User-Centric Anomaly Detection in Activities of Daily Living

Salisu Wada Yahaya

Department of Computer Science

A thesis submitted in partial fulfilment of the requirements of Nottingham Trent University for the degree of

Doctor of Philosophy

May 2021

This thesis is dedicated to my beloved parents, wife and family. You have been my sources of inspiration. I am sure you are proud of this achievement which wouldn't have been possible without your support and prayers.

Acknowledgements

The work described in this thesis was carried out at Nottingham Trent University between October 2017 and May 2021, while I was working as a full-time doctoral research student.

All praises and glory to Almighty Allah for making this a reality.

My utmost gratitude and appreciation to my director of studies, Prof. Ahmad Lotfi, for his guidance throughout the course of my research. Your consistent support, kindness and motivation is what made this achievement possible and therefore, will forever be remembered. Many thanks for making your office accessible despite your busy schedules. Also, I will like to thank my co-supervisor, Dr. Mufti Mahmud. I am grateful for the creative ideas you always had to offer and for the many corrections you made to the publications we share. Lastly, to Dr. Caroline Langensiepen for the supervision given at the early stage of my research.

My sincere appreciation to my family for their support and care through the course of my research. To my Dad and Mom, words cannot express how blessed I am to be your son. Your consistent words of encouragement and prayers have contributed in my journey to achieving this feat. I pray that you live to benefit from this achievement live to enjoy the fruit of your labour. To my wife, I sincerely thank you for your support and encouragement. To my siblings, relatives and friends, I owe you a lot and pray that you live long to benefit from this achievement.

Finally, I would like to appreciate my friends and colleagues at NTU, Specifically, my fellow PhD researchers in the Computational Intelligence and Applications research group, and to everyone who in one way or the other contributed to the success of my research, I am most grateful.

Lastly, I would like to appreciate the sponsors of this research work. This research was supported by Nottingham Trent University (United Kingdom) through Vice-Chancellor Studentship Scheme and Petroleum Technology Development Fund (PTDF, Nigeria) through the Overseas Scholarship Scheme.

Salisu Wada Yahaya
May 2021

Abstract

The current system for providing care to older adults is not sustainable due to its excessive cost. It places an unbearable financial burden on the government and families and pressure on the workforce due to the demand for human carers. Studies have also shown that older adults prefer to be looked after in their homes rather than in a care facility. An automated system of monitoring can provide much-needed support at a lower cost and give peace of mind to relatives.

The focus of the research reported in this thesis is to investigate the concept of abnormality detection in activities of daily living. More precisely, this work is aimed at proposing a dynamic approach for anomaly detection capable of adapting to changes in human behaviour. Abnormalities in daily activities can be an early indication of health decline. Therefore, early detection can inform the families of the need for intervention. Anomalies are often detected by modelling the existing activity data representing the usual behavioural routine of an individual to serve as a baseline model. Subsequent activities deviating from the baseline are then classified as outliers or anomalies. However, existing approaches suffer from a high rate of false prediction due to the static nature and the inability of the approaches to adapt to the changing human behaviour.

The contributions of the research are reported in four main categories. First, a novel ensemble approach termed "Consensus Novelty Detection Ensemble" is proposed. The outlying activities are predicted by computing their normality score using the internal and external consensus vote and the estimated weights of the models

in the ensemble. Activities with a score exceeding a threshold estimated using a statistical method based on data distribution are predicted as outliers and vice versa.

Secondly, a similarity measure approach for identifying the likely sources of the ADL anomalies is proposed. While the models can detect anomalous activities, they are unable to identify the source (cause) of the anomaly. Identifying the anomaly source allows for the development of an adaptive system. The approach is based on a pairwise distance measurement of the features extracted from the activity data. Two approaches for performing the similarity measures are presented, namely, One vs One Similarity Measure (OOSM) and One vs All Similarity Measure (OASM). Features of the data with a higher dissimilarity rate are predicted as the source.

To make the proposed model adaptive to the changes in human behaviour, a novel adaptive approach is proposed based on the concept of forgetting factors. This allows the model to forget (discard) outdated activity data and adapt to the current behavioural patterns by incorporating newly verified data. The data verification can be performed by incorporating human feedback into the system. Two forgetting factor approaches are proposed namely; Forgetting Factor based on Data Ageing (FFDD) and Forgetting Factor based on Data Dissimilarity (FFDA). The data ageing forgetting factor discard old behavioural routine based on the age of the activity data, while in the data dissimilarity approach, this is achieved by measuring the similarity of the activity data.

