Towards Reliable, Stable and Fast Learning for Smart Home Activity Recognition

Rebeen Ali Hamad

Supervisors:
Thorsteinn Rögnvaldsson
Mohamed-Rafik Bouguelia
Eric Järpe
Jens Lundström

LICENTIATE THESIS | Halmstad University Dissertations no. 85

Towards Reliable, Stable and Fast Learning for Smart Home Activity Recognition

Halmstad University Dissertations no. 85
ISBN 978-91-88749-79-6 (printed)
ISBN 978-91-88749-80-2 (pdf)
Publisher: Halmstad University Press, 2022 | www.hh.se/hup
Printer: Media-Tryck, Lund

Towards Reliable, Stable and Fast Learning for Smart Home Activity Recognition

Abstract

The current population age grows increasingly in industrialized societies and calls for more intelligent tools to monitor human activities. The aims of these intelligent tools are often to support senior people in their homes, to keep track of their daily activities, and to early detect potential health problems to facilitate a long and independent life. The recent advancements of smart environments using miniaturized sensors and wireless communications have facilitated unobtrusively human activity recognition.

Human activity recognition has been an active field of research due to its broad applications in different areas such as healthcare and smart home monitoring. This thesis project develops work on machine learning systems to improve the understanding of human activity patterns in smart home environments. One of the contributions of this research is to process and share information across multiple smart homes to reduce the learning time, reduce the need and effort to recollect the training data, as well as increase the accuracy for applications such as activity recognition. To achieve that, several contributions are presented to pave the way to transfer knowledge among smart homes that includes the following studies. Firstly, a method to align manifolds is proposed to facilitate transfer learning. Secondly, we propose a method to further improve the performance of activity recognition over the existing methods. Moreover, we explore imbalanced class problems in human activity recognition and propose a method to handle imbalanced human activities. The summary of these studies are provided below.

In our work, it is hypothesized that aligning learned low-dimensional manifolds from disparate datasets could be used to transfer knowledge between different but related datasets. The t-distributed Stochastic Neighbor Embedding(t-SNE) is used to project the high-dimensional input dataset into low-dimensional manifolds. However, since t-SNE is a stochastic algorithm and there is a large variance of t-SNE maps, a thorough analysis of the stability is required before applying Transfer learning. In response to this, an extension to Local Procrustes Analysis called Normalized Local Procrustes Analysis (NLPA) is proposed to non-linearly align manifolds by using locally linear mappings to test the stability of t-SNE low-dimensional manifolds. Experiments show that the disparity from using NLPA to align low-dimensional manifolds decreases by order of magnitude compared to the disparity obtained by Procrustes Analysis (PA). NLPA outperforms PA and provides much better alignments for the

low-dimensional manifolds. This indicates that t-SNE low-dimensional manifolds are locally stable, which is the part of the contribution in this thesis.

Human activity recognition in smart homes shows satisfying recognition results using existing methods. Often these methods process sensor readings that precede the evaluation time (where the decision is made) to evaluate and deliver real-time human activity recognition. However, there are several critical situations, such as diagnosing people with dementia where "preceding sensor activations" are not always sufficient to accurately recognize the resident's daily activities in each evaluated time. To improve performance, we propose a method that delays the recognition process to include some sensor activations that occur after the point in time where the decision needs to be made. For this, the proposed method uses multiple incremental fuzzy temporal windows to extract features from both preceding and some oncoming sensor activations. The proposed method is evaluated with two temporal deep learning models: one-dimensional convolutional neural network (1D CNN) and long short-term memory (LSTM) on a binary sensor dataset of real daily living activities. The experimental evaluation shows that the proposed method achieves significantly better results than the previous state-of-the-art.

Further, one of the main problems of activity recognition in a smart home setting is that the frequency and duration of human activities are intrinsically imbalanced. The huge difference in the number of observations for the categories means that many machine learning algorithms focus on the classification of the majority examples due to their increased prior probability while ignoring or misclassifying minority examples. This thesis explores well-known class imbalance approaches (synthetic minority over-sampling technique, cost-sensitive learning and ensemble learning) applied to activity recognition data with two temporal data pre-processing for the deep learning models LSTM and 1D CNN. This thesis proposes a data level perspective combined with a temporal window technique to handle imbalanced human activities from smart homes in order to make the learning algorithms more sensitive to the minority class. The experimental results indicate that handling imbalanced human activities from the data-level outperforms algorithm level and improved the classification performance.

Acknowledgements

I would like to express my gratitude and appreciation to my supervisors team:

- Principal supervisor: Thorsteinn Rögnvaldsson
- Co-supervisors: Mohamed-Rafik Bouguelia and Eric Järpe
- Previous supervisors: Jens Lundström

for providing me with this opportunity to be a student in Halmstad University. This opportunity could help me to explore science and attend research venues and present my research work. I would like to thank my colleagues in the lab, particularly Saeed Gholami Shahbandi. Finally, I would like to thank my parents.

This research project in the Center for Applied Intelligent Systems Research (CAISR) was supported by the Knowledge Foundation (Grant Agreement No. 20100271).

Contents

List of Figures

List of Tables

List of Algorithms

Chapter 1
Introduction

Activity Recognition (AR) is a challenging and highly dynamic research field aiming at recognizing human activities based on sensor observation data [4]. AR is fundamental for different application including smart homes [5], healthcare [6] and surveillance [7]. AR systems can be used to continuously and remotely monitor the physical human activities of a resident to alleviate the caregiver's burden and to reduce the economic pressures of families [8]. Hence, AR systems can help caregivers to monitor movements of residents in their own homes and send alerts to the caregiver whenever abnormal activities occur that may cause physical damage to the residents [9].

1.1 Context and motivation

The aging and dependent population have been recognized as a major social and economic challenge for the coming decades. According to the World Health Organization (WHO), the population of elderly people has increased drastically in the past decades, and estimated that there will be nearly two billion people aged 60 and older by 2050 [10]. In Europe, it is estimated that the elderly population above 65 years of age will rise to nearly 30% in 2060 [11]. Elders who are dependent and vulnerable in several perspectives due to physical and cognitive impairment require assistance in their activities of daily living (ADLs) [12]. ADLs are the normal daily activities that we perform for self-care such as eating, drinking, and bathing [13]. The increase of the elderly population and the rising cost of healthcare may bring a major issue and stress to the society [14]. One of the promising solutions to this challenge is ambient assisted living (AAL) systems. Such systems aim to reduce the costs of healthcare and would enable elders to live independently in their home [15]. AAL system consists of sensors, actuation, and networking technology, and data processing techniques to assist elderly people with their daily physical activities to help them stay safe and healthy while living independently. One of the most important roles and components of the AAL system is Activity Recognition (AR). AR is an active and challenging research field using sensors and artificial intelligence (AI) methods that can stimulate different applications such as healthcare monitoring, resident situation assessment, and behaviour pattern recognition in pro-active home care [9, 15, 16, 17]. Moreover, AR systems can support or even replace human operators to improve the performance and effectiveness of the observation and analysis process. For instance, an AR system can keep track of the health condition of older adults and inform the health staff in urgent cases [18, 19, 20]. Hence, AR could be used to perform recognition of dangerous situations and detects deviations of behavior to improve elderly-care alert systems [21].

The improvement in sensor technology and wireless communication networks regarding cost efficiency, capacity increase, and power efficiency made them feasible for AR [22]. Since these miniaturized sensors have been deployed in smart home environments, a vast amount of data has been produced. While the data supply is increased, the demand for methods to process and extract useful information from such a huge amount of data in a reasonable amount of time is also increased. To meet this demand, data-driven methods that are easily applicable to novel settings have been used. Moreover, to infer human activities, machine learning methods have been used on the recorded data from smart homes, however, these methods need labeled data to be trained on. Labeling recorded data is expensive since it requires time and human effort. Although the labeled data is essential, it is rarely useful when recorded in laboratory settings following predefined scenarios since the recorded data does not reflect the normal human activity [23].

1.2 Challenges of activity recognition

The AR of daily physical activities from smart home environment data is a challenging field. There exist several practical limitations regarding the layout of the smart home environment such as the number, location, and type of sensors. In addition to these obvious limitations, several additional issues immediately affect the success of human AR systems. Factors that contribute to the complexity of the AR task can be categorized into the following types.

1.2.1 Labeling sensor readings

A particular difficulty for employing machine learning techniques is to label a substantial amount of sensor readings in each smart home for training a model [24]. Properly labeling a large number of sensor readings for their correct corresponding human activities needs a domain knowledge expert and is an extremely time-consuming process [25, 26]. Mainly, AR systems are profoundly developed to model a single smart home data. Moreover, such systems rely on the assumption that the distribution of the training data and testing data are the same [27]. However, this assumption is seldomly valid in real-world applications. Hence, developing a technique to reduce the need for labeled data and share knowledge across different but related smart homes to further increase the accuracy for AR systems is demanding and will be an important contribution.

1.2.2 Real-time constraints

Real-time systems play a vital role in various applications such as credit card fraud detection [28] to business analytics[29], and healthcare monitoring [30]. Machine learning can continuously support learning applications by updating the system's learning model whenever new data is fed to the system. Hence, real-time learning system has become more demanding particularly in surveillance and healthcare system. Assessing a patient's health in real-time could lead to enhanced diagnoses and treatment. Moreover, in surveillance systems, real-time detection enables humans to respond to abnormal behavior immediately. Several abnormal activities, including illegal parking, traffic rule-breaking, shooting in public places can be detected by a real-time intelligent surveillance system. A real-time system can be used to detect fire, to prevent terrorism and theft at public places by generating an alarm automatically [31].

Real-time AR that demands to recognize activities in a real-time manner has the following challenges. Firstly, the real-time application often needs one-pass algorithms to process input data streams with a short real-time delay. Secondly, to keep performance high as an essential demand for real-time systems, a minimum number of features are needed. However, often numerous mutual exclusive training features are required to enhance the prediction accuracy of a system.

Real-time AR systems could be used by family members or caregivers to monitor the activities of elderly people or people with Alzheimer to assist them when it is needed. Existing approaches for AR in smart homes perform predictions in real-time based on sensor activations that precede the evaluation time. Due to only relying on previous activations, such real-time approaches may lack precision in recognizing some daily life activities [30]. To overcome this problem, it becomes necessary to know which sensor activations are generated later since the activity to be recognized will depend on the subsequent sensors.

1.2.3 Diversity and frequency of human activity

Accurately modeling human activities is difficult due to the complex and varying nature of human activities [32]. The diversity in human activities regarding duration, interactions with the smart home environment and the differences in the order of the activities of different elderly will make the problem even more complicated [33]. For instance, an activity like preparing breakfast consists of many actions such as turning on the coffee maker, turning on the toaster and getting cheese out of the fridge. The order of these actions may vary for different occasions of the same activity or some of the steps may completely disappear.

Not only are human activities highly diverse in the form of different sensor activations but the frequency of activities themselves are inherently imbalanced and hence accurate AR is challenging from a machine learning perspective [34]. Large differences in the number of examples for the classes to learn can make the machine learning algorithm emphasize learning majority classes and thereby partially or completely ignore minority classes [25]. As an example, cooking may occur with a higher frequency than grooming. Another more prominent example is the vast difference between the total samples of eating and sleeping activities. The latter occurs with a much higher frequency in data sets collected over a long duration.

1.2.4 Number of activities

A large number and different activities are performed in the daily routine. Hence, AR systems from smart homes should be able to recognize a different and large set of human activities [35]. However, recognizing a small set of daily life activities by AR systems is normally more straightforward than recognizing a large set of human activities. The reason for this could be attributed to the fact that as the number of human activities increases, the classifier has to discriminate among a larger set of human activities, which may add further difficulty for the classifier to correctly distinguish among activities [25].

1.2.5 Types of activities

Activities which are highly similar such as snack with lunch, dinner or breakfast are very hard to discriminate as they overlap significantly in the feature space. However, these kitchen activities are very different from bedroom activities such as sleeping and living-room activity such as spare time -tv [22]. Furthermore, recognizing a large number of human activities having both very similar and different characteristics at the same time from smart home environments makes the recognition problem even harder by AR systems. Moreover, discriminating overlapped activities is also a challenging task for an AR system [32]. For instance, while a person is preparing lunch in the kitchen the phone rings, the person will stop cooking for a while until the person finishes the call. Performing more than one activity at the same time by a resident in a smart home such as watching TV and preparing a snack is another problem of AR systems.

1.2.6 Sensor challenges

Sensors are fundamental components of a smart home environment to make critical infrastructure monitoring human behavior systems. Sensors play a crucial role in supporting the reliability of continuous activity monitoring and the safety of the smart home system. However, in smart homes, variability in sensor characteristics is one of the practical challenges of implementing AR due to external and internal factors of sensors. External factors that can create variability in sensors may include loose straps or changes in the operating temperature, while internal factors comprise sensor drift, hardware errors, or complete sensor failures [25]. A failure in one of the sensors in a smart home can cause misleading results of AR and downgrading the performance of a smart home system [36]. In a remote healthcare monitoring system, this could make dramatic consequences on the health of the resident. Moreover, the power supply of sensors is a bottleneck, which is a problem in demanding long-term and continuous monitoring of residential behaviors. Therefore the activation of sensors in a smart home environment influences the sensor power consumption (battery lifetime) which is a problem. Replacing batteries on devices and sensors is a nontrivial task since sensors are often embedded into the objects of the smart home such as bed, and door frames [37].

1.3 Research question and summary of contributions

Considering the above challenges of human AR within smart home environments, this thesis addresses the following research questions.

i. ***How stable are low-dimensional maps of human activities in a smart home?***

ii. ***How AR could be improved at the expense of real-time recognition?***

iii. ***How to handle imbalanced class problems in the context of AR?***

These research questions were investigated and relevant contributions were made in the following sections and papers :

1.3.1 Paper 1: Stability analysis of the t-SNE algorithm for human activity pattern data

The t-distributed Stochastic Neighbor Embedding (t-SNE) mapping stability of human activity patterns in smart homes via the analysis of the reproducibility of low-dimensional manifolds is investigated. Importantly, we propose an extension of the Local Procrustes Analysis (LPA) technique to non-linearly align manifolds by using locally linear mappings which we called Normalized Local Procrustes Analysis (NLPA). Experiments show that our proposed technique offers a much better result compared to Procrustes Analysis. Stability investigation is a key step towards transferring knowledge across similar domains. Results from this work were presented and published at The 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC2018). This paper is in Appendix A.

I contributed to the study design, implemented the proposed method (NLPA), performed the experiments, and wrote the majority of the manuscript.

1.3.2 Paper 2: Efficient activity recognition in smart homes using delayed fuzzy temporal windows on binary sensors

Existing AR systems in smart homes have obtained encouraging results. Commonly these systems evaluate real-time recognition of human activities using only sensor activations that precede the evaluation time (where the decision is made). However, in many situations, such as diagnosing people with dementia. To improve performance, we propose a method that would delay the recognition process to include some sensor activations that occur after the point in time where the decision needs to be made.

For this, the proposed method uses multiple incremental fuzzy temporal windows to extract features from both preceding and oncoming sensor activations. The proposed method is evaluated with two temporal deep learning models: Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM), on a binary sensor dataset of real daily living activities. The experimental evaluation shows that the proposed method achieves significantly better results than the real-time approach and that the representation with fuzzy temporal windows enhances performance within Deep Learning models. This paper is published in the Journal of Biomedical and Health Informatics. This paper is in Appendix B.

I contributed to the conceptualization, design, and formulation, implemented the proposed approach using the two temporal models (LSTM and CNN). I performed the experiments and wrote the majority of the manuscript.

1.3.3 Paper 3: Efficacy of imbalanced data handling methods on deep learning for smart homes environments

This paper focuses on investigating the particularly problematic aspect of learning imbalanced activities. Human activity data sets are typically highly imbalanced because certain activities occur more frequently than others. Consequently, it is challenging to train classifiers from imbalanced human activity data sets. With the spread of deep learning methods in recent years, numerous deep learning-based recognition methods are also being explored to improve classification performance. Deep learning algorithms perform well on balanced data sets, yet their performance is not satisfactory on imbalanced data sets. Therefore, we aim to address the problem of class imbalance in deep learning for smart home data. We assess it with ADL recognition using binary sensors data set. This paper proposes a data level perspective, combined with a temporal window technique, to handle imbalanced human activities from smart homes to make the learning algorithms more sensitive to the minority class. The experimental results indicate that handling imbalanced human activities from the data level outperforms the algorithm level and improved the classification performance. This paper is published in the SN Computer Science Journal. This paper is in Appendix C.

I contributed to the conceptualization, design, and formulation, performed the experiments, and wrote the majority of the manuscript.

1.3.4 Potential questions for future research

How to transfer knowledge between smart homes with a different layout, sensor setting, and resident ? The aim of this research question is to exploit what has been learned in one smart home to improve generalization in different but related smart homes to reduce the need for labeling data.

Chapter 2
Survey of Human Activity Recognition Research

2.1 Motivation

To enhance the performance of AR systems and enable its wide applications including healthcare, smart home monitoring, and surveillance in real-world scenarios, different sensing technologies have been explored to conduct considerable research and proposing several approaches to model and recognize human activities. In the literature, pervasive computing technologies have been exploited to propose numerous approaches to devise effective AR systems. AR is a key component of smart home technology that makes independent living as a viable solution for people, and thus enhances and maintains the quality of life and care. In this chapter, AR based on a smart home setting is comprehensively reviewed. As opposed to AR based on smart home, AR based on Wearable sensors briefly reviewed. Therefore, the review is classified into AR based on smart environment and wearable sensors as shown in figure 2.1.

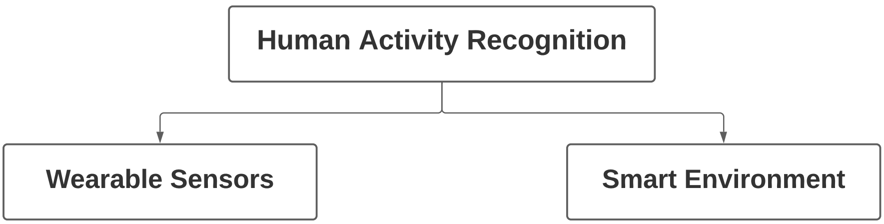


Figure 2.1: Taxonomy of Human Activity Recognition System

2.2 Wearable Sensors based Human Activity Recognition

The small-sized sensors that can be worn on the human body are referred to as wearable sensors. These types of sensors are mainly embedded in mobile devices, clothes, belts, wristwatches, glasses, or shoes to provide a continuous stream of information and to make them more comfortable to wear. Wearable sensors are used by people to generate more information about their interaction with their physical surroundings, motion, posture, and location. Wearable sensors including accelerometer, gyroscope, GPS, and RFID-readers (used together with RFID tags) have been used to collect information about users' movement e.g. walking, running, and sleeping. [38, 39, 40, 15].

With the rapid development of communication, design, and increasing process of the mobile devices, most of smartphones with embedded built-in sensors (accelerometers, GPS, and gyroscopes) can be used to recognize daily physical activities since they do not need any further equipment to collect and process data. However, solutions based on wearable sensors are not convenient and comfortable for users, and the success of the methods relies on users' involvement (e.g., wearing battery-powered sensors) [41]. In addition to the adversity or disaffection of wearing wearable sensors by people, there is no guarantee that the wearable devices such as bracelet sensors are continually worn. Commonly these devices are not very practical in real-world situations (e.g., elderly people may forget to wear the sensors or may not be able to wear the sensors). Moreover, AR systems based on wearable sensors are not well suited to recognize or distinguish a part of the daily activities that are characterized by the interaction of the user with several objects (e.g., sitting in the bathroom or the living room, sleeping in the bedroom or the living room).

Smartphones as part of the wearable sensors have emerged and have been widely used for human activity recognition with an increasing trend over the past few years [42, 43, 44]. Instead of attaching extra sensors on individuals, it is more practical to exploit the already embedded sensors from smartphones which we often carry around. Besides, smartphone devices can be placed on several parts of the human body, ranging from the lower such as leg or ankle to arm or wrist in the upper. Also, they are suitable for both outdoor and indoor settings. Furthermore, users can use more than one smartphone to record daily physical activities [45]. Smartphones contain different built-in sensing units such as accelerometers, gyroscopes, cameras, Global Positioning System (GPS) sensors. Among the sensors on a smartphone, the accelerometers are most often used for AR purposes. Mainly, smartphone devices have been employed in AR for sportive activities such as soccer, bicycling, nordic walking, rowing, and daily activities such as going to work. Besides, smartphone devices were used for detecting emergency and dangerous situations such as falls.

Recording activities-related raw data using these sensors has allowed researchers to use smartphones as an alternative and economic way for human AR [46]. However, orientation is one of the major problems of AR based on smartphones. Since smartphones can be carried by users in different positions such as in the bag or the pocket or their hands, accurate and robust AR becomes challenging even for simple activities such as walking.

These sensors can be classified into inertial (e.g. Accelerometers and Gyroscopes) and vital signs sensors (e.g. Bio-sensors) [47, 11]. Wearable inertial sensors are able to provide useful information about the user's movement and body posture. Vital signals that can be obtained from wearable bio-sensors such as skin temperature, heart rate, blood oxygen level, blood pressure play a crucial role in monitoring elderly people's health condition. Most of the commonly used wearable sensors that have been used for monitoring ambulatory activities and vital signs will be further discussed and summarized below.

2.2.1 Inertial Sensors

The most common inertial sensors that have been used for ambulatory activity monitoring are accelerometers that can be used to measure the value of acceleration along an axis. Accelerometers have been used for monitoring activities associated with body motion such as sitting, walking, walking downstairs and upstairs, standing, and doing exercise [48]. Accelerometers have been used for various purposes including detection of fall [49, 50], analysis of body motion and movement [51, 50, 52], early diagnosis of people with Parkinson's disease [53], and individual's postural orientation [54, 55, 56].

Accelerometers provide four attributes: time and acceleration along three axes. These attributes can present information to show human movement responding to tilt and frequency, which is important to assess the posture. Accelerometers can be embedded into belts wrist bands, bracelets, and watches due to their small size to monitor human activities. Accelerometers can send data wirelessly to mobile computing devices. From the collected data, it is possible to infer a context that can be used to monitor human activities for a long time and also detect an emergency such as fall detection [48, 57, 58, 59].

Accelerometers can be placed on the different parts of the human body to represent most human motions and obtain optimal performance of the human activity monitoring system [60]. Commonly wearable sensors are placed on waist [61], sternum [62] and lower back [63]. Figure 2.2 [1] shows different places of the human body that can wear sensors. Moreover, Table 2.1 shows recent studies about different placement of accelerometers.

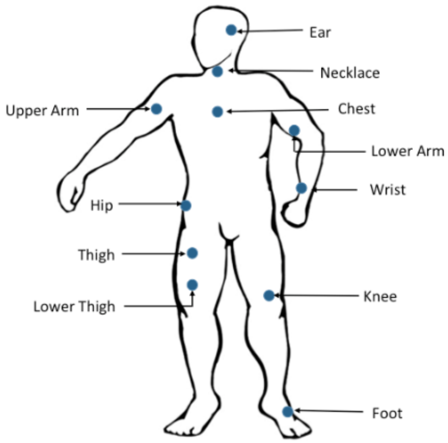


Figure 2.2: Graphical illustration of wearable sensor placement [1].

Table 2.1: Accelerometer placements on human body for activity monitoring

Reference	No.Accelerometers	Placements	Activities
Gjoreski and Gams, 2011 [64]	7	Chest, left thigh,right ankle	Standing,sitting,lying,goingdown,standing up,sitting on the ground, on all fours
Jiang et al., 2011 [65]	4	Left forearm, right forearm, left shank and right shank	Standing straight, sitting on a chair,lying on a bed, walking, jogging, cycling, walking on an elliptical machine, running on an elliptical machine, rowing and weight lifting
Jennifer et al., 2011 [66]	1	Smartphone	Walking, jogging, upstairs, downstairs, standing, sitting
Chun and Weihua, 2011 [67]	1	Right thigh	Sitting, standing, lying, walking, sit-to-stand, stand-to-sit, lie-to-sit, sit-to-lie
Siirtola and Roning, 2012 [68]	1	Smartphone placed in trousers' front pocket	Walking, cycling, sitting, standing, driving a car
Sweetlin 2013 [69]	1	Chest	Standing, walking, sitting, lying, fall
Mannini et al., 2013 [70]	1	Wrist/ankle	26 daily activities
Zheng 2008[71]	1	Wrist/hip/waist pocket	Lying, sitting, standing, walking, running, dancing, jogging, upstairs, downstairs, skipping
Lei et al., 2014[1]	4	Chest, left under-arm, waist and thigh	Lying, sitting, standing, flat walking and up and downstairs, lie-to-stand, stand-to-lie, sit-to-stand, stand-to-sit

2.3 Smart Home Based Human Activity Recognition

A smart environment is more suitable than wearable sensors for human AR when privacy and user acceptance are concerned due to its non-intrusive and non-wearable nature [72, 73]. A smart environment is designed to provide ambient assisted living (AAL) and can be used to detect multi-resident activities [22]. Smart environments are equipped with sensors that unobtrusively monitor and detect resident's interaction with objects.

During the past decades, the demand for using smart home setting technology to record and monitor human activities of elderly people has increased due to the potential supports that smart homes can provide to people [74]. Furthermore, the interest to support and assist elderly people over the age of 65 using smart homes is due to the predicted increase of 25% of the population by 2050 in the whole world. Besides, elderly people have been facing healthcare and social trend crises due to the prevalence of insufficient physical activity, cognitive decline, and sensory deficiency among them [75]. Healthcare and social resources are overstretched by these conditions which increase resource needs and investment. Therefore, It has become necessary for the healthcare systems to face the demographic changes of the elderly people to adapt and enable them to live in the best possible conditions [74]. Since the mobility and efficiency in the daily life of elderly people have typically decreased, the provision of care and assistance must be tailored based on elderly people's requirements and personalized to the individual. In addition to this, older adults may need frequent, urgent medical and care intervention to avoid fatal consequences [11]. Such emergencies could be alleviated by monitoring the physical activities of older adults in a continuous fashion. In many emergency cases, elderly people may require very expensive in-patient care if the hospital stay is prolonged, which may result in a serious financial burden on the patient [76]. New contributions to community healthcare and social care services are demanded since the requirement of having new models for healthcare and social delivery with affordable costs has increased. Recently, remotely monitoring health in a smart environment setting as one of the solutions has enabled the elderly to stay in their comfortable homes rather than in hospitals [77]. Figure 2.3 shows an example of a smart home equipped with many sensors.

Smart home environments are designed for intelligent service and to facilitate remote monitoring of residents. Smart homes have several key aspects. The heterogeneous devices can communicate in the network, manage and control the systems, sensors, and actuators that record information based on the interaction of the residents and objects. Consequently, due to the mentioned capabilities, human activity recognition has been studied based on smart home data for human activity recognition. Many smart home environment projects have been established in research labs that are listed and described in 2.3.1.

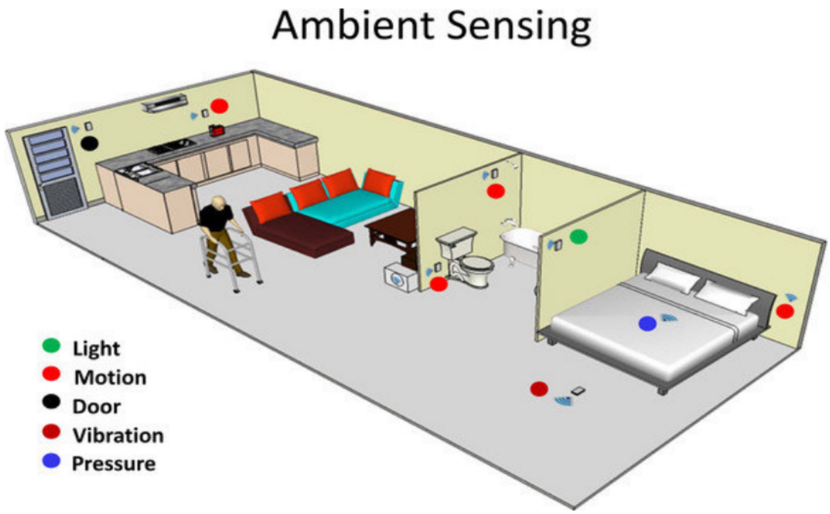


Figure 2.3: An example of smart homes [2].

2.3.1 Smart Home Projects

This section reviews the most notable smart home projects. The projects have been used to develop and evaluate technology that can be used to enhance assisted living for elderly people. Chronological order is determined for this presentation to provide a more transparent view of how the expectations and research issues associated with smart homes have evolved with time. At the beginning of opening these smart home projects, the main addressed problems were associated with the logical and physical connectivity of the sensors. Later the smart home projects evolved to multidisciplinary approaches particularly focusing on enhancing the usability of the interaction of sensor devices with the residents. The form of interaction from these smart homes demands technologies such as Artificial Intelligence (AI) that has a high grade of abstraction and encouraged the evolution to become more natural interaction approaches. Moreover, the technology of AI from these smart home projects enable independent living. A summary of the broadly reusable datasets obtained by the different smart home projects is also presented to make selecting between them easier for researchers.

- Halmstad Intelligent Home (HINT) [78] is a sensor-equipped home environment able to capture occupancy, movement, and interactions at the Halmstad University. HINT is a one-bedroom fully functional apartment of 50 m^2 con-

Figure 2.4: Halmstad Intelligent Home (HINT)

structed to give researchers, students and industrial partners with a technology-equipped realistic home environment. The layout of the HINT is shown in figure 2.4 and HINT facilitate experiments and studies within the areas of intelligent environments, Ambient Assisted Living (AAL), and social robots. Moreover, HINT is expected also to facilitate longitudinal studies by allowing residents to stay in the apartment for an extended period of time.

- GatorTech is a smart home project [3] built at the University of Florida, which is equipped with a wide range of sensors and devices to provide services such as activity recognition and tracking as well as voice recognition. The layout of the GatorTech is shown in figure 2.5. The purpose of this project is to create assistive environments such as homes that can sense and record the interaction of residents with the environment. In addition to this, the smart home can enact mappings between the real world and remote monitoring and intervention services. Besides, there are some early smart home projects including the following: PlaceLab [79] from the MIT, Adaptive Versatile home (MavHome)

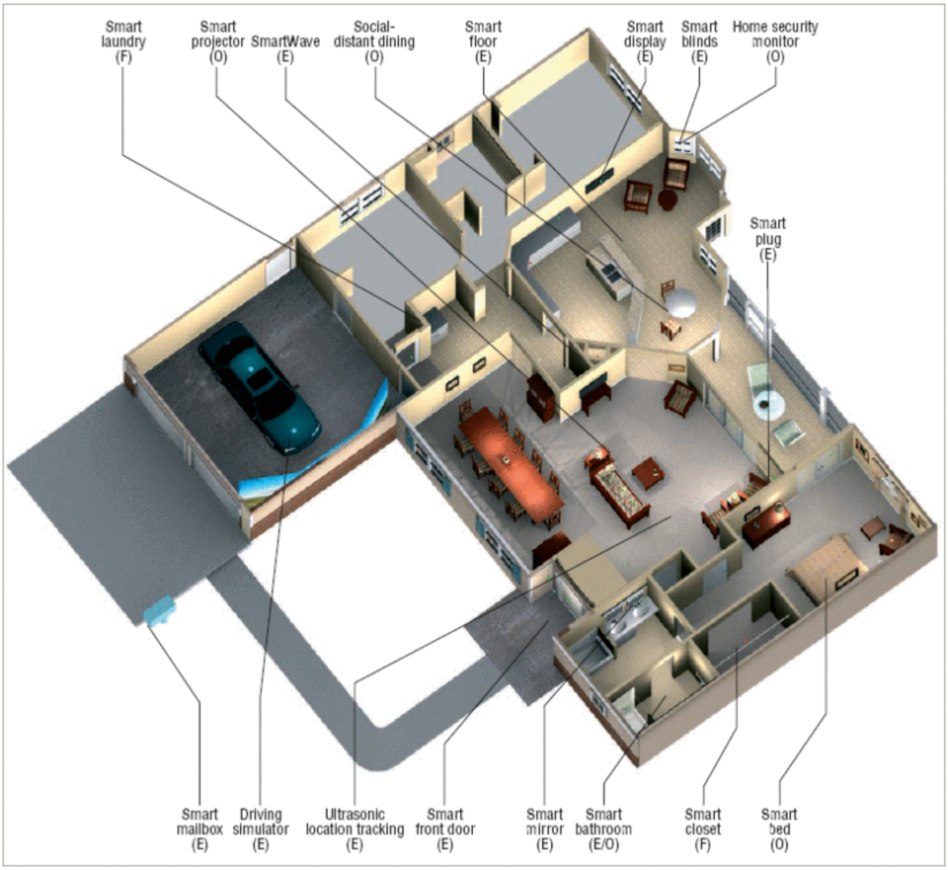


Figure 2.5: Gator tech smart house [3]

[80] from the University of Texas at Arlington and Intelligent System Lab (ISL) [81] from the University of Amsterdam.

