Unsupervised Human Activity Recognition Using the Clustering Approach: A Review

Paola Ariza Colpas [1],*, Enrico Vicario [2], Emiro De-La-Hoz-Franco [1] , Marlon Pineres-Melo [3], Ana Oviedo-Carrascal [4] and Fulvio Patara [2],*

[1] Department of Computer Science and Electronics, Universidad de la Costa CUC,
   Barranquilla 080002, Colombia; edelahoz@cuc.edu.co
[2] Department of Information Engineering, University of Florence, 50139 Firenze, Italy; enrico.vicario@unifi.it
[3] Department of Systems Engineering, Universidad del Norte, Barranquilla 081001, Colombia;
   pineresm@uninorte.edu.co
[4] Faculty of Engineering in Information and Communication Technologies, Universidad Pontificia Bolivariana,
   Medellín 050031, Colombia; ana.oviedo@upb.edu.co
* Correspondence: pariza1@cuc.edu.co (P.A.C.); fulvio.patara@unifi.it (F.P.)

Abstract: Currently, many applications have emerged from the implementation of software development and hardware use, known as the Internet of things. One of the most important application areas of this type of technology is in health care. Various applications arise daily in order to improve the quality of life and to promote an improvement in the treatments of patients at home that suffer from different pathologies. That is why there has emerged a line of work of great interest, focused on the study and analysis of daily life activities, on the use of different data analysis techniques to identify and to help manage this type of patient. This article shows the result of the systematic review of the literature on the use of the Clustering method, which is one of the most used techniques in the analysis of unsupervised data applied to activities of daily living, as well as the description of variables of high importance as a year of publication, type of article, most used algorithms, types of dataset used, and metrics implemented. These data will allow the reader to locate the recent results of the application of this technique to a particular area of knowledge.

1. Introduction

1.1. Focus of this Survey

Clustering, also called grouping, aims to divide the data into groups of objects with similar characteristics. Through this technique, simplification of data information is achieved. This technique can be applied in both supervised and unsupervised learning, being currently used in different contexts, such as: data recovery, spatial data analysis, marketing, support in medical diagnosis, computational biology, among others. In the specific case of this review of the literature, the results obtained from the application of Clustering to Human Activity Recognition will be analyzed.

- Phase 1. Extraction or Selection of Feature: In this phase, it is necessary to define the characteristics or similarities to be analyzed. This feature can be selected or extracted; the difference in the processes is that, when the selected, the feature is chosen [1], whereas in the second option, the feature is transformed by different techniques to prepare it for new characteristic extracts [2].

The main purpose of this phase is to find the patterns belonging to different clusters, without noise, which are easy to analyze and are known [3,4].

- Phase 2. Clustering Algorithm selection: After extracting the feature, it is necessary to define the clustering algorithm to be applied. In addition to this important selection, defining a corresponding proximity measure and the construction of a criterion function is also indispensable. When the proximity measure function was built, it became an optimization problem with several case studies in the literature [5]. The clustering approach is now applicable to different areas, and for this reason, is very important in order to understand the characteristic of the problem to correctly decide correctly algorithm for solving the identified problem.

- Phase 3. Cluster Validation: In a group of data, the algorithms selected show the different partitions. The big difficulty is to understand and know the quality of the results—the results are defined by the clustering quality metrics [6]. These metrics are divided into two groups: externals and internal. The more useful internal metrics are: cohesion and separation [7], SSW (Sum of Squared Within) [8], SSB (Sum of Squared Between) [7], Sum of Squared base Indexes [6], Davies Bouldin [9], Silhouette coefficient [10] and Dunn-index [11]. The more useful external metrics are: Precision [12], Recall [13], F-Measure [14], Entropy [15], Purity [16], Mutual Information [17,18], and Rand-Index [19]

- Phase 4: Result Interpretation: The purpose of using clustering is to show new information extracted from the original data to solve the initial problem. In some occasions, in order to understand the results, it is necessary to contact an expert in order to explain the cluster's resultant characteristics. Additionally, additional experiments can be applied in order to explain and prove the extracted knowledge.

1.2. The Big Picture: Human Activity Recognition Using Learning Techniques Approach

With the increasing popularity and power of mobile technologies, there is an increasing interest in activity recognition from data collected from sensors in mobile devices [20,21]. One of the challenges of mobile-based activity recognition is the limited computational resource available on mobile devices. Furthermore, activity recognition within smart environments is presented with a number of general challenges. Individuals often carry out activities of daily living (ADL) differently, utilizing a high degree of freedom in relation to the sequential order and duration in which activities are performed. Additionally, a range of sensors within a smart environment typically generates heterogeneous data in terms of both formats and semantics. Consequently, it is often necessary to fuse and interpret sensor data from multiple sources in order to establish the context of an on-going ADL.

Data-driven approaches [22,23] learn activity models from pre-existent large-scale datasets of users' behaviors using data mining and machine learning techniques. Such techniques are advantageous as they are capable of handling uncertain and temporal information; however, they suffer from the problems of data scarcity, scalability and reusability. By contrast, knowledge-driven approaches build activity models [24–27] by exploiting rich prior knowledge in the domain of interest to construct models directly using knowledge engineering and management technologies.

While such models have the advantages of being semantically clear, logical and easy to initialize, they suffer from an inability to handle uncertainty and temporal information. Despite the enormous efforts applied to activity recognition, at present, there exists a gap between the potential of data generation and the aspiration of advanced assistance provision in which context-aware, personalized ADL assistance can be provided whenever needed. Taking into account the reasons explained above, the analysis of the activities of daily life is an area of knowledge rich in challenges where various research groups converge around and have generated, through their experience, different datasets for the experimentation of various learning techniques that are automatic for the analysis and extraction of knowledge. As a result of the different experiments in this line of work, a series of datasets have been built, which will be explained in more detail in the following sections, of which the following stand out: Kasteren [28], CASAS (Kyoto [29], Aruba [30], Multiresident [31]), UCI Human Activity Recognition (HAR) [32], Opportunity [33], and mHealth [34].

1.3. Outline

This article is a review of the literature concerning the use of the Clustering technique to discover information in Human Activity Recognition Unsupervised Datasets. First, a taxonomy of the Clustering approach to analyze Human Activity recognition is defined (Section 2). Second, the conceptual information is shown (Section 3). Third, the type of clustering methods for human activity recognition are explained (Section 4). Fourth, the methodology used for the selection and analysis of the results is detailed (Section 5). Fifth, the results of the scientometric analysis and technical analysis are specified, based on the articles selected for the systematic review of the literature (Sections 6 and 7). Finally, the conclusions and future works are shown (Sections 8 and 9).

2. Taxonomy

For the development of the systematic search in the literature, a conceptual or domain model was defined as the fundamental basis to understand the most relevant aspects to be analyzed within the research analyzed, which can be seen in detail in Figure 1. Below is the description and relationships of each of the concepts explained in the diagram.
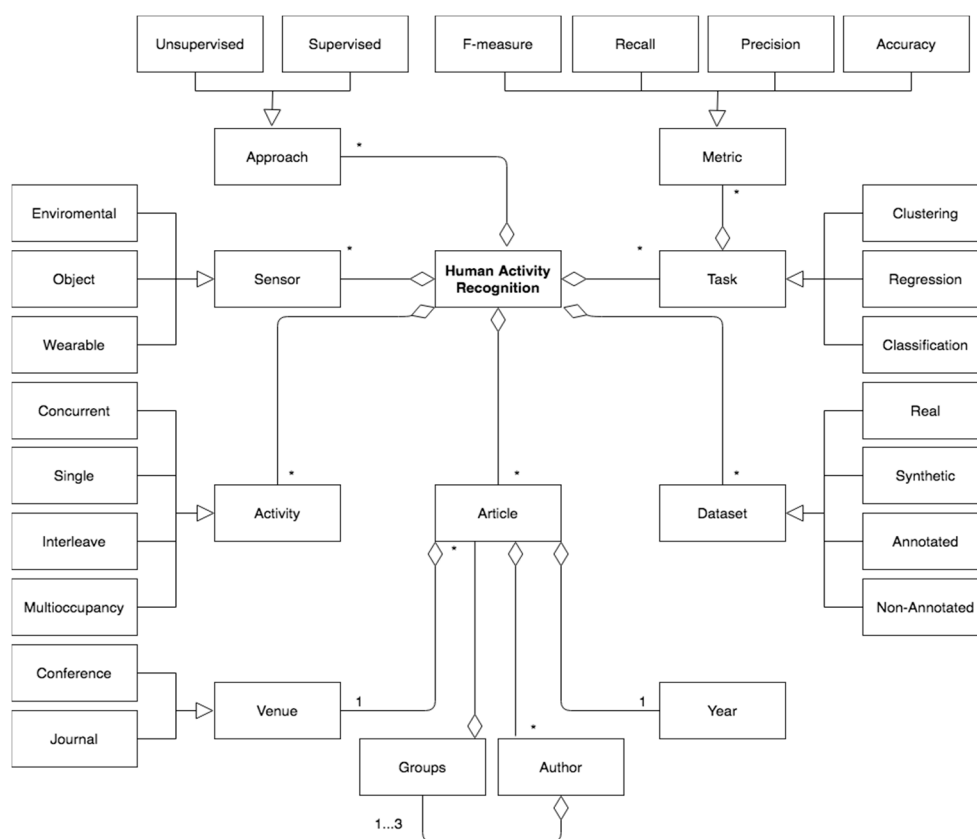

Figure 1. Conceptual model using UML Class Diagram formalism to represent the review of literature concepts.

Human Activity Recognition (HAR) (see Section 3.2) are socialized through Articles that are made at one Venue, whose proceedings can be Conferences or in academic Journals. An Article is published in a specific Year and can be written by many Authors who may belong to different research Groups. To perform the analysis of Human Activity Recognition (HAR), many Datasets are structured. The Dataset can be generated through the diffracted interaction of the users in the houses, which are Real or can be generated in Synthetic form through specialized software. The Dataset can be Annotated

when the inhabitants of the houses report, through logbooks or through devices, the activities they perform, or Non-Annotated when this information is not available (see Section 3.2.3).

To analyze Human Activity Recognition (HAR), different tasks of machine learning are used. There are different types of machine learning Tasks that can be used to extract information from these Datasets, of which we can highlight: Classification, Regression, and Clustering. A Classification task consists of being given a fact and is able to identify which predefined class it belongs. Clustering tasks allow you to group objects according to their similarity (see Section 3.1). Regression tasks are used when what you want to find out is a numerical value of a continuous variable. Additionally, the results of the analysis of the quality metrics associated with the implementation of the techniques, both internal and external, can be shown in an Article.

Human Activity Recognition (HAR), can be analyzed by different Approaches taken in both Supervised and Unsupervised scenarios. A scenario is Supervised when the class criteria information is available. On the contrary, it is Unsupervised when it is devoid of this class criteria information. Human Activity Recognition (HAR) data, is collected through the interaction of people or pets with different types of Sensors, such as: Environmental, Object, and Wearable (see Section 3.2.2). A Sensor is Environmental when it allows for the capturing of different changes in the house, such as: temperature, luminosity, etc. A Sensor is an Object when it is placed on the objects with which people have interactions, such as refrigerators, refrigerators, beds, sofa, etc. Finally, a Sensor is Wearable when it is located directly in the body of the person in order to be able to identify what action it takes, for example, to place sensors in the neck, arms, legs, waist, etc.

The Activities that are captured through the Sensors can be: Single, Interleaves and Multioccupancy. An Activity is Single when in the house of a single inhabitant who performs various activities in a given time; these activities have a special characteristic, i.e., to be able to perform an activity that the previous activity must have culminated. An Activity is Interleave when it is captured through the interaction of a single individual in the house; the main characteristic of this type of activity is that the individual can develop an activity without having completed an activity that has previously started. An Activity is Multioccupancy when it is captured through the interaction of several people in a house that can carry out various activities simultaneously, without having completed the previously initiated activity. Additionally, an Activity is Concurrent when an individual can develop several activities in parallel (see Section 3.2.1).

