Chapter 2

Preliminaries and Related Work

This chapter presents the preliminaries that are necessary for understanding our approaches, discussions, and conclusions. We also present state-of-the-art work on the problems of early dementia detection, future ADL prediction, online ADL recognition, and recognizing ADL through transfer learning.

2.1 Preliminaries

This section presents the terminologies used in this work. It includes the concept of smart home, particularly as a *supporting system* to assist older adults in independent living. Subsequently, in Section 2.1.2, we introduce different sensor modalities used to gather human activities information. Finally, in Section 2.1.3 we define the activities of daily living.

2.1.1 Smart home

A smart home is an application of ubiquitous computing that involves incorporating smartness into dwellings for comfort, healthcare, safety, security, and energy conservation [42]. Smart home is defined by Satpathy as a home which is smart enough to assist the inhabitants to live independently and comfortably with the help of technology. All the mechanical and digital devices of a smart home are interconnected to form a network, which can communicate with each other and with the user to create an interactive space [15]. Smart homes offer a better quality of life by introducing automated appliance control and assistive services. The main objective of a smart home is to enhance comfort, energy-saving, and security for the inhabitants in the home. In the earlier development, the idea was oriented to building a smart home environment for ordinary non-disabled persons with the simple purpose of enhancing home comfort [43]. Recently, the same technology has become a bright perspective for people with special need. They may help enhance the quality of life, prolong independent living, and reduce healthcare costs and caregivers' burdens [15,42,43]. Among all the possible applications of a smart home, we will focus our attention on the possibilities of using smart home to support the independent living of the elderly.

In recent years various projects have been developed throughout the world to prolong independent living for this population segment, such as Center for Advanced Studies in Adaptive Systems (CASAS) [44] at Washington State University, Adaptive Versatile home (MavHome) [45] at the University of Texas, PlaceLab [46] at MIT, GatorTech Smart House [47] at University of Florida, Aware Home [48] at Georgia Tech University, iDorm [49] at University of Essex, Adaptive House [50] at University of Colorado, and Intelligent System Lab [51] at University of Amsterdam. These projects aim to provide ambient assisted living for older people, remote monitoring, early detection of problems, and safety. *Ambient assisted living* involves the use of information and communication technologies in a person's daily living and working environment to enable them to stay active longer, remain socially connected and live independently into old age [52].

Datasets are significantly crucial for machine learning algorithms testing and validation. The collection of real data is challenging due to the involved budget, human resources, and annotation cost; that's why most researchers prefer to utilize existing datasets for evaluation purposes. In this work for empirical evaluation, we consider real datasets collected at Washington State University's CASAS smart home testbeds [41]. We chose CASAS datasets because they are well-established and open to external usage (*i.e.* datasets are publicly available). Also, these datasets perfectly fulfill all our assumptions stated in Section 1.2.

2.1.2 Sensor modalities

With the recent advancements of sensor technology, various kinds of low-cost, low-power, and miniaturized sensors have been utilized in our daily life to gather human activity information. The extensively used sensors largely include the *wearable sensors* and the *environmental sensors.* They generate massive amounts of sensor readings containing information about human activities [53].

Wearable sensors are sensors worn on the body of inhabitants. These sensors can be embedded into wristwatches, shoes, and clothes or placed directly or indirectly on the body. Various wearable sensors have been used in existing literature, such as accelerometers, gyroscopes, and magnetometers. They can provide reasonable information about the pose and the movement of the humans. Therefore, they are suitable for recognizing activities that are mainly characterized by ambulatory movements, like running, walking, sitting, lying down, standing, and falling. As they are worn on the body of humans, they are not restricted to the environment and can be carried by humans to different environments. However, this characteristic also makes the wearable sensors can capture little information about the environment. Thus, they are not suitable for recognizing activities that involve many interactions with the environment [53]. They have other natural flaws also, such as short battery life, require regular maintenance, and could be easily lost or forgotten. Moreover, wearable sensors may be uncomfortable and inconvenient for inhabitants, and also their accuracy relies upon the body attachment position [6, 30].

