



# Perspective The Lifespan of Human Activity Recognition Systems for Smart Homes

Shruthi K. Hiremath <sup>†</sup> and Thomas Plötz \*,<sup>†</sup>

School of Interactive Computing, Georgia Institute of Technology, Atlanta, GA 30308, USA; shiremath9@gatech.edu

- \* Correspondence: thomas.ploetz@gatech.edu
- <sup>+</sup> These authors contributed equally to this work.

**Abstract:** With the growing interest in smart home environments and in providing seamless interactions with various smart devices, robust and reliable human activity recognition (HAR) systems are becoming essential. Such systems provide automated assistance to residents or to longitudinally monitor their daily activities for health and well-being assessments, as well as for tracking (longterm) behavior changes. These systems thus contribute towards an understanding of the health and continued well-being of residents. Smart homes are personalized settings where residents engage in everyday activities in their very own idiosyncratic ways. In order to provide a fully functional HAR system that requires minimal supervision, we provide a systematic analysis and a technical definition of the lifespan of activity recognition systems for smart homes. Such a designed lifespan provides for the different phases of building the HAR system, where these different phases are motivated by an application scenario that is typically observed in the home setting. Through the aforementioned phases, we detail the technical solutions that are required to be developed for each phase such that it becomes possible to derive and continuously improve the HAR system through data-driven procedures. The detailed lifespan can be used as a framework for the design of state-of-the-art procedures corresponding to the different phases.

**Keywords:** smart homes; lifespan; data-driven procedures; minimal supervision; human activity recognition

# 1. Introduction

Developing human activity recognition (HAR) systems is at the core of ubiquitous computing systems. With the recent resurgence of interest in smart home environments and with the availability of technology to make smart devices helpful and engaging for their end users [1,2], numerous applications that provide assistance to residents are now available. Owing to the reduced cost of sensors and the advancements in Internet of Things (IoT) technologies, as well as the accessibility of reliable and inexpensive sensing and computing technology, instrumenting homes with sensors for everyday activity recognition in real-world living environments is now a realistic option for many. This has encouraged, for example, the widespread use of pervasive sensing devices in ambient assisted living (AAL) environments. Initial works in AAL were focused on providing care for the elderly population [3,4], but, with the aforementioned advancements, the availability of such systems for diverse populations is now possible. Although systems developed earlier faced technical challenges corresponding to data collection and automated analysis [5–8], with the present advancements, this is not an issue that current deployments have to encounter. Even when such advancements have made the data collection process seamless and straightforward, substantial challenges remain in developing and deploying HAR systems in smart homes.

"Off-the-shelf" human activity recognition systems are desirable in smart homes, since they promise to be deployable "as-is" and immediately into any environment while



**Citation:** Hiremath, S.K.; Plötz, T. The Lifespan of Human Activity Recognition Systems for Smart Homes. *Sensors* **2023**, *23*, 7729. https://doi.org/10.3390/s23187729

Academic Editors: Anthony Fleury, Paolo Crippa and Giorgio Biagetti

Received: 10 July 2023 Revised: 15 August 2023 Accepted: 5 September 2023 Published: 7 September 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/). not requiring additional resources in terms of cost, effort, or time. However, given the individualized settings and the idiosyncratic behaviors of residents, it is not realistic to assume that such systems will work without the need for adaptation to such individualistic environments. Prior work shows that these adaptive procedures require experts in the loop to identify sensor grouping based on the location, function, and mapping of similar types of sensors between the source and target domains, without guarantees of obtaining optimal performance in the target environment [9–11]. Additionally, the goal to provide a fully functional system, requiring sample-precise annotations and the development of tailored HAR systems, that cater to specific home settings and individuals is realistic.

In this "perspectives" article, we define the lifespan of a HAR system for smart homes, i.e., we conceptualize how such "bespoke" systems can be derived for a range of practical application scenarios of sensor-based human activity recognition in smart homes. We do so by analyzing relevant related work and contextualizing the derived concept—the "lifespan"—with respect to realistic application scenarios. This addresses the concerns arising from the unavailability of an "off-the-shelf" system, which is often hard to achieve. Thus we accomplish the following: (i) discuss the challenges in developing HAR systems for smart homes and state-of-the-art procedures; (ii) detail a typical application scenario in the smart home, which serves as a guideline for detailing the lifespan of HAR system; and (iii) provide the technical details of the different components that make up the lifespan of the HAR system.

### 2. Background

With decreasing sensor costs, automating 'regular' homes has become a possibility for many. IoT or environmental sensors can collect data for extended periods of time without concerns regarding battery recharging or privacy. Such sensors capture data on the detecting motion [12] or on the interaction with objects that are instrumented with sensing capabilities [13]. As such, these advances have enabled the recognition of the activities of daily living for ambient assisted living (AAL) environments. Such automated assessment systems are required for home automation [14].

#### 2.1. Smart Homes

Mark Weiser, in his seminal paper 'The Computer for the 21st Century', stated

"The most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it." [15]

Toward this goal, researchers in ubiquitous computing aim to enable and provide computing away from the desktop. A unifying research theme that arose out of this was to focus on the computing needs in everyday lives, especially those that were away from work or office spaces. This initiated the research effort towards investigating computing in the home [16]. Smart Homes are a branch of ubiquitous computing that involve incorporating "intelligence" into living spaces. Smart homes for health care have been described as "a home equipped with smart sensors such as Bluetooth, Wi-Fi, or similar technology, not restricted to IoT, to automate, regulate, and monitor home occupants' physical health, mental health, and environments within the home. The focus must be on convenience, safety, and improvement of one's quality of life, to address the needs of the individuals, caregivers, and health professionals" [17]. Through ambient intelligence systems, the goal is to monitor smart homes and provide the control of home appliances and devices to users such that it enables them to execute tasks automatically. Ambient intelligence provides for instrumented environments that are sensitive and responsive to the presence of people [18] through providing intelligent monitoring and access control [19].

The services facilitated through automating homes can be broadly classified into three major categories: (i) Comfort; (ii) Healthcare; and (iii) Security [19]. Through the provision of comfort, smart homes provide ease in daily life. "Optimization for comfort in the environment is possible through the identification and automation procedure, learning user behavior, tracking user location, identifying the user, and automating tasks. Remote access and control enable users to remotely access, monitor and control their home environment" [19].

Smart homes also support providing healthcare for their inhabitants. As stated in [19], a significant portion of the world's population would be considered the elderly by the year 2050. To aid such elderly individuals in living independently, maintaining the safety of such individuals is of paramount importance, which can be achieved by detecting and preventing accidents, such as detecting an event such as a fall and calling for emergency services when such an accident occurs. Supporting aging in place (also termed as 'Gerontechnology' [20]) is another important aspect of helping senior adults with daily living activities, such as reminding them to take medications. One such research initiative at Georgia Tech devoted to the multidisciplinary exploration of emerging technologies and services based in the home is that of the Aware Home Research Initiative (AHRI). Supporting busy families in the generation where adults work full-time jobs is another scenario where smart home automation provides support. Often times, such families will have both elderly parents as well as young children that they must care for and thus the term coined for such populations is the "Sandwich Generation" [1].

Research into smart homes gained interest in academic domains, where the focus was on providing context awareness and smart decision making in automated environments. A number of initiatives have been focused on providing the following: (i) home automation, where providing comfort and convenience are of importance along with the goal of saving energy and resources; (ii) facilitating safety and security, where the goal is to provide monitoring of the inhabitants' well-being and aim for the provision of safety and security such as burglary recognition; and (iii) for entertainment, where the goal is to connect users and media with each other and to facilitate communication. A number of research works [1,2] were introduced with the goal of providing such services. One of the major applications was to provide assistance and support to the elderly to make their lives easier and provide support in daily living activities.

Toward the goal of automating homes to make human lives easier and more comfortable, automation efforts were initiated as early as the 19th century [21], where home appliances were designed to automate chores. Echo IV [22] was the first smart device that was used to manage shopping lists, control the home temperature and humidity, and provide tips for cooking. With technological advancements such as microcontrollers [23] and transistors [24], as well as reductions in the costs required for instrumenting homes with sensors, new application fields of computing—such as smart home automation—have become a possibility.

A few of the smart home automation research developed were the MavHome [8], GatorTech [2], and the Ambient Kitchen project [25]. MavHome [8] was set up with the goal for the environment to understand the resident's activities and respond accordingly to assist them in their daily living routines. GatorTech [2] (an extension of the Matilda Smart Home [26]) created an actual live-in environment aimed at assisting older individuals and those with special needs. Similarly, the Ambient Kitchen project [25] was aimed at developing a high-fidelity prototype by instrumenting objects used in the home to design applications that assist in everyday environments.

During the early 2000's, there was a push from the industry towards automating smart homes, and a number of appliances were developed that aided in this automation process. This was possible due to the advent of technology that aimed at making smart homes more accessible, engaging, and helpful to their end users [27–29]. With recent advancements in IoT-based technologies and cloud computing practices, there has been a renewed interest in research and development efforts. The focus has also now shifted toward private home automation. With the increasing number of smart home devices and vendors, sophisticated industry standards and platforms, and the number of assistants embedded in the environment, interest in developing smart home-based devices is evident.