Lastly, the means of utilising an assistive robot as a communication intermediary is explored for incorporating human feedback into the learning process using hand gestures as a communication modality. Experimental data used for the gesture recognition model is collected using a wearable sensor and a 2D camera. The feasibility of utilising the robotic platform as an exercise coach to encourage physical activity and promote a healthy lifestyle is explored. To this end, an exercise training solution is developed for the robotic platform to coach, motivate and assess the older adults in the recommended physical activities.

Publications

As a result of the research presented in this thesis, the following publications are already accomplished.

Journal Papers

Yahaya S.W., Lotfi A., Mahmud M., Towards a Data-Driven Adaptive Anomaly Detection System for Human Activity. Pattern Recognition Letters. 2021; 145:200-207.

Yahaya S.W., Lotfi A., Mahmud, M., Detecting Anomaly and Its Sources in Activities of Daily Living. SN Computer Science. 2021; 2:14.

Yahaya S.W., Lotfi A., Mahmud M., A consensus novelty detection ensemble approach for anomaly detection in activities of daily living. Applied Soft Computing. 2019; 83:105613.

Lotfi A., Langensiepen C., Yahaya S.W., Socially Assistive Robotics: Robot Exercise Trainer for Older Adults. Technologies. 2018; 6(1):32.

Conference Proceedings

Yahaya S.W., Lotfi A., Mahmud M., Towards the Development of an Adaptive System for Detecting Anomaly in Human Activities, 2020 IEEE Symposium Series on Computational Intelligence (SSCI), Canberra, Australia, 2020, pp.

534-541.

Yahaya S.W., Lotfi A., Mahmud M., Machado P., Kubota N., Gesture Recognition Intermediary Robot for Abnormality Detection in Human Activities, 2019 IEEE Symposium Series on Computational Intelligence (SSCI), Xiamen, China, 2019, pp. 1415-1421.

Yahaya S.W., Lotfi A., Mahmud M., A similarity measure approach for identifying causes of anomaly in activities of daily living. In: Proceedings of the 12th ACM International Conference on PErvasive Technologies Related to Assistive Environments (PETRA '19), Greece, 2019.

Yahaya, S. W., Lotfi, A., Mahmud, M., A Framework for Anomaly Detection in Activities of Daily Living using an Assistive Robot. UK-RAS19 Conference: "Embedded Intelligence: Enabling & Supporting RAS Technologies" Proceedings, 2019, pp 131-134.

Yahaya S.W., Langensiepen C., Lotfi A. (2019) Anomaly Detection in Activities of Daily Living Using One-Class Support Vector Machine. In: Advances in Computational Intelligence Systems. UKCI 2018. Advances in Intelligent Systems and Computing, 2018, 840.

Lotfi A., Langensiepen C., Yahaya S.W., Active and Healthy Ageing: Development of a Robotic Platform as an Exercise Trainer. In: Proceedings of the 10th International Conference on PErvasive Technologies Related to Assistive Environments (PETRA '17). Greece, 2017.

Contents

Nomenclature

Acronyms

ADL Activities of Daily Living

AAL Ambient Assisted Living

CNDE Consensus Novelty Detection Ensemble

CNN Convolutional Neural Network

EDCV Ensemble of Detectors with Correlated Votes

EDVV Ensemble of Detectors with Variability Votes

FFDA Forgetting Factor bases on Data Ageing

FFDD Forgetting Factor based on Data Dissimilarity

HAR Human Activity Recognition

HRI Human-Robot Interaction

iForest Isolation Forest

IoT Internet of Things

LOF Local Outlier Factor

MCI Mild Cognitive Impairment

OASM One vs All Similarity Measure

OC-SVM One-Class Support Vector Machine

OOSM One vs One Similarity Measure

PCA Principle component analysis

PIR Passive Infrared Sensor

RCE Robust Covariance Estimation

RNN Recurrent Neural Networks

WSN Wireless Sensor Network

YOLO You Only Look Once

List of Figures

List of Tables

Chapter 1

Introduction

Pervasive computing has become a fabric of our daily life due to the recent advancement in Internet of Things (IoT) and miniaturisation of sensing technologies. This enabled the easy collection and transmission of data for utilisation in various application areas such as health care, home automation, behaviour modelling, industrial monitoring process, etc. [1, 5, 80, 174]. Pervasive computing has reached a level where it will be around us without even noticing its presence [170]. These advancements form the basis for the realisation of smart homes that can cater for the convenience and security of the inhabitants, giving rise to research work in ambient intelligence, especially in Ambient Assisted Living (AAL), which is largely driven by the need to improve the quality of life and wellbeing of the increasing ageing population [35, 138, 155]. Smart living environments, also known as smart homes, are now equipped with devices, and the data generated by these devices are being utilised for task relating to Human Activity Recognition (HAR) and abnormality detection [148, 155].