3. Conceptual Information

3.1. Clustering Techniques

Clustering is frequently used in any discipline that involves multivariate data analysis. Therefore, there is currently a large amount of literature documenting the importance of grouping together data analysis. Likewise, numerous scientific and application fields that have used clustering techniques, as well as thousands of published algorithms, have also been evidenced. Among the most relevant problems where clustering has been used, we can find: Image Segmentation [27–36], Document processing [37–39], Customer behavior analysis [40,41], Biological analysis of human behavior [42], Grouping and analysis of daily life activities [43–45], etc. The use of clustering normally serves the following purposes:

- **Underlying structure**: to depend the data, generate hypotheses, detect anomalies, and identify the most prominent characteristics.
- **Natural classification**: to identify the degree of similarity between the forms of organisms (phylogenetic relationship).
- **Compression**: as a method to organize data and complement it through clustering prototypes.

3.1.1. Clustering Methods

There are currently different methods for the application of clustering. The following will be described: Partitional, Hierarchical, Diffuse, Graphical Based, Evolutionary, Kernel-Based, and Neural Network-Based, see Table 1.

Table 1. Clustering's methods and algorithms.

Method	Algorithm
Partitional Method	K-means algorithm [46,47]
Hierarchical Method [48]	COBWEB [49,50]
Diffuse Method [13,51]	Fuzzy C Means [52,53]
Method Based on Neural Networks [54,55]	SOM [56]
Evolutionary Methods [57–59]	Genetic Algorithms [57–62]
Kernel-Based methods [63,64]	Kernel K-means Algorithms [65,66]
Spectral Methods [36]	Standard Spectral Clustering [59]

3.1.2. Clustering Methods Descriptions

Partitional Method: The method of partition ordering has, as its main objective, the single partition of data without requiring another additional sub-partition. The result of this type of method is the separation between the groups in the form of hypersurfaces. The main achievement of the partition algorithms is to analyze the distances between the objects that are processed, which have a wide applicability in solving different problems.

Hierarchical Method: This method aims to optimize a specific function. That is, objects in the same cluster should be like each other, while those in different groups should be as different as possible. The different algorithms framed within this method mainly vary in the measure of similarity and the criteria to evaluate the overall quality of the grouping performed.

Diffuse Method: Clustering algorithms framed in the diffuse method have the purpose of identifying the different classes that represent the functional states that are present in a system, considering the historical dataset that is available. In order to establish the degree of association of the classes that have been identified, it is necessary to have expert staff in the process. By achieving the completeness of these two main tasks, the learning stage of the Diffuse Classifier is completed.

Evolutionary Method: Evolutionary methods are based on employing heuristic targets, which in turn employ other types of computational models for the analysis of the evolutionary process. There are different ways to perform the analysis, based on evolutionary methods, but the following can be highlighted: Genetic algorithms [57,58], Evolutionary Programming Evolutionary Programming [61,67], Evolutionary Strategies [68,69], and Genetic Programming [70].

Spectral Method: Spectral clustering has become one of the most modern ways of performing grouping processes; they are simple to implement and can efficiently solve problems under linear algebraic standards with very good performance, as compared to traditional algorithms, such as K-means [71].

Kernel-based method: Spectral clustering methods arise from concepts based on the spectral graph theory. The fundamental idea is to construct a weighted graph for the initial dataset where each of the nodes represent patterns and each weighted edge that changes the values based on each of the patterns. Many of the representations of kernel-based algorithms are based on Laplacian graphs [63,64].

Method based on Neural Networks: Clustering methods, based on neural networks, have become an alternative for solving many real problems [54,55]. The unsupervised neural networks, particularly the SOM self-organizing maps [69,72], provide a more robust and secure use for clustering large amounts of data.

3.2. Human Activity Recognition

Human Activity Recognition (HAR) is the process of automatically detecting human actions from the data collected from different types of sensors [73]. This sensor detects a set of common activities depending of the kind of sensors and the place of occupancy. In fact, this area is an important area of research because it can help solve different problems in sectors, such as: healthcare, energy, computer vision, and others. In this section, the descriptions of the feature of HAR is shown: activities, types of sensors, common datasets, occupancy features and supervised and unsupervised technique examples.

3.2.1. Activities

The main objective of this area of study is to identify the type of activity that is being developed by one or more individuals within a place equipped with a set of sensors that measure different types of parameters. Among the activities that are normally identified are shown in Table 2.

Table 2. Useful activity in Human Activity Recognition.

#	Activity' Name	Description
1	Make a phone call [74]	The participant moves to the phone in the dining room, looks up a specific number in the phone book, dials the number, and listens to the message.
2	Wash hands [74]	The participant moves into the kitchen sink and washes his/her hands in the sink, using hand soap and drying their hands with a paper towel.
3	Cook [74]	The participant cooks using a pot.
4	Eat [74]	The participant goes to the dining room and eats the food.
5	Clean [74]	The participant takes all the dishes to the sink and cleans them with water and dish soap in the kitchen.
6	Fill medication dispenser [32]	The participant retrieves a pill dispenser and bottle of pills.
7	Watch DVD [32]	The participant moves to the living room, puts a DVD in the player, and watches a news clip on TV.
8	Water plants [32]	The participant retrieves a watering can from the kitchen supply closet and waters three plants.
9	Answer the phone [32]	The phone rings, and the participant answer it.
10	Prepare birthday card [32]	The participant fills out a birthday card with a check to a friend and addresses the envelope.
11	Prepare soup [32]	The participant moves to the kitchen and prepares a cup of noodle soup in the microwave.
12	Choose outfit [32]	The participant selects an outfit from the clothes closet that their friend will wear for a job interview.
13	Hang up clothes in the hallway closet [75]	The clothes are laid out on the couch in the living room.
14	Move the couch and coffee table to the other side of the living room [75]	Request help from another person in multioccupancy experimentation.
15	Sit on the couch and read a magazine [75]	The participant sits down in the living room and reads a magazine.
16	Sweep the kitchen floor [75]	Sweep the kitchen floor using the broom and dustpan located in the kitchen closet.
17	Play a game [75]	Play a game of checkers for a maximum of five minutes in a multioccupancy context.
18	Simulate paying an electric bill [75]	Retrieve a check, a pen, and an envelope from the cupboard underneath the television in the living room.
19	Walking [76]	Using body sensors, define if the participant is performing the walking action.

Table 2. *Cont.*

#	Activity' Name	Description
20	Sitting [76]	Using body sensors, define if the participant is performing the sitting action.
21	Sleeping [76]	Using body sensors, define if the participant is performing the sleeping action.
22	Using a computer [76]	The participant is in the position of use of the computer for a certain time.
23	Showering [76]	Detection of environmental sensors of the participant's stay in the shower.
24	Toileting [76]	Detection of environmental sensors of the participant's stay in the bathroom.
25	Oral hygiene [76]	Using the object and body sensors, the oral hygiene action is identified.
26	Making Coffee [76]	Detection of objects and environmental sensors of the action of making coffee by the participant.
27	Walking upstairs [76]	The participant performs the action of climbing the stairs, being detected by the body sensors.
28	Walking down stairs [76]	The participant performs the action of going down the stairs, being detected by the body sensors.

Four types of activities have been identified in the literature review: single [32], interleaved [77], multioccupancy [75] and concurrent. The conceptual difference between these types of activities is that a single activity is constituted that is completely terminated before a new activity can be initiated. An interleaved activity refers to one that can be carried out in parallel with another activity and the multioccupancy activities refer to activities carried out by different people simultaneously (see Figures 2–5).
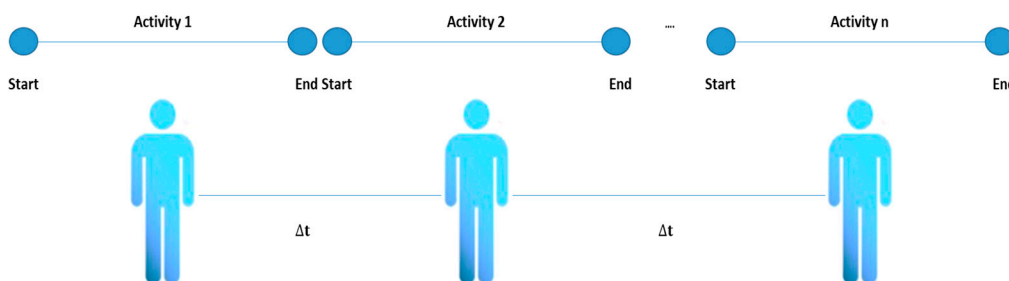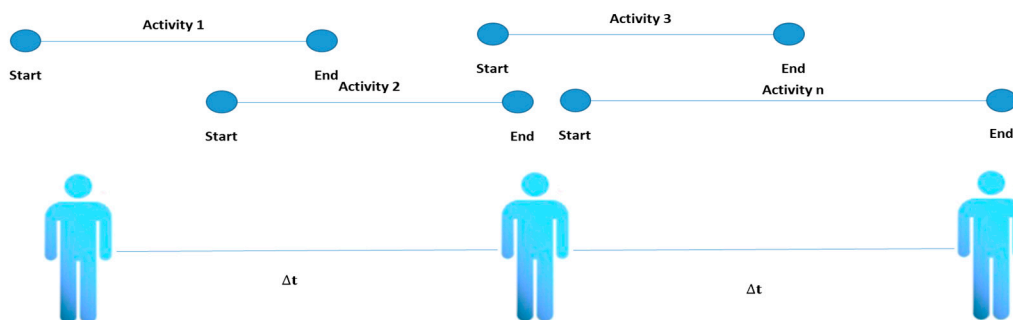

Figure 2. Single activity representation.

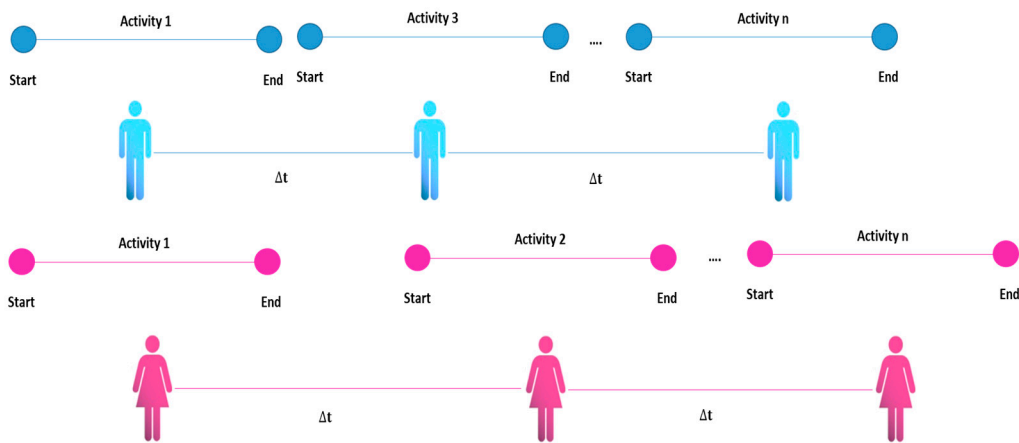

Figure 3. Interleaved activity representation.

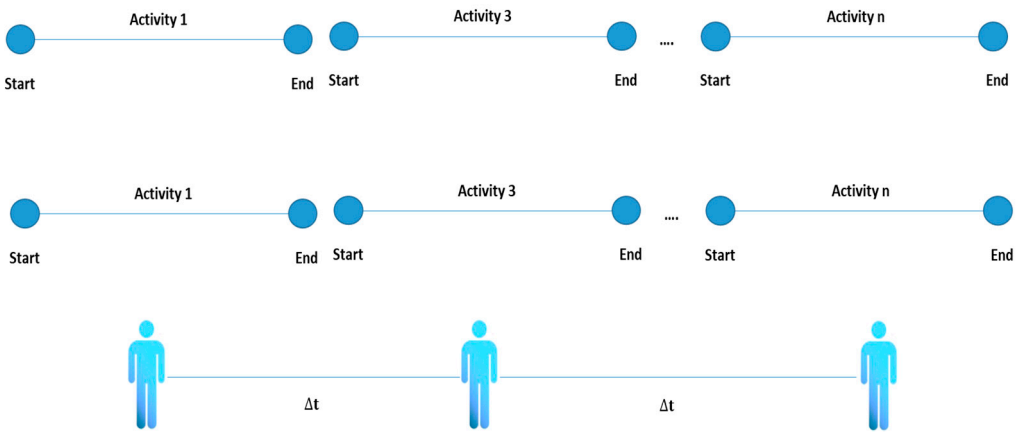**Figure 4.** Multioccupancy activity representation.