Environmental or external or ambient sensors are not affixed to the inhabitants performing the activity but are deployed in the environment surrounding them. Environmental sensors are Passive InfraRed (PIR) motion, magnetic door, ambient light, temperature, ambient sound, pressure, vibration sensors, *etc.* They can be attached to the objects, such as the magnetic sensors attached to doors and cabinets to detect door/cabinet openings and closings. Also, they can be deployed at some fixed locations in the house, such as the PIR motion sensors attached to the wall or on the ceiling to detect motion. Since these sensors are positioned in the environment, information about the environment can be obtained from them. Therefore, they are suitable for identifying activities that have interactions with the environment. In applications related to smart homes, we aim to recognize daily living activities like cooking, having a meal, taking medicine, bathing, personal hygiene, and housekeeping. These activities involve many interactions with the environment. Moreover, they are small in size, low-cost, suitable to supply constant supervision, easy to install, and flexible. Apart from these, their non-intrusive characteristics specifically make them suitable for home environments where user acceptance and privacy are needed [54]. Thus, environmental sensors are considered acceptable solutions for sensing smart homes and are widely used in existing works for gathering activity information [6, 30, 31, 33–35, 39, 40]. Among the various types of environmental sensors, a large portion of them belong to binary sensors. Binary sensors are sensors that have binary values. They can be state-change sensors, like reed switch-based magnetic door sensors and PIR motion sensors. The limitation of environmental sensors is that they are deployed in a fixed environment and can identify activity within it [53]. Another limitation of these sensors is the uncertainty of providing and separating detailed information about the subject [55]. Thus, in a multi-resident scenario, it is challenging to recognize different residents from each other and assign each resident's performed activities. Wearable sensors or a combination of environmental and wearable sensors are used in some existing works to overcome this limitation.

Other widely used devices for gathering activity information are cameras. The cameras record a wealth of information about human activities in the generated videos and are the best in terms of accuracy; however, camera sensors are not practical in smart home settings due to privacy-invasiveness. In addition, cameras are often considered recording devices. When using them to collect information about activities, humans may behave unnaturally. This may affect the quality of the information entered [53]. Finally, different sensor modalities can also be combined to gather human activity information.

2.1.3 Activities of daily living

An important class of healthcare and assisted living activities is the so-called Activities of Daily Living. ADL include necessary self-care activities that individuals perform daily, such as meal preparation, eating, bathing, toileting, personal hygiene, sleeping, or taking medicine. Individuals need to be able to complete ADL to function independently at home. Whether or not an individual can perform ADL on their own or if they rely on a caregiver to perform the ADL provides a comparative measure of their functional status. Many older adults with cognitive illnesses and functional disabilities have difficulties in performing ADL tasks and require intervention. Such interventions may range from minor infrastructural modifications to the provision of additional assistive technologies; in extreme cases, such people will need institutionalized care. Smart homes envisage the provision of technological services that will aid people to live independently for as long as possible.

2.2 Related Work

In this section, we discuss and review the state-of-the-art solution approaches for the problems of early dementia detection, future ADL prediction, online ADL recognition addressing the long-tailed class distribution problem, and recognizing ADL through transfer learning. Section 2.2.1 discusses the related works on travel pattern classification approaches for early dementia detection. Section 2.2.2 reviews related works on future ADL labels, locations, and starting time prediction. Section 2.2.3 reviews related works on ADL recognition from both wearable and environmental perspectives, further identifying the main pros and cons of existing environmental sensors-based systems and highlighting our system's advantages. Finally, Section 2.2.4 concludes the chapter by reviewing transfer learning approaches for ADL recognition.

2.2.1 Detection of dementia at an early stage

In this section, we present the literature on travel pattern classification approaches for early dementia detection using data gathered through environmental sensors. We also present the existing work on feature extraction of sensory data.

Over the last decades, different approaches have targeted early detection of dementia, such as neuroimaging, behavioral markers, *etc.* Neuroimaging (Medical brain imaging techniques), *e.g.*, Magnetic Resonance Imaging (MRI), Computed Tomography (CT), Positron Emission Tomography (PET), or Single-Photon Emission Computed Tomography (SPECT) scans are standard methods of diagnosing dementia. However, by the time individuals receive the diagnosis; cognitive impairments will generally have progressed. Furthermore, both the subtle changes produced in the brain and the lack of a complete understanding of the disease development still pose challenges to the diagnosis assistance through neuroimaging [56]. Also, due to the high costs of neuroimaging and their other associated side effects, it is crucial to explore behavioral markers to predict the development of dementia successfully [57]. Consequently, current research aims to determine prodromal markers for early detection of dementia, for example, changes in movement patterns [57].