Designing HAR systems has been made possible through the availability of data repositories such as the CASAS smart home datasets [30] and virtual smart home simulators

such as the Home I/O simulator [31] or the VirtualHome simulator [32–35]. Smart homes in [30] are aimed at identifying analysis procedures that aid in discovering user patterns. With the advances made in machine learning and data analysis techniques, it has become possible to analyze the data collected in such environments [36].

#### 2.2. Activity Recognition in Smart Homes

Human activity recognition (HAR) is aimed at identifying activities that are performed by a person as a result of analyzing the data collected from various sensing mechanisms [37]. Equipped with sensors and actuators, smart homes aim to not only detect movement within the home [38], but also to identify interaction with objects, devices, and appliances. All such devices, sensors, actuators, appliances, and objects in the home are interconnected through communication protocols [39].

To utilize the services provided through the smart home, it is essential for the home to understand and recognize the activities of the residents. Activity recognition systems in smart homes are typically designed and deployed to provide such recognition capabilities. Through the process of logging identified daily living activities, changes in regular routines can be indicative of health-related concerns that can be used to inform residents and their caregivers [6,7,30,40–42].

Sensing modalities that are used to detect movement in smart homes primarily belong to either vision-based or sensor-based systems. Vision-based approaches use perception-based sensing mechanisms, also known as optical sensors, to capture data for analysis [43,44]. These optical sensors aid in the collection of 2D images, 3D images, and video data. The use of depth-video-based HAR designed for elderly health care monitoring utilizes skeleton joint features to analyze behaviors and their changes therein [45], and a depth-silhouette-based human activity recognition system has been used for the real-time logging of performed activities [46]. However, this sensing modality comes with privacy concerns, where residents may not be willing to accept the information collected through the vision-based sensing mechanisms [47,48].

Sensor-based HAR systems comprise on-body or wearable sensing [49], sensors placed on objects [25,50], and ambient or in-the-environment sensors [51]. HAR methods comprise various sensors that are networked and connected with numerous devices to track the resident's activity or behavior. Since these modalities either record data through state changes (ambient sensors) or more continuous-valued data (wearables), the data recorded provides for a time series analysis problem. The sensor-based HAR system is less privacy intrusive and has thus been widely accepted to monitor the activities of daily living [52,53].

#### 2.2.1. Human Activity Recognition Systems

The data collected in smart homes is obtained through recording the values of sensors that are used to capture the way residents interact with their respective environments. Recordings that capture the actions and interactions of the resident are then used to analyze the activities of daily living of the resident [39,43,48]. Analysis procedures either make use of contextual knowledge such as the location, time and frequency of activities, spatiotemporal information, and interactions of the residents with objects [39,54]. Such procedures have been termed as 'Knowledge-Driven Approaches' and require knowledge from domain experts to design the system [39]. Ontology-based approaches are used to build these context-aware applications [55]. Complementary to this kind of analysis is 'Data-Driven Approaches', where statistical models are built using the data recorded in the smart home. These require large amounts of annotated data to learn probabilistic machine learning models, such as hidden Markov models, K-nearest neighbors, etc., that can recognize the activities of residents [39,56,57]. Although, both 'Knowledge-Driven Approaches' and 'Data-Driven Approaches' can be used to learn activity models, due to the individualistic environments of smart homes that require knowledge of the idiosyncratic behaviors or residents, either complete knowledge of the environment or large amounts of annotated data would be required for building activity models [54]. This either requires

domain expertise or large amounts of wait time (to collect annotated data) for an activity model to be available to residents.

Since ambient sensors record state changes, the time series data recorded results in an irregular sampling rate, which is unlike the data recorded from wearables or videos, where data collected from these modalities have a more consistent sampling rate. In order to identify the points in time where behavior changes occur in the time series analysis problem, algorithms belonging to the family of change point detection (CPD) methods are used [58]. A number of CPD-based methods have been applied to segment time series data into activities of interest-also known as 'Activity Segmentation'. Algorithms are developed to identify the activity segments automatically and then the activity in those segments is identified through a recognition procedure [59]. Real-time activity recognition can be essential when identifying activities that require immediate care such as fall detection [60] or to automatically log behaviors for health monitoring. Some CPD methods that have been developed are suitable to provide real-time activity recognition systems, whereas others recognize activity segments after a delay from the time of occurrence. Different windowing procedures have been compared in [61] to estimate activity boundaries. Identifying explicit windows (EW) corresponds to the 'pre-segment' technique [62,63] where a given window contains all the sensor events corresponding to a given activity. Annotations are required to identify the beginning and end points of the activities of interest. When no such annotations are available, change points corresponding to when changes occur in activities are identified using the statistical and probability-based measures described in [64,65]. In order to perform activity recognition, it becomes essential to identify all segments corresponding to the activity instances, since the whole segment is used to predict a given activity. Thus, it is not straightforward to use the 'pre-segment' technique in real-time analysis. For real-time analysis, the time window (TW)—where windows span over a specific duration of time [66–70]—the sensor event window (SEW)—where windows span over a specific number of sensor event triggers [10,71]—has been used. Although, the time-based windowing technique is favorable for regularly or continuously sampled data over time when commonly used with wearable sensors, the data collected through ambient sensors can be sampled at regular intervals through a forward-filling procedure. In both cases, however, identifying the ideal window length of the sliding window can be challenging and requires domain-specific knowledge to estimate the window length [72]. Heuristic approaches such as rules, thresholds, and dissimilarity measures between window embeddings have been used to identify window lengths dynamically [67,73,74].

A number of traditional approaches have been proposed for sensor-based HAR systems in smart homes. Classification approaches such as random forests, naive Bayes, decision trees, and conditional random fields have been explored [75]. Some of the traditional feature representations look into the number of sensor firings, the time spent at a given location, and the time spent moving between locations [76]. Another work made use of the features that model contextual information by considering the mutual information between sensor events and decay in sensor event triggers [10]. Spatiotemporal features explored with multilayer perceptron, hidden Markov models, decision trees, etc. have been shown to outperform traditionally used feature representations [77]. The use of SVMs and variants of incremental SVMs have been employed to improve the performance of the HAR systems [78]. In [79], the authors made use of sensor data contribution significance analysis and spatial distance matrices to identify (a) the relevant sensors that are most informative of the activities in the home; and (b) the noise caused by various factors such as pets and visitors that are not relevant to the activity being monitored. Eliminating the noise and sensors that do not contribute to activity recognition has been shown to improve performance scores. Inspired from the cluster then classify paradigm, in [4], the location was used as the contextual information to cluster data points that were assumed to belong to the same activity. Some of these conventional procedures require handcrafted feature extraction methods to learn relevant information from the sensor event triggers [39] or require large amounts of labeled data to build knowledge [79]. Recent work [80] looked

at different metrics that could be employed to obtain features that are of relevance for classification. This still requires domain knowledge to initialize the entire list of features that are used as a starting point. A subset of these features is then deemed to be important for the classification procedure. Similar drawbacks exist in activity recognition systems that make use of ontologies where detailed information about activity interactions is required for generating feature vectors [55].

With the advancements in deep learning techniques, it has become possible to model high-level abstractions from complex data [81]. Deep learning techniques can be used to learn good high-level feature representations from raw signals by utilizing unsupervised learning procedures, without requiring any manual engineering efforts. Such procedures can also be used for end-to-end learning systems where the models perform classification using the automatically extracted features [82,83]. Convolutional neural networks (CNNs) have been used to capture local dependencies in time series data. Since they are invariant to scale and translation, they are able to capture the local temporal dynamics between data points [84]. Deep convolutional neural networks have been used to analyze sequences of binary sensors that are converted into gray-scale images [85,86]. A 1D-CNN structure developed on raw data sequences was used in [87] to extract high-level features, to learn mappings between sensor event triggers, and for activities in an end-to-end learning procedure. To capture the temporal information in sequential data modeling procedures, RNNs are employed. LSTM networks have become popular due to their ability to capture long-term dependencies [84]. Thus, different variants of LSTMs such as bidirectional and cascading LSTMs have been used to automatically learn temporal information from raw sensor sequences and achieve reasonable performance outcomes [88]. Using frequency encodings to capture the co-occurrence of sensor event triggers in a sequence has been used to model and learn good feature presentations that are then passed through a fully convolutional network [89]. Other embedding procedures inspired by language-based modeling techniques such as ELMO have been used to learn better feature representations [90]. Activity2Vec [11] is a sequence-to-sequence model that is aimed at learning feature representations for the activities of daily living and activities that occur rarely such as fall detection. Although most of these deep-learning-based approaches achieved state-of-the-art performances, they have a major drawback—they make use of 'pre-segmented' activity instances as inputs to the classification procedures. The use of the pre-segmentation technique to identify explicit windows (EWs) is not ideal, since systems that require knowledge of the start and end points of activities cannot be used in the deployment scenario.

