user activities and interactions. Below is a detailed description of each subfolder and its contents.

1) *Heatmaps of User Localization (`other.user.activity`):* This section contains visual representations of user locations during specific activities, providing a spatial distribution of the user within the environment on a specific day.
   a) *Data Format:* The files are in PNG format, representing heatmaps that highlight areas of frequent presence during different activities (e.g., `0001.cooking.png`).
2) *Activity Interval Logs (`INTERVAL.user.all`):* These files contain TSV with start and end times of activities throughout the day, along with activity type and index ranges.
   a) *Data Format:* Each row in the TSV files includes the following.
      i) *Activity Type:* The specific activity being recorded (e.g., `kitchen`, `cooking`, and `sleep`).
      ii) *Index Range:* The range of indices representing the sequence of timestamps related to the activity [e.g., `(1, 5)`].
      iii) *Start Time:* The start date and time of the activity in a human-readable format (e.g., `2024-07-31 02:01:00`).
      iv) *End Time:* The end date and time of the activity (e.g., `2024-07-31 02:05:00`).

[1]https://www.kaggle.com/datasets/aurorapolorodriguez/multi-occupancy-activity-recognition

[2]https://github.com/AuroraPR/Multiuser-HAR-Fusion

b) *Example (`INTERVAL.0001.all`):* Represents the specific intervals when user 0001 was engaged in cooking.

3) *"Object_sensor":* Contains aggregated data from environmental sensors, aggregated by minute.
   a) *Data Format:* The files are in TSV format, including:
      i) *Timestamp:* Human-readable date and time.
      ii) *Sensor Values:* For binary sensors, a value of "1" is applied if the sensor was activated during that minute; otherwise, it is "0." For sensors with values other than "0" or "1," the median of the values obtained during that minute is calculated (e.g., power consumption in watts, humidity percentage, or temperature in degrees Celsius).
4) *"Act":* Includes files that document the processing of activities, including labels, location data, and predictions. The files are in TSV format, capturing details as follows.
   a) *Activity Labels (`LABEL.user.activity`):* Defines ground-truth labels for users and activities per minute.
      i) *Data Format:*
         A) *Timestamp:* Human-readable date and time.
         B) *Activity Value:* Binary value indicating whether activity occurred (`True`) or not (`False`).
   b) *Location Data (`loc.user.activity`):* Represents the spatial interaction degree of a user in a zone associated with an activity.
      i) *Data Format:*
         A) *Timestamp:* Human-readable date and time.
         B) *Interaction Degree:* A numeric value representing the level of interaction or presence within the specified zone.
      ii) *Example:* `loc.0001.cooking.tsv` shows the calculated interaction degree of user 0001 in the cooking zone.
   c) *Predictions (`PREDICTION.user.activity`):* Contains model predictions of user activities, used to compare against labeled data for evaluation.
      i) *Data Format:*
         A) *Timestamp:* Human-readable date and time.
         B) *Prediction Value:* A numeric value between 0 and 1 indicating the prediction by the model of the degree of activity.
5) *"User_agg"*: Contains heatmaps of user localization minute by minute.
   a) *Data Format:* The files are in PNG format, each representing a heatmap showing the user's location within the environment per minute.

D. Evaluation

In this work, we used a one-day cross-validation approach to evaluate the performance of the models. This method assesses the ability of the model to generalize and recognize activities on different days and users, capturing the variability in the data. The evaluation involves using one day as the test set while the remaining days are used for training. This process is repeated so that each day in the dataset serves as a test day once, ensuring a thorough and unbiased evaluation of the model performance under varying temporal conditions.

For each iteration, the users who are active on the selected test day and their activities are identified. Training examples are then created from the data of all other days, generating both positive examples (where specific activities are present) and negative examples (where those activities are absent). This random selection captures sequences of sensor and location data that represent the natural variability in human behavior. The generated training data is then used to train the models, which optimizes their parameters to recognize the activities from the sensor data, adapting to the diverse patterns in the training set. To evaluate the model, test data are prepared using only data from the selected test day. These test sequences are similar to the training data but entirely new to the model, testing its ability to predict activities in an unseen context. The predictions of the model on the test data are compared with the actual labels, and its performance is evaluated using standard metrics, such as accuracy, precision, recall, and F1-score.

Several models have been evaluated: [GRU], [LSTM], [CNN + GRU + ATT], [CNN + ATT + GRU + ATT], [Transformer+multihead], and [Transformer+position encoding], which are described in detail in Section III-D. In addition, two temporal windows are proposed for including the performance of real-time and delayed temporal windows [40]: *40 + 0*) a temporal window of 45 past minutes, which configures a sliding sequence of 45 elements for each input, *30 + 15*) a delayed temporal window of 30 past minutes and 15 ongoing minutes, which also configure a sliding sequence of 45 elements for each input. We note that the configuration *40 + 0* provides real-time recognition performance and *30 + 15* a delayed 15-min performance in activity recognition.

In addition, a parameter N is used to define a tolerance range in the evaluation of the prediction metrics, allowing matches between the prediction and the ground truth not only at the exact index but also within the N positions forward and backward. This introduces flexibility in the evaluation by capturing nearby matches that reflect the prediction quality when slight temporal misalignments are acceptable. In our analysis, the evaluation was carried out with tolerances $N = 0$ and $N = 3$ min. The time margin N is relevant for removing prediction delays that are unavoidable in some cases of user movement, offering dispersed trajectories, or sensor activations that are triggered or deactivated with delay. For example, it is especially sensitive in the "shower" activity where the activation of the humidity sensor takes several minutes to activate, so the prediction of the activity may take a few minutes, even if it is performed correctly.

We describe the results of evaluating the proposed models with inputs of ambient sensor, user location, and collective location in the next tables.

1) Tables I and II describe the performance of [GRU] with [30 + 15] and [40 + 5] min of time windows, respectively.

TABLE I
PERFORMANCE OF THE GRU MODEL WITH A 30 + 15 TIME WINDOW (30 MIN IN THE PAST AND 15 MIN INTO THE FUTURE). METRICS INCLUDE ACCURACY (ACC), PRECISION (PRE), RECALL (REC), AND F1 SCORE (F1)

GRU 30+15				
Activity	ACC N=0 - N=3	PRE N=0 - N=3	REC N=0 - N=3	F1 N=0 - N=3
toileting	*0,97 - 0,98*	*0,83 - 0,87*	*0,95 - 0,94*	*0,87 - 0,88*
resting	*0,97 - 0,97*	*0,96 - 0,97*	*0,98 - 0,98*	*0,97 - 0,97*
exit	*0,99 - 1,00*	*0,80 - 0,94*	*0,94 - 0,99*	*0,86 - 0,96*
cooking	*1,00 - 1,00*	*0,80 - 0,88*	*0,89 - 0,96*	*0,79 - 0,90*
shower	*1,00 - 1,00*	*0,90 - 0,95*	*0,99 - 0,99*	*0,92 - 0,96*
pc	*0,98 - 0,98*	*0,99 - 1,00*	*0,99 - 1,00*	*0,99 - 1,00*
sleep	*0,97 - 0,98*	*0,68 - 0,78*	*0,94 - 0,98*	*0,76 - 0,84*
kitchen	*0,99 - 1,00*	*0,79 - 0,90*	*0,97 - 1,00*	*0,87 - 0,94*
TOTAL	***0,98 - 0,99***	***0,85 - 0,91***	***0,96 - 0,98***	***0,88 - 0,93***

TABLE II
PERFORMANCE OF THE GRU MODEL WITH A 45 + 0 TIME WINDOW (45 MIN IN THE PAST ONLY). METRICS INCLUDE ACCURACY (ACC), PRECISION (PRE), RECALL (REC), AND F1 SCORE (F1)

GRU 45+0				
Activity	ACC N=0 - N=3	PRE N=0 - N=3	REC N=0 - N=3	F1 N=0 - N=3
toileting	*0,96 - 0,97*	*0,72 - 0,80*	*0,86 - 0,95*	*0,78 - 0,86*
resting	*0,96 - 0,97*	*0,93 - 0,94*	*0,95 - 0,96*	*0,93 - 0,95*
exit	*0,95 - 0,97*	*0,31 - 0,55*	*0,67 - 0,97*	*0,41 - 0,69*
cooking	*0,99 - 1,00*	*0,68 - 0,80*	*0,86 - 0,97*	*0,71 - 0,84*
shower	*0,99 - 1,00*	*0,90 - 0,93*	*0,96 - 0,99*	*0,92 - 0,94*
pc	*0,97 - 0,98*	*0,97 - 0,99*	*0,98 - 0,99*	*0,98 - 0,99*
sleep	*0,95 - 0,97*	*0,60 - 0,69*	*0,83 - 0,94*	*0,67 - 0,76*
kitchen	*0,95 - 0,97*	*0,27 - 0,51*	*0,66 - 0,98*	*0,37 - 0,66*
TOTAL	***0,97 - 0,98***	***0,68 - 0,78***	***0,85 - 0,97***	***0,73 - 0,84***