This research reported in this thesis investigates the concept of abnormality detection in an AAL context, specifically, a user-centric approach due to behavioural variation among individuals. *Abnormality* or *anomaly* in this context is defined as any significant deviation from the usual or normal behaviour routine of an individual [17, 23].

To better understand the relative nature of abnormality detection in this context, a scenario is given as follows: Suppose Beth and Thompson are older adults of age 77 and 64 respectively, living alone in their separate homes. Beth has a routine of sleeping for $8-9$ hours and waking up at around 8am. Followed by activities such as breakfast, bathing, watching TV and shopping where she usually returns home at noon when she often takes her lunch thereafter. She visits the restroom on average $4-5$ times a day with an average duration of $3-10$ minutes. She goes to bed at around midnight without any significant sleep disturbance. Thompson, on the other hand, has a relatively different routine as he sleeps for a short duration of 6 hours daily and visits the restroom frequently both during the day and night. He also goes to bed early but visits the restroom $2-3$ times during bedtime. Given the above scenario, if Beth sleeps for 6 hours a day or visits the restroom frequently, especially during bedtime, her behaviour can be considered abnormal since it deviates from her usual norm, while the same routine could be considered normal for Thompson.

This chapter provides an introduction to this thesis, while in Section 1.1, the background and motivation of the research work are provided. Section 1.2 gives an overview of the research methodology. The identified research questions are listed in Section 1.3 while in Section 1.4, the aim and objectives of the research are outlined. The contributions of this work are highlighted in Section 1.5. Lastly, the structure of this thesis with a summary of the chapters are given in Section 1.6.

1.1 Background and Motivation

The world ageing population (i.e. people between the age of 65 and above) is increasing, and it is estimated to be over 1.91 billion constituting around 20% of the global population by 2050 [32, 138]. However, this is accompanied by challenges such as the cognitive decline of the older adults, increase in the cost of care, and impact on the society workforce due to demand for caregivers. In the UK alone, according to the most recent data published in 2021 by the Office for National Statistics (ONS), 12.4 million people (i.e. 18.5% of the population) are between the age of 65 and above. This number is projected to rise to 19.8 million, amounting to over 26.3% of the total population by 2069 [58]. A study by the UK National Institute for Health Research (NIHR) in 2018 and data from the

UK Chief Medical Officer's annual report of 2020 shows that over 36.1 million of the older adults have at least one health challenge (morbidity). This amounts to over 54% of the ageing population. This percentage of older adults with health challenges is projected to rise to approximately 44.8 million (i.e. 67%) by 2035 with each individual having at least $2 - 3$ morbidities [57, 94, 124].

The demand for carers, as well as the cost of care, is increasing and local authorities and governments are unable to meet the financial demands [102, 120]. According to the UK National Health Service (NHS), annual spending by local authorities in the UK on social care rose by 556 million in $2016 - 2017$ thereby, resulting in a total of 17.5 billion pounds [46]. Research study has also shown that it is always a preferred option for the older adults to stay in their homes for as long as possible instead of being looked after in care home facilities [138].

Considering the fragility and susceptibility of the older adults to health-related challenges, as well the need to save the cost of care and improve their quality of life, regular monitoring of their Activities of Daily Living (ADL) becomes necessary. These are essential activities individuals must be able to perform without any intervention in order to live independently, such as mobility, maintenance of personal hygiene and continence etc. [32]. Regular monitoring can guarantee an improved quality of life for the older adults and peace of mind for the relatives. However, the current model for providing social care is becoming unbearable due to its financial burden and the demand for caregivers, which is having a significant impact on the larger economy. Automation can cost-effectively provide much-needed relief. This can be achieved by using computational intelligence techniques with data collected from the living environment to model human behaviour for different purposes such as activity recognition, fall detection, self-neglect identification and abnormality detection.

This research work is motivated by the limitations of the current systems for abnormality detection such as the high rate of false prediction and the inability of the systems to adapt to human behaviour over time [70]. Additionally, the need to promote independent living and improve the quality of life of the increasing ageing population in a cost-effective manner through the early identification of ADL anomalies has played an important role in the selection of this research. These abnormalities are significant deviations from the usual routine of the individuals which could be an early indication of health decline, and therefore, may be detrimental to wellbeing [18, 23]. For example, studies have shown that early indicators of Mild Cognitive Impairment (MCI) such as Dementia are identifiable through changes in individuals' behavioural routine including having a frequently interrupted sleep, having high activity level during the night and less activity during the day, forgetfulness and confusion in carrying out daily routines etc. [18, 23, 44, 182].