More recently, graph attention networks have been used as a classification approach due to their increasing prevalence in various fields [91]. By utilizing a graph structure to model human activities, a sequence is converted into a sequence graph, where each graph node is connected to one or more nodes [92]. Information is aggregated through messages received from neighboring nodes in the graph and is then transmitted. An attention mechanism is used to produce node representations that allow for distinguishing the contribution of each node on the target node. The work in [93] used sensor events as nodes, and the edges between the nodes were represented by the intensity of the connections, thus forming a graph. The graph attention network proposed in this work makes use of the generated features to learn a location-oriented and time-oriented graph attention network that is further passed through convolutional layers and subsequently through a fully connected layer to identify the activity performed.

#### 2.2.2. Active Learning

The successful implementation of a human activity recognition system requires the sensor data gathered in the smart home to be accurately mapped to human behavior. Developing supervised training approaches requires large amounts of annotated data to achieve reasonable performance scores [94]. An annotation procedure is employed to provide labels to activities that are identified over the duration of data collection in the

homes. A significant challenge in obtaining a large number of such annotations is that it is cumbersome and requires time and effort [95]. Annotator expertise determines the quality of the annotations obtained, which can be expensive based on the expertise of the annotator. Annotations can be obtained either in situ by residents while the activity is being performed or retrospectively, wherein either the resident or an external annotator provides the annotations by observing the data collected. Obtaining such labels could cause disruptions by interrupting users frequently or requiring too much effort and time when providing such labels retrospectively. As the recognition system evolves, new sets of activities may be picked up by residents, such as picking up dancing as a hobby, or the initially recognized set of activities can be performed differently, such as relaxing as an activity that could change in definition from reading a book to taking a nap. Thus capturing activities of interest and variants of these, without burdening the resident, requires the use of a semisupervised machine learning paradigm—active learning.

In contrast to supervised approaches that require annotations for all the collected data, active learning (AL) queries only those data points that are deemed informative to the learning procedure. As such, two different components for identifying such data points exist—(i) sampling strategies and (ii) query strategies. Sampling strategies determine the procedure by which the data points to be picked for the annotation procedure are identified. Two popular sampling strategies corresponding to pool-based active learning and stream-based active learning are used [96,97]. In the pool-based active learning procedure, which is suitable for an offline learning procedure, a large unlabeled dataset and a small labeled dataset are utilized. The algorithm then selects the best data points that are most informative from the unlabeled dataset to be queried. Similarly, in the stream-based active learning procedure, which is suitable for an online learning procedure, streams of data are analyzed for obtaining an annotation based on the informativeness of the incoming data point. Query strategies determine the specific data points to be identified based on the informativeness of the data points for the learning procedure. Two popular query strategies correspond to the uncertainty sampling and the margin-based uncertainty sampling, where, for the first strategy, the model queries for those data points that it is least confident about [98]. In the margin-based uncertainty sampling, the model queries for those samples for which the margin between the two most probable class prediction probabilities is small. Other query strategies such as query by committee [98] and logistic margin sampling [99] have been explored. To determine the number of data points to be queried for a budget, this process is predetermined. The size of the budget determines the performance gains that can be achieved and how close the AL procedure reaches the performance measures corresponding to supervised methods. Various strategies for how this determined budget is spent in obtaining the annotations are discussed in [100]. Active learning procedures have been utilized in smart home settings to learn about resident activities and behaviors, wherein they use the sampling and query procedures described in [101]. Markov decision processes that make use of gestures and vocal expressions to obtain feedback from the resident, which are incorporated into the modeling procedure. Positive and negative responses to their interactions with the system are provided as feedback to the modeling procedure [102]. Other works have used contextual information such as location to annotations corresponding to informative clusters [4,103].

#### 3. Need for Bespoke HAR Systems in Smart Homes

By critically analyzing the state-of-the-art approaches developed for activity recognition in smart homes, we discuss the drawbacks of said systems. Although the recognition scores obtained through systems in [88–90] wer high, such systems are not ideal for providing an HAR solution, since they require event-based sequences, where knowledge of the start and end times corresponding to a given activity sequence is needed. These systems, when deployed in smart home scenarios, require the resident to provide information regarding this, which is often cumbersome. The resident will have to either be continuously engaged with the system to provide these activity boundaries or provide this information retrospectively, which may be subject to recall bias [104]. Thus, techniques developed for the 'pre-segmented' activity instances are not readily deployable in real-world scenarios. Although the recent literature makes use of robust sequential models such as LSTMs and graph attention networks, the reliance on the identification of 'pre-segments' limits the deployable capabilities of the systems [88,90]. Some such systems are listed in Table 1, which correspond to the category of 'Event-based analysis'. As observed from the literature, most other procedures to identify the appropriate data sequence length for analysis require some knowledge of the activities performed by the resident [10,66–71]. Other approaches require the procedure to be context-aware (e.g., location) through incorporating knowledge of the environment [93]. This requires domain expertise to build such knowledge for every environment analyzed. Examples of such systems are listed in Table 1, which correspond to the category of 'Requires domain knowledge'.

The methods developed in [11] have been tested on some of the collected CASAS datasets but not on others; thus, it is not clear if the given method generalizes across different smart home settings. The works in [8,9] developed methods for specific homes [5] by making use of object and appliance interactions (developed through ontologies) that do not generalize to other smart homes, which consist of different objects. Differences in the sensor positions in different home settings lead to changes in sequence patterns, which were utilized to model activities in [8]. This is a typical issue with some of the related work. Similarly, the works in [8,9] developed systems that are specific to the given smart home of PlaceLab [5] and MavHome [8], respectively. Contextual information, which is specific to individualized settings, has been utilized in developing ontologies [105] and in building recognition procedures that require context-sensitive embeddings (the examples in Table 1 corresponding to the category of 'Requires domain knowledge') [106–108].

The HAR system developed in [30] used specific information in the home for feature engineering such as the number of sensor events triggered, recent sensor events, and contextual information such as the day, week, or the hour of the day, as well as information related to the resident's activity patterns such as the elapsed time for each sensor event. These details change across different smart home settings, and, thus, the HAR system would require such details for every individual environment. As seen in [11], the developed sequence-to-sequence model was developed for a given smart home of the CASAS dataset— HH101 [30]. Similarly, in [109], the activity boundaries were identified using statistical measures such as the Pearson coefficient, but it is unclear if it is applicable to smart homes that are more complex. Such handcrafted features require large amounts of annotated data, which are required to be collected in the smart home [94] to learn representations that are suitable for a given home. Some such systems have been listed in Table 1, which correspond to the category of 'Requires annotated data'. These would lead to building robust supervised learning procedures. As discussed, such analysis procedures require large amounts of resources in terms of annotations, time, and effort [95], and the developed system becomes available only after long periods of wait time (required to collect data in the given home), which might not be acceptable to the resident [110].

To summarize, the design of HAR systems for the application scenario of smart homes is challenging and not straightforward. Existing HAR systems cannot be used "as-is" when deploying them to a new smart home setting, and each deployment requires domainspecific modifications in order to be usable and effective. Detailed knowledge of the target environment such as activities performed in the home, extensive ground truth annotations for such activities, and precise activity boundaries are required for building most state-ofthe-art HAR techniques in smart home scenarios. Such knowledge is obtained in the form of detailed extensive ground truth annotation and frame-precise activity segmentation. Thus, developing a supervised HAR system for specific settings requires large amounts of annotated data. Obtaining such annotations from residents is burdensome and requires considerable effort [98,99]. We identify and list the drawbacks of the aforementioned HAR systems as follows: (i) most systems are not developed for providing automated recognition; (ii) idiosyncratic environments require the development of HAR systems that are specific to the settings of smart homes; and (iii) the approaches developed for a given smart home environment do not necessarily generalize to other smart homes. Hence, "off-the-shelf" systems that will be universally useful without tuning them for individual smart homes do not exist. To address the mentioned concerns, we motivate the development of bespoke HAR systems for such individualized settings that learn from scratch in a given environment, with minimal involvement from the residents. Such a system is developed through a data-driven procedure and does not make use of specific (contextual) information in a given setting.