TABLE III
PERFORMANCE OF THE LSTM MODEL WITH A 30 + 15 TIME WINDOW (30 MIN IN THE PAST AND 15 MIN INTO THE FUTURE). METRICS INCLUDE ACCURACY (ACC), PRECISION (PRE), RECALL (REC), AND F1 SCORE (F1)

LSTM 30+15				
Activity	ACC N=0 - N=3	PRE N=0 - N=3	REC N=0 - N=3	F1 N=0 - N=3
toileting	*0,96 - 0,97*	*0,77 - 0,82*	*0,95 - 0,96*	*0,82 - 0,86*
resting	*0,98 - 0,98*	*0,99 - 0,99*	*0,98 - 0,99*	*0,98 - 0,99*
exit	*0,99 - 1,00*	*0,71 - 0,88*	*0,96 - 1,00*	*0,81 - 0,93*
cooking	*1,00 - 1,00*	*0,67 - 0,75*	*0,89 - 0,97*	*0,70 - 0,81*
shower	*0,99 - 0,99*	*0,92 - 0,93*	*0,94 - 0,95*	*0,91 - 0,93*
pc	*0,98 - 0,98*	*0,99 - 1,00*	*0,99 - 0,99*	*0,99 - 1,00*
sleep	*0,97 - 0,98*	*0,70 - 0,81*	*0,92 - 0,95*	*0,77 - 0,86*
kitchen	*0,99 - 1,00*	*0,79 - 0,90*	*0,97 - 1,00*	*0,86 - 0,94*
TOTAL	***0,98 - 0,99***	***0,82 - 0,88***	***0,95 - 0,98***	***0,86 - 0,91***

TABLE IV
PERFORMANCE OF THE LSTM MODEL WITH A 45 + 0 TIME WINDOW (30 MIN IN THE PAST AND 15 MIN INTO THE FUTURE). METRICS INCLUDE ACCURACY (ACC), PRECISION (PRE), RECALL (REC), AND F1 SCORE (F1)

LSTM 45+0				
Activity	ACC N=0 - N=3	PRE N=0 - N=3	REC N=0 - N=3	F1 N=0 - N=3
toileting	*0,95 - 0,96*	*0,70 - 0,76*	*0,86 - 0,94*	*0,74 - 0,81*
resting	*0,97 - 0,97*	*0,97 - 0,98*	*0,97 - 0,98*	*0,97 - 0,98*
exit	*0,94 - 0,97*	*0,28 - 0,53*	*0,67 - 0,99*	*0,38 - 0,68*
cooking	*0,99 - 1,00*	*0,72 - 0,77*	*0,85 - 0,96*	*0,75 - 0,83*
shower	*0,99 - 0,99*	*0,88 - 0,93*	*0,98 - 0,98*	*0,90 - 0,93*
pc	*0,97 - 0,97*	*0,96 - 0,98*	*0,99 - 0,99*	*0,97 - 0,99*
sleep	*0,94 - 0,97*	*0,63 - 0,70*	*0,86 - 0,93*	*0,71 - 0,77*
kitchen	*0,95 - 0,97*	*0,28 - 0,52*	*0,65 - 0,99*	*0,38 - 0,66*
TOTAL	***0,96 - 0,98***	***0,68 - 0,77***	***0,85 - 0,97***	***0,73 - 0,83***

TABLE V
PERFORMANCE OF THE CNN+GRU+ATT MODEL WITH A 30 + 15 TIME WINDOW (30 MIN IN THE PAST AND 15 MIN INTO THE FUTURE). METRICS INCLUDE ACCURACY (ACC), PRECISION (PRE), RECALL (REC), AND F1 SCORE (F1)

CNN+GRU+ATT 30+15				
Activity	ACC N=0 - N=3	PRE N=0 - N=3	REC N=0 - N=3	F1 N=0 - N=3
toileting	*0,98 - 0,98*	*0,84 - 0,89*	*0,93 - 0,98*	*0,89 - 0,93*
resting	*0,98 - 0,98*	*0,99 - 0,99*	*0,98 - 0,99*	*0,98 - 0,99*
exit	*0,99 - 1,00*	*0,84 - 0,93*	*0,95 - 0,99*	*0,88 - 0,96*
cooking	*1,00 - 1,00*	*0,82 - 0,93*	*0,91 - 0,96*	*0,82 - 0,94*
shower	*1,00 - 1,00*	*0,91 - 0,94*	*0,95 - 0,97*	*0,90 - 0,94*
pc	*0,98 - 0,98*	*0,99 - 0,99*	*0,99 - 1,00*	*0,99 - 1,00*
sleep	*0,98 - 0,99*	*0,76 - 0,84*	*0,91 - 0,96*	*0,82 - 0,88*
kitchen	*1,00 - 1,00*	*0,84 - 0,92*	*0,97 - 1,00*	*0,90 - 0,95*
TOTAL	***0,99 - 0,99***	***0,88 - 0,93***	***0,95 - 0,98***	***0,90 - 0,95***

TABLE VI
PERFORMANCE OF THE CNN+GRU+ATT MODEL WITH A 45 + 0 TIME WINDOW (45 MIN IN THE PAST ONLY). METRICS INCLUDE ACCURACY (ACC), PRECISION (PRE), RECALL (REC), AND F1 SCORE (F1)

CNN+GRU+ATT 45+0				
Activity	ACC N=0 - N=3	PRE N=0 - N=3	REC N=0 - N=3	F1 N=0 - N=3
toileting	*0,95 - 0,97*	*0,74 - 0,77*	*0,91 - 0,97*	*0,79 - 0,83*
resting	*0,97 - 0,98*	*0,96 - 0,97*	*0,98 - 0,99*	*0,96 - 0,98*
exit	*0,96 - 0,98*	*0,36 - 0,64*	*0,62 - 0,97*	*0,44 - 0,76*
cooking	*1,00 - 1,00*	*0,63 - 0,77*	*0,80 - 0,98*	*0,67 - 0,83*
shower	*1,00 - 1,00*	*0,88 - 0,93*	*0,97 - 0,99*	*0,90 - 0,94*
pc	*0,97 - 0,98*	*0,97 - 0,99*	*0,99 - 1,00*	*0,98 - 0,99*
sleep	*0,95 - 0,97*	*0,59 - 0,70*	*0,84 - 0,97*	*0,67 - 0,79*
kitchen	*0,96 - 0,98*	*0,36 - 0,59*	*0,60 - 0,99*	*0,44 - 0,73*
TOTAL	***0,97 - 0,98***	***0,69 - 0,80***	***0,84 - 0,98***	***0,74 - 0,86***

2) Tables III and IV describe the performance of the model [LSTM] with [30 + 15] and [40 + 5] min of time windows, respectively.
3) Tables V and VI describe the performance of the model [CNN+GRU+ATT] with [30 + 15] and [40 + 5] min of time windows, respectively.
4) Tables VII and VIII describe the performance of the model [CNN+ATT+GRU+ATT] with [30 + 15] and [40 + 5] min of time windows, respectively.
5) Tables IX and X describe the performance of the model [Transformer+multihead] with [30 + 15] and [40 + 5] min of time windows, respectively.
6) Tables XI and XII describe the performance of the model [Transformer+position encoding] with [30 + 15] and [40 + 5] min of time windows, respectively.

1) Evaluation Using Uniquely Individual Location: This section presents a targeted evaluation to assess the impact of integrating the individual location of the user being evaluated, without utilizing the aggregated location data of other users. For this purpose, the model with the best performance has been employed, as analyzed in the previous section.

This evaluation is crucial to account for the influence of other individuals on the recognition of patterns that affect the classification of each individual. To achieve this, we have removed the input to the model that captures the location patterns of other users, leaving only the individual location and sensor activation data. Furthermore, it is important because it allows us to understand the potential loss of precision and recall in configurations where, due to privacy concerns, only the location of the user being evaluated can be considered.

