

## Article

# An Unsupervised Method to Recognise Human Activity at Home Using Non-Intrusive Sensors

Raúl Gómez-Ramos <sup>1,2</sup>, Jaime Duque-Domingo <sup>2</sup>, Eduardo Zalama <sup>1,2,\*</sup> and Jaime Gómez-García-Bermejo <sup>1,2</sup>

<sup>1</sup> CARTIF, Technological Center, 47151 Valladolid, Spain; raugom@cartif.es (R.G.-R.);

jaigom@eii.uva.es (J.G.-G.-B.)

<sup>2</sup> ITAP-DISA, University of Valladolid, 47002 Valladolid, Spain; jaime.duque@uva.es

\* Correspondence: ezalama@uva.es

**Abstract:** As people get older, living at home can expose them to potentially dangerous situations when performing everyday actions or simple tasks due to physical, sensory or cognitive limitations. This could compromise the residents' health, a risk that in many cases could be reduced by early detection of the incidents. The present work focuses on the development of a system capable of detecting in real time the main activities of daily life that one or several people can perform at the same time inside their home. The proposed approach corresponds to an unsupervised learning method, which has a number of advantages, such as facilitating future replication or improving control and knowledge of the internal workings of the system. The final objective of this system is to facilitate the implementation of this method in a larger number of homes. The system is able to analyse the events provided by a network of non-intrusive sensors and the locations of the residents inside the home through a Bluetooth beacon network. The method is built upon an accurate combination of two hidden Markov models: one providing the rooms in which the residents are located and the other providing the activity the residents are carrying out. The method has been tested with the data provided by the public database SDHAR-HOME, providing accuracy results ranging from 86.78% to 91.68%. The approach presents an improvement over existing unsupervised learning methods as it is replicable for multiple users at the same time.

**Keywords:** smart home; IoT; HAR; HMM; sensors; SDHAR-HOME



**Citation:** Gómez-Ramos, R.; Duque-Domingo, J.; Zalama, E.; Gómez-García-Bermejo, J. An Unsupervised Method to Recognise Human Activity at Home Using Non-Intrusive Sensors. *Electronics* **2023**, *12*, 4772. <https://doi.org/10.3390/electronics12234772>

Academic Editor: Duc Tai Le

Received: 27 October 2023

Revised: 16 November 2023

Accepted: 23 November 2023

Published: 24 November 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Human activity recognition (HAR) has become a major field of research in the health-care sector over the last two decades due to its application in domains such as entertainment, security and wellbeing improvement [1]. HAR is usually accompanied by the use of Internet-of-Things (IoT) technologies [2], such as sensors, images or smartphones [3], in order to extract information from the environment and converge towards a specific activity of daily life [4]. However, identifying patterns of behaviour and actions performed by individuals is a challenge due to the speed and duration of certain activities, the variability and different ways of performing a given activity or the spatial distribution of the household itself [5]. All this requires the creation of a generalist model that can be applicable to the great diversity of cases that may occur when carrying out HAR [6].

Moreover, the percentage of older people has been increasing in recent years, thus creating a need for assistance due to their natural loss of autonomy [7]. As people grow older, the probability of mental [8,9] and physical decline increases [10,11]. This may result in dangerous situations, such as accidental falls, malnutrition or sleeping problems, among others [12]. For this reason, a correct recognition of activities in real time is important for an appropriate intervention, when required [13]. All these circumstances represent the need to establish a mechanism that monitors the behaviour patterns of this group of people [14]

through activity recognition technologies and algorithms [15]. Detecting an absence of activity is also very important as it can provide very important information about the user. This can be an indication that the user is down or has had a problem [16].

An HAR system is composed of three functional subsystems [17]: (i) a sensing module responsible for continuously collecting information from the environment [18], (ii) a processing module responsible for extracting the main features from the sensor signals to discriminate between activities [19] and (iii) a classification module to identify the activity from the key features extracted by the previous module [20]. Concerning (i), the use of computer vision technology has been proposed [21,22] due to its high reliability [23]. A large amount of temporal information about the development of an activity can be extracted by studying the person's postures and movements, the objects he/she is using or the environment where the activity is taking place [24]. Sensor-based methods [25,26] through discrete event analysis represent a valuable alternative that can be advantageous because it is not perceived as intrusive [27]. These methods provide information about the environment and home conditions at the time the activity takes place [28] or about the person's postural parameters [29,30].

Concerning (ii) and (iii), information can be extracted from sensors by applying machine learning (ML) methods [31,32]. These learning methods can be supervised or unsupervised. Supervised learning methods need to have labelled outputs corresponding to the inputs crossing the model [33]. This requirement is necessary when using neural networks, such as convolutional neural networks (CNN) [34,35] or recurrent neural networks (RNN) [36]. For example, there are networks based on different cascading CNN layers, which can be combined with other DL modules, such as attention modules [37]. Learning is performed by tagging the user's activities, which has to be performed manually by the user or a collaborating expert. This can be costly and tedious because it usually requires collecting data over an extended period of time to achieve successful learning (weeks or even months). So, while this learning method is possible in theory and has been used in previous research, it can be difficult to implement in practice. Moreover, when the system is moved to another house, with other people and other sensors, a retraining process is usually required. The situation is further complicated in the case of more than one user living in the same house, as the sensors are not able to collect information inherent to a particular user. In contrast, unsupervised methods only need input data, and, by means of different classification and clustering algorithms, the system is able to generate a set of outputs [38]. An example of unsupervised methods could be hidden Markov models (HMM) [39]. The application of unsupervised methods is computationally more complex than the application of supervised methods (regardless of training) [40]. Other methods are known as semi-supervised learning methods. These methods perform learning by combining a small set of labelled data (typical of supervised learning) with a larger set of unlabelled data (typical of unsupervised learning). For example, the broad learning system (BLS) is a semi-supervised learning method with high computational efficiency that presents good feature extraction behaviour [41]. The main problem with supervised methods is the need to have a specific output for a set of inputs, and this does not occur in real cases where real-time learning is desired [42]. Moreover, in HMMs, more precise control over the system's operation is available compared to neural network-based methods, where knowledge is diffused throughout the system. This advantage facilitates the adaptation of HMMs for a larger number of residents. HMMs are also advantageous over rule-based models [43] or decision trees [44], which are stricter and less flexible methods. Moreover, the case of two people living together is frequent and should be addressed. However, it is challenging since, on many occasions, both users are in close proximity within the home, doing similar activities, which hinders HAR. Likewise, the system has to be fast enough to anticipate possible dangerous situations, which represents a further challenge.

The main purpose of this paper is the development of an unsupervised learning method for recognising, with high reliability, the activities being performed by several people inside a house. Although supervised learning methods usually offer better results,

their implementation in real environments is much more complex than unsupervised methods due to labelling during the training phase. The paper aims to perform HAR by analysing discrete events provided by sensors distributed around the home, through the use of an unsupervised method based on HMMs. In order to validate the proposed approach, it is tested on a public database obtained from a two-month real case of two people living together. Likewise, the learning model is intended to be as generalist as possible in order to be applicable in homes other than the one under analysis.

The main contributions of the paper are the following: (i) A method for recognising activities carried out by several people simultaneously inside a house is developed by applying unsupervised learning methods. (ii) The system is able to distinguish between a total of 16 different activities inside a house (e.g., sleeping, eating or taking medicine). (iii) The technology chosen to detect the activities does not compromise the privacy of the users of the house. Sensors collect the events that occur in the house. For this reason, the use of cameras and microphones has been avoided [45], since this type of technology captures information linked to the residents. (iv) The system predictions are carried out in real time, minimising the time required to execute a response action (e.g., if the user has taken more medication than recommended, the health services should be notified as soon as possible). (v) The predictions are independent of the actual time, which gives greater flexibility to the solution. (vi) Models of the inhabitants are linked through behavioural patterns, reinforcing the effectiveness of the system.

The article is structured as follows: Section 2 gives an overview of the main existing learning methods to perform HAR, both supervised and unsupervised, looking at hit rates and databases used to verify the approach reliability. Section 3 addresses an overview of the method and the system developed to perform HAR (Section 3.2) and a description of the database chosen to evaluate it (Section 3.1). Section 4 details the different experiments that have been carried out to verify the validity of the system developed in this paper. Section 5 describes the evidence obtained during the experimentation stage. Finally, Section 6 summarises the main strengths of the developed system as well as possible future lines of research.

## 2. Related Work

In this section, the latest technologies used to perform activity recognition by supervised and unsupervised methods are discussed.

### 2.1. Summary of the Main HAR Supervised Learning Methods

In recent years, activity recognition has become a target of study for many research collectives due to its multiple fields of application. Within activity recognition and DL systems, one group widely used is supervised methods. These methods can reach high recognition accuracy given that the model learns as it analyses the input data, knowing the output it has to generate. In [46], the authors developed a system capable of analysing the main joints of users and their movements using computer vision during different daily activities. To build the prediction model, the authors used an Inception-ResNet neural network [47]. Moreover, they evaluated the approach against three public databases: UTD\_MHAD [48], HDM05 [49] and NTU RGB+D 60 (NTU60) [50], obtaining accuracies between 85.45% and 98.13%. In general, computer-vision-based approaches can achieve high recognition rates, but may compromise users' privacy significantly. Other studies, such as the one carried out by [51], also analyse video sequences to obtain the postures performed during gymnastics sessions recorded by themselves, obtaining an accuracy rate of over 98%. To achieve these results, the authors used a model based on multilayer perceptron (MLP) networks [52], integrated with an unsupervised HMM model. In [53], the authors present a system capable of analysing signals provided by MEMS inertial sensors by applying long short-term memory (LSTM) RNNs [54]. The effectiveness of their system was evaluated with two open-access databases, WISDM [55] and PAMAP2 [56], obtaining accuracies over 97%. In [57], the authors developed a robust classification model

to perform HAR by analysing signals from wearable sensors. For this purpose, they used CNN networks combined with bidirectional LSTM (BiLSTM) layers in order to extract the main features of the signals before converging to the resulting activities. The authors obtained their results by applying their model to three databases: UCI-HAR [58], WISDM and PAMAP2, resulting in hit rates ranging from 94% to 96%. The same databases were used in the work by [59]. In this case, the authors combined CNN networks with gated recurrent unit (GRU) networks [60], giving results that reach 97% accuracy. The work by [61] also used bidirectional LSTM networks to obtain a set of 16 activities with an accuracy of 95.42%. In this case, the authors used the Milan [62] database, which is a subset of the CASAS [63] database group. Other authors proposed the use of supervised support vector machine (SVM) models [64], such as the work developed by [65]. The authors managed to detect six ambulation activities (e.g., walking, climbing stairs or sitting) with an accuracy of 94.66%. The work carried out by [66] presents the development of a skeleton-based action recognition model by applying a semantics-guided neural network (SGN) composed of the combination of CNNs with graph convolutional networks (GCNs) [67]. The model was tested against three different databases: NTU RGB+D 60 (NTU60), NTU RGB+D 120 (NTU120) [68] and SYSU 3D Human-Object Interaction Dataset (SYSU) [69], with results ranging from a 79.2% to 90.6% hit rate. Another work using environmental sensors is the one developed by [70], in which they created an HAR algorithm based on CNN networks to analyse the main features of the time domain combined with a series of dense layers [71]. The authors used the CASAS database (Cairo, Milan, Kyoto7, Kyoto8 and Kyoto11) to obtain an accuracy between 86.68% and 97.08%. In [72], the authors developed a database (SDHAR-HOME) with a total of 18 activities, whose measurements correspond to signals from environmental sensors located in a real home inhabited by two people. The authors applied three DL approaches: RNN [73], LSTM and GRU, obtaining a hit rate of 90.91%. Finally, there is the work by [74], in which they used a novel DL method known as Transformer models [75]. One of the main objectives of a Transformer model is to reduce the computational load by applying temporal inference algorithms to the information that contributes most value to the model while spending less time and load on less important information. The authors applied this model to three databases: Penn-Action [76], NTU60 and NTU120, obtaining results above 90%. A summary of all the works presented is given in Table 1.