- The Centre for Advanced Studies in Adaptive Systems (CASAS) is a smart home project that was developed at Washington State University in 2007. CASAS is a multi-disciplinary research project focused on building an intelligent home environment by using unobtrusive sensors and actuators [82]. CASAS could be used for research in several areas including machine learning, assistive technology, and activity recognition. The CASAS inventors have recently developed new research called "smart home in a box" [83]. The smart home in a box is a lightweight toolkit smart home layout that can be effortlessly installed to provide smart home capabilities out of the box with no customization or training needed. This toolkit was installed in 32 smart homes and generated many publicly available datasets. One of the capabilities is activity recognition which renders real-time activity labeling as sensor activations arrive in a stream and activity discovery for unlabeled data by employing unsupervised learning algorithms.

- SWEET-HOME [84] is a new smart home environment designed based on audio technology that is supported by a nationally French research project. This smart home project aims to have three main goals: proving an audio-based interaction technology that allows the residents to have full control over their smart home, detecting dangerous situations (e.g. fall at home) and detect deviations of behavior to enhance elderly-care alert systems, finally improving the social inclusion of the frail population and older people. The context-aware decision process is one of the interesting research paths from this smart home project, that employs a dedicated Markov Logic Network method to improve the experience to cope with uncertain events indicated from generated sensor data[85].

- In 2011, a smart home project "Unobtrusive Smart Environments for Independent Living" (USFIL) started which strives to provide affordable and excellent healthcare assistance utilizing the smart home system. Different types of sensors are installed from this smart home including a wrist wearable unit, camera, microphone, and Kinect sensor to recognize the basic daily physical activities (lying, sitting, walking, standing, cycling, running, ascending and descending stairs) of elderly people [86]. The goal of this project is to create applications addressing the gap between advanced technologies and the aging population.

- A smarter and safer home was proposed by [87] at CSIRO to improve elderly people's quality of life. To accomplish this, numerous smart home environments are installed in different locations to recognize human activities and movements. Based on this project, a smart assisted living was proposed by [88] to allow older adults to remain independent as long as possible in their homes. The sensors distributed in this smart home are expected to generate a continuous data stream to indicate the residence's movements. Extracting features and analyzing the generated data from the smart home using AI mechanisms are helpful to conduct diagnosis and decision making by health caregivers and clinical experts.

These smart home projects have produced many datasets some of which are publicly available and can be used for further studies by researchers. The smart home datasets used in most of the studies of activity recognition are reviewed in section 2.3.2.

2.3.2 Smart Home Datasets

In this section, many smart home datasets are reviewed which are mainly used for human AR. The first two datasets are used in this thesis. The other seven datasets can possibly be used in our future work. The long-term goal of our research is to explore transfer learning techniques over different smart homes.

i. Halmstad Intelligent Home [78] has a smart home dataset. The dataset was recorded by 11 residents from 37 binary sensors. Sensor events are represented by a particular ID of the activated sensor, the associated binary state, and a time-stamp of when the event occurred. This dataset has 8 activities: 1. go to bed 2. use bathroom 3. prepare breakfast 4. leave house 5. get cold drink 6. office7. get hot drink 8. prepare dinner.

ii. Two public smart home datasets ("OrdonezA" and "OrdonezB") is provided in which residents from both smart homes perform their daily living activities [89]. In OrdonezA, the resident performed 9 daily activities in 14 days over 19932 minutes from 12 binary sensors. In OrdonezB, the resident performed 10 daily activities in 22 days over 30495 minutes from 12 binary sensors. Both datasets are fully and manually labeled and details are shown in table 2.2.

Table 2.2: Details of recorded Ordonez datasets

	Home A	Home B
Setting	Home	Home
Rooms	4	5
Duration	14 days	21 days
Sensors	12	12
Number of Activities	9	10
Home setting	4 rooms house	5 rooms house
Number of days	14 days	21 days
Activities	Leaving,Toileting,Showering, Sleeping,Breakfast, Lunch,Dinner,SpareTime/TV, Snack, Grooming	Leaving,Toileting,Lunch, Showering,Sleeping, Snack, Breakfast, Dinner, SpareTime/TV, Grooming
Number of sensors	12 sensors	12 sensors
Sensors	PIR(Shower,Basin,Cooktop), Magnetic (Maindoor,Fridge, Cabinet, Cupboard), Flush(Toilet), Pressure(Seat, Bed), Electric(Microwave, Toaster)	PIR(Shower, Basin, Door Kitchen, Door Bathroom, Door Bedroom) Magnetic(Maindoor, Fridge, Cupboard) Flush(Toilet) Pressure(Seat, Bed) Electric(Microwave)

iii. A study conducted on older adults [90] based on collected data from passive sensors networks that were placed in 17 flats within an eldercare facility. Various sensors were employed to collect data such as motion sensors and pressure sensors for a long time (e.g. two years) in some of the smart homes.

iv. The Intelligent System Laboratory (ISL) [91] introduced datasets from three smart homes that were collected from an individual activity daily livings. The datasets were recorded for two months using 14, 23 and 21 sensors from the smart home A, B, and C respectively. The details of activities and sensors of these smart homes are illustrated from table 2.3.

Table 2.3: Details of recorded datasets of the ISL

	House A	House B	House C
Age	26	28	57
Gender	Male	Male	Male
Setting	Apartment	Apartment	House
Rooms	3	2	6
Duration	25 days	14 days	19 days
Sensors	14	23	21
Activities	10	13	16
Activities A	***Brush-Teeth, Drink, Snack, Go-to-Bed, Leave-house , Prepare-Breakfast, Prepare-Dinner, Shower, Use-Toilet***		
Activities B	***Brush-Teeth, Eat-Brunch, Eat-Dinner, Drink, Dressed Go-to-Bed, Leaving-house, Prepare-Brunch, Prepare-Dinner ,take-shower, Use-Toilet, Wash-Dishes***		
Activities C	***Eating, Brush-Teeth, Get-Dressed, Get-Drink, Get-Snack, Go-to-Bed, Leave-House, Prepare-Breakfast, Prepare-Dinner, Prepare-Lunch, Shave, Take-Medication, Take-Shower, Use-Toilet-Downstairs, Use-Toilet-Upstairs***		

v. Besides the given smart home datasets, [92] provided a dataset that contains gestures and acceleration data in addition to data from equipped sensors. [93] provided a public benchmark dataset and broadly used for activity recognition. It contains recorded data from nine inertial sensors attached to different parts of the body. 17 volunteers participated to record this motion dataset that related to 33 fitness activities.

vi. A real smart home dataset was published by [94] for complex scenarios of multi-residents which is called ARAS (Activity Recognition with Ambient

Sensing). The dataset was recorded from two smart homes for two residents' activities in each home for two months.

vii. ContextAct@A4H is a most recent and real-life daily living dataset in the Amiqual4Home smart apartment [95]. The smart home is equipped with different types of sensors to collect data. The dataset was recorded while a resident was living in the home during two periods in June and November (summer and fall respectively). The experiment for recording the ContextAct@A4H dataset was conducted in the frame of a collaboration between LIG, Universidad de Los Andes (Colombia), and Amiqual4Home. The inclusion of context variables (humidity, weather, noise, presence of visitors, temperature, etc) is one of the main contributions fo this dataset.

viii. CASAS project collected datasets from smart homes [82]. 19 datasets are collected from single-resident homes, 4 datasets are recorded from two-residents homes, the rest of the datasets are collected from larger families or residents with pets. Different types of binary sensors deployed from the CASAS smart homes including motion sensors and temperature sensors at various locations. The sensors were installed on multiple objects from the smart home such as doors, TV lounge, kitchen stove burners, bathroom, toilet, and bedrooms as well as other places in the home environment. These datasets could be used for human activity recognition [83]. The datasets were collected based on the generated sensor events for long durations of time while daily living activities are performed by an individual or multiple residents. Table 2.4 shows the characteristics of the CASAS datasets regarding activity count, sensor count, duration, and the number of residents.

Table 2.4: Characteristics of CASAS smart home datasets

Datasets	Activity count	Sensor count	Duration	Inhabitants
HH101	11	7		1
Tulum2010	16	37	149 days	2
Tworsummer2009	12	86	55 days	2
Tulum2009	10	20	83 days	2
Twor2009	13	71	46 days	2

ix. MIT announced two publicly available smart home datasets using a set of simple state-change sensors from two different homes that were collected for two different residents [79]. The datasets used different sensor types including those coming from motion, switch, and RFID sensors. It is one of the largest datasets collected from a real-world environment. Table 2.5 shows the number of performed activities, deployed sensors, and durations with residents' information.

The common activities of both datasets are "preparing breakfast", "preparing a snack", "toileting", "preparing dinner", "washing dishes", "preparing lunch".

Table 2.5: Characteristics of MIT smart home datasets

Datasets	Activity count	Sensor count	Duration	Inhabitants
Subject 1	28	28	16 days	30 year old woman
Subject 2	9	20	16 days	80 year old woman

2.3.3 Recent Surveys on Smart Homes

In this section, most of the recent survey and review papers on smart homes are reviewed and summarized. In 2012 Alam et al. [96] provided a review on smart homes and a comprehensive summary of prior developments, present situations, and also future challenges. In this review, smart home projects were reviewed according to three main desired services which are healthcare, security, and comfort as well as the objectives of the smart homes. Moreover, detailed information about all the requirements of smart homes is given such as sensors, algorithms, communications protocols, and also multimedia devices. Rashidi and Mihailidis, 2012 [97] conducted a survey on ambient assisted living technologies for elderly people. In this survey, ambient assisted living technologies are summarized regarding smart homes, wearable sensors, and assistive robotics. Besides, healthcare applications in this survey are also explored which focus on activity recognition algorithms and context modeling.

A review is presented by Salih and Abraham, 2013 [98] on ambient intelligence assisted healthcare monitoring and summarized the recent development of ambient assisted living for elderly people. The review focuses on the use of machine learning and data mining techniques to ambient assisted living using wearable sensors from smart homes for older adults and patients with chronic diseases.

Peetoom et al., 2015 [99] performed a systematic investigation on monitoring technology to activity recognition or a significant event such as falls at home or change in health status for older adults at home. The review aims to present the latest technologies that have been used in the smart home to monitor the activities of older adults. The review shows five main types of monitoring technologies: PIR motion sensors, body-worn sensors, pressure sensors, video monitoring, and sound recognition. This review also demonstrates the functionalities and the results of using these technologies to prolong the independent living of older adults. Positive effects as results of this review are suggested to both residents and caregivers to use monitoring technologies.

In 2013, a study on automated methods for real-time human ambulation, activity, and physiological monitoring was presented. The study aims to address the demands of assisted living, clinical observations, and rehabilitation, as well as the evaluation through sensor-based monitoring [100]. Three main areas of sensor-based monitoring systems are reviewed which are various types of sensors, frameworks and applications, data collection techniques, data processing and analysis, and research gaps, limitations, and challenges.

In 2018, Emiro, et al. [101] provided an exhaustive systematic review analysis of the sensor-based datasets used in human activity recognition. The contributions of the review are presented as follows. First, the most proper and well-known datasets of human activity recognition are identified with respect to the type of activities, information about residents and data acquisition devices. Secondly, analyzing the segmentation approaches used to extract features, the classification methods used for activity recognition based on the identified datasets. Furthermore, the number of samples of data used for training and testing phases as well as the metrics such as accuracy or f-measure for each dataset are presented. Lastly, several suggestions are provided regarding different techniques including segmentation, feature extraction, sampling and balancing of the datasets for the experimentation processes.

2.3.4 Environmental and unobtrusive sensing

Smart home Environments that are equipped with various non-invasive binary sensors can be used to unobtrusively monitor the interaction of a resident with the physical surrounding objects in the smart home and resident's movements [102]. The interactions are the Activities of daily life (ADL) which have been used to recognize activity or action. Most of the ADLs in smart homes are performed by a resident in a particular location with the specific object in the smart home [103]. For example, showering activity usually takes place in the bathroom, cooking activity usually takes place in the kitchen. Therefore, the activity could be recognized from the interaction between a resident and an object combined with environment observation. These sensors are employed to detect a resident's presence and interaction with the objects and furniture (e.g., dining table and chairs, bed, sofa) [104]. For example, if sensors indicate that the microwave or stove is on and the refrigerator or the cupboard is opened or there is water usage in the kitchen, this strongly shows that the activity of cooking is taking place. Therefore, it has been found the smart home sensors data could constitute important information to recognize human activities within smart home environments [105]. These sensors are being developed and becoming increasingly cheaper and smaller, more efficient and accurate, flexible and reliable, data-driven intelligent system, responsive and also increasing communication capability [106]. Due to these factors and also the availability of new technologies that contribute to the growth of sensor-based devices, the costs of sensors are being gradually reduced. The integration of networking technologies and ubiquitous sensing enables the evolution of many new applications in various fields such as s intelligent transport systems and smart home. The measurements generated by those sensors within the smart home are continuously transferred to the home gateway, in order to be used by activity monitoring systems [106].

2.3.5 Embedded Sensors

In this section, common embedded sensors used in smart homes for recognizing activity daily life will be summarized and also their advantages and disadvantages will be discussed. The most common environmental sensors are cheap devices and a binary sensor outputs 0 or 1. Binary sensors measure simple changes such as "door open" or "door closed" [107]. Simple binary sensors like motion sensors, state-change sensors, pressure sensors, and contact switches could be deployed on various objects in the environment to monitor residents' movement and locations [106].

- Motion sensors such as Passive InfraRed (PIR) are being broadly used [98, 29, 33, 72, 32] to detect and monitor the presence of the users in particular home locations (e.g., living, kitchen, or bathroom) and to track user's motion patterns. PIR sensors were used for stress monitoring and support to clinical decisions [108], and for moving targets detection and localization [107], and security [109].
- A simple state-change sensor has been employed for detecting any change of the state of an object which indicates the human-object interactions [110, 111, 112, 113, 114]. For example, a state-change sensor could be attached to the handset of a telephone in order to detect the human-object interaction when the handset will be lifted by a resident from the telephone base station in the smart home.
- Pressure sensors have been commonly used to detect when the resident lying on the bed or sits on a chair, sofa, or floor. The pressure sensor can be used to track the resident's movement and location [38, 115, 116, 117]. For example, a pressure sensor could be placed on the floor in front of the kitchen stove to detect cooking activity with the help of other sensors.
- Contact switch sensors have been mainly installed on the doors of rooms, fridge, and cabinets for detecting particular human-object interactions that the resident performs with these objects [118, 119, 120, 23, 121, 122, 123, 124]. Examples of such sensors are magnetic sensors that have been used to detect when drawers or doors are closed or opened [125].

In real-world scenarios, a single binary sensor is not sufficient for activity detection. Therefore, activity monitoring systems usually use multiple sensors such as motion detectors, break-beam sensors, pressure sensors, and contact switches to provide more information associated with activity monitoring. These sensors could be activated by gross movement, gross manipulation, point movement, and point manipulation in order to monitor and track resident's activities [126]. Ordonez et al.,2013 [89] studied activity recognition for seven common daily activities in a smart home. Activities are Leaving, Toileting, Shower, Sleeping, Breakfast, Dinner, and Drinking. Three different types of sensors were used to record activities: PIR sensors were used to detect motions in a particular area of the home, for example, kitchen area or living room, open or close states of cupboards and doors were detected by reed switch sensors, flushing toilet detected using float sensors. The binary sensors have been distinguished by these properties: low-cost, long-live and easy installation and replacing [11]. Furthermore, binary sensors can be used for unobtrusive human-object interaction recognition in a privacy-preserving way and also they require minimal computation resources to collect data. However, since the binary sensors provide very limited information, the main drawback is they may not be suitable for detecting activities of multi-resident from a smart home [11]. Radio-Frequency Identification (RFID) can be used based on a combination of a wearable sensor and smart home sensor technologies. Monitoring or human activity recognition systems based on sensor fusion that uses data from disparate sensors and sources leads to reduce uncertainty than using these different types of sensors individually. RFID comprises a reader that will be used by the resident and an electronic tag attached to a physical surrounding object in a smart home environment [127, 128, 129]. The tag will respond to a unique identifier when interrogated by a reader from the resident and then it will electronically be stored in memory [130, 131, 132, 133].

In a smart home environment, both active and passive RFID tags can be deployed to detect different types of activities. Fujinami et al.,2015 [134] performed research to track the long-term daily life activities of older adults with dementia using RFID tags attached to slippers from a set smart home in japan. Philipose et al.,2004 [122] developed an activity recognition system based on RFID tags attached to the physical objects in a smart home and presented activity information by a probabilistic sequence of the used physical objects. Kim et al.,2013 [135] adopted RFID technology to propose a real-time indoor ubiquitous-healthcare system (U-healthcare) in order to precisely locate and track elderly people. The U-healthcare system analyses the elderly's location based on the time slots and length of time that the elderly stay in the same place to recognize the user's movement and activities and determine elderly people's well-being. This shows that the advantages of RFID in healthcare systems particularly to track and monitor the elderly to stay safe and live independently at home [136]. Although the RFID has been used successfully in the aforementioned studies, RFID suffers from collisions in both the reader and tag when several RFID readers simultaneously interrogate the RFID tag. Thus, the RFID tag cannot distinguish among the unique identifier sent by the RFID reader.

2.3.6 Sensor Data Processing

Raw data provided by various sensors from smart homes need to be pre-processed for the later phases such as activity recognition modeling. This process plays an important role in achieving the research's target, output accuracy as well as recognition results. Existing research from smart home environments mainly focuses on learning different algorithms and reasoning approaches. Yet not sufficient researches published to pre-process smart home data which takes place between the raw sensors and activity recognition systems.

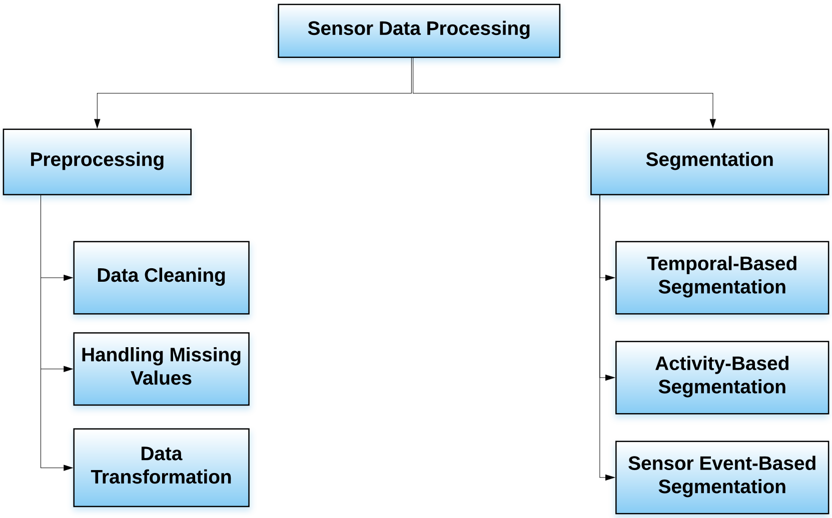


Figure 2.6: Sensor Data Processing

As a part of the activity recognition overview, processing sensor data from smart homes will be reviewed and discussed. Figure 2.6 shows sensor data processing formalization which are pre-processing and segmentation. These approaches will be further explained and analyzed to review the existing literature and extracting important information to make the processing sensors data of human activity recognition systems quite clear.

1. Data Pre-processing

   Collected data from different types of sensors in smart homes or wearable is inherently noisy. The operations of data preprocessing essentially involve data cleaning to remove irrelevant samples, data interpolation to cope with missing values in the dataset, and data transformation to build the proper data format.

   (a) Data Cleaning

   Raw data collected from fusion sensors contain redundant information, noise, and errors caused by batteries discharged, sensor failure detection, and sensor network failure. Data cleaning is required to compensate for these issues. Wilson and Atkeson [126] preprocessed four binary sensors and RFID sensor data using both Bayes and Particle filters. The outcomes of the Bayes filter work properly in a noisy environment on tracking an individual user, whereas the particle filter is better in scenarios with group users. Noury and Hadidi [2] removed nonlinear artifacts using a median filter and the redundant information was removed by the first-order-hold filter from the sensors data installed in various places in a smart home.

Guettari et al.,2014 [137] used a median filter to remove abnormal measurements that generated from passive presence sensors to detect human presence within a smart home environment.

(b) Handling Missing Values

Recorded data from sensors may contain missing values, particularly from RFID sensors [11]. Data interpolation can be used to solve missing value issues. Parlak and Marsic, 2013 [138] detected object motions for trauma resuscitation using passive RFID tags where an issue about irregular intervals tag readings appeared. The issue was generated by both signal transmission delay and irregularities in data arrival. Linear interpolation was used to solve the missing values within each window. Recorded data from sensors may contain missing values, particularly from RFID sensors [11]. Data interpolation can be used to solve missing value issues. Muaaz and Mayrhofer, 2013 [139] used mobile phones to measure human walking activity. The data is recorded using accelerometer sensors that cannot render data at equal intervals. Linear interpolation was conducted to reshape the data generated by the acceleration sensor into equal intervals.

(c) Data Transformation

For further analysis, data has to be prepared according to the system requirements. Rodner and Litz,2013 [140] applied association rule mining methods to model users' activities in smart homes. The recorded data of this study was generated by motion sensors and also integrated with a lux meter. The format of the data is (timestamp [numeric], motion [binominal] and lux [numeric]) including numeric values that could not be used in rule mining. Therefore, the numeric values are converted to nominal values in order to be used by the rule mining. Sun and Zhang,2014 [141] converted recorded electrocardiogram (ECG) data from analog signals to digital signals for further analysis. Moreover, normalization as a common data format transformation is used to obtain representation formal [142].

2. Data Segmentation

Data segmentation is a preprocessing technique to divide the raw data into small blocks of information since sensors data are normally generated as a continuous flow of raw data. The following three segmentation approaches are used in current literature.

(a) Temporal-Based Segmentation. Temporal based segmentation or sliding window segmentation is commonly used that divides raw sensor data into chunks of equal time duration. Ordóñez et al.,2013 [113] conducted a one-minute time interval to divide the raw data generated from binary sensors equipped in a smart home environment. This interval length is determined by considering proper labeling and activity discrimination. Selecting the optimal value for the time interval to segment raw sensors data from the smart home environment is highly critical [143, 144, 145, 146]. A short-time interval may generate duplicate activities, particularly for a long-duration activity such as sleeping, which leads to creating an imbalanced dataset. Besides, a long interval may combine several activities into the same segment resulting in losing important information [147, 148, 149]. Therefore, data segmentation requires effective heuristics for selecting an optimal value of the time-interval. Dynamic sliding windows are also used in several studies to segment raw sensor data based on activities or sensors ID [145, 146, 150]. Banos et al.,2014 [144] studied different window sizes with a non-overlap sliding window to segment sensor data of human activity recognition. The study revealed that the short window size normally leads to better recognition performance. However, based on human interpretation these approaches lack a longer temporal representation which is crucial for AR and has been recognized as a significant aspect of performance of sliding window approaches. Therefore multiple incremental fuzz temporal windows (FTWs) [151] is proposed to segment the timeline of human activities and to capture long-term and short-term activities. FTWs are compared with other approaches such as Equally Sized Temporal Windows(ESTWs), Raw and Last Activation (RLA), and Raw and Last Next Activation (RLNA) [151, 152].

(b) Activity-Based Segmentation

Segmentation of data based on activities consists of splitting the raw sensor data distinguishing the start time and end time of each activity. Selecting the correct boundary of the activities is the main issue of this approach [11]. A method to distinguish static activities such as sleeping from movement-related activities such as leaving home is proposed by Yoshizawa et al.,2013 [153]. The method detects the change points based on a threshold for the static activities and the starting time and ending time of the movement-related activities are determined by the analysis of variations in the frequency domain. This method is mainly possible in laboratory environments.

(c) Sensor Event-Based Segmentation

Segmentation of sensor data based on sensor events has been used to identify activities that have a sequence of movements, events, or actions that occur in specific time order or are interleaved with other activities' events, for example, cleaning or cooking. Different from the temporal-based segmentation approach, activities could not be split uniformly in the time since events happen sporadically, hence the size of the time windows is not fixed. Raw sensor data segmented based on the duration of each sensor event [154, 155, 156].

Techniques of sensor data processing are used in this research project to generate the input datasets from the raw sensor data for experiments. In this research project, multiple incremental FTWs is used to generating input datasets since temporal models i.e., LSTM and 1D CNN have achieved better performance based on FTWs compared to other techniques such as ESTWs, RLA, and RLNA for AR.

Chapter 3
Addressing activity recognition challenges

3.1 Motivation

The research project Situation Awareness for Ambient Assisted Living (SA3L) focused on the development of robust machine learning systems to improve the understanding of resident behaviour patterns, health and needs within sensor-equipped home environments. The long term goal of the systems is to process and share information across multiple smart homes to reduce the learning time and data collection as well as increase accuracy for applications such activity recognition. One solution through machine learning development is to use transfer learning to enhance systems ability. Knowledge transfer could be achieved by aligning learned manifolds to build a correspondence between different disparate data sets. Accordingly, knowledge transfer using machine learning(i.e. transfer learning), is gained during the transition from one learned domain to another, aiming to improve learning in the target task by leveraging knowledge from the source task [163]. In our work, it is hypothesized that learned manifolds from disparate data sets could be used for transfer learning. Therefore, it is crucial to investigate the stability of t-SNE maps in order to properly align manifolds for the purpose of transfer learning. Therefore, the first contribution of this thesis is investigating stability of t-SNE maps, which is described in section 3.2. The second contribution is proposing an efficient AR method in smart homes to further improve AR by including feature sensor readings in addition to preceding sensor reading, which is described in section 3.4. The third contribution is investigating handling imbalanced class problems in AR from data-level and algorithms level. We propose a data level prospective combined with a temporal window technique to handle imbalanced data in AR, which is described in section 3.5.

I contributed to the conceptualization, design and formulation, and methodology of the papers that are presented in the next sections. I implemented the proposed methods and performed experiments, validation, formal analysis as well as investigation. I wrote the majority of the manuscripts, reviewing and editing the manuscripts.

3.2 Paper 1: Stability analysis of the t-SNE algorithm for human activity pattern data.

This paper was presented and published in the 2018 IEEE international conference on Systems, Man, and Cybernetics (SMC2018)

3.2.1 Background

Exploratory data analysis based on all available dimensions of high dimensional datasets (HDD) is generally intractable to predict future insight and to make valuable decisions. Dimensionality reduction (DR) methods are used to derive the underlying manifold structure of datasets with low intrinsic dimensionality that could be considered as the minimum number of parameters needed for explaining the observed properties of the data. Furthermore, the global and local structure of the HDD in the lower-dimensional representative space could be preserved by such parameters [157].

Methods that linearly map HDD into a lower-dimensional representation such as Principal Component Analysis (PCA) [158] and multidimensional scaling (MDS) [159] are widely used particularly in business and marketing applications for DR and data visualization. PCA and MDS reduce the dimensions of HDD by computing low-dimensional maps where dissimilar data points are far apart. However, presenting similar data points close together in a low-dimensional map is crucial for HDD that lie on or near a low dimensional manifold, which is often difficult using linear mapping techniques [160]. Hence, the linear techniques for DR such as PCA and MDS are often not recommended for many complex and non-linear datasets [161]. Instead, a common non-linear DR technique such as the t-SNE [160] is used to render low dimensional representations of the high-dimensional input data [162, 163] presumably close to the sought real low-dimensional manifold. In contrast to PCA and MDS, t-SNE as a non-linear DR method can cope with complex data sets that are likely to lie on a low-dimensional non-linear manifold [164].

The primary problem studied in this work is how to analyze the stability of the t-SNE algorithm output. The proposed approach utilizes comparisons of several output maps as a whole and partially by clustered low-dimensional data points. This is performed by using smart home data of human activity patterns as input data.

Learned manifolds can be aligned to build a correspondence between different disparate data sets and thereby provide knowledge transfer across the data sets from different domains [165]. Accordingly, knowledge transfer using machine learning (i.e. transfer learning), is gained during the transition from one learned domain to another, aiming to improve learning in the target task by leveraging knowledge from the source task [166]. In our work, it is hypothesized that aligning learned manifolds (by a data-driven model such as t-SNE) from disparate data sets could be used for transfer learning.

3.2.2 Contribution

To explore stability of human activities when observed in t-SNE low dimensional maps, Normalized Local Procrustes Analysis (NLPA) proposed to non-linearly align manifolds by using locally linear mappings. NLPA is an extension of Local Procrustes Analysis (LPA) [167] to compare locally aligned clusters alignments to a complete map, which can be compared to Procrustes Analysis (PA) [168] as well as being used for stability analysis. Results of NLPA outperforms the results obtained by PA. Algorithm 1 shows the NLPA procedure. NLPA applies PA on each cluster and then

Algorithm 1 Normalized Local Procrustes Analysis (NLPA)

1: **Input:** M_1, M_2 — M_1, M_2 are the input t-SNE maps
2: $n_c \leftarrow n$ — number of clusters created using LPA
3: **for** $i \leftarrow 0$ **to** n_c **do**
4: $\quad \{M_1, M_2, -, \text{norm}_1, \text{norm}_2, \mu M_1\} \leftarrow \text{PA} M_1, M_2$
5: $\quad M_1 \leftarrow M_1 \cdot \text{norm}_1\ \mu M_1$
6: $\quad M_2 \leftarrow M_2 \cdot \text{norm}_2\ \mu M_1$
7: $\quad \text{templist}_1 \leftarrow M_1$
8: $\quad \text{templist}_2 \leftarrow M_2$
9: **end for**
10: $\text{Map}_1 \leftarrow \text{templist}_1$
11: $\text{Map}_2 \leftarrow \text{templist}_2$
12: $\text{disparity} \leftarrow \sum_{i0}^{n} \left(\text{Map}_{1,i} - \text{Map}_{2,i}\right)^2$ — dissimilarity between the two sets
13: **Output:** disparity

normalizes the clusters so that the transformed data-points in each cluster are mapped back to the original space of the data after alignment. Normalization is done by multiplying the clusters with the norms of the aligned clusters that are produced by PA and then adding the mean of the first cluster, as shown in lines 5-6 of Algorithm 1. This normalization allocates the combined clusters of data-points to the same space as to original data. Finally, NLPA computes the disparity and estimated mean probability of obtaining the true corresponding data point within the aligned five nearest neighbors for the combined aligned clusters.

3.2.3 Results & Discussion

t-SNE stability has gained a lot of interest in projecting high-dimensional data into a low-dimensional manifold with the aim of transferring knowledge using manifold alignment. However, since t-SNE is a stochastic algorithm and since there is a large variance of t-SNE maps, a thorough analysis of the stability is required before applying transfer learning. Exhaustive scenarios are considered for an investigation about the t-SNE stability through manifold alignment using PA and NLPA. Table 3.1 shows the estimated disparity and probability of obtaining the correct correspondence observation within five nearest neighbors of the PA and the NLPA methods respectively for different values of perplexity. It turns out that for all perplexity values considered the disparity values from using NLPA are less than 4% of the corresponding disparity value from using PA. In other words, the NLPA method is 25 times better than the PA method in terms of disparity. Also, the disparity from using PA decreases slightly for perplexity ranging from 5 to 20 while it increases for perplexity values from 25 to 50. The disparity values from using NLPA increases almost monotonically for all perplexity values considered.

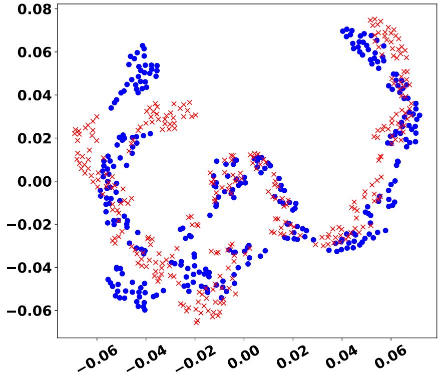
With respect to the probability of obtaining the correct correspondence within the five nearest neighbors, the PA correct correspondence mean values commonly increase by increasing perplexity, especially as the perplexity value reaches 25. On the other hand, the probability of obtaining correspondence within the five nearest neighbors slightly decreases by increasing perplexity ranging from 5 to 50 for NLPA.