TABLE VII
PERFORMANCE OF THE CNN+ATT+GRU+ATT MODEL WITH A 30 + 15 TIME WINDOW (30 MIN IN THE PAST AND 15 MIN INTO THE FUTURE). METRICS INCLUDE ACCURACY (ACC), PRECISION (PRE), RECALL (REC), AND F1 SCORE (F1)

CNN+ATT+GRU+ATT 30+15				
Activity	ACC N=0 - N=3	PRE N=0 - N=3	REC N=0 - N=3	F1 N=0 - N=3
toileting	*0,97 - 0,98*	*0,84 - 0,84*	*0,93 - 0,95*	*0,86 - 0,87*
resting	*0,98 - 0,98*	*0,98 - 0,98*	*0,98 - 0,99*	*0,98 - 0,98*
exit	*0,99 - 1,00*	*0,81 - 0,92*	*0,94 - 0,99*	*0,86 - 0,96*
cooking	*1,00 - 1,00*	*0,69 - 0,83*	*0,95 - 0,98*	*0,76 - 0,88*
shower	*1,00 - 1,00*	*0,93 - 0,92*	*0,99 - 0,95*	*0,94 - 0,92*
pc	*0,98 - 0,98*	*0,99 - 1,00*	*0,99 - 1,00*	*0,99 - 1,00*
sleep	*0,98 - 0,99*	*0,73 - 0,82*	*0,92 - 0,96*	*0,79 - 0,86*
kitchen	*0,99 - 1,00*	*0,79 - 0,90*	*0,97 - 1,00*	*0,87 - 0,94*
TOTAL	***0,98 - 0,99***	***0,84 - 0,90***	***0,96 - 0,98***	***0,88 - 0,93***

TABLE VIII
PERFORMANCE OF THE CNN+ATT+GRU+ATT MODEL WITH A 45 + 0 TIME WINDOW (30 MIN IN THE PAST AND 15 MIN INTO THE FUTURE). METRICS INCLUDE ACCURACY (ACC), PRECISION (PRE), RECALL (REC), AND F1 SCORE (F1)

CNN+ATT+GRU+ATT 45+0				
Activity	ACC N=0 - N=3	PRE N=0 - N=3	REC N=0 - N=3	F1 N=0 - N=3
toileting	*0,96 - 0,97*	*0,74 - 0,80*	*0,86 - 0,95*	*0,77 - 0,84*
resting	*0,97 - 0,97*	*0,95 - 0,97*	*0,97 - 0,99*	*0,96 - 0,97*
exit	*0,96 - 0,98*	*0,35 - 0,63*	*0,59 - 0,97*	*0,43 - 0,75*
cooking	*0,99 - 1,00*	*0,63 - 0,77*	*0,87 - 0,98*	*0,67 - 0,84*
shower	*0,99 - 1,00*	*0,87 - 0,89*	*0,96 - 0,97*	*0,88 - 0,91*
pc	*0,97 - 0,98*	*0,98 - 0,99*	*0,99 - 1,00*	*0,98 - 0,99*
sleep	*0,95 - 0,97*	*0,60 - 0,66*	*0,80 - 0,95*	*0,67 - 0,74*
kitchen	*0,96 - 0,98*	*0,34 - 0,60*	*0,62 - 0,99*	*0,43 - 0,74*
TOTAL	***0,97 - 0,98***	***0,68 - 0,79***	***0,83 - 0,97***	***0,72 - 0,85***

TABLE IX
PERFORMANCE OF THE TRANSFORMER+MULTIHEAD MODEL WITH A 30 + 15 TIME WINDOW (30 MIN IN THE PAST AND 15 MIN INTO THE FUTURE). METRICS INCLUDE ACCURACY (ACC), PRECISION (PRE), RECALL (REC), AND F1 SCORE (F1)

Transformer+multihead 30+15				
Activity	ACC N=0 - N=3	PRE N=0 - N=3	REC N=0 - N=3	F1 N=0 - N=3
toileting	*0,92 - 0,93*	*0,66 - 0,66*	*0,84 - 0,92*	*0,69 - 0,72*
resting	*0,94 - 0,95*	*0,92 - 0,93*	*0,88 - 0,92*	*0,89 - 0,92*
exit	*0,97 - 0,98*	*0,44 - 0,64*	*0,87 - 0,97*	*0,56 - 0,76*
cooking	*1,00 - 1,00*	*0,75 - 0,79*	*0,87 - 0,98*	*0,77 - 0,85*
shower	*0,98 - 0,99*	*0,94 - 0,91*	*0,95 - 0,96*	*0,93 - 0,91*
pc	*0,96 - 0,97*	*0,98 - 0,99*	*0,96 - 0,98*	*0,97 - 0,98*
sleep	*0,96 - 0,98*	*0,63 - 0,73*	*0,85 - 0,98*	*0,70 - 0,81*
kitchen	*0,97 - 0,98*	*0,45 - 0,63*	*0,90 - 0,99*	*0,59 - 0,76*
TOTAL	***0,96 - 0,97***	***0,72 - 0,78***	***0,89 - 0,96***	***0,76 - 0,84***

We describe the results of evaluating the proposed models with inputs of ambient sensor and user location without the collective location of other users in the next tables.

1) Tables XIII and XIV describe the performance of [GRU] with [30 + 15] and [40 + 5] min of time windows, respectively.
2) Tables XV and XVI describe the performance of the model [LSTM] with [30 + 15] and [40 + 5] min of time windows, respectively.
3) Tables XVII and XVIII describe the performance of the model [CNN+GRU+ATT] with [30 + 15] and [40 + 5] min of time windows, respectively.

TABLE X
PERFORMANCE OF THE TRANSFORMER+MULTIHEAD MODEL WITH A 45 + 0 TIME WINDOW (30 MIN IN THE PAST AND 15 MIN INTO THE FUTURE). METRICS INCLUDE ACCURACY (ACC), PRECISION (PRE), RECALL (REC), AND F1 SCORE (F1)

Transformer+multihead 45+0				
Activity	ACC N=0 - N=3	PRE N=0 - N=3	REC N=0 - N=3	F1 N=0 - N=3
toileting	*0,89 - 0,90*	*0,53 - 0,57*	*0,75 - 0,89*	*0,58 - 0,64*
resting	*0,93 - 0,93*	*0,87 - 0,89*	*0,88 - 0,92*	*0,87 - 0,90*
exit	*0,92 - 0,95*	*0,21 - 0,45*	*0,72 - 0,99*	*0,31 - 0,60*
cooking	*0,99 - 1,00*	*0,78 - 0,70*	*0,90 - 0,95*	*0,81 - 0,76*
shower	*0,98 - 0,98*	*0,92 - 0,88*	*0,93 - 0,96*	*0,92 - 0,90*
pc	*0,95 - 0,96*	*0,95 - 0,96*	*0,94 - 0,97*	*0,94 - 0,96*
sleep	*0,93 - 0,95*	*0,52 - 0,56*	*0,86 - 0,96*	*0,61 - 0,66*
kitchen	*0,93 - 0,95*	*0,19 - 0,41*	*0,65 - 0,97*	*0,29 - 0,56*
TOTAL	***0,94 - 0,95***	***0,62 - 0,68***	***0,83 - 0,95***	***0,66 - 0,75***

TABLE XI
PERFORMANCE OF THE TRANSFORMER+POSITION ENCODING MODEL WITH A 30 + 15 TIME WINDOW (30 MIN IN THE PAST AND 15 MIN INTO THE FUTURE). METRICS INCLUDE ACCURACY (ACC), PRECISION (PRE), RECALL (REC), AND F1 SCORE (F1)

Transformer+position encoding 30+15				
Activity	ACC N0-N3	PRE N0-N3	REC N0-N3	F1 N0-N3
toileting	*0,96 - 0,97*	*0,76 - 0,79*	*0,92 - 0,96*	*0,81 - 0,84*
resting	*0,98 - 0,98*	*0,99 - 0,99*	*0,96 - 0,98*	*0,97 - 0,98*
exit	*0,99 - 1,00*	*0,71 - 0,89*	*0,94 - 1,00*	*0,80 - 0,94*
cooking	*1,00 - 1,00*	*0,73 - 0,86*	*0,88 - 0,97*	*0,73 - 0,88*
shower	*1,00 - 1,00*	*0,92 - 0,94*	*0,98 - 1,00*	*0,93 - 0,96*
pc	*0,98 - 0,98*	*0,99 - 1,00*	*0,99 - 1,00*	*0,99 - 1,00*
sleep	*0,97 - 0,98*	*0,66 - 0,77*	*0,94 - 0,98*	*0,74 - 0,84*
kitchen	*0,99 - 1,00*	*0,77 - 0,89*	*0,95 - 1,00*	*0,84 - 0,93*
TOTAL	***0,98 - 0,99***	***0,82 - 0,89***	***0,95 - 0,98***	***0,85 - 0,92***

TABLE XII
PERFORMANCE OF THE TRANSFORMER+POSITION ENCODING MODEL WITH A 45 + 0 TIME WINDOW (45 MIN IN THE PAST ONLY). METRICS INCLUDE ACCURACY (ACC), PRECISION (PRE), RECALL (REC), AND F1 SCORE (F1)

Transformer+position encoding 45+0				
Activity	ACC N0-N3	PRE N0-N3	REC N0-N3	F1 N0-N3
toileting	*0,95 - 0,96*	*0,72 - 0,75*	*0,89 - 0,95*	*0,77 - 0,81*
resting	*0,97 - 0,98*	*0,97 - 0,98*	*0,97 - 0,99*	*0,97 - 0,98*
exit	*0,95 - 0,97*	*0,30 - 0,59*	*0,72 - 0,99*	*0,42 - 0,73*
cooking	*0,99 - 1,00*	*0,60 - 0,76*	*0,81 - 0,97*	*0,63 - 0,82*
shower	*0,99 - 0,99*	*0,85 - 0,90*	*0,97 - 0,99*	*0,88 - 0,92*
pc	*0,97 - 0,98*	*0,97 - 0,98*	*0,99 - 1,00*	*0,98 - 0,99*
sleep	*0,94 - 0,96*	*0,56 - 0,59*	*0,90 - 0,97*	*0,65 - 0,69*
kitchen	*0,93 - 0,95*	*0,22 - 0,45*	*0,77 - 0,99*	*0,33 - 0,61*
TOTAL	***0,96 - 0,97***	***0,65 - 0,75***	***0,88 - 0,98***	***0,70 - 0,82***

4) Tables XIX and XX describe the performance of the model [CNN+ATT+GRU+ATT] with [30 + 15] and [40 + 5] min of time windows, respectively.
5) Tables XXI and XXII describe the performance of the model [Transformer+multihead] with [30 + 15] and [40 + 5] min of time windows, respectively.
6) Tables XXIII and XXIV describe the performance of the model [Transformer+position encoding] with [30 + 15] and [40 + 5] min of time windows, respectively.