**Table 1.** Overview of the main HAR supervised learning methods.

Development	DL Method	Database	Activities	Accuracy
[46]	Inception-ResNet	UTD_MHAD	27	98.13%
		HDM05	130	90.67%
		NTU RGB+D 60	60	85.45%
[51]	MLP + HMM	Own development	19	98.05%
[53]	LSTM	WISDM	6	97.32%
		PAMAP2	18	97.15%
[57]	CNN + BiLSTM	UCI-HAR	6	96.37%
		WISDM	6	96.05%
		PAMAP2	18	94.29%
[59]	CNN + GRU	UCI-HAR	6	96.20%
		WISDM	6	97.21%
		PAMAP2	18	95.27%
[61]	BiLSTM	Milan	16	95.42%
[65]	SVM	Own development	6	94.66%
[66]	SGN -> GCN + CNN	NTU RGB+D 60	60	89%
		NTU RGB+D 120	120	79.20%
		SYSU	12	90.60%

Table 1. Cont.

Development	DL Method	Database	Activities	Accuracy
[70]	CNN	Cairo	13	91.99%
		Milan	15	95.35%
		Kyoto7	13	86.68%
		Kyoto8	12	97.08%
		Kyoto11	25	90.27%
[72]	RNN , LSTM and GRU	SDHAR-HOME	18	90.91%
[74]	Transformer	Penn-Action	15	98.7%
		NTU RGB+D 60	60	92%
		NTU RGB+D 120	120	90.3%

## 2.2. Summary of the Main HAR Unsupervised Learning Methods

Unlike the supervised methods discussed in Section 2.1, unsupervised methods do not require prior knowledge of the output to be generated by the model during the training stage [77]. This is an advantage when the model is implemented in real environments, since in many cases there is no set of outputs in accordance with a set of inputs. This advantage improves the replicability of the system, giving the possibility of using the same model in different residences with different users and different living habits. The work developed by [78] proposes the use of statistical methods to extract the main characteristics of the signals generated by the sensors of a smartphone from the WEKA database [79]. The chosen method is the application of fast Fourier transform (FFT) [80]. Another unsupervised algorithm for HAR is the HMM method [81]. This algorithm is used in the work developed by [82] to perform a regression with respect to the information from a set of nine accelerometers. The authors were able to recognise a total of 12 activities with a hit rate of 89%. Another system implemented upon HMM is the one elaborated by [83]. In this case, the authors previously extracted the characteristics of the input signals by means of a micro-Doppler method [84]. The method was tested against a database generated during experimentation, with five wandering activities and an accuracy of 69% was obtained. In [85], the authors propose the use of an unsupervised method known as the variational autoencoder (VAE) [86]. This method takes care of compressing the input data vector into a representation vector and, in subsequent steps, decompressing it back into an output data vector with the same dimensions, thus exploiting the truly useful information in the input data [87]. The authors evaluated their method against an unlabelled database known as HHAR [88], obtaining a hit rate of 87%. The work carried out by [89] presents a comparison of three unsupervised methods:  $k$ \_Means [90], Gaussian mixture models (GMM) [91] and HMM. The authors developed their own database from information provided by a series of accelerometers distributed over the users' bodies, resulting in a hit rate of between 75% and 84%. They compared the results of these unsupervised methods with supervised methods, such as SVM, obtaining a hit rate of 95.55%. This comparison shows that supervised learning methods usually tend to be more accurate although less replicable [92]. Another development that works with the accelerometers of wearable sensors is the one carried out by [93]. In this work, the authors designed a method called unsupervised deep learning (DL)-assisted reconstructed coder (UDR-RC). This algorithm is based on feature extraction methods using FFT. The authors used the WISDM database to analyse a total of six activities and obtained a hit rate of around 97.28%. In [94], the authors developed an algorithm known as MaxGap, which is based on the implication of the input signals to the model in the generation of a given output, and the algorithm computes a positive or negative effect depending on the weight it has on the output. To test its veracity, the authors compared the results against an experimental database and a known algorithm, such as HMM. As a result, they obtained a hit rate for MaxGap of 91.4% versus 93.5% for HMM. This reflects the fact that the algorithm converges well, but the HMM still gives better results. Other authors have opted to apply rule-based approaches to recognise activities [95]. In [96],

the authors have developed a deep rule-based method (DRB). A total of 50 activities can be analysed using 1000 images from the UCF50 database. Several authors have opted to use decision-tree-based mechanisms to perform HAR [97]. In [98], the authors have used decision trees to analyse the database “Activity of Daily Living” (ADL) [99] obtaining an accuracy of 88.02% for 8 activities. A summary of all the work presented is given in Table 2.

**Table 2.** Overview of the main HAR unsupervised learning methods.

Development	DL Method	Database	Activities	Accuracy
[78]	FFT	WEKA	5	79.98%
[82]	HMM	Own development	12	89%
[83]	HMM	Own development	5	69%
[85]	VAE	HHAR	9	87%
[89]	k_Means	Own development	12	72.95%
	GMM			75.60%
	HMM			83.89%
[93]	UDR-RC	WISDM	6	97.28%
[94]	MaxGap	Own development	17	91.40%
	HMM			93.50%
[96]	DRB	UCF50	50	82.00%
[98]	Decision trees	ADLs	8	88.02%
Proposed	Proposed	SDHAR-HOME	18	91.68%

The contributions reviewed in Section 2.1 have the advantage of a high hit rate in HAR applications. However, it is costly to bring these models to a real environment due to their low replicability. These models are strict in terms of the training conditions in which they are carried out, giving worse results if the observed conditions are modified (e.g., if it is decided to take the model to another house). In contrast, the papers reviewed in Section 2.2 are more easily replicable and applicable in real life. They have the advantage that they do not require the activities performed by household residents to be labelled at all times, which facilitates implementation in a larger number of households. However, unsupervised models often give worse results. Therefore, the approach proposed in this paper aims to obtain an unsupervised method that provides a high hit rate and can be replicated in households other than the one analysed.

### 3. The Proposed HAR Approach

In this section, a description of the unsupervised model developed to perform activity recognition is given. First, the structure of the database chosen for the study is explained. Finally, the mathematical principle of the approach and its application are described.

#### 3.1. SDHAR-HOME Dataset: Analysis and Description

The database chosen to perform the study is SDHAR-HOME [72]. It was built in a real home where two people live together. This fact represents a challenge to be solved with HMMs, since it is necessary to take into account the activity that each resident is carrying out at all times. This database has the recorded data of 18 activities over a two month period, giving a wide time margin to incorporate different situations and scenarios to deal with the approach developed in this paper. This database has the following technology groups:

1. Non-intrusive sensor network: The database has real-time measurements of the events provided by a sensor network implemented in the house. This network is composed of the following sensors: 8 motion sensors, 8 door contact sensors, 2 temperature and humidity sensors, 11 vibration sensors, 2 power consumption sensors and 2 lighting sensors. This sensor network provides low-power Zigbee signals that are collected

by a central hub, which is responsible for storing the information and managing the devices.

2. Bluetooth beacon triangulation: Each resident wears a smart band, and by using a network of beacons deployed in the house (one beacon in each room), the power of the Bluetooth signal is continuously measured to each of the smart bands in order to locate the user within the home. This information helps to differentiate which user is performing a specific activity.
3. Wearable devices: As mentioned in the previous section, each resident wears a smart band to be able to locate him or her in the house. This smart band is able to provide information linked to the physical activity of each user (e.g., heart rate, calories, steps or data from the device's gyroscopes).

The occurrence of events collected by the sensors is asynchronous, which represents a challenge for the data processing. This database contains discrete signals, such as those provided by a motion sensor, and signals with a temporary inertia typical of temperature or humidity sensors.

During the two month period, there were moments in which the users were not performing any actions, even though some events were provided by the sensors. These events may correspond, for example, to transitions between rooms or periods between two activities that are not close in time. In the dataset, these periods of inactivity are labelled under the activity "Other".

### 3.2. Mathematical Principles and Application

The mathematical principles used for the approach development are based on HMMs. HMM models correspond to an unsupervised learning method commonly used in the field of ML based on dynamic Bayesian networks. This type of model is typically used for speech recognition, natural language processing and image processing [100]. For this work, HMMs have been chosen because of their advantages in terms of replicability and setup effort. This type of model, because it is unsupervised, can be implemented in any home without the need for a training stage, avoiding the need for manual labelling of activities. This facilitates its acceptance in a larger number of homes [101].

HMMs can be divided into two hierarchical levels [102]:

1. Hidden states  $S$ : These correspond to the system variables that are unknown and that are desired to be recognised. In the case being addressed, they correspond to the activities carried out by the users within the household. These hidden states can be represented as follows:

$$S = \{s_1, s_2, s_3, \dots, s_N\}. \quad (1)$$

In (1),  $N$  corresponds to the total number of hidden states of the system and each  $s_i$  refers, in this case, to each of the activities analysed.

2. Observations  $O$ : The sequence of observations corresponds to the observable (and measurable) facts from which information can be extracted from the environment in which the system is located. In the case under consideration, this sequence of observations corresponds to the information provided by the technology with which the household is equipped (e.g., sensors or imagery). The sequence of observations can be represented as follows:

$$O = \{o_1, o_2, o_3, \dots, o_T\}. \quad (2)$$

In (2), each observation  $o_i$  is related to the time at which it occurs.  $T$  corresponds to the total number of observations to be analysed. The sequence of observations  $O$  includes all possible measurements obtained from the environment where the model is located. Thus, the total set of possible observations  $V$  can be represented as follows:

$$V = \{v_1, v_2, v_3, \dots, v_M\}. \quad (3)$$

In (3),  $M$  represents the total number of different signals entering the system and each  $v_i$  represents each of the sensors.

Once the hidden states  $S$  that the Markov network includes and the observations  $O$  that trigger the state changes have been defined, it is necessary to introduce the probability matrices that parameterise the model:

1. State transition matrix  $A$ : This corresponds to the probability matrix of transitions between the different hidden states of the previously defined Markov network. For this reason, it is an  $N \times N$  matrix (4). Moreover, the sum of all transition probabilities in the same row is equal to one (5). The probability of remaining in the same state corresponds to the values on the diagonal, while the probabilities of moving from one hidden state to another correspond to the remaining probabilities.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1N} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2N} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3N} \\ \dots & \dots & \dots & \dots & \dots \\ a_{N1} & a_{N2} & a_{N3} & \dots & a_{NN} \end{bmatrix}. \quad (4)$$

$$\sum_{j=1}^N a_{ij} = 1, \forall i \in [1, N]. \quad (5)$$

2. Emission matrix  $B$ : This corresponds with the probability matrix of change to a certain hidden state depending on the concrete observation entering the system. For this reason, it is a matrix whose dimensions are  $N \times M$  (6).

$$B = \begin{bmatrix} b_1(1) & b_1(2) & b_1(3) & \dots & b_1(M) \\ b_2(1) & b_2(2) & b_2(3) & \dots & b_2(M) \\ b_3(1) & b_3(2) & b_3(3) & \dots & b_3(M) \\ \dots & \dots & \dots & \dots & \dots \\ b_N(1) & b_N(2) & b_N(3) & \dots & b_N(M) \end{bmatrix}. \quad (6)$$

3. Initial probability vector: At the beginning of the execution of a Markov network, the different hidden states must start with an output probability ( $\pi_i$ ). This probability is given by (7).

$$\pi = \{\pi_1, \pi_2, \pi_3, \dots, \pi_N\}. \quad (7)$$

Since the database contains information on two users cohabitating in the home simultaneously, a scenario has been considered where two HMMs are running in parallel (one network for each user). Another aspect to be taken into account is the nature of the information provided by each sensor in the database. For this purpose, the non-intrusive sensors have been categorised into two groups:

1. Event sensors: These sensors provide information about the different events that occur in the house. For example, an event could be the opening of a cupboard, the vibration of a chair or the presence of a user in a specific room. For this reason, the sensors that belong to this set are the following: presence, contact and vibration sensors.
2. Environmental sensors: These sensors provide information about the conditions of the home over time. Therefore, their activation is not enough to establish a transition between activities unless they are accompanied by an action provided by an event sensor. For example, recognising that the TV is powered on upon its energy consumption does not imply that any user is actually watching it at that moment, as he or she may be wandering around the room or in another room. For this reason, it is necessary to complement this energy consumption with an event sensor, such as vibration on the sofa, in order to know that the user is actually watching TV. The sensors that



belong to this group are the following: temperature and humidity, consumption and luminosity sensors.