1.2 Overview of the Research

Numerous research studies have been reported in the past utilising various computational models to detect anomalies in ADL data [7, 16, 43]. Machine learning models such as One-Class Support Vector Machine (OC-SVM) [83], Recurrent Neural Network (RNN) [17], Echo State Network (ESN) [102], Convolutional Neural Network (CNN) [18] are trained on the collected ADL data. Once the ADLs are identified, they can be classified into normal and abnormal activities. The limitation of these approaches is their inability to adapt to human behaviour, which is subject to changes due to seasonal or physiological factors. Moreover, the variability of activities among individuals is not taken into consideration as human behaviour differs from one individual to another, leading to the lack of a standardised dataset representing normal and abnormal behaviours. These identified limitations affect the performance of the existing approaches leading to a high rate of false alarms, thereby making such systems of monitoring unacceptable [70, 73].

To address the above mentioned underlying shortcomings, this research proposed an adaptive approach by including human in the loop as shown in the schematic diagram in Figure 1.1. This approach differs from existing methods because it is data-driven and environment invariant. The anomaly detection model is trained on an individual basis, making the system user-centric since activities differ among individuals. The proposed approach contains different layers of components, which are described as follows:

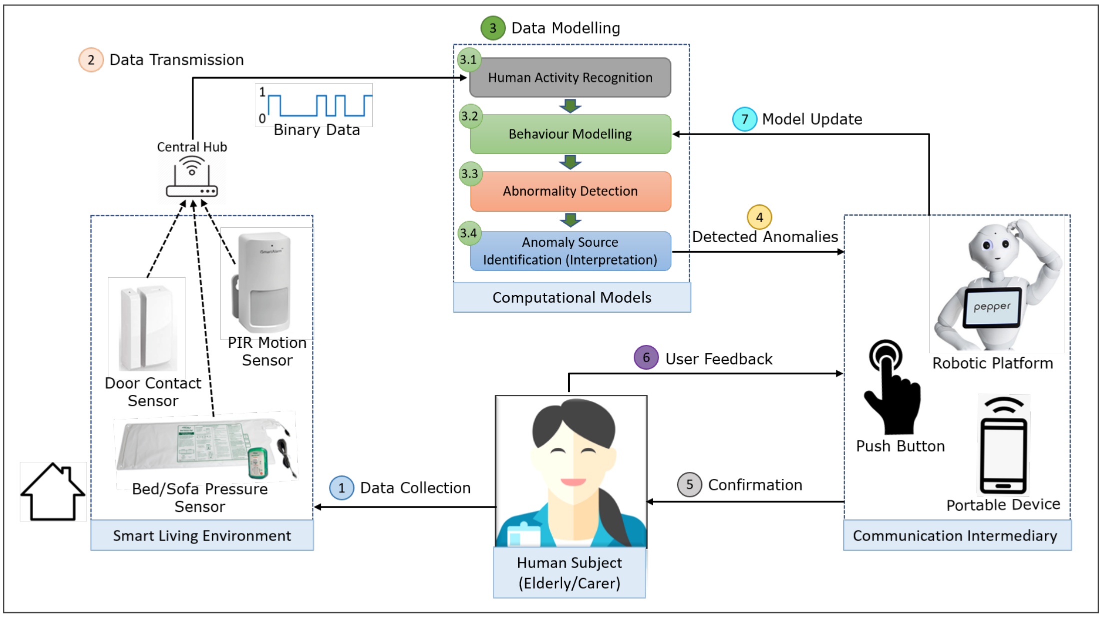


Figure 1.1: Research overview showing the different components of the methodology and the order of execution.

- The smart living environment layer deals with data collection from the living environments using ambient sensors such as motion, contact and pressure sensors. The collected data (often in binary format) representing the activity of the residents' is consolidated on a data collection hub for storage before further processing at the computational models' layer.

- The computational models' layer performs the core functionality of processing and interpretation of the collected data. Different components of this layer include a human activity recogniser that interprets the binary data into human-readable activities, with each activity having metadata such as name, location, start time and end time. The behaviour modelling component uses computational models to create a baseline model of the individuals' behaviour using extracted features of the activity data. To detect anomalies in subsequent activities, the data representing the said activity is compared to the baseline model and a significant deviation from the baseline are predicted as outliers. The outcome of the anomaly detector for any given activity is either normal or abnormal. Therefore, a component is introduced to identify the sources or reasons for the abnormality by exploring the outlying data entries. The predicted anomalies and their sources can be communicated for human re-validation.