A prominent symptom of dementia is the change in movement patterns (*i.e.,* wandering or aimless movements) that arise as a consequence of forgetfulness and confusion.

Previous literature suggests that changes in movement start years before the onset of dementia, suggesting that monitoring the movement could be useful to enable early diagnosis [58]. Observational methods are the only approach to measure wandering behavior in clinical settings [59]. Nursing scientists apply both direct and videotape observation techniques to study the wandering of People with Dementia (PwD). In both cases, vigilant observers or coders are hired to manually capture and document wandering episodes according to a predefined protocol. Direct observation or cameras in people's homes, in particular, are considered highly invasive. These drawbacks make this approach unsuitable for long-term monitoring [59]. Smart homes are an emerging technological solution for monitoring inhabitant's behavior using various sensors modalities.

In existing works, different sensor modalities used to monitoring inhabitants and gather movement information include wearable sensors, video sensors and environmental sensors [60]. Vuong *et al.* [59] considered movement data gathered from RFID tags for the detection of travel patterns in people with dementia. In paper [61], authors presented a grid-based layout representation strategy to identify wandering using GPS for outdoor, and an ultra-wideband radio tag for indoor localization. Similarly, the authors in [62] used GPS to detect mobility patterns associated with wandering and disorientation in outdoor scenarios. In the video sensors, stationary cameras are placed in the home environment to monitor inhabitant movements. For example, paper [36] proposed a video-based intelligent monitoring system. The proposed system was equipped with a camera network and an automatic elopement detection algorithm to reduce the risks of un-witnessed elopements from dementia. Very few existing work [2, 63] considered environmental sensors to gather movement data for recognizing the travel patterns.

Some existing work [36, 59] uses hand-crafted features for recognizing the travel patterns. However, such features are always extracted via a heuristic and hand-crafted way, which is usually limited by the domain knowledge of human [64, 65]. Recent advances in deep learning enable the automatic extraction of high-level features from the raw sensor data without any specific domain knowledge. in many tasks such as object recognition, speech recognition, natural language processing, machine translation and many other areas. The authors in [2] proposed a deep convolutional neural network classifier for indoor travel patterns. They consider the Aruba dataset presented by CASAS. Furthermore, this paper also highlighted the importance of device-free non-privacy invasive PIR motion sensors for movement data collection.

Many researchers [66, 67] have found that the Martino-Saltzman (MS) model is a useful tool to detect wandering patterns of PWD. Studies [59, 61–63] have employed MS model to monitor older adults for outdoor or indoor wandering detection purposes. Vuong *et al.* [59] proposed a deterministic predefined tree-based travel pattern classification algorithm based on the MS model. This algorithm is simple and accurate for detecting MS travel patterns [2]. We consider this algorithm in our study to prepare the ground truth.

2.2.2 Future ADL prediction

This section covers the related work on future ADL prediction. Although the activity recognition problem is heavily explored, less attention has been given to predicting the future activities that the inhabitant is likely to perform. In the machine literature, *prediction* primarily refers to *sequential prediction*, where the goal is to predict the activity that will happen next based on a known limited history of past activities. Associated with the future ADL prediction problem are the problems of predicting the label, the location, and the starting time of the next ADL.

2.2.2.1 Future ADL label prediction

There have been a few works that predict the label of the future unobserved activity, such as [68–70]. In [68] authors proposed an algorithm called sequence prediction via enhanced episode discovery (SPEED) based on Markov models, which uses the prediction by partial matching (PPM) to predict the most probable next activity. In [69] SPEED was modified to improve accuracy by including a time component. In [70], authors proposed an activity prediction model called Current activity and features to predict next features (CRAFFT) that uses Bayesian networks. CRAFFT first predicts the next activity features and then predicts the next activity label using the features predicted earlier and the current activity label.