From this division into two sensor groups, it is possible to say that the event sensors are the sensors that influence the emission matrix  $B$ . On the other hand, for the environmental sensors, it is necessary to create another matrix, called the Environmental Matrix  $C$ :

$$C = \begin{bmatrix} c_1(1) & c_1(2) & c_1(3) & \dots & c_1(E) \\ c_2(1) & c_2(2) & c_2(3) & \dots & c_2(E) \\ c_3(1) & c_3(2) & c_3(3) & \dots & c_3(E) \\ \dots & \dots & \dots & \dots & \dots \\ c_N(1) & c_N(2) & c_N(3) & \dots & c_N(E) \end{bmatrix}. \quad (8)$$

The structure of the  $C$  matrix can be seen in (8). It is an  $N \times E$  matrix, where  $N$  corresponds to the total number of hidden states and  $E$  to the total number of environmental sensors.

Regarding the data provided by the Bluetooth beacons for positioning the users, a study has been carried out on the quality of the signals. Due to the precision of the Bluetooth signal, the signals can be affected by the walls of the house or electromagnetic noise that can occur at any given moment. This has an impact on possible transition moments between the different house rooms without the user making that movement. With all this, the following conclusions can be drawn: it is impossible to make a transition between rooms that are not connected to each other, and it is impossible to change rooms without the movement sensors sending the presence event. Due to these two particularities, a Markov network has been developed to filter out impossible transitions between rooms. In this case, the hidden states correspond to the seven different rooms of the house.

Figure 1 shows a schematic graphic of the SDHAR-HOME house morphology. As can be seen, it is not possible to make a transition between rooms without first passing through the hall. For this reason, in the  $A$  matrix of this Markov network, the probability of transition between rooms that are not interconnected is close to 0 (e.g., bedroom–lounge transition). With respect to the emission matrix  $B$ , the observations  $O$  taken into account are the motion sensors in the rooms and the data provided by the positioning beacons.

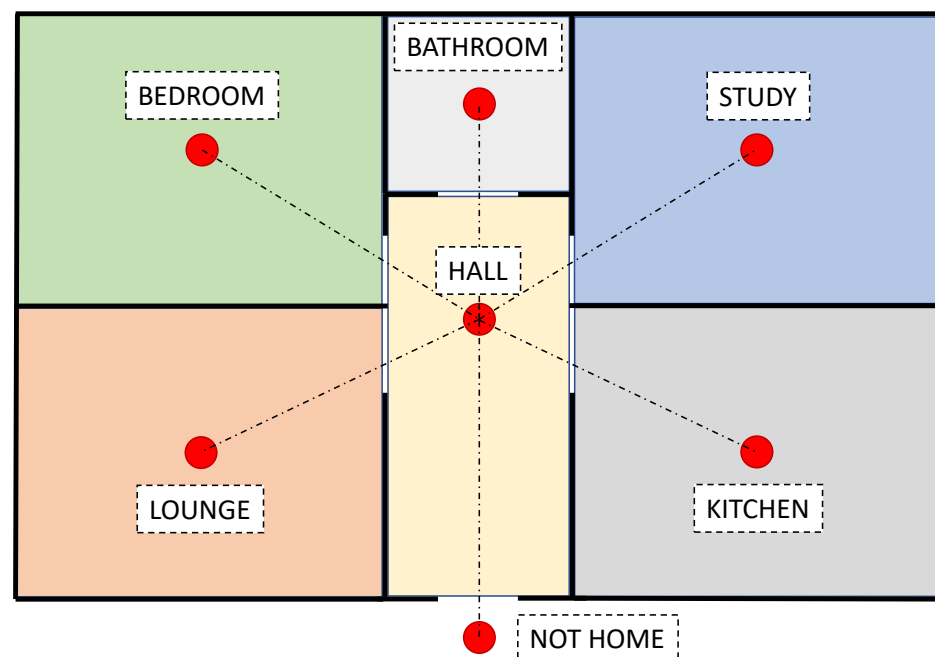


Figure 1. Generic scheme of the house morphology.

Figure 2 shows the structure of the Markov network developed for indoor positioning. Figure 2a shows the possible relationships between the different hidden states and the probabilities associated with the transition matrix  $A$ . In the figure, the nodes correspond to the possible rooms that the system is able to detect. Furthermore, Figure 2b shows the relationship between the chain of hidden states and an example of possible observations, together with the probabilities associated with the emission matrix  $B$ . Motion sensor nodes are shown in yellow, the wristband signals are shown in blue, while room nodes are shown in green.

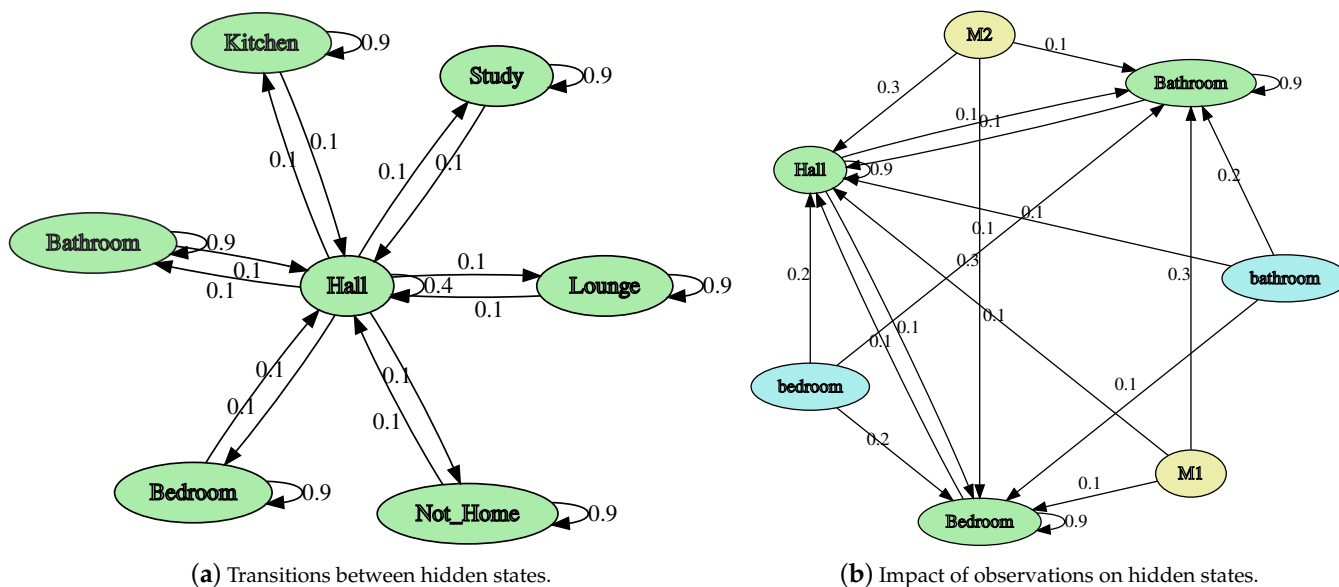


Figure 2. Hidden Markov model for indoor positioning.

Once the correct rooms where the users are when an event occurs in the home are known, a new matrix can be defined that relates the room to the possible activity that is taking place. For example, the probability that a resident is eating in the bathroom is very low. On the other hand, the probability that the resident is sleeping in the bedroom is high. For this reason, it is necessary to create the matrix  $M$ :

$$M = \begin{bmatrix} m_1(1) & m_1(2) & m_1(3) & \dots & m_1(R) \\ m_2(1) & m_2(2) & m_2(3) & \dots & m_2(R) \\ m_3(1) & m_3(2) & m_3(3) & \dots & m_3(R) \\ \dots & \dots & \dots & \dots & \dots \\ m_N(1) & m_N(2) & m_N(3) & \dots & m_N(R) \end{bmatrix}. \tag{9}$$

In (9), the structure of the  $M$  matrix is reflected. This matrix has dimensions  $N \times R$ , where  $R$  is the number of the possible rooms of the house provided by the indoor positioning Markov network.

Figure 3a shows the different relationships between the hidden states of the activity recognition Markov network. In the figure, the nodes correspond to the possible activities that the system is able to detect. Figure 3b shows how the different event sensors affect the Markov network. Sensor nodes are shown in yellow, while activity nodes are shown in green. In both figures, the approximate probabilities defined in the  $A$  and  $B$  matrices are also shown.

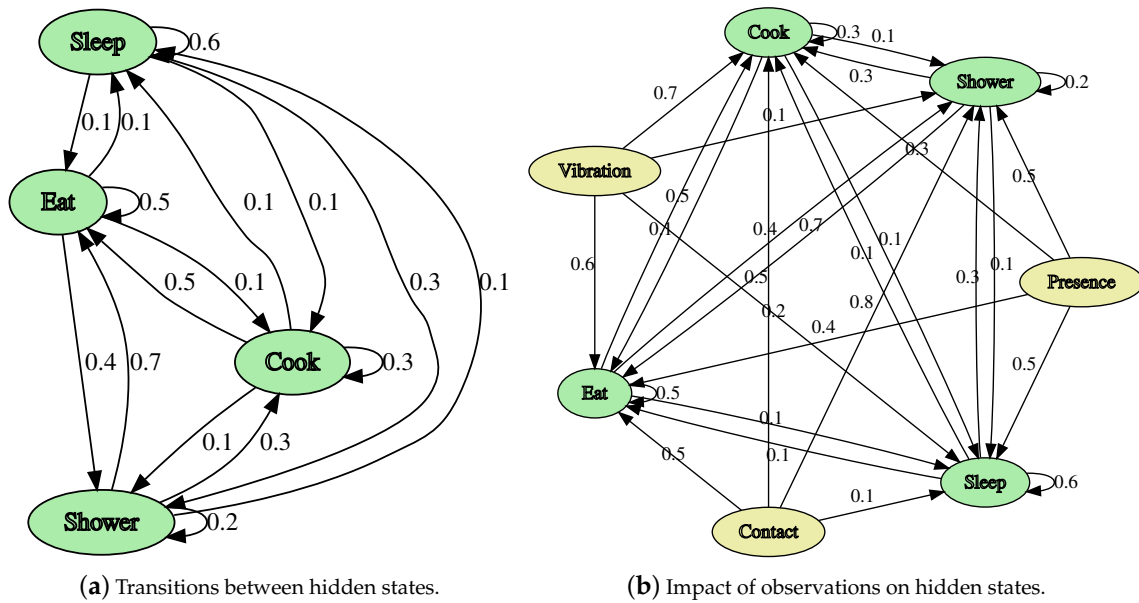


Figure 3. Hidden Markov model for HAR.

Having defined and parametrised the system architecture, it is time to develop and apply the algorithm responsible for obtaining the paths within the system where the highest probability is obtained. We have chosen a forward–backward approach [103].

1. Forward algorithm: The forward algorithm is in charge of calculating the different activities and rooms successively each time a new event sensor record or beacon position arrives. Therefore, it has a progression in favour of the timeline. Hidden state probabilities are calculated by combining propagation stages (applying the  $A$  matrix) and update stages (applying the  $B$  matrix) in a similar way to the behaviour of Bayesian filters [104]. The algorithm for the user’s location system is the following (see Algorithm 1):

---

**Algorithm 1** Location system: forward algorithm

---

**Input:** HMM and Observations sequence  $O = o_1 \dots o_T$

**Output:**  $\delta_t(s_i) = P(X_t = s_i, O) \forall s_i$

- 1: **if**  $t = 1$  **then**
  - 2:  $\delta_1(s_i) = b_i(o_1) \cdot \pi_i, i \in [1, n]$
  - 3: **else**
  - 4: **for**  $k = 2$  **to**  $t$  **do**
  - 5: **for**  $j = 1$  **to**  $n$  **do**
  - 6:  $\delta_k(s_j) = b_j(o_k) \sum_{i=1}^n (a_{ij} \cdot \delta_{k-1}(s_i))$
  - 7: **end for**
  - 8: **end for**
  - 9: **end if**
  - 10: **return**  $\delta_t(s_i), i \in [1, n]$
- 

In this and the following algorithms, the parameter  $n$  corresponds to the total number of different hidden states. In addition, the parameter  $t$  refers to the instant at which the algorithm is applied. This algorithm is the one used for the Markov network that is responsible for locating users within the home. In contrast, for the Markov network that performs activity recognition, a variant has been created to implement the environmental sensors. The variant is as follows (see Algorithm 2):

**Algorithm 2** HAR: forward algorithm with environmental sensors**Input:** HMM, Observations sequence  $O = o_1 \dots o_T$  and Environment sequence  $E = e_1 \dots e_T$ **Output:**  $\alpha_t(s_i) = P(X_t = s_i, O, E) \forall s_i$ 

```

1: if  $t = 1$  then
2:    $\alpha_1(s_i) = b_i(o_1) \cdot \pi_i, i \in [1, n]$ 
3: else
4:   for  $k = 2$  to  $t$  do
5:     for  $j = 1$  to  $n$  do
6:        $\alpha_k(s_j) = b_j(o_k) \sum_{i=1}^n (a_{ij} \cdot \alpha_{k-1}(s_i))$ 
7:       if  $e_k \neq 0$  then
8:          $\alpha_k(s_j) = c_j(e_k) \cdot \alpha_k(s_j)$ 
9:       end if
10:    end for
11:  end for
12: end if
13: return  $\alpha_t(s_i), i \in [1, n]$ 

```

2. Backward algorithm: The backward algorithm is in charge of calculating the different  $\beta$  as successive sensor events arrive at the system. For this reason, the flow of this algorithm goes against the timeline, reinforcing the value provided by the forward algorithm with information from subsequent events. A total of  $k = 2$  subsequent events has been chosen so that the system does not become too slow. This part of the proposed system is shown in Algorithm 3.