Figure 3.1 shows the alignment of t-SNE maps that has been used to investigate the t-SNE stability using PA and NLPA. The alignment maps in Figure 3.1 indicate that the t-SNE maps are stable locally compared to globally aligned t-SNE maps using PA.

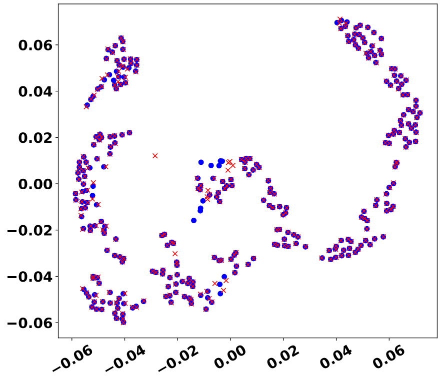
Table 3.1: Estimated expected disparity and estimated probability of obtaining the correspondence observation within five nearest neighbors.

Perplexity	**Disparity**		**Probability of obtaining the correspondence within the 5 nearest neighbors %**	
	Mean (SE)			
	PA	***NLPA***	***PA***	***NLPA***
5	0.5954 (0.0206)	0.0010 (0.0002)	7.363 (0.0074)	98.541 (0.0099)
10	0.3135 (0.0248)	0.0006 (0.0001)	24.527 (0.0226)	98.541 (0.0099)
15	0.1097 (0.0167)	0.0007 (0.0001)	44.618 (0.0235)	98.476 (0.0099)
20	0.0291 (0.0022)	0.0006 (0.0001)	60.365 (0.0163)	98.450 (0.0099)
25	0.0296 (0.0026)	0.0007 (0.0001)	63.873 (0.0189)	98.485 (0.0099)
30	0.0447 (0.0058)	0.0009 (0.0002)	56.693 (0.0261)	98.435 (0.0099)
35	0.1483 (0.0152)	0.0029 (0.0005)	28.216 (0.0241)	98.126 (0.0098)
40	0.1648 (0.0129)	0.0046 (0.0006)	26.186 (0.0223)	97.869 (0.0098)
45	0.1699 (0.0123)	0.0049 (0.0006)	31.522 (0.0256)	97.637 (0.0098)
50	0.1501 (0.0116)	0.0056 (0.0006)	38.485 (0.0296)	97.744 (0.0098)

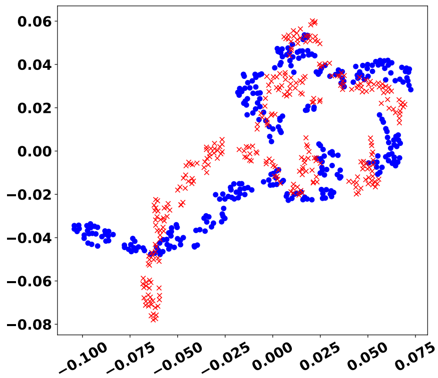
Despite showing improvements to align manifolds using NLPA over PA, using only one datasets is a limitation in this study that could be addressed in future research. Future work will explore extensions of NLPA for aligning low-dimensional manifolds of disparate data sets. Then t-SNE low-dimensional manifolds of disparate data sets will be compared using NLPA to discover the common manifolds of the disparate data sets to be used for transfer learning.



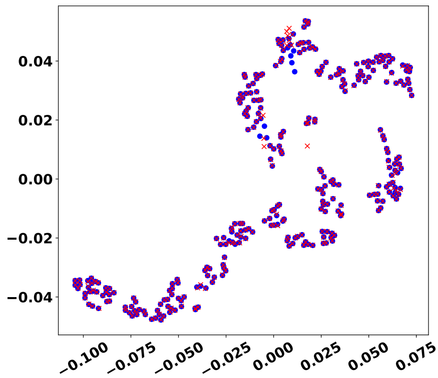
(a) PA for perplexity 20



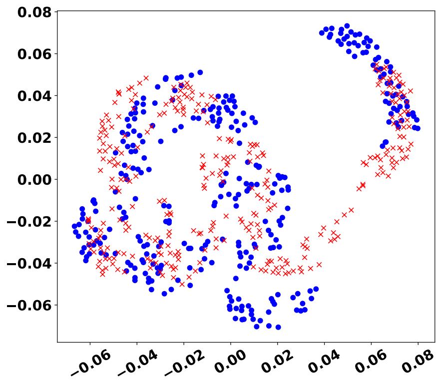
(b) NLPA for perplexity 20



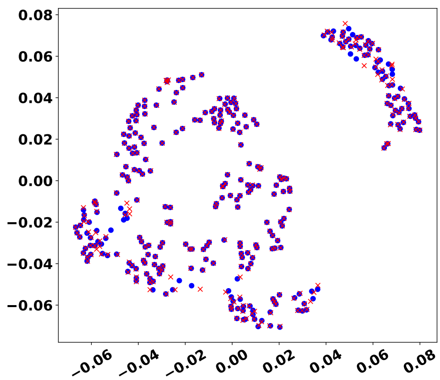
(c) PA for perplexity 15



(d) NLPA for perplexity 15



(e) PA for perplexity 35



(f) NLPA for perplexity 35

Figure 3.1: Manifold alignment using PA and NLPA for three cases

3.3 Paper 2: Efficient activity recognition in smart homes using delayed fuzzy temporal windows on binary sensors.

This paper is published in IEEE Journal of Biomedical and Health Informatics.

3.3.1 Background

Existing approaches for activity recognition in smart homes perform predictions in real time based on sensor activations that precede the evaluation time. Due to only relying on previous activations, such real-time approaches may lack precision in recognizing some daily life activities. To overcome this problem, it becomes necessary to know which sensor activations are generated later since the activity to be recognized will depend on the subsequent sensors.

Although real-time approaches enable permanent interaction with users in smart environments, it has a lower recognition rate [169, 170]. This is not admissible in some critical cases such as diagnosing dementia, as it requires more accurate activity recognition to correctly detect abnormal behaviours in the inhabitant, [171]. These cases, although less frequent, prevent AR from being a high-precision tool for assessing the conditions of inhabitants in smart homes.

The contribution of this paper takes the above-mentioned issues into consideration. We propose a data-driven approach that aims to increase precision and sensitivity in daily activity recognition by means of i) delaying the activity recognition, ii) extracting representations of binary sensor activations that occur before and after the time where the prediction is made, iii) evaluating Deep Learning methods for classification, and iv) analyzing the impact of the delayed AR process on precision and sensitivity.

3.3.2 Contribution

The contribution of the study is to propose a method that delays the recognition process and includes sensor activations that occur after the point in time where the decision is made. For this, the proposed method uses multiple incremental fuzzy temporal windows (FTWs) to extract features from both preceding and partial oncoming sensor activations. In order to avoid the human configuration of FTWs we have modeled their shapes with the Fibonacci sequence, which has been defined to model incremental sequences in a harmonic way under the fields of mathematics, science, and engineering [172]. The proposed method is evaluated with three temporal deep learning models (CNN and LSTM Network as well as hybrid models combining CNN and LSTM), on a binary sensor dataset of real daily living activities. The experimental evaluation shows that the proposed method achieves significantly better results than the real-time approach and that the representation with fuzzy temporal windows enhances performance within Deep Learning models. Analyzing the impact of delaying the activity recognition by using oncoming sensor activations in real daily activities. Specifically, showing the impact of including oncoming sensor activations in activity recognition to improve the recognition of some rare activities such as *leaving, snack, grooming, and dinner* that have been poorly recognized using only real-time activity recognition.

3.3.3 Results & discussion

In a real-time activity recognition scenario, only past sensor data are considered for each evaluation time. Evaluation time is the time when the decision-making of each activity has been developed. To predict what activity has been performed in a specific time *T*, different time delays of oncoming sensors after the time *T* in addition to the preceding sensor activations of the time *T* are included in the feature extraction based on FTWs. For example, now (Evaluation time), we evaluate what activity was developed 4 hours ago (Evaluated time). Evaluated time. It is the time which is evaluated by the classifier to recognize which activity has been developed in this point of time based on the preceding and oncoming sensor data. In the case of real time, evaluation time is equal to the evaluated time. In the case of delays in time that consider oncoming sensor data, evaluation time is higher (delayed) from evaluated time.

In the scenario where the AR is delayed, preceding sensor activations with different time delays, particularly 5 minutes, 20 minutes, 1 hour, and 4 hours, are tested to improve the recognition process. The results show that delaying time with LSTM in decision-making leads to building more accurate models. The results are significantly improved when considering oncoming sensor activations and increasing delays in the evaluated time. In house A, for example, the total results of the F1-score of the model in real-time is 89.05, while the results of the model are improved notably, up to 96.44, when considering oncoming sensor activations. Tables 3.2 and 3.3 show the results of the F1-score and training time of LSTM, CNN, and the hybrid CNN LSTM based on FTWs from home A and B respectively. The results indicate that the F1-score of the models improves substantially by increasing time delays with a slight increase of training time. This means that delaying the decision-making of human activity recognition yields better and more accurate models.

In summary, the proposed method of this study has enhanced the models for recognizing all the activities performed in homes A and B while maintaining a low time cost. We highlight that the proposed model with FTWs and Deep Learning achieves encouraging performance particularly in the activities that real-time models have difficulties recognizing accurately, such as *Leaving*, *Snack*, *Grooming*, and *Toileting* from home A. Regarding home B, the results of the same activities in addition to *Dinner* are significantly improved. This refers to the fact that taking the oncoming sensor activations into account is important in order to enhance the learning process of the models.

Table 3.2: F1-score and training time (minutes) of LSTM, CNN, and CNN LSTM with different delays in time based on FTWs from Ordonez Home A (m=minutes, h=hour)

Activity	LSTM				CNN				CNN LSTM			
	5 m delay	**20 m delay**	**1 h delay**	**4 h delay**	**5 m delay**	**20 m delay**	**1 h delay**	**4 h delay**	**5 m delay**	**20 m delay**	**1 h delay**	**4 h delay**
Breakfast	95.18	95.87	96.87	**96.95**	92.91	93.24	93.29	**94.69**	95.22	95.87	96.78	**97.99**
Grooming	66.85	67.96	71.24	**86.19**	66.71	71.28	76.81	**83.68**	72.76	75.28	81.26	**86.19**
Leaving	91.17	91.85	94.15	**99.85**	94.31	96.71	96.83	**99.49**	97.41	98.21	98.78	**99.85**
Lunch	96.88	97.96	97.82	**99.00**	95.76	96.73	97.22	**98.92**	96.11	96.42	96.57	**99.07**
Showering	95.21	95.53	96.89	**98.21**	93.23	93.53	95.11	**96.49**	94.52	95.62	95.79	**99.98**
Sleeping	99.68	99.72	99.33	**99.85**	95.42	96.38	96.87	**99.73**	99.73	99.57	99.67	**99.85**
Snack	93.72	94.99	96.82	**99.31**	89.62	92.21	95.14	**98.42**	97.99	98.36	98.72	**99.34**
Spare Time	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Toileting	73.52	73.68	74.73	**90.76**	65.23	67.27	71.62	**86.97**	78.75	82.51	82.98	**90.76**
Total	**90.24**	**90.84**	**91.97**	**96.44**	**88.13**	**89.70**	**91.43**	**95.37**	**92.49**	**93.53**	**94.52**	**96.97**
Train-time	3.91	4.1	4.6	5.1	2.3	2.7	3.1	4.01	3.2	3.5	3.9	4.6

Table 3.3: F1-score and training time (minutes) of LSTM, CNN, and CNN LSTM with different delays in time from Ordonez Home B (m=minutes, h=hour)

Activity	LSTM				CNN				CNN LSTM			
	5 m delay	**20 m delay**	**1 h delay**	**4 h delay**	**5 m delay**	**20 m delay**	**1 h delay**	**4 h delay**	**5 m delay**	**20 m delay**	**1 h delay**	**4 h delay**
Breakfast	93.31	94.11	94.81	**99.69**	91.27	91.97	**95.65**	99.25	94.33	94.87	96.43	**99.54**
Grooming	78.37	78.89	79.23	**91.97**	73.14	76.36	82.85	**87.07**	78.91	83.22	85.23	**90.67**
Leaving	93.62	98.72	98.72	**99.50**	94.28	98.51	98.63	**99.39**	98.84	98.86	98.89	**99.55**
Lunch	95.41	95.54	96.22	**99.02**	92.84	93.14	95.87	**98.43**	96.96	96.94	97.12	**99.14**
Showering	89.42	90.1	90.32	**98.85**	87.12	87.98	92.65	**99.42**	89.31	94.84	96.42	**99.42**
Sleeping	99.41	99.48	99.67	**99.80**	99.48	99.51	99.63	**99.68**	99.64	99.65	99.68	**99.73**
Snack	84.67	85.42	85.67	**96.39**	81.57	83.63	89.71	**94.49**	86.53	88.94	92.74	**97.63**
Spare Time	96.98	96.93	96.98	**99.33**	97.14	97.52	97.58	**98.95**	96.92	97.12	97.54	**99.29**
Toileting	62.37	66.37	67.12	**86.23**	62.73	65.72	69.43	**76.17**	67.63	72.78	76.68	**85.19**
Dinner	81.24	84.24	86.24	**97.39**	79.16	83.16	88.83	**95.62**	78.83	83.74	87.68	**97.03**
Total	**87.48**	**88.98**	**89.49**	**96.82**	**85.87**	**87.75**	**91.05**	**94.84**	**88.79**	**91.09**	**92.84**	**96.72**
Train-time	4.9	5.23	5.41	5.93	3.17	3.33	3.84	4.31	4.36	4.51	4.82	5.71

3.4 Paper 3: Efficacy of Imbalanced Data Handling Methods on Deep Learning for Smart Homes Environments.

This paper is published in SN Computer Science Journal.

3.4.1 Background and Contribution

Human activities are highly diverse not only in the form of different sensor activations but the frequency of activities themselves are inherently imbalanced and hence accurate AR is challenging from a machine learning perspective. Applying a machine learning model on an imbalanced dataset, it tends to partially or completely ignore the minority classes. As an example cooking may occur with a higher frequency than grooming. Another example is the vast difference in the number of examples between eating and sleeping where the latter occurs with a much higher frequency in data sets collected over a long duration. This study focuses on investigating the particularly problematic aspect of learning activities over days or even months which are imbalanced.

Despite many past efforts of research on the class imbalance problem and approaches to cope with this general problem, there is a lack of empirical work on targeting machine learning beyond shallow methods [173]. Traditional machine learning algorithms such as decision tree, support vector machine, naive Bayes, and hidden Markov models have been used to minimize the recognition error [174, 175]. Satisfying recognition results have been achieved by adopting these approaches. However, such algorithms may heavily depend on classical heuristic and hand-crafted feature extraction which might be limited by human domain knowledge [176]. A natural variation within each activity is often present in collected smart home datasets and is not unlikely to fluctuate even more between different residents. These variations are also influenced by contextual factors such as time of the day and location of where the activity is performed. Given these conditions as well as considering the multitude of choices at sensor installation (e.g. sensor types and sensor locations) AR based on shallow learning where features are hand-crafted can be challenging. Therefore, discovering more systematic methods to obtain features has drawn increasing research interests [177]. The influence of deep learning has been demonstrated in many areas not only in image classification such as speech recognition, natural language processing as surveyed in [176]. Consequently, studies of activity recognition using deep learning have multiplied because the number of elderly smart-home health-care services has steadily increased for the last few years and all reporting state of the art performances achieved on diverse activity recognition benchmark datasets [178, 179]. Especially two methods have brought promising results of AR, LSTM and CNN when using data prepared with a Fuzzy-based approach to represent temporal components of the data [151, 180]. However, these two machine learning algorithms for AR have not been studied from the context of different temporal preprocessing methods along with traditional methods for handling class imbalance in order to improve recognition accuracy.

The study described in this paper is therefore designed to fill parts of such a knowledge gap and also put a particular focus on the classes representing activities with a relatively low number of observations (i.e., minority classes). Thus, the main contribution of this paper is the study of well-known class imbalance approaches (Synthetic Minority Oversampling Technique, Cost-Sensitive learning, and Ensemble learning) applied to activity recognition data with various temporal data preprocessing for the deep learning models LSTM and 1D CNN.

3.4.2 Results & discussion

The results of the experiments using LSTM and CNN are presented and discussed in the aspect of different methods of handling imbalanced classes and different feature extraction approaches. FTWs and ESTWs are used to pre-process data and build the datasets for training. SMOTE, Cost-sensitive and Ensemble learning methods are used for handling the class imbalance present in the datasets. Table 3.4 shows the results of the F1-score of the LSTM and CNN models from the home A for the imbalanced dataset, with Cost-Sensitive corrections and minority sampling using SMOTE. The F1-score of the minority classes which are *Breakfast, Grooming, Lunch, Showering, Toileting*, and *Snack* from the home A are improved using SMOTE based on both approaches of extracting features and both models.

Table 3.4: F1-score Home A

	FTWs							ESTWs						
	Imbalanced data		Cost-Sensitive		SMOTE		Ensemble	Imbalanced data		Cost-Sensitive		SMOTE		Ensemble
Activity	**CNN**	**LSTM**	**CNN**	**LSTM**	**CNN**	**LSTM**		**CNN**	**LSTM**	**CNN**	**LSTM**	**CNN**	**LSTM**	
Snack	0.00	0.00	0.00	0.00	0.28	0.39	0.00	0.00	0.00	0.00	0.00	0.27	0.42	0.01
Showering	0.36	0.48	0.43	0.47	0.70	0.70	0.51	0.79	0.81	0.82	0.81	0.89	f0.89	0.82
Grooming	0.00	0.00	0.00	0.00	0.25	0.28	0.12	0.55	0.53	0.54	0.55	0.56	0.57	0.57
Breakfast	0.61	0.67	0.65	0.68	0.71	0.73	0.38	0.71	0.72	0.76	0.74	0.73	0.77	0.67
Toileting	0.00	0.00	0.00	0.00	0.31	0.37	0.17	0.00	0.00	0.00	0.00	0.28	0.29	0.17
Lunch	0.75	0.80	0.81	0.82	0.80	0.84	0.64	0.81	0.80	0.82	0.85	0.86	0.86	0.81
Leaving	0.76	0.86	0.75	0.83	0.88	0.89	0.83	0.85	0.86	0.86	0.86	0.87	0.87	0.84
Sleeping	0.96	0.96	0.96	0.96	0.92	0.90	0.92	0.99	0.99	0.99	0.99	0.99	0.99	0.99
Spare Time	0.91	0.91	0.90	0.91	0.92	0.93	0.76	0.98	0.98	0.98	0.98	0.99	0.99	0.98
Average	**0.44**	**0.48**	**0.46**	**0.47**	**0.63**	**0.67**	**0.48**	**0.60**	**0.62**	**0.62**	**0.63**	**0.71**	**0.73**	**0.65**

Table 3.5: F1-score Home B

	FTWs							ESTWs						
	Imbalanced data		Cost-Sensitive		SMOTE		Ensemble	Imbalanced data		Cost-Sensitive		SMOTE		Ensemble
Activity	**CNN**	**LSTM**	**CNN**	**LSTM**	**CNN**	**LSTM**		**CNN**	**LSTM**	**CNN**	**LSTM**	**CNN**	**LSTM**	
Dinner	0.00	0.00	0.00	0.00	0.31	0.34	0.06	0.00	0.01	0.00	0.00	0.26	0.27	0.13
Snack	0.00	0.00	0.02	0.08	0.27	0.29	0.22	0.00	0.00	0.00	0.00	0.26	0.28	0.07
Showering	0.0 0	0.22	0.00	0.21	0.26	0.36	0.24	0.73	0.80	0.71	0.79	0.82	0.84	0.53
Grooming	0.13	0.30	0.09	0.30	0.39	0.36	0.42	0.62	0.61	0.61	0.61	0.64	0.65	0.54
Breakfast	0.50	0.47	0.51	0.51	0.52	0.58	0.36	0.26	0.23	0.24	0.19	0.30	0.35	0.29
Toileting	0.00	0.00	0.00	0.00	0.31	0.32	0.32	0.23	0.04	0.23	0.10	0.26	0.27	0.14
Lunch	0.39	0.35	0.31	0.38	0.41	0.42	0.37	0.00	0.00	0.00	0.00	0.36	0.38	0.00
Leaving	0.90	0.90	0.89	0.89	0.90	0.90	0.84	0.66	0.66	0.66	0.66	0.66	0.66	0.66
Sleeping	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97
Spare Time	0.83	0.82	0.84	0.84	0.85	0.86	0.79	0.90	0.90	0.90	0.90	0.89	0.90	0.90
Average	**0.33**	**0.36**	**0.36**	**0.41**	**0.51**	**0.54**	**0.45**	**0.40**	**0.40**	**0.40**	**0.40**	**0.54**	**0.56**	**0.42**

The results also show the majority classes which are *Leaving* and *Spare-Time* activities (except Sleeping) which are also improved based on both approaches of extracting features for both models using the SMOTE method. The average results of the LSTM and CNN for all activities are improved using the SMOTE method based on both FTWs and ESTWs. Regarding home B, the F1-score of the minority classes (*Breakfast, Grooming, Lunch, Showering, Toileting, Snack,* and *Dinner*) are considerably improved which are shown in the table 3.5. Moreover, only the results of the *Spare-Time* as the majority classes are improved based on FTWs. The average results of home B indicate that the SMOTE method substantially improved the recognition, particularly for the minority classes. The F1-scores in tables 3.4 and 3.5 indicate that the results of the models based on both feature extraction approaches using SMOTE are better (higher F1-score) than the results of models based on Cost-Sensitive and class imbalanced datasets. Moreover, the F1-score results based on SMOTE with ESTWs can be seen to be higher than F1-scores based on SMOTE with FTWs from both homes of both models on average. Moreover, the obtained results based on the SMOTE technique with both feature extraction method (FTW and ESTW) and with both temporal models (LSTM and CNN) are better than the results obtained by balanced ensemble learning as shown in tables 3.4 and 3.5. Therefore, the proposed data level solution (SMOTE and ESTWs) to handle imbalanced human activities from smart homes is more promising than algorithms level (Cost-sensitive and Ensemble learning).

Chapter 4
Conclusion and Future work

The t-SNE mapping stability of human activity patterns in smart homes via the analysis of reproducibility of low-dimensional manifolds is investigated. One could claim that any two data sets could be aligned via a non-linear mapping function with enough degrees of freedom. However, this study aims at analyzing parts of a map in order to investigate the stability of t-SNE. Therefore, the choice of linear and local transformations gives human intuition about the stability of t-SNE. Procrustes Analysis (PA) is used for linearly aligning low-dimensional manifolds in order to compute disparity and correct correspondence observation within the five nearest neighbors. An extension to Local Procrustes Analysis called Normalized Local Procrustes Analysis (NLPA) is proposed to non-linearly align manifolds by using locally linear mappings. Experiments show that the disparity from using NLPA decreases by magnitudes compared to the disparity from using PA. Also, the probabilities of obtaining the correct corresponding observation within the five nearest neighbors from the second set of data points for each point in the first set of data points are radically increased by using NLPA compared to PA. For instance when the t-SNE parameter is 20, the disparity mean value decreases from 0.2913 in the case of using PA to a mere 0.00066 upon using NLPA. The probability of obtaining the correct corresponding observation within the five nearest neighbors for the same comparison, increases from 60.37 when using PA to 98.45 in case of using NLPA. In conclusion, NLPA outperforms PA by providing much better alignments for the low-dimensional manifolds on the same data set. This indicates that t-SNE low-dimensional manifolds are locally stable which is the part of the achievements of this research project.

Since, human activity recognition is a highly dynamic and challenging research field that plays a crucial role in diverse applications such as health care, elderly care, emergencies, security, smart environments, surveillance and context-aware-systems, a novel method is proposed to improve understanding of human activity recognition. The method is a new data-driven approach that aims to increase precision and sensitivity for human activity recognition systems from smart home environment. The proposed method considers the partial oncoming sensor activations in addition to preceding sensor activations. With the use of oncoming sensor activation, we can take the benefits of enhancing the learning process that leads to improved recognition performance compared with the approaches using only the preceding sensor activations in the intelligent environment. Multiple and incremental fuzzy temporal windows were used to extract features from both preceding and partial oncoming sensor activations. Defining multiple and incremental fuzzy temporal windows from long-term to short-term has provided suitable semantics to determine a sequence of temporal features that boosts learning using LSTM sequence models and CNN.

Further, one of the main problem of home-bases systems activity recognition is the frequency and duration of human activities are intrinsically imbalanced. The huge difference in the number of observations for the classes to learn will make many machine learning algorithms to focus on the classification of the majority examples due to its increased prior probability while ignoring or misclassifying minority examples. In this research project, SMOTE, and cost-sensitive learning are applied to temporal models and compared with ensemble learning to handle the class imbalance problem as well as to study the relation to two data pre-processing methods. Experiments show that f-measures of the minority classes are increased when using SMOTE with both temporal models (LSTM and CNN) and based on both ways of extracting features (FTWs and ESTWs). For example, the recognition measurement of the *Snack and Dinner* as one of the minority classes is notably improved in both homes, using both models and based on both feature extraction methods. The experimental results indicate that handling imbalanced data is more important than selecting machine learning algorithms and improves classification performance. Moreover, handling imbalanced class problem from data level using SMOTE and ESTWs for these activity datasets outperforms the algorithm level.

Figure 4.1 shows a framework for AR systems based on the above integrated results. Firstly, the multiple incremental FTWs technique is applied to generate the input dataset. Then SMOTE is applied to handle the imbalanced class problem of human activities and to generate a balanced dataset. However, SMOTE in generating synthetic samples does not consider neighbouring samples that can be from other classes. This can introduce additional noise and can increase the overlapping of classes. To address this limitation, we compute K nearest neighbours of each generated synthetic sample to make sure the generated samples are correctly labelled. The generated sample for each of the minority classes with the K nearest neighbours must have the same class. For instance, generated synthetic samples of *Snack* activity must have K *Snack* activity as nearest neighbors. Over-sampling with accurate labelling technique is used to minimize the focus of the proposed network on learning only the majority activities or partially neglect the minority activities. The balanced dataset is fed into the proposed delayed temporal models i.e. LSTM and 1D CNN to accurately classify human activities.

For future work, firstly, we will explore extensions of NLPA for aligning low-dimensional manifolds of disparate data sets. Then t-SNE low-dimensional manifolds of disparate data sets will be compared using NLPA to discover the common manifolds of the disparate data sets to be used for Transfer learning.

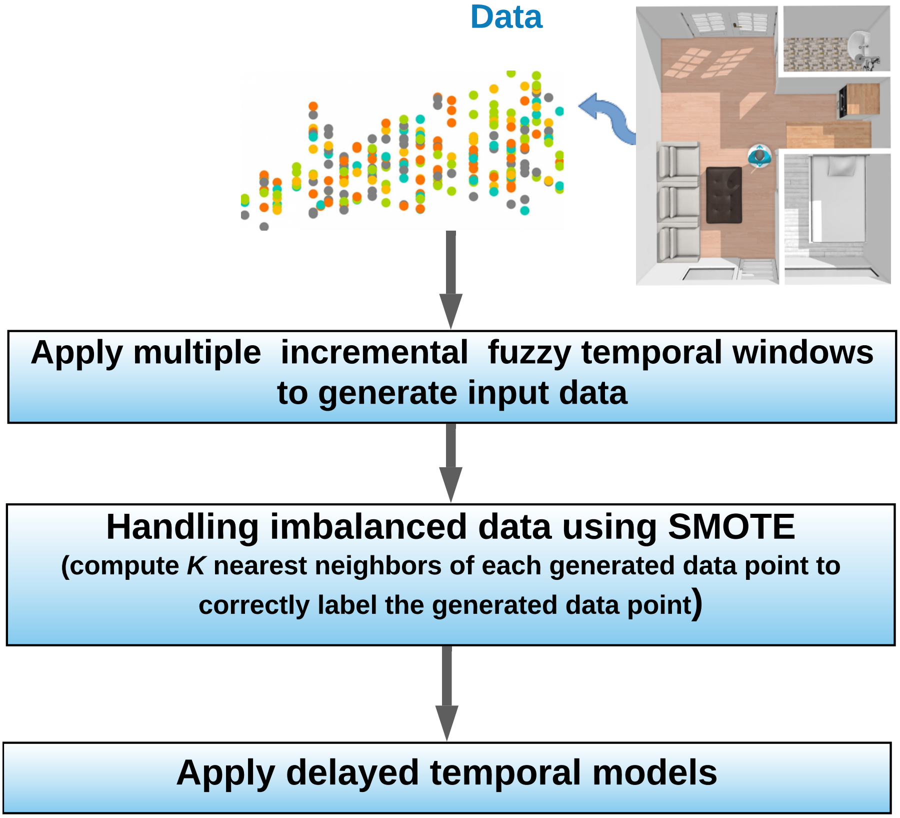


Figure 4.1: Framework for AR

Secondly, future work will explore a newly proposed approach to handle the imbalanced class problem by integrating SMOTE with weak supervision. This approach will use SMOTE only to generate observations from minority classes and use weak supervision to correctly and properly label the new observations. The idea is designed to target the challenge of correctly labeling samples created in an over-sampling context.

Finally, this project mainly will work on boosting learning across different smart homes aiming to perform robust recognition of dangerous situations and detect behavior deviations in order to enhance elderly-care alert systems. This will be conducted by transferring knowledge over different smart homes in terms of layout, resident and sensor configuration.

References

[1] Lei Gao, AK Bourke, and John Nelson. Evaluation of accelerometer based multi-sensor versus single-sensor activity recognition systems. *Medical engineering & physics*, 36(6):779–785, 2014.

[2] Norbert Noury and Tareq Hadidi. Computer simulation of the activity of the elderly person living independently in a health smart home. *Computer methods and programs in biomedicine*, 108(3):1216–1228, 2012.

[3] Sumi Helal, William Mann, Hicham El-Zabadani, Jeffrey King, Youssef Kaddoura, and Erwin Jansen. The gator tech smart house: A programmable pervasive space. *Computer*, (3):50–60, 2005.

[4] Suneth Ranasinghe, Fadi Al Machot, and Heinrich C Mayr. A review on applications of activity recognition systems with regard to performance and evaluation. *International Journal of Distributed Sensor Networks*, 12(8):1550147716665520, 2016.

[5] Shoya Ishimaru, Kensuke Hoshika, Kai Kunze, Koichi Kise, and Andreas Dengel. Towards reading trackers in the wild: detecting reading activities by eog glasses and deep neural networks. In *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*, pages 704–711, 2017.

[6] Xinyu Li, Yanyi Zhang, Mengzhu Li, Ivan Marsic, JaeWon Yang, and Randall S Burd. Deep neural network for rfid-based activity recognition. In *Proceedings of the Eighth Wireless of the Students, by the Students, and for the Students Workshop*, pages 24–26, 2016.

[7] Wei Niu, Jiao Long, Dan Han, and Yuan-Fang Wang. Human activity detection and recognition for video surveillance. In *2004 IEEE International Conference on Multimedia and Expo (ICME)(IEEE Cat. No. 04TH8763)*, volume 1, pages 719–722. IEEE, 2004.

[8] Chun-Fang Su, Li-Chen Fu, Yi-Wei Chien, and Ting-Ying Li. Activity recognition system for dementia in smart homes based on wearable sensor data. In *2018 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 463–469. IEEE, 2018.

[9] Jiho Park, Kiyoung Jang, and Sung-Bong Yang. Deep neural networks for activity recognition with multi-sensor data in a smart home. In *2018 IEEE 4th World Forum on Internet of Things (WF-IoT)*, pages 155–160. IEEE, 2018.

[10] 10 facts on ageing and health. `https://www.who.int/features/factfiles/ageing/en/`. Accessed: 2020-07-20.

[11] Qin Ni, Ana García Hernando, and Iván de la Cruz. The elderly's independent living in smart homes: A characterization of activities and sensing infrastructure survey to facilitate services development. *Sensors*, 15(5):11312–11362, 2015.

[12] Thinagaran Perumal, YL Chui, Mohd Anuaruddin Bin Ahmadon, and Shingo Yamaguchi. Iot based activity recognition among smart home residents. In *2017 IEEE 6th Global Conference on Consumer Electronics (GCCE)*, pages 1–2. IEEE, 2017.

[13] Adls (activities of daily living). `https://www.medicinenet.com/script/main/art.asp?articlekey=2152`. Accessed: 2020-07-19.

[14] Ruijiao Li, Bowen Lu, and Klaus D McDonald-Maier. Cognitive assisted living ambient system: a survey. *Digital Communications and Networks*, 1(4):229–252, 2015.

[15] Jindong Wang, Yiqiang Chen, Shuji Hao, Xiaohui Peng, and Lisha Hu. Deep learning for sensor-based activity recognition: A survey. *Pattern Recognition Letters*, 119:3–11, 2019.