Reference	Category of HAR System	Activity Recognition System	Datasets(s)
A sequential deep learning application for recognizing human activities [88]	Event-based analysis; requires annotated data	Variations in sequential modeling techniques (LSTMs) were used on event-based data instances	CASAS datasets (Milan, Cairo, Kyoto7, Kyoto8, and Kyoto11) [30]
Fully convolutional network bootstrapped by word encoding and embedding for activity recognition in smart homes [89]	Event-based analysis; requires annotated data	A word2vec encoding was applied to sensor event-based windows, which were then passed through a fully convolutional network for classification	CASAS datasets (Aruba and Milan) [30]
Using the language model to bootstrap human activity recognition that utilized ambient sensors Based in smart homes [90]	Event-based analysis; requires annotated data	Different embedding techniques were used to obtain the learned features followed by a sequential modeling procedure (LSTM) on the event-based data instances	CASAS datasets (Aruba, Milan, and Cairo) [30]
Activity2vec: Learning adl embeddings from sensor data with a sequence-to-sequence model [11]	Requires annotated data	A sequence-to-sequence model was used to generate features followed by a random forest model for classification	CASAS dataset (HH101) [30]
Enhancing activity recognition using CPD-based activity segmentation [71]	Requires annotated data	A heuristic function followed by a dissimilarity-based approach were used to identify change points. Handcrafted features were extracted. A random-forest-based modeling procedure was employed to perform the classification	CASAS dataset (Apt 101-130) [30]
Using ontologies in case-based activity [9]	Requires domain knowledge	Rules case-based reasoning, where the information gained for each feature for a given activity is used to provide the predictions	PlaceLab [5]
Activity recognition on streaming sensor data [10]	Requires annotated data	Handcrafted features generated over sliding windows. An SVM-based classification model was employed.	Smart home TestBeds—B1, B2, and B3 [62]
MavHome: An agent-based smart home [8]	Requires domain knowledge	Episode discovery algorithm that identified significant episodes in the sequence of patterns mined	MavHome [8]

Table 1. Activity recognition systems developed for smart homes.

We designed bespoke HAR systems for these individual environments by defining the various components that they should encompass, where the goal was to develop a fully functional HAR system. Such a system was derived through the assumption of a cold start scenario, where (initially) unlabeled sensor data were passively observed in the smart home. Thus, the design of the system did not require any contextual information or knowledge specific to a given home setting. An initial functional HAR system becomes quickly available to the resident without requiring long periods of wait time. As the system observed data in the home, it built and incorporated knowledge hierarchically. Minimal supervision provided by the residents was used to develop these HAR systems and their updates therein in order to reduce the effort and burden on the resident. To motivate the technical aspects of the said lifespan, we detailed an application scenario that observed the different scenarios that are typical in a smart home. Technical aspects were designed through various components to address these different scenarios in the home.

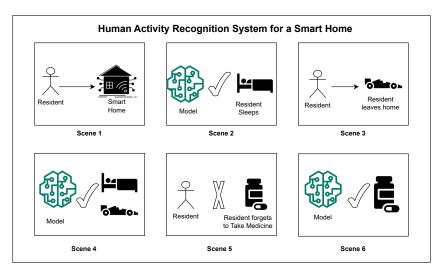
#### 4. The Lifespan of Human Activity Recognition Systems for Smart Homes

To motivate the need for defining the lifespan of HAR systems for smart homes, we provide an illustration of a typical application scenario for activity recognition in these settings, which is shown in Figure 1. It details the scenario in a smart home and the need for a functional HAR system that addresses these. In the beginning (Scene 1), a resident moves into a new smart home, with installed motion and door sensors to track movement patterns that are essential for activity monitoring. Given that no actual HAR system exists at this time ("cold-start"), because "off-the-shelf" HAR systems will not work in highly individualized and situated environments such as a private home.

As such, an initial bootstrapping procedure was deployed that first collected raw sensor readings while the resident conducted their regular activities (Scene 2). The initial HAR system aimed at detecting the most prominent or frequently occurring activities in the home (e.g., 'sleeping', as shown in the figure). However, the resident also performs less-frequent activities such as 'leaving home' (Scene 3), which do not get modeled yet. 'Leaving Home' can be considered as a less-frequent activity in the case of an elderly living environment, where the activity is not one of the prominent activities. The HAR system would now be continuously updated to capture both the more and less prominent activities in the smart home (Scene 4).

As the resident continues to live in the smart home, the model additionally focuses on assessing activity routines. Most developed HAR systems are utilized to identify activities that help with logging behaviors performed by residents. However, there is more to understanding these behaviors than just recognizing individual instances. Thus, the assessment of routines helps in the analysis of (any) changes that occur in a resident's life at a level higher than just analyzing individual activity patterns, for example, when the resident is forgetting to "take medicine" (Scene 5). Identifying *regular* routines can inform any deviations that will be used to inform the resident or caregivers (Scene 6).

In order to detail the technical concepts that aim at capturing the scenarios illustrated in Figure 1, we provide detailed technical specifications of the various components of this overarching technical concept in what follows. The development of individualized HAR systems for smart homes requires minimal human intervention, and the focus is on the rapid availability of essential functionality to the resident. As part of the technical details, we also ensure that the HAR system continuously improves and adapts to changes in the home system over time.



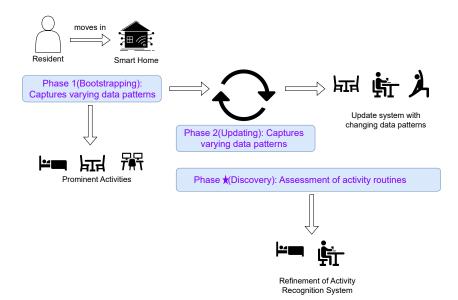
**Figure 1.** The lifespan of a human activity recognition system for smart homes in terms of requirements evolving over time and the needed responses of such systems to a life that is ever-changing. See text for description.

#### 4.1. Phase 1: Bootstrapping

Figure 2 illustrates the various components of the HAR system as we defined it to provide the basis for the design of the lifespan of an HAR system for smart homes. At the beginning of the HAR system (Phase 1 in Figure 2) there is a bootstrapping procedure, which aims at getting the first working version of the HAR system in place, thereby focusing on minimal user involvement and rapid deployment. As such, it collects unlabeled sensor readings as soon as the resident moves into the home, thereby targeting the most frequently occurring and most prominent activities in a specific home. Identifying such prominent activities, albeit certainly not all of them, has more practical value than designing a sophisticated system that would require large training data and a longer wait time for the resident, thus, overall requiring more resident involvement. Residents are expected to be involved at a minimal level primarily to confirm the identified sequences of movement patterns, which correspond to the aforementioned most frequent and prominent activities, which will be modeled by the HAR system. The resident is asked to provide annotations in an active learning scenario [98]. Annotations corresponding to only the most prominent movement patterns are requested from the resident. These patterns can be identified through a set of designed filtering procedures that are implemented during the construction of the initial analysis pipeline.

The design of the state-of-the-art analysis pipelines faces technical challenges due to the sparse set of annotations obtained from the smart homes provided by residents. Additionally, there is a lack of availability of high-quality annotations, since these annotations come from the resident, who provides them retrospectively. Thus, the goal of this stage, as described in previous work, is "to provide a system that "jump-starts" the activity recognition pipeline for the smart home" [111].

At the end of this initial phase, the HAR system should be able to recognize the most prominent activities with satisfactory accuracy for a subset of activity classes observed in the smart home. Also, there is typically more happening in a home than those limited sets of activities, which corresponds to those activities that are not captured by the initial system and variants of the initial set of captured activity classes that are missed [112,113]. However, with such an automatically derived ("bootstrapped"), functional HAR system, the smart home can already fulfill a range of routine operations. For example, it would be able to monitor and track sleeping patterns or the work-life balance of its residents. This phase corresponds to Scenes 1 and 2 in Figure 1. Previous work has presented prototypes that cover Phase 1 of the lifespan of a smart home's HAR system as we have defined in this paper (e.g., [111]).



**Figure 2.** Illustration of the three components of HAR system for smart homes consisting of *bootstrapping*, *updating*, *and assessment*. Phase 1: An initial, fully functional HAR system is bootstrapped from scratch in a data-driven procedure. Phase 2: Maintenance of the sensor-based HAR system. Phase **\***: Aims at the assessment of activity routines.

#### 4.2. Phase 2: Updating

In the next phase of the HAR system (Phase 2 in Figure 2), an initial HAR system (from Phase 1) will be updated and extended in an incremental fashion—as more data is captured in the smart home, this likely increases the variability of the activities themselves, as well as their individual appearances. At the technical level, updating and extending an initial HAR system is targeted by the second phase of the lifespan, and it requires a *class-* as well as *style–incremental* approach. Through the former, the HAR system learns to identify new activities, and through the latter, the HAR system learns to identify variants of already identified activities, and, as such, refines them with regard to, for example, more accurate segmentation and/ or covering specific variants of individualized activities. An example corresponding to a new class corresponds to an activity that the HAR system cannot recognize at the end of Phase 1, for example, 'leaving home', which occurs infrequently as shown in Figure 1. As the resident continues to stay at home, they may adopt different ways to relax, which, for example, may change from reading a book to taking a nap. This corresponds to identifying changes in activities already identified through Phase 1. This phase corresponds to Scenes 3 and 4 in Figure 1.