**Algorithm 3** Backward algorithm**Input:** HMM and Observations sequence  $O = o_1 \dots o_T$  $k \in [1, t]$ **Output:**  $\beta_k(s_i) = P(o_{k+1}, \dots, o_t | X_k = s_i) \forall s_i$ 

```

1: if  $t = k$  then
2:    $\beta_t(s_i) = 1, i \in [1, n]$ 
3: else
4:   for  $r = t - 1$  to  $k$  do
5:     for  $j = 1$  to  $n$  do
6:        $\beta_r(s_j) = \sum_{i=1}^n b_i(o_{r+1}) \beta_{r+1}(s_i) a_{ji}$ 
7:     end for
8:   end for
9: end if
10: return  $\beta_k(s_i), i \in [1, n]$ 

```

A study of the activities performed by both users throughout the database has been carried out. From this study, it has been deduced that there are activities whose probability of being executed jointly is higher than the rest. For example, when user 1 is sleeping, it is very likely that user 2 is also sleeping. Another example could happen with the activity of eating, since both users are used to eating at the same time. For this reason, an implemented  $R$  matrix aims to relate the outputs of both users' Markov networks in order to reinforce the activities that are performed jointly (see (10)).

$$R = \begin{bmatrix} r_1(1) & r_1(2) & r_1(3) & \dots & r_1(N_2) \\ r_2(1) & r_2(2) & r_2(3) & \dots & r_2(N_2) \\ r_3(1) & r_3(2) & r_3(3) & \dots & r_3(N_2) \\ \dots & \dots & \dots & \dots & \dots \\ r_{N_1}(1) & r_{N_1}(2) & r_{N_1}(3) & \dots & r_{N_1}(N_2) \end{bmatrix}. \quad (10)$$

$$\gamma(U_1) = R \cdot (\alpha_i(U_2) \cdot \beta_i(U_2) \cdot \delta_i(U_2)). \quad (11)$$

Equation (11) shows the way the  $\gamma$  parameter is calculated, which is used to collect the information from the output of the other resident’s network and then to apply it in its own inference engine.

Finally, once the parameters  $\alpha, \beta, \delta$  and  $\gamma$  of each user are known, the activity that each user performs at a specific time  $t$  can be obtained as:

$$ACT(s_i) = \alpha(s_i) \cdot \beta(s_i) \cdot \gamma(s_i) \cdot \delta(s_i). \tag{12}$$

In (12), the inference process is shown, from which a vector of probabilities  $ACT(s_i)$  is obtained for each of the possible system activities. Once this vector is known, it is only necessary to select the activity with the highest probability to know the activity being performed by the chosen user.

The overall HAR system architecture is summarised in Figure 4. This diagram shows the two HMMs implemented in the proposed method: on the left (HMM HAR) is the model that recognises the activity and on the right (HMM LOCATION) is the model that provides the room in which the user is located. It can be seen that HMM HAR has been implemented using a combination of forward and backward algorithms, while HMM LOCATION has only a forward stage. It can also be deduced that each algorithm is accompanied by its corresponding probability distribution matrices. The inputs of the HMM HAR forward stage are the activity distribution for the previous time  $t - 1$ , the action provided by an event sensor and the home conditions provided by the environmental sensors. In contrast, the backward stage relies only on the signals from the subsequent event sensors  $t + 1$  and  $t + 2$ . The inputs of the HMM LOCATION forward algorithm are the signals provided by the Bluetooth beacons and the events from the motion sensors (together with the probability distribution of the user’s position in the house at the previous time  $t - 1$ ). All outputs are multiplied in the inference block, together with the influence of the other user’s activity, to produce the most likely activity at time  $t$ .

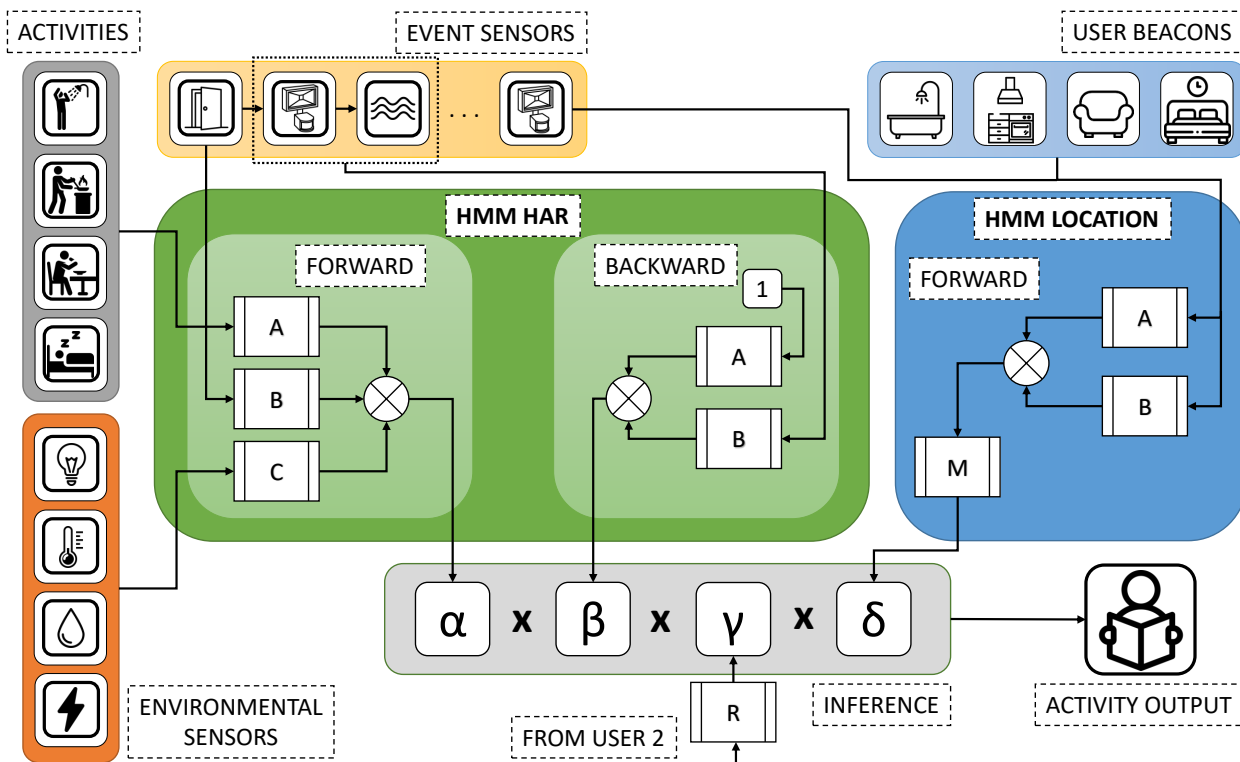
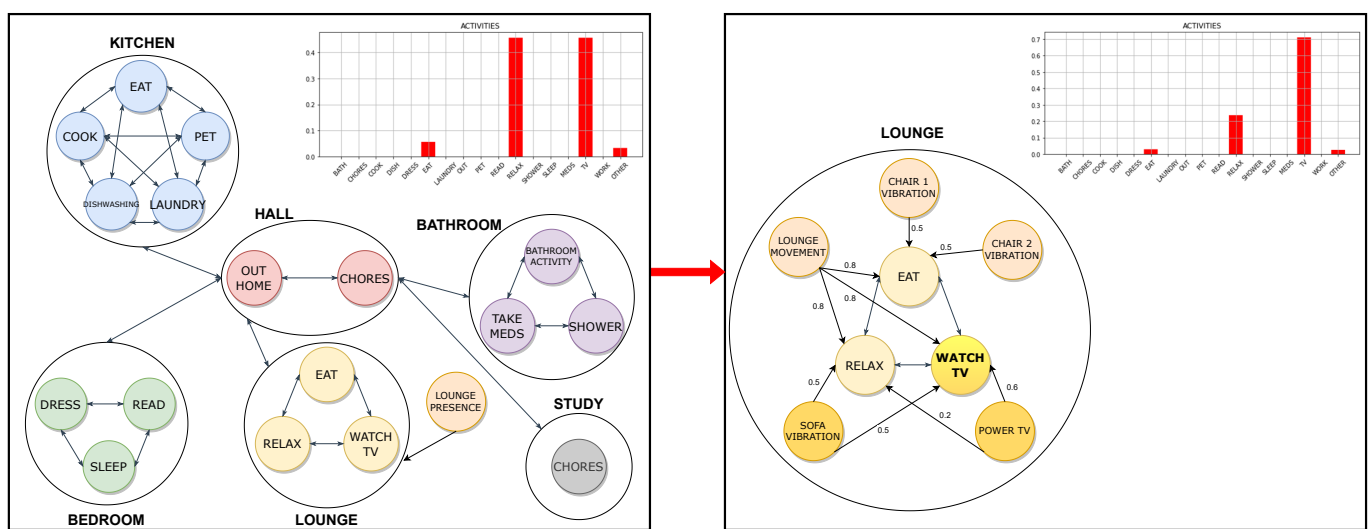


Figure 4. Generic scheme of the solution for HAR.

### 4. Experiments

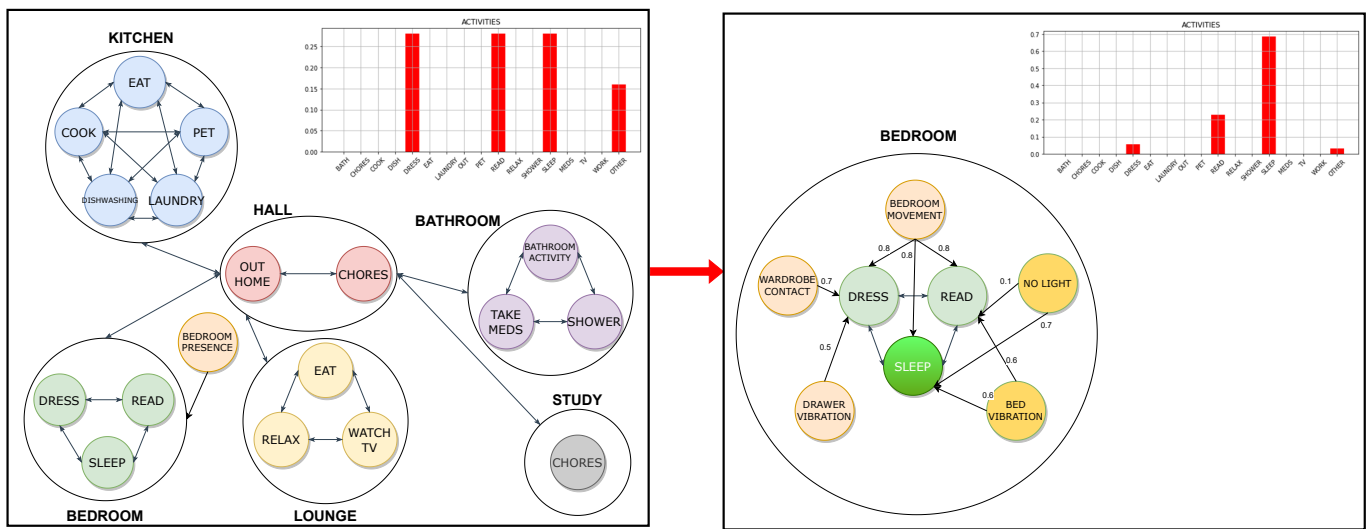
In order to evaluate the efficiency and accuracy of the unsupervised system presented in Section 3.2, data from the SDHAR-HOME database have been used, in conjunction with the system, in order to obtain the corresponding activities and compare them with the real activities from the database. Since the frequencies of the predicted and real activities are different (as the system provides predictions every time data arrive from the event sensors), an output sampling module has been used to compare them every second. Concerning error, it must be considered that a short time delay in recognising an activity is acceptable. There are activities with a certain temporal inertia, such as the “Shower” activity. When one of the users labels “Shower”, he/she takes some time to get into the shower and turn on the water, and it takes some time for the humidity in the room to increase. This time interval is not considered an error if the system has finally recognised the “Shower” activity with some delay. This happens with other activities, as the labelling of an activity does not coincide exactly with the start of the activity. For this reason, a module has been developed to analyse the time interval that the system takes to recognise the activity. If this time interval is less than 5 min, the prediction is considered correct. “Make Simple Food” and “Cook” have been grouped together as the sensors involved in the recognition are the same. The computational total time duration is 70 s to process the entire database. All activities labelled in the SDHAR-HOME database have been evaluated (16 activities in total). All the tests and experiments were conducted on an Intel(R) Core(TM) i7-10875H CPU @ 2.30 GHz/16 Gb with NVIDIA GeForce RTX 2060 GPU.