[16] Godwin Ogbuabor and Robert La. Human activity recognition for healthcare using smartphones. In *Proceedings of the 2018 10th International Conference on Machine Learning and Computing*, pages 41–46, 2018.

[17] Duckki Lee and Sumi Helal. From activity recognition to situation recognition. In *International Conference on Smart Homes and Health Telematics*, pages 245–251. Springer, 2013.

[18] Henry Friday Nweke, Ying Wah Teh, Mohammed Ali Al-Garadi, and Uzoma Rita Alo. Deep learning algorithms for human activity recognition using mobile and wearable sensor networks: State of the art and research challenges. *Expert Systems with Applications*, 2018.

[19] Liang Cao, Yufeng Wang, Bo Zhang, Qun Jin, and Athanasios V Vasilakos. Gchar: An efficient group-based context—aware human activity recognition on smartphone. *Journal of Parallel and Distributed Computing*, 118:67–80, 2018.

[20] Luyang Jing, Taiyong Wang, Ming Zhao, and Peng Wang. An adaptive multi-sensor data fusion method based on deep convolutional neural networks for fault diagnosis of planetary gearbox. *Sensors*, 17(2):414, 2017.

[21] Flávia Alves, Martin Gairing, Frans A Oliehoek, and Thanh-Toan Do. Sensor data for human activity recognition: Feature representation and benchmarking. *arXiv preprint arXiv:2005.07308*, 2020.

[22] Kaixuan Chen, Dalin Zhang, Lina Yao, Bin Guo, Zhiwen Yu, and Yunhao Liu. Deep learning for sensor-based human activity recognition: overview, challenges and opportunities. *arXiv preprint arXiv:2001.07416*, 2020.

[23] Dan Ding, Rory A Cooper, Paul F Pasquina, and Lavinia Fici-Pasquina. Sensor technology for smart homes. *Maturitas*, 69(2):131–136, 2011.

[24] Diane Cook, Kyle D Feuz, and Narayanan C Krishnan. Transfer learning for activity recognition: A survey. *Knowledge and information systems*, 36(3):537–556, 2013.

[25] Andreas Bulling, Ulf Blanke, and Bernt Schiele. A tutorial on human activity recognition using body-worn inertial sensors. *ACM Computing Surveys (CSUR)*, 46(3):1–33, 2014.

[26] Derek Hao Hu and Qiang Yang. Transfer learning for activity recognition via sensor mapping. In *Twenty-second international joint conference on artificial intelligence*, 2011.

[27] Kyle D Feuz and Diane J Cook. Collegial activity learning between heterogeneous sensors. *Knowledge and information systems*, 53(2):337–364, 2017.

[28] Jon TS Quah and M Sriganesh. Real-time credit card fraud detection using computational intelligence. *Expert systems with applications*, 35(4):1721–1732, 2008.

[29] Alejandro Vera-Baquero, Ricardo Colomo-Palacios, and Owen Molloy. Real-time business activity monitoring and analysis of process performance on big-data domains. *Telematics and Informatics*, 33(3):793–807, 2016.

[30] OS Albahri, AS Albahri, KI Mohammed, AA Zaidan, BB Zaidan, M Hashim, and Omar H Salman. Systematic review of real-time remote health monitoring system in triage and priority-based sensor technology: Taxonomy, open challenges, motivation and recommendations. *Journal of medical systems*, 42(5):80, 2018.

[31] Rajesh Kumar Tripathi, Anand Singh Jalal, and Subhash Chand Agrawal. Suspicious human activity recognition: a review. *Artificial Intelligence Review*, 50(2):283–339, 2018.

[32] Barnan Das, Narayanan C Krishnan, and Diane J Cook. Handling class overlap and imbalance to detect prompt situations in smart homes. In *2013 IEEE 13th International Conference on Data Mining Workshops*, pages 266–273. IEEE, 2013.

[33] Monica-Andreea Dragan and Irina Mocanu. Human activity recognition in smart environments. In *2013 19th International Conference on Control Systems and Computer Science*, pages 495–502. IEEE, 2013.

[34] Lamya Fergani, Belkacem Fergani, Anthony Fleury, et al. Improving human activity recognition in smart homes. *International Journal of E-Health and Medical Communications (IJEHMC)*, 6(3):19–37, 2015.

[35] Jubil T Sunny, Sonia Mary George, Jubilant J Kizhakkethottam, Jubil T Sunny, Sonia Mary George, and Jubilant J Kizhakkethottam. Applications and challenges of human activity recognition using sensors in a smart environment. *IJIRST Int. J. Innov. Res. Sci. Technol*, 2:50–57, 2015.

[36] Nancy E ElHady and Julien Provost. A systematic survey on sensor failure detection and fault-tolerance in ambient assisted living. *Sensors*, 18(7):1991, 2018.

[37] Per Lynggaard. Smart home wireless sensor nodes addressing the challenges using smart objects and artificial intelligence. *arXiv preprint arXiv:1904.01504*, 2019.

[38] Liming Chen, Jesse Hoey, Chris D Nugent, Diane J Cook, and Zhiwen Yu. Sensor-based activity recognition. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 42(6):790–808, 2012.

[39] Li Liu, Yuxin Peng, Ming Liu, and Zigang Huang. Sensor-based human activity recognition system with a multilayered model using time series shapelets. *Knowledge-Based Systems*, 90:138–152, 2015.

[40] Wei Wang, Alex X Liu, Muhammad Shahzad, Kang Ling, and Sanglu Lu. Understanding and modeling of wifi signal based human activity recognition. In *Proceedings of the 21st annual international conference on mobile computing and networking*, pages 65–76. ACM, 2015.

[41] Iram Fatima, Muhammad Fahim, Young-Koo Lee, and Sungyoung Lee. A unified framework for activity recognition-based behavior analysis and action prediction in smart homes. *Sensors*, 13(2):2682–2699, 2013.

[42] Charissa Ann Ronao and Sung-Bae Cho. Human activity recognition with smartphone sensors using deep learning neural networks. *Expert systems with applications*, 59:235–244, 2016.

[43] Muhammad Shoaib, Stephan Bosch, Ozlem Incel, Hans Scholten, and Paul Havinga. Complex human activity recognition using smartphone and wrist-worn motion sensors. *Sensors*, 16(4):426, 2016.

[44] Yufei Chen and Chao Shen. Performance analysis of smartphone-sensor behavior for human activity recognition. *Ieee Access*, 5:3095–3110, 2017.

[45] Ozlem Durmaz Incel, Mustafa Kose, and Cem Ersoy. A review and taxonomy of activity recognition on mobile phones. *BioNanoScience*, 3(2):145–171, 2013.

[46] Reza Akhavian and Amir H Behzadan. Smartphone-based construction workers' activity recognition and classification. *Automation in Construction*, 71:198–209, 2016.

[47] Samaneh Zolfaghari and Mohammad Reza Keyvanpour. Sarf: Smart activity recognition framework in ambient assisted living. In *2016 Federated Conference on Computer Science and Information Systems (FedCSIS)*, pages 1435–1443. IEEE, 2016.

[48] Subhas Chandra Mukhopadhyay. Wearable sensors for human activity monitoring: A review. *IEEE sensors journal*, 15(3):1321–1330, 2014.

[49] Tal Shany, Stephen J Redmond, Michael R Narayanan, and Nigel H Lovell. Sensors-based wearable systems for monitoring of human movement and falls. *IEEE Sensors Journal*, 12(3):658–670, 2011.

[50] Lorenzo T D'Angelo, Jakob Neuhaeuser, Yan Zhao, and Tim C Lueth. Simple-use—sensor set for wearable movement and interaction research. *IEEE Sensors Journal*, 14(4):1207–1215, 2013.

[51] Yao-Chiang Kan and Chun-Kai Chen. A wearable inertial sensor node for body motion analysis. *IEEE Sensors Journal*, 12(3):651–657, 2011.

[52] Yindar Chuo, Marcin Marzencki, Benny Hung, Camille Jaggernauth, Kouhyar Tavakolian, Philip Lin, and Bozena Kaminska. Mechanically flexible wireless multisensor platform for human physical activity and vitals monitoring. *IEEE transactions on biomedical circuits and systems*, 4(5):281–294, 2010.

[53] Benoit Mariani, Mayté Castro Jiménez, François JG Vingerhoets, and Kamiar Aminian. On-shoe wearable sensors for gait and turning assessment of patients with parkinson's disease. *IEEE transactions on biomedical engineering*, 60(1):155–158, 2012.

[54] Edward S Sazonov, George Fulk, James Hill, Yves Schutz, and Raymond Browning. Monitoring of posture allocations and activities by a shoe-based wearable sensor. *IEEE Transactions on Biomedical Engineering*, 58(4):983–990, 2010.

[55] Søren Brage, Niels Brage, Paul W Franks, Ulf Ekelund, and Nicholas J Wareham. Reliability and validity of the combined heart rate and movement sensor actiheart. *European journal of clinical nutrition*, 59(4):561, 2005.

[56] Long Yan, Joonsung Bae, Seulki Lee, Taehwan Roh, Kiseok Song, and Hoi-Jun Yoo. A 3.9 mw 25-electrode reconfigured sensor for wearable cardiac monitoring system. *IEEE Journal of Solid-State Circuits*, 46(1):353–364, 2010.

[57] Bruce H Dobkin and Andrew Dorsch. The promise of mhealth: daily activity monitoring and outcome assessments by wearable sensors. *Neurorehabilitation and neural repair*, 25(9):788–798, 2011.

[58] Che-Chang Yang and Yeh-Liang Hsu. A review of accelerometry-based wearable motion detectors for physical activity monitoring. *Sensors*, 10(8):7772–7788, 2010.

[59] Young-Dong Lee and Wan-Young Chung. Wireless sensor network based wearable smart shirt for ubiquitous health and activity monitoring. *Sensors and Actuators B: Chemical*, 140(2):390–395, 2009.

[60] Ferhat Attal, Samer Mohammed, Mariam Dedabrishvili, Faicel Chamroukhi, Latifa Oukhellou, and Yacine Amirat. Physical human activity recognition using wearable sensors. *Sensors*, 15(12):31314–31338, 2015.

[61] Dean M Karantonis, Michael R Narayanan, Merryn Mathie, Nigel H Lovell, and Branko G Celler. Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring. *IEEE transactions on information technology in biomedicine*, 10(1):156–167, 2006.

[62] Alberto G Bonomi, Annelies HC Goris, Bin Yin, and Klaas R Westerterp. Detection of type, duration, and intensity of physical activity using an accelerometer. *Medicine & Science in Sports & Exercise*, 41(9):1770–1777, 2009.

[63] Bijan Najafi, Kamiar Aminian, Anisoara Paraschiv-Ionescu, François Loew, Christophe J Bula, and Philippe Robert. Ambulatory system for human motion analysis using a kinematic sensor: monitoring of daily physical activity in the elderly. *IEEE Transactions on biomedical Engineering*, 50(6):711–723, 2003.

[64] Hristijan Gjoreski and Matjaž Gams. Accelerometer data preparation for activity recognition. In *Proceedings of the International Multiconference Information Society, Ljubljana, Slovenia*, volume 1014, page 1014, 2011.

[65] Ming Jiang, Hong Shang, Zhelong Wang, Hongyi Li, and Yuechao Wang. A method to deal with installation errors of wearable accelerometers for human activity recognition. *Physiological measurement*, 32(3):347, 2011.

[66] Jennifer R Kwapisz, Gary M Weiss, and Samuel A Moore. Activity recognition using cell phone accelerometers. *ACM SigKDD Explorations Newsletter*, 12(2):74–82, 2011.

[67] Chun Zhu and Weihua Sheng. Motion-and location-based online human daily activity recognition. *Pervasive and Mobile Computing*, 7(2):256–269, 2011.

[68] Pekka Siirtola and Juha Röning. User-independent human activity recognition using a mobile phone: Offline recognition vs. real-time on device recognition. In *Distributed computing and artificial intelligence*, pages 617–627. Springer, 2012.

[69] C Sweetlin Hemalatha and Vijay Vaidehi. Frequent bit pattern mining over tri-axial accelerometer data streams for recognizing human activities and detecting fall. *Procedia Computer Science*, 19:56–63, 2013.

[70] Andrea Mannini, Stephen S Intille, Mary Rosenberger, Angelo M Sabatini, and William Haskell. Activity recognition using a single accelerometer placed at the wrist or ankle. *Medicine and science in sports and exercise*, 45(11):2193, 2013.

[71] Yonglei Zheng, Weng-Keen Wong, Xinze Guan, and Stewart Trost. Physical activity recognition from accelerometer data using a multi-scale ensemble method. In *Twenty-Fifth IAAI Conference*, 2013.

[72] Aiguo Wang, Guilin Chen, Cuijuan Shang, Miaofei Zhang, and Li Liu. Human activity recognition in a smart home environment with stacked denoising autoencoders. In *International conference on web-age information management*, pages 29–40. Springer, 2016.

[73] Guilin Chen, Aiguo Wang, Shenghui Zhao, Li Liu, and Chih-Yung Chang. Latent feature learning for activity recognition using simple sensors in smart homes. *Multimedia Tools and Applications*, 77(12):15201–15219, 2018.

[74] Pireh Pirzada, Neil White, and Adriana Wilde. Sensors in smart homes for independent living of the elderly. In *2018 5th International Multi-Topic ICT Conference (IMTIC)*, pages 1–8. IEEE, 2018.

[75] Pavle Skocir, Petar Krivic, Matea Tomeljak, Mario Kusek, and Gordan Jezic. Activity detection in smart home environment. In *KES*, pages 672–681, 2016.

[76] Ioan Susnea, Luminita Dumitriu, Mihai Talmaciu, Emilia Pecheanu, and Dan Munteanu. Unobtrusive monitoring the daily activity routine of elderly people living alone, with low-cost binary sensors. *Sensors*, 19(10):2264, 2019.

[77] Yegang Du, Yuto Lim, and Yasuo Tan. A novel human activity recognition and prediction in smart home based on interaction. *Sensors*, 19(20):4474, 2019.

[78] Jens Lundström, Wagner O De Morais, Maria Menezes, Celeste Gabrielli, João Bentes, Anita Sant'Anna, Jonathan Synnott, and Chris Nugent. Halmstad intelligent home-capabilities and opportunities. In *International Conference on IoT Technologies for HealthCare*, pages 9–15. Springer, 2016.

[79] Beth Logan, Jennifer Healey, Matthai Philipose, Emmanuel Munguia Tapia, and Stephen Intille. A long-term evaluation of sensing modalities for activity recognition. In *International conference on Ubiquitous computing*, pages 483–500. Springer, 2007.

[80] Diane J Cook, Michael Youngblood, Edwin O Heierman, Karthik Gopalratnam, Sira Rao, Andrey Litvin, and Farhan Khawaja. Mavhome: An agent-based smart home. In *Proceedings of the First IEEE International Conference on Pervasive Computing and Communications, 2003.(PerCom 2003).*, pages 521–524. IEEE, 2003.

[81] TLM Van Kasteren, Gwenn Englebienne, and Ben JA Kröse. Activity recognition using semi-markov models on real world smart home datasets. *Journal of ambient intelligence and smart environments*, 2(3):311–325, 2010.

[82] Chao Chen and Prafulla Dawadi. Casasviz: Web-based visualization of behavior patterns in smart environments. In *2011 IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)*, pages 301–303. IEEE, 2011.

[83] Diane J Cook, Aaron S Crandall, Brian L Thomas, and Narayanan C Krishnan. Casas: A smart home in a box. *Computer*, 46(7):62–69, 2012.

[84] Michel Vacher, Dan Istrate, François Portet, Thierry Joubert, Thierry Chevalier, Serge Smidtas, Brigitte Meillon, Benjamin Lecouteux, Mohamed Sehili, Pedro Chahuara, et al. The sweet-home project: Audio technology in smart homes to improve well-being and reliance. In *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 5291–5294. IEEE, 2011.

[85] Pedro Chahuara, François Portet, and Michel Vacher. Making context aware decision from uncertain information in a smart home: A markov logic network approach. In *International Joint Conference on Ambient Intelligence*, pages 78–93. Springer, 2013.

[86] Panagiotis E Antoniou, Evdokimos I Konstantinidis, Antonis S Billis, and Panagiotis D Bamidis. Integrating the usefil assisted living platform; observation from the field. In *6th European conference of the international federation for medical and biological engineering*, pages 657–660. Springer, 2015.

[87] Qing Zhang, Mohan Karunanithi, Rajib Rana, and Jiajun Liu. Determination of activities of daily living of independent living older people using environmentally placed sensors. In *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 7044–7047. IEEE, 2013.

[88] Qing Zhang, Ying Su, and Ping Yu. Assisting an elderly with early dementia using wireless sensors data in smarter safer home. In *International Conference on Informatics and Semiotics in Organisations*, pages 398–404. Springer, 2014.

[89] Fco Ordóñez, Paula de Toledo, Araceli Sanchis, et al. Activity recognition using hybrid generative/discriminative models on home environments using binary sensors. *Sensors*, 13(5):5460–5477, 2013.

[90] Marjorie Skubic, Gregory Alexander, Mihail Popescu, Marilyn Rantz, and James Keller. A smart home application to eldercare: Current status and lessons learned. *Technology and Health Care*, 17(3):183–201, 2009.

[91] TLM Van Kasteren, Gwenn Englebienne, and Ben JA Kröse. Transferring knowledge of activity recognition across sensor networks. In *International Conference on Pervasive Computing*, pages 283–300. Springer, 2010.

[92] Daniel Roggen, Alberto Calatroni, Mirco Rossi, Thomas Holleczek, Kilian Förster, Gerhard Tröster, Paul Lukowicz, David Bannach, Gerald Pirkl, Alois Ferscha, et al. Collecting complex activity datasets in highly rich networked sensor environments. In *2010 Seventh international conference on networked sensing systems (INSS)*, pages 233–240. IEEE, 2010.

[93] Oresti Baños, Miguel Damas, Héctor Pomares, Ignacio Rojas, Máté Attila Tóth, and Oliver Amft. A benchmark dataset to evaluate sensor displacement in activity recognition. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, pages 1026–1035. ACM, 2012.

[94] Hande Alemdar, Halil Ertan, Ozlem Durmaz Incel, and Cem Ersoy. Aras human activity datasets in multiple homes with multiple residents. In *Proceedings of the 7th International Conference on Pervasive Computing Technologies for Healthcare*, pages 232–235. ICST (Institute for Computer Sciences, Social-Informatics and ..., 2013.

[95] Paula Lago, Fréderic Lang, Claudia Roncancio, Claudia Jiménez-Guarín, Radu Mateescu, and Nicolas Bonnefond. The contextact@ a4h real-life dataset of daily-living activities. In *International and Interdisciplinary Conference on Modeling and Using Context*, pages 175–188. Springer, 2017.

[96] Muhammad Raisul Alam, Mamun Bin Ibne Reaz, and Mohd Alauddin Mohd Ali. A review of smart homes—past, present, and future. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 42(6):1190–1203, 2012.

[97] Parisa Rashidi and Alex Mihailidis. A survey on ambient-assisted living tools for older adults. *IEEE journal of biomedical and health informatics*, 17(3):579–590, 2012.

[98] Abdelhamid Salih and A Abraham. A review of ambient intelligence assisted healthcare monitoring. *International Journal of Computer Information Systems and Industrial Management (IJCISIM)*, 5:741–750, 2013.

[99] Kirsten KB Peetoom, Monique AS Lexis, Manuela Joore, Carmen D Dirksen, and Luc P De Witte. Literature review on monitoring technologies and their outcomes in independently living elderly people. *Disability and Rehabilitation: Assistive Technology*, 10(4):271–294, 2015.

[100] Rinat Khusainov, Djamel Azzi, Ifeyinwa E Achumba, and Sebastian D Bersch. Real-time human ambulation, activity, and physiological monitoring: Taxonomy of issues, techniques, applications, challenges and limitations. *Sensors*, 13(10):12852–12902, 2013.

[101] Emiro De-La-Hoz-Franco, Paola Ariza-Colpas, Javier Medina Quero, and Macarena Espinilla. Sensor-based datasets for human activity recognition–a systematic review of literature. *IEEE Access*, 6:59192–59210, 2018.

[102] Eftim Zdravevski, Petre Lameski, Vladimir Trajkovik, Andrea Kulakov, Ivan Chorbev, Rossitza Goleva, Nuno Pombo, and Nuno Garcia. Improving activity recognition accuracy in ambient-assisted living systems by automated feature engineering. *Ieee Access*, 5:5262–5280, 2017.

[103] Jin-Hyuk Hong, Julian Ramos, Choonsung Shin, and Anind K Dey. An activity recognition system for ambient assisted living environments. In *International Competition on Evaluating AAL Systems Through Competitive Benchmarking*, pages 148–158. Springer, 2012.

[104] Holger Storf, Thomas Kleinberger, Martin Becker, Mario Schmitt, Frank Bomarius, and Stephan Prueckner. An event-driven approach to activity recognition in ambient assisted living. In *European conference on ambient intelligence*, pages 123–132. Springer, 2009.

[105] Geetika Singla, Diane J Cook, and Maureen Schmitter-Edgecombe. Incorporating temporal reasoning into activity recognition for smart home residents. In *Proceedings of the AAAI workshop on spatial and temporal reasoning*, pages 53–61, 2008.

[106] Geetika Singla, Diane J Cook, and Maureen Schmitter-Edgecombe. Tracking activities in complex settings using smart environment technologies. *International journal of biosciences, psychiatry, and technology (IJBSPT)*, 1(1):25, 2009.

[107] Zhiqiang Zhang, Xuebin Gao, Jit Biswas, and Jian Kang Wu. Moving targets detection and localization in passive infrared sensor networks. In *2007 10th International Conference on Information Fusion*, pages 1–6. IEEE, 2007.

[108] Gennaro Tartarisco, Giovanni Baldus, Daniele Corda, Rossella Raso, Antonino Arnao, Marcello Ferro, Andrea Gaggioli, and Giovanni Pioggia. Personal health system architecture for stress monitoring and support to clinical decisions. *Computer Communications*, 35(11):1296–1305, 2012.

[109] Pema Chodon, Devi Maya Adhikari, Gopal Chandra Nepal, Rajen Biswa, Sangay Gyeltshen, et al. Passive infrared (pir) sensor based security system. *International Journal of Electrical, Electronics & Computer Systems*, 14(2), 2013.

[110] Dipak Surie, Olivier Laguionie, and Thomas Pederson. Wireless sensor networking of everyday objects in a smart home environment. In *2008 International Conference on Intelligent Sensors, Sensor Networks and Information Processing*, pages 189–194. IEEE, 2008.

[111] Xin Hong and Chris D Nugent. Partitioning time series sensor data for activity recognition. In *2009 9th international conference on information technology and applications in biomedicine*, pages 1–4. IEEE, 2009.

[112] Huiru Zheng, Haiying Wang, and Norman Black. Human activity detection in smart home environment with self-adaptive neural networks. In *2008 IEEE International Conference on Networking, Sensing and Control*, pages 1505–1510. IEEE, 2008.

[113] Fco Javier Ordóñez, José Antonio Iglesias, Paula De Toledo, Agapito Ledezma, and Araceli Sanchis. Online activity recognition using evolving classifiers. *Expert Systems with Applications*, 40(4):1248–1255, 2013.

[114] Chao Chen, Barnan Das, and Diane J Cook. A data mining framework for activity recognition in smart environments. In *2010 Sixth International Conference on Intelligent Environments*, pages 80–83. IEEE, 2010.

[115] Nicholas Foubert, Anita M McKee, Rafik A Goubran, and Frank Knoefel. Lying and sitting posture recognition and transition detection using a pressure sensor array. In *2012 IEEE International Symposium on Medical Measurements and Applications Proceedings*, pages 1–6. IEEE, 2012.

[116] Jonathan Lester, Tanzeem Choudhury, and Gaetano Borriello. A practical approach to recognizing physical activities. In *International conference on pervasive computing*, pages 1–16. Springer, 2006.

[117] Joon-Ho Lim, Hyunchul Jang, Jaewon Jang, and Soo-Jun Park. Daily activity recognition system for the elderly using pressure sensors. In *2008 30th annual international conference of the ieee engineering in medicine and biology society*, pages 5188–5191. IEEE, 2008.

[118] Norbert Noury, Thierry Hervé, Vicent Rialle, Gilles Virone, Eric Mercier, Gilles Morey, Aldo Moro, and Thierry Porcheron. Monitoring behavior in home using a smart fall sensor and position sensors. In *1st Annual International IEEE-EMBS Special Topic Conference on Microtechnologies in Medicine and Biology. Proceedings (Cat. No. 00EX451)*, pages 607–610. IEEE, 2000.

[119] Michael A Mills and Thomas J Gerhardt. Patient cable sensor switch, December 7 1999. US Patent 5,997,343.

[120] Alan K Forsythe, Andrew F Thompson, Charles H Simmons, and John M Morris. Contactless stalk mounted control switch, January 16 2001. US Patent 6,175,290.

[121] Xin Hong, Chris Nugent, Maurice Mulvenna, Sally McClean, Bryan Scotney, and Steven Devlin. Evidential fusion of sensor data for activity recognition in smart homes. *Pervasive and Mobile Computing*, 5(3):236–252, 2009.

[122] Matthai Philipose, Kenneth P Fishkin, Mike Perkowitz, Donald J Patterson, Dieter Fox, Henry Kautz, and Dirk Hahnel. Inferring activities from interactions with objects. *IEEE pervasive computing*, (4):50–57, 2004.

[123] Joshua R Smith, Kenneth P Fishkin, Bing Jiang, Alexander Mamishev, Matthai Philipose, Adam D Rea, Sumit Roy, and Kishore Sundara-Rajan. Rfid-based techniques for human-activity detection. *Communications of the ACM*, 48(9):39–44, 2005.

[124] James Fogarty, Carolyn Au, and Scott E Hudson. Sensing from the basement: a feasibility study of unobtrusive and low-cost home activity recognition. In *Proceedings of the 19th annual ACM symposium on User interface software and technology*, pages 91–100. ACM, 2006.

[125] Nadia Zouba, Francois Bremond, Monique Thonnat, and Van Thinh Vu. Multi-sensors analysis for everyday activity monitoring. *Proc. of SETIT*, pages 25–29, 2007.

[126] Daniel H Wilson and Chris Atkeson. Simultaneous tracking and activity recognition (star) using many anonymous, binary sensors. In *International Conference on Pervasive Computing*, pages 62–79. Springer, 2005.

[127] Nils Krahnstoever, Jens Rittscher, P Tu, Kevin Chean, and T Tomlinson. Activity recognition using visual tracking and rfid. In *2005 Seventh IEEE Workshops on Applications of Computer Vision (WACV/MOTION'05)-Volume 1*, volume 1, pages 494–500. IEEE, 2005.

[128] Siddika Parlak, Ivan Marsic, and Randall S Burd. Activity recognition for emergency care using rfid. In *Proceedings of the 6th International Conference on Body Area Networks*, pages 40–46. ICST (Institute for Computer Sciences, Social-Informatics and . . . , 2011.

[129] Liang Wang, Tao Gu, Hongwei Xie, Xianping Tao, Jian Lu, and Yu Huang. A wearable rfid system for real-time activity recognition using radio patterns. In *International Conference on Mobile and Ubiquitous Systems: Computing, Networking, and Services*, pages 370–383. Springer, 2013.

[130] Dany Fortin-Simard, Jean-Sébastien Bilodeau, Kevin Bouchard, Sebastien Gaboury, Bruno Bouchard, and Abdenour Bouzouane. Exploiting passive rfid technology for activity recognition in smart homes. *IEEE Intelligent systems*, 30(4):7–15, 2015.

[131] Yunhao Liu, Yiyang Zhao, Lei Chen, Jian Pei, and Jinsong Han. Mining frequent trajectory patterns for activity monitoring using radio frequency tag arrays. *IEEE Transactions on Parallel and Distributed Systems*, 23(11):2138–2149, 2011.

[132] Christian Meißner, Jürgen Meixensberger, Andreas Pretschner, and Thomas Neumuth. Sensor-based surgical activity recognition in unconstrained environments. *Minimally Invasive Therapy & Allied Technologies*, 23(4):198–205, 2014.

[133] Sara Amendola, Luigi Bianchi, and Gaetano Marrocco. Movement detection of human body segments: Passive radio-frequency identification and machine-learning technologies. *IEEE Antennas and Propagation Magazine*, 57(3):23–37, 2015.

[134] Tsutomu Fujinami, Motoki Miura, Ryozo Takatsuka, and Taro Sugihara. A study of long term tendencies in residents' activities of daily living at a group home for people with dementia using rfid slippers. In *International Conference on Smart Homes and Health Telematics*, pages 303–307. Springer, 2011.

[135] Soo-Cheol Kim, Young-Sik Jeong, and Sang-Oh Park. Rfid-based indoor location tracking to ensure the safety of the elderly in smart home environments. *Personal and ubiquitous computing*, 17(8):1699–1707, 2013.

[136] Liang Wang, Tao Gu, Xianping Tao, Hanhua Chen, and Jian Lu. Recognizing multi-user activities using wearable sensors in a smart home. *Pervasive and Mobile Computing*, 7(3):287–298, 2011.

[137] Toufik Guettari, Jérôme Boudy, B E Benkelfat, Gérard Chollet, J L Baldinger, Pascal Doré, and Dan Istrate. Thermal signal analysis in smart home environment for detecting a human presence. In *2014 1st International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, pages 334–339. IEEE, 2014.

[138] Siddika Parlak and Ivan Marsic. Detecting object motion using passive rfid: A trauma resuscitation case study. *IEEE Transactions on Instrumentation and Measurement*, 62(9):2430–2437, 2013.

[139] Muhammad Muaaz and René Mayrhofer. An analysis of different approaches to gait recognition using cell phone based accelerometers. In *Proceedings of International Conference on Advances in Mobile Computing & Multimedia*, page 293. ACM, 2013.

[140] Thorsten Rodner and Lothar Litz. Data-driven generation of rule-based behavior models for an ambient assisted living system. In *2013 IEEE Third International Conference on Consumer Electronics¿ Berlin (ICCE-Berlin)*, pages 35–38. IEEE, 2013.

[141] Xudong Sun and Yue Zhang. Design and implementation of portable ecg and body temperature monitor. In *2014 International Symposium on Computer, Consumer and Control*, pages 910–913. IEEE, 2014.

[142] HB Mitchell. Sensor value normalization. In *Multi-Sensor Data Fusion*, pages 97–112. Springer, 2007.

[143] Fadi Al Machot, Heinrich C Mayr, and Suneth Ranasinghe. A windowing approach for activity recognition in sensor data streams. In *2016 Eighth International Conference on Ubiquitous and Future Networks (ICUFN)*, pages 951–953. IEEE, 2016.

[144] Oresti Banos, Juan-Manuel Galvez, Miguel Damas, Hector Pomares, and Ignacio Rojas. Window size impact in human activity recognition. *Sensors*, 14(4):6474–6499, 2014.

[145] Qin Ni, Timothy Patterson, Ian Cleland, and Chris Nugent. Dynamic detection of window starting positions and its implementation within an activity recognition framework. *Journal of biomedical informatics*, 62:171–180, 2016.

[146] Javier Ortiz Laguna, Angel García Olaya, and Daniel Borrajo. A dynamic sliding window approach for activity recognition. In *International Conference on User Modeling, Adaptation, and Personalization*, pages 219–230. Springer, 2011.

[147] Nawel Yala, Belkacem Fergani, and Anthony Fleury. Feature extraction for human activity recognition on streaming data. In *2015 International Symposium on Innovations in Intelligent SysTems and Applications (INISTA)*, pages 1–6. IEEE, 2015.

[148] Anzah H Niazi, Delaram Yazdansepas, Jennifer L Gay, Frederick W Maier, Lakshmish Ramaswamy, Khaled Rasheed, and Matthew P Buman. Statistical analysis of window sizes and sampling rates in human activity recognition. In *HEALTHINF*, pages 319–325, 2017.

[149] Jonathan Liono, A Kai Qin, and Flora D Salim. Optimal time window for temporal segmentation of sensor streams in multi-activity recognition. In *Proceedings of the 13th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services*, pages 10–19. ACM, 2016.

[150] Narayanan C Krishnan and Diane J Cook. Activity recognition on streaming sensor data. *Pervasive and mobile computing*, 10:138–154, 2014.

[151] Javier Medina-Quero, Shuai Zhang, Chris Nugent, and M Espinilla. Ensemble classifier of long short-term memory with fuzzy temporal windows on binary sensors for activity recognition. *Expert Systems with Applications*, 114:441–453, 2018.