Building on the initial HAR system (from Phase 1) shifts the manual efforts, which are still required for the occasional annotation of significant, new movement patterns, which are away from the activity patterns that are already known (from Phase 1) and thus keeps the burden on the resident at a reasonable level. A suitable approach to accommodating novel activity class instances and potential concept drift in already modeled activities is through utilizing continual learning (CL) models [114–120] when such instances are observed. Continual learning as a concept corresponds to situations when a model—any model not restricted to the recent surge in CL research in the deep learning community, but rather which adopts the general concept of continuously updating existing HAR models to changed circumstances—learns sequentially from data or tasks without forgetting knowledge obtained from preceding tasks. Throughout Phase 2—which, essentially spans from the end of Phase 1, i.e., when the first HAR system is available—the HAR system is continuously updated and extended with the goal of keeping up with the ever-changing circumstances of life, which, however, does not forget the already learned concepts (activities and their styles).

#### 4.3. Phase **\***: Routine Discovery

Phases 1 and 2 aim at developing a functional HAR system that recognizes regular activities in the home. To complete the lifespan of the HAR, we next detail the final phase to consist in assessing activity routines. Thus, Phase \* (occurring in parallel to Phase 2) is aimed at using the developed activity recognition system in assessing these routines. Monitoring these routines helps in establishing regularity (or not) of the behavior patterns of the resident. Thus, this phase is aimed at identifying these different routines in the home.

The previously described phases detailed the recognition system that can be utilized for monitoring the activities of the resident. In addition to providing such monitoring, assessing the activity routines of the residents is beneficial in the long-term goal of providing assistance in smart homes, for example, in providing care to the resident. Activity routines are patterns of behavior constituting sequences of activities [121]. Monitoring the regularity of such routine patterns can inform of the resident's health and be utilized in, for example, observing circadian rhythms [122]. Aberrations from *regular* routines are indicative of deviant behaviors which may be a cause for concern. Hence, changes in behavior patterns can be informed to caregivers, especially in the case of the elderly. Thus, this phase—Phase  $\star$  of the lifespan of the HAR (Figure 2)—is aimed at assessing activity routines from the recognition models identified previously. Refinement of the identified activity routines can be further used in refining the designed recognition system through the previous phases. This phase corresponds to Scenes 5 and 6 in Figure 1.

# 5. Scalability of the Proposed Conceptual System

Developing an activity recognition system for the settings of smart homes is not straightforward. Several challenges to developing HAR systems for providing activity monitoring in smart homes exist. Most of the developed state-of-the-art procedures are not aimed at providing automated recognition, and they often-times make use of 'presegmented' activity instanced for analysis. Smart homes are individualized environments with the idiosyncratic behaviors of residents. Thus, HAR systems are required to be developed specifically for the settings of a given home. This leads to a lack of an "off-theshelf" HAR system that can be used "as-is" in a different smart home setting. The lack of such universally usable systems, which require at least some modifications in terms of tuning to different smart homes, motivates the need to develop "bespoke" HAR systems for such environments.

To aid in developing the "bespoke" HAR system, in this work, we propose a conceptual system. This conceptual system details the lifespan of a HAR system for smart homes. The designed lifespan provides for the different phases of building the fully-functional HAR system, which require minimal resident supervision. Such a designed lifespan provides for the different phases of building the system, which is motivated by a typical application scenario in the home setting. Through the use of data-driven procedures, the technical solutions developed for each of these phases makes it possible to derive and continuously improve the HAR system. We discuss the scalability of the proposed conceptual system with regard to (i) the performance in new and unseen environments and (ii) the performance in multi-resident smart homes.

Developing machine learning models on a generic population, where data is collected from various data sources, and fine-tuning these models to specific populations or individuals is of interest [123–125]. At its core, this idea of personalization is aimed at adapting to specific persons to provide for better activity monitoring. In order to achieve this goal, in a fine-tuning procedure, a small portion of data from the specific population or individual is used to fine-tune a model built using generic data. The expectation of building such a personalized model is to adapt to the given specific population. The phases corresponding to Phase 1 and Phase 2 of the conceptual system—the lifespan of the HAR—mirror this phenomenon. In Phase 1, the procedure starts in a cold-start scenario, to "jump-start" the activity recognition system. Since such a system learns from passively observing data in a given home and builds knowledge hierarchically, thus, the procedure is designed to learn the activity patterns of any given setting. Through the update and maintenance procedure, in Phase 2, the initial HAR system is updated through a continual learning procedure. This HAR system learns a refined model that is personalized to a given home setting and the idiosyncrasies of the resident occupying the home. Thus, the conceptual system is a generic approach and can be used "as-is" in any new and unseen smart home. The HAR system developed through the phases of the conceptual system is personalized to the individual home and to the behaviors of its residents.

The design of the conceptual system, with its various phases, does not differentiate between a single-resident household and a multi-resident household. Differentiating between the activities performed by different inhabitants in a multi-resident household is possible using a person identifier, which helps distinguish between the different residents. Various studies have looked into such identification techniques such as, but not limited to, radio frequency signals, WiFi, and using videos. Fusing data from different modalities to analyze activities pertaining to a given individual has also been developed [126]. In the absence of a way to identify different residents, the HAR system learns of the activity patterns in the home without making a distinction between who performed a given activity. However, this is not a limitation of the conceptual system, and these methods can be incorporated into the HAR system obtained from the implementation of the conceptual system. Such integration will result in activity monitoring that learns activity patterns for the different residents in the home.

# 6. Conclusions

A number of activity recognition systems have been developed for smart homes. Analyzing these systems developed across the various smart home datasets [5,8,30] provides insights into the complex and challenging application scenario, which requires such systems to be developed for specific environments [4,8,11].

The proposed lifespan for the HAR is aimed at capturing three important and essential technical challenges that building recognition systems in smart homes face: (i) requiring large amounts of in situ data to build (supervised) activity recognition systems, (ii) requiring a large number of annotations from residents or experts over the entire duration of data collected, to provide labels, and; (iii) requiring a substantial amount of wait time (by the resident) before the system becomes available for use in the home.

In order to primarily address these challenges as observed with the current state-ofthe-art procedures, in this "perspectives" track we introduced the lifespan of the HAR systems for smart homes. The lifespan of the HAR addresses the aforementioned challenges by providing a fully functional HAR system that becomes available quickly to its residents after observing data for an initial period of time and obtaining annotations for the most prominent movement patterns during Phase 1. The update procedure detailed in Phase 2 serves to identify movement patterns that are not observed initially yet require minimal supervision from the resident. Since updates to the system are available after the update procedure, the "new" HAR system becomes available as soon as it is ready for resident usage. Lastly, this developed recognition procedure serves as a tool to understand user activity patterns and regularities in *daily* routines.

Author Contributions: Conceptualization, S.K.H. and T.P.; Data curation, S.K.H. and T.P.; Methodology, S.K.H.; Software, S.K.H.; Validation, S.K.H.; Formal analysis, S.K.H.; Investigation, S.K.H.; Writing—original draft preparation, S.K.H.; Writing—review and editing, S.K.H. and T.P.; Visualization, S.K.H.; Supervision, T.P.; Project administration, T.P.; Funding acquisition, T.P. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was partially supported by KDDI Research. We would also like to thank the AI-CARING team for their input on the project. We gratefully acknowledge Nvidia Corporation for supporting any computation resources that were required for the project.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analysis, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