In Figure 5, two examples of the system’s operation can be observed. For instance, in Figure 5a, presence is detected within the living room. With this information, the probabilities of carrying out any possible action in the living room become equal. Shortly thereafter, presence on the sofa is detected and the television is turned on, leading the system to deduce that the user is currently watching TV. A similar scenario unfolds in the example in Figure 5b. The motion sensor detects presence in the bedroom, resulting in the equalization of probabilities for all possible bedroom activities. Upon detecting that the light is off and someone is lying in bed, the system infers that the user is sleeping.



(a) Watching TV activity detection process.

Figure 5. Cont.



(b) Sleeping activity detection process.

Figure 5. Examples of activity detection based on sensor signals.

In Figures 6 and 7, the confusion matrices of the hidden Markov networks can be observed for each user. In these confusion matrices, the percentage and distribution of the system’s success can be observed, since the equivalence between the real activities and the activities predicted by the system can be seen. The hit rates can be observed along the diagonal of the matrices, while the off-diagonal probabilities correspond to system failures.

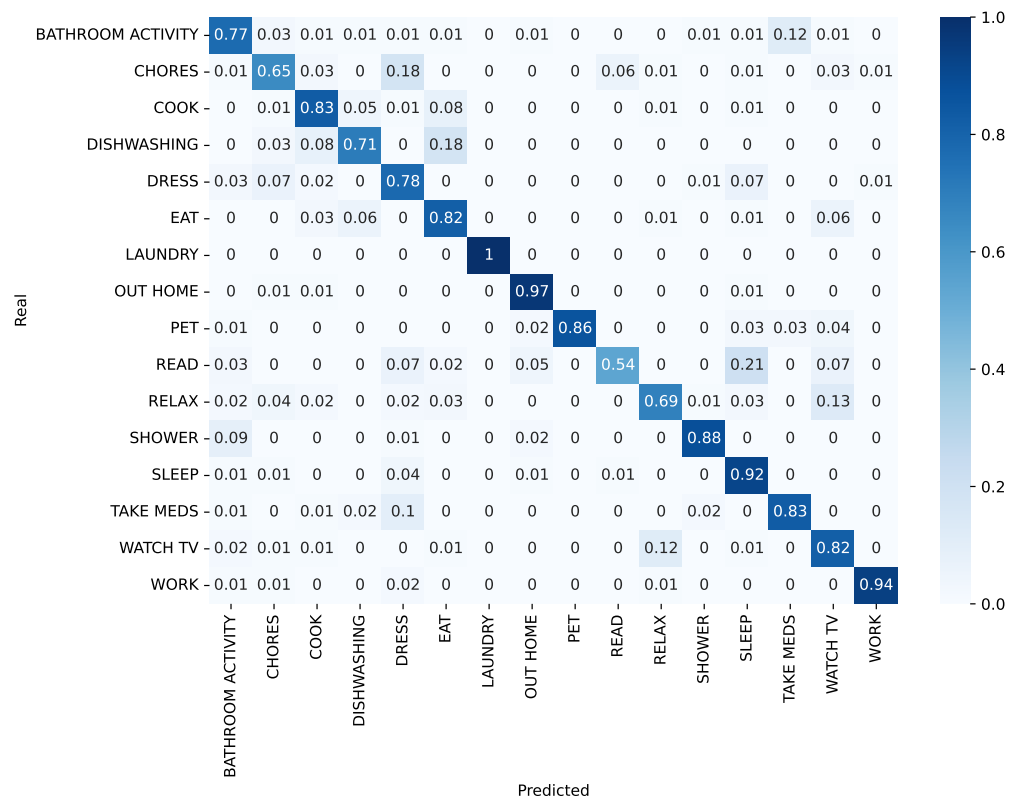


Figure 6. HMM User 1: Confusion matrix.

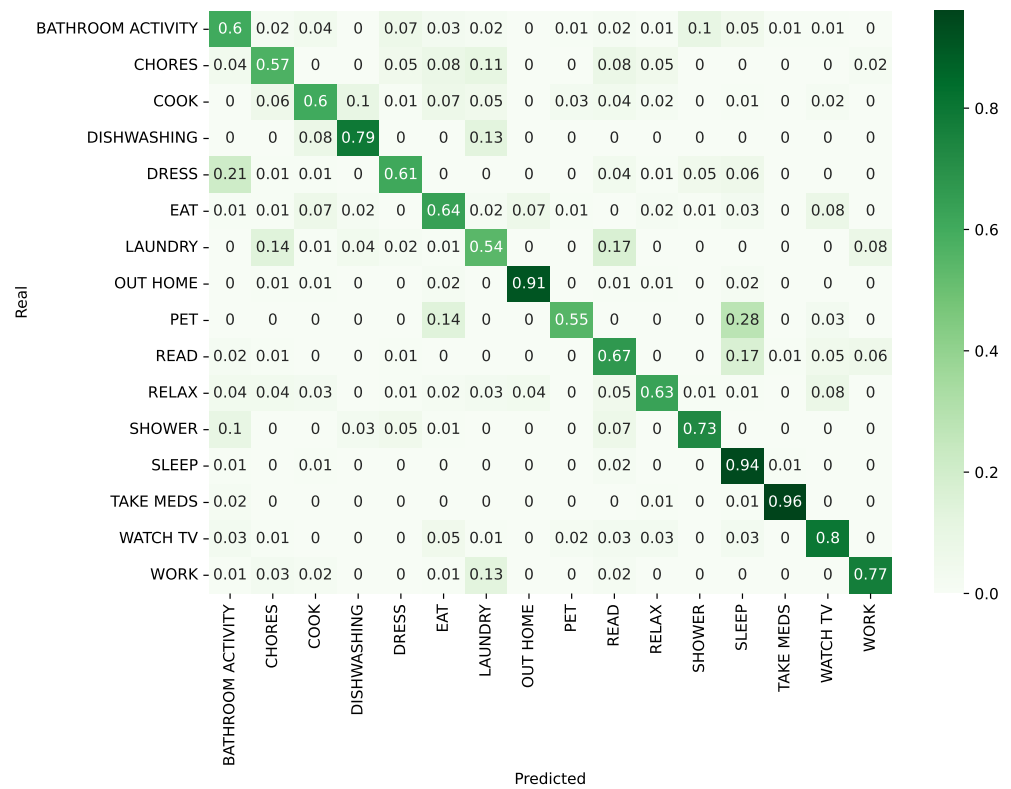


Figure 7. HMM User 2: Confusion matrix.

From the confusion matrices, the most accurate activities of the system are deduced. For example, the “Sleep” activity or the “Out of Home” activity have the highest accuracy rate (both are over 88% accurate). On the other hand, the system may confuse different pairs of activities. For example, the “Cook” activity is focused on the use of the glass–ceramic hob, and this behaviour is analysed by a temperature and humidity sensor. These two variables have a high inertia, so the activity may be sustained longer than it should be. Another pair of activities that can be faulty is “Read” with “Sleep”. Both activities take place in the bed, the only difference is whether the light is on or off. Finally, the “Chores” activity can fail because it is a very generic activity that takes place all over the house. This results in a large number of sensor events, which leads to a decrease in the probability of “Chores” due to the forward propagation effect.

In Table 3, the average duration of each of the activities in the database are shown. All durations in the table are reported in seconds (s). From this table, a substantial variability in duration between each activity can be deduced. In addition, activities with longer durations have a greater number of events provided by the sensors. During the experimentation, the HAR approach obtained an average response time of around 6 min. Although the forwarding stage is immediate, as it takes place at the moment an event arrives from a sensor, the backtracking effect needs to rely on nearby future signals to infer activity. Moreover, the experimentation was carried out continuously for two months.

In order to evaluate the robustness of the system, the following set of metrics has been analysed from the confusion matrices: precision, recall, F1-score and accuracy [105].

$$Precision = \frac{TP}{TP + FP} \tag{13}$$

$$Recall = \frac{TP}{TP + FN} \tag{14}$$



$$F1Score = \frac{2}{\frac{1}{Recall} + \frac{1}{Precision}}. \quad (15)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}. \quad (16)$$

From (13)–(16), the calculation methods of the previously mentioned metrics can be observed. *TP* is true positive, *FP* is false positive, *TN* is true negative and *FN* is false negative.

**Table 3.** Average duration of SDHAR-HOME activities for both users.

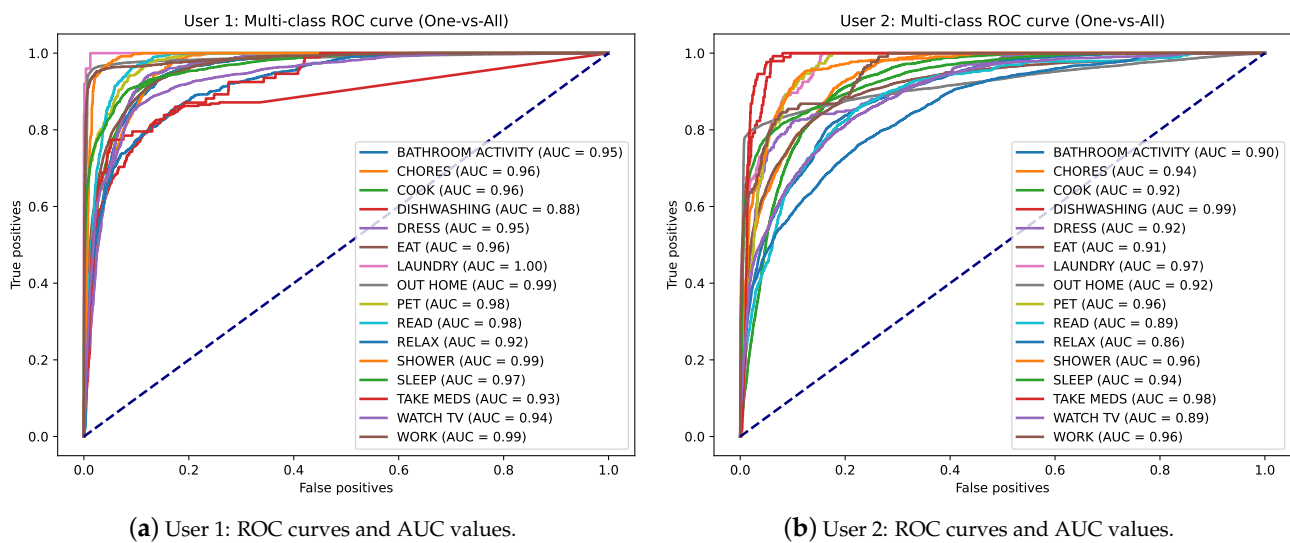
	Bath. Act.	Chores	Cook	Dish	Dress	Eat	Laundry	Out of Home
User 1	505	1594	630	723	268	1649	224	24,003
User 2	914	1887	604	631	607	2213	832	27,261
	Pet	Read	Relax	Shower	Sleep	Take Meds	Watch TV	Work
User 1	146	1997	1960	634	25,268	103	4319	5202
User 2	304	8924	3788	1602	28,907	60	4262	4244

Table 4 shows the precision, recall and F1-score values for each of the SDHAR-HOME activities for each user.