Figure 5. Concurrent activity representation.

3.2.2. Type of Sensor

To complete the process of detecting activities, it is necessary to have sensors of different types. Among which, we can highlight the environmental sensors, object sensors, and wearable sensors. The environmental sensor in an activity recognition context is a passive sensor that are integrated into the environment itself [78,79], such as: robot sensor, combination of audio and video, vision based system, eye tracking sensor, pressure sensors, passive infrared sensor, etc. The object's sensor shows the state of the objects in an activity recognition context; in some cases, it uses radio frequency identification (RFID). Figure 6 illustrates an example of the activities that can be identified using environmental and object sensors in a common apartment (Table 3) and describes some work related to the use of this type of sensor.

The wearable sensors, as the name implies, are integrated into wearable objects or directly with the body in order to help monitor health and/or provide clinically relevant data for care [87]; Table 4 and Figure 7 describe the related work and the use of this type of sensor.
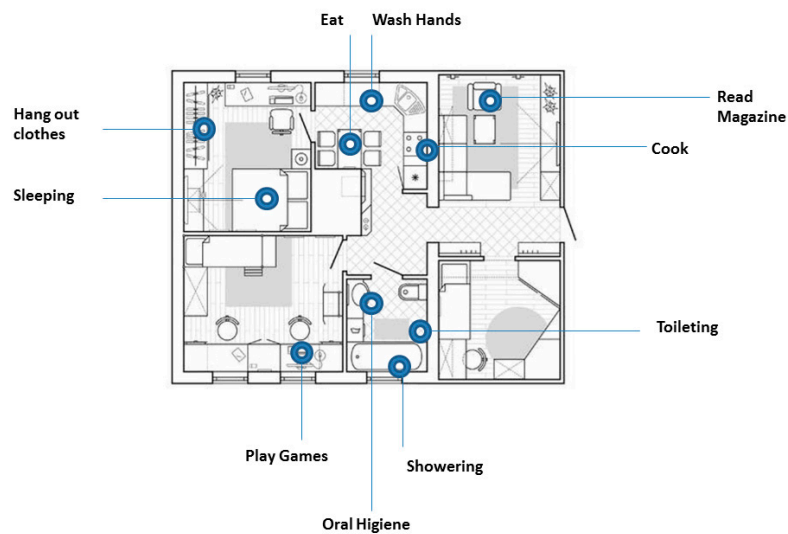
Figure 6. Environmental and object sensor representation related to activities of daily living (ADL).

Table 3. Body sensor analysis.

#	Type of Sensor	Sensor	Type of Activities	Reference
1	Environmental and Object sensors	Motion detectors, break-beam, pressure mats, contact switches, water flow, and wireless object movement	Eat, drink, housework, toileting, cooking, using a computer, watching TV, and call by phone	[80]
2		motion, temperature and humidity sensors, contacts switches in the doors, and item sensors on key items	phone call, cooking, wash hands, and clean up.	[81]
3		Binary sensors on doors and objects	Toileting, bathing, and grooming	[82]
4	Object sensors	Shake sensors	Leaving, toileting, showering, sleeping, drinking, and eating	[83]
5		radio frequency identification (RFID)	Toileting, oral hygiene, washing, telephone use, taking medication, etc.	[84]
6			Using bathroom, making meals/drinks, telephone use, set/clean table, eat, and take out trash	[85]
7			Making coffee	[86]

Table 4. Body sensor analysis.

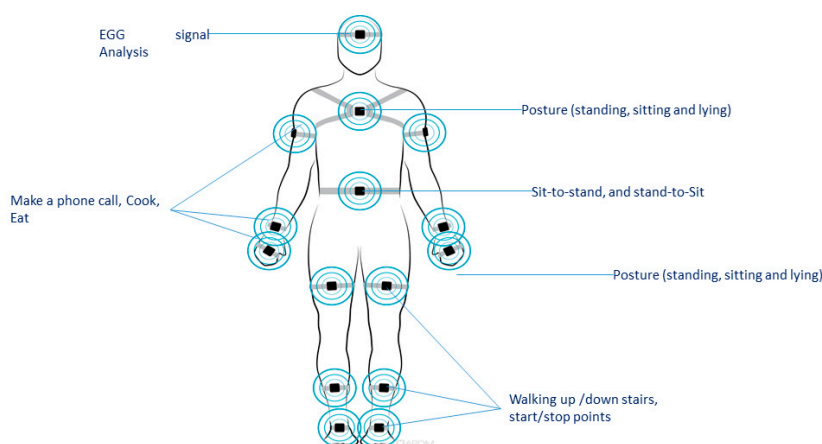| Number	Sensor Location	Type of Activities
1	Chest [88]	Standing, sitting and lying.
2	Waist [89]	Sit-to-stand, stand-to-sit, walking.
3	Upper arm, wrist, thigh and ankle [90]	Posture and some ADLs.
4	Wrist [91]	Sport movement.
5	Wrist, waist, and shoulder [92]	Riding elevator, walking up stairs.
6	On the belt [93,94]	Walking upstairs, walking downstairs, start or stop points.

Figure 7. Body sensor representation related to ADL.

3.2.3. Dataset for Human Activity Recognition

In the literature, many authors developed or used different datasets to identify human activities. In this section, the most useful datasets are explained, taking into account different aspects, such as: occupancy, number of individuals that interact, number of activities, and the type of sensor, with the following acronyms being used. E—Environmental Sensor, O—Object Sensor, A—Accelerometer, G—Gyroscope, and M—Magnetometer (see Table 5).

Table 5. Dataset's descriptions.

Number	Dataset's Name	Occupancy	# Subjects	# Activities	Sensor's Type
1	Vankasteren [28]	Single	1	8	E
2	Opportunity [33]	Multioccupancy	4	16	O, A
3	CASAS- Daily Life Kyoto [29]	Single	1	10	O, A
4	UCI SmartPhone [32]	Multioccupancy	30	6	A, G
5	CASAS Aruba [30]	Single	1	11	E, O
6	PAMAP2 [95]	Multioccupancy	9	18	A, G, M
7	CASAS Multiresident [31]	Multioccupancy	2	8	A, O, E
8	USC-HAD [96]	Multioccupancy	14	12	A, G
9	mHeath [34]	Multioccupancy	10	12	A, G
10	WISDM [97]	Multioccupancy	29	6	A
11	MIT PlaceLab [98]	Single	1	10	A, O, G
12	DSADS [99]	Multioccupancy	8	19	A, G, M
13	DOMUS [100]	Single	1	15	A, G, O
14	Smart Environment- Ulster University [101]	Single	1	9	A, G, M
15	UJAmI SmartLab [102]	Single	1	7	O, E

The Van Kasteren dataset [28] recorded the data of single occupancy in houses with one man. He stays alone in the apartment with three rooms. In this apartment, there were 14 wireless sensors. The data was collected for 28 days; the results of the interaction of the man with the sensors were: 2120 sensors events and 245 instances of activities. The Opportunity dataset [33] shows the interaction of four people in a house. In this house, there were 23 body-worn sensors, 12 object sensors, 21 ambient sensors and 4 recordings. Finally, this dataset contains 242 attributes and 2551 instances of activities.

The CASAS Daily Life Kyoto [29] is a single occupancy dataset that was collected from 25 October 2007–31 May 2008. In this apartment, there were 40 sensors consisting of: 37 motion and 3 temperature sensors; this dataset contains 4 attributes and 1530 instances of activities. The UCI Human Activity Recognition Using Smartphones DatasSet [32] details the interaction of 30 subjects, within an age bracket of 19–48 years. The subjects perform 6 activities: walking, walking upstairs, walking downstairs, sitting, standing and laying. This dataset contains 561 attributes and 10,299 instances.

The CASAS Aruba [30] is a single occupancy dataset; the volunteer woman interacts with 4 temperature sensors, 31 motion sensors, and 4 door closure sensors. This dataset contains 5 attributes,

and 6438 activities were identified. The PAMAP2 dataset [95] shows the interaction of 9 people for approximately 1 month. The records detail 19 physical activities, such as: standing, walking, ascending stairs, and descending stairs.

The CASAS Kyoto Multiresident [31] shows the interaction of 2 individuals in an apartment at the same time, with 78 sensors: 51 motion, 8 item, 12 cabinets, 2 water, 1 burner, 1 phone, and 3 temperature sensors, distributed in several places. The records show 5 attributes: date Time, SensorID, Value ResidentID, and TaskID. The first three are generated automatically for the smart interactions and the last two are annotated by the residents. The USC-HAD [96] shows the interaction of 14 people (7 male and 7 female), performing 12 activities: walking forward, walking left, walking right, walking upstairs, walking downstairs, running forward, jumping, sitting, standing, sleeping, elevator up, and elevator down. The record shows 14 attributes, which were generate using different types of sensors: accelerometer and a motion binary sensor.

The mHealth dataset [34] shows the interaction of 10 people in an apartment. During the experiment, 3 wearable sensors sent information concerning the activities carried out by the people. These activities consisted of standing still, sitting and relaxing, lying down, walking, climbing stairs, waist bends forward, front elevation of arms, knees bending (crouching), cycling, jogging, running, and jump front and back. The record contains 23 attribute and 120 instances. The WISDM dataset [99] details the activities of 29 people in an apartment performing 6 activities: walking, jogging, sitting, standing, and climbing stairs. The data was collected using mobile phone with the Android operating system.

The MIT PlaceLab [98] is a single occupancy dataset. The data was collected on Friday, 4 March 2005 from 9 a.m. to 1 p.m., performing 10 activities. In the apartment, there was a variety of sensor as switch sensors, light sensors, and current sensors, etc. The data were labeled by the people who participated in the experimentation. The Daily and Sports Activities Dataset (DADS) [99] shows the interaction of 8 people (4 males and 4 females between the ages 20 and 30) for 5 min. The persons performing 19 activities: sitting, standing, lying on back and on right side, ascending and descending stairs, standing in an elevator still, moving around in an elevator, walking in a parking lot, walking on a treadmill with a speed of 4 Km/h, walking on a treadmill with a speed of 8 Km/h, exercising on a stepper, exercising on a cross trainer, cycling on an exercise bike, rowing, jumping, and playing basketball. For recovery of the data, 9 sensors were used: x,y,z accelerometers, x,y,z gyroscopes, and x,y,z magnetometers. This dataset contains 5625 attributes and 9120 instances.

Finally, three datasets are described in the DOMUS dataset [100]; the data was collected in an apartment with 40 square meters in order to detect 15 activities. This apartment was provided by a variety of sensors types: electricity counter, water counters, ceiling spots, presence detector, power plugs, dimmed plug, temperature sensor, external shutters, internal blinds, luminosity sensor, air quality sensor, controllable curtains, dimmed light, and luminosity sensor. The Smart Environment-Ulster University [101] dataset detects the following activities: drink a glass of water, prepare tea, prepare a hot chocolate, drink a glass milk, call by phone, prepare hot snack, prepare cold snack, watch TV, and washing dishes. To detect these types of activities, different types of sensor were used: kitchen door sensor, living room door sensor, cutlery cupboard sensor, dishes cupboard sensor, glasses and cup cupboard sensor, pantry cupboard sensor, fridge door sensor, chair sensor, sofa sensor, television sensor, phone sensor, water sensor, and kettle sensor. The UJAmI SmartLab [102], by Universidad of Jaén in Spain, shows the interaction of different types of sensors: environmental wearable sensors; actuators; smart device; low cost devices, like raspberry Pi and arduino; indoor locations; vision cameras; screen; health devices; brain interfaces; human computer interfaces; and robots.