# References

- 1. Kientz, J.A.; Patel, S.N.; Jones, B.; Price, E.; Mynatt, E.D.; Abowd, G.D. The georgia tech aware home. In Proceedings of the CHI'08 Extended Abstracts on Human Factors in Computing Systems, Florence, Italy, 5–10 April 2008; pp. 3675–3680.
- 2. Helal, S.; Mann, W.; El-Zabadani, H.; King, J.; Kaddoura, Y.; Jansen, E. The gator tech smart house: A programmable pervasive space. *Computer* **2005**, *38*, 50–60. [CrossRef]
- Bagaveyev, S.; Cook, D.J. Designing and evaluating active learning methods for activity recognition. In Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication, Seattle, WA, USA, 13–17 September 2014; pp. 469–478.
- Hoque, E.; Stankovic, J. AALO: Activity recognition in smart homes using Active Learning in the presence of Overlapped activities. In Proceedings of the 2012 6th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshops, IEEE, San Diego, CA, USA, 21–24 May 2012; pp. 139–146.
- Intille, S.S.; Larson, K.; Beaudin, J.; Tapia, E.M.; Kaushik, P.; Nawyn, J.; McLeish, T.J. The PlaceLab: A live-in laboratory for pervasive computing research (video). In Proceedings of the PERVASIVE 2005 Video Program, Online, May 2005.
- Hooper, C.J.; Preston, A.; Balaam, M.; Seedhouse, P.; Jackson, D.; Pham, C.; Ladha, C.; Ladha, K.; Plötz, T.; Olivier, P. The french kitchen: Task-based learning in an instrumented kitchen. In Proceedings of the 2012 ACM Conference on Ubiquitous Computing, Pittsburgh, PA, USA, 21–24 May 2012; pp. 193–202.
- Alemdar, H.; Ertan, H.; Incel, O.D.; Ersoy, C. ARAS human activity datasets in multiple homes with multiple residents. In Proceedings of the 2013 7th International Conference on Pervasive Computing Technologies for Healthcare and Workshops, IEEE, Venice, Italy, 5–8 May 2013; pp. 232–235.
- Cook, D.J.; Youngblood, M.; Heierman, E.O.; Gopalratnam, K.; Rao, S.; Litvin, A.; Khawaja, F. MavHome: An agent-based smart home. In Proceedings of the First IEEE International Conference on Pervasive Computing and Communications (PerCom 2003), IEEE, Fort Worth, TX, USA, 26 March 2003; pp. 521–524.
- Knox, S.; Coyle, L.; Dobson, S. Using Ontologies in Case-Based Activity Recognition. In Proceedings of the FLAIRS Conference, Daytona Beach, FL, USA, 19–21 May 2010; pp. 1–6.
- 10. Krishnan, N.C.; Cook, D.J. Activity recognition on streaming sensor data. *Pervasive Mob. Comput.* **2014**, *10*, 138–154. [CrossRef] [PubMed]
- 11. Ghods, A.; Cook, D.J. Activity2vec: Learning adl embeddings from sensor data with a sequence-to-sequence model. *arXiv* 2019, arXiv:1907.05597.
- 12. Skubic, M.; Alexander, G.; Popescu, M.; Rantz, M.; Keller, J. A smart home application to eldercare: Current status and lessons learned. *Technol. Health Care* 2009, 17, 183–201. [CrossRef] [PubMed]
- 13. Ding, D.; Cooper, R.A.; Pasquina, P.F.; Fici-Pasquina, L. Sensor technology for smart homes. *Maturitas* **2011**, *69*, 131–136. [CrossRef] [PubMed]
- Tragos, E.Z.; Foti, M.; Surligas, M.; Lambropoulos, G.; Pournaras, S.; Papadakis, S.; Angelakis, V. An IoT based intelligent building management system for ambient assisted living. In Proceedings of the 2015 IEEE International Conference on Communication Workshop (ICCW), IEEE, London, UK, 8–12 June 2015; pp. 246–252.
- 15. Weiser, M. The computer for the 21st century. ACM SIGMOBILE Mob. Comput. Commun. Rev. 1999, 3, 3–11. [CrossRef]
- Kidd, C.D.; Orr, R.; Abowd, G.D.; Atkeson, C.G.; Essa, I.A.; MacIntyre, B.; Mynatt, E.; Starner, T.E.; Newstetter, W. The aware home: A living laboratory for ubiquitous computing research. In Proceedings of the Cooperative Buildings, Integrating Information, Organizations, and Architecture: Second International Workshop, CoBuild'99, Pittsburgh, PA, USA, 1–2 October 1999; Proceedings 2; Springer: Pittsburgh, PA, USA, 1999; pp. 191–198.
- 17. Morita, P.P.; Sahu, K.S.; Oetomo, A. Health Monitoring Using Smart Home Technologies: Scoping Review. *JMIR MHealth UHealth* **2023**, *11*, e37347. [CrossRef]
- 18. Chan, M.; Estève, D.; Escriba, C.; Campo, E. A review of smart homes—Present state and future challenges. *Comput. Methods Programs Biomed.* **2008**, *91*, 55–81. [CrossRef]
- 19. Alam, M.R.; Reaz, M.B.I.; Ali, M.A.M. A review of smart homes—Past, present, and future. *IEEE Trans. Syst. Man, Cybern. Part C (Appl. Rev.)* 2012, 42, 1190–1203. [CrossRef]
- 20. Bouma, H.; Graafmans, J.A. Gerontechnology; IOS Press: Amsterdam, The Netherlands, 1992; Volume 3.
- 21. Chatting, D. Automated Indifference. Interactions 2023, 30, 22–26. [CrossRef]
- Ray, A.K.; Bagwari, A. Study of smart home communication protocol's and security & privacy aspects. In Proceedings of the 2017 7th International Conference on Communication Systems and Network Technologies (CSNT), IEEE, Nagpur, India, 11–13 November 2017; pp. 240–245.
- 23. Spasov, P. Microcontroller Technology: The 68HC11; Prentice-Hall, Inc.: Hoboken, NJ, USA, 1993.
- 24. Riordan, M.; Hoddeson, L.; Herring, C. The invention of the transistor. Rev. Mod. Phys. 1999, 71, S336. [CrossRef]

- Olivier, P.; Xu, G.; Monk, A.; Hoey, J. Ambient kitchen: Designing situated services using a high fidelity prototyping environment. In Proceedings of the 2nd International Conference on Pervasive Technologies Related to Assistive Environments, Corfu, Greece, 9–13 June 2009; pp. 1–7.
- Schneps-Schneppe, M.; Maximenko, A.; Namiot, D.; Malov, D. Wired Smart Home: Energy metering, security, and emergency issues. In Proceedings of the 2012 IV International Congress on Ultra Modern Telecommunications and Control Systems, IEEE, St. Petersburg, Russia, 3–5 October 2012; pp. 405–410.
- 27. Smart Home—United States: Statista Market Forecast, 2021. Available online: www.statista.com/outlook/dmo/smart-home/ united-states (accessed on 29 April 2022).
- 28. Smart Home Market Size, Share, Ecosystems, and Dynamics (Driver, Restrain, Opportunities) 2030. Available online: www. verifiedmarketresearch.com/product/global-smart-home-market-size-and-forecast (accessed on 29 April 2022).
- 29. Global Smart Home Market Research Report (2021 to 2026)—By Product, Technologies, Service, Deployment and Region. 2021. Available online: www.researchandmarkets.com (accessed on 29 April 2022).
- 30. 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]
- Philippot, A.; Riera, B.; Koza, M.; Pichard, R.; Saddem, R.; Gellot, F.; Annebicque, D.; Emprin, F. HOME I/O and FACTORY I/O: 2 Pieces of innovative PO simulation software for automation education. In Proceedings of the 2017 27th EAEEIE Annual Conference (EAEEIE), IEEE, Grenoble, France, 7–9 June 2017; pp. 1–6.
- Puig, X.; Ra, K.; Boben, M.; Li, J.; Wang, T.; Fidler, S.; Torralba, A. Virtualhome: Simulating household activities via programs. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 8494–8502.
- Liao, Y.H.; Puig, X.; Boben, M.; Torralba, A.; Fidler, S. Synthesizing environment-aware activities via activity sketches. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 6291–6299.
- 34. Puig, X.; Shu, T.; Li, S.; Wang, Z.; Liao, Y.H.; Tenenbaum, J.B.; Fidler, S.; Torralba, A. Watch-and-help: A challenge for social perception and human-ai collaboration. *arXiv* **2020**, arXiv:2010.09890.
- Li, S.; Puig, X.; Du, Y.; Wang, C.; Akyurek, E.; Torralba, A.; Andreas, J.; Mordatch, I. Pre-trained language models for interactive decision-making. *arXiv* 2022, arXiv:2202.01771.
- 36. Ramasamy Ramamurthy, S.; Roy, N. Recent trends in machine learning for human activity recognition—A survey. Wiley Interdiscip. Rev. Data Min. Knowl. Discov. 2018, 8, e1254. [CrossRef]
- González, S.; Sedano, J.; Villar, J.R.; Corchado, E.; Herrero, Á.; Baruque, B. Features and models for human activity recognition. *Neurocomputing* 2015, 167, 52–60. [CrossRef]
- Chan, M.; Campo, E.; Estève, D.; Fourniols, J.Y. Smart homes—current features and future perspectives. *Maturitas* 2009, 64, 90–97. [CrossRef] [PubMed]
- 39. Bouchabou, D.; Nguyen, S.M.; Lohr, C.; LeDuc, B.; Kanellos, I. A survey of human activity recognition in smart homes based on IoT sensors algorithms: Taxonomies, challenges, and opportunities with deep learning. *Sensors* **2021**, *21*, 6037. [CrossRef]
- 40. Plötz, T.; Moynihan, P.; Pham, C.; Olivier, P. Activity recognition and healthier food preparation. *Act. Recognit. Pervasive Intell. Environ.* **2011**, 313–329.\_14. [CrossRef]
- Tapia, E.M.; Intille, S.S.; Larson, K. Activity recognition in the home using simple and ubiquitous sensors. In Proceedings of the Pervasive Computing: Second International Conference, PERVASIVE 2004, Linz/Vienna, Austria, 21–23 April 2004; Proceedings 2; Springer: Linz/Vienna, Austria, 2004; pp. 158–175.
- 42. 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, Republic of Korea, 21–24 September 2008; pp. 1–9.
- 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]
- Beddiar, D.R.; Nini, B.; Sabokrou, M.; Hadid, A. Vision-based human activity recognition: A survey. *Multimed. Tools Appl.* 2020, 79, 30509–30555. [CrossRef]
- 45. Kim, K.; Jalal, A.; Mahmood, M. Vision-based human activity recognition system using depth silhouettes: A smart home system for monitoring the residents. *J. Electr. Eng. Technol.* **2019**, *14*, 2567–2573. [CrossRef]
- 46. Jalal, A.; Kamal, S. Real-time life logging via a depth silhouette-based human activity recognition system for smart home services. In Proceedings of the 2014 11th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), IEEE, Seoul, Republic of Korea, 26–29 August 2014; pp. 74–80.
- Singh, D.; Psychoula, I.; Kropf, J.; Hanke, S.; Holzinger, A. Users' perceptions and attitudes towards smart home technologies. In Proceedings of the Smart Homes and Health Telematics, Designing a Better Future: Urban Assisted Living: 16th International Conference, ICOST 2018, Singapore, 10–12 July 2018; Proceedings 16; Springer: Singapore, 2018; pp. 203–214.
- 48. 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. *Neural Comput. Appl.* **2022**, *34*, 18289–18324. [CrossRef]
- 49. Ordóñez, F.J.; Roggen, D. Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition. *Sensors* **2016**, *16*, *115*. [CrossRef] [PubMed]
- 50. Li, X.; Zhang, Y.; Marsic, I.; Sarcevic, A.; Burd, R.S. Deep learning for rfid-based activity recognition. In Proceedings of the 14th ACM Conference on Embedded Network Sensor Systems CD-ROM, Stanford, CA, USA, 14–16 November 2016; pp. 164–175.