**Table 4.** Summary table of results by activity (User 1–User 2).

Activity	Precision	Recall	F1-Score
Bathroom Activity	0.81–0.63	0.77–0.6	0.79–0.61
Chores	0.28–0.25	0.65–0.57	0.39–0.35
Cook	0.68–0.37	0.83–0.6	0.75–0.46
Dishwashing	0.24–0.15	0.71–0.79	0.36–0.25
Dress	0.12–0.19	0.78–0.61	0.21–0.29
Eat	0.91–0.65	0.82–0.64	0.86–0.64
Laundry	0.85–0.11	1.0–0.54	0.92–0.18
Out of Home	0.99–0.99	0.97–0.91	0.98–0.95
Pet	0.82–0.11	0.86–0.55	0.84–0.18
Read	0.64–0.58	0.54–0.67	0.59–0.62
Relax	0.42–0.76	0.69–0.63	0.52–0.69
Shower	0.87–0.47	0.88–0.73	0.87–0.57
Sleep	0.96–0.94	0.92–0.94	0.94–0.94
Take Meds	0.09–0.09	0.83–0.96	0.16–0.16
Watch TV	0.89–0.79	0.82–0.8	0.85–0.79
Work	0.99–0.57	0.94–0.77	0.96–0.66
Accuracy			0.92–0.87
Macro avg.	0.66–0.48	0.81–0.71	0.69–0.52
Weighted avg.	0.94–0.9	0.92–0.87	0.93–0.88

Figure 8 shows the different receiver operating characteristic (ROC) [106] curves obtained from the experimentation stage for both users. In both figures, the ROC curves for all possible activities can be observed. Furthermore, the value of the area under the curve (AUC) [107] for each of the activities can be checked. The ROC curves and the AUC value represent the sensitivity of the system to false detections, showing the degree of randomness of the predictions. From the AUC values, it can be concluded that the system is robust as the majority of activities are over 90%.



**Figure 8.** One-vs.-All ROC curves resulting from the experimentation stage and AUC values.

## 5. Discussion

From Table 4, it can be deduced that the precision of the Markov network of user 1 is 91.68%, while the precision of user 2 is slightly lower with a value of 86.78%. This accuracy difference may be due to a few different factors. For example, it may be due to the fact that each user has made a different labelling of the actual activity. It may also be due to the adjustment of the system matrices ( $A$ ,  $B$ ,  $C$ ,  $M$  and  $R$ ) on which the probability apportionment is based. The adjustment of the system matrices can be performed easily with a little prior knowledge from the sensors that affect each activity. For example, the “Take Meds” activity is closely linked to the medicine drawer. Therefore, the contact sensor located in this drawer must have a high probability value within the  $B$  matrix. Another example could be the humidity sensor in the bathroom, whose measurement is linked to the “Shower” activity. Therefore, this humidity sensor must have a high probability value within the matrix  $C$ . There is a possibility that both users are in the same room. In addition, both may be doing the same activity (e.g., sleeping) or different activities (e.g., user 1 sleeping and user 2 reading). In the second case, the system cannot infer which activity each user is doing, as there is not enough information available to discriminate which of the two users is performing the action.

The possible weaknesses in the HAR system could be due to the fact that the system itself is not able to distinguish which of the two users is performing the action, even though the activity is being recognised satisfactorily. For this reason, a prediction metric has been extracted regardless of the user to whom the activity is granted. The system has a recognition capability of 95.51%. This metric indicates that the system is able to recognise activities with good accuracy but may fail to identify the user.

From the results obtained, the following deductions can be drawn:

1. Unsupervised models based on HMMs can be used to process time series data provided by discrete event and environmental sensors together with indoor positioning signals.
2. The results obtained with our HMM-based model are close to results obtained using supervised methods, such as RNN, LSTM or GRU [72].
3. The system obtained in the present paper is highly generalist, since the internal parameters that determine the model can be freely modified to analyse another household, with other residents and in other circumstances than those used in the present experimentation.
4. False positives obtained during the testing phase are mainly due to failures between activities of a similar nature or failures between activities within the same room.

5. The use of beacon installation for indoor positioning is unavoidable in cases where two or more people cohabit within the same house, as the system needs to know where the residents are located for probability allocation. In the case where there is only one person, the position could be determined from the signals of the PIR motion sensors.
6. Results obtained using an unsupervised method such as HMMs are high compared to the previous work discussed in Section 2.2 using similar unsupervised methods.

The code containing the HMM developed in this paper and the experimental method is available at the following link: <https://github.com/raugom13/Unsupervised-Human-Activity-Recognition-approach-in-Multi-User-Households>, accessed on 24 November 2023.

## 6. Conclusions

In this paper, a novel unsupervised method has been proposed that is able to recognise in real time the activities that two people perform simultaneously in the same house. This method is based on HMMs, with a number of relevant adaptations to improve its performance and adaptability to several simultaneous users. The scenario that has been analysed corresponds to the public database SDHAR-HOME, which provides information on the activities performed by two users in parallel to the collection of information from three technological groups: non-intrusive sensor events, indoor positioning using Bluetooth beacons and physiological data provided by activity bracelets. The method developed in this paper corresponds to a probabilistic method that takes into account the events captured by the non-intrusive sensor network and the positions of both users inside the house.

The structure of the developed method consists of the following two subsystems: an HMM that is responsible for reliably providing the room in which each person is located during the whole experimentation from the house's motion sensors and beacon data, together with another HMM that takes into account all the events happening inside the house (using the forward-backward algorithm), the room in which the user is located and the task that the other resident of the house is performing. The output of this method is compared second by second with the corresponding real activity in order to establish an accuracy metric to report on the accuracy and reliability of the system. The architecture of the system is easily replicable for a larger number of users, as it would be necessary to incorporate as many Markov models as users in the house.

The prediction system is able to recognise a total of 16 activities of daily living with an overall accuracy of 91.68% for user 1 and 86.78% for user 2. These accuracy values are considered satisfactory compared to the results obtained using RNNs in previous studies for the same SDHAR-HOME database [72]. Moreover, these results are high in comparison with the values obtained by other authors using unsupervised learning methods with other databases (see Section 2.2). The drawback of HMMs lies a priori in their effectiveness compared to supervised methods, such as neural networks. Supervised methods may perform more effective and optimised learning for the input data set [108]. However, our experiments have shown that our method provides highly satisfactory results and improves the replicability of the system.

The system proposed in this paper is useful for the non-intrusive monitoring of elderly people living alone in order to establish an alarm mechanism if there are conditions or situations that are dangerous to their health. Furthermore, as an unsupervised method, the system is replicable for use in other homes with different residents and behavioural patterns. This is an improvement on supervised methods that use DL approaches, such as RNNs, as these systems are more strict and rigid, and their internal model cannot be manipulated. An advantage of the model presented in this paper is that it allows new residents to be incorporated into the model, as the internal probability distribution is fully controlled.

As future research, it is proposed to infer risk situations, such as lack of nutrition or hygiene, with greater accuracy, as well as to detect anomalous behaviours, such as night wandering or an excessively sedentary life. A further proposal is to reduce the time range

taken into account in the experimentation phase (see Section 4), as it would be interesting to reduce the 5 min that have been used to increase the speed of the model in the case of certain events (e.g., falls). Finally, another suggested future line corresponds to using the output information of the model to elaborate a set of rules so as to determine which actions are recommended based on the behaviour patterns of the residents (e.g., by means of supervised learning by an expert).

**Author Contributions:** R.G.-R., J.D.-D., E.Z. and J.G.-G.-B. conceived and designed the experiments; R.G.-R., J.D.-D., E.Z. and J.G.-G.-B. performed the experiments, analysed the data and wrote the paper. All authors have read and agreed to the published version of the manuscript.

**Funding:** The research leading to these results has received funding from projects ROSOGAR PID2021-123020OB-I00 funded by MCIN/AEI/10.13039/501100011033/FEDER, UE, and EIAROB funded by Consejería de Familia of the Junta de Castilla y León—Next Generation EU.

**Data Availability Statement:** Publicly available data sets were analysed in this study. These data can be found here: <https://github.com/raugom13/SDHAR-HOME-A-Sensor-Dataset-for-Human-Activity-Recognition-at-Home>, accessed on 24 November 2023.

**Conflicts of Interest:** The authors declare that there are no conflict of interest regarding the publication of this manuscript.

## Abbreviations

The following abbreviations are used in this manuscript:

DL	Deep Learning
HAR	Human Activity Recognition
IoT	Internet of Things
ML	Machine Learning
CNN	Convolutional Neural Network
RNN	Recurrent Neural Network
HMM	Hidden Markov Model
BLS	Broad Learning System
LSTM	Long Short-Term Memory
GRU	Gated Recurrent Unit
SVM	Support Vector Machine
SGN	Semantics-Guided Neural Network
GCN	Graph Convolutional Network
FFT	Fast Fourier Transform
VAE	Variational Autoencoder
DRB	Deep Rule-Based
ADL	Activity of Daily Living
ROC	Receiver Operating Characteristic
AUC	Area Under the Curve

## References

1. Li, Q.; Gravina, R.; Li, Y.; Alsamhi, S.H.; Sun, F.; Fortino, G. Multi-user activity recognition: Challenges and opportunities. *Inf. Fusion* **2020**, *63*, 121–135. [[CrossRef](#)]
2. Dhiman, C.; Vishwakarma, D.K. A review of state-of-the-art techniques for abnormal human activity recognition. *Eng. Appl. Artif. Intell.* **2019**, *77*, 21–45. [[CrossRef](#)]
3. Jobanputra, C.; Bavishi, J.; Doshi, N. Human Activity Recognition: A Survey. *Procedia Comput. Sci.* **2019**, *155*, 698–703. [[CrossRef](#)]
4. Wan, S.; Qi, L.; Xu, X.; Tong, C.; Gu, Z. Deep learning models for real-time human activity recognition with smartphones. *Mob. Netw. Appl.* **2020**, *25*, 743–755. [[CrossRef](#)]
5. Kulsoom, F.; Narejo, S.; Mehmood, Z.; Chaudhry, H.N.; Butt, A.; Bashir, A.K. A review of machine learning-based human activity recognition for diverse applications. In *Neural Computing and Applications*; Springer: Berlin/Heidelberg, Germany, 2022; pp. 1–36.
6. Xia, K.; Huang, J.; Wang, H. LSTM-CNN architecture for human activity recognition. *IEEE Access* **2020**, *8*, 56855–56866. [[CrossRef](#)]
7. Tun, S.Y.Y.; Madanian, S.; Mirza, F. Internet of things (IoT) applications for elderly care: A reflective review. *Aging Clin. Exp. Res.* **2021**, *33*, 855–867. [[CrossRef](#)] [[PubMed](#)]