- The communication intermediary layer interfaces the computation model layer with the human agent to communicate the detected anomalies. Different intermediaries can be incorporated, such as smart screen devices, push buttons and robotic platforms using different communication modalities ranging from voice, gesture and touch. Feedback from the human layer is returned to the computational model layer to fine-tune and reinforces the computational models.

The proposed approach has the potential of addressing the limitation of the existing methodologies as it is capable of adapting to the changes in human behavioural routine over time, thereby reducing the rate of false predictions. Additionally, due to the adaptive nature of the approach, the anomaly detection system can be effective irrespective of the environment and limited availability of training data.

1.3 Research Questions

Given the presented research overview, the major questions serving as the basis for this research work are identified as follows:

- Can human activities be modelled efficiently with an acceptable degree of accuracy while taking into consideration the variability in the behavioural routine among individuals?

- Can abnormality in human activities be identified given the data representing the usual behaviour routine of the individual in a user-centric manner? Activities predicted as anomalous for one individual could be considered normal of another individual based on their behavioural routine requiring the personalisation of the computational model.

- Since human behaviour is dynamic and subject to changes due to different factors, can the anomaly detection system be data-driven and adapt to the changes in the routine of individuals over time?

- Can humans be incorporated into the system decision-making process to reinforce the model prediction by collecting feedback through a communication intermediary? The communication mechanism also must be efficient and non-intrusive.

To address these outlined research questions, the subsequent section presents the aim and objectives of this thesis research.

1.4 Aim and Objectives

This research aims to investigate a user-centric approach to anomaly detection in human activities. This involves the recognition of human activities, behaviour modelling, and detection of changes in the modelled behaviour that could constitute an abnormality. However, instead of predicting all behavioural changes as anomalies, a human verification of detected anomalies is carried out to enable the model to adapt to changing behaviours. The activity taken as a use case for the anomaly detection relates to the bedtime (sleeping) routine of the monitored individuals to model relevant features of the activity such as the duration, pattern and irregularities in the activity. Research studies have shown that abnormalities in bedtime activity routine have a direct correlation to early signs of cognitive decline in older adults [18, 23, 44, 182]. The verification of detected anomalies can be performed using a communication intermediary such as an assistive robot platform

To achieve the above stated aim, the following objectives are identified:

1. To investigate the different approaches for interpreting human activities from data generated by low-cost ambient sensors.

2. To conduct an extensive investigation on existing methodologies for anomaly detection specifically in an AAL context to detect behavioural changes in human activities that could constitute an abnormality.

3. To propose a reliable and efficient computational model for detecting anomalies in human activities from an ambient data stream.
4. To propose a computational approach for in-depth analysis and understanding of anomalous activities and for identifying the anomaly sources.
5. To investigate a user-centric approach to anomaly detection with support for incorporating human feedback into the system for model optimisation.
6. To propose an adaptive learning model for anomaly detection utilising human feedback to adapt to changes in human behaviour.
7. To investigate and propose an efficient approach for incorporating a robotic platform as a communication intermediary, as well as other relevant use cases of the robotic intermediary.

1.5 Major Contributions

The major contributions of this thesis research are summarised as follows:

- A novel ensemble approach for anomaly detection models based on a consensus score for aggregating heterogeneous models. The output of the ensemble approach is a normality score such that activities with a lower score are predicted as outliers and vice versa. Additionally, the weights of the models are estimated based on their performance using a weight estimation algorithm, thereby giving an insight into the usability of the models on the given dataset.
- A novel similarity measure based approach for identifying the sources of anomalies in human activities. This approach explores the anomalous data to identify the extracted features that are the likely reasons for the abnormality.
- A novel adaptive approach for anomaly detection based on forgetting factors is proposed that is able to adapt to new behavioural routines and discard outdated routines.

In addition to the outlined major contributions, an approach for incorporating a communication intermediary in the anomaly detection system is explored. This is in line with the identified research question in Section 1.3 for investigating the possibility of realising an adaptive monitoring system capable of incorporating human feedback in the learning process. Although a robotic platform is selected as the communication intermediary with hand gestures as a communication modality, the most suitable platform can best be ascertained after a thorough evaluation of user preferences. Additionally, a means of utilising the robotic platform to promote physical activities that could improve the overall health of older adults is explored.

1.6 Thesis Outline

The outlined contributions of the thesis are presented in eight distinct but related chapters. The diagram in Figure 1.2 presents the organisation and dependencies of the chapters to give better insight. A summary of the chapters is given below.