[152] Miguel Angel Lopez Medina, Macarena Espinilla, Cristiano Paggeti, and Javier Medina Quero. Activity recognition for iot devices using fuzzy spatio-temporal features as environmental sensor fusion. *Sensors*, 19(16):3512, 2019.

[153] Minoru Yoshizawa, Wataru Takasaki, and Ren Ohmura. Parameter exploration for response time reduction in accelerometer-based activity recognition. In *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication*, pages 653–664. ACM, 2013.

[154] Muhammad Fahim. *Evolutionary learning models for indoor and outdoor human activity recognition*. PhD thesis, PhD Thesis, 2014.

[155] Xin Hong and Chris D Nugent. Segmenting sensor data for activity monitoring in smart environments. *Personal and ubiquitous computing*, 17(3):545–559, 2013.

[156] Emmanuel Munguia Tapia, Stephen S Intille, and Kent Larson. Activity recognition in the home using simple and ubiquitous sensors. In *International conference on pervasive computing*, pages 158–175. Springer, 2004.

[157] Mohammed Shukur, T Rani, S Bhavani, G Sastry, and Surampudi Raju. Local and global intrinsic dimensionality estimation for better chemical space representation. *Multi-disciplinary Trends in Artificial Intelligence*, pages 329–338, 2011.

[158] Harold Hotelling. Analysis of a complex of statistical variables into principal components. *Journal of educational psychology*, 24(6):417, 1933.

[159] Warren S Torgerson. Multidimensional scaling: I. theory and method. *Psychometrika*, 17(4):401–419, 1952.

[160] Laurens van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of Machine Learning Research*, 9(Nov):2579–2605, 2008.

[161] Kyoungok Kim and Jaewook Lee. Sentiment visualization and classification via semi-supervised nonlinear dimensionality reduction. *Pattern Recognition*, 47(2):758–768, 2014.

[162] Francisco Javier García Fernández, Michel Verleysen, John A Lee, and Ignacio Díaz Blanco. Stability comparison of dimensionality reduction techniques attending to data and parameter variations. In *Eurographics Conference on Visualization (EuroVis)(2013)*. The Eurographics Association, 2013.

[163] Andreas C Müller and Sarah Guido. Introduction to machine learning with python: a guide for data scientists. 2017.

[164] Francisco J García-Fernández, Michel Verleysen, John Aldo Lee, and Ignacio Díaz Blanco. Sensitivity to parameter and data variations in dimensionality reduction techniques. In *ESANN*, 2013.

[165] Chang Wang and Sridhar Mahadevan. Manifold alignment without correspondence. In *IJCAI*, volume 2, page 3, 2009.

[166] Lisa Torrey and Jude Shavlik. Transfer learning. *Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques*, 1:242, 2009.

[167] Ndivhuwo Makondo, Benjamin Rosman, and Osamu Hasegawa. Knowledge transfer for learning robot models via local procrustes analysis. In *Humanoid Robots (Humanoids), 2015 IEEE-RAS 15th International Conference on*, pages 1075–1082. IEEE, 2015.

[168] Amy Ross. Procrustes analysis. *Course report, Department of Computer Science and Engineering, University of South Carolina*, 2004.

[169] Jozsef Suto, Stefan Oniga, Claudiu Lung, and Ioan Orha. Comparison of offline and real-time human activity recognition results using machine learning techniques. *Neural Computing and Applications*, pages 1–14, 2018.

[170] Javier Medina, Maria Fernandez-Olmo, M.D. Peláez, and Macarena Espinilla. Real-time monitoring in home-based cardiac rehabilitation using wrist-worn heart rate devices. *Sensors*, 17(2892), 12/2017 2017.

[171] KS Gayathri, Susan Elias, and Balaraman Ravindran. Hierarchical activity recognition for dementia care using markov logic network. *Personal and Ubiquitous Computing*, 19(2):271–285, 2015.

[172] AP Stakhov. The generalized principle of the golden section and its applications in mathematics, science, and engineering. *Chaos, Solitons & Fractals*, 26(2):263–289, 2005.

[173] Justin M. Johnson and Taghi M. Khoshgoftaar. Survey on deep learning with class imbalance. *Journal of Big Data*, 6(1):27, Mar 2019.

[174] KG Manosha Chathuramali and Ranga Rodrigo. Faster human activity recognition with svm. In *Advances in ICT for Emerging Regions (ICTer), 2012 International Conference on*, pages 197–203. IEEE, 2012.

[175] Oscar D Lara, Miguel A Labrador, et al. A survey on human activity recognition using wearable sensors. *IEEE Communications Surveys and Tutorials*, 15(3):1192–1209, 2013.

[176] Jindong Wang, Yiqiang Chen, Shuji Hao, Xiaohui Peng, and Lisha Hu. Deep learning for sensor-based activity recognition: A survey. *Pattern Recognition Letters*, 2018.

[177] Frédéric Li, Kimiaki Shirahama, Muhammad Adeel Nisar, Lukas Köping, and Marcin Grzegorzek. Comparison of feature learning methods for human activity recognition using wearable sensors. *Sensors*, 18(2):679, 2018.

[178] Jianbo Yang, Minh Nhut Nguyen, Phyo Phyo San, Xiaoli Li, and Shonali Krishnaswamy. Deep convolutional neural networks on multichannel time series for human activity recognition. In *Ijcai*, volume 15, pages 3995–4001, 2015.

[179] Nils Y Hammerla, Shane Halloran, and Thomas Ploetz. Deep, convolutional, and recurrent models for human activity recognition using wearables. *arXiv preprint arXiv:1604.08880*, 2016.

[180] Fernando Moya Rueda, René Grzeszick, Gernot Fink, Sascha Feldhorst, and Michael ten Hompel. Convolutional neural networks for human activity recognition using body-worn sensors. In *Informatics*, volume 5, page 26. Multidisciplinary Digital Publishing Institute, 2018.

Appendix A

Paper I
Stability analysis of the t-SNE algorithm for human activity pattern data

Rebeen Ali Hamad, Eric Järpe, Jens Lundström

2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)
Miyazaki, Japan, 7-10 Oct. 2018

2018 IEEE International Conference on Systems, Man, and Cybernetics

Stability analysis of the t-SNE algorithm for human activity pattern data

Rebeen Ali Hamad
Intelligent Systems and Digital Design
Halmstad University, Sweden
rebeen.ali_hamad@hh.se

Eric Järpe
Intelligent Systems and Digital Design
Halmstad University, Sweden
eric.jarpe@hh.se

Jens Lundström
JeCom Consulting
Halmstad, Sweden
jens.lundstrom@jecom-consulting.com

I. Abstract

Health technological systems learning from and reacting on how humans behave in sensor equipped environments are today being commercialized. These systems rely on the assumptions that training data and testing data share the same feature space, and residing from the same underlying distribution - which is commonly unrealistic in real-world applications. Instead, the use of transfer learning could be considered. In order to transfer knowledge between a source and a target domain these should be mapped to a common latent feature space. In this work, the dimensionality reduction algorithm t-SNE is used to map data to a similar feature space and is further investigated through a proposed novel analysis of output stability. The proposed analysis, Normalized Linear Procrustes Analysis (NLPA) extends the existing Procrustes and Local Procrustes algorithms for aligning manifolds. The methods are tested on data reflecting human behaviour patterns from data collected in a smart home environment. Results show high partial output stability for the t-SNE algorithm for the tested input data for which NLPA is able to detect clusters which are individually aligned and compared. The results highlight the importance of understanding output stability before incorporating dimensionality reduction algorithms into further computation, e.g. for transfer learning.

II. Introduction

Exploratory data analysis based on all available dimensions of a high-dimensional data set (HDD) is generally intractable. The underlying manifold structure could have a low intrinsic dimensionality and is therefore often explored using dimensionality reduction (DR) techniques. The intrinsic dimensionality of an HDD could be considered as the minimum number of parameters needed for explaining the observed properties of the data. Moreover, such parameters could preserve both global and local structure of the HDD in the lower-dimensional representative space [28]. Principal Component Analysis (PCA) [9] and multidimensional scaling (MDS) [30] techniques that linearly map an HDD into a lower dimensional representation are broadly used in business and marketing applications for DR and data visualization. PCA and MDS compute low-dimensional maps where dissimilar data points are far apart. However, keeping similar data points close together in a low-dimensional map is crucial for an HDD that lies on or near a low-dimensional manifold, which is often difficult using linear mapping techniques [19]. Therefore PCA and MDS are not suitable for many complex and non-linear datasets [12]. Instead, a popular non-linear DR algorithm is the t-distributed Stochastic Neighbor Embedding (t-SNE) [19] commonly known for producing low-dimensional representations of the input data [5], [22] presumably close to the sought real low-dimensional manifold. In contrast to PCA and MDS, t-SNE as a non-linear DR method has the ability to cope with complex data sets that are likely to lie on a low-dimensional non-linear manifold [4]. t-SNE works in an unsupervised fashion, can utilize any distance metric and commonly adapts to both sparse and dense input data [16], better than ISOMAP [29] and kernel-PCA as it seems [27]. Algorithms for manifold learning, such as t-SNE are used across a broad range of information processing applications including immunology [3], and data compression [32].

Commonly many DR algorithms such as t-SNE are used for visualization of HDD data [23] or for further data processing such as when modelling of human activity patterns [18] by mapping data to a low-dimensional representation. A suitable representation of *how*, *when* and *where* humans performing activities in their own home opens up for various health technology applications such as systems for anomaly detection (e.g. falls) or tracking progression of diseases (e.g. early-warning of dementia). In this paper, data from a sensor equipped smart home is used.

Although t-SNE is presented to be a suitable method for data visualization, t-SNE has a few potential weaknesses [19]. One is the uncertainty of convergence. Due to the non-convex cost function there is no guarantee that the mapped output results are similar even for different runs of the algorithm given identical input data, especially since the initialization of the map points is randomized. Despite the non-deterministic setup of the algorithm, the visual interpretation of different runs is easily compared by humans and far more simple to perform than automatic machine-based comparisons which are necessary for accurate evaluation of the algorithm. The primary problem studied in this work is how to analyze the stability of the t-SNE algorithm output. The proposed approach utilizes comparisons of several output maps as a whole and partially by clustered low-dimensional data points. This is performed by using smart home data of human activity

2577-1655/18/$31.00 ©2018 IEEE
DOI 10.1109/SMC.2018.00318

patterns as input data.

Learned manifolds can be aligned to build a correspondence between different disparate data sets and thereby provide knowledge transfer across the data sets from different domains [34]. Accordingly, knowledge transfer using machine learning (i.e. transfer learning), is gained during the transition from one learned domain to another, aiming to improve learning in the target task by leveraging knowledge from the source task [31]. In our work, it is hypothesized that learned manifolds (by a data-driven model such as t-SNE) from disparate data sets could be used for transfer learning. Therefore, it is crucial to investigate the stability of t-SNE output in order to use this algorithm further for aligning manifolds for the purpose of transfer learning.

Random Forests (RF) are commonly used as a one-class classifier, e.g. in order to model human behaviour patterns in an unsupervised fashion from sensor data collected in a smart home environment [18]. Then the proximity matrix from RF from such a model is fed to the t-SNE algorithm in order to map human behaviour patterns to a lower dimensional manifold. However, since t-SNE is a stochastic algorithm and since there is a large variance of t-SNE maps, a thorough analysis of the stability is required before applying TL.

The contribution of this work is the development of methods and tools for studying t-SNE output stability on smart home data used for modelling human activity patterns. The long-term goal of this research is to achieve automatic knowledge transfer between related data sets from different smart homes using manifold comparisons.

The rest of the paper is organized as follows. In Section 2, related work will be described and in Section 3, methods for testing stability of t-SNE for different runs is proposed. In Section 4, experiment results will be presented and discussed. Finally, the findings and opportunities of further research will be summarized in Section 5, Conclusions.

III. Related Work

Dimensionality reduction techniques have been widely used in many application domains to map HDD onto a low-dimensional manifold in order to produce a meaningful and visualizable representation. In [34] manifold alignment are used to construct connections (low-dimensional mappings) between different but related data sets by aligning their underlying learned manifolds for transferring knowledge across the data sets. Several comparative studies outlining the various DR techniques have been addressed in the literature [32], [33], [14]. In various research studies a set of quality assessment criteria has been considered based on local and global geometry preservation concepts [21], [7], [35]. However, these studies are mostly based on artificial data sets and the assessment of ability to find a good representation often relies on visual interpretation. To our knowledge an exhaustive empirical investigation analysis of the stability of non-linear DR techniques has not been carried out. Moreover, several studies have been done for investigating the performance of the non-linear DR methods in artificial and real tasks [2], [15], [24].

The stability of unsupervised DR techniques was studied by Garcia et. al who studied the parameter and data variations on several artificial data sets [5]. The study concluded by visual inspection that parameter variations in the resulting embeddings did not render instability.

Moreover, Laplacian Eigenmaps (LE) and Local Linear Embedding (LLE) were tested for stability when small or minor parameter variations which led to the conclusion that local methods (LE and LLE) are more likely to be affected by small modifications in parameter variations and therefore less stable than t-SNE.

Khoder et al performed a comparative study [11] to investigate the stability of unsupervised dimensionality reduction techniques using perturbed data of large images. The authors presented a new method for measuring the stability of non-linear and linear methods based on the noise variance at various scales. Results showed that PCA and MDS are limited by their linear character or are difficult to use when working on HDD because of their complexity.

A method for comparing DR techniques in terms of loss of quality with the aim to preserve the geometry of data sets has also been proposed [6]. Results revealed that the best results on all data sets are obtained by t-SNE.

Moreover, the accuracy of non-linear DR methods has been under review using real and synthetic data sets [32]. The experiment results show that non-linear DR methods perform well on the preferred synthetic data set, while this strong performance is not proved to extend to the real data sets.

Consequently, to the best of our knowledge, the stability analysis of t-SNE using manifold alignment on partial data points in the output maps has not be attempted before. Stability of low-dimensional manifolds plays a key role to align manifolds properly for the purpose of transfer learning. Thus, the contribution of this paper is important for transferring knowledge in a multi-smart home environment based on aligning manifolds.

IV. Methods & Proposed Approach

To measure the stability of t-SNE output an approach based on partially aligning t-SNE maps is proposed. The following steps (see Algorithm 1) constitutes the method for computing stability measures for a set of t-SNE maps. Firstly, T maps are produced by repeatedly computing t-SNE output maps with random initialization of low-dimensional data points given identical HDD input, X (see lines 2-4). Secondly, the resulting maps are pair-wise aligned into a modified target space, creating T^2 alignments (where $\frac{1}{2}T^2 - T$ are unique alignments due to the commutativity of the align operator: $\text{align}(\text{map\_list}[i], \text{map\_list}[j]) = \text{align}(\text{map\_list}[j], \text{map\_list}[i])$. The variability of these alignments is then evaluated by two measures of point-cloud alignment: mean disparity ($\overline{d_{ij}}$, where disparity is the sum of squared pair-wise differences between observations) and the estimated mean probability of obtaining the true corresponding data point within the aligned five nearest neighbors, $\overline{p^{5\text{NN}}}$.

Algorithm 1 Compute t-SNE output stability

```
Input: HDD, X
for i ← 0 to T do
    map_list ← compute_tSNE(X)
end for
for i ← 0 to T do
    for j ← 0 to T do
        (am_i, am_j) ← align(map_list[i], map_list[j])
        p_ij^5NN ← prob_5NN(am_i, am_j)
        d_ij ← disparity(am_i, am_j)
    end for
end for
Output: mean(p^5NN), mean(d_ij)
```

Besides giving an explanation of the t-SNE algorithm this section describes three different methods of alignment: Procrustes Analysis (PA), Local Procrustes Analysis (LPA) and the proposed extension Normalized Local Procrustes Analysis (NLPA).

A. t-SNE

t-SNE, introduced by van der Maaten and Hinton [19], is a nonlinear dimensionality reduction algorithm that maps high-dimensional data-points into a lower dimensional space. t-SNE utilizes embedding, which is constructed such that data-points in the vicinity of each other (i.e. similar data-points) in a high-dimensional space will remain in the vicinity by embed-points in a lower-dimensional space. Mainly, the t-SNE technique consists of two phases. Firstly, a joint probability over pairs of the data-points will be computed so that similar data points from the original (high-dimensional) data set have a large probability of being picked by each other for the embedding space. This results in dissimilar data-points to having a smaller probability of being picked. Accordingly, t-SNE is preserving the local geometry of the original high-dimensional data [19]. Secondly, over the map-points a probability distribution will be determined by t-SNE that fits data-point positions in the map in order to minimize the Kullback-Leibler divergence between both high and low dimensional distributions. Furthermore, t-SNE algorithm has originated from Stochastic neighbor Embedding (SNE) [8] and aimed to alleviate the main SNE challenges related to the thin tails of the normal distribution resulting in a data representation where even dissimilar data-points are *crushed* together which is known as the crowding problem [19]. t-SNE use a heavy-tailed distribution (Student-t distribution) to compute the similarity between two points. Moreover a symmetric version of the SNE cost function is implemented by t-SNE with simpler gradients.

B. Aligning by Procrustes Analysis

PA is one of the most popular rigid shape analysis algorithms. PA applies translation, scaling and rotation to two identically sized data sets in a multivariate Euclidean space to find the optimal alignment and to minimize the disparity [25]. Algorithm 2 shows the PA process. Firstly, PA translates the data sets to their origin. Secondly, PA normalizes the data sets using the Frobenius norm. Finally, PA rotates the second dataset to fit the first dataset in order to minimize the disparity. In this work PA is used to align the two-dimensional manifolds of the smart home dataset produced by t-SNE. PA performs well for data sets which are linear transformations of each other. However, PA works poorly on aligning non-linear smart home maps produced by t-SNE.

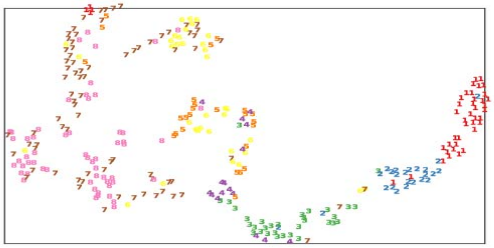


Fig. 1: t-SNE map for human behaviour patterns data

Algorithm 2 Procrustes Analysis (PA)

Input: M_1, M_2 — M_1, M_2 are the input t-SNE maps

$M_1 \leftarrow M_1 - \mu(M_1)$

$M_2 \leftarrow M_2 - \mu(M_2)$ — translate both data sets to their origin

$M_1 \leftarrow M_1 \,/\, \|M_1\|_F$ — scaling of M_1 and M_2

$M_2 \leftarrow M_2 \,/\, \|M_2\|_F$ — where $\|\cdot\|_F$ denotes Frobenius norm

$M_2 \leftarrow \text{Rotation}\,(M_1, M_2)$ — rotation M_2 with respect to M_1 to minimize disparity

disparity $\leftarrow \sum_{i=0}^{n} (M_{1_i} - M_{2_i})^2$ — measure the dissimilarity between the two data sets

Output: M_1, M_2, disparity, $\|M_1\|_F$, $\|M_2\|_F$, $\mu(M_1)$

C. Aligning by Local Procrustes Analysis

LPA was introduced by [20] to non-linearly align manifolds by using locally linear mappings. This algorithm comprises two main steps. Firstly, it follows a divisive approach to cluster datasets. The algorithm starts by considering a cluster of all data points and keeps on splitting into two sub-clusters recursively and terminates if the diversity of a cluster is below a predetermined threshold. Secondly, at each stage PA is applied to all clusters in the first data set and the corresponding cluster in the second dataset to compute disparity. If the disparity falls short of the threshold, the clustering process stops for these clusters at this stage. LPA uses K-means to create clusters. K-means is a non-deterministic algorithm that gives different results for different runs in terms of number of data points in each cluster and centroids location. For the analysis in this paper, K-means is replaced by a deterministic clustering algorithm, *hierarchical agglomerative clustering*, that plays a significant role in having a proper stability investigation of the t-SNE which is a non-deterministic algorithm. Figure 2 shows the process of clustering data. At initialization, each point cloud (which in this paper is the t-SNE map) is assigned to one cluster and the disparity of Procrustes Analysis is computed for the entire two maps. If the disparity is greater than a threshold, then the cluster will be divided into two sub-clusters if both clusters have at least two data points each. For this work, this condition has been modified to make sure that there are at least two distinct data points in order to meet the criteria of applying PA. This process will be applied to the sub-clusters for better manifold alignment with respect to disparity. Cluster alignments that offer a disparity lower than the threshold will not be further clustered. It is worth noting that disparities of LPA and PA cannot be fairly compared. This is due to the fact that the space of the data-clusters for the LPA and the space of the entire dataset for PA are in normalized for each cluster and map respectively.

For instance, ten disparities will be obtained if ten data-clusters in the first low-dimensional dataset are compared with their corresponding data-clusters in the second dataset. On the other hand, only one disparity will be obtained by applying PA on the two low dimensional dataset. Therefore, it is not reasonable to compare the disparities obtained using the PA and LPA. In this paper LPA is only applied to create clusters. In the next section, we propose an extension for the LPA in order to have the same space with the PA for the sake of the disparity comparison.

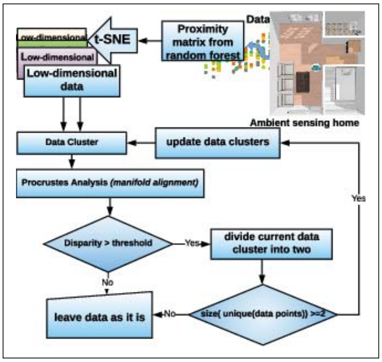


Fig. 2: Clustering data using LPA, this is from Halmstad Intelligent Home [17].

D. Aligning by Normalized Local Procrustes Analysis

In this work, LPA is used for creating clusters of t-SNE low-dimensional maps. We propose an extension to LPA which we call Normalized Local Procrustes Analysis (NLPA) in order to compare locally aligned clusters alignments to a complete map, which can be compared to PA as well as being used for stability analysis. Algorithm 3 shows the NLPA procedure. NLPA applies PA on each cluster and then normalizes the clusters so that the transformed data-points in each cluster are mapped back to the original space of the data after alignment.

Algorithm 3 Normalized Local Procrustes Analysis (NLPA)

1: **Input:** M_1, M_2 — M_1, M_2 are the input t-SNE maps
2: $n_c \leftarrow n$ — number of clusters created using LPA
3: **for** $i \leftarrow 0$ **to** n_c **do**
4: $\{M_1, M_2, -, \text{norm}_1, \text{norm}_2, \mu(M_1)\} \leftarrow \text{PA}(M_1, M_2)$ — call PA from Algorithm 2
5: $M_1 \leftarrow M_1 \cdot \text{norm}_1 + \mu(M_1)$
6: $M_2 \leftarrow M_2 \cdot \text{norm}_2 + \mu(M_1)$
7: $\text{templist}_1 \leftarrow M_1$
8: $\text{templist}_2 \leftarrow M_2$
9: **end for**
10: $\text{Map}_1 \leftarrow \text{templist}_1$
11: $\text{Map}_2 \leftarrow \text{templist}_2$
12: $\text{disparity} \leftarrow \sum_{i=0}^{n} \left(\text{Map}_{1,i} - \text{Map}_{2,i}\right)^2$ — dissimilarity between the two sets
13: **Output:** disparity

Normalization is done by multiplying the clusters with the norms of the aligned clusters that are produced by PA and then adding the mean of the first cluster, as shown in lines 5-6 of Algorithm 3. This normalization allocates the combined clusters of data-points to the same space as to original data. Finally, NLPA computes the disparity and estimated mean probability of obtaining the true corresponding data point within the aligned five nearest neighbors for the combined aligned clusters.

The changes of the proposed method NLPA compared to the LPA procedure can be summarized

1) Modification: Firstly the clustering algorithm is modified from k-means to agglomerative clustering.
2) Improvement: Secondly the creating clustering criteria is improved to have two distinct data points in each cluster and the threshold is minimized to render better alignment.
3) Extension: Finally, the NLPA is extended on LPA to normalize the transformed clusters in order to the combined clusters with NLPA and the whole dataset with PA have a same space.

V. Experimental Setup

A. Data

Data from Halmstad Intelligent Home [17] (HINT) was acquired for this work. HINT is a sensor-equipped home able to capture occupancy, movement, and interactions. In this home, 8 activities were performed by 11 individuals. The data were generated by an incoming stream of binary events from 37 sensors of the home. The events are represented by a particular ID of the triggered sensor, the associated binary state, and a time-stamp of when the event occurred. The observations of the data set are equal to the number of time windows over the measurement time period for the 11 individuals (310 observations). One observation (over a time window of 30 seconds) holds R number of features where R is equal to the time resolution (1 second) within a moving window times the number of events. The data pre-processing also involves a convolution over time in order to create a *memory* between sensor interactions over time (a process which has been demonstrated to be successful when modelling human behaviour [18]. Moreover are the observations being fed to a Random Forest which discriminates between the observations class and a class randomized from the data itself to train a one-class classifier. At last, the proximity matrix is extracted from the forests (a detailed explanation of the process can be found in [18]) and used as input for the t-SNE stability analysis.

Fig 1 is an example of a low-dimensional representation of human activities computed using t-SNE. The numbers in Fig 1 indicate the following activities 1. go to bed 2. use bathroom 3. prepare breakfast 4. leave house 5. get cold drink 6. office 7. get hot drink 8. prepare dinner.

B. Measurements & Parameter selection

To have an exhaustive analysis of the stability of the t-SNE, 100 low-dimensional data are mapped using 10 different configurations of t-SNE. Each configuration of t-SNE has a specific value of the t-SNE perplexity parameter from a set of numbers which are 5, 10, 15, 20, 25, 30, 35, 40, 45, 50 as a typical value between 5 and 50 recommended in the original paper of the t-SNE [19]. The default perplexity parameter is 20. The perplexity parameter controls how far to *look* for neighbors around a data-point and is related to the width of the t-distribution used in t-SNE. Each t-SNE map is compared with the rest maps for each t-NSE configuration to compute disparity and probability of obtaining the correct correspondences observation within five nearest neighbors for the transformed data by PA and NLPA. Therefore, 10000 comparisons are conducted in the experiments for each t-SNE configuration. Figure 3 shows the experiments procedure of the t-SNE stability analysis. Lastly, PA and NLPA are applied on the low dimensional manifolds to compute disparities. The experimental results show that the smaller disparity threshold renders better results. Therefore, the threshold is decreased from 0.001 of LPA to 0.00001 for NLPA. Besides, for every point in the first input dataset, the correct correspondence observation is found within five nearest neighbors in the second input dataset to compute the probability of obtaining the correct correspondence observation within five nearest neighbors.

VI. Results & Discussion

Recently, t-SNE stability has gained a lot of interest for projecting high-dimensional data into a low-dimensional manifold with the aim of transferring knowledge using manifold alignment. However, since t-SNE is a stochastic algorithm and since there is a large variance of t-SNE maps, a thorough analysis of the stability is required before applying TL.

Exhaustive scenarios are considered for an investigation about the t-SNE stability through manifold alignment using PA and NLPA. Table I shows the estimated disparity and probability of obtaining the correct correspondence observation within five nearest neighbors of the PA and the NLPA methods respectively for different values of perplexity. It turns out that

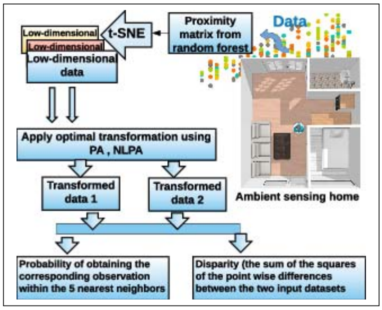


Fig. 3: Experiments process, ambient sensing home is from Halmstad Intelligent Home [17].

TABLE I: Estimated expected disparity and estimated probability of obtaining the correspondence observation within five nearest neighbors. After each estimate the standard error is given.

Perplexity	Disparity		Probability of obtaining the correspondence within the 5 nearest neighbors %	
	Mean (SE)			
	PA	*NLPA*	*PA*	*NLPA*
5	0.5954 (0.0206)	0.0010 (0.0002)	7.363 (0.0074)	98.541 (0.0099)
10	0.3135 (0.0248)	0.0006 (0.0001)	24.527 (0.0226)	98.541 (0.0099)
15	0.1097 (0.0167)	0.0007 (0.0001)	44.618 (0.0235)	98.476 (0.0099)
20	0.0291 (0.0022)	0.0006 (0.0001)	60.365 (0.0163)	98.450 (0.0099)
25	0.0296 (0.0026)	0.0007 (0.0001)	63.873 (0.0189)	98.485 (0.0099)
30	0.0447 (0.0058)	0.0009 (0.0002)	56.693 (0.0261)	98.435 (0.0099)
35	0.1483 (0.0152)	0.0029 (0.0005)	28.216 (0.0241)	98.126 (0.0098)
40	0.1648 (0.0129)	0.0046 (0.0006)	26.186 (0.0223)	97.869 (0.0098)
45	0.1699 (0.0123)	0.0049 (0.0006)	31.522 (0.0256)	97.637 (0.0098)
50	0.1501 (0.0116)	0.0056 (0.0006)	38.485 (0.0296)	97.744 (0.0098)

for all perplexity values considered the disparity values from using NLPA are less than 4% of the corresponding disparity value from using PA. In other words, the NLPA method is 25 times better than the PA method in terms of disparity. Also, the disparity from using PA decreases slightly for perplexity ranging from 5 to 20 while it increases for perplexity values from 25 to 50. The disparity values from using NLPA increases almost monotonically for all perplexity values considered.

With respect to the probability of obtaining the correct correspondence within the five nearest neighbors, the PA correct correspondence mean values commonly increase by increasing perplexity, especially as the perplexity value reaches 25. On the other hand, the probability of obtaining correspondence within the five nearest neighbors slightly decreases by increasing perplexity ranging from 5 to 50 for NLPA.

Figure 6 shows PA and NLPA histograms of disparity and probability of obtaining the correct correspondence observation within the five nearest neighbors where perplexity equals 20. Both the table and the histograms show that NLPA consistently outperforms PA for all t-SNE configurations, see Figures 4 and 5 respectively. Based on these results, it is concluded that t-SNE maps are stable locally for human behavior data that reflects the indicated property of t-SNE and which preserves local structure of data. Figure 7 shows the alignment of t-SNE maps that has been used to investigate the t-SNE stability using PA and NLPA. The alignment maps in Figure 7 indicate that the t-SNE maps are stable locally compared to globally aligned t-SNE maps using PA.