2.2.2.2 Future ADL location prediction

In [71], authors presented a LeZi Update algorithm to track user mobility using the $LZ78$ data compression algorithm [72]. It constructs a decision tree by employing $LZ78$ dictionary contexts and predicts the likely user location based on a prediction by partial matching (PPM). It creates an order $k-1$ Markov model, where k is the length of the longest dictionary words. Authors in [73] improved the LeZi Update, proposed the Active LeZi algorithm to predict device usage in the smart home. Active LeZi employs a variable-length window to reduce the data loss across phase boundaries, and it converges faster than LeZi Update algorithm since it gathers more data.

2.2.2.3 Future ADL starting time prediction

Predicting the starting times of future unobserved activities is a new research problem. To the best of our knowledge, there are only two relevant works [74,75] in the domain of environmental sensor data analysis. Authors in [74] presented a regression-tree-based activity forecasting method to predict the occurrence of future activities for prompting initiation of such activities. Another regression method was addressed in [75] to predict the future activity occurrence times from real sensor data collected from smart home testbeds.

2.2.3 ADL recognition

This section presents the literature on ADL Recognition (ADLR) and the existing approaches for handling long-tailed class distribution problems. ADLR aims to identify human ADL from data gathered through various sensors. From a machine learning viewpoint, ADLR maps a sequence of sensor events to a corresponding ADL label. ADLR is crucial for developing advanced assistive services and is a well-researched problem, and thus, there exist many approaches for ADLR. These approaches differ mainly depending on the underlying sensor modalities for gathering the ADL data and different feature extraction methods used to recognize activities. Table 2.1 shows a comparison summary of the features considered in the existing work.

2.2.3.1 Sensor modalities for ADLR

A wide variety of sensor modalities have been utilized to gather environmental and human activity information to recognize ADL. The widely used sensors include the *wearable sensors* worn on the body of the humans and/or the *environmental/external sensors* deployed in the home or working environments.

In existing works, various types of wearable sensors have been considered, such as accelerometers, gyroscopes, and magnetometers [19–28]. They are implanted into wristwatches, shoes, and clothes or placed directly or indirectly on the human body. Thus, they are commonly used for recognizing activities that are primarily defined by ambulatory movements such as running, walking, sitting, lying down, and standing [20, 22, 23, 27]. However, wearable sensors can capture little information about the environment. Thus, they are not suitable for recognizing complex activities that involve many interactions with the environment, such as daily living activities [53]. Moreover, wearable sensors have other natural flaws, such as the discomfort of wearing, short battery life, require regular maintenance, could be easily lost or forgotten, and also their accuracy relies upon the body attachment position [6, 30].

Table 2.1: Comparative summary of some of the existing work on human activity recognition, where **W, V, E, Hc** and **Hi** indicate wearable sensors, video sensors, environmental sensors, hand-crafted features, and high-level features, respectively.

Paper ID	Year	Sensors			Imbalanced learning	Features		Online
		W	V	E		Hc	Hi	
[36]	2007		√				√	
[31]	2014			√		√		√
[37]	2016		√				√	
[38]	2017		√				√	
[26]	2018	√				√		
[34]	2018			√		√		√
[35]	2018			√		√		
[30]	2018			√			√	
[21]	2019	√					√	
[22]	2019	√				√		
[23]	2019	√				√		
[6]	2019			√			√	
[24]	2019	√		√		√		
[32]	2019			√		√		√
[25]	2019	√					√	
[27]	2019	√				√		
[28]	2019	√				√		
[33]	2019			√			√	√
[76]	2020	√				√		
[77]	2020	√					√	√
[78]	2021			√			√	

Environmental sensors such as infrared motion, magnetic door, light, temperature, or pressure sensors are considered acceptable solutions for recognizing ADL without adding the extra burden of wearing sensors [6,30,33]. As these sensors are placed in the environment, information about the environment can be obtained from them, making them suitable for recognizing complex activities involving many interactions with the environment. Moreover, due to their suitability to supply constant supervision, low cost, flexibility, and rapid deployment, these sensors are considered one of the most promising technologies for enabling health monitoring. Their non-intrusive characteristics make them suitable for home environments where user acceptance and privacy are needed [54]. Many existing works [6, 30, 31, 33–35, 40] are considered smart homes in which environmental sensors are embedded on different objects to gather information about the daily life activities of inhabitants for ADLR.