3.2.4. Supervised and Unsupervised

In the construction of the data, it is important to specify that it is possible to perform different types of analyses, taking into account whether or not you have the class criteria that defines the activity that the individual is doing inside the house. In the case of having information on the kind of activity that is being developed, either one or several individuals or even pets, it is determined that the learning that can

be done on the data is supervised, or also in terms of the area of knowledge annotated. These types of annotations are usually made through the completion of activity logbooks, or alternatively, using different devices, either audio or video. When there is no information on the kind of activity that is being carried out by either an individual or several individuals in the home, the learning that can be done on this data can be called unsupervised, or it can also be called non-annotated. In other occasions, the same dataset allows the two types of both supervised and unsupervised learning to be carried out, so it is partially annotated.

3.2.5. Single or Multioccupancy

A highly important factor in the field of recognition of daily life activities is the definition of the number of people or animals that interact within the house. This variable is really very important because the habits of the activities in each of the individuals vary. For example, in some datasets, the participation of the experimentation of older adults and children is identified. Additionally, interaction with pets that can activate the action of a sensor inside the house for a certain time without performing any activity can be included in the data. All aforementioned variables complicate the model and therefore generate new research topics in this area of knowledge.

4. Type of Clustering Methods for Human Activity Recognition

Taking into account the fact that clustering is one of the techniques mostly used in issues of extraction and the selection of similarities and common characteristics between objects, different architectures or forms of implementation have been developed that allow its correct application in the recognition of activities of daily life. The similarities of the implementation of different clustering techniques allow its view to become the basis for the subsequent implementation of other techniques or tasks of machine learning, thus showing the versatility of applications that clustering can have in this context. These architectures or forms of knowledge extraction are shown in the extensions below:

With the generation of new datasets, the analysis of the data quality and the beginnings of implementation arises. Many of the datasets generated can be analyzed through either the execution of a technique or by comparing several existing techniques to later analyze the results of the quality metrics associated with the implementation. Other more mature experiments focus on the selection of a different HAR dataset, which contain similar characteristics associated with: occupancy, type of sensors, among others, in order to identify the results of different techniques when being implemented in the selected dataset.

5. Methodology

To carry out this review of the literature effectively, different tools and supplies were considered. The searches were carried out in the specialized databases: Scopus, Web of Science, Science Direct, IEEE, and ACM. The methodology for the presentation and analysis of the results is referenced by Kitchenham [103]. Figure 8 illustrates the different concepts that were considered to perform the search.

It is important to specify that the area of study in question is limited to the analysis of the different activities that you can perform, either an individual or several, in a smart place. That is why different acronyms or keywords have been defined in the literature to identify this area of study. Among which, the following can be identified: Human Activity Recognition (HAR), Activity Daily Living (ADL), and Ambient Assisting Living (AAL). Additionally, only articles published between 2015 and 2020 were selected. Other very relevant aspects are the identification of the approach that can be supervised or unsupervised for the specific case of this review, the work carried out on human activity recognition with a "supervised" approach, and the "images and audio" feature was excluded. In the same way, there are different tasks of data mining that have been used to recognize the activities of daily life; punctually in this review refers to those that, in any of their phases of knowledge discovery, have made use of "Clustering" to support the interpretation or improvement of the results.
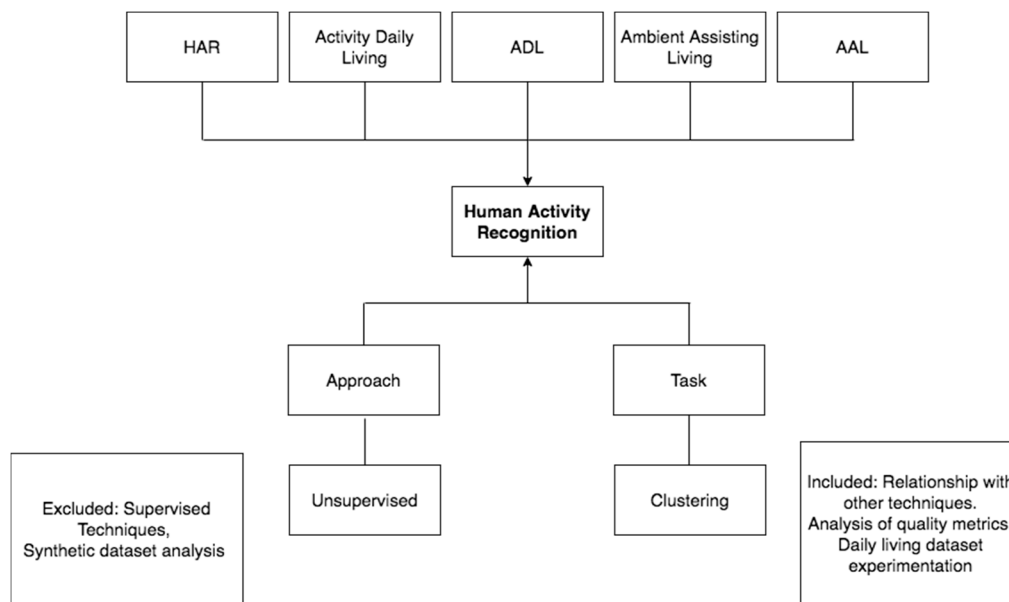
Figure 8. Search features of review. HAR—Human Activity Recognition; ADL—Activities of Daily Living; AAL—Ambient Assisting Living.

With the combination of keywords mentioned above, a total of 190 records were obtained, which were subsequently filtered through the relevance of the appearance of the keywords in the document, the application of the technique in the discovery of knowledge and the contribution of the combination of clustering with other techniques to solve problems associated with this study area, and finally obtaining a set of 64 records that will be analyzed in later sections of this document. These records were analyzed using a meta-analytical matrix that has the following fields: year of publication, title of the paper, name of the journal or event where the article was published, ISSN, venue type in which the article was published as either a journal or conference, quartile if applicable, country where the journal originated from or where the conference was held, country of the first author, university to which the first author is a part, dataset used for experimentation, type of dataset whether synthetic or real, type of task performed, algorithm used, and metrics extracted from the application of the selected algorithm. With these aspects, it is possible to detail the results of the various experiments carried out.

In this matrix, general information of the publications are analyzed (see Table 6). Using the information of general publications, different conclusions of high relevance can be made, such as: the year of the greatest number of publications on the specific subject, which can determine the trend line in the publications, and the database where the majority of the authors selected to make the publications related to the topic. The titles that are constrained within the selected keywords to perform the search can also be identified. In the journal column, the journal, conference, or book series where the results regarding this line of work are usually published can be identified.

In the same way, by using this matrix, the quartile of the publications, as well as the country where the subject is being published and the university and country of the first author, can be identified. All this information supports the identification of the research groups or researchers that are supporting the advances in the development of this line of knowledge.

Among the most relevant technical aspects of the meta-analytical matrix is the identification of the clustering algorithm and the different analyses of the quality of the techniques associated with the different experiments (see Table 7).

**Table 6.** General information of publication analyzed in a meta-analytical matrix.

Identifier	Year	Paper Title	Journal	ISSN	Proceedings or Book	Quartile	Journal Country	First Author´s Country	University
Art1	2015	Towards unsupervised physical activity recognition using Smartphone accelerometers	Multimedia Tools and Applications	1380–7501	Book Series	Q1	Netherlands	China	Langhou University

Table 7. General information of publication analyzed in a meta-analytical matrix.

Identifier	Dataset	Type	Methods	Metrics				Approach
				Accuracy	Precision	Recall	F-Measure	
Art1	Kasteren	Real	Calculating neighborhood radius	86	76	80	76	Unsupervised
Art2	WISDM	Real	MCODE-Based	85	77	83	77	Unsupervised

6. Scientometric Analysis

Considering the analysis of the 64 records, which were processed based on the identification of the use of the clustering technique in the recognition of unsupervised daily life activities, the following scientific data could be identified, as detailed below. In Figure 9, it can be noted that 55% of the publications have been made in the Journal, while 30% of the publications are from participation in conferences where those that have been organized by IEEE stand out. Only 10% correspond to the Book Series.
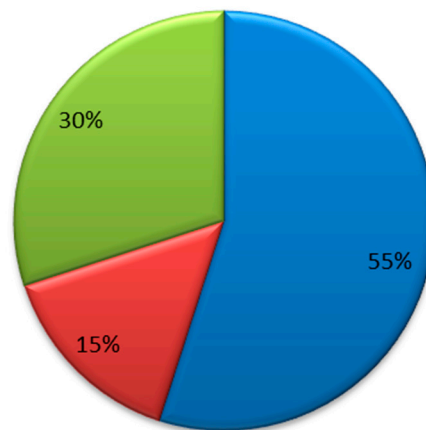
Figure 9. Origin of the publications by venue.

Regarding the years of publication of the manuscripts, it is important to specify that the search was confined to between 2015 and 2020. The year 2018 had the highest concentration of publications, and the year 2017 was consolidated as the year where the lowest number of publications was found, as can be seen in Figure 10.
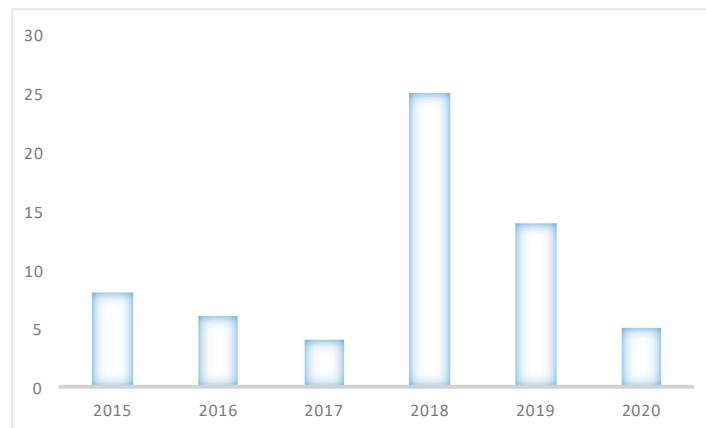
Figure 10. Analysis of the articles according to the year of publication.

In the same way, when performing an analysis of the articles published only in journals, we excluded those published in conferences and book series. The Sensors journal is consolidated as the journal with the highest number of works in this work area being published, based on the articles of this review, followed by the journal of Ambient Intelligence and Humanized Computing and the Engineering Applications of Artificial Intelligence (see Figure 11).
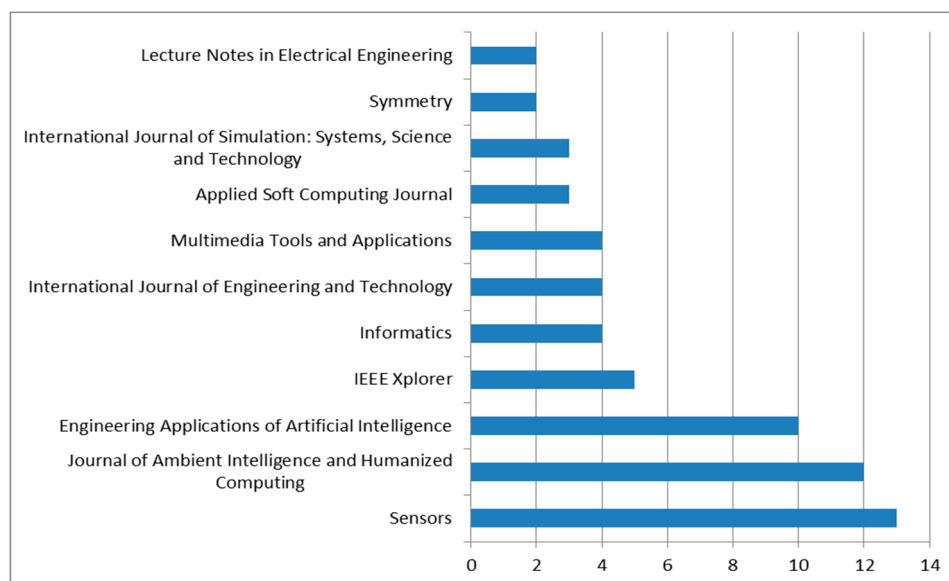

Figure 11. Distribution of articles according to the journal in which they were published.