E. Discussion

Comparative analysis of CNN + GRU + ATT and GRU models in the 30 + 15 and 45 + 0 time window configurations reveals significant differences in the performance and classification accuracy of various activities. The results obtained,

TABLE XIII
PERFORMANCE OF THE GRU (WITHOUT COLLECTIVE LOCATIONS) MODEL WITH A 30 + 15 TIME WINDOW (30 MIN IN THE PAST AND 15 MIN INTO THE FUTURE). METRICS INCLUDE ACCURACY (ACC), PRECISION (PRE), RECALL (REC), AND F1 SCORE (F1)

GRU (without collective locations) 30+15				
Activity	ACC N=0 - N=3	PRE N=0 - N=3	REC N=0 - N=3	F1 N=0 - N=3
toileting	*0,96 - 0,97*	*0,79 - 0,83*	*0,92 - 0,95*	*0,83 - 0,87*
resting	*0,98 - 0,98*	*0,98 - 0,99*	*0,98 - 0,99*	*0,98 - 0,99*
exit	*0,99 - 1,00*	*0,80 - 0,92*	*0,93 - 0,99*	*0,85 - 0,95*
cooking	*1,00 - 1,00*	*0,75 - 0,89*	*0,88 - 0,97*	*0,77 - 0,91*
shower	*1,00 - 1,00*	*0,92 - 0,95*	*1,00 - 0,98*	*0,93 - 0,95*
pc	*0,98 - 0,98*	*0,99 - 1,00*	*0,99 - 1,00*	*0,99 - 1,00*
sleep	*0,97 - 0,98*	*0,66 - 0,77*	*0,93 - 0,97*	*0,73 - 0,82*
kitchen	*0,99 - 1,00*	*0,80 - 0,90*	*0,97 - 1,00*	*0,86 - 0,94*
TOTAL	***0,98 - 0,99***	***0,84 - 0,90***	***0,95 - 0,98***	***0,87 - 0,93***

TABLE XIV
PERFORMANCE OF THE GRU (WITHOUT COLLECTIVE LOCATIONS) MODEL WITH A 45 + 0 TIME WINDOW (45 MIN IN THE PAST ONLY). METRICS INCLUDE ACCURACY (ACC), PRECISION (PRE), RECALL (REC), AND F1 SCORE (F1)

GRU (without collective locations) 45+0				
Activity	ACC N=0 - N=3	PRE N=0 - N=3	REC N=0 - N=3	F1 N=0 - N=3
toileting	*0,95 - 0,97*	*0,73 - 0,79*	*0,86 - 0,95*	*0,77 - 0,84*
resting	*0,97 - 0,97*	*0,95 - 0,97*	*0,98 - 0,99*	*0,96 - 0,98*
exit	*0,95 - 0,97*	*0,31 - 0,56*	*0,66 - 0,97*	*0,40 - 0,69*
cooking	*1,00 - 1,00*	*0,67 - 0,80*	*0,87 - 0,96*	*0,70 - 0,85*
shower	*0,99 - 0,99*	*0,86 - 0,91*	*0,96 - 1,00*	*0,88 - 0,94*
pc	*0,97 - 0,98*	*0,97 - 0,98*	*0,99 - 1,00*	*0,98 - 0,99*
sleep	*0,94 - 0,97*	*0,61 - 0,68*	*0,85 - 0,95*	*0,68 - 0,75*
kitchen	*0,94 - 0,97*	*0,26 - 0,48*	*0,67 - 0,98*	*0,36 - 0,64*
TOTAL	***0,96 - 0,98***	***0,67 - 0,77***	***0,86 - 0,98***	***0,72 - 0,83***

TABLE XV
PERFORMANCE OF THE LSTM (WITHOUT COLLECTIVE LOCATIONS) MODEL WITH A 30 + 15 TIME WINDOW (30 MIN IN THE PAST AND 15 MIN INTO THE FUTURE). METRICS INCLUDE ACCURACY (ACC), PRECISION (PRE), RECALL (REC), AND F1 SCORE (F1)

LSTM (without collective locations) 30+15				
Activity	ACC N=0 - N=3	PRE N=0 - N=3	REC N=0 - N=3	F1 N=0 - N=3
toileting	*0,96 - 0,97*	*0,81 - 0,83*	*0,93 - 0,94*	*0,85 - 0,86*
resting	*0,98 - 0,98*	*0,98 - 0,99*	*0,98 - 0,99*	*0,98 - 0,99*
exit	*0,99 - 1,00*	*0,79 - 0,92*	*0,95 - 0,99*	*0,86 - 0,95*
cooking	*1,00 - 1,00*	*0,75 - 0,81*	*0,88 - 0,96*	*0,75 - 0,86*
shower	*0,99 - 0,99*	*0,90 - 0,93*	*0,99 - 1,00*	*0,92 - 0,95*
pc	*0,98 - 0,98*	*0,99 - 1,00*	*0,99 - 1,00*	*0,99 - 1,00*
sleep	*0,97 - 0,98*	*0,74 - 0,79*	*0,92 - 0,96*	*0,79 - 0,84*
kitchen	*0,99 - 1,00*	*0,72 - 0,85*	*0,97 - 1,00*	*0,82 - 0,91*
TOTAL	***0,98 - 0,99***	***0,84 - 0,89***	***0,95 - 0,98***	***0,87 - 0,92***

TABLE XVI
PERFORMANCE OF THE LSTM (WITHOUT COLLECTIVE LOCATIONS) MODEL WITH A 45 + 0 TIME WINDOW (30 MIN IN THE PAST AND 15 MIN INTO THE FUTURE). METRICS INCLUDE ACCURACY (ACC), PRECISION (PRE), RECALL (REC), AND F1 SCORE (F1)

LSTM (without collective locations) 45+0				
Activity	ACC N=0 - N=3	PRE N=0 - N=3	REC N=0 - N=3	F1 N=0 - N=3
toileting	*0,96 - 0,97*	*0,72 - 0,79*	*0,87 - 0,94*	*0,78 - 0,84*
resting	*0,97 - 0,98*	*0,96 - 0,98*	*0,97 - 0,98*	*0,96 - 0,98*
exit	*0,95 - 0,97*	*0,28 - 0,54*	*0,68 - 0,98*	*0,39 - 0,68*
cooking	*0,99 - 1,00*	*0,65 - 0,76*	*0,87 - 0,96*	*0,69 - 0,82*
shower	*0,99 - 1,00*	*0,90 - 0,94*	*0,97 - 0,98*	*0,90 - 0,94*
pc	*0,97 - 0,98*	*0,97 - 0,98*	*0,99 - 1,00*	*0,98 - 0,99*
sleep	*0,94 - 0,97*	*0,53 - 0,63*	*0,82 - 0,94*	*0,61 - 0,72*
kitchen	*0,95 - 0,97*	*0,28 - 0,50*	*0,65 - 0,99*	*0,38 - 0,66*
TOTAL	***0,96 - 0,98***	***0,66 - 0,77***	***0,85 - 0,97***	***0,71 - 0,83***

TABLE XVII
PERFORMANCE OF THE CNN-GRU-ATT (WITHOUT COLLECTIVE LOCATIONS) MODEL WITH A 30 + 15 TIME WINDOW (30 MIN IN THE PAST AND 15 MIN INTO THE FUTURE). METRICS INCLUDE ACCURACY (ACC), PRECISION (PRE), RECALL (REC), AND F1 SCORE (F1)