- 51. Gomes, L.; Sousa, F.; Vale, Z. An intelligent smart plug with shared knowledge capabilities. *Sensors* 2018, *18*, 3961. [CrossRef]
- 52. Hussain, Z.; Sheng, M.; Zhang, W.E. Different approaches for human activity recognition: A survey. arXiv 2019, arXiv:1906.05074.
- Chen, L.; Hoey, J.; Nugent, C.D.; Cook, D.J.; Yu, Z. Sensor-based activity recognition. *IEEE Trans. Syst. Man, Cybern. Part C (Appl. Rev.)* 2012, 42, 790–808. [CrossRef]
- 54. Chen, L.; Nugent, C.D. Human Activity Recognition and Behaviour Analysis; Springer: Berlin/Heidelberg, Germany, 2019.
- 55. Chen, L.; Nugent, C. Ontology-based activity recognition in intelligent pervasive environments. *Int. J. Web Inf. Syst.* 2009, 5, 410–430. [CrossRef]
- 56. Cook, D.J. Learning setting-generalized activity models for smart spaces. IEEE Intell. Syst. 2012, 27, 32. [CrossRef]
- Fahad, L.G.; Tahir, S.F.; Rajarajan, M. Activity recognition in smart homes using clustering based classification. In Proceedings of the 2014 22nd International Conference on Pattern Recognition, IEEE, Stockholm, Sweden, 24–28 August 2014; pp. 1348–1353.
- 58. Aminikhanghahi, S.; Cook, D.J. A survey of methods for time series change point detection. *Knowl. Inf. Syst.* **2017**, *51*, 339–367. [CrossRef]
- Khan, M.A.; Mittal, M.; Goyal, L.M.; Roy, S. A deep survey on supervised learning based human detection and activity classification methods. *Multimed. Tools Appl.* 2021, 80, 27867–27923. [CrossRef]
- Yu, M.; Rhuma, A.; Naqvi, S.M.; Wang, L.; Chambers, J. A posture recognition-based fall detection system for monitoring an elderly person in a smart home environment. *IEEE Trans. Inf. Technol. Biomed.* 2012, *16*, 1274–1286. [CrossRef] [PubMed]
- 61. Quigley, B.; Donnelly, M.; Moore, G.; Galway, L. A comparative analysis of windowing approaches in dense sensing environments. *Proceedings* **2018**, *2*, 1245.
- 62. Cook, D.J.; Krishnan, N.C.; Rashidi, P. Activity discovery and activity recognition: A new partnership. *IEEE Trans. Cybern.* 2013, 43, 820–828. [CrossRef] [PubMed]
- Yala, N.; Fergani, B.; Fleury, A. Feature extraction for human activity recognition on streaming data. In Proceedings of the 2015 International Symposium on Innovations in Intelligent SysTems and Applications (INISTA), IEEE, Madrid, Spain, 2–4 September 2015; pp. 1–6.
- Aminikhanghahi, S.; Cook, D.J. Using change point detection to automate daily activity segmentation. In Proceedings of the 2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), IEEE, Kona, Hawaii, 13–17 March 2017; pp. 262–267.
- 65. Aminikhanghahi, S.; Wang, T.; Cook, D.J. Real-time change point detection with application to smart home time series data. *IEEE Trans. Knowl. Data Eng.* **2018**, *31*, 1010–1023. [CrossRef] [PubMed]
- van Kasteren, T.L.M. Activity recognition for health monitoring elderly using temporal probabilistic models. In Proceedings of the ASCI, Seoul, Republic of Korea, 22–25 August 2011.
- 67. Medina-Quero, J.; Zhang, S.; Nugent, C.; Espinilla, M. Ensemble classifier of long short-term memory with fuzzy temporal windows on binary sensors for activity recognition. *Expert Syst. Appl.* **2018**, *114*, 441–453. [CrossRef]
- Hamad, R.A.; Hidalgo, A.S.; Bouguelia, M.R.; Estevez, M.E.; Quero, J.M. Efficient activity recognition in smart homes using delayed fuzzy temporal windows on binary sensors. *IEEE J. Biomed. Health Inform.* 2019, 24, 387–395. [CrossRef]
- 69. Hamad, R.A.; Yang, L.; Woo, W.L.; Wei, B. Joint learning of temporal models to handle imbalanced data for human activity recognition. *Appl. Sci.* **2020**, *10*, 5293. [CrossRef]
- 70. Hamad, R.A.; Kimura, M.; Yang, L.; Woo, W.L.; Wei, B. Dilated causal convolution with multi-head self attention for sensor human activity recognition. *Neural Comput. Appl.* **2021**, *33*, 13705–13722. [CrossRef]
- Aminikhanghahi, S.; Cook, D.J. Enhancing activity recognition using CPD-based activity segmentation. *Pervasive Mob. Comput.* 2019, 53, 75–89. [CrossRef]
- Li, H.; Abowd, G.D.; Plötz, T. On specialized window lengths and detector based human activity recognition. In Proceedings of the 2018 ACM International Symposium on Wearable Computers, Singapore, 8–12 October 2018; pp. 68–71.
- Al Machot, F.; Mayr, H.C.; Ranasinghe, S. A windowing approach for activity recognition in sensor data streams. In Proceedings of the 2016 Eighth International Conference on Ubiquitous and Future Networks (ICUFN), IEEE, Vienna, Austria, 5–8 July 2016; pp. 951–953.
- 74. Bermejo, U.; Almeida, A.; Bilbao-Jayo, A.; Azkune, G. Embedding-based real-time change point detection with application to activity segmentation in smart home time series data. *Expert Syst. Appl.* **2021**, *185*, 115641. [CrossRef]
- 75. Jethanandani, M.; Sharma, A.; Perumal, T.; Chang, J.R. Multi-label classification based ensemble learning for human activity recognition in smart home. *Internet Things* **2020**, *12*, 100324. [CrossRef]
- 76. Petersen, J.; Larimer, N.; Kaye, J.A.; Pavel, M.; Hayes, T.L. SVM to detect the presence of visitors in a smart home environment. In Proceedings of the 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE, San Diego, CA, USA, 28 August–1 September 2012; pp. 5850–5853.
- SEDKY, M.; HOWARD, C.; Alshammari, T.; Alshammari, N. Evaluating machine learning techniques for activity classification in smart home environments. *Int. J. Inf. Syst. Comput. Sci.* 2018, 12, 48–54.
- Nawal, Y.; Oussalah, M.; Fergani, B.; Fleury, A. New incremental SVM algorithms for human activity recognition in smart homes. J. Ambient. Intell. Humaniz. Comput. 2022, 1–18. [CrossRef]
- 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]