Figure 12 shows the quartiles in which the set of publications is delimited, excluding those belonging to conference proceedings. The publications that have been taken as the basis for this review are in quartile Q1, followed, respectively, by quartile Q2, having, as a minority, those published in journals Q4.
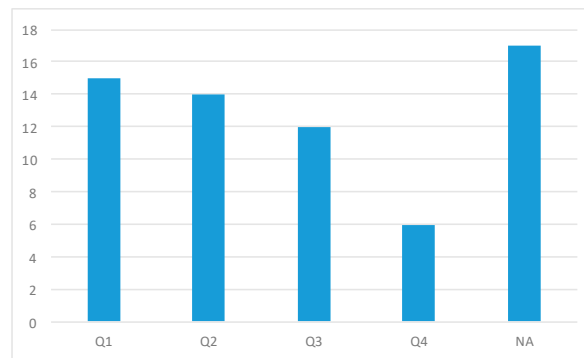
Figure 12. Quartile of journal publications.

When carrying out an analysis of the most used dataset for experimentation with clustering in an unsupervised environment, the wide use of the VanKasteren dataset with 30% can be highlighted, followed by the Casas Kyoto dataset with 13%. The Casas Aruba and Opportunity dataset have also been used in 7% of experiments (see Figure 13).
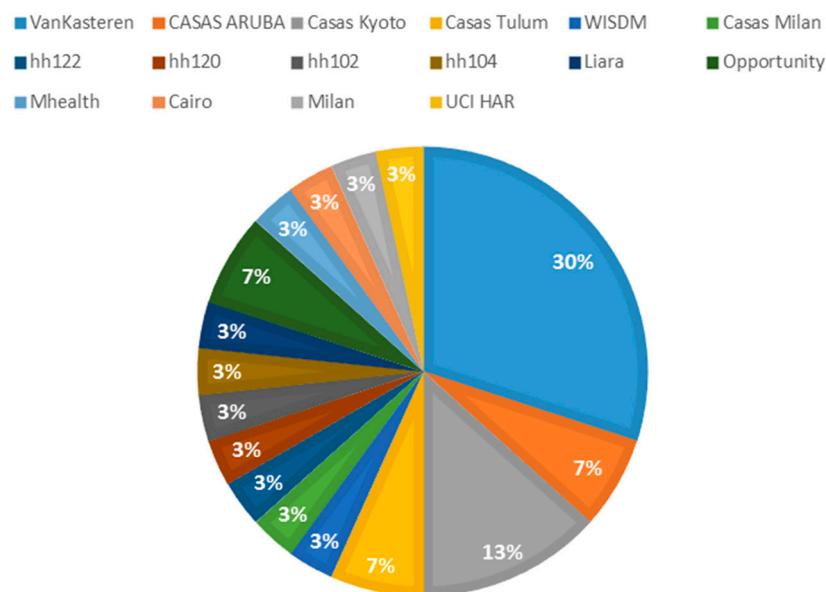

Figure 13. Search features of the review.

7. Technical Analysis

There is a wide range of clustering techniques within the literary review, which have been used to perform various experiments in different contexts. In this review, we will concentrate on three techniques that have been most used to analyze the recognition of activities of daily living: K-NN, K-means, and Sub-clustering. While K-means and sub-clustering are concerned with finding similarities between the data and forming groups, the K-NN technique is used to predict new data to which cluster they belong. The combination of K-means to find groups and K-NN to predict new instances is usual because both algorithms are based on distance measurements. In Tables 8–10, each of the respective techniques have been discriminated, considering the high relevance of variables, such as the reference to the author who performed the experimentation, the dataset used, and the quality metrics associated with the experimentation. For the specific case of this analysis, the metrics of quality accuracy, precision, recall and F-measure have been used, the data being entered in accordance with the data provided by the authors.

One of the mostly used algorithms is the K-NN algorithm. The experiments of this algorithm have been carried out with the dataset widely detailed in the literature, such as: VanKasteren, Casas Aruba, Casas Kyoto, Casas Tulum, Hh102, Hh1034, UCI HAR, MIT PlaceLab, and PAMAP2. In Reference [104], the best results of the K-NN implementation in the VanKasteren dataset can be seen, obtaining results, such as: 97.2% accuracy, 88.25% accuracy, 83.6% recall, and 84% F-measure. In the same way [104], the results of the implementation of this technique in the Casas Aruba and Casas Kyoto dataset are detailed, with the following quality metrics: Accuracy 98.14%, precision 74.73%, Recall 76.29%, and F-measure 72%. Another experimentation carried out in References [104] in the Casas Tulum dataset, shows the following results: Accuracy 86.15%, Accuracy 59.18%, Recall 57.12%, and F-measure 57%. Similarly, the experiments carried out in the following datasets stand out, which show the following results of the metric accuracy: Casas Tulum [104] 86.15%, Hh102 [105] 66%, Hh104 [105] 78%, UCI HAR [106] 71%, MIT PlaceLab [107] 94.5%, and PAMAP2 [108] 62%. Table 7 shows the result of the implementation of the K-means algorithm in different datasets, in which the following accuracy results can be highlighted: VanKasteren [109] 88.6%, WISDM [110] 71%, Liara [111] 86%, MHealth [112] 71.66%, Opportunity [113] 86.8%, and UCI HAR [114] 52.1%.

Table 8. Details of the K-NN experimentation.

References	Dataset	Accuracy	Precision	Recall	F-Measure
[104]		97.2%	88.25%	83.66%	84%
[116]		96.67%	97.33%	96.67%	97%
[117]		93.55%	92.97%	91.3%	91%
[109]		–	95%	100%	97%
[118]		88.14%	–	–	–
[107]	Van Kasteren [28]	97%	–	–	–
[119]		92%	–	–	–
[120]		78.9%	–	–	–
[121]		84%	–	–	–
[122]		89.5%	–	–	–
[123]		82%	–	–	–
[104]		98.14%	74.73%	76.29%	72%
[123]		77.10%	–	–	–
[124]	Casas Aruba [30]	74%	–	–	–
[125]		78%	–	–	–
[126]		98.93%	–	–	–
[127]		73.44%	–	–	–
[104]		98.14%	74.73%	76.29%	72%
[116]		94.21%	90.10%	93.11%	91%
[117]		94.62%	93.21%	94.62%	93%
[128]		91%	–	–	–
[129]	Casas Kyoto [29]	89%	–	–	–
[130]		81.1%	–	–	–
[131]		–	83.26%	–	–
[132]		87.45%	86.12%	–	–
[125]		78%	–	–	–
[104]		86.15%	59.18%	57.12%	57%
[131]		–	–	–	72%
[132]	Casas Tulum [67]	–	–	–	74%
[125]		–	65.3%	82%	–
[104]		75.45%	–	78%	–
[105]	Hh102 [68]	66%	–	–	53%
[105]	Hh104 [68]	78%	–	–	60%
[115]	UCI Human Activity Recognition (HAR) [90]	71%	–	–	–
[107]	MIT PlaceLab [97]	94.5%	–	–	–
[118]	PAMAP2 [92]	62%	–	–	–

Table 9. Details of the K-means experimentation.

References	Dataset	Accuracy	Precision	Recall	F-Measure
[46]		–	88.6%	95.48%	91.91%
[133]		87.21%	–	–	–
[134]	VanKasteren	82%	–	–	–
[135]	[28]	–	–	72%	85%
[136]		–	–	–	82.78%
[137]		–	76.23%	–	–
[138]					
[139]	WISDM [96]	71%	–	–	–
[140]	Liara [93]	86%	–	–	–
[109]		79%	–	–	–
[128]		80%	–	–	–
[115]		86.8%	–	–	–
[138]	Opportunity	–	79.67%	–	–
[141]	[33]	–	82.45%	–	–
[142]		–	–	75.45%	–
[143]		–	–	–	87.32%
[144]		–	–	–	85.45%
[139]		71.66%	–	–	–
[112]		71%	–	–	–
[145]		78.45%	–	–	–
[146]	MHealth [34]	–	–	–	78.56%
[147]		–	–	–	77.56%
[148]		73.45%	–	–	–
[149]		78.63%%	–	–	–
[150]		52.1%	–	–	–
[151]		76.32%	–	–	–
[152]	UCI HAR	–	–	–	77.22%
[153]	[32]	–	–	–	78.45%
[154]		79.37%	–	–	–
[155]		75.31%	–	–	–

Table 10. Details of the sub-clustering experimentation.

References	Dataset	Accuracy
[156]		94.3%
[157]		78.5%
[158]	VanKasteren [28]	75.42%
[159]		81.65%
[160]		86.32%
[161]		89.45%
[156]		91.88%
[161]		88.32%
[162]	Casas Aruba [30]	89.78%
[163]		87.67%
[164]		86.43%
[165]		89.12%
[156]		96.67%
[166]		86.32%
[167]	Casas Kyoto [29]	76.45%
[168]		89.12%
[169]		85.34%
[156]	Casas Tulum [67]	99.28%
[156]	Milan [68]	95.20%
[156]	Cairo [68]	94.17%

In Reference [115], the author showed the experimentation of the sub-clustering technique at different datasets, obtaining the following accuracy results: VanKasteren 94.3%, Casas Aruba 91.88%, Casas Kyoto 96.67%, Casas Tulum 99.28%, Milan 95.20%, and Cairo 94.17.

8. Conclusions

The objective of this article is to show to the community of researchers of recognition of human activities, or HAR, some recommendations about the different Clustering techniques used for the analysis of different types of dataset in an unsupervised experiment. It is important to point out that this review is framed in articles that were published between 2015 and 2020. The database that is currently publishing the largest number of articles focused on clustering applications to HAR is IEEE Xplorer, with 66% of the publications when comparing results with other databases, such as Scopus, Science Direct, Web of Science, and ACM.

Additionally, it is important to point out that the vast majority of the works, 70%, come from conference articles, while the rest of the articles have been published in journals or Book Series. The USA is the country that has accepted the largest number of publications in this area of knowledge, followed by Germany and the Netherlands. One of the most published journals in this line of knowledge is Sensors and the Journal of Ambient Intelligence and Humanized Computing. Among the institutions that stand out the most for the development of this theme are: WSU (USA) and the University of Ulster (Ireland).

Considering the datasets that have been used to carry out the different unsupervised experiments, VanKasteren has been the dataset most used by the authors (30%), followed by Washington State University with its CASAS project, where the datasets: Kyoto (13%), Tulum (7%), and Aruba (3%) are the ones that have been used by the authors.

Within the experiments analyzed, it is possible to identify the effectiveness of the clustering techniques or algorithms in unsupervised environments, as detailed in Table 11. Of all the datasets implemented, Casas Aruba, VanKasteren, and Casas Tulum have been the ones that have shown the best results in the unsupervised clustering implementation processes, with the following accuracies, 98.14%, 88.6%, and 99.28%, respectively.

Table 11. Details of the best results of the experiments.

References	Dataset	Technique	Accuracy
[110]	Casas Aruba [30]	K-NN	98.14%
[119]	VanKasteren [28]	K-means	88.6%
[125]	Casas Tulum [67]	Sub-Clustering	99.28%

The analysis described above allows for the analyzing of the wide usability of partition algorithms to analyze the dataset of activities of daily life, as well as their effectiveness in the evaluation of the quality metrics of the results of the experiments.

9. Future Works

Among the future works planned, after this systematic review of the literature, is to explore other types of techniques that have not been used as methods, based on hierarchies or kernel for the identification of this type of activity; likewise, is the objective to perform different experiments with different methods that support, with greater precision, the identification of the activities of daily life. Other aspects or challenges of this line or area of research can be highlighted:

- Usability of clustering techniques in conjunction with other techniques or algorithms, such as HMM, which support the unsupervised detection of daily life activities.
- Generation and use of new techniques to analyze temporal space support to improve the results of the identification of activities of daily life.

- Other challenges within the clustering application can identify the behavioral analysis of each of the groups generated. This analysis is called Multiclustering Methods, which creates multiple groupings and then combines them into a single result (see Figure 14).
- Exploration of different experimentation scenarios with multi-level applications that include the behavior of unidentified activities.
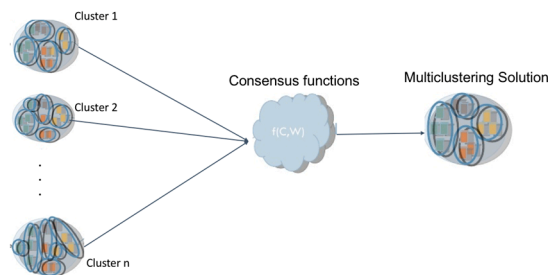

Figure 14. Multiclustering application architecture.