Chapter 2: Literature Review - In this chapter, a comprehensive literature review is provided on anomaly detection and activity recognition in AAL context, an overview of the sensing devices used for data collection, as well as a review of gesture recognition models for utilisation on the robotic intermediary. The review covers aspects of the computational models and their performance metrics. A summary of the reviewed works, an overview of existing research gaps and a description of how the research gaps are addressed in this thesis are provided.

Chapter 3: Computational Models and Dataset - This chapter presents a general overview of existing computational models for anomaly detection and activity recognition for ADL data as well as the supervised models for gesture recognition. The second segment of the chapter presents the dataset employed for the validation of the novel methodologies proposed in this thesis. This includes details of the sensing and data collection modalities, preprocessing techniques and the evaluation metrics used to assess the performance of the models.

Chapter 4: Ensemble Model for Anomaly Detection - This chapter presents a novel ensemble model termed as Consensus Novelty Detection Ensemble (CNDE) for anomaly detection. This ensemble model allows for the

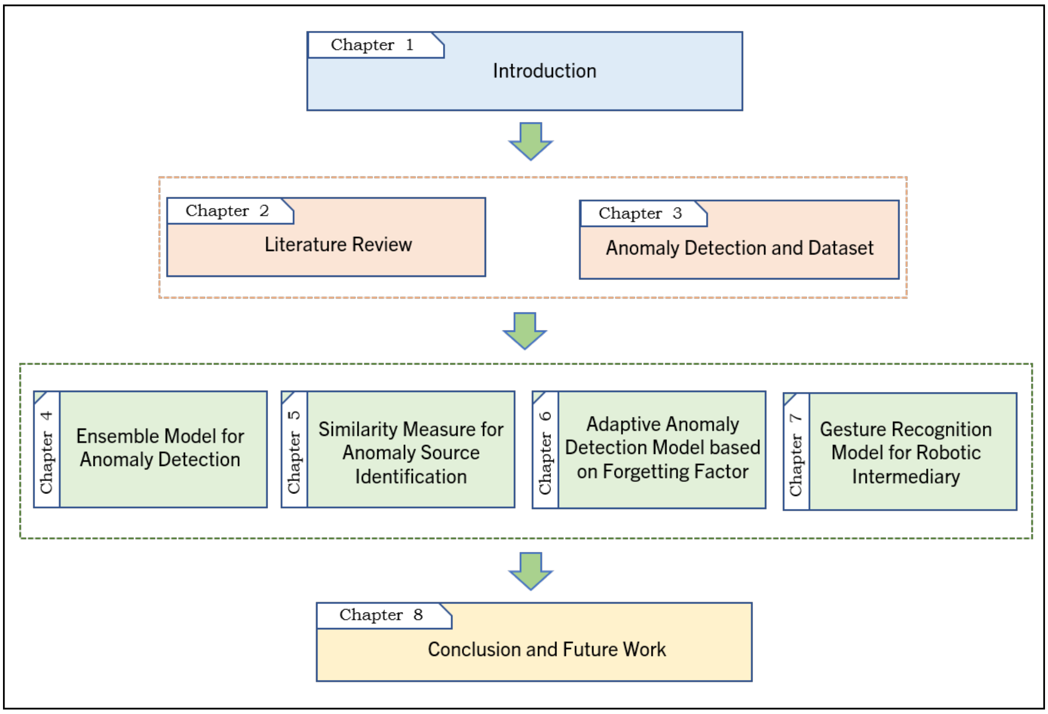


Figure 1.2: Thesis structure showing the organisation of the chapters.

aggregation of heterogeneous models based on a consensus approach and generates a normality score for the given activities. Activities with a normality score exceeding a defined threshold are predicted as outliers.

Chapter 5: Similarity Measure for Identifying Anomaly Sources - This chapter presents an approach for exploring data entries that are predicted as outliers to identify the sources of the anomaly, thereby addressing the limitation of the existing models. The proposed approach is based on distance measurement of the pairwise features of the data. Two approaches proposed for the similarity measure are termed as One vs One Similarity Measure (OOSM) and One vs All Similarity Measure (OASM).

Chapter 6: Adaptive Anomaly Detection Model based on Forgetting Factors - This chapter presents an approach for realising an adaptive anomaly detection system. This is achieved by incorporating human feedback to reinforce the model prediction. A forgetting factor concept is proposed, which allows the system to discard outdated data entries based on human feedback and adapt to new activities. Forgetting Factor bases on Data Ageing (FFDA) and Forgetting

Factor based on Data Dissimilarity (FFDD) are presented.