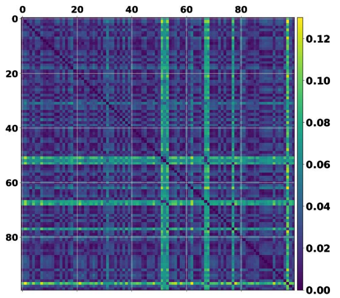


Fig. 4: PA disparity similarity matrix of size 100×100

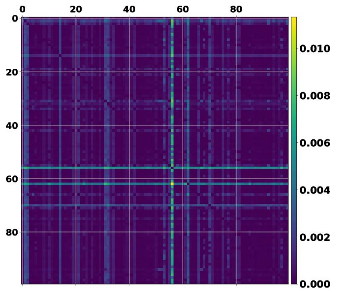
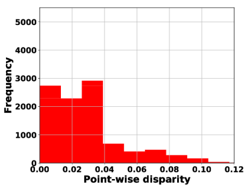
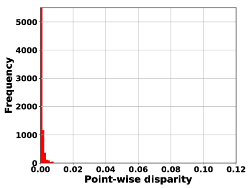


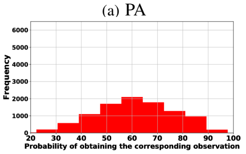
Fig. 5: NLPA disparity similarity matrix of size 100×100



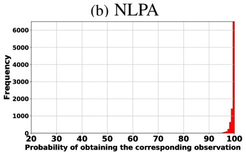
(a) PA



(b) NLPA

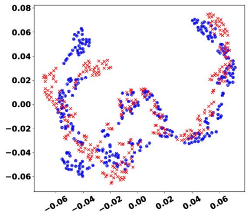


(c) PA

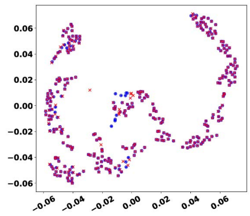


(d) NLPA

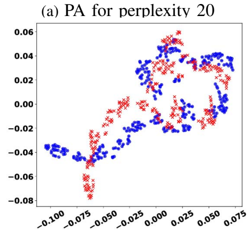
Fig. 6: PA and NLPA histograms of disparity and probability of obtaining the correspondence observation within five nearest neighbors where perplexity is equal to 20



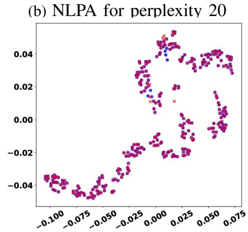
(a) PA for perplexity 20



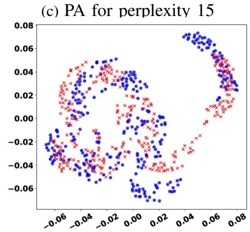
(b) NLPA for perplexity 20



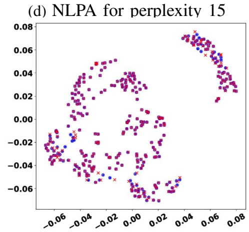
(c) PA for perplexity 15



(d) NLPA for perplexity 15



(e) PA for perplexity 35



(f) NLPA for perplexity 35

Fig. 7: Manifold alignment using PA and NLPA for three cases

VII. Conclusion

The t-SNE mapping stability of human activity patterns in smart homes via the analysis of reproducibility of low-dimensional manifolds is investigated. One could claim that any two data sets could be aligned via a non-linear mapping function with enough degrees of freedom. However, this study aims at analyzing parts ofa map in order to investigate the stability of t-SNE. Therefore, the choice of linear and local transformations gives human intuition about the stability of t-SNE. Procrustes Analysis (PA) is used for linearly aligning low-dimensional manifolds in order to compute disparity and correct correspondence observation within the five nearest neighbors. An extension to Local Procrustes Analysis called Normalized Local Procrustes Analysis (NLPA) is proposed to non-linearly align manifolds by using locally linear mappings. Experiments show that the disparity from using NLPA decreases by magnitudes compared to the disparity from using PA. Also, the probabilities of obtaining the correct corresponding observation within the five nearest neighbors from the second set of data points for each point in the first set of data points are radically increased by using NLPA compared to PA. For instance when the t-SNE parameter is 20, the disparity mean value decreases from 0.2913 in the case of using PA to a mere 0.00066 upon using NLPA. The probability of obtaining the correct corresponding observation within the five nearest neighbors for the same comparison, increases from 60.37 when using PA to 98.45 in case of using NLPA. In conclusion, NLPA outperforms PA by providing much better alignments for the low-dimensional manifolds on the same data set. This indicates that t-SNE low-dimensional manifolds are locally stable which is the main achievement of this study.

Future work will explore extensions of NLPA for aligning low-dimensional manifolds of disparate data sets. Then t-SNE low-dimensional manifolds of disparate data sets will be compared using NLPA to discover the common manifolds of the disparate data sets to be used for Transfer learning.

REFERENCES

[1] Mikhail Belkin and Partha Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural computation*, 15(6):1373–1396, 2003.

[2] Pedro Latorre Carmona and Reiner Lenz. Performance evaluation of dimensionality reduction techniques for multispectral images. *International Journal of Imaging Systems and Technology*, 17(3):202–217, 2007.

[3] Michael Thomas Wong et al. A high-dimensional atlas of human t cell diversity reveals tissue-specific trafficking and cytokine signatures. *Immunity*, 45(2):442 – 456, 2016.

[4] Francisco J García-Fernández, Michel Verleysen, John Aldo Lee, and Ignacio Díaz Blanco. Sensitivity to parameter and data variations in dimensionality reduction techniques. In *ESANN*, 2013.

[5] Francisco Javier García Fernández, Michel Verleysen, John A Lee, and Ignacio Díaz Blanco. Stability comparison of dimensionality reduction techniques attending to data and parameter variations. In *Eurographics Conference on Visualization (EuroVis)(2013)*. The Eurographics Association, 2013.

[6] Antonio Gracia, Santiago González, Victor Robles, and Ernestina Menasalvas. A methodology to compare dimensionality reduction algorithms in terms of loss of quality. *Information Sciences*, 270:1–27, 2014.

[7] Hisashi Handa. On the effect of dimensionality reduction by manifold learning for evolutionary learning. *Evolving Systems*, 2(4):235–247, 2011.

[8] Geoffrey E Hinton and Sam T Roweis. Stochastic neighbor embedding. In *Advances in neural information processing systems*, pages 857–864, 2003.

[9] Harold Hotelling. Analysis of a complex of statistical variables into principal components. *Journal of educational psychology*, 24(6):417, 1933.

[10] Chenping Hou, Changshui Zhang, Yi Wu, and Yuanyuan Jiao. Stable local dimensionality reduction approaches. *Pattern Recognition*, 42(9):2054–2066, 2009.

[11] Jihan Khoder, Rafic Younes, and Fethi Ben Ouezdou. Stability of dimensionality reduction methods applied on artificial hyperspectral images. In *International Conference on Computer Vision and Graphics*, pages 465–474. Springer, 2012.

[12] Kyoungok Kim and Jaewook Lee. Sentiment visualization and classification via semi-supervised nonlinear dimensionality reduction. *Pattern Recognition*, 47(2):758–768, 2014.

[13] John A Lee and Michel Verleysen. *Nonlinear dimensionality reduction*. Springer Science & Business Media, 2007.

[14] John A Lee and Michel Verleysen. Scale-independent quality criteria for dimensionality reduction. *Pattern Recognition Letters*, 31(14):2248–2257, 2010.

[15] Jiaxi Liang, Shojaeddin Chenouri, and Christopher G Small. A new method for performance analysis in nonlinear dimensionality reduction. *arXiv preprint arXiv:1711.06252*, 2017.

[16] Pedro E Lopez-de Teruel, Oscar Canovas, and Felix J Garcia. Visualization of clusters for indoor positioning based on t-sne. In *Proceedings of the Indoor Positioning and Indoor Navigation Conference*, 2016.

[17] Jens Lundström, Wagner O De Morais, Maria Menezes, Celeste Gabrielli, João Bentes, Anita SantAnna, Jonathan Synnott, and Chris Nugent. Halmstad intelligent home-capabilities and opportunities. In *International Conference on IoT Technologies for HealthCare*, pages 9–15. Springer, 2016.

[18] Jens Lundström, Eric Järpe, and Antanas Verikas. Detecting and exploring deviating behaviour of smart home residents. *Expert Systems with Applications*, 55:429–440, 2016.

[19] Laurens van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of Machine Learning Research*, 9(Nov):2579–2605, 2008.

[20] Ndivhuwo Makondo, Benjamin Rosman, and Osamu Hasegawa. Knowledge transfer for learning robot models via local procrustes analysis. In *Humanoid Robots (Humanoids), 2015 IEEE-RAS 15th International Conference on*, pages 1075–1082. IEEE, 2015.

[21] Deyu Meng, Yee Leung, and Zongben Xu. A new quality assessment criterion for nonlinear dimensionality reduction. *Neurocomputing*, 74(6):941–948, 2011.

[22] Andreas C Müller and Sarah Guido. Introduction to machine learning with python: a guide for data scientists. 2017.

[23] Nicola Pezzotti, Boudewijn PF Lelieveldt, Laurens van der Maaten, Thomas Höllt, Elmar Eisemann, and Anna Vilanova. Approximated and user steerable tsne for progressive visual analytics. *IEEE transactions on visualization and computer graphics*, 23(7):1739–1752, 2017.

[24] Hassan Radvar-Esfahlan and S-A Tahan. Performance study of dimensionality reduction methods for metrology of nonrigid mechanical parts. *International Journal of Metrology and Quality Engineering*, 4(3):193–200, 2013.

[25] Amy Ross. Procrustes analysis. *Course report, Department of Computer Science and Engineering, University of South Carolina*, 2004.

[26] Sam T Roweis and Lawrence K Saul. Nonlinear dimensionality reduction by locally linear embedding. *science*, 290(5500):2323–2326, 2000.

[27] Bernhard Schölkopf, John C Platt, John Shawe-Taylor, Alex J Smola, and Robert C Williamson. Estimating the support of a high-dimensional distribution. *Neural computation*, 13(7):1443–1471, 2001.

[28] Mohammed Shukur, T Rani, S Bhavani, G Sastry, and Surampudi Raju. Local and global intrinsic dimensionality estimation for better chemical space representation. *Multi-disciplinary Trends in Artificial Intelligence*, pages 329–338, 2011.

[29] Joshua B Tenenbaum, Vin De Silva, and John C Langford. A global geometric framework for nonlinear dimensionality reduction. *science*, 290(5500):2319–2323, 2000.

[30] Warren S Torgerson. Multidimensional scaling: I. theory and method. *Psychometrika*, 17(4):401–419, 1952.

[31] Lisa Torrey and Jude Shavlik. Transfer learning. *Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques*, 1:242, 2009.

[32] Laurens Van Der Maaten, Eric Postma, and Jaap Van den Herik. Dimensionality reduction: a comparative. *J Mach Learn Res*, 10:66–71, 2009.

[33] Jarkko Venna et al. *Dimensionality reduction for visual exploration of similarity structures*. Helsinki University of Technology, 2007.

[34] Chang Wang and Sridhar Mahadevan. Manifold alignment without correspondence. In *IJCAI*, volume 2, page 3, 2009.

[35] Peng Zhang, Yuanyuan Ren, and Bo Zhang. A new embedding quality assessment method for manifold learning. *Neurocomputing*, 97:251–266, 2012.

Appendix B

Paper II
Efficient activity recognition in smart homes using delayed fuzzy temporal windows on binary sensors

Rebeen Ali Hamad, Alberto Salguero Hidalgo,
Mohamed-Rafik Bouguelia, Macarena Espinilla Estevez,
Javier Medina Quero

IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS

Efficient Activity Recognition in Smart Homes Using Delayed Fuzzy Temporal Windows on Binary Sensors

Rebeen Ali Hamad, Alberto Salguero Hidalgo, Mohamed-Rafik Bouguelia, Macarena Espinilla Estevez, and Javier Medina Quero

***Abstract*—Human activity recognition has become an active research field over the past few years due to its wide application in various fields such as health-care, smart home monitoring, and surveillance. Existing approaches for activity recognition in smart homes have achieved promising results. Most of these approaches evaluate real-time recognition of activities using only sensor activations that precede the evaluation time (where the decision is made). However, in several critical situations, such as diagnosing people with dementia, "preceding sensor activations" are not always sufficient to accurately recognize the inhabitant's daily activities in each evaluated time. To improve performance, we propose a method that delays the recognition process in order to include some sensor activations that occur after the point in time where the decision needs to be made. For this, the proposed method uses multiple incremental fuzzy temporal windows to extract features from both preceding and some oncoming sensor activations. The proposed method is evaluated with two temporal deep learning models (convolutional neural network and long short-term memory), on a binary sensor dataset of real daily living activities. The experimental evaluation shows that the proposed method achieves significantly better results than the real-time approach, and that the representation with fuzzy temporal windows enhances performance within deep learning models.**

***Index Terms*—Activity recognition, fuzzy temporal windows, deep learning, temporal evaluation.**

I. Introduction

SMART homes make use of distributed sensor networks with high processing capabilities and low power consumption, and have the ability to record information about the behavior of an inhabitant who interacts with the environment [1], [2]. These environments are made to perceive the user's context in order to help people in their daily living activities and provide smart solutions to address some of the problems associated with the growing size of the population [3], [4].

Activity Recognition (AR) in smart homes is an active research topic as well as a useful means to assess the circumstances of a person's daily living. It aims to predict human activities within the smart environment by continuously observing a series of actions and environmental conditions. In the health context, AR plays an important role in assessing the patient's condition. Examples of this include the identification of abnormal behaviours for elderly dementia sufferers [5], and the early detection of Alzheimer disease [6].

In data-driven AR, an activity model is built by training machine learning models on sensor data corresponding to various activities of the inhabitant. Then, a testing stage is carried out to evaluate the built model on a stream of sensor activations [7]. In particular, binary sensors are deemed to be one of the useful means for unobtrusive monitoring within the home environment. One of the advantages of such data-driven AR approaches consists in the capability of handling uncertainty and temporal information. Among these approaches, Deep Learning methods have been used for activity recognition, and evaluated on binary sensors [8] by using Convolutional Neural Networks (CNN) [9] and Long Short-Term Memories (LSTM) [10].

Existing approaches for activity recognition in smart homes perform predictions in real time based on sensor activations that precede the evaluation time. Due to only relying on previous activations, such real-time approaches may lack precision in recognizing some daily life activities. To overcome this problem, it becomes necessary to know which sensor activations are generated later, since the activity to be recognized will depend on the subsequent sensors. In order to illustrate the limitation of the real-time approaches for activity recognition, let's consider the following two scenarios. In the first scenario, if a binary sensor deployed on the front door generates an open activation, the real-time approach will recognize the activity as "the inhabitant has left the house". However, it can happen that the inhabitant opens the main door only to talk to another person in the entrance of the house and returns to the house again without leaving it. In this scenario, it is important to consider the activations generated after opening the door. Indeed, if nothing happens after a short time (e.g., two minutes), then the inhabitant has performed the activity "leaving the house"; however, if sensor activations

Manuscript received February 14, 2019; revised April 23, 2019; accepted May 15, 2019. Date of publication May 22, 2019; date of current version February 6, 2020. This work was supported by the REMIND project Marie Sklodowska-Curie EU Framework for Research and Innovation Horizon 2020 under Grant 734355. *(Corresponding author: Rebeen Ali Hamad.)*

R. A. Hamad and M.-R. Bouguelia are with the Department of Intelligent Systems and Digital Design, Halmstad University, Halmstad 30118, Sweden (e-mail: rebali@hh.se; mohbou@hh.se).

A. S. Hidalgo is with the Department of Computer Science, University of Cádiz, Cádiz 11001, Spain (e-mail: alberto.salguero@uca.es).

M. E. Estevez and J. M. Quero are with the Department of Computer Science, University of Jaén, Jaén 23071, Spain (e-mail: mestevez@ujaen.es; jmquero@ujaen.es).

Digital Object Identifier 10.1109/JBHI.2019.2918412

are generated during the next two minutes after opening the door, then the inhabitant has only opened the door without leaving the house. The second scenario that illustrates the problem of real-time activity recognition approaches is as follows. An inhabitant goes to the kitchen and opens the refrigerator. Until this action is followed up with interactions with other objects (and therefore the activation of other sensors) it is not possible to recognize exactly whether the inhabitant has gone to open the refrigerator to have a snack and eat, or to perform some other activity, for example, to take out certain products from the refrigerator for cooking. Therefore, the accuracy of a real-time (or online) activity recognition approach is confronted with the use of oncoming sensor activations with a slight delay.

Although real-time approaches enable permanent interaction with users in smart environments, it has a lower recognition rate [11], [12]. This is not admissible in some critical cases such as diagnosing dementia, as it requires more accurate activity recognition to correctly detect abnormal behaviours in the inhabitant, [13]. These cases, although less frequent, prevent AR from being a high-precision tool for assessing the conditions of inhabitants in smart homes.

The contribution of this paper takes the above-mentioned issues into consideration. We propose a data-driven approach that aims to increase precision and sensitivity in daily activity recognition by means of i) delaying the activity recognition, ii) extracting representations of binary sensor activations that occur before and after the time where the prediction is made, iii) evaluating Deep Learning methods for classification, and iv) analyzing the impact of the delayed AR process on precision and sensitivity.

The remainder of this paper is structured as follows: Section II describes previous work in activity recognition related to our proposal, highlighting the main novelty of the proposed methodology. Section III presents the proposed methodology for activity recognition as a high-precision tool. In Section IV, the proposed methodology is evaluated on a popular dataset with daily activities developed in a real smart environment [4]. Finally, a conclusion and future works are proposed in Section V.

II. Related Works

Activity recognition based on the use of binary sensors is a useful approach to evaluate the conditions of daily living within a sensorized environment in an unobtrusive manner. Binary sensors are small and light devices which are installed in everyday objects to register human interaction, such as, passive infrared sensors, motion detectors, contact switches, break-beam sensors, and pressure mats [14]. They have been proposed as suitable devices for describing daily human activities within a smart environment setting. Their main advantages are that they are: i) easy to install, ii) small in size, iii) low-cost and iv) minimally invasive in comparison to videos and microphones [15].

The real-time capabilities have become a key challenge in activity recognition to offer a tool that meets real-world conditions, and enable *adequate assistance services* which can be offered within Ambient Assisted Living [16]. Previously, a combination of human-defined binary sensor features, such as the last activation within a fixed time period, or the current raw activation, were proposed as suitable representations in real-time AR using windowing approaches [17]. In [18], the combination of multiple and incremental fuzzy temporal windows has shown an increase in the performance of real-time AR compared to previous binary sensor representations. Most of the algorithms carry out successful analyses on stored datasets, but in real time, data should be quickly pre-processed for proper action to be taken if needed [19]. Using smart home settings, activity recognition systems attempt to recognize daily routine activities. [20] recognizes a set of activities such as *Breakfast, Brushing teeth, Drinking, Showering*, using machine learning algorithms including Hidden Markov mode (HMM), condition random field (CRF), hidden semi-Markov models (HSMM) and semi-Markov conditional random field (SMCRF). [21] recognizes daily living activities *bedtoilet, transition, leaving, eating, cooking, relaxing, sleeping and working* using Naive Bays, HMM, SVM, CRF, SVMS, which consistently obtains 84% average weighed accuracy. The increasingly large amounts of smart home datasets have lead to the use of deep learning approaches. [22] and [8] applied Long Short Term Memory (LSTM) on three real world smart home datasets where the number of activities are 10, 13, 16 for the first, second and third houses respectively. The results of LSTM outperforms Naive Bays, HMM, HSMM, CRF in terms of accuracy and performance. However, the features of the datasets for the input models were computed using only equally-sized temporal windows and the techniques of separating data for training and testing phases to avoid over-fitting are not described. Moreover, the accuracy of the individual classes are not computed, rather only average accuracy of the models is shown and the training time of each model is not computed. Hence, the recognition of which activities are improved is not clear.

As explained previously, since real-time activity recognition approaches use only the preceding sensor activations, they can perform poorly in recognizing some activities [11]. From an operational perspective, there are situations where one needs to achieve higher precision while tolerating a short time delay (i.e. real time is not necessary). For example, when monitoring a person with early dementia using binary sensors, it is necessary to provide high accuracy even if it means introducing a short delay in recognizing the activity. The proposed data-driven approach differs from the previously proposed ones by taking into consideration the aforementioned issues, and aiming to further increase the precision and sensitivity of the activity recognition by means of:

- 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. In order to avoid the human configuration of FTWs we have modeled their shapes with the Fibonacci sequence, which has been defined to model incremental sequences in a harmonic way under the fields of mathematics, science, and engineering [23].
- Evaluating the temporal models using a Deep Learning approach has presented significant differences compared to other contexts of activity recognition [24]. Hybrid models combining CNN and LSTM are evaluated as Deep Learning architectures. This is compared with other models

under delayed representation of sensor activations. Moreover, we have evaluated the configuration for LSTM and CNN described in [8] based on equal-sized temporal windows for extracting features, showing the impact of FTWs in extracting features to represent suitable long-term performance of sensor activations with a low number of features and achieving better F1-score.

- Analyzing the impact of delaying the activity recognition by using oncoming sensor activations in real daily activities. Specifically, showing the impact of including oncoming sensor activations in activity recognition to improve the recognition of some rare activities such as *leaving, snack, grooming and dinner* that have been poorly recognized using only real-time activity recognition.

Consequently, to the best of our knowledge, including and evaluating both preceding and oncoming sensor activations for activity recognition has not been attempted before. Human activity recognition considering oncoming sensors activation plays a key role in recognizing activities properly for the purpose of enhancing elderly care alert systems. In concrete terms, in building robust AR for diagnosing patients based on their user activities in smart homes in a highly accurate way.

III. Methodology

In order to increase the precision and sensitivity of activity recognition in smart homes, in this paper we propose a new data-driven methodology that applies multiple and incremental fuzzy temporal windows to extract features from both preceding and partial oncoming sensor activations. This results in a sequence of temporal features that boosts learning using LSTM sequence models as well as CNN.

A. Representation of Activities and Binary Sensors

A set of binary sensors is represented by $S = \{S_1, \ldots, S_{|S|}\}$ whose related set of daily activities is represented by $A = \{A_1, \ldots, A_{|A|}\}$, where $|S|$ and $|A|$ are the number of sensors and daily activities respectively. Each binary sensor and each daily activity are described by a set of binary activations within a range of time, defined by a starting and ending point of time as shown by Eq. (1):

$$\begin{aligned} S_i &= \{S_{i_0}, \ldots, S_{|S_i|}\}, S_{i_j} = \{S^0_{i_j}, S^+_{i_j}\} \\ A_i &= \{A_{i_0}, \ldots, A_{|A_i|}\}, A_{i_j} = \{A^0_{i_j}, A^+_{i_j}\} \end{aligned} \quad (1)$$

where i in $|S_i|$ and $|A_i|$ is the total number of activations for a given binary sensor S_i and a daily activity respectively, and ii) $S^0_{i_j}, S^+_{i_j}$ the starting and ending point of a given time of activation respectively. Finally, the timeline is determined by the range of time $T = \{min(S^0_{i_j}), \max(S^+_{i_j})\}$.

B. Feature Sequence of Binary Sensors With Incremental Fuzzy Temporal Windows

In this Section, a binary-sensor representation by means of multiple and incremental FTWs is described. A FTW describes a membership degree from a given time t^* to a point of time t_i by means of their temporal distance $\Delta t^*_i = t^* - t_i, t^* > t_i$. In case $\Delta t^*_i > 0$, we are evaluating a preceding point of time, and in case $\Delta t^*_i < 0$ an oncoming point of time. A FTW T_k is defined by a fuzzy set $T_k(\Delta t^*_i)$, which is characterized by a membership function $\mu_{\tilde{T}_k}(\Delta t_i)$. In sake of simplicity we note $T_k(\Delta t^*_i)$ instead of $\mu_{\tilde{T}_k}(\Delta t^*_i)$.

In this work, each FTW T_k is described by a fuzzy set characterized with a membership function and its shape corresponds to a trapezoidal function $T_k(\Delta t^*_i)[l_1, l_2, l_3, l_4]$. The well-known trapezoidal membership functions are defined by a lower limit l_1, an upper limit l_4, a lower support limit l_2, and an upper support limit l_3 according to Eq. (2):

$$TS(x)[l_1, l_2, l_3, l_4] = \begin{cases} 0 & x \leq l_1 \\ (x - l_1)/(l_2 - l_1) & l_1 < x < l_2 \\ 1 & l_2 \leq x \leq l_3 \\ (l_4 - x)/(l_4 - l_3) & l_3 < x < l_4 \\ 0 & l_4 \leq x \end{cases} \quad (2)$$

Once a FTW T_k is defined, the activation degree of a binary activation S_{i_j} from a sensor S_i at evaluated time t^* is computed by Eq. (3).

$$T_k(S_{i_j}, t^*) = \begin{cases} \max(T_k(\Delta t^*_i))\forall t_i \in S_{i_j} & \exists t_i \in [S^0_{i_j}, S^+_{i_j}] \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

So, in order to obtain the degree of the binary activation S_{i_j} within the FTW T_k, we compute the maximal degree of T_k for each point t_i of time within the time interval from sensor activation.

In Fig. 1, some examples of activation degree from a FTW T_k and three different binary activations S_{i_j} are shown.

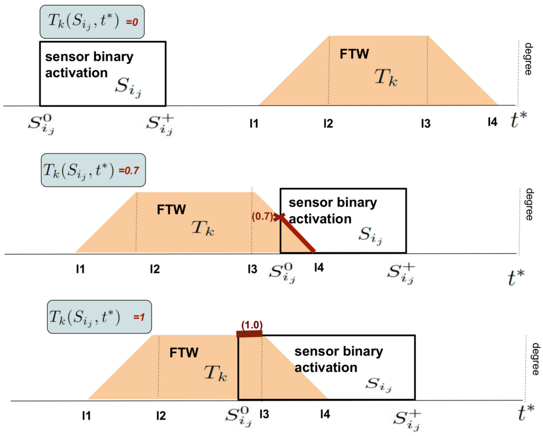


Fig. 1. Activation degree from a FTW T_k and different binary activations S_{i_j}, which is computed as the maximal degree of T_k for each point t_i of time within the sensor activation. Top) No overlapping between T_k and S_{i_j}, middle) partial overlapping between T_k and S_{i_j} and bottom) partial overlapping within support limits of T_k.

Next, in order to evaluate a FTW T_k and a sensor S_i in the complete timeline, we aggregate $T_k(S_i, t^*)$ computing the maximal value in the time-line according to Eq. (4).

$$T_k(S_i, t^*) = \max(T_k(S_{i_j}, t^*)), \forall S_{i_j} \in S_i \tag{4}$$

This representation includes an updated model from previous works [18], [25], where a fuzzy aggregation of the sensor activation in wearable and binary sensors was initially proposed. In this work, we have simplified and improved the computing of the activation degree by evaluating the sensor activation and FTW in a continuous way using Eq. (3). In Fig. 2, a representation of preceding and oncoming FTWs defined by the Fibonacci sequence in an sample timeline is shown.

C. Preceding and Oncoming Sensor Activation With FTWs

In this work multiple and incremental FTWs are defined to collect i) long-term to short-term temporal sensor activations, and ii) the preceding and oncoming sensor activations. To do so, first, the maximal temporal size is defined for evaluating preceding and oncoming sensor activations, which are defined by $L^*_{-;}$ and L^*_{+} respectively. We note:

- $L^*_{-;}$ must include a value which determines a wide range of preceding evaluation of sensors, as it has been demonstrated that long-term activations of sensors are key to improve AR [18], [26]. For example, the difference between some activities carried out in the kitchen, such as breakfast, lunch, dinner or snack, is determined by the evaluation of the bed sensor activation in short, middle and long term. In addition, due to temporal aggregation with incremental FTWs, defining hours of sensor evaluation are reduced to a short number of features [18]. The definition of preceding FTWs is determined by $T_k(\Delta t^*_i)$ where $\Delta t^*_i > 0, t^* - L^*_{-;} \leq t_i \leq t^*$.
- L^*_{+} represents the maximal time to evaluate oncoming sensor activations. A long-term evaluation will improve the results. For this reason, in the same way as $L^*_{-;}$, with preceding sensor activations, the results also improved. It generates a delay in comparison with real-time AR. However, the results obtained will be more precise. The more time is spent to evaluate oncoming sensor activations, the more time is needed to obtain the recognized activity. So, t^+ defines the evaluating time which is defined as the point of time where the AR evaluates the time t^*. The definition of oncoming FTWs is determined by $T_k(\Delta t^*_i)$ where $\Delta t^*_i < 0, t^* \leq t_i \leq t^+$.

Second, in order to generate multiple and incremental FTWs from $L^*_{-;}$ and L^*_{+} in a simple manner, the following process is proposed:

- A set of incrementally ordered evaluation times $L = \{L_1, \ldots, L_{|L|}\}, L_{i-1} < L_i$, which define the limits of the trapezoidal functions, are calculated according to the temporal window index $T_k = T_k(\Delta t^*_i)[L_k, L_{k-1}, L_{k-2}, L_{k-3}]$.
- L is defined by the Fibonacci sequence [27] $L = \{1, 2, 3, 5, 8, \ldots\}$, which has been previously shown as a successful sequence for defining FTWs without requiring expert knowledge definition [26].
- L_+ and $L_{-;}$ are established from the maximal temporal size $L^*_{-;}$ and L^*_{+} respectively. In both cases, the number of FTWs defined by Fibonacci is a set to the closer value of $L^*_{-;}$ and L^*_{+} in the sequence L. For example, if $L^*_{+} = 240m$, $L_+ = \{0m, 1m, 1m, 2m, 3m, 5m, \ldots, 144m, 240m\}$.

D. Configuring Sequence Features With FTWs, Time Segmentation and Labeling

The use of FTWs T_k defines a sequence representation of preceding and oncoming features for each sensor S_i in a given point of time t_i. For evaluation purposes of the AR model, the activity is estimated for each evaluation time step of the timeline [17].

So, the timeline $T = \{min(S^0_{i_j}), max(S^+_{i_j})\}$ is divided in time steps t^* defined by a temporal granularity ΔT where $t^{*-1} = t^* - \Delta T, t^{*+1} = t^* + \Delta T, \forall t^*$, which configure the points of time where the AR is evaluated.

For each evaluated time step t^* and sensor S_i, the preceding and oncoming FTWs are computed to determine the sequence features. It is worth noting that the FTW parameters are defined according to distance to the evaluated time t^*, so a sliding window configuration [28] is developed in the timeline T to compute the features.

Finally, in order to label each evaluated time step t^*, the activity A_i is selected, which is developed if and only if $\exists t^* \in [A^0_{i_j}, A^+_{i_j}] \forall A_{i_j}$.

IV. Experimental Setup and Evaluation

A. Data

Data from two real intelligent homes A and B in which participants perform their daily routine were acquired for this work [4]. These two homes are equipped with sensors that are able to capture the movements and interactions of the resident. In home A, 9 human daily activities that were performed in 14 days over a period of 19,932 minutes were described by an incoming stream of binary events from 12 sensors in the home. In home B, 10 human daily activities that were performed in 22 days over a period of 30,495 minutes were described by 12 binary sensors. The timeline of the activities is segmented in time slots using the window size Δ t=60 s, based on the standard reference from [8], [17], [24]. The activities of homes A and B are *Breakfast, Grooming, Leaving, Lunch, Showering, Sleeping, Snack, Spare Time, Toileting*; in addition to these, home B has the activity *Dinner*.

The leave-one-day-out cross-validation is used for evaluation, where a single day of the activities is used for the test set and the rest of the days are used for the training set. This process is repeated until the data from all the days is involved in the training set and the testing set [15]. The average F-score is computed from the results of the cross-validation as done in [8], [24].

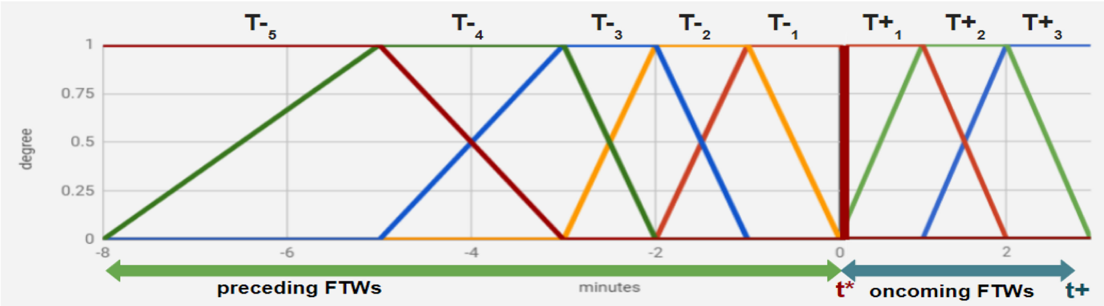


Fig. 2. Example of preceding and oncoming FTWs $L_+ = \{0,1,1,2,3\}$ and $L_{-;} = \{0,1,1,2,3,5,8\}$. Evaluation time t^+ and the evaluated time t^* in the timeline are shown.

Since the classes of the datasets are imbalanced, the sampling approach is used by randomly oversampling observations of the minority class and undersampling observations of the majority class to have an equal number of samples (5000 samples) for each class. This allows us to create a more balanced dataset and avoid having models biased toward one class or the other [18], [29].

B. Feature Vector

Features are computed by applying 15 FTWs on the raw data from all 12 binary sensors in each minute for both datasets. The datasets A and B have 19,932 and 30,495 samples respectively, where each sample represents one minute of data with $12 \times 15 = 180$ features. The resulting datasets are used for real-time activity recognition. Feature extraction based on FTWs is evaluated and compared with Equally-sized (1 minute) temporal windows (ESTWs) [8] and Raw Last sensor Activation (RAW) in one-minute windows [26]. Regarding the feature extraction with delays in time, different delays in time are considered, which are 5 minutes, 20 minutes, 1 hour, and 4 hours. For example, when a 5-minutes delay is used over real time, 4 FTWs in addition to 15 FTWs are computed, where the 4 FTWs represent the 5-minute delay in time, hence, the number of features becomes $12 sensors \times 19 FTWs$. For other delays in time, 20 minutes, 1 hour, and 4 hours, the number of features for each sensor becomes 22, 24, and 27 respectively.

C. Model Selection and Architecture

The models that have been used in this study are described below.