References

1. Jain, A.; Duin, R.; Mao, J. Statistical pattern recognition: A review. *IEEE Trans. Pattern Anal. Mach. Intell.* **2000**, *22*, 4–37. [CrossRef]
2. Jain, A.; Murty, M.; Flynn, P. Data clustering: A review. *ACM Comput. Surv.* **1999**, *31*, 264–323. [CrossRef]
3. Bishop, C. *Neural Networks for Pattern Recognition*; Oxford University Press: New York, NY, USA, 1995.
4. Sklansky, J.; Siedlecki, W. Large-scale feature selection. In *Handbook of Pattern Recognition and Computer Vision*; Chen, C., Pau, L., Wang, P., Eds.; World Scientific: Singapore, 1993; pp. 61–124.
5. Kleinberg, J. An impossibility theorem for clustering. In Proceedings of the 2002 15th International Conference on Neural Information Processing Systems, Vancouver, BC, Canada, 9–14 December 2002; Volume 15, pp. 463–470.
6. Jain, A.; Dubes, R. *Algorithms for Clustering Data*; Prentice-Hall: Englewood Cliffs, NJ, USA, 1988.
7. Gordon, A. Cluster validation. In *Data Science, Classification, and Related Methods*; Hayashi, C., Ohsumi, N., Yajima, K., Tanaka, Y., Bock, H., Bada, Y., Eds.; Springer: New York, NY, USA, 1998; pp. 22–39.
8. Dubes, R. Cluster analysis and related issue. In *Handbook of Pattern Recognition and Computer Vision*; Chen, C., Pau, L., Wang, P., Eds.; World Scientific: Singapore, 1993; pp. 3–32.
9. Bandyopadhyay, S.; Maulik, U. Nonparametric genetic clustering: Comparison of validity indices. *IEEE Trans. Syst. Man Cybern. C Appl. Rev.* **2001**, *31*, 120–125. [CrossRef]
10. Bezdek, J.; Pal, N. Some new indexes of cluster validity. *IEEE Trans. Syst. Man Cybern. B Cybern.* **1998**, *28*, 301–315. [CrossRef]
11. Dunn, J.C. A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. *J. Cybern.* **1973**, *3*, 32–57. [CrossRef]
12. Halkidi, M.; Batistakis, Y.; Vazirgiannis, M. Cluster validity methods: Part I & II. *SIGMOD Rec.* **2002**, *31*, 40–45.

13. Leung, Y.; Zhang, J.; Xu, Z. Clustering by scale-space filtering. *IEEE Trans. Pattern Anal. Mach. Intell.* **2000**, *22*, 1396–1410. [CrossRef]

14. Levine, E.; Domany, E. Resampling method for unsupervised estimation of cluster validity. *Neural Comput.* **2001**, *13*, 2573–2593. [CrossRef] [PubMed]

15. Davé, R.; Krishnapuram, R. Robust clustering methods: A unified view. *IEEE Trans. Fuzzy Syst.* **1997**, *5*, 270–293. [CrossRef]

16. Geva, A. Hierarchical unsupervised fuzzy clustering. *IEEE Trans. Fuzzy Syst.* **1999**, *7*, 723–733. [CrossRef]

17. Hammah, R.; Curran, J. Validity measures for the fuzzy cluster analysis of orientations. *IEEE Trans. Pattern Anal. Mach. Intell.* **2000**, *22*, 1467–1472. [CrossRef]

18. Rand, W.M. Objective criteria for the evaluation of clustering methods. *J. Am. Stat. Assoc.* **1971**, *66*, 846–850. [CrossRef]

19. Lane, N.; Miluzzo, E.; Lu, H.; Peebles, D.; Choudhury, T.; Campbell, A. A survey of mobile phone sensing. *IEEE Commun. Mag.* **2010**, *48*, 140–150. [CrossRef]

20. bin Abdullah, M.F.A.; Negara, A.F.P.; Sayeed, M.S.; Choi, D.J.; Muthu, K.S. Classification algorithms in human activity recognition using smartphones. *World Acad. Sci. Eng. Technol.* **2012**, *68*, 422–430.

21. Stikic, M.; Schiele, B. Activity Recognition from Sparsely Labeled Data Using Multi-Instance Learning. In Proceedings of the 4th International Symposium Location and Context Awareness, Tokyo, Japan, 7–8 May 2009; Volume 5561, pp. 156–173.

22. Chen, L.; Nugent, C.D. Ontology-based activity recognition in intelligent pervasive environments. *Int. J. Web Inf. Syst.* **2009**, *5*, 410–430. [CrossRef]

23. Palmes, P.; Pung, H.K.; Gu, T.; Xue, W.; Chen, S. Object relevance weight pattern mining for activity recognition and segmentation. *Pervasive Mob. Comput.* **2010**, *6*, 43–57. [CrossRef]

24. Chen, L.; Nugent, C.D.; Wang, H. A Knowledge-Driven Approach to Activity Recognition in Smart Homes. *IEEE Trans. Knowl. Data Eng.* **2011**, *24*, 961–974. [CrossRef]

25. Ye, J.; Stevenson, G.; Dobson, S. A top-level ontology for smart environments. *Pervasive Mob. Comput.* **2011**, *7*, 359–378. [CrossRef]

26. Jain, A.K.; Flynn, P. Image segmentation using clustering. In *Advances in Image Understanding*; IEEE Computer Society Press: Piscataway, NJ, USA, 1996.

27. Van Kasteren, T.L.M.; Englebienne, G.; Kröse, B.J.A. Activity recognition using semi-Markov models on real world smart home datasets. *J. Ambient Intell. Smart Environ.* **2010**, *2*, 311–325. [CrossRef]

28. Cook, D.; Crandall, A.S.; Thomas, B.L.; Krishnan, N.C. CASAS: A smart home in a box. *Computer* **2013**, *46*, 62–69. [CrossRef]

29. Cook, D. Learning setting-generalized activity models for smart spaces. *IEEE Intell. Syst.* **2012**, *27*, 32–38. [CrossRef] [PubMed]

30. Singla, G.; Cook, D.J.; Schmitter-Edgecombe, M. Recognizing independent and joint activities among multiple residents in smart environments. *J. Ambient Intell. Hum. Comput.* **2010**, *1*, 57–63. [CrossRef] [PubMed]

31. Almaslukh, B.; AlMuhtadi, J.; Artoli, A. An effective deep autoencoder approach for online smartphone-based human activity recognition. *Int. J. Comput. Sci. Netw. Secur.* **2017**, *17*, 160–165.

32. Ordóñez, F.; Roggen, D. Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition. *Sensors* **2016**, *16*, 115. [CrossRef] [PubMed]

33. Ha, S.; Choi, S. Convolutional neural networks for human activity recognition using multiple accelerometer and gyroscope sensors. In Proceedings of the 2016 International Joint Conference on Neural Networks (IJCNN), Vancouver, BC, Canada, 24–29 July 2016; pp. 381–388.

34. Frigui, H.; Krishnapuram, R. A robust competitive clustering algorithm with applications in computer vision. *IEEE Trans. Pattern Anal. Mach. Intell.* **1999**, *21*, 450–465. [CrossRef]

35. Shi, J.; Malik, J. Normalized cuts and image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2000**, *22*, 888–905.

36. Iwayama, M.; Tokunaga, T. Cluster-based text categorization: A comparison of category search strategies. In Proceedings of the 18th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Seattle, WA, USA, 9–13 July 1995; pp. 273–280.

37. Sahami, M. Using Machine Learning to Improve Information Access. Ph.D. Thesis, Stanford University, Stanford, CA, USA, 15 December 1998.

38. Bhatia, S.K.; Deogun, J.S. Conceptual clustering in information retrieval. *IEEE Trans. Syst. Man Cybern. Part B Cybern.* **1998**, *28*, 427–436. [CrossRef]

39. Hubert, L.; Arabie, P. The analysis of proximity matrices through sums of matrices having (anti-) Robinson forms. *Br. J. Math. Stat. Psychol.* **1994**, *47*, 1–40. [CrossRef]

40. Hu, W.Y.; Scott, J.S. Behavioral obstacles in the annuity market. *Financ. Anal. J.* **2007**, *63*, 71–82. [CrossRef]

41. Hung, S.P.; Baldi, P.; Hatfield, G.W. Global Gene Expression Profiling in *Escherichia coli*K12 THE EFFECTS OF LEUCINE-RESPONSIVE REGULATORY PROTEIN. *J. Biol. Chem.* **2002**, *277*, 40309–40323. [CrossRef]

42. 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]

43. Rawassizadeh, R.; Dobbins, C.; Akbari, M.; Pazzani, M. Indexing multivariate mobile data through spatio-temporal event detection and clustering. *Sensors* **2019**, *19*, 448. [CrossRef]

44. Bouchard, K.; Lapalu, J.; Bouchard, B.; Bouzouane, A. Clustering of human activities from emerging movements. *J. Ambient Intell. Hum. Comput.* **2019**, *10*, 3505–3517. [CrossRef]

45. Jain, A.K. Data clustering: 50 years beyond K-means. *Pattern Recognit. Lett.* **2010**, *31*, 651–666. [CrossRef]

46. Drineas, P.; Frieze, A.M.; Kannan, R.; Vempala, S.; Vinay, V. Clustering in Large Graphs and Matrices. In Proceeding of the Symposium on Discrete Algorithms (SODA), Baltimore, MD, USA, 17–19 January 1999; Volume 99, pp. 291–299.

47. Gonzalez, T.F. Clustering to minimize the maximum intercluster distance. *Theor. Comput. Sci.* **1985**, *38*, 293–306. [CrossRef]

48. Fisher, D.H. Knowledge acquisition via incremental conceptual clustering. *Mach. Learn.* **1987**, *2*, 139–172. [CrossRef]

49. Gennari, J.H.; Langley, P.; Fisher, D. Models of incremental concept formation. *Artif. Intel.* **1989**, *40*, 11–61. [CrossRef]

50. Aguilar-Martin, J.; De Mantaras, R.L. The Process of Classification and Learning the Meaning of Linguistic Descriptors of Concepts. *Approx. Reason. Decis. Anal.* **1982**, *1982*, 165–175.

51. Omran, M.G.; Engelbrecht, A.P.; Salman, A. An overview of clustering methods. *Intell. Data Anal.* **2007**, *11*, 583–605. [CrossRef]

52. Bezdek, J.C. *Pattern Recognition with Fuzzy Objective Function Algorithms*; Plenum Publishing Corporation: New York, NY, USA, 1981.

53. Tamayo, P.; Slomin, D.; Mesirov, J.; Zhu, Q.; Kitareewan, S.; Dmitrovsky, E.; Lander, E.S.; Golub, T.R. Interpreting patterns of gene expresión with self-organizing map: Methos and application to hematopoietic differentiation. *Proc. Natl Acad. Sci. USA* **1999**, *96*, 2901–2912. [CrossRef]

54. Toronen, P.; Kolehmainen, M.; Wong, G.; Catrén, E. Analysis of gene expresión data using self-organizing maps. *FEBS Lett.* **1999**, *451*, 142–146. [CrossRef]

55. Kohonen, T. *Self-Organizing Maps*; Springer: Berlin, Germany, 1997.

56. Goldberg, D. *Genetic Algorithms in Search, Optimization, and Machine Learning*; Addison Wesley: Reading, MA, USA, 1989.

57. Holland, J.H. *Hidden Orderhow Adaptation Builds Complexity*; Helix Books: Totowa, NJ, USA, 1995.

58. Fogel, L.J.; Owens, A.J.; Walsh, M.J. *Artificial Intelligence through Simulated Evolution*; Wiley: Chichester, WS, UK, 1966.

59. Fogel, D.B. *Evolutionary Computation: Toward a New Philosophy of Machine Intelligence*; John Wiley & Sons: Hoboken, NJ, USA, 2006.