CNN-GRU-ATT (without collective locations) 30+15				
Activity	ACC N=0 - N=3	PRE N=0 - N=3	REC N=0 - N=3	F1 N=0 - N=3
toileting	*0,97 - 0,97*	*0,79 - 0,83*	*0,90 - 0,96*	*0,81 - 0,86*
resting	*0,97 - 0,98*	*0,98 - 0,98*	*0,98 - 0,99*	*0,98 - 0,98*
exit	*0,99 - 1,00*	*0,81 - 0,93*	*0,95 - 1,00*	*0,87 - 0,96*
cooking	*1,00 - 1,00*	*0,80 - 0,92*	*0,95 - 0,99*	*0,83 - 0,95*
shower	*1,00 - 1,00*	*0,90 - 0,94*	*0,96 - 0,98*	*0,91 - 0,95*
pc	*0,98 - 0,98*	*0,99 - 1,00*	*0,99 - 1,00*	*0,99 - 1,00*
sleep	*0,98 - 0,99*	*0,75 - 0,83*	*0,95 - 0,98*	*0,81 - 0,88*
kitchen	*1,00 - 1,00*	*0,86 - 0,94*	*0,97 - 1,00*	*0,91 - 0,97*
TOTAL	***0,99 - 0,99***	***0,86 - 0,92***	***0,96 - 0,99***	***0,89 - 0,94***

TABLE XVIII
PERFORMANCE OF THE CNN-GRU-ATT (WITHOUT COLLECTIVE LOCATIONS) MODEL WITH A 45 + 0 TIME WINDOW (45 MIN IN THE PAST ONLY). METRICS INCLUDE ACCURACY (ACC), PRECISION (PRE), RECALL (REC), AND F1 SCORE (F1)

CNN-GRU-ATT (without collective locations) 45+0				
Activity	ACC N=0 - N=3	PRE N=0 - N=3	REC N=0 - N=3	F1 N=0 - N=3
toileting	*0,96 - 0,97*	*0,73 - 0,79*	*0,89 - 0,96*	*0,77 - 0,84*
resting	*0,97 - 0,98*	*0,93 - 0,95*	*0,95 - 0,97*	*0,94 - 0,96*
exit	*0,96 - 0,98*	*0,36 - 0,64*	*0,63 - 0,98*	*0,45 - 0,77*
cooking	*0,99 - 1,00*	*0,62 - 0,70*	*0,85 - 0,97*	*0,67 - 0,78*
shower	*0,99 - 1,00*	*0,87 - 0,92*	*0,97 - 1,00*	*0,89 - 0,94*
pc	*0,97 - 0,98*	*0,98 - 0,99*	*0,99 - 1,00*	*0,98 - 0,99*
sleep	*0,95 - 0,98*	*0,62 - 0,66*	*0,83 - 0,97*	*0,70 - 0,74*
kitchen	*0,96 - 0,98*	*0,32 - 0,58*	*0,62 - 1,00*	*0,42 - 0,72*
TOTAL	***0,97 - 0,98***	***0,68 - 0,78***	***0,84 - 0,98***	***0,73 - 0,84***

discussed below, provide an interpretation of the factors that may have influenced the performance of the models.

The 30 + 15 configuration, which includes 30 min in the past and 15 min in the future, shows markedly better performance compared to the 45 + 0 configuration, which only considers 45 min in the past. The ability of the 30 + 15 configuration to capture information from both the recent past and the near future provides a considerable advantage in activity classification, allowing models to better interpret temporal transitions and patterns in the data. N-tolerance also plays a crucial role, as it allows the model to better handle temporal variations and adapt to changes in sensor data. In the CNN+GRU+ATT model, this configuration results in high accuracy values and F1 scores for activities, such as "resting" and "pc." Integrating a time window into the future allows models to more effectively anticipate and differentiate between activities that may have similar characteristics in the past time window. This anticipatory capability is crucial for accurate identification of activities involving long patterns or subtle changes in the data. The $N = 3$ tolerance also improves the model's ability to handle transitions and variations in activities with complex or less predictable patterns.

In contrast, the 45 + 0 configuration, which only uses data from the past without considering the future, shows a decrease in performance, especially in activities, such as "exit" and "kitchen." This reduction in accuracy and F1 score suggests that the lack of information about the future limits the ability of the model to capture changes and transitions

TABLE XIX
PERFORMANCE OF THE CNN+ATT+GRU+ATT (WITHOUT COLLECTIVE LOCATIONS) MODEL WITH A 30 + 15 TIME WINDOW (30 MIN IN THE PAST AND 15 MIN INTO THE FUTURE). METRICS INCLUDE ACCURACY (ACC), PRECISION (PRE), RECALL (REC), AND F1 SCORE (F1)

CNN+ATT+GRU+ATT (without collective locations) 30+15				
Activity	ACC N=0 - N=3	PRE N=0 - N=3	REC N=0 - N=3	F1 N=0 - N=3
toileting	*0,97 - 0,97*	*0,80 - 0,86*	*0,92 - 0,95*	*0,84 - 0,89*
resting	*0,98 - 0,98*	*0,98 - 0,98*	*0,98 - 0,99*	*0,98 - 0,98*
exit	*0,99 - 1,00*	*0,82 - 0,93*	*0,95 - 0,99*	*0,88 - 0,96*
cooking	*1,00 - 1,00*	*0,70 - 0,89*	*0,93 - 0,98*	*0,76 - 0,93*
shower	*1,00 - 1,00*	*0,91 - 0,94*	*0,98 - 1,00*	*0,92 - 0,96*
pc	*0,97 - 0,98*	*0,99 - 0,99*	*0,99 - 1,00*	*0,99 - 0,99*
sleep	*0,98 - 0,99*	*0,77 - 0,86*	*0,93 - 0,97*	*0,82 - 0,90*
kitchen	*0,99 - 1,00*	*0,83 - 0,92*	*0,98 - 1,00*	*0,89 - 0,95*
TOTAL	***0,98 - 0,99***	***0,85 - 0,92***	***0,96 - 0,99***	***0,88 - 0,95***

TABLE XX
PERFORMANCE OF THE CNN+ATT+GRU+ATT (WITHOUT COLLECTIVE LOCATIONS) MODEL WITH A 45 + 0 TIME WINDOW (30 MIN IN THE PAST AND 15 MIN INTO THE FUTURE). METRICS INCLUDE ACCURACY (ACC), PRECISION (PRE), RECALL (REC), AND F1 SCORE (F1)

CNN+ATT+GRU+ATT (without collective locations) 45+0				
Activity	ACC N=0 - N=3	PRE N=0 - N=3	REC N=0 - N=3	F1 N=0 - N=3
toileting	*0,96 - 0,97*	*0,73 - 0,79*	*0,89 - 0,96*	*0,77 - 0,84*
resting	*0,97 - 0,97*	*0,95 - 0,97*	*0,97 - 0,99*	*0,96 - 0,97*
exit	*0,96 - 0,98*	*0,35 - 0,65*	*0,62 - 0,99*	*0,44 - 0,77*
cooking	*0,99 - 1,00*	*0,64 - 0,73*	*0,84 - 0,98*	*0,67 - 0,80*
shower	*0,99 - 1,00*	*0,87 - 0,92*	*0,95 - 0,99*	*0,88 - 0,94*
pc	*0,97 - 0,98*	*0,98 - 0,99*	*0,99 - 1,00*	*0,98 - 0,99*
sleep	*0,95 - 0,97*	*0,58 - 0,69*	*0,79 - 0,96*	*0,65 - 0,77*
kitchen	*0,96 - 0,97*	*0,28 - 0,50*	*0,63 - 0,99*	*0,38 - 0,66*
TOTAL	***0,97 - 0,98***	***0,67 - 0,78***	***0,84 - 0,98***	***0,72 - 0,84***

TABLE XXI
PERFORMANCE OF THE TRANSFORMER+MULTIHEAD (WITHOUT COLLECTIVE LOCATIONS) MODEL WITH A 30 + 15 TIME WINDOW (30 MIN IN THE PAST AND 15 MIN INTO THE FUTURE). METRICS INCLUDE ACCURACY (ACC), PRECISION (PRE), RECALL (REC), AND F1 SCORE (F1)

Transformer+multihead (without collective locations) 30+15				
Activity	ACC N=0 - N=3	PRE N=0 - N=3	REC N=0 - N=3	F1 N=0 - N=3
toileting	*0,93 - 0,94*	*0,66 - 0,68*	*0,87 - 0,95*	*0,70 - 0,74*
resting	*0,95 - 0,95*	*0,93 - 0,93*	*0,90 - 0,93*	*0,90 - 0,93*
exit	*0,97 - 0,98*	*0,43 - 0,64*	*0,88 - 0,97*	*0,56 - 0,76*
cooking	*1,00 - 1,00*	*0,73 - 0,80*	*0,86 - 0,97*	*0,75 - 0,84*
shower	*0,98 - 0,99*	*0,90 - 0,86*	*0,93 - 0,90*	*0,89 - 0,86*
pc	*0,96 - 0,97*	*0,98 - 0,98*	*0,95 - 0,98*	*0,96 - 0,98*
sleep	*0,96 - 0,97*	*0,57 - 0,68*	*0,91 - 0,98*	*0,66 - 0,77*
kitchen	*0,97 - 0,98*	*0,43 - 0,65*	*0,89 - 0,98*	*0,57 - 0,77*
TOTAL	***0,96 - 0,97***	***0,70 - 0,78***	***0,90 - 0,96***	***0,75 - 0,83***