- Long-Short Term Memory (LSTM): a type of recurrent neural network (RNN) that includes a memory and is designated to learn from sequence data, such as sequences of observations over time. LSTM is most widely used in natural language processing and speech recognition that can model temporal dependence between observations [30]. LSTM has obtained satisfying results in activity recognition [31], [32]. Hence, in this study LSTM is designed to be used in the experiments by stacking a LSTM layer with 40% dropout rate and 0.001 learning rate followed by a fully-connected, i.e., a dense layer and a soft-max layer. For all the models in this study the batch size and training epochs are equal to 10, which is a total of 100 batches during the entire training process. Commonly, large batch sizes result in quicker progress in training but mostly do not always converge as fast. On the other hand, smaller batch sizes train slower but could converge faster, therefore it is mostly an independent problem [33]. Regarding the 40% dropout, which is a regularization technique for preventing deep learning models from overfitting [34], the dropout ignores randomly selected neurons during the training phase. Those ignored neurons are temporally removed on the forward pass and their weights are not updated on the backward pass.
- Convolutional Neural Network (CNN): used in the experiments because CNN is competent in extracting features from signals. CNN has obtained promising results in image classification, text analysis and speech recognition [32]. CNN has two advantages for human activity recognition which are local dependency and scale invariance. Local dependency refers to the nearby observations in human activity recognition that are likely to be correlated, while scale invariance means the scale is invariant for different paces or frequencies. CNN can learn hierarchical data representations which leads to rendering promising results in human activity recognition [32]. In this study, a one dimensional (1D) CNN architecture is developed which can extract local 1D sub-sequences from the sequence data. The 1D CNN could be competitive with RNNs on some sequence-processing applications such as audio generation and machine translation with a cheaper computation cost compared to RNN. The model is designed by stacking two convolutional layers with 40% dropout rate and 0.001 learning rate followed by a max-pooling layer and followed by a fully-connected, i.e., a dense layer and a soft-max layer.
- Hybrid model: since input sub-sequences are processed by 1D CNN independently, unlike LSTM, the sub-sequences are not sensitive to the time step order. However, many

convolution layers and pooling layers could be stacked to recognize longer-term patterns. This leads to long chunks of the original inputs to be considered by the upper layers. Yet this is not a fairly strong way to induce order sensitivity, when order sensitivity is key to activity recognition, since RNNs for processing very long sequences are highly expensive, while 1D CNNs are cheap. Hence, designing a hybrid model by combining the speed and lightness of CNNs with the order-sensitivity of LSTM consists in using a 1D CNN as a pre-processing step before an LSTM. This strategy is particularly important because the 1D CNN turns the long input sequence data into much shorter sequences of high-level features. The sequence data of extracted features by 1D CNN then becomes the input to the LSTM part of the network. The hybrid model is the combination of two deep learning models, particularly Convolutional neural network and Long-Short Term Memory (CNN LSTM). CNN LSTM is able to systematically learn feature representation and model the temporal tendencies between their activations [24]. The model is designed by stacking a convolutional layer that was followed by a max-pooling layer. Then, a LSTM Layer with 40% dropout rate and 0.001 learning rate is followed by a fully-connected, i.e., a dense layer and a soft-max layer.

D. Measurement

The following well-known metrics are used to evaluate the models.

- Precision represents the proportion of correct observations of a class to the entire observations classified as that class, with high precision indicating low false positives. Precision is equal to $\frac{TP}{TP+FP}$:
- Sensitivity (recall) shows the proportion of observations correctly classified as a given class to the actual total observations in that class calculated using the formula $\frac{TP}{TP+FN}$:
- F1-score (F1-sc) shows an insight into the balanced between precision ($\frac{TP}{TP+FP}$) and sensitivity (recall) ($\frac{TP}{TP+FN}$). This metric is also used in Activity recognition [17].

E. Results and Discussion

In this section, results of different models are shown and discussed based on different feature extraction approaches. In a real-time activity recognition scenario, only past sensor data are considered for each evaluation time. Evaluation time is the time when the decision-making of each activity has been developed. Here, two different feature extraction approaches based on several machine algorithms have been compared with FTWs. First approaches described in [4] included an ad-hoc feature vector with Raw and Last sensor Activation (RLA) within a one-minute window as a suitable configuration for learning AR with C4.5 and SVM for the homes A and B. Second, we include the results from approach [8], where equally-sized temporal windows (ESTWs) of one minute sensor activation are described as a suitable feature extraction approach for AR binary sensors with LSTM and CNN models. Table I shows the results of the F1-score and training time for real-time recognition of different models based on extracting features with ESTWs, RLA and FTWs. The training time is the average of 10 runs of the training model. The F1-score results of the models based on FTWs are significantly higher than the results of the models based on ESTWs and RLA. Therefore, FTWs have been used for this paper as a contribution to AR research, with different delays in time in addition to real time to improve the recognition process. To predict what activity has been performed in a specific time *T*, different time delays of oncoming sensors after the time *T* in addition to the preceding sensor activations of the time *T* are included in the feature extraction based on FTWs. For example, now (Evaluation time), we evaluate what activity was developed 4 hours ago (Evaluated time). Evaluated time. It is the time which is evaluated by the classifier to recognize which activity has been developed in this point of time based on the preceding and oncoming sensor data. In the case of real time, evaluation time is equal to evaluated time. In the case of delays in time that consider oncoming sensor data, evaluation time is higher (delayed) from evaluated time.

In the scenario where the activity recognition is delayed, preceding sensor activations with different time delays, particularly 5 minutes, 20 minutes, 1 hour, and 4 hours, are tested to improve the recognition process. The results show that delaying time with LSTM in decision-making leads to building more accurate models. The results are significantly improved when considering oncoming sensor activations and increasing delay in the evaluated time. In house A, for example, the total results of F1-score of the model in real time is 89.05, while the results of the model are improved notably, up to 96.44, when considering oncoming sensor activations. Tables II and III show the results of the F1-score and training time of LSTM, CNN, and the hybrid CNN LSTM based on FTWs from home A and B respectively.

The results indicate that the F1-score of the models improves substantially by increasing time delays with a slight increase of training time. This means that delaying the decision-making of human activity recognition yields better and more accurate models. In addition to the F1-score, a precision summary of the models is also shown in Fig. 3 for homes A and B respectively. The figures show that the precision of the models is increased when considering oncoming sensor activations over real time.

Finally, the results of the models with 4-hour time delays are based on extracting features using three approaches. Firstly, Raw and Last Next Activation (RLNA) for learning AR with C4.5 and SVM. Secondly, Equally-sized Temporal Windows (ESTWs) [8] for learning AR with LSTM. Thirdly, Fuzzy Temporal Windows (FTWs) for learning AR with LSTM. Table IV shows the results of the F1-score and training time of the models based on extracting features with ESTWs, RLNA and FTWs. The results indicate that LSTM based on FTWs has obtained the highest F1-score with suitable and reasonable training time compared to LSTM based on ESTWs or C4.5 and CVM based on RLNA.

In summary, the proposed method of this paper has enhanced the models for recognizing all the activities performed in homes A and B while maintaining a low time cost. We highlight that

TABLE I
F1-SCORE AND TRAINING TIME FOR REAL-TIME RECOGNITION OF DIFFERENT MODELS BASED ON EXTRACTING FEATURES WITH EQUAL-SIZED TEMPORAL WINDOWS (ESTWs), RAW AND LAST ACTIVATION (RLA) AND FUZZY TEMPORAL WINDOWS (FTWs)

(a) Home A

Activity	ESTWs		RLA		FTWs		
	LSTM	**CNN**	**C4.5**	**SVM**	**LSTM**	**CNN**	**CNN LSTM**
Breakfast	72.35	69.75	82.71	81.75	95.08	92.69	94.85
Grooming	52.10	49.13	75.21	75.22	65.96	56.25	69.31
Leaving	88.66	87.98	98.56	98.56	88.97	94.31	87.18
Lunch	77.77	84.00	86.49	82.26	96.20	95.76	95.27
Showering	91.59	79.50	96.43	96.42	94.47	93.23	94.47
Sleeping	99.82	99.66	99.99	99.98	99.68	95.42	99.62
Snack	0.0	0.0	00.00	40.00	90.23	89.62	91.63
Spare Time	96.24	96.19	98.53	98.54	100.00	100.00	100.00
Toileting	32.93	29.26	35.19	38.09	70.37	64.62	63.39
Total	**67.94**	**66.16**	**74.79**	**79.43**	**89.05**	**85.38**	**88.41**
Train-time	5.1m	3.9m	4.1s	1.92s	3.65m	2.1m	3.1m

(b) Home B

Activity	ESTWs		RLA		FTWs		
	LSTM	**CNN**	**C4.5**	**SVM**	**LSTM**	**CNN**	**CNN LSTM**
Breakfast	41.83	46.34	41.64	29.36	92.32	89.43	91.08
Grooming	61.65	59.92	90.79	90.31	74.33	69.89	72.61
Leaving	70.08	71.16	98.66	98.68	88.59	88.25	89.11
Lunch	0.0	01.38	13.20	13.28	94.66	92.51	94.34
Showering	74.94	52.79	95.54	96.00	88.6	86.86	88.20
Sleeping	96.78	96.68	98.41	98.34	99.38	99.45	99.39
Snack	0.05	09.00	13.54	17.26	83.57	79.54	84.72
Spare Time	88.06	87.48	90.97	91.15	96.98	96.00	96.82
Toileting	27.70	23.25	88.00	84.78	60.82	56.37	59.68
Dinner	25.30	22.50	11.11	00.00	79.15	74.64	75.28
Total	**51.84**	**49.83**	**64.19**	**61.99**	**85.83**	**83.29**	**85.12**
Train-time	6.4m	5.1m	5.8s	3.2s	4.8m1	2.9m	4.2m

TABLE II
F1-SCORE AND TRAINING TIME (MINUTES) OF LSTM, CNN, AND CNN LSTM WITH DIFFERENT DELAYS IN TIME BASED ON FTWs FROM ORDONEZ HOME A

Activity	LSTM				CNN				CNN LSTM			
	5 Minutes delay	**20 Minutes delay**	**1 Hour delay**	**4 Hours delay**	**5 Minutes delay**	**20 Minutes delay**	**1 Hour delay**	**4 Hours delay**	**5 Minutes delay**	**20 Minutes delay**	**1 Hour delay**	**4 Hours delay**
Breakfast	95.18	95.87	96.87	96.95	92.91	93.24	93.29	94.69	95.22	95.87	96.78	96.69
Grooming	66.85	67.96	71.24	86.19	66.71	71.28	76.81	83.68	72.76	75.28	81.26	86.19
Leaving	91.17	91.85	94.15	99.85	94.31	96.71	96.83	99.49	97.41	98.21	98.78	99.85
Lunch	96.88	97.96	97.82	99.00	95.76	96.73	97.22	98.92	96.11	96.42	96.57	99.07
Showering	95.21	95.53	96.89	98.21	93.23	93.53	95.11	96.49	94.52	95.62	95.79	99.98
Sleeping	99.68	99.72	99.33	99.85	95.42	96.38	96.87	99.73	99.73	99.57	99.86	99.85
Snack	93.72	94.99	96.82	99.31	89.62	92.21	95.14	98.42	97.99	98.36	98.72	99.34
Spare Time	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Toileting	73.52	73.68	74.73	90.76	65.23	67.27	71.62	86.97	78.75	82.51	82.98	90.76
Total	**90.24**	**90.84**	**91.97**	**96.44**	**88.13**	**89.70**	**91.43**	**95.37**	**92.49**	**93.53**	**94.52**	**96.97**
Train-time	3.91	4.1	4.6	5.1	2.3	2.7	3.1	4.01	3.2	3.5	3.9	4.6

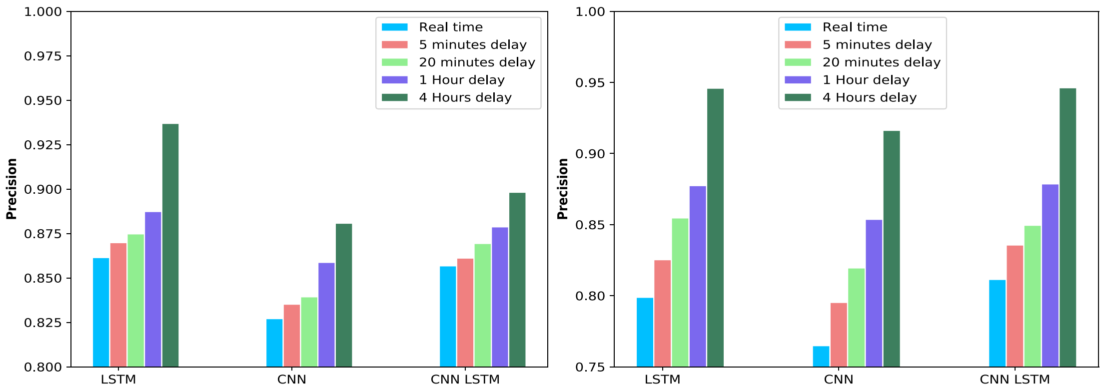
TABLE III
F1-SCORE AND TRAINING TIME (MINUTES) OF LSTM, CNN, AND CNN LSTM WITH DIFFERENT DELAYS IN TIME FROM ORDONEZ HOME B

Activity	LSTM				CNN				CNN LSTM			
	5 Minutes delay	**20 Minutes delay**	**1 Hour delay**	**4 Hours delay**	**5 Minutes delay**	**20 Minutes delay**	**1 Hour delay**	**4 Hours delay**	**5 Minutes delay**	**20 Minutes delay**	**1 Hour delay**	**4 Hours delay**
Breakfast	93.31	94.11	94.81	99.69	91.27	91.97	95.65	99.25	94.33	94.87	96.43	99.54
Grooming	78.37	78.89	79.23	91.97	73.14	76.36	82.85	87.07	78.91	83.22	85.23	90.67
Leaving	93.62	98.72	98.72	99.50	94.28	98.51	98.63	99.39	98.84	98.86	98.89	99.55
Lunch	95.41	95.54	96.22	99.02	92.84	93.14	95.87	98.43	96.96	96.94	97.12	99.14
Showering	89.42	90.1	90.32	98.85	87.12	87.98	92.65	99.42	89.31	94.84	96.42	99.42
Sleeping	99.41	99.48	99.67	99.80	99.48	99.51	99.63	99.68	99.64	99.65	99.68	99.73
Snack	84.67	85.42	85.67	96.39	81.57	83.63	89.71	94.49	86.53	88.94	92.74	97.63
Spare Time	96.98	96.93	96.98	99.33	97.14	97.52	97.58	98.95	96.92	97.12	97.54	99.29
Toileting	62.37	66.37	67.12	86.23	62.73	65.72	69.43	76.17	67.63	72.78	76.68	85.19
Dinner	81.24	84.24	86.24	97.39	79.16	83.16	88.83	95.62	78.83	83.74	87.68	97.03
Total	**87.48**	**88.98**	**89.49**	**96.82**	**85.87**	**87.75**	**91.05**	**94.84**	**88.79**	**91.09**	**92.84**	**96.72**
Train-time	4.9	5.23	5.41	5.93	3.17	3.33	3.84	4.31	4.36	4.51	4.82	5.71

the proposed model with FTWs and Deep Learning achieves encouraging performance regarding ad-hoc classical approaches and sensor representations, as well as previous approaches based on Deep Learning with Equally-sized temporal windows.

This is particularly so in the activities that real-time models have difficulty recognizing accurately, such as *Leaving*, *Snack*, *Grooming*, and *Toileting* from home A. Regarding home B, the results of the same activities in addition to *Dinner* are significantly improved. This refers to the fact that taking oncoming sensor activations into account is crucial in order to enhance the learning process of the models.

Finally, we note the learning time in Deep Learning is in the magnitude of minutes compared with classical approaches which develop learning in seconds. However, the evaluation of Deep Learning under this approach has been developed with low learning time requirements (10 epochs which take less than 6 minutes). In addition, one limitation of this work is the difficulty in handling interleaved activities due to a single classifier



(a) Home A (b) Home B

Fig. 3. Precision of LSTM, CNN and CNN LSTM for real time and different delays in time.

TABLE IV
F1-SCORE AND TRAINING TIME FROM 4 HOURS DELAYED IN ADDITION TO REAL-TIME RECOGNITION FOR DIFFERENT MODELS WITH EQUAL-SIZED TEMPORAL WINDOWS (ESTWs), RAW AND LAST NEXT ACTIVATION (RLNA) AND FUZZY TEMPORAL WINDOWS(FTWs)

(a) Home A

Activity	RLNA		FTWs	ESTWs
	C4.5	**SVM**	**LSTM**	**LSTM**
Breakfast	82.87	77.66	96.95	78.32
Grooming	75.65	75.14	86.19	59.17
Leaving	99.94	99.94	99.85	91.21
Lunch	70.90	78.00	99.00	81.31
Showering	96.42	96.42	98.21	94.53
Sleeping	99.98	99.98	99.85	99.41
Snack	00.00	38.46	99.31	48.37
Spare Time	98.51	98.55	100.00	97.41
Toileting	62.68	61.57	90.76	51.32
Total	**78.33**	**80.63**	**96.44**	**77.89**
Train-time	2.78 s	2.21 s	5.1 m	13.3 m

(b) Home B

Activity	RLNA		FTWs	ESTWs
	C4.5	**SVM**	**LSTM**	**LSTM**
Breakfast	49.28	47.33	99.69	48.78
Grooming	95.10	94.06	91.97	69.26
Leaving	98.88	98.83	99.50	81.75
Lunch	29.42	31.87	99.02	41.61
Showering	95.54	95.54	98.85	96.13
Sleeping	99.39	99.34	99.80	99.48
Snack	46.52	45.14	96.39	39.69
Spare Time	91.58	91.74	99.33	92.91
Toileting	91.71	92.69	86.23	51.32
Dinner	21.51	00.00	97.39	42.83
Total	**71.89**	**69.65**	**96.82**	**66.37**
Train-time	6.77 s	12.18 s	5.93 m	18.42 m

being developed for learning with multi-class labels. In order to enable this methodology to support interleaved activities, a configuration using an ensemble architecture (where each activity is represented by a classifier) is necessary [26]; however, we note that the learning time and learning data size required is multiplied by the number of classes.

V. CONCLUSION AND FUTURE WORK

Human activity recognition is a highly dynamic and challenging research field that plays a crucial role in diverse applications such as health care, elderly care, emergencies, security, smart environments, surveillance and context-aware-systems. In this study, we have proposed a new data-driven approach that aims to increase precision and sensitivity in human activity recognition applied in a smart home setting. The proposed method considers the partial oncoming sensor activations in addition to preceding sensor activations. With the use of oncoming sensor activation, we can take the benefits of enhancing the learning process that leads to improved recognition performance compared with the approaches using only the preceding sensor activations in the intelligent environment. Multiple and incremental fuzzy temporal windows were used to extract features from both preceding and partial oncoming sensor activations. Defining multiple and incremental fuzzy temporal windows from long-term to short-term has provided suitable semantics to determine a sequence of temporal features that boosts learning using LSTM sequence models and CNN.

Experiments show that precision and sensitivity increase by magnitudes when using preceding as well as partial oncoming sensor activations compared to precision and sensitivity when using only preceding sensor activations. The results of the experiment indicate that the more partial oncoming sensors are included, the better results are achieved. However, results are also improved when considering partial oncoming in addition to preceding sensor activations within a short amount of time, for example 5 minutes. The proposed approach of this paper has enhanced the models for recognizing all the activities performed in houses A and B. Particularly the activities that real-time models have difficulty to recognize accurately such as *Leaving*, *Snack*, *Grooming*, and *Toileting* from house A. Regarding house B, the results of the same activities in addition to *Dinner* are significantly improved. This refers to the fact that taking partial oncoming sensor activations into account is crucial in order to enhance the learning process of the models.

Future work will explore the proposed approach in two main directions. First, we will work on boosting learning over different smart homes aiming to perform robust recognition of dangerous situations and detect behaviour deviations in order to enhance elderly-care alert systems. The long-term goal of our project based on the proposed approach will be key to transferring knowledge over different smart homes in terms of layout, resident and sensor configuration. Second, we will integrate heterogeneous sensors from wearable and location sources using fuzzy logic and scales in learning AR of patients in smart homes, which have been previously demonstrated as suitable fusion representation [35].

References

[1] R. A. Hamad, E. Järpe, and J. Lundström, "Stability analysis of the t-SNE algorithm for human activity pattern data," in *Proc. IEEE Int. Conf. Syst., Man Cybern.*, 2018, pp. 1839–1845.

[2] J. Bravo, R. Hervas, J. Fontecha, and I. Gonzalez, "m-health: Lessons learned by m-experiences," *Sensors*, vol. 18, no. 5, p. 1569, 2018.

[3] K. Bouchard *et al.*, "The cornerstones of smart home research for healthcare," *Smart Innov., Syst. Technol.*, vol. 93, pp. 185–200, 2018.

[4] F. J. Ordónez, P. de Toledo, and A. Sanchis, "Activity recognition using hybrid generative/discriminative models on home environments using binary sensors," *Sensors*, vol. 13, no. 5, pp. 5460–5477, 2013.

[5] A. Lotfi, C. Langensiepen, S. M. Mahmoud, and M. J. Akhlaghinia, "Smart homes for the elderly dementia sufferers: Identification and prediction of abnormal behaviour," *J. Ambient Intell. Humanized Comput.*, vol. 3, no. 3, pp. 205–218, 2012.

[6] R. Varatharajan, G. Manogaran, M. K. Priyan, and R. Sundarasekar, "Wearable sensor devices for early detection of alzheimer disease using dynamic time warping algorithm," *Cluster Comput.*, vol. 21, pp. 681–690, 2018.

[7] L. Á. San Martín, V. M. Peláez, R. González, A. Campos, and V. Lobato, "Environmental user-preference learning for smart homes: An autonomous approach," *J. Ambient Intell. Smart Environ.*, vol. 2, no. 3, pp. 327–342, 2010.

[8] D. Singh, E. Merdivan, S. Hanke, J. Kropf, M. Geist, and A. Holzinger, "Convolutional and recurrent neural networks for activity recognition in smart environment," in *Proc. Towards Integrative Mach. Learn. Knowl. Extraction*, Springer, 2017, pp. 194–205.

[9] J. Yang, M. N. Nguyen, P. P. San, X. Li, and S. Krishnaswamy, "Deep convolutional neural networks on multichannel time series for human activity recognition," in *Proc. 24th Int. Joint Conf. Artif. Intell. ,* 2015, vol. 15, pp. 3995–4001.

[10] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.

[11] J. Suto, S. Oniga, C. Lung, and I. Orha, "Comparison of offline and real-time human activity recognition results using machine learning techniques," *Neural Comput. Appl.*, pp. 1–14, Mar. 2018, doi: 10.1007/s00521-018-3437-x.

[12] J. Medina, M. Fernandez-Olmo, M. D. Peláez, and M. Espinilla, "Real-time monitoring in home-based cardiac rehabilitation using wrist-worn heart rate devices," *Sensors*, vol. 17, 2017, Art. no. 2892.

[13] K. S. Gayathri, S. Elias, and B. Ravindran, "Hierarchical activity recognition for dementia care using Markov logic network," *Pers. Ubiquitous Comput.*, vol. 19, no. 2, pp. 271–285, 2015.

[14] D. H. Wilson and C. Atkeson, "Simultaneous tracking and activity recognition (star) using many anonymous, binary sensors," in *Proc. Int. Conf. Pervasive Comput.*, Springer, 2005, pp. 62–79.

[15] T. Van Kasteren, A. Noulas, G. Englebienne, and B. Kröse, "Accurate activity recognition in a home setting," in *Proc. 10th Int. Conf. Ubiquitous Comput.*, 2008, pp. 1–9.

[16] H. Storf, T. Kleinberger, M. Becker, M. Schmitt, F. Bomarius, and S. Prueckner, "An event-driven approach to activity recognition in ambient assisted living," in *Proc. Eur. Conf. Ambient Intell.*, Springer, 2009, pp. 123–132.

[17] T. L. M. Van Kasteren, G. Englebienne, and B. J. A. Kröse, "An activity monitoring system for elderly care using generative and discriminative models," *Pers. Ubiquitous Comput.*, vol. 14, no. 6, pp. 489–498, 2010.

[18] J. Medina-Quero, S. Zhang, C. Nugent, and M. Espinilla, "Ensemble classifier of long short-term memory with fuzzy temporal windows on binary sensors for activity recognition," *Expert Syst. Appl.*, vol. 114, pp. 441–453, 2018.

[19] P. Skocir, P. Krivic, M. Tomeljak, M. Kusek, and G. Jezic, "Activity detection in smart home environment," *Procedia Comput. Sci.*, vol. 96, pp. 672–681, 2016.

[20] T. L. M. Van Kasteren, G. Englebienne, and B. J. A. Kröse, "Activity recognition using semi-Markov models on real world smart home datasets," *J. Ambient Intell. Smart Environ.*, vol. 2, no. 3, pp. 311–325, 2010.

[21] D. J. Cook, A. S. Crandall, B. L. Thomas, and N. C. Krishnan, "Casas: A smart home in a box," *Computer*, vol. 46, no. 7, pp. 62–69, 2013.

[22] D. Singh *et al.*, "Human activity recognition using recurrent neural networks," in *Proc. Int. Cross-Domain Conf. Mach. Learn. Knowl. Extraction*, Springer, 2017, pp. 267–274.

[23] A. P. Stakhov, "The generalized principle of the golden section and its applications in mathematics, science, and engineering," *Chaos, Solitons Fractals*, vol. 26, no. 2, pp. 263–289, 2005.

[24] F. J. Ordóñez and D. Roggen, "Deep convolutional and LSTM recurrent neural networks for multimodal wearable activity recognition," *Sensors*, vol. 16, no. 1, p. 115, 2016.

[25] J. Medina, L. Martinez, and M. Espinilla, "Subscribing to fuzzy temporal aggregation of heterogeneous sensor streams in real-time distributed environments," *Int. J. Commun. Syst.*, vol. 30, no. 5, p. e3238, 2017.

[26] J. Medina-Quero, C. Orr, S. Zang, C. Nugent, A. Salguero, and M. Espinilla, "Real-time recognition of interleaved activities based on ensemble classifier of long short-term memory with fuzzy temporal windows," *Multidisciplinary Digit. Publishing Inst. Proc.*, vol. 2, p. 1225, 2018.

[27] P. F. Byrd, "Expansion of analytic functions in polynomials associated with Fibonacci numbers," *Fibonacci Quart.*, vol. 1, no. 1, pp. 16–27, 1963.

[28] N. C. Krishnan and D. J. Cook, "Activity recognition on streaming sensor data," *Pervasive Mobile Comput.*, vol. 10, pp. 138–154, 2014.

[29] G. E. Batista, R. C. Prati, and M. C. Monard, "A study of the behavior of several methods for balancing machine learning training data," *ACM Sigkdd Explorations Newslett.*, vol. 6, no. 1, pp. 20–29, 2004.

[30] J. Collins, J. Sohl-Dickstein, and D. Sussillo, "Capacity and trainability in recurrent neural networks," *CoRR*, vol. abs/1611.09913, 2017.

[31] A. Murad and J.-Y. Pyun, "Deep recurrent neural networks for human activity recognition," *Sensors*, vol. 17, no. 11, 2017, Art. no. 2556.

[32] N. Y. Hammerla, S. Halloran, and T. Plötz, "Deep, convolutional, and recurrent models for human activity recognition using wearables," in *Proc. 25th Int. Joint Conf. Art. Intell.*, New York, NY, USA, 2016, pp. 1533–1540.

[33] A. Devarakonda, M. Naumov, and M. Garland, "Adabatch: Adaptive batch sizes for training deep neural networks," *CoRR*, vol. abs/1712.02029, 2017.

[34] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, 2014.

[35] J. M. Quero, M. Á. L. Medina, A. S. Hidalgo, and M. Espinilla, "Predicting the urgency demand of copd patients from environmental sensors within smart cities with high-environmental sensitivity," *IEEE Access*, vol. 6, pp. 25081–25089, 2018.

[36] L. Wang, T. Gu, X. Tao, and J. Lu, "A hierarchical approach to real-time activity recognition in body sensor networks," *Pervasive Mobile Comput.*, vol. 8, no. 1, pp. 115–130, 2012.

Appendix C

Paper III
Efficacy of Imbalanced Data Handling Methods on Deep Learning for Smart Homes Environments

Rebeen Ali Hamad , Masashi Kimura , Jens Lundstrom

published in SN Computer Science.

SN Computer Science (2020) 1:204
https://doi.org/10.1007/s42979-020-00211-1

ORIGINAL RESEARCH

Efficacy of Imbalanced Data Handling Methods on Deep Learning for Smart Homes Environments

Rebeen Ali Hamad[1] · Masashi Kimura[2] · Jens Lundström[3]

Received: 28 February 2020 / Accepted: 29 May 2020

Abstract
Human activity recognition as an engineering tool as well as an active research field has become fundamental to many applications in various fields such as health care, smart home monitoring and surveillance. However, delivering sufficiently robust activity recognition systems from sensor data recorded in a smart home setting is a challenging task. Moreover, human activity datasets are typically highly imbalanced because generally certain activities occur more frequently than others. Consequently, it is challenging to train classifiers from imbalanced human activity datasets. Deep learning algorithms perform well on balanced datasets, yet their performance cannot be promised on imbalanced datasets. Therefore, we aim to address the problem of class imbalance in deep learning for smart home data. We assess it with Activities of Daily Living recognition using binary sensors dataset. This paper proposes a data level perspective combined with a temporal window technique to handle imbalanced human activities from smart homes in order to make the learning algorithms more sensitive to the minority class. The experimental results indicate that handling imbalanced human activities from the data-level outperforms algorithms level and improved the classification performance.

Keywords Activity recognition · Smart home · Imbalanced class

Introduction

By equipping environments such as ordinary homes with binary sensors for monitoring resident activities, a vast area of different applications is made possible, including smart monitoring of energy utilization and assessing resident situation and behavior pattern for proactive home care. In the case of monitoring for home care, independent living solutions have been provided for older adults in their own homes by smart home technology to improve and maintain the quality of life and care [2, 27, 33]. Smart homes that are used for transparently represent how, when and where humans perform activities opens up diverse health technology applications such as anomaly detection (e.g., falls) or tracking progression of diseases or recovery. Activity recognition (AR) has progressed by the recent advancement of machine learning to enhance elderly care alert systems and improve assistance in emergency situations from smart home data [12]. Another example of an application requiring AR includes smart medication reminders [40] which utilize the contexts in which to send a reminder. Similar to medication reminders is the application of assisting people with cognitive impairments to complete tasks [9]. These applications relying on AR would potentially benefit from a more accurate recognition. Moreover, by tracking the characteristics of activities related to basic needs and their change over time renders a possibility to assess parts of the progression of a persons *functional ability*, which is a focus concept for how WHO defines healthy aging. Activities of in-home mobility as showering, watching TV, cooking, eating, sleeping and grooming are therefore of importance to monitor and track in order to assess the functional health status of older adults. Moreover, the framework of AR using machine learning methods provides enough mechanisms to detect both

✉ Rebeen Ali Hamad
rebeen.ali_hamad@hh.se

Masashi Kimura
kimura@convergence-lab.com

Jens Lundström
jens@convergia-consulting.io

1 Intelligent Systems and Digital Design, Halmstad University, Halmstad, Sweden

2 Convergence Lab, Tokyo, Japan

3 Convergia Consulting, Halmstad, Sweden

Published online: 18 June 2020

ambulatory and postural activities, actions of residents and body movements using different multimodal data generated by heterogeneous sensors [5, 19, 31].

Not only are human activities highly diverse in the form of different sensor activations but the frequency of activities themselves is inherently imbalanced and hence accurate AR is challenging from a machine learning perspective. Large differences in the number of examples for the classes to learn can make the machine learning algorithm to put emphasis on learning majority classes and thereby partially or completely neglect minority classes. As an example, cooking may occur with a higher frequency than grooming. Another more prominent example is the vast difference in the number of examples between eating and sleeping where the latter occurs with a much higher frequency in datasets collected over a long duration. This paper focuses on investigating the particularly problematic aspect of learning activities over days or even months which are imbalanced.