8. Lentzas, A.; Vrakas, D. Non-intrusive human activity recognition and abnormal behavior detection on elderly people: A review. *Artif. Intell. Rev.* **2020**, *53*, 1975–2021. [[CrossRef](#)]
9. Erickson, S.R.; Williams, B.C.; Gruppen, L.D. Relationship between symptoms and health-related quality of life in patients treated for hypertension. *Pharmacother. J. Hum. Pharmacol. Drug Ther.* **2004**, *24*, 344–350. [[CrossRef](#)]
10. Bhattacharya, D.; Sharma, D.; Kim, W.; Ijaz, M.F.; Singh, P.K. Ensem-HAR: An ensemble deep learning model for smartphone sensor-based human activity recognition for measurement of elderly health monitoring. *Biosensors* **2022**, *12*, 393. [[CrossRef](#)]
11. Sun, H.; Chen, Y. Real-Time Elderly Monitoring for Senior Safety by Lightweight Human Action Recognition. In Proceedings of the 2022 IEEE 16th International Symposium on Medical Information and Communication Technology (ISMICT), Lincoln, NE, USA, 2–4 May 2022; pp. 1–6.
12. Gudur, G.K.; Sundaramoorthy, P.; Umaashankar, V. Activeharnet: Towards on-device deep bayesian active learning for human activity recognition. In Proceedings of the 3rd International Workshop on Deep Learning for Mobile Systems and Applications, Seoul, Korea, 19 June 2019; pp. 7–12.
13. Shalaby, E.; ElShennawy, N.; Sarhan, A. Utilizing deep learning models in CSI-based human activity recognition. In *Neural Computing and Applications*; Springer: Berlin/Heidelberg, Germany, 2022; pp. 1–18.
14. Zimmermann, L.C. Elderly Activity Recognition Using Smartphones and Wearable Devices. Ph.D. Thesis, Universidade de São Paulo, São Paulo, Brazil, 2019.
15. Subasi, A.; Fllatah, A.; Alzobidi, K.; Brahimi, T.; Sarirete, A. Smartphone-based human activity recognition using bagging and boosting. *Procedia Comput. Sci.* **2019**, *163*, 54–61. [[CrossRef](#)]
16. Demrozi, F.; Turetta, C.; Pravadelli, G. B-HAR: An open-source baseline framework for in depth study of human activity recognition datasets and workflows. *arXiv* **2021**, arXiv:2101.10870.
17. Bibbò, L.; Carotenuto, R.; Della Corte, F. An Overview of Indoor Localization System for Human Activity Recognition (HAR) in Healthcare. *Sensors* **2022**, *22*, 8119. [[CrossRef](#)] [[PubMed](#)]
18. Muangprathub, J.; Sriwichian, A.; Wanichsombat, A.; Kajornkasirat, S.; Nillaor, P.; Boonjing, V. A novel elderly tracking system using machine learning to classify signals from mobile and wearable sensors. *Int. J. Environ. Res. Public Health* **2021**, *18*, 12652. [[CrossRef](#)] [[PubMed](#)]
19. Li, X.; He, Y.; Fioranelli, F.; Jing, X. Semisupervised human activity recognition with radar micro-Doppler signatures. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 1–12. [[CrossRef](#)]
20. Popescu, A.C.; Mocanu, I.; Cramariuc, B. Fusion mechanisms for human activity recognition using automated machine learning. *IEEE Access* **2020**, *8*, 143996–144014. [[CrossRef](#)]
21. Zhou, X.; Liang, W.; Kevin, I.; Wang, K.; Wang, H.; Yang, L.T.; Jin, Q. Deep-learning-enhanced human activity recognition for Internet of healthcare things. *IEEE Internet Things J.* **2020**, *7*, 6429–6438. [[CrossRef](#)]
22. Franco, A.; Magnani, A.; Maio, D. A multimodal approach for human activity recognition based on skeleton and RGB data. *Pattern Recognit. Lett.* **2020**, *131*, 293–299. [[CrossRef](#)]
23. Ke, S.R.; Thuc, H.L.U.; Lee, Y.J.; Hwang, J.N.; Yoo, J.H.; Choi, K.H. A review on video-based human activity recognition. *Computers* **2013**, *2*, 88–131. [[CrossRef](#)]
24. Aggarwal, J.; Xia, L. Human activity recognition from 3D data: A review. *Pattern Recognit. Lett.* **2014**, *48*, 70–80. [[CrossRef](#)]
25. Dang, L.M.; Min, K.; Wang, H.; Piran, M.J.; Lee, C.H.; Moon, H. Sensor-based and vision-based human activity recognition: A comprehensive survey. *Pattern Recognit.* **2020**, *108*, 107561. [[CrossRef](#)]
26. San-Segundo, R.; Blunck, H.; Moreno-Pimentel, J.; Stisen, A.; Gil-Martín, M. Robust Human Activity Recognition using smartwatches and smartphones. *Eng. Appl. Artif. Intell.* **2018**, *72*, 190–202. [[CrossRef](#)]
27. Janidarman, M.; Roshan Fekr, A.; Radecka, K.; Zilic, Z. A comprehensive analysis on wearable acceleration sensors in human activity recognition. *Sensors* **2017**, *17*, 529. [[CrossRef](#)]
28. De-La-Hoz-Franco, E.; Ariza-Colpas, P.; Quero, J.M.; Espinilla, M. Sensor-based datasets for human activity recognition—a systematic review of literature. *IEEE Access* **2018**, *6*, 59192–59210. [[CrossRef](#)]
29. Wang, A.; Chen, G.; Yang, J.; Zhao, S.; Chang, C.Y. A comparative study on human activity recognition using inertial sensors in a smartphone. *IEEE Sens. J.* **2016**, *16*, 4566–4578. [[CrossRef](#)]
30. Bi, S.; Hu, Z.; Zhao, M.; Zhang, H.; Di, J.; Sun, Z. Continuous frame motion sensitive self-supervised collaborative network for video representation learning. *Adv. Eng. Inform.* **2023**, *56*, 101941. [[CrossRef](#)]
31. Ann, O.C.; Theng, L.B. Human activity recognition: A review. In Proceedings of the 2014 IEEE International Conference on Control System, Computing and Engineering (ICCSCE 2014), Penang, Malaysia, 28–30 November 2014; pp. 389–393.
32. Singh, Y.; Bhatia, P.K.; Sangwan, O. A review of studies on machine learning techniques. *Int. J. Comput. Sci. Secur.* **2007**, *1*, 70–84.
33. Pramanik, R.; Sikdar, R.; Sarkar, R. Transformer-based deep reverse attention network for multi-sensory human activity recognition. *Eng. Appl. Artif. Intell.* **2023**, *122*, 106150. [[CrossRef](#)]
34. Xu, C.; Chai, D.; He, J.; Zhang, X.; Duan, S. InnoHAR: A deep neural network for complex human activity recognition. *IEEE Access* **2019**, *7*, 9893–9902. [[CrossRef](#)]
35. Liu, T.; Zheng, H.; Zheng, P.; Bao, J.; Wang, J.; Liu, X.; Yang, C. An expert knowledge-empowered CNN approach for welding radiographic image recognition. *Adv. Eng. Inform.* **2023**, *56*, 101963. [[CrossRef](#)]
36. Hibat-Allah, M.; Ganahl, M.; Hayward, L.E.; Melko, R.G.; Carrasquilla, J. Recurrent neural network wave functions. *Phys. Rev. Res.* **2020**, *2*, 023358. [[CrossRef](#)]

37. Li, M.; Zhang, W.; Hu, B.; Kang, J.; Wang, Y.; Lu, S. Automatic assessment of depression and anxiety through encoding pupil-wave from HCI in VR scenes. *ACM Trans. Multimed. Comput. Commun. Appl.* **2023**, *20*, 1–22. [[CrossRef](#)]
38. Zhang, H.; Fritts, J.E.; Goldman, S.A. Image segmentation evaluation: A survey of unsupervised methods. *Comput. Vis. Image Underst.* **2008**, *110*, 260–280. [[CrossRef](#)]
39. Manouchehri, N.; Bouguila, N. Human Activity Recognition with an HMM-Based Generative Model. *Sensors* **2023**, *23*, 1390. [[CrossRef](#)] [[PubMed](#)]
40. Bouchabou, D.; Nguyen, S.M.; Lohr, C.; LeDuc, B.; Kanellos, I. Using language model to bootstrap human activity recognition ambient sensors based in smart homes. *Electronics* **2021**, *10*, 2498. [[CrossRef](#)]
41. Zhao, H.; Zheng, J.; Deng, W.; Song, Y. Semi-supervised broad learning system based on manifold regularization and broad network. *IEEE Trans. Circuits Syst. I Regul. Pap.* **2020**, *67*, 983–994. [[CrossRef](#)]
42. Chen, K.; Yao, L.; Zhang, D.; Wang, X.; Chang, X.; Nie, F. A semisupervised recurrent convolutional attention model for human activity recognition. *IEEE Trans. Neural Netw. Learn. Syst.* **2019**, *31*, 1747–1756. [[CrossRef](#)] [[PubMed](#)]
43. Ahmim, A.; Maglaras, L.; Ferrag, M.A.; Derdour, M.; Janicke, H. A novel hierarchical intrusion detection system based on decision tree and rules-based models. In Proceedings of the 2019 15th International Conference on Distributed Computing in Sensor Systems (DCOSS), Santorini, Greece, 29–31 May 2019; pp. 228–233.
44. Daghero, F.; Pagliari, D.J.; Poncino, M. Two-stage Human Activity Recognition on Microcontrollers with Decision Trees and CNNs. In Proceedings of the 2022 17th Conference on Ph.D Research in Microelectronics and Electronics (PRIME), Villasimius, Italy, 12–15 June 2022; pp. 173–176.
45. Kelly, P.; Marshall, S.J.; Badland, H.; Kerr, J.; Oliver, M.; Doherty, A.R.; Foster, C. An ethical framework for automated, wearable cameras in health behavior research. *Am. J. Prev. Med.* **2013**, *44*, 314–319. [[CrossRef](#)]
46. Basak, H.; Kundu, R.; Singh, P.K.; Ijaz, M.F.; Woźniak, M.; Sarkar, R. A union of deep learning and swarm-based optimization for 3D human action recognition. *Sci. Rep.* **2022**, *12*, 5494. [[CrossRef](#)]
47. Szegedy, C.; Ioffe, S.; Vanhoucke, V.; Alemi, A. Inception-v4, inception-resnet and the impact of residual connections on learning. In Proceedings of the AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 4–9 February 2017; Volume 31.
48. Chen, C.; Jafari, R.; Kehtarnavaz, N. UTD-MHAD: A multimodal dataset for human action recognition utilizing a depth camera and a wearable inertial sensor. In Proceedings of the 2015 IEEE International Conference on Image Processing (ICIP), Quebec City, QC, Canada, 27–30 September 2015; pp. 168–172.
49. Müller, M.; Röder, T.; Clausen, M.; Eberhardt, B.; Krüger, B.; Weber, A. *Mocap Database HDM05*; Technical report, No. CG-2007-2; Universität Bonn: Bonn, Germany, 2007; ISSN 1610-8892.
50. Shahroudy, A.; Liu, J.; Ng, T.T.; Wang, G. Ntu rgb+ d: A large scale dataset for 3d human activity analysis. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 1010–1019.
51. Domingo, J.D.; Gómez-García-Bermejo, J.; Zalama, E. Visual recognition of gymnastic exercise sequences. Application to supervision and robot learning by demonstration. *Robot. Auton. Syst.* **2021**, *143*, 103830. [[CrossRef](#)]
52. Taud, H.; Mas, J. Multilayer perceptron (MLP). In *Geomatic Approaches for Modeling Land Change Scenarios*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 451–455.
53. Li, Y.; Wang, L. Human activity recognition based on residual network and BiLSTM. *Sensors* **2022**, *22*, 635. [[CrossRef](#)]
54. Su, T.; Sun, H.; Ma, C.; Jiang, L.; Xu, T. HDL: Hierarchical deep learning model based human activity recognition using smartphone sensors. In Proceedings of the 2019 International Joint Conference on Neural Networks (IJCNN), Budapest, Hungary, 14–19 July 2019; pp. 1–8.
55. Kwapisz, J.R.; Weiss, G.M.; Moore, S.A. Activity recognition using cell phone accelerometers. *ACM SigKDD Explor. Newsl.* **2011**, *12*, 74–82. [[CrossRef](#)]
56. Reiss, A.; Stricker, D. Introducing a new benchmarked dataset for activity monitoring. In Proceedings of the 2012 16th International Symposium on Wearable Computers, Newcastle, UK, 18–22 June 2012; pp. 108–109.
57. Challa, S.K.; Kumar, A.; Semwal, V.B. A multibranch CNN-BiLSTM model for human activity recognition using wearable sensor data. *Vis. Comput.* **2022**, *38*, 4095–4109. [[CrossRef](#)]
58. Anguita, D.; Ghio, A.; Oneto, L.; Parra, X.; Reyes-Ortiz, J.L. A public domain dataset for human activity recognition using smartphones. In Proceedings of the European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, Bruges, Belgium, 24–26 April 2013; Volume 3, p. 3.
59. Dua, N.; Singh, S.N.; Semwal, V.B. Multi-input CNN-GRU based human activity recognition using wearable sensors. *Computing* **2021**, *103*, 1461–1478. [[CrossRef](#)]
60. Cho, K.; Van Merriënboer, B.; Bahdanau, D.; Bengio, Y. On the properties of neural machine translation: Encoder-decoder approaches. *arXiv* **2014**, arXiv:1409.1259.
61. Ramos, R.G.; Domingo, J.D.; Zalama, E.; Gómez-García-Bermejo, J. Daily human activity recognition using non-intrusive sensors. *Sensors* **2021**, *21*, 5270. [[CrossRef](#)]
62. Liciotti, D.; Bernardini, M.; Romeo, L.; Frontoni, E. A sequential deep learning application for recognising human activities in smart homes. *Neurocomputing* **2020**, *396*, 501–513. [[CrossRef](#)]
63. Cook, D.J.; Crandall, A.S.; Thomas, B.L.; Krishnan, N.C. CASAS: A smart home in a box. *Computer* **2012**, *46*, 62–69. [[CrossRef](#)] [[PubMed](#)]