Chapter 7: Gesture Recognition Model for Robotic Intermediary - This chapter presents the proposed gesture recognition model for communication with the robotic intermediary. Two modalities are presented utilising a wearable tri-axial sensor and a 2D camera. Conventional classifiers and a deep learning model are used for gesture recognition. Additionally, the ability of the robotic intermediary is to serve as an exercise trainer is explored by the implementation of an application for administering physical activities.

Chapter 8: Conclusion - This chapter presents a summary of this thesis and its findings. A discussion is presented in line with the identified research questions. Recommendations for future research direction are also outlined.

Chapter 2

Literature Review

2.1 Introduction

Improving the wellbeing and quality of life of the increasing ageing population is necessary due to the ageing-related challenges, the need to provide the needed support and reduce the accompanying financial burden. Assistive technologies have the potentials to mitigate this challenge. Hence, research in the field of ambient intelligence has been getting so much attention in recent years [1, 32, 36, 97, 155]. This chapter presents a comprehensive review of related work and state-of-the-art approaches in Human Activity Recognition (HAR) and abnormality detection in an AAL context. A particular emphasis will be put on abnormality detection in ADL, including a thorough review of the challenges and research gaps in the field that paves the way for the justification of the research carried out in this thesis. Additionally, a review of approaches for gesture recognition and utilisation of assistive robot intermediary is carried out to identify its role in the proposed anomaly detection approach.

The remainder of this chapter is structure as follows: Section 2.2 presents a review of research works in HAR highlighting the different sensing modalities and some of the applied computational models. Section 2.3 gives a detailed evaluation of anomaly detection approaches specifically in an AAL context, while Section 2.4 establishes the research gaps and opportunities in this area and how they are addressed in this thesis. A conclusion of the chapter is provided in Section 2.5.

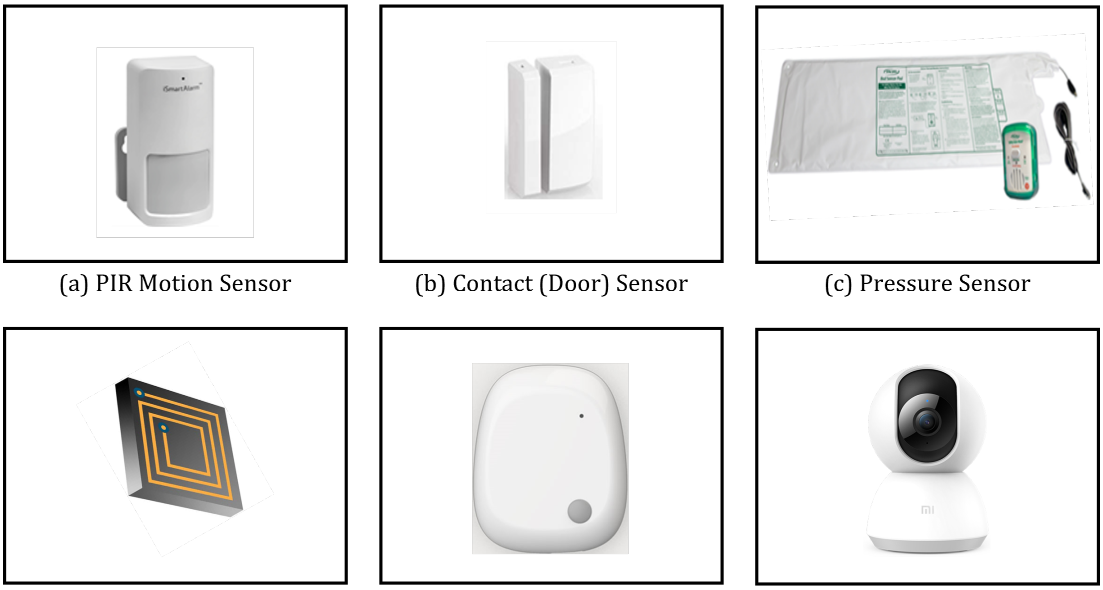
2.2 Human Activity Recognition

The application of computational models on data collected using various sensing modalities to interpret the activities performed by humans is referred to as HAR or Human Activity Recognition [72, 85, 95, 138, 163]. Obtaining accurate information on the nature of activities performed by individuals has immeasurable benefits, and its application can be found in different areas, such as medicine, entertainment, and security [1, 95]. This is achieved through the extraction of relevant descriptive features from the data, followed by its interpretation using computational models [35, 163]. Since activity recognition is an integral part of behaviour modelling and abnormality detection, a proper understanding of the different sensing modalities is required. Therefore, a review of the existing sensing modalities and computational models applicable to HAR is carried out in the subsequent subsections. Additionally, research works utilising robotic platforms for activity recognition, specifically for hand gesture recognition in line with the methodology proposed in this thesis are reviewed.