TABLE XXII
PERFORMANCE OF THE TRANSFORMER+MULTIHEAD (WITHOUT COLLECTIVE LOCATIONS) MODEL WITH A 45 + 0 TIME WINDOW (30 MIN IN THE PAST AND 15 MIN INTO THE FUTURE). METRICS INCLUDE ACCURACY (ACC), PRECISION (PRE), RECALL (REC), AND F1 SCORE (F1)

Transformer+multihead (without collective locations) 45+0				
Activity	ACC N=0 - N=3	PRE N=0 - N=3	REC N=0 - N=3	F1 N=0 - N=3
toileting	*0,88 - 0,89*	*0,56 - 0,59*	*0,83 - 0,91*	*0,61 - 0,66*
resting	*0,94 - 0,95*	*0,94 - 0,95*	*0,93 - 0,96*	*0,92 - 0,94*
exit	*0,91 - 0,94*	*0,19 - 0,42*	*0,76 - 1,00*	*0,29 - 0,57*
cooking	*0,99 - 0,99*	*0,66 - 0,67*	*0,91 - 0,98*	*0,70 - 0,74*
shower	*0,98 - 0,98*	*0,90 - 0,90*	*0,94 - 0,98*	*0,91 - 0,92*
pc	*0,95 - 0,96*	*0,95 - 0,96*	*0,94 - 0,96*	*0,94 - 0,96*
sleep	*0,92 - 0,94*	*0,43 - 0,52*	*0,87 - 0,98*	*0,52 - 0,62*
kitchen	*0,91 - 0,94*	*0,17 - 0,38*	*0,69 - 0,98*	*0,27 - 0,53*
TOTAL	***0,93 - 0,95***	***0,60 - 0,67***	***0,86 - 0,97***	***0,65 - 0,74***

TABLE XXIII
PERFORMANCE OF THE TRANSFORMER+POSITION ENCODING (WITHOUT COLLECTIVE LOCATIONS) MODEL WITH A 30 + 15 TIME WINDOW (30 MIN IN THE PAST AND 15 MIN INTO THE FUTURE). METRICS INCLUDE ACCURACY (ACC), PRECISION (PRE), RECALL (REC), AND F1 SCORE (F1)

Transformer+position encoding (without collective locations) 30+15				
Activity	ACC N=0 - N=3	PRE N=0 - N=3	REC N=0 - N=3	F1 N=0 - N=3
toileting	*0,96 - 0,97*	*0,80 - 0,82*	*0,90 - 0,94*	*0,83 - 0,85*
resting	*0,98 - 0,98*	*0,99 - 0,99*	*0,97 - 0,98*	*0,98 - 0,99*
exit	*0,99 - 1,00*	*0,72 - 0,89*	*0,94 - 1,00*	*0,80 - 0,93*
cooking	*1,00 - 1,00*	*0,75 - 0,86*	*0,87 - 0,96*	*0,76 - 0,89*
shower	*1,00 - 1,00*	*0,92 - 0,93*	*0,96 - 0,98*	*0,92 - 0,94*
pc	*0,98 - 0,98*	*0,98 - 0,99*	*0,99 - 1,00*	*0,99 - 0,99*
sleep	*0,97 - 0,98*	*0,68 - 0,77*	*0,93 - 0,98*	*0,75 - 0,83*
kitchen	*0,99 - 1,00*	*0,67 - 0,83*	*0,98 - 1,00*	*0,78 - 0,90*
TOTAL	***0,98 - 0,99***	***0,81 - 0,88***	***0,94 - 0,98***	***0,85 - 0,92***

TABLE XXIV
PERFORMANCE OF THE TRANSFORMER+POSITION ENCODING (WITHOUT COLLECTIVE LOCATIONS) MODEL WITH A 45 + 0 TIME WINDOW (45 MIN IN THE PAST ONLY). METRICS INCLUDE ACCURACY (ACC), PRECISION (PRE), RECALL (REC), AND F1 SCORE (F1)

Transformer+position encoding (without collective locations) 45+0				
Activity	ACC N0-N3	PRE N0-N3	REC N0-N3	F1 N0-N3
toileting	*0,96 - 0,97*	*0,73 - 0,79*	*0,86 - 0,94*	*0,77 - 0,83*
resting	*0,97 - 0,98*	*0,97 - 0,98*	*0,97 - 0,98*	*0,97 - 0,98*
exit	*0,94 - 0,97*	*0,27 - 0,54*	*0,78 - 1,00*	*0,39 - 0,69*
cooking	*0,99 - 1,00*	*0,65 - 0,74*	*0,86 - 0,94*	*0,69 - 0,80*
shower	*0,99 - 1,00*	*0,88 - 0,90*	*0,98 - 0,98*	*0,90 - 0,93*
pc	*0,97 - 0,98*	*0,97 - 0,98*	*0,99 - 1,00*	*0,98 - 0,99*
sleep	*0,93 - 0,95*	*0,46 - 0,56*	*0,91 - 0,99*	*0,56 - 0,67*
kitchen	*0,94 - 0,96*	*0,24 - 0,49*	*0,74 - 0,99*	*0,36 - 0,64*
TOTAL	***0,96 - 0,97***	***0,65 - 0,75***	***0,89 - 0,98***	***0,70 - 0,82***

between activities. Activities with more subtle patterns or rapid transitions may not be adequately differentiated with past information alone, resulting in a higher classification error rate. The N-tolerance also affects this configuration, as a lower tolerance may not be sufficient to capture subtle variations in the data.

Another important factor in interpreting the results is the influence of sensors and environmental parameters, such as humidity and temperature, on the classification of activity. Activities, such as "showering" and "cooking," are associated with sensors that measure specific changes in humidity and temperature, which is clearly reflected in the classification results. These activities have more detailed and specific rules that use multiple sensors to capture changes in these parameters, which contributes to a more accurate identification in the 30 + 15 configuration. N-tolerance also helps to better capture these variations, improving the classification accuracy. On the other hand, activities, such as "exit" and "kitchen," present classification challenges due to the generality of the rules and the lack of specific environmental parameters. The "exit" activity, for example, may involve less distinct changes in sensor data, which can make it difficult to identify correctly. The 45+0 configuration, by not considering information from the future, may exacerbate these problems by not capturing

subtle transitions or patterns that could help better differentiate these activities.

The CNN + GRU + ATT model generally outperforms other models in terms of accuracy and F1 score, particularly in the 30 + 15 configuration. This can be attributed to the ability of the CNN + GRU + ATT to take advantage of both spatial (captured by the convolutional layers) and temporal (captured by the recurrent and ATT layers) features. The combination of these capabilities allows the model to perform a more accurate and robust classification of complex activities. The N-tolerance in this configuration also improves the model's ability to handle subtle temporal patterns and transitions.

In the CNN+GRU+ATT configuration of 45+0 (Table VI) with $N = 0$, the overall performance is lower, with an average F1 of 0.74. Although "resting" and "pc" still show high scores (0.96 and 0.98), activities, such as "exit" and "kitchen," have low F1-s (0.44). When using $N = 3$, overall performance improves to an average F1 of 0.86, with better scores for activities, such as "exit" and "kitchen," although still lower than "resting" and "pc."

The CNN + GRU + ATT configuration with a temporal window of 30 + 15 (Table V) and $N = 0$, the model presents an outstanding performance with an average F1 of 0.9. Activities, such as "resting" and "pc," stand out with F1 close to 0.99, while "cooking" and "sleep" have lower results, with F1 of 0.82. With $N = 3$, performance improves overall, reaching an average F1 of 0.94. F1 scores for "resting" and "pc" remain high, and activities, such as "cooking" and "sleep," also improve markedly.

In comparison, the GRU and LSTM model, while also showing good performance in the 30 + 15 configuration, tends to be less effective in accurately identifying activities with complex or subtle patterns, especially in the 45 + 0 configuration. This suggests that the GRU model, lacking the ATT component, may face greater difficulties in interpreting data sequences with rapid temporal changes or less obvious patterns. The GRU-based model achieved slightly better results than LSTM in terms of both complexity and F1-score.

The GRU configuration of 30 + 15 (Table I) with $N = 0$, the model achieves an average F1 of 0.88. Activities, such as "resting," "pc," and "shower," have high F1 scores (0.97, 0.99, and 0.92), while "sleep" and "cooking" present lower values (0.76 and 0.79). With $N = 3$, performance improves to an average F1 of 0.93, with high scores for most activities, particularly for "resting," "pc," and "shower."

The GRU configuration of 45 + 0 (Table II) with $N = 0$, overall performance is moderate, with an average F1 of 0.73. Activities, such as "resting" and "pc," obtain good F1 scores (0.93 and 0.98), but "exit" and "kitchen" have low F1-s (0.41 and 0.37). With $N = 3$, performance improves to an average F1 of 0.84, with notable improvement in activities, such as "exit" and "kitchen," although they are still lower than "resting" and "pc."