64. Sazonov, E.; Hegde, N.; Browning, R.C.; Melanson, E.L.; Sazonova, N.A. Posture and activity recognition and energy expenditure estimation in a wearable platform. *IEEE J. Biomed. Health Inform.* **2015**, *19*, 1339–1346. [[CrossRef](#)] [[PubMed](#)]
65. D'Arco, L.; Wang, H.; Zheng, H. Assessing impact of sensors and feature selection in smart-insole-based human activity recognition. *Methods Protoc.* **2022**, *5*, 45. [[CrossRef](#)]
66. Zhang, P.; Lan, C.; Zeng, W.; Xing, J.; Xue, J.; Zheng, N. Semantics-guided neural networks for efficient skeleton-based human action recognition. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 1112–1121.
67. Yan, S.; Xiong, Y.; Lin, D. Spatial temporal graph convolutional networks for skeleton-based action recognition. In Proceedings of the AAAI Conference on Artificial Intelligence, New Orleans, LA, USA, 2–7 February 2018; Volume 32.
68. Liu, J.; Shahroudy, A.; Perez, M.; Wang, G.; Duan, L.Y.; Kot, A.C. Ntu rgb+ d 120: A large-scale benchmark for 3d human activity understanding. *IEEE Trans. Pattern Anal. Mach. Intell.* **2019**, *42*, 2684–2701. [[CrossRef](#)]
69. Hu, J.F.; Zheng, W.S.; Lai, J.; Zhang, J. Jointly learning heterogeneous features for RGB-D activity recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 5344–5352.
70. Li, Y.; Yang, G.; Su, Z.; Li, S.; Wang, Y. Human activity recognition based on multienvironment sensor data. *Inf. Fusion* **2023**, *91*, 47–63. [[CrossRef](#)]
71. Ruan, D.; Wang, J.; Yan, J.; Gühmann, C. CNN parameter design based on fault signal analysis and its application in bearing fault diagnosis. *Adv. Eng. Inform.* **2023**, *55*, 101877. [[CrossRef](#)]
72. Ramos, R.G.; Domingo, J.D.; Zalama, E.; Gómez-García-Bermejo, J.; López, J. SDHAR-HOME: A sensor dataset for human activity recognition at home. *Sensors* **2022**, *22*, 8109. [[CrossRef](#)]
73. Medsker, L.R.; Jain, L. Recurrent neural networks. *Des. Appl.* **2001**, *5*, 64–67.
74. Ahn, D.; Kim, S.; Hong, H.; Ko, B.C. STAR-Transformer: A Spatio-temporal Cross Attention Transformer for Human Action Recognition. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Waikoloa, HI, USA, 2–7 January 2023; pp. 3330–3339.
75. Dosovitskiy, A.; Beyer, L.; Kolesnikov, A.; Weissenborn, D.; Zhai, X.; Unterthiner, T.; Dehghani, M.; Minderer, M.; Heigold, G.; Gelly, S.; et al. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv* **2020**, arXiv:2010.11929.
76. Zhang, W.; Zhu, M.; Derpanis, K.G. From actemes to action: A strongly-supervised representation for detailed action understanding. In Proceedings of the IEEE International Conference on Computer Vision, Sydney, NSW, Australia, 1–8 December 2013; pp. 2248–2255.
77. Jiang, W.; Zhou, K.; Xiong, C.; Du, G.; Ou, C.; Zhang, J. KSCB: A novel unsupervised method for text sentiment analysis. *Appl. Intell.* **2023**, *53*, 301–311. [[CrossRef](#)]
78. Kwon, Y.; Kang, K.; Bae, C. Unsupervised learning for human activity recognition using smartphone sensors. *Expert Syst. Appl.* **2014**, *41*, 6067–6074. [[CrossRef](#)]
79. Hall, M.; Frank, E.; Holmes, G.; Pfahringer, B.; Reutemann, P.; Witten, I.H. The WEKA data mining software: An update. *ACM SIGKDD Explor. Newsl.* **2009**, *11*, 10–18. [[CrossRef](#)]
80. Cooley, J.W.; Tukey, J.W. An algorithm for the machine calculation of complex Fourier series. *Math. Comput.* **1965**, *19*, 297–301. [[CrossRef](#)]
81. Lin, J.F.S.; Kulic, D. Automatic human motion segmentation and identification using feature guided hmm for physical rehabilitation exercises. In Proceedings of the Robotics for Neurology and Rehabilitation, Workshop at IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), San Francisco, CA, USA, 25–30 September 2011.
82. Trabelsi, D.; Mohammed, S.; Chamroukhi, F.; Oukhellou, L.; Amirat, Y. An unsupervised approach for automatic activity recognition based on hidden Markov model regression. *IEEE Trans. Autom. Sci. Eng.* **2013**, *10*, 829–835. [[CrossRef](#)]
83. Li, W.; Xu, Y.; Tan, B.; Piechocki, R.J. Passive wireless sensing for unsupervised human activity recognition in healthcare. In Proceedings of the 2017 13th International Wireless Communications and Mobile Computing Conference (IWCMC), Valencia, Spain, 26–30 June 2017; pp. 1528–1533.
84. Kim, Y.; Ling, H. Human activity classification based on micro-Doppler signatures using a support vector machine. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 1328–1337.
85. Bai, L.; Yeung, C.; Efstratiou, C.; Chikomo, M. Motion2Vector: Unsupervised learning in human activity recognition using wrist-sensing data. In Proceedings of the Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers, London, UK, 9–13 September 2019; pp. 537–542.
86. Kingma, D.P.; Welling, M. An introduction to variational autoencoders. *Found. Trends<sup>®</sup> Mach. Learn.* **2019**, *12*, 307–392. [[CrossRef](#)]
87. Valarezo, A.E.; Rivera, L.P.; Park, H.; Park, N.; Kim, T.S. Human activities recognition with a single wrist IMU via a Variational Autoencoder and android deep recurrent neural nets. *Comput. Sci. Inf. Syst.* **2020**, *17*, 581–597. [[CrossRef](#)]
88. Stisen, A.; Blunck, H.; Bhattacharya, S.; Prentow, T.S.; Kjærgaard, M.B.; Dey, A.; Sonne, T.; Jensen, M.M. Smart devices are different: Assessing and mitigating mobile sensing heterogeneities for activity recognition. In Proceedings of the 13th ACM Conference on Embedded Networked Sensor Systems, Seoul, Korea, 1–4 November 2015; pp. 127–140.
89. Attal, F.; Mohammed, S.; Dedabrivili, M.; Chamroukhi, F.; Oukhellou, L.; Amirat, Y. Physical human activity recognition using wearable sensors. *Sensors* **2015**, *15*, 31314–31338. [[CrossRef](#)] [[PubMed](#)]
90. Sinaga, K.P.; Yang, M.S. Unsupervised K-means clustering algorithm. *IEEE Access* **2020**, *8*, 80716–80727. [[CrossRef](#)]

91. Zong, B.; Song, Q.; Min, M.R.; Cheng, W.; Lumezanu, C.; Cho, D.; Chen, H. Deep autoencoding gaussian mixture model for unsupervised anomaly detection. In Proceedings of the International Conference on Learning Representations, Vancouver, Canada, 30 April–3 May 2018.
92. Yang, X.; Song, Z.; King, I.; Xu, Z. A survey on deep semi-supervised learning. *IEEE Trans. Knowl. Data Eng.* **2022**, *35*, 8934–8954. [[CrossRef](#)]
93. Janarthanan, R.; Doss, S.; Baskar, S. Optimized unsupervised deep learning assisted reconstructed coder in the on-nodule wearable sensor for human activity recognition. *Measurement* **2020**, *164*, 108050. [[CrossRef](#)]
94. Gu, T.; Chen, S.; Tao, X.; Lu, J. An unsupervised approach to activity recognition and segmentation based on object-use fingerprints. *Data Knowl. Eng.* **2010**, *69*, 533–544. [[CrossRef](#)]
95. Ezeiza, N.; Alegria, I.; Arriola, J.M.; Urizar, R.; Aduriz, I. Combining stochastic and rule-based methods for disambiguation in agglutinative languages. In *COLING 1998 Volume 1: The 17th International Conference on Computational Linguistics*; Association for Computational Linguistics: Stroudsburg, PA, USA, 1998.
96. Sargano, A.B.; Gu, X.; Angelov, P.; Habib, Z. Human action recognition using deep rule-based classifier. *Multimed. Tools Appl.* **2020**, *79*, 30653–30667. [[CrossRef](#)]
97. Nurwulan, N.; Selamaj, G. Human daily activities recognition using decision tree. *Proc. J. Phys. Conf. Ser.* **2021**, *1833*, 012039. [[CrossRef](#)]
98. Sánchez, V.G.; Skeie, N.O. Decision Trees for Human Activity Recognition in Smart House Environments. *Linköping Electron. Conf. Proc.* **2018**, *153*, 222–229 [[CrossRef](#)]
99. Ordóñez, F.J.; De Toledo, P.; Sanchis, A. Activity recognition using hybrid generative/discriminative models on home environments using binary sensors. *Sensors* **2013**, *13*, 5460–5477. [[CrossRef](#)] [[PubMed](#)]
100. Zeng, Y. Evaluation of physical education teaching quality in colleges based on the hybrid technology of data mining and Hidden Markov Model. *Int. J. Emerg. Technol. Learn. (IJET)* **2020**, *15*, 4–15. [[CrossRef](#)]
101. Wang, X.; Liu, J.; Moore, S.J.; Nugent, C.D.; Xu, Y. A behavioural hierarchical analysis framework in a smart home: Integrating HMM and probabilistic model checking. *Inf. Fusion* **2023**, *95*, 275–292. [[CrossRef](#)]
102. Chadza, T.; Kyriakopoulos, K.G.; Lambotharan, S. Analysis of hidden Markov model learning algorithms for the detection and prediction of multi-stage network attacks. *Future Gener. Comput. Syst.* **2020**, *108*, 636–649. [[CrossRef](#)]
103. Yu, S.Z.; Kobayashi, H. An efficient forward-backward algorithm for an explicit-duration hidden Markov model. *IEEE Signal Process. Lett.* **2003**, *10*, 11–14.
104. Valdiviezo-Diaz, P.; Ortega, F.; Cobos, E.; Lara-Cabrera, R. A collaborative filtering approach based on Naïve Bayes classifier. *IEEE Access* **2019**, *7*, 108581–108592. [[CrossRef](#)]
105. Nica, I.; Alexandru, D.B.; Craciunescu, S.L.P.; Ionescu, S. Automated Valuation Modelling: Analysing Mortgage Behavioural Life Profile Models Using Machine Learning Techniques. *Sustainability* **2021**, *13*, 5162. [[CrossRef](#)]
106. Ekström, J.; Åkerrén Ögren, J.; Sjöblom, T. Exact Probability Distribution for the ROC Area under Curve. *Cancers* **2023**, *15*, 1788. [[CrossRef](#)]
107. Mingote, V.; Miguel, A.; Ortega, A.; Lleida, E. Optimization of the area under the ROC curve using neural network supervectors for text-dependent speaker verification. *Comput. Speech Lang.* **2020**, *63*, 101078. [[CrossRef](#)]
108. Khosravani Pour, L.; Farrokhi, A. Language recognition by convolutional neural networks. *Sci. Iran.* **2023**, *30*, 116–123.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.