60. Holland, J. Genetic algorithms. *Sci. Am.* **1992**, *267*, 66–73. [CrossRef]

61. Davis, L. (Ed.) *Handbook of Genetic Algorithms*; Van Nostrand Reinhold: New York, NY, USA, 1991; p. 385.

62. Chung, F.R.K. *Spectral Graph Theory, CBMS Regional Conference Series in Mathematics*; American Mathematical Society: Providence, RI, USA, 1997; Volume 92.

63. Fiedler, M. Algebraic connectivity of graphs. *Czechoslov. Math. J.* **1973**, *23*, 298–305.

64. Schölkopf, B.; Smola, A.J.; Müller, K.R. Nonlinear component analysis as a kernel eigenvalue problem. *Neural Comput.* **1998**, *10*, 1299–1319. [CrossRef]

65. Girolami, M. Mercer kernel based clustering in feature space. *IEEE Trans. Neural Netw.* **2002**, *13*, 780–784. [CrossRef]

66. Ng, A.Y.; Jordan, M.I.; Weiss, Y. On Spectral Clustering: Analysis and an algorithm. In *Advances in Neural Information Processing Systems*; MIT Press: Cambridge, MA, USA, 2001.

67. Rechenber, I. Evolution strategy. In *Computational Intelligence: Imitating Life*; Zurada, J.M., Marks, R.J., Robinson, C., Eds.; IEEE Press: Piscataway, NJ, USA, 1994.

68. Schwefel, H.-P. *Evolution and Optimum Seeking*; Wiley: New York, NY, USA, 1995.

69. Koza, J.R. *Genetic Programming: On the Programming of Computers by Means of Natural Selection*; MIT Press: Cambridge, MA, USA, 1992.

70. Von Luxburg, U. A tutorial on spectral clustering. *Stat. Comput.* **2007**, *17*, 395–416. [CrossRef]

71. Kohonen, T. The self-organizing map. *Proc. IEEE* **1990**, *78*, 1464–1480. [CrossRef]

72. Howedi, A.; Lotfi, A.; Pourabdollah, A. Exploring Entropy Measurements to Identify Multi-Occupancy in Activities of Daily Living. *Entropy* **2019**, *21*, 416. [CrossRef]

73. Cook, D.; Schmitter-Edgecombe, M. Assessing the quality of activities in a smart environment. *Methods Inf. Med.* **2009**, *48*, 480–485.

74. Singla, G.; Cook, D.; Schmitter-Edgecombe, M. Tracking activities in complex settings using smart environment technologies. *Int. J. BioSci. Psychiatry Technol.* **2009**, *1*, 25–35.

75. Cook, D.J.; Youngblood, M.; Das, S.K. A multi-agent approach to controlling a smart environment. In *Designing Smart Homes*; Springer: Berlin/Heidelberg, Germany, 2006; pp. 165–182.

76. Dernbach, S.; Das, B.; Krishnan, N.C.; Thomas, B.L.; Cook, D.J. Simple and complex activity recognition through smart phones. In Proceedings of the 2012 Eighth International Conference on Intelligent Environments, Guanajuato, Mexico, 26–29 June 2012.

77. Sahaf, Y. Comparing Sensor Modalities for Activity Recognition. Master's Thesis, Washington State University, Pullman, WA, USA, August 2011.

78. Rawassizadeh, R.; Keshavarz, H.; Pazzani, M. Ghost imputation: Accurately reconstructing missing data of the off period. *IEEE Trans. Knowl. Data Eng.* **2019**. [CrossRef]

79. Wilson, D.H. Assistive Intelligent Environments for Automatic Health Monitoring. Ph.D. Thesis, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, USA, September 2005.

80. Singla, G.; Cook, D.J.; Schmitter-Edgecombe, M. Incorporating temporal reasoning into activity recognition for smart home residents. In Proceedings of the AAAI Workshop on Spatial and Temporal Reasoning, Chicago, IL, USA, 13 July 2008; pp. 53–61.

81. Wren, C.R.; Tapia, E.M. Hierarchical Processing in Scalable and Portable Sensor Networks for Activity Recognition. U.S. Patent No. 7359836, 15 April 2008.

82. Van Kasteren, T.; Noulas, A.; Englebienne, G.; Kröse, B. Accurate activity recognition in a home setting. In Proceedings of the 10th International Conference on Ubiquitous Computing, Seoul, Korea, 21–24 September 2008; pp. 1–9.

83. Philipose, M.; Fishkin, K.P.; Perkowitz, M.; Patterson, D.J.; Fox, D.; Kautz, H.; Hahnel, D. Inferring activities from interactions with objects. *IEEE Pervasive Comput.* **2004**, *3*, 50–57. [CrossRef]

84. Patterson, D.J.; Fox, D.; Kautz, H.; Philipose, M. Fine-grained activity recognition by aggregating abstract object usage. In Proceedings of the Ninth IEEE International Symposium on Wearable Computers (ISWC'05), Osaka, Japan, 18–21 October 2005; pp. 44–51.

85. Hodges, M.R.; Newman, M.W.; Pollack, M.E. Object-Use Activity Monitoring: Feasibility for People with Cognitive Impairments. In Proceedings of the AAAI Spring Symposium: Human Behavior Modeling, Stanford, CA, USA, 23–25 March 2009; pp. 13–18.

86. Fang, F.; Aabith, S.; Homer-Vanniasinkam, S.; Tiwari, M.K. High-resolution 3D printing for healthcare underpinned by small-scale fluidics. In *3D Printing in Medicine*; Woodhead Publishing: Cambrigde, MA, USA, 2017; pp. 167–206.

87. Veltink, P.H.; Bussmann, H.J.; De Vries, W.; Martens, W.J.; Van Lummel, R.C. Detection of static and dynamic activities using uniaxial accelerometers. *IEEE Trans. Rehab. Eng.* **1996**, *4*, 375–385. [CrossRef] [PubMed]

88. Mathie, M.J.; Coster, A.C.F.; Lovell, N.H.; Celler, B.G. Detection of daily physical activities using a triaxial accelerometer. *Med. Biol. Eng. Comput.* **2003**, *41*, 296–301. [CrossRef] [PubMed]

89. Bao, L.; Intille, S.S. Activity recognition from user-annotated acceleration data. In Proceedings of the International Conference on Pervasive Computing, Linz/Vienna, Austria, 21–23 April 2004; Springer: Berlin/Heidelberg, Germany, 2004; pp. 1–17.

90. Chambers, G.S.; Venkatesh, S.; West, G.A.; Bui, H.H. Hierarchical recognition of intentional human gestures for sports video annotation. In *Object Recognition Supported by User Interaction for Service Robots*; IEEE: Quebec, PQ, Canada, 2002; pp. 1082–1085.

91. Lester, J.; Choudhury, T.; Borriello, G. A practical approach to recognizing physical activities. In Proceedings of the 4th International Conference on Pervasive Computing, Dublin, Ireland, 7–10 May 2006; Springer: Berlin/Heidelberg, Germany, 2006; pp. 1–16.

92. Mantyjarvi, J.; Himberg, J.; Seppanen, T. Recognizing human motion with multiple acceleration sensors. In Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, Tucson, AZ, USA, 7–10 October 2001; pp. 747–752.

93. Al-Ani, T.; Le Ba, Q.T.; Monacelli, E. On-line automatic detection of human activity in home using wavelet and hidden markov models scilab toolkits. In Proceedings of the 2007 IEEE International Conference on Control Applications, Singapore, 1–3 October 2007; pp. 485–490.

94. Zheng, Y.; Liu, Q.; Chen, E.; Ge, Y.; Zhao, J.L. Time series classification using multi-channels deep convolutional neural networks. In Proceedings of the International Conference on Web-Age Information Management, Macau, China, 16–18 June 2014; Springer: Cham, Switzerland, 2014; pp. 298–310.

95. Jiang, W.; Yin, Z. Human activity recognition using wearable sensors by deep convolutional neural networks. In Proceedings of the 23rd ACM International Conference on Multimedia, Brisbane, Australia, 26–30 October 2015; pp. 1307–1310.

96. Alsheikh, M.A.; Selim, A.; Niyato, D.; Doyle, L.; Lin, S.; Tan, H.P. Deep activity recognition models with triaxial accelerometers. In Proceedings of the Workshops at the Thirtieth AAAI Conference on Artificial Intelligence, Phoenix, AZ, USA, 12–13 February 2016.

97. Hu, D.H.; Yang, Q. CIGAR: Concurrent and Interleaving Goal and Activity Recognition. In Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence, Chicago, IL, USA, 13–17 July 2008; Volume 8, pp. 1363–1368.

98. Zhang, L.; Wu, X.; Luo, D. Recognizing human activities from raw accelerometer data using deep neural networks. In Proceedings of the 2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA), Miami, FL, USA, 9–11 December 2015; pp. 865–870.

99. Using the Multicom Domus Dataset 2011. Available online: https://hal.archives-ouvertes.fr/hal-01473142/ (accessed on 25 January 2020).

100. Zhang, S.; McCullagh, P.; Nugent, C.; Zheng, H. Activity monitoring using a smart phone's accelerometer with hierarchical classification. In Proceedings of the 2010 Sixth International Conference on Intelligent Environments, Kuala Lumpur, Malaysia, 19–21 July 2010; pp. 158–163.

101. Espinilla, M.; Martínez, L.; Medina, J.; Nugent, C. The experience of developing the UJAmI Smart lab. *IEEE Access* **2018**, *6*, 34631–34642. [CrossRef]

102. Kitchenham, B.; Brereton, O.P.; Budgen, D.; Turner, M.; Bailey, J.; Linkman, S. Systematic literature reviews in software engineering—A systematic literature review. *Inf. Softw. Technol.* **2009**, *51*, 7–15. [CrossRef]

103. Fahad, L.G.; Ali, A.; Rajarajan, M. Learning models for activity recognition in smart homes. In *Information Science and Applications*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 819–826.

104. Nguyen, D.; Le, T.; Nguyen, S. An Algorithmic Method of Calculating Neighborhood Radius for Clustering In-home Activities within Smart Home Environment. In Proceedings of the 2016 7th International Conference on Intelligent Systems, Modelling and Simulation (ISMS), Bangkok, Thailand, 25–27 January 2016; pp. 42–47.

105. Nguyen, D.; Le, T.; Nguyen, S. A Novel Approach to Clustering Activities within Sensor Smart Homes. *Int. J. Simul. Syst. Sci. Technol.* **2016**, *17*. [CrossRef]

106. Sukor, A.S.A.; Zakaria, A.; Rahim, N.A.; Setchi, R. Semantic knowledge base in support of activity recognition in smart home environments. *Int. J. Eng. Technol.* **2018**, *7*, 67–72. [CrossRef]

107. Jänicke, M.; Sick, B.; Tomforde, S. Self-adaptive multi-sensor activity recognition systems based on gaussian mixture models. *Informatics* **2018**, *5*, 38. [CrossRef]

108. Honarvar, A.R.; Zaree, T. Frequent sequence pattern based activity recognition in smart environment. *Intell. Decis. Technol.* **2018**, *12*, 349–357. [CrossRef]

109. Chen, W.H.; Chen, Y. An ensemble approach to activity recognition based on binary sensor readings. In Proceedings of the 2017 IEEE 19th International Conference on e-Health Networking, Applications and Services (Healthcom), Dalian, China, 12–15 October 2017; pp. 1–5.

110. Khan, M.A.A.H.; Roy, N. Transact: Transfer learning enabled activity recognition. In Proceedings of the 2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), Kona, HI, USA, 13–17 March 2017; pp. 545–550.

111. Fahad, L.G.; Tahir, S.F.; Rajarajan, M. Feature selection and data balancing for activity recognition in smart homes. In Proceedings of the 2015 IEEE International Conference on Communications (ICC), London, UK, 8–12 June 2015; pp. 512–517.