- 80. Fang, H.; Tang, P.; Si, H. Feature selections using minimal redundancy maximal relevance algorithm for human activity recognition in smart home environments. *J. Healthc. Eng.* **2020**, *2020*, 8876782. [CrossRef]
- 81. Pouyanfar, S.; Sadiq, S.; Yan, Y.; Tian, H.; Tao, Y.; Reyes, M.P.; Shyu, M.L.; Chen, S.C.; Iyengar, S.S. A survey on deep learning: Algorithms, techniques, and applications. *ACM Comput. Surv.* (*CSUR*) **2018**, *51*, 1–36. [CrossRef]
- Fang, H.; He, L.; Si, H.; Liu, P.; Xie, X. Human activity recognition based on feature selection in smart home using back-propagation algorithm. *ISA Trans.* 2014, 53, 1629–1638. [CrossRef]
- 83. Irvine, N.; Nugent, C.; Zhang, S.; Wang, H.; Ng, W.W. Neural network ensembles for sensor-based human activity recognition within smart environments. *Sensors* 2019, 20, 216. [CrossRef]
- 84. Hammerla, N.Y.; Halloran, S.; Plötz, T. Deep, convolutional, and recurrent models for human activity recognition using wearables. *arXiv* **2016**, arXiv:1604.08880.
- 85. Gochoo, M.; Tan, T.H.; Liu, S.H.; Jean, F.R.; Alnajjar, F.S.; Huang, S.C. Unobtrusive activity recognition of elderly people living alone using anonymous binary sensors and DCNN. *IEEE J. Biomed. Health Inform.* **2018**, *23*, 693–702. [CrossRef] [PubMed]
- Mohmed, G.; Lotfi, A.; Pourabdollah, A. Employing a deep convolutional neural network for human activity recognition based on binary ambient sensor data. In Proceedings of the 13th ACM International Conference on PErvasive Technologies Related to Assistive Environments, Corfu, Greece, 30 June–3 July 2020; pp. 1–7.
- Singh, D.; Merdivan, E.; Hanke, S.; Kropf, J.; Geist, M.; Holzinger, A. Convolutional and recurrent neural networks for activity recognition in smart environment. In Proceedings of the Towards Integrative Machine Learning and Knowledge Extraction: BIRS Workshop, Banff, AB, Canada, 24–26 July 2015; Revised Selected Papers; Springer: Banff, AB, Canada, 2017; pp. 194–205.
- 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]
- Bouchabou, D.; Nguyen, S.M.; Lohr, C.; Leduc, B.; Kanellos, I. Fully convolutional network bootstrapped by word encoding and embedding for activity recognition in smart homes. In Proceedings of the Deep Learning for Human Activity Recognition: Second International Workshop, DL-HAR 2020, Held in Conjunction with IJCAI-PRICAI 2020, Kyoto, Japan, 8 January 2021; Proceedings 2; Springer: Kyoto, Japan, 2021; pp. 111–125.
- 90. 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]
- 91. Zhou, J.; Cui, G.; Hu, S.; Zhang, Z.; Yang, C.; Liu, Z.; Wang, L.; Li, C.; Sun, M. Graph neural networks: A review of methods and applications. *AI Open* **2020**, *1*, 57–81. [CrossRef]
- Li, L.; Gan, Z.; Cheng, Y.; Liu, J. Relation-aware graph attention network for visual question answering. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019; pp. 10313–10322.
- 93. Ye, J.; Jiang, H.; Zhong, J. A Graph-Attention-Based Method for Single-Resident Daily Activity Recognition in Smart Homes. Sensors 2023, 23, 1626. [CrossRef] [PubMed]
- 94. Jaiswal, A.; Babu, A.R.; Zadeh, M.Z.; Banerjee, D.; Makedon, F. A survey on contrastive self-supervised learning. *Technologies* 2020, *9*, 2. [CrossRef]
- 95. Hiremath, S.K.; Plötz, T. Deriving effective human activity recognition systems through objective task complexity assessment. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **2020**, *4*, 1–24. [CrossRef]
- 96. Atlas, L.; Cohn, D.; Ladner, R. Training connectionist networks with queries and selective sampling. *Adv. Neural Inf. Process. Syst.* **1989**, 2.
- 97. Lewis, D.D. A sequential algorithm for training text classifiers: Corrigendum and additional data. In *Acm Sigir Forum*; ACM: New York, NY, USA, 1995; Volume 29, pp. 13–19.
- Settles, B. Active Learning Literature Survey. 2009. Avialable online: http://digital.library.wisc.edu/1793/60660 (accessed on 9 July 2023).
- 99. Adaimi, R.; Thomaz, E. Leveraging active learning and conditional mutual information to minimize data annotation in human activity recognition. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **2019**, *3*, 1–23. [CrossRef]
- Miu, T.; Plötz, T.; Missier, P.; Roggen, D. On strategies for budget-based online annotation in human activity recognition. In Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication, Seattle, WA, USA, 13–17 September 2014; pp. 767–776.
- Alemdar, H.; Van Kasteren, T.; Ersoy, C. Active learning with uncertainty sampling for large scale activity recognition in smart homes. J. Ambient Intell. Smart Environ. 2017, 9, 209–223. [CrossRef]
- Karami, A.B.; Fleury, A.; Boonaert, J.; Lecoeuche, S. User in the loop: Adaptive smart homes exploiting user feedback—state of the art and future directions. *Information* 2016, 7, 35. [CrossRef]
- 103. Li, J.; Shlizerman, E. Iterate & cluster: Iterative semi-supervised action recognition. arXiv 2020, arXiv:2006.06911.
- Mikelsons, G.; Smith, M.; Mehrotra, A.; Musolesi, M. Towards deep learning models for psychological state prediction using smartphone data: Challenges and opportunities. arXiv 2017, arXiv:1711.06350.
- 105. Asghari, P.; Soleimani, E.; Nazerfard, E. Online human activity recognition employing hierarchical hidden Markov models. J. Ambient Intell. Humaniz. Comput. 2020, 11, 1141–1152. [CrossRef]
- 106. Peters, M.E.; Neumann, M.; Iyyer, M.; Gardner, M.; Clark, C.; Lee, K.; Zettlemoyer, L. Deep contextualized word representations. *arXiv* **1802**, arXiv:1802.05365.

- 107. Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv* **2018**, arXiv:1810.04805.
- 108. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. *Adv. Neural Inf. Process. Syst.* 2017, 30.
- 109. Najeh, H.; Lohr, C.; Leduc, B. Convolutional Neural Network Bootstrapped by Dynamic Segmentation and Stigmergy-Based Encoding for Real-Time Human Activity Recognition in Smart Homes. *Sensors* **2023**, *23*, 1969. [CrossRef]
- 110. Codispoti, J.; Khamesi, A.R.; Penn, N.; Silvestri, S.; Shin, E. Learning from non-experts: An interactive and adaptive learning approach for appliance recognition in smart homes. *ACM Trans. Cyber-Phys. Syst. (TCPS)* **2022**, *6*, 1–22. [CrossRef]
- 111. Hiremath, S.K.; Nishimura, Y.; Chernova, S.; Plötz, T. Bootstrapping Human Activity Recognition Systems for Smart Homes from Scratch. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **2022**, *6*, 1–27. [CrossRef]
- 112. Cook, D.J.; Schmitter-Edgecombe, M. Assessing the quality of activities in a smart environment. *Methods Inf. Med.* 2009, 48, 480–485. [PubMed]
- 113. 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]
- 114. Diethe, T.; Borchert, T.; Thereska, E.; Balle, B.; Lawrence, N. Continual learning in practice. arXiv 2019, arXiv:1903.05202.
- 115. Farquhar, S.; Gal, Y. Towards robust evaluations of continual learning. arXiv 2018, arXiv:1805.09733.
- Ho, S.; Liu, M.; Du, L.; Gao, L.; Xiang, Y. Prototype-Guided Memory Replay for Continual Learning. *IEEE Trans. Neural Netw. Learn. Syst.* 2023. [CrossRef] [PubMed]
- 117. Jha, S.; Schiemer, M.; Ye, J. Continual learning in human activity recognition: An empirical analysis of regularization. *arXiv* 2020, arXiv:2007.03032.
- 118. Jha, S.; Schiemer, M.; Zambonelli, F.; Ye, J. Continual learning in sensor-based human activity recognition: An empirical benchmark analysis. *Inf. Sci.* 2021, 575, 1–21. [CrossRef]
- Kim, C.D.; Jeong, J.; Kim, G. Imbalanced continual learning with partitioning reservoir sampling. In Proceedings of the Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, 23–28 August 2020; Proceedings, Part XIII 16; Springer: Virtual, 2020; pp. 411–428.
- Mai, Z.; Li, R.; Kim, H.; Sanner, S. Supervised contrastive replay: Revisiting the nearest class mean classifier in online classincremental continual learning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 3589–3599.
- 121. Fatima, I.; Fahim, M.; Lee, Y.K.; Lee, S. A unified framework for activity recognition-based behavior analysis and action prediction in smart homes. *Sensors* **2013**, *13*, 2682–2699. [CrossRef]
- 122. Monk, T.H.; Petrie, S.R.; Hayes, A.J.; Kupfer, D.J. Regularity of daily life in relation to personality, age, gender, sleep quality and circadian rhythms. *J. Sleep Res.* **1994**, *3*, 196–205. [CrossRef]
- 123. Banovic, N.; Krumm, J. Warming up to cold start personalization. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2018, 1, 1–13. [CrossRef]
- 124. Mazankiewicz, A.; Böhm, K.; Bergés, M. Incremental real-time personalization in human activity recognition using domain adaptive batch normalization. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2020, *4*, 1–20. [CrossRef]
- 125. Xu, M.; Qian, F.; Mei, Q.; Huang, K.; Liu, X. Deeptype: On-device deep learning for input personalization service with minimal privacy concern. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **2018**, *2*, 1–26. [CrossRef]
- 126. 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]

**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.