Researchers have also used data from video cameras for monitoring and recognizing different activities [36–38]. The use of video cameras for activity recognition is widespread in security-related applications. However, they are not practical in smart home settings where privacy is crucial since they are often regarded to be intrusive. Secondly, video cameras are considered recording devices. When using them to collect information about activities, humans may behave unnaturally. In contrast, this problem is less severe when wearable or environmental sensors are used for collecting human activity information. Thirdly, the videos generated by the cameras are generally a type of data with high dimensions and have very high computation costs. In contrast, the cost of dealing with wearable or environmental sensors readings is relatively very low [53]. With the above advantages, wearable and environmental sensors have got increasingly wide adoption in human activity recognition [6, 19–28, 30, 31, 33–35, 39, 40]. For sensing smart homes, environmental sensors are considered primarily, and they are widely used in the existing works [6, 30, 31, 33–35, 39, 40] for gathering human activity and environmental information. For non-intrusive ADLR, this work consider data collected from environmental sensors and mainly from binary PIR motion and magnetic door sensors that can be easily embedded in a smart home environment.

2.2.3.2 Feature extraction methods

Based upon feature extraction methods used to recognize activities, existing ADLR approaches can be categorized into conventional Hand-crafted features-based and High-level features-based techniques. (1) *Hand-crafted features:* Many existing work [22–24, 27, 28, 31, 32, 34, 35] uses hand-crafted features for recognizing the ADL. The hand-crafted features based ADLR system heavily relies on human domain knowledge or experience. Furthermore, only shallow features can be learned according to human expertise [64]. (2) *High-level features*: Recent advances in deep learning enable the automatic extraction of high-level features from the raw sensor data without any specific domain knowledge. Thus deep neural networks achieve promising performance in many areas, including activity recognition [64, 65].

• **ADLR from environmental sensor data:** The authors in [6] proposed the ADLR system where the sensor events are segmented and sliding windowed, then transformed into activity images. Then, CNN is used to obtain spatial information from the activity images to recognize them. Hamed *et al.* [40] proposed a method including multiple and incremental Fuzzy Temporal Windows (FTWs) to compute the features from a given sensor in both preceding activations, as well as oncoming activations. The proposed method used CNN and LSTM to recognize the activities. Yan *et al.* [33] proposed an approach that learns the latent knowledge from explicit-activity windows and determines the prediction for a given sliding window. The method then feeds the prediction with other sliding window features into a classifier to produce the final class prediction for each sliding window. The authors in [30] proposed a deep learning classifier based on RGB activity image for the unobtrusive classification of the multi-resident activities. The authors used CNN for high-level feature extraction.

2.2.3.3 Class imbalanced learning

It aims to lift the importance of minority classes. Existing approaches include (1) *Data-level:* These methods change the class representations in the original data and do not require any change to the learning algorithms. Typical methods include over-sampling the minority classes, down-sampling the majority classes, or both [79, 80]. However, down-sampling can lose valuable information about the majority class data; over-sampling can cause model overfitting due to frequently visiting duplicated training samples and makes the training computationally burdensome. (2) *Algorithm-level:* Modifying existing learning algorithms to improve the sensitivity of the classifier toward minority classes [81–83]. Chung *et al.* [82] proposed a loss function, which substitutes conventional softmax with a regression loss. Similarly, the authors in [83] proposed a loss function, which gives equal emphasis to mistakes in the majority and minority classes. The main limitation of algorithm-level approaches is, in general, challenging to optimize the cost matrix or relationships. It is problem-specific and non-scalable since usually given by experts.

2.2.4 Transfer learning for ADL recognition

In this section, we present the related work on ADL recognition with unsupervised heterogeneous transfer learning technique.

Transfer learning has been successfully employed in several applications such as natural language processing [84], visual object recognition [85], and Wi-Fi localization [86]. As given in the literature survey [87], transfer learning techniques can be grouped into three categories: instance-based, parameter-based, and feature-based techniques. Instance-based techniques perform knowledge transfer essentially through instance reweighting techniques. Parameter-based techniques first train a model using the labeled source domain, then perform clustering on the target domain. Feature-based techniques learn a feature transformation between domains when the distance can be minimized.