It is observed that the definition of the temporal window and the tolerance parameter of recognition has a significant impact on performance. The CNN + GRU + ATT configuration with a temporal window of 30 + 15 and $N = 3$ offers the best overall performance, with an average F1 of 0.95. This model shows a noticeable improvement in classifying activities as "cooking" and "sleep" as the tolerance of N increases, highlighting its ability to handle temporal variability in predictions. In contrast, the same configuration with N = 0 performs poorly, with an average F1 of 0.90, indicating that a larger tolerance range in the evaluation may be beneficial in capturing more complex patterns in the data.

In comparison, the CNN + GRU + ATT configuration with the 45 + 0 window shows a decrease in performance, with an average F1 of 0.74 and 0.86 for $N = 0$ and $N = 3$, respectively. Providing real-time recognition for delayed recognition is the cause of slightly decreasing performance. Although the addition of $N = 3$ improves the results, the performance is still lower than the 30 + 15 window configuration. This suggests that a wider time window without proper tolerance may introduce noise or dilute the model's ability to make accurate predictions. However, the GRU configuration with the 30 + 15 window shows consistent performance with an average F1 of 0.88 and 0.93 for $N = 0$ and $N = 3$, respectively, indicating that the GRU model is robust to changes in the parameter N. The 45 + 0 GRU configuration also improves with $N = 3$, achieving an average F1 of 0.84 versus 0.73 with $N = 0$. However, it remains less effective compared to the CNN+GRU+ATT configurations.

The Transformer-based models, while demonstrating good overall performance, did not surpass the CNN + GRU + ATT model. This suggests that the explicit extraction of spatial and temporal features by the CNN and GRU layers, combined with the ATT mechanism, provides a more effective representation for this specific task. In particular, the Transformer model with positional encoding exhibited competitive results, indicating the importance of incorporating temporal order information for activity recognition. However, it appears that the benefits of positional encoding, while contributing to competitive performance, were not sufficient to outperform the CNN + GRU + ATT approach. Furthermore, the CNN + GRU + ATT model generally exhibited lower computational complexity compared to the Transformer-based models, making it a more efficient choice for resource-constrained environments.

The results obtained highlight the importance of properly selecting the time window configuration and input parameters to maximize activity classification performance. The 30 + 15 configuration proves to be the most effective for most activities, especially those involving complex patterns and subtle transitions. For activities with less distinctive patterns or more general rules, such as "exit" and "kitchen," integrating additional information or improving classification rules can be beneficial. The combination of an appropriate time window and the integration of information about environmental parameters and temporal patterns is key to improving activity classification performance. Choosing models that can effectively capture and utilize these features, such as CNN+GRU+ATT, can offer significant advantages in practical activity classification applications.

On the evaluation of integrating only individual location versus individual and collective location, the analysis reveals that incorporating location information from other users significantly improves the recognition accuracy of specific activities,

such as toileting and showering, in confined spaces with an area like the toilet. This improvement likely stems from the added contextual information provided by the location of other occupants, which helps to disambiguate activities with similar sensor interaction patterns in close proximity.

However, the impact of integrating other users' location data was less pronounced and slightly reduced the performance for most activities in other rooms, which tend to be more solitary and less influenced by the presence or actions of other occupants. This suggests that the benefits of incorporating such contextual information may vary depending on the nature of the activity and the spatial dynamics of the environment.

V. Conclusion and Ongoing Works

This study proposes an innovative framework for multi-occupancy activity recognition, combining UWB technology and binary sensors to enable accurate, privacy-preserving monitoring. The integration of localization heatmaps with nearby sensor data significantly improved recognition accuracy, achieving up to 90% F1 scores by leveraging temporal windows that incorporate both past and future data. The inclusion of future-aware windows proved particularly effective, enhancing model performance in scenarios with overlapping actions and diverse sensor inputs. Advanced deep learning models, such as CNNs with GRU and ATT mechanisms, excelled in capturing complex activity patterns. Notably, the analysis highlighted the critical role of individual location data in confined spaces, such as toileting, where it consistently enhanced performance, often surpassing results achieved with collective data. These findings underscore the scalability and adaptability of the proposed approach, paving the way for smart environments and assistive applications.

Future research will focus on enriching datasets by incorporating diverse multioccupant environments with overlapping activities and routine interruptions in residential contexts. Additional efforts will explore adaptive time window configurations that dynamically respond to activity patterns and integrate complementary sensors, such as sound [41], light, and air quality, to improve contextual understanding. Furthermore, the development of personalization techniques and robust privacy-preserving mechanisms will be essential to tailor models to individual behaviors while ensuring secure deployment. These advancements will support scalable, user-centered solutions for multioccupancy smart environments.

References

[1] S. Qiu et al., "Multi-sensor information fusion based on machine learning for real applications in human activity recognition: State-of-the-art and research challenges," *Inf. Fusion*, vol. 80, pp. 241–265, Apr. 2022.

[2] P. Franco, J. M. Martinez, Y.-C. Kim, and M. A. Ahmed, "IoT based approach for load monitoring and activity recognition in smart homes," *IEEE Access*, vol. 9, pp. 45325–45339, 2021.

[3] M. Lupión, J. Medina-Quero, J. F. Sanjuan, and P. M. Ortigosa, "DOLARS, a distributed on-line activity recognition system by means of heterogeneous sensors in real-life deployments—A case study in the smart lab of the university of Almería," *Sensors*, vol. 21, no. 2, p. 405, 2021.

[4] M. Elsanhoury et al., "Precision positioning for smart logistics using ultra-wideband technology-based indoor navigation: A review," *IEEE Access*, vol. 10, pp. 44413–44445, 2022.

[5] D. Gnaś and P. Adamkiewicz, "Precise indoor location system using ultra-wideband technology," *Przeglad Elektrotechniczny*, vol. 99, no. 1, p. 270, 2023.

[6] D. Trivedi and V. Badarla, "Occupancy detection systems for indoor environments: A survey of approaches and methods," *Indoor Built Environ.*, vol. 29, no. 8, pp. 1053–1069, 2020.

[7] S. J. Hayward, K. van Lopik, C. Hinde, and A. A. West, "A survey of indoor location technologies, techniques and applications in industry," *Internet Things*, vol. 20, Nov. 2022, Art. no. 100608.

[8] X. Zhu, W. Qu, T. Qiu, L. Zhao, M. Atiquzzaman, and D. O. Wu, "Indoor intelligent fingerprint-based localization: Principles, approaches and challenges," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 4, pp. 2634–2657, 4th Quart., 2020.

[9] S. C. Bastidas, M. Espinilla, and J. M. Quero, "Review of ultra wide band (UWB) for indoor positioning with application to the elderly," in *Proc. 55th Hawaii Int. Conf. Syst. Sci.*, 2022, pp. 2145–2154.

[10] L. Yao, L. Yao, and Y.-W. Wu, "Analysis and improvement of indoor positioning accuracy for UWB sensors," *Sensors*, vol. 21, no. 17, p. 5731, 2021.

[11] P. S. Farahsari, A. Farahzadi, J. Rezazadeh, and A. Bagheri, "A survey on indoor positioning systems for IoT-based applications," *IEEE Internet Things J.*, vol. 9, no. 10, pp. 7680–7699, May 2022.

[12] J. Gómez, J. R. Casas, and S. Villalba, "Structural health monitoring with distributed optical fiber sensors of tunnel lining affected by nearby construction activity," *Autom. Constr.*, vol. 117, Sep. 2020, Art. no. 103261.

[13] A. Polo-Rodríguez, F. Cavallo, C. Nugent, and J. Medina-Quero, "Human activity mining in multi-occupancy contexts based on nearby interaction under a fuzzy approach," *Internet Things*, vol. 25, Apr. 2024, Art. no. 101018.

[14] Q. Li, R. Gravina, Y. Li, S. H. Alsamhi, F. Sun, and G. Fortino, "Multi-user activity recognition: Challenges and opportunities," *Inf. Fusion*, vol. 63, pp. 121–135, Nov. 2020.

[15] K. Chen, D. Zhang, L. Yao, B. Guo, Z. Yu, and Y. Liu, "Deep learning for sensor-based human activity recognition: Overview, challenges, and opportunities," *ACM Comput. Surv.*, vol. 54, no. 4, pp. 1–40, 2021.

[16] I. Khan, O. Zedadra, A. Guerrieri, and G. Spezzano, "Occupancy prediction in IoT-enabled smart buildings: Technologies, methods, and future directions," *Sensors*, vol. 24, no. 11, p. 3276, 2024.

[17] Y. Zhan and H. Haddadi, "MoSen: Activity modelling in multiple-occupancy smart homes," 2021, *arXiv:2101.00235*.

[18] A. Naser, A. Lotfi, and J. Zhong, "Adaptive thermal sensor array placement for human segmentation and occupancy estimation," *IEEE Sensors J.*, vol. 21, no. 2, pp. 1993–2002, Jan. 2021.

[19] L. Hiley, A. Preece, Y. Hicks, S. Chakraborty, P. Gurram, and R. Tomsett, "Explaining motion relevance for activity recognition in video deep learning models," 2020, *arXiv:2003.14285*.