112. Fahad, L.G.; Khan, A.; Rajarajan, M. Activity recognition in smart homes with self verification of assignments. *Neurocomputing* **2015**, *149*, 1286–1298. [CrossRef]

113. Bota, P.; Silva, J.; Folgado, D.; Gamboa, H. A Semi-Automatic Annotation Approach for Human Activity Recognition. *Sensors* **2019**, *19*, 501. [CrossRef] [PubMed]

114. Zhang, S.; Ng, W.W.; Zhang, J.; Nugent, C.D.; Irvine, N.; Wang, T. Evaluation of radial basis function neural network minimizing L-GEM for sensor-based activity recognition. *J. Ambient Intell. Hum. Comput.* **2019**. [CrossRef]

115. Wen, J.; Zhong, M. Activity discovering and modelling with labelled and unlabelled data in smart environments. *Expert Syst. Appl.* **2015**, *42*, 5800–5810. [CrossRef]

116. Fahad, L.G.; Rajarajan, M. Integration of discriminative and generative models for activity recognition in smart homes. *Appl. Soft Comput.* **2015**, *37*, 992–1001. [CrossRef]

117. Ihianle, I.; Naeem, U.; Islam, S.; Tawil, A.R. A hybrid approach to recognising activities of daily living from object use in the home environment. *Informatics* **2018**, *5*, 6. [CrossRef]

118. Chua, S.L.; Foo, L.K. Sensor selection in smart homes. *Procedia Comput. Sci.* **2015**, *69*, 116–124. [CrossRef]

119. Shahi Soozaei, A. Human Activity Recognition in Smart Homes. Ph.D. Thesis, University of Otago, Dunedin, New Zealand, January 2019.

120. Caldas, T.V. From Binary to Multi-Class Divisions: Improvements on Hierarchical Divisive Human Activity Recognition. Master's Thesis, Universidade do Porto, Oporto, Portugal, July 2019.

121. Fang, L.; Ye, J.; Dobson, S. Discovery and recognition of emerging human activities using a hierarchical mixture of directional statistical models. *IEEE Trans. Knowl. Data Eng.* **2019**. [CrossRef]

122. Guo, J.; Li, Y.; Hou, M.; Han, S.; Ren, J. Recognition of Daily Activities of Two Residents in a Smart Home Based on Time Clustering. *Sensors* **2020**, *20*, 1457. [CrossRef]

123. Kavitha, R.; Binu, S. Performance Evaluation of Area-Based Segmentation Technique on Ambient Sensor Data for Smart Home Assisted Living. *Procedia Comput. Sci.* **2019**, *165*, 314–321. [CrossRef]

124. Akter, S.S. Improving Sensor Network Predictions through the Identification of Graphical Features. Ph.D. Thesis, Washington State University, Washington, DC, USA, August 2019.

125. Oukrich, N. Daily Human Activity Recognition in Smart Home based on Feature Selection, Neural Network and Load Signature of Appliances. Ph.D. Thesis, Mohammed V University In Rabat, Rabat, Morocco, April 2019.

126. Yala, N. Contribution aux Méthodes de Classification de Signaux de Capteurs dans un Habitat Intelligent. Ph.D. Thesis, The University of Science and Technology—Houari Boumediene, Bab-Ezzouar, Algeria, October 2019.

127. Lyu, F.; Fang, L.; Xue, G.; Xue, H.; Li, M. Large-Scale Full WiFi Coverage: Deployment and Management Strategy Based on User Spatio-Temporal Association Analytics. *IEEE Internet Things J.* **2019**, *6*, 9386–9398. [CrossRef]

128. Chetty, G.; White, M. Body sensor networks for human activity recognition. In Proceedings of the 2016 3rd International Conference on Signal Processing and Integrated Networks (SPIN), Noida, India, 11–12 February 2016; pp. 660–665.

129. Singh, T.; Vishwakarma, D.K. Video benchmarks of human action datasets: A review. *Artif. Intell. Rev.* **2019**, *52*, 1107–1154. [CrossRef]

130. Senda, M.; Ha, D.; Watanabe, H.; Katagiri, S.; Ohsaki, M. Maximum Bayes Boundary-Ness Training for Pattern Classification. In Proceedings of the 2019 2nd International Conference on Signal Processing and Machine Learning, Hangzhou, China, 27–29 November 2019; pp. 18–28.

131. Petrovich, M.; Yamada, M. Fast local linear regression with anchor regularization. *arXiv* **2020**, arXiv:2003.05747.

132. Yadav, A.; Kumar, E. *A Literature Survey on Cyber Security Intrusion Detection Based on Classification Methods of Supervised Machine Learning*; Bloomsbury: New Delhi, India, 2019.

133. Marimuthu, P.; Perumal, V.; Vijayakumar, V. OAFPM: Optimized ANFIS using frequent pattern mining for activity recognition. *J. Supercomput.* **2019**, *75*, 5347–5366. [CrossRef]

134. Raeiszadeh, M.; Tahayori, H.; Visconti, A. Discovering varying patterns of Normal and interleaved ADLs in smart homes. *Appl. Intell.* **2019**, *49*, 4175–4188. [CrossRef]

135. Hossain, H.S.; Roy, N. Active Deep Learning for Activity Recognition with Context Aware Annotator Selection. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Anchorage, AK, USA, 4–8 August 2019; pp. 1862–1870.

136. Kong, D.; Bao, Y.; Chen, W. Collaborative learning based on centroid-distance-vector for wearable devices. *Knowl.-Based Syst.* **2020**. [CrossRef]

137. 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]

138. Arya, M. Automated Detection of Acute Leukemia Using K-Means Clustering Algorithm. Master's Thesis, North Dakota State University, Fargo, ND, USA, May 2019.

139. Chetty, G.; Yamin, M. Intelligent human activity recognition scheme for eHealth applications. *Malays. J. Comput. Sci.* **2015**, *28*, 59–69.

140. Soulas, J.; Lenca, P.; Thépaut, A. Unsupervised discovery of activities of daily living characterized by their periodicity and variability. *Eng. Appl. Artif. Intell.* **2015**, *45*, 90–102. [CrossRef]

141. Rojlertjanya, P. Customer Segmentation Based on the RFM Analysis Model Using K-Means Clustering Technique: A Case of IT Solution and Service Provider in Thailand. Master's Thesis, Bangkok University, Bangkok, Thailand, 16 August 2019.

142. Zhao, B.; Shao, B. Analysis the Consumption Behavior Based on Weekly Load Correlation and K-means Clustering Algorithm. In Proceedings of the International Conference on Advanced Intelligent Systems and Informatics, Cairo, Egypt, 26–28 October 2019; Springer: Cham, Switzerland, 2019; pp. 70–81.

143. Zahi, S.; Achchab, B. Clustering of the population benefiting from health insurance using K-means. In Proceedings of the 4th International Conference on Smart City Applications, Casablanca, Morocco, 2–4 October 2019; pp. 1–6.

144. Dana, R.D.; Dikananda, A.R.; Sudrajat, D.; Wanto, A.; Fasya, F. Measurement of health service performance through machine learning using clustering techniques. *J. Phys. Conf. Ser.* **2019**, *1360*, 012017. [CrossRef]

145. Baek, J.W.; Kim, J.C.; Chun, J.; Chung, K. Hybrid clustering based health decision-making for improving dietary habits. *Technol. Health Care* **2019**, *27*, 459–472. [CrossRef] [PubMed]

146. Rashid, J.; Shah, A.; Muhammad, S.; Irtaza, A. A novel fuzzy k-means latent semantic analysis (FKLSA) approach for topic modeling over medical and health text corpora. *J. Intell. Fuzzy Syst.* **2019**, *37*, 6573–6588. [CrossRef]

147. Lütz, E. Unsupervised Machine Learning to Detect Patient Subgroups in Electronic Health Records. Master's Thesis, KTH Royal Institute of Technology, Stockholm, Sweden, January 2019.

148. Maturo, F.; Ferguson, J.; Di Battista, T.; Ventre, V. A fuzzy functional k-means approach for monitoring Italian regions according to health evolution over time. *Soft Comput.* **2019**. [CrossRef]

149. Wang, S.; Li, M.; Hu, N.; Zhu, E.; Hu, J.; Liu, X.; Yin, J. K-means clustering with incomplete data. *IEEE Access* **2019**, *7*, 69162–69171. [CrossRef]

150. Long, J.; Sun, W.; Yang, Z.; & Raymond, O.I. Asymmetric Residual Neural Network for Accurate Human Activity Recognition. *Information* **2019**, *10*, 203. [CrossRef]

151. Yuan, C.; Yang, H. Research on K-value selection method of K-means clustering algorithm. *J. Multidiscip. Sci. J.* **2019**, *2*, 226–235. [CrossRef]

152. Wang, P.; Shi, H.; Yang, X.; Mi, J. Three-way k-means: Integrating k-means and three-way decision. *Int. J. Mach. Learn. Cybern.* **2019**, *10*, 2767–2777. [CrossRef]

153. Sadeq, S.; Yetkin, G. Semi-Supervised Sparse Data Clustering Performance Investigation. In Proceedings of the International Conference on Data Science, MachineLearning and Statistics, Van, Turkey, 26–29 June 2019; p. 463.

154. Boddana, S.; Talla, H. Performance Examination of Hard Clustering Algorithm with Distance Metrics. *Int. J. Innov. Technol. Explor. Eng.* **2019**, *9*. [CrossRef]

155. Xiao, Y.; Chang, Z.; Liu, B. An efficient active learning method for multi-task learning. *Knowl. Based Syst.* **2020**, *190*, 105137. [CrossRef]

156. Yao, L.; Nie, F.; Sheng, Q.Z.; Gu, T.; Li, X.; Wang, S. Learning from less for better: Semi-supervised activity recognition via shared structure discovery. In Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, Heidelberg, Germany, 12–16 September 2016; pp. 13–24.

157. Tax, N.; Sidorova, N.; van der Aalst, W.M. Discovering more precise process models from event logs by filtering out chaotic activities. *J. Intell. Inf. Syst.* **2019**, *52*, 107–139. [CrossRef]

158. Artavanis-Tsakonas, K.; Karpiyevich, M.; Adjalley, S.; Mol, M.; Ascher, D.; Mason, B.; van der Heden van Noort, G.; Laman, H.; Ovaa, H.; Lee, M. Nedd8 hydrolysis by UCH proteases in Plasmodium parasites. *PLoS Pathog.* **2019**, *15*, e1008086.

159. Koole, G. *An Introduction to Business Analytics*; MG Books: Amsterdam, The Netherlands, 2019.

160. Oh, H.; Jain, R. Detecting Events of Daily Living Using Multimodal Data. *arXiv* **2019**, arXiv:1905.09402.

161. Caleb-Solly, P.; Gupta, P.; McClatchey, R. Tracking changes in user activity from unlabelled smart home sensor data using unsupervised learning methods. *Neural Comput. Appl.* **2020**. [CrossRef]

162. Patel, A.; Shah, J. Sensor-based activity recognition in the context of ambient assisted living systems: A review. *J. Ambient Intell. Smart Environ.* **2019**, *11*, 301–322. [CrossRef]

163. Leotta, F.; Mecella, M.; Sora, D. Visual process maps: A visualization tool for discovering habits in smart homes. *J. Ambient Intell. Hum. Comput.* **2019**. [CrossRef]

164. Ferilli, S.; Angelastro, S. Activity prediction in process mining using the WoMan framework. *J. Intell. Inf. Syst.* **2019**, *53*, 93–112. [CrossRef]

165. Wong, W. Combination Clustering: Evidence Accumulation Clustering for Dubious Feature Sets. *OSF Prepr.* **2019**. [CrossRef]

166. Wong, W.; Tsuchiya, N. *Evidence Accumulation Clustering Using Combinations of Features*; Center for Open Science: Victoria, Australia, 2019.

167. Zhao, W.; Li, P.; Zhu, C.; Liu, D.; Liu, X. Defense Against Poisoning Attack via Evaluating Training Samples Using Multiple Spectral Clustering Aggregation Method. *CMC-Comput. Mater. Cont.* **2019**, *59*, 817–832. [CrossRef]

168. Yang, Y.; Zheng, K.; Wu, C.; Niu, X.; Yang, Y. Building an effective intrusion detection system using the modified density peak clustering algorithm and deep belief networks. *Appl. Sci.* **2019**, *9*, 238. [CrossRef]

169. Cuzzocrea, A.; Gaber, M.M.; Fadda, E.; Grasso, G.M. An innovative framework for supporting big atmospheric data analytics via clustering-based spatio-temporal analysis. *J. Ambient Intell. Hum. Comput.* **2019**, *10*, 3383–3398. [CrossRef]