Our work is related to the feature-based category, which brings the source and target domain features into the common subspace. Manual mapping approaches have been considered to overcome differences in the source and target domain feature spaces by many researchers [88–90]. These approaches do not need any labeled data in the target domain. Still, they need the manual definition of sensor locations (kitchen, bathroom, bedroom, *etc.*) to map sensors from one smart home to another. In [90] authors group sensors by their function/location. Then the sensors of the source domain are mapped to identical sensors of the target domain. In paper [90], authors introduce a manual mapping between sensors in different smart homes and learn the target model parameters using the EM algorithm to transit probabilities of HMM models from source to target. Similarly, in [89], authors learn sensor mappings based on their activity models' locations and roles. The role is defined in mutual information, measuring the mutual dependence between a sensor and an activity and implying the importance of using the sensor to predict the corresponding activity. Paper [91] proposes a data-driven approach to map sensors according to their meta-features, essentially about when a sensor reports and time intervals between events reported by this sensor and other sensors. We have used manual mapping similar to the work mentioned above to link sensor features between target and source domains.

2.3 Conclusion

This chapter presented the existing work on assisting older adults in independent living to get a quick understanding of the notable contributions that have been made over the years. To assist older adults, we considered problems of early dementia detection, future ADL prediction, online ADL recognition, and recognizing ADL through transfer learning.

After reviewing the existing literature on early dementia detection through travel pattern classification, we observed that most of the researchers have focused on either *wearable sensors* or *video sensors* to gather movement data. As discussed earlier, wearable sensors may be inconvenient for inhabitants, and video sensors are not practical due to privacy issues. Furthermore, multimodality-based solutions which consider two or more types of sensors (*e.g.*, wearable and video sensors) consist of the limitations of both types of sensors. Also, the existing works are either based on conventional feature engineering, which is usually limited by the domain knowledge of humans or requires transforming the travel patterns into image-like representations. Thus, further research is required to address these issues in the problem of early dementia detection.

Contrasted with the significant research done to recognize ADL, less consideration has been given to predicting the future ADL that the inhabitant is likely to perform. There are some relevant works [92–94] in the computer vision field, but sensor-based future ADL prediction is a new research problem. Associated with the future ADL prediction problem are the problems of predicting label, location, and starting time of the future ADL. The existing literature on future ADL label prediction [68, 70], location prediction [71–73], and starting time prediction [74, 75] use individual models for predicting the upcoming activity, location, and its starting time, respectively. These models do not employ the advantage of multi-task learning. Existing literature suggests that multi-task learning has been successfully applied in a multitude of machine learning problems [95–97]. The intuition behind multi-task learning is that if two or more tasks are correlated, the joint model can learn effectively from the shared representations. To the best of our knowledge, we are the first to address the problem of jointly predicting the activity labels, location, and starting time.

After reviewing the existing work on ADL recognition, we observed that most researchers have mainly focused on one or more of the following three parameters: *sensor modalities* for gathering the ADL data, different *feature extraction methods* used to recognize ADL and *real-time/online* or *offline* ADL recognition. Table 2.1 shows that most existing works considered wearable sensors, and some of the existing approaches attempted to recognize ADL in real-time/online. Online ADLR requires dealing with un-segmented activity streams and recognizing and discarding or labeling data unrelated to predefined activity classes. Furthermore, the unpredictability nature of future events and the requirement of dealing with the un-segmented activity streams make online ADLR very challenging [31, 33]. Until now, research on ADLR has mostly centered around offline recognition. Also, existing work either considered hand-crafted features or high-level features, but none attempted to combine them both. Moreover, there exists no work that can handle the issues of class imbalanced learning. In contrast, ADLR is intrinsically an imbalanced classification problem since various ADL may occur at different frequencies. Thus, further research is required to address these issues in online ADL recognition.

In contrast to the notable research that has been done to recognize ADL through supervised learning, little attention has been paid to recognizing ADL through transfer learning with environmental sensor data. Thus, further research is needed to address the problem of recognizing ADL through transfer learning effectively. This work aims to deal with the rarely addressed case of ADL recognition with data obtained through environmental binary sensors embedded in smart homes using unsupervised heterogeneous transfer learning techniques.