Despite many past efforts of research on the class imbalance problem and approaches to cope with this general problem, there is a lack of empirical work on targeting machine learning beyond shallow methods [20]. Traditional machine learning algorithms such as decision tree, support vector machine, naive Bayes and hidden Markov models have been used to minimize the recognition error [6, 23]. Satisfying recognition results have been achieved by adopting these approaches. However, such algorithms may heavily depend on classical heuristic and hand-crafted feature extraction which might be limited by human domain knowledge [39]. A natural variation within each activity is often present in collected smart home datasets and is not unlikely to fluctuate even more between different residents. These variations are also influenced by contextual factors such as time of the day and location of where the activity is performed. Given these conditions as well as considering the multitude of choices at sensor installation (e.g., sensor types and sensor locations), AR based on shallow learning where features are hand-crafted can be challenging. Therefore, discovering more systematic methods to obtain features has drawn increasing research interests [24]. The influence of deep learning has been demonstrated in many areas not only in image classification such as speech recognition and natural language processing as surveyed in [39]. Consequently, studies of activity recognition using deep learning have multiplied because the number of elderly smart-home healthcare services has steadily increased for the last few years and all reporting state-of-the-art performances achieved on diverse activity recognition benchmark datasets [16, 43]. Particularly, two methods have brought promising results of AR, long short-term memory (LSTM) and convolutional neural networks (CNNs) when using data prepared with a fuzzy-based approach to represent temporal components of the data [15, 26, 28]. However, to the best of our knowledge, these two machine learning algorithms for AR have not been studied from the context of different temporal preprocessing methods along with traditional methods for handling class imbalance in order to improve recognition accuracy. The study described in this paper is therefore designed to fill parts of such a knowledge gap and also put a particular focus on the classes representing activities with a relatively low number of observations (i.e., minority classes). Thus, the main contribution of this paper is the study of well-known class imbalance approaches (synthetic minority over-sampling technique, cost-sensitive learning and ensemble learning) applied to activity recognition data with various temporal data preprocessing for the deep learning models LSTM and 1D CNN.

The rest of the paper is organized as follows. In Sect. 2, related work is described, and in Sect. 3 Methodology, the outline and details of the study are described, whereas in Sect. 4, experiment results are presented and discussed. Finally, the findings and opportunities of further research are summarized in Sect. 5, Conclusion and future work.

Related Work

Elements of the class imbalance problem are widely studied, especially from a shallow learning perspective. Extensive work by [18] outlined three important factors of the problem: the complexity of concept (or underlying distributions), training set size and degree of imbalance. It was shown that problems with low concept complexity were insensitive to class imbalances but with an increased concept complexity the models (C5.0 & MLP) performed poorly, even when a low-class imbalance was present. Moreover, Japkowicz and Stephen concluded that a severe complex problem could be handled with a good performance given a sufficiently large amount of training data [18]. Finally, their conclusion that over-sampling and cost-modifying methods for improving model performance are preferred over an undersampling strategy, is a direction explored in this paper for deep learning models.

The intrinsic property of classes representing human activities to be imbalanced makes the topic of AR learning algorithms for imbalance handling crucial to study, especially since the arrival of deep learning which typically requires a larger dataset. Different strategies for dealing with class imbalance for deep learning were recently surveyed by [20]. The survey revealed that the number of research studies containing empirical work on targeting the class imbalance problem for deep learning is limited. However, the same survey showed that classical methods for handling imbalance (e.g., random over-sampling of minority classes and cost-sensitive target function to avoid skewed learning toward majority classes) applied in deep learning situations show promising results.

Most past works on handling class imbalance for deep neural networks focus on computer vision tasks where image classification dominates the reviewed papers and hence not directly translatable to an AR setting. A modified cost-sensitive learning scheme was proposed by [22] with good results compared to standard cost-sensitive (when the target function is weighted toward the size or importance of classes) approaches and sampling methods (where the majority classes are undersampled or minority classes are over-sampled). However, the evaluation was based on data for image classification tasks. Another novel approach (focusing on a vision classification problem) combined sampling and a modified hinge loss to render tighter constraints between classes for a better discriminative deep representation [17]. The focus of this paper is class imbalance handling for activity recognition in a deep learning context which has earlier been approached by Nguyen et al. who proposed an extension to the random over-sampling method SMOTE called BLL-SMOTE which improved the classification results drastically [30]. However, the study was limited to mobile phone sensors which is only a subset of the type of sensors available as smart home technology.

Besides handling imbalanced activity classes, the domain of activity recognition often needs alignment to the use of a carefully selected temporal window size. In the case of mobile sensing devices, the use of a temporal window size needs a thorough analysis to properly and correctly segment the data [4]. Shallow learning schemes such as support vector machines (SVMs), decision tree or hidden Markov model based on the dynamic or sliding windows have previously been evaluated [11, 36, 38, 42]. These studies have aimed to adjust dynamic or fixed window size to enhance the performance of the classifiers. Binary stream sequence data are mostly split into subsequences called windows, where every window is related to a broader activity by a sliding window technique. Binary sensor data segmentation using only one window for deploying HAR cannot provide accurate results since the duration of human activities differ and the exact boundaries of activities are difficult to specify. Intuitively, decreasing the window size has led to increasing the performance of activity recognition in addition to minimizing resources and energy needs [4]. It has been found that the window size of 60 s extracts satisfactory features for activity recognition from smart home [26, 32].

Consequently, thorough comparisons of the use of fixed window size and fuzzy temporal windows (of particularly one hour) are important to study. The contribution of this paper is therefore significant to alleviate the complexity of defining the window size and to correctly, easily and rapidly recognize real-time imbalanced activities.

Methodology

In this study, aspects of how to approach the class imbalance problem are considered. This section describes the relevant key components: window methods for pre-processing, machine learning algorithms used and class imbalance strategies.

Methods to Handle Imbalanced Class Problem

The following two methods are used to handle the imbalanced class problem in activity recognition from algorithm level and data level.

Cost-Sensitive

Cost-sensitive is one of the commonly used algorithm level methods to handle classification problems with imbalanced data in machine learning and data mining setting [44]. Cost-sensitive evaluates the cost associated with misclassifying samples. Cost-sensitive is not creating balanced data distribution; rather, this method assigns the training samples of different classes with different weights, where the weights will be in proportion to the misclassification costs. Then, the weighted samples will be fed to learning algorithms [45].

SMOTE

Synthetic minority over-sampling technique (SMOTE) is a commonly used data-level method to handle imbalanced data and is based on sampling. This method over-samples the minority classes by creating synthetic samples rather than by over-sampling with replacement [7]. The minority classes will be over-sampled by selecting each minority class sample and generating synthetic observations along the line segments joining any/all of the k minority class nearest neighbor. Neighbors will be randomly chosen from the k nearest neighbors depending on the amount of required over-sampling. Commonly five nearest neighbors are used in practice. For example, if 200% is the amount needed to be over-sampled, only two neighbors are selected from the five nearest neighbors and one sample will be created in the direction of each. Synthetic samples are created by taking the difference between the sample and its nearest neighbor. The difference will be multiplied by a random number between 0 and 1 and added to the feature vector. This procedure will effectively force the decision region of the minority class to become more general. The synthetic samples will be generated in a less application-specific manner by operating in feature space instead of data space to alleviate the issues with class imbalanced distribution. Despite the common use of SMOTE at data level, the method is less studied in deep learning contexts nor is it, to the best of our knowledge, studied together with the effect of windowing pre-processing techniques (described in section 3.3). Thereby, this paper aims to explore the potential enhancements of class imbalance approaches (where SMOTE is one of the tested methods) together with two deep learning models (1D CNN and LSTM) and several pre-processing methods described in later sections.

Ensemble Techniques

Ensemble techniques combine several based models into one single model to enhance prediction and decrease bias and variance. The decision of several estimators on a different randomly selected subset of data will be combined to improve overall performance [14, 41]. However, commonly the subsets of data are not balanced as input to the classifiers in the ensemble. Therefore, the classifiers may favor the majority classes and generate a biased model during the training phase on the input imbalanced datasets. To overcome this problem and to reasonably compare the results of the ensemble model with the cost-sensitive and SMOTE, balanced ensemble learning is used in this study which is introduced in [13]. Balanced ensemble learning will first balance the data and then will combine the decision of multiple classifiers to avoid bias and to render better performance. Decision trees as the base models with bootstrap aggregation (Bagging) are used to build the ensemble learning.

Smart Home Data for Evaluation

We used the activities of daily living (ADLs) for recognition using binary sensors dataset, which were acquired in two real intelligent homes A and B in which residents perform their daily routine [32]. These two homes are equipped with sensors that are able to capture the movements and interactions of the inhabitants. The binary sensors are passive infrared (PIR) motion detectors to identify movement in a specific area, pressure sensors on beds and couches to detect the user's presence, reed switches on cupboards and doors to measure open or close status and float sensors in the bathroom to measure toilet being flushed or not. The use of PIR sensors as well as pressure sensors is limited in their ability to capture details compared to other sensors such as cameras or accelerometers. However, low-resolution sensors such as PIR and pressure sensors may preserve the privacy and integrity of residents to a greater extent than for example cameras. Table 1 shows details of the two homes with information of the resident, number of activities and sensors. In home A, 9 human daily activities that were performed in 14 days over a period of 19,932 min were described by an incoming stream of binary events from 12 sensors in the home. In home B, ten human daily activities that were performed in 22 days over a period of 30,495 min were described by 12 binary sensors. The timeline of the activities is segmented in time slots using the window size $\Delta t = 1\,\mathrm{min}$. The activities of homes A and B that were manually labeled are *Breakfast, Grooming, Idle, Leaving, Lunch, Showering, Sleeping, Snack, Spare Time/TV, Toileting*; in addition to these, home B has the activity *Dinner*.

Leave-one-out cross-validation is used and repeated this for every day and for both homes. Deep learning models (described in the next section) are trained for each home since the number of sensors varies and a different user resides in each home. Sensors are recorded at one-minute interval for 24 h , which totals in 1440 length input in minutes for each day. The average F-score is computed from the results of the cross-validation. Since the classes of the datasets are imbalanced, we propose synthetic minority oversampling technique (SMOTE) as input data for the deep learning model. This allows us to handle the imbalanced activities and avoid having models biased toward one class or the other (Table 2).

Table 1 Details of recorded datasets

	Home A	Home B
Setting	Home	Home
Rooms	4	5
Duration	14 days	21 days
Sensors	12	12
Activities	10	11

Table 2 Number of observations for each activity in the datasets

Activity	Home A	Home B
Spare Time/ TV	8555	8984
Sleeping	7866	10763
Leaving	1664	5268
Idle	1598	3553
Lunch	315	395
Toileting	138	167
Breakfast	120	309
Grooming	98	427
Showering	96	75
Snack	6	408
Dinner	–	120

Data Pre-Processing

Multiple and incremental fuzzy temporal windows (FTWs) are used to extract features. Each FTW T_k is defined by a fuzzy set characterized with a membership function, and its shape corresponds to a trapezoidal function $T_k[l_1, l_2, l_3, l_4]$. The well-known trapezoidal membership functions are defined by a lower limit l_1, an upper limit l_4, a lower support limit l_2 and an upper support limit l_3. The values of l_1, l_2, l_3, l_4 are defined by the Fibonacci sequence which was previously shown as a successful sequence for defining FTWs without requiring expert knowledge definition [15, 25]. Figure 2 shows nine windows of FTWs created based on Fibonacci sequence. To extract features, the FTWs are slided over sensors activations x in every minute according to Eq. (1): Features are computed by applying 15 FTWs on the raw data from all 12 binary sensors in each minute for both datasets. The datasets A and B have 19,932 and 30,495 samples, respectively, where each sample represents one minute of data with $12 \times 15 = 180$ features. The resulting datasets are used for real-time activity recognition. Algorithm 1 shows the procedure of computing FTWs. Feature extraction based on FTWs is evaluated and compared with equally sized (1 min) temporal windows (ESTWs) [34] as shown in Fig. 1.

$$T_k(x)[l_1, l_2, l_3, l_4] = \begin{cases} 0 & x \le l_1 \\ (x - l_1)/(l_2 - l_1) & l_1 < x < l_2 \\ 1 & l_2 \le x \le l_3 \\ (l_4 - x)/(l_4 - l_3) & l_3 < x < l_4 \\ 0 & l_4 \le x \end{cases} . \tag{1}$$

Algorithm 1 Extracting Features using FTWs

```
1: Input: Raw_data                     ▷ Home A and B are the input Raw data
2: FTWs ← Fibonacci                    ▷ FTWs get values from Fibonacci
3: Sensor_intervals ← Raw_data         ▷ sensor intervals data
4: for ftw ← FTWs do
5:    for sen_intv ← Sensor_intervals do
6:       apply ftw on sen_intv
7:    end for
8:    features ← max(ftw)
9: end for
10: dataset ← features
11: Output: dataset
```

Algorithm 2 shows the process of handling imbalanced class problem where firstly data preprocessed by FTWs or ESTWs and then infrequent classes are over-sampled by SMOTE to be used as the input data of the models (Fig. 2).

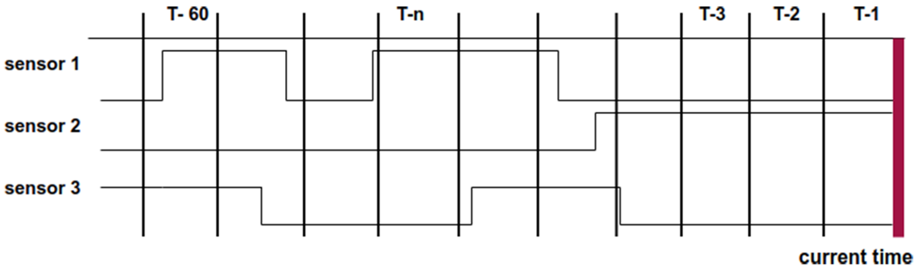


Fig. 1 Example of temporal segmentation on time series of three sensors by the equally sized temporal window method

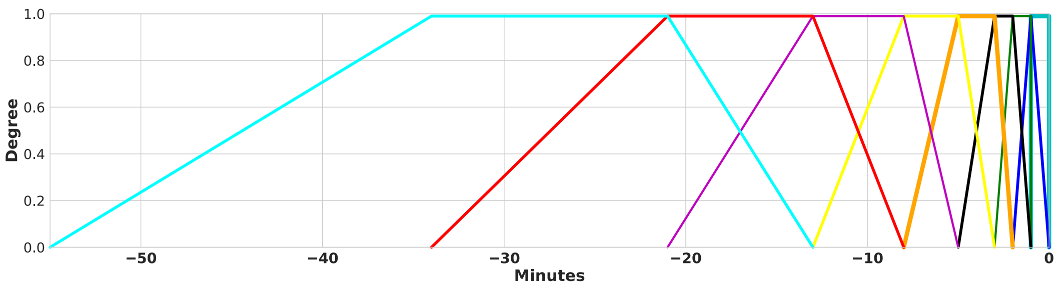


Fig. 2 Example of temporal segmentation on sensors time series by the fuzzy incremental temporal windows method

Algorithm 2 Process of Handling imbalanced Data

1: **Input:** *Raw_data* input Raw data
2: *FTW*, ESTW FTW, ESTW to extract features and build datasets
3: $SMOTE \leftarrow datasets$ oversamples infrequent classes
4: $LSTM, 1D\ CNN \leftarrow datasets$ Apply temporal models

Model Selection and Architecture

In this study, we investigate two types of neural networks: One is based on LSTM (long short-term memory) and another is based on CNN (convolutional neural network). The architecture and parameters of the temporal models are described in the following.

LSTM

LSTM is the extended form of the recurrent neural network (RNN) that is designated to learn from temporal sequential pattern data. We expect an LSTM architecture to handle the activity timeline of a smart home. LSTM solves the vanishing gradient problem of a simple RNN which cannot learn long-term sequences and lose the effect of initial dependencies in the sequence. LSTM is most widely used in natural language processing, stock market prediction and speech recognition that can model temporal dependence between observations [8]. LSTM has obtained satisfying results in activity recognition [16, 29]. Hence, in this study LSTM is used in the experiments by stacking two LSTM layers with 40% dropout rate and 0.001 learning rate followed by a fully connected, i.e., dense layer and softmax layer. For all the models in this study, the batch size and training epochs are equal to 10, which is a total of 100 batches during the entire training process. While large batch size commonly results in faster training, it is unable to converge as fast. On the other hand, smaller batch sizes train slower but could converge faster; therefore, it is mostly an independent problem [10]. Regarding the 40% dropout, which is a regularization technique for preventing deep learning models from overfitting [35], the dropout ignores randomly selected neurons during the training phase. Those ignored neurons are temporally removed on the forward pass and their weights are not updated on the backward pass (Fig. 3).

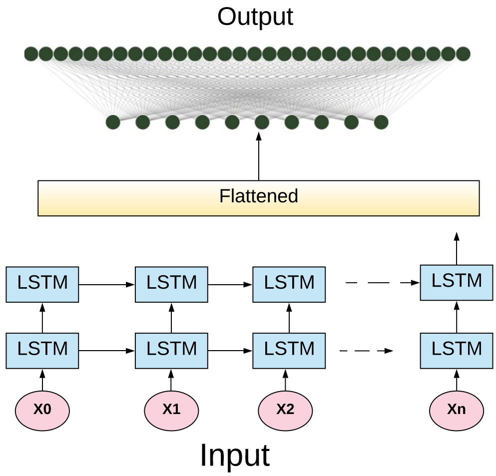


Fig. 3 Architecture of LSTM

1D CNN

Convolutional neural network (CNN) is used in the experiments because it is competent in extracting features from signals. CNN has obtained promising results in image classification, text analysis and speech recognition [16]. CNN has two advantages for human activity recognition which are local dependency and scale invariance. Local dependency refers to the nearby observations in human activity recognition that are likely to be correlated, while scale invariance means the scale is invariant for different paces or frequencies. CNN can learn hierarchical data representations which lead to rendering promising results in human activity recognition [16]. In this study, a one-dimensional (1D) CNN architecture is used and can extract local 1D subsequences from the sequence data. The 1D CNN could be competitive with RNN on some sequence-processing applications such as audio generation and machine translation with a cheaper computation cost compared to RNN [3, 15]. The model is designed by stacking two convolutional layers each with 64 filters, kernel size 3 and stride 1 with 40% dropout rate and 0.001 learning rate followed by a max-pooling layer and followed by a fully connected, i.e., dense layer and softmax layer (Fig. 4).

Measure Evaluation

How the classification performance is evaluated plays an important role in this study. Without proper measures, no

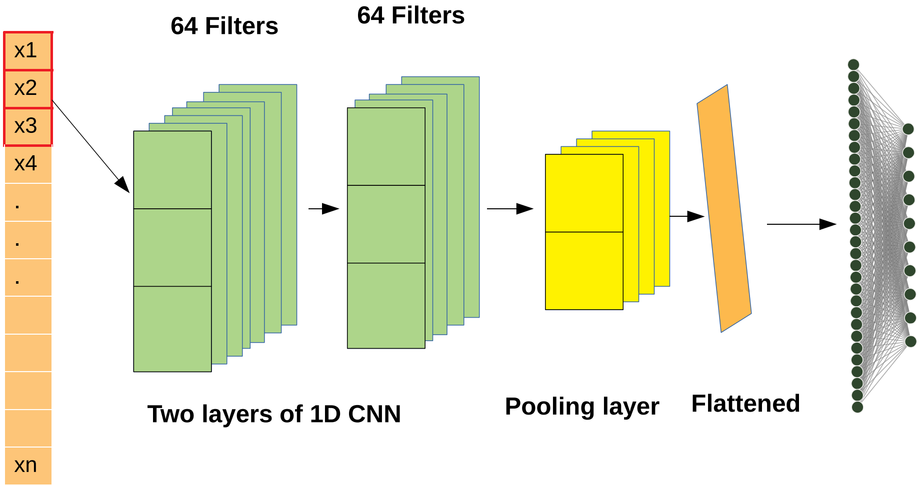


Fig. 4 Architecture of 1D CNN

deeper insight could be achieved. Traditionally, accuracy was commonly used to measure the performance of classifiers. However, for classification with the imbalanced class distribution problem, accuracy is no longer a appropriate measure since the minority classes have a very little impact on the accuracy compared to the majority classes [37]. Therefore, in this study, the F1-score is used to evaluate the models because the F1-score ($2\frac{\text{precision}\times\text{recall}}{\text{prscecision}+\text{recall}}$) shows an insight into the balance between sensitivity (recall) ($\frac{TP}{TP+FN}$) and precision ($\frac{TP}{TP+FP}$). This metric is also widely used in activity recognition [15, 21]

Results and Discussion

In this section, the results of the experiments using LSTM and CNN are presented and discussed in the aspect of different methods of handling imbalanced classes and different feature extraction approaches. FTWs and ESTWs are used to pre-process data and build the datasets for training. SMOTE, cost-sensitive and ensemble learning methods are used for handling the class imbalance present in the datasets. Table 3 shows the results of the F1-score of the LSTM and CNN models from the home A for the imbalanced dataset, with cost-sensitive corrections and minority sampling

Table 3 F1-score Home A

Activity	FTWs							ESTWs						
	Imbalanced data		Cos-Sensitive		SMOTE		Ensemble	Imbalanced Data		Cos-Sensitive		SMOTE		Ensemble
	CNN	LSTM	CNN	LSTM	CNN	LSTM		CNN	LSTM	CNN	LSTM	CNN	LSTM	
Snack	0.00	0.00	0.00	0.00	0.28	0.39	0.00	0.00	0.00	0.00	0.00	0.27	0.42	0.01
Showering	0.36	0.48	0.43	0.47	0.70	0.70	0.51	0.79	0.81	0.82	0.81	0.89	0.89	0.82
Grooming	0.00	0.00	0.00	0.00	0.25	0.28	0.12	0.55	0.53	0.54	0.55	0.56	0.57	0.57
Breakfast	0.61	0.67	0.65	0.68	0.71	0.73	0.38	0.71	0.72	0.76	0.74	0.73	0.77	0.67
Toileting	0.00	0.00	0.00	0.00	0.31	0.37	0.17	0.00	0.00	0.00	0.00	0.28	0.29	0.17
Lunch	0.75	0.80	0.81	0.82	0.80	0.84	0.64	0.81	0.80	0.82	0.85	0.86	0.86	0.81
Leaving	0.76	0.86	0.75	0.83	0.88	0.89	0.83	0.85	0.86	0.86	0.86	0.87	0.87	0.84
Sleeping	0.96	0.96	0.96	0.96	0.92	0.90	0.92	0.99	0.99	0.99	0.99	0.99	0.99	0.99
Spare Time	0.91	0.91	0.90	0.91	0.92	0.93	0.76	0.98	0.98	0.98	0.98	0.99	0.99	0.98
Average	**0.44**	**0.48**	**0.46**	**0.47**	**0.63**	**0.67**	**0.48**	**0.60**	**0.62**	**0.62**	**0.63**	**0.71**	**0.73**	**0.65**

using SMOTE. The F1-score of the minority classes which are *Breakfast, Grooming, Lunch, Showering, Toileting* and *Snack* from the home A are improved using SMOTE based on both approaches of extracting features and both models. The results also show the majority classes which are *Leaving* and *Spare-Time* activities (except Sleeping) which are also improved based on both approaches of extracting features for both models using the SMOTE method. The average results of the LSTM and CNN for all activities are improved using the SMOTE method based on both FTWs and ESTWs. Regarding home B, the F1-score of the minority classes (*Breakfast, Grooming, Lunch, Showering, Toileting, Snack,* and *Dinner*) is considerably improved, which are shown in Table 4. Moreover, only the results of the *Spare-Time* as the majority classes are improved based on FTWs. The average results of home B indicate that the SMOTE method substantially improved the recognition, particularly for the minority classes. The F1-scores in Tables 3 and 4 indicate that the results of the models based on both feature extraction approaches using SMOTE are better (higher F1-score) than the results of models based on cost-sensitive and class imbalanced datasets. Moreover, the F1-score results based on SMOTE with ESTWs can be seen to be higher than F1-scores based on SMOTE with FTWs from both homes of both models on average. Moreover, the obtained results based on the SMOTE technique with both feature extraction method (FTW and ESTW) and with both temporal models (LSTM and CNN) are better than the results obtained by balanced ensemble learning as shown in Tables 3 and 4. Therefore, the proposed data-level solution (SMOTE and ESTWs) to handle imbalanced human activities from smart homes is more promising than algorithms level (cost-sensitive and ensemble learning).

Conclusion and Future Work

Human activity recognition is a dynamic and challenging research area that plays an important role in diverse applications such as smart environments, security, health care, elderly care, emergencies, surveillance and context-aware systems. The frequency and duration of human activities are intrinsically imbalanced. The huge difference in the number of observations for the classes to learn will make many machine learning algorithms to focus on the classification of the majority examples due to its increased prior probability while ignoring or misclassifying minority examples. In this study, SMOTE and cost-sensitive learning are applied to temporal models and compared with ensemble learning to handle the class imbalance problem as well as to study the relation to two data pre-processing methods. Experiments show that f-measures of the minority classes are increased when using SMOTE with both temporal models (LSTM and CNN) and based on both ways of extracting features (FTWs and ESTWs). For example, the recognition measurement of the *Snack and Dinner* as one of the minority classes is notably improved in both homes, using both models and based on both feature extraction methods. The experimental results indicate that handling imbalanced data is more important than selecting machine learning algorithms and improves classification performance. Moreover, handling imbalanced class problem from data level using SMOTE and ESTWs for these activity datasets outperforms the algorithm level.

Future work will explore a newly proposed approach to handle the imbalanced class problem by integrating SMOTE with weak supervision. This approach will use SMOTE only to generate observations from minority classes and use weak supervision to correctly and properly label the new

Table 4 F1-score Home B

Activity	FTWs							ESTWs						
	Imbalanced data		Cos-Sensitive		SMOTE		Ensemble	Imbalanced data		Cos-Sensitive		SMOTE		Ensemble
	CNN	LSTM	CNN	LSTM	CNN	LSTM		CNN	LSTM	CNN	LSTM	CNN	LSTM	
Dinner	0.00	0.00	0.00	0.00	0.31	0.34	0.06	0.00	0.01	0.00	0.00	0.26	0.27	0.13
Snack	0.00	0.00	0.02	0.08	0.27	0.29	0.22	0.00	0.00	0.00	0.00	0.26	0.28	0.07
Showering	0.0 0	0.22	0.00	0.21	0.26	0.36	0.24	0.73	0.80	0.71	0.79	0.82	0.84	0.53
Grooming	0.13	0.30	0.09	0.30	0.39	0.36	0.42	0.62	0.61	0.61	0.61	0.64	0.65	0.54
Breakfast	0.50	0.47	0.51	0.51	0.52	0.58	0.36	0.26	0.23	0.24	0.19	0.30	0.35	0.29
Toileting	0.00	0.00	0.00	0.00	0.31	0.32	0.32	0.23	0.04	0.23	0.10	0.26	0.27	0.14
Lunch	0.39	0.35	0.31	0.38	0.41	0.42	0.37	0.00	0.00	0.00	0.00	0.36	0.38	0.00
Leaving	0.90	0.90	0.89	0.89	0.90	0.90	0.84	0.66	0.66	0.66	0.66	0.66	0.66	0.66
Sleeping	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97
Spare Time	0.83	0.82	0.84	0.84	0.85	0.86	0.79	0.90	0.90	0.90	0.90	0.89	0.90	0.90
Average	**0.33**	**0.36**	**0.36**	**0.41**	**0.51**	**0.54**	**0.45**	**0.40**	**0.40**	**0.40**	**0.40**	**0.54**	**0.56**	**0.42**

observations. The idea is designed to target the challenge of correctly labeling samples created in an over-sampling context. The long-term goal of our project will work on boosting learning across different smart homes aiming to perform robust recognition of dangerous situations and detect behavior deviations in order to enhance elderly care alert systems. This will be conducted by transferring knowledge over different smart homes in terms of layout, resident and sensor configuration.

Acknowledgements Open access funding provided by Halmstad University. This research is supported by the Knowledge Foundation under the project of the Center for Applied Intelligent Systems, under Grant Agreement No. 20100271.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. What is healthy ageing? https://www.who.int/ageing/healthy-ageing/en/. Accessed: 2019-08-10.
2. Ali Hamad Rebeen, Järpe Eric, Lundström Jens. Stability analysis of the t-sne algorithm for humanactivity pattern data. In *The 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC2018)*, 2018.
3. Bai Shaojie, Zico Kolter J, Koltun Vladlen. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. arXiv preprint arXiv:1803.01271, 2018.
4. Banos Oresti, Galvez Juan-Manuel, Damas Miguel, Pomares Hector, Rojas Ignacio. Window size impact in human activity recognition. Sensors. 2014;14(4):6474–99.
5. Cao Liang, Wang Yufeng, Zhang Bo, Jin Qun, V Vasilakos Athanasios. Gchar: An efficient group-based context–aware human activity recognition on smartphone. Journal of Parallel and Distributed Computing. 2018;118:67–80.
6. Manosha Chathuramali KG, Rodrigo Ranga . Faster human activity recognition with svm. In *Advances in ICT for Emerging Regions (ICTer), 2012 International Conference on*, pages 197–203. IEEE, 2012.
7. Chawla Nitesh V, Bowyer Kevin W, Hall Lawrence O, Philip Kegelmeyer W. Smote: synthetic minority over-sampling technique. Journal of artificial intelligence research. 2002;16:321–57.
8. Collins Jasmine, Sohl-Dickstein Jascha, Sussillo David. Capacity and trainability in recurrent neural networks. stat. 2017;28:1050.
9. Das Barnan , Seelye Adriana M, Thomas Brian L, Cook Diane J, Holder Larry B, Schmitter-Edgecombe Maureen. Using smart phones for context-aware prompting in smart environments. In *2012 IEEE Consumer Communications and Networking Conference (CCNC)*, pages 399–403. IEEE, 2012.
10. Devarakonda Aditya, Naumov Maxim, Garland Michael. Adabatch: Adaptive batch sizes for training deep neural networks. arXiv preprint arXiv:1712.02029, 2017.
11. Espinilla M, Medina J, Hallberg J, Nugent C. A new approach based on temporal sub-windows for online sensor-based activity recognition. J Ambient Intell Human Comput. 2018. https://doi.org/10.1007/s12652-018-0746-y.
12. Fatima Iram, Fahim Muhammad, Lee Young-Koo, Lee Sungyoung. Analysis and effects of smart home dataset characteristics for daily life activity recognition. The Journal of Supercomputing. 2013;66(2):760–80.
13. Fung Gabriel Pui Cheong, Yu Jeffrey Xu, Wang Haixun, Cheung David W, Liu Huan. A balanced ensemble approach to weighting classifiers for text classification. In *Sixth International Conference on Data Mining (ICDM'06)*, pages 869–873. IEEE, 2006.
14. Galar Mikel, Fernandez Alberto, Barrenechea Edurne, Bustince Humberto, Herrera Francisco. A review on ensembles for the class imbalance problem: bagging-, boosting-, and hybrid-based approaches. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews). 2011;42(4):463–84.
15. Hamad R. A, Salguero A. G, Bouguelia M, Espinilla M, Quero J. M. Efficient activity recognition in smart homes using delayed fuzzy temporal windows on binary sensors. *IEEE Journal of Biomedical and Health Informatics*, pages 1–1, 2019. ISSN 2168-2194. https://doi.org/10.1109/JBHI.2019.2918412.
16. Hammerla Nils Y, Halloran Shane, Ploetz Thomas. Deep, convolutional, and recurrent models for human activity recognition using wearables. arXiv preprint arXiv:1604.08880, 2016.
17. Huang Chen, Li Yining, Change Loy Chen, Tang Xiaoou. Learning deep representation for imbalanced classification. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 5375–5384, 2016.
18. Japkowicz Nathalie, Stephen Shaju. The class imbalance problem: A systematic study. Intelligent data analysis. 2002;6(5):429–49.
19. Jing Luyang, Wang Taiyong, Zhao Ming, Wang Peng. An adaptive multi-sensor data fusion method based on deep convolutional neural networks for fault diagnosis of planetary gearbox. Sensors. 2017;17(2):414.
20. Johnson Justin M, Khoshgoftaar Taghi M. Survey on deep learning with class imbalance. *Journal of Big Data*, 6(1):27, Mar 2019. ISSN 2196-1115. https://doi.org/10.1186/s40537-019-0192-5.
21. Kasteren TL, Englebienne Gwenn, Kröse BJ. An activity monitoring system for elderly care using generative and discriminative models. Personal and ubiquitous computing. 2010;14(6):489–98.
22. Khan Salman H, Hayat Munawar, Bennamoun Mohammed, Sohel Ferdous A, Togneri Roberto. Cost-sensitive learning of deep feature representations from imbalanced data. IEEE transactions on neural networks and learning systems. 2017;29(8):3573–87.
23. Lara Oscar D, Labrador Miguel A, et al. A survey on human activity recognition using wearable sensors. IEEE Communications Surveys and Tutorials. 2013;15(3):1192–209.
24. Li Frédéric, Shirahama Kimiaki, Nisar Muhammad Adeel, Köping Lukas, Grzegorzek Marcin. Comparison of feature learning methods for human activity recognition using wearable sensors. Sensors. 2018;18(2):679.
25. Medina-Quero Javier, Orr Claire, Zang Shuai, Nugent Chris, Salguero Alberto, Espinilla Macarena. Real-time recognition of interleaved activities based on ensemble classifier of long short-term memory with fuzzy temporal windows. In *Multidisciplinary*