[20] H. Yuan et al., "Self-supervised learning for human activity recognition using 700,000 person-days of wearable data," *NPJ Digit. Med.*, vol. 7, no. 1, p. 91, 2024.

[21] B. Fu, N. Damer, F. Kirchbuchner, and A. Kuijper, "Sensing technology for human activity recognition: A comprehensive survey," *IEEE Access*, vol. 8, pp. 83791–83820, 2020.

[22] S. Zhang et al., "Deep learning in human activity recognition with wearable sensors: A review on advances," *Sensors*, vol. 22, no. 4, p. 1476, 2022.

[23] A. Polo-Rodriguez, M. A. Anguita-Molina, I. Rojas, and J. Medina-Quero. "Enhanced multi-occupant tracking via fusion of low-intrusive radar and wearable ultra-wideband devices using Autoencoders." SSRN.com. 2024. [Online]. Available: http://dx.doi.org/10.2139/ssrn.4892704

[24] T. Yang, A. Cabani, and H. Chafouk, "A survey of recent indoor localization scenarios and methodologies," *Sensors*, vol. 21, no. 23, p. 8086, 2021.

[25] H. Obeidat, W. Shuaieb, O. Obeidat, and R. Abd-Alhameed, "A review of indoor localization techniques and wireless technologies," *Wireless Pers. Commun.*, vol. 119, pp. 289–327, Feb. 2021.

[26] M. Javaid, A. Haleem, S. Rab, R. P. Singh, and R. Suman, "Sensors for daily life: A review," *Sens. Int.*, vol. 2, Jan. 2021, Art. no. 100121.

[27] R. Mutegeki and D. S. Han, "A CNN-LSTM approach to human activity recognition," in *Proc. Int. Conf. Artif. Intell. Inf. Commun. (ICAIIC)*, 2020, pp. 362–366.

[28] I. U. Khan, S. Afzal, and J. W. Lee, "Human activity recognition via hybrid deep learning based model," *Sensors*, vol. 22, no. 1, p. 323, 2022.

[29] E. Ramanujam, T. Perumal, and S. Padmavathi, "Human activity recognition with smartphone and wearable sensors using deep learning techniques: A review," *IEEE Sensors J.*, vol. 21, no. 12, pp. 13029–13040, Jun. 2021.
[30] M. J. Rodrigues, O. Postolache, and F. Cercas, "Physiological and behavior monitoring systems for smart healthcare environments: A review," *Sensors*, vol. 20, no. 8, p. 2186, 2020.
[31] 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," *Knowl.-Based Syst.*, vol. 223, Jul. 2021, Art. no. 106970.
[32] P. Kumar, S. Chauhan, and L. K. Awasthi, "Human activity recognition (HAR) using deep learning: Review, methodologies, progress and future research directions," *Arch. Comput. Methods Eng.*, vol. 31, no. 1, pp. 179–219, 2024.
[33] S. Bian, M. Liu, B. Zhou, and P. Lukowicz, "The state-of-the-art sensing techniques in human activity recognition: A survey," *Sensors*, vol. 22, no. 12, p. 4596, 2022.
[34] A. Polo-Rodríguez, S. Rotbei, S. Amador, O. Baños, D. Gil, and J. Medina, "Smart architectures for evaluating the autonomy and behaviors of people with autism spectrum disorder in smart homes," in *Neural Engineering Techniques for Autism Spectrum Disorder*. Amsterdam, The Netherlands Elsevier, 2021, pp. 55–76.
[35] A. Polo-Rodríguez, J. C. Valera, J. Peral, D. Gil, and J. Medina-Quero, "Tracking daily paths in home contexts with RSSI fingerprinting based on UWB through deep learning models," *Multimedia Tools Appl.*, pp. 1–25, Aug. 2024. [Online]. Available: https://doi.org/10.1007/s11042-024-19914-1
[36] A. Polo-Rodríguez et al., "Discovering social interactions between caregivers and frail individuals using indoor localization," in *Proc. UCAMI Conf.*, 2024, pp. 319–331.
[37] J. Medina-Quero, S. Zhang, C. Nugent, and M. Espinilla, "Ensemble classifier of long short-term memory with fuzzy temporal windows on binary sensors for activity recognition," *Expert Syst. Appl.*, vol. 114, pp. 441–453, Dec. 2018.
[38] J. Medina-Quero, C. Orr, S. Zang, C. Nugent, A. Salguero, and M. Espinilla, "Real-time recognition of interleaved activities based on ensemble classifier of long short-term memory with fuzzy temporal windows," *Proceedings*, vol. 2, no. 19, p. 1225, 2018.
[39] A. Polo-Rodríguez, J. Medina-Quero, and M. Á. Anguita-Molina, 2024, "Multi-occupancy activity recognition," Dataset. [Online]. Available: https://www.kaggle.com/ds/5788723
[40] R. A. Hamad, A. S. Hidalgo, M.-R. Bouguelia, M. E. Estevez, and J. M. Quero, "Efficient activity recognition in smart homes using delayed fuzzy temporal windows on binary sensors," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 2, pp. 387–395, Feb. 2020.
[41] A. Polo-Rodriguez, J. M. V. Chiachio, C. Paggetti, and J. Medina-Quero, "Ambient sound recognition of daily events by means of convolutional neural networks and fuzzy temporal restrictions," *Appl. Sci.*, vol. 11, no. 15, p. 6978, 2021.

Miguel Ángel Anguita-Molina received the B.S. degree in computer engineering from the University of Jaén, Jaén, Spain, in 2021, and the M.S. degree in biomedical engineering and digital health from the University of Seville, Seville, Spain, in 2022. He is currently pursuing the Ph.D. degree in information and communication technologies with the University of Granada, Granada, Spain.

His research interests include IoT, activity recognition, smart homes, deep learning, and eHealth.

Mr. Anguita-Molina received special recognition at the 7th Ada Lovelace Awards from the University of Jaén for his thesis degree.

Pedro J. S. Cardoso received the B.Sc. degree in mathematics/computer science from the University of Coimbra, Coimbra, Portugal, in 1996, the M.Sc. degree in computational mathematics from the University of Minho, Braga, Portugal, in 1999, and the Ph.D. degree in mathematics/operations research from the University of Seville, Seville, Spain, in 2007.

Serving as a Coordinator Professor and a Researcher with the Universidade do Algarve, Faro, Portugal, for over two decades, he is also an integrated member of the NOVA LINCS research center. Throughout his career, he has been involved in over a dozen scientific and development projects and has published more than 90 papers. His primary research interests lie in machine learning, particularly its real-world applications, affective computing, anomaly detection, and energy efficiency.

João M. F. Rodrigues received the fiver years bachelor's degree (Licenciatura) in electrical engineering (electronics, instrumentation and computation) from the University of Trás os Montes e Alto Douro, Vila Real, Portugal, in 1993, the M.S. degree in computer systems engineering (major in computer systems) and the Ph.D. degree in electronics and computer engineering (major in computer sciences) from the University of the Algarve, Faro, Portugal, in 1998 and 2008, respectively, and the Habilitation degree (Agregação) in electrical and computer engineering from the University of Trás os Montes e Alto Douro in 2021.

He is a Professor with the Institute of Engineering, University of the Algarve, Faro, Portugal, where he has lectured in Computer Science and Computer Vision Curricular Units since 1994. He is the Vice-Rector for Transference, Innovation and Digital University, and a member of NOVA LINCS—Algarve research center. He participated in more than 35 financed scientific projects, several as a PI, and co-authored more than 250 scientific publications. His research interests lie in computer vision, human–computer interaction, and human-centered AI.

Javier Medina-Quero received the M.Sc. and Ph.D. degrees in computer science from the University of Granada, Granada, Spain, in 2007 and 2010, respectively.

He is an Associate Professor of Computer Science with the University of Granada. He has published more than 40 articles in impact index journals related to Health and Computer Sciences. Furthermore, as a Principal Investigator and a Collaborating Researcher, he has participated in several European, national, and regional research projects focused on e-health. His research interests include eHealth, machine learning, and ubiquitous computing.

Prof. Medina-Quero has participated as a guest editor in several JCR journals.

Aurora Polo-Rodríguez received the B.Sc. and M.Sc. degrees in computer science in 2020 and 2022, respectively, and the Ph.D. degree in information and communication technologies in 2024 from the University of Jaén, Jaén, Spain.

She is a Postdoctoral Researcher with the University of Granada, Granada, Spain, and has actively contributed to various European, national, and regional research projects, such as PHArA-ON, which focuses on improving the quality of life for the elderly using advanced technologies. Her research interests include smart environments, IoT, e-health, and machine learning.

Dr. Polo-Rodríguez received the best final degree awards for both her bachelor's and master's degrees, with her doctoral thesis earning "cum laude" distinction. She also won the "Platinum Curriculum Distinction" for academic excellence, along with a competitive research collaboration scholarship.