2.2.1 Sensing Modalities

The choice of sensing devices for activity recognition in an AAL context has a significant effect on the entire monitoring process and is determined by several factors, such as the nature of the monitored activities, acceptability of the modality and other physical constraints. These sensing modalities are broadly classified into two categories relative to their placement, namely; *ambient sensors* and *wearable sensors* [1].

As the name implies, *ambient sensing devices* are ubiquitous sensors installed in the living environment to monitor the users' activity without requiring any intervention [36]. A Wireless Sensor Network (WSN) is often required to interconnect the devices to a central hub for data aggregation and processing [36, 155]. This sensing modality approach is further subdivided into visual-based sensors and non-visual sensors [138]. In the vision-based sensing approach, devices capable of collecting visual data, such as cameras, are used for data collection. The recognition of activities is achieved through a series of



(a) PIR Motion Sensor (b) Contact (Door) Sensor (c) Pressure Sensor

(f) RFID Tag (f) Wearable Sensor (MMR) (g) Visual Sensor (Camera)

Figure 2.1: Some sensing devices for human activity recognition.

processing which includes background subtraction using models such as Gaussian Mixture Model (GMM) [66, 123], extraction of features from the silhouette image vector [6, 126], and the training of the classification models with the extracted features to identify the performed activity. Approaches based on deep learning tend to eliminate the problems of manual background subtraction and feature extraction by learning the appropriate data encoding and relevant features at the training phase [129]. The major downside of the visual-based sensors is that they are affected by poor lighting conditions, occlusion, and are computationally expensive [1]. Although depth cameras tend to mitigate some of the listed shortcomings (e.g. the problem of poor lighting conditions), visual-based sensors are generally not acceptable due to ethical and privacy and concerns [1, 35]. This sensing modality, however, has been employed by researchers for the identification of different human activities [2, 3, 129].

The non-visual sensing devices consist of a broad range of ambient sensors that often generates a binary stream of data. Most commonly, these devices are embedded within the home environment at a fixed position while in some cases, attached (tagged) to objects within the living environment. Examples of these sensors are: Passive Infrared (PIR) sensors, door switch sensors, pressure mats, Radio Frequency Identification (RFID) tags etc [1]. Other non-visual ambient sensors such as temperature and light intensity sensor generate numeric values representing the degree of the physical quantities they measure from the environment while devices like microphones generate acoustic waves stream [1]. These devices often form a cluster and are interconnected using a WSN. The PIR sensor allows for the identification of the location of the subjects, and in conjunction with other sensors, has been utilised for the identification of different ADLs in both single and multi-occupancy settings [49, 75, 116]. Research study has shown that non-visual ambient sensors are generally more acceptable due to the non-invasive nature of the modality [35]. Additionally, they have low power consumption, are generally cost-effective and data processing is computationally less expensive [37, 161]. Figure 2.1 shows some of the sensing devices used for activity recognition.

The *Wearable sensors* on the hand are required to be carried or worn by the monitored individuals. Devices such as RFID tags play the dual role of being an ambient and wearable sensor [1, 32]. Most commonly, wearable devices are equipped with an accelerometer and a gyroscope [1, 13, 95]. This allows for the provision of fine-grained data (usually a high-frequency time series) for the recognition of motion-related activities such as gestures and body movement [95, 169]. Modern wearable sensors are able to measure physiological properties such as heart rate and environmental properties such as temperature and humidity [95, 169]. Some major area of application for wearable devices in AAL context include food and drink intake identification [13, 14], fall detection [74], gesture recognition [11, 69], sleep analysis [44], and other daily activities of interest [19, 32]. The major downside of this sensing modality is that individuals are required to wear or carry the devices at all time. In general, wearable sensors are more favoured for the recognition of motion-based activities and activity that requires the measurement of physiological quantities, while ambient sensors are more often utilised for recognition of activities involving interaction with the physical environment or objects. There exist a hybrid approach that incorporates both modalities in the literature [10, 114].

2.2.2 Computational Models

Numerous computational models have been applied to various datasets for the interpretation of the performed activities [24, 82, 84, 117]. Given that this is merely a classification problem, supervised learning models are often utilised, while data pre-processing techniques such as data representation, segmentation and feature extraction are applied irrespective of the computational intelligence technique prior to the model training [86, 106]. Signal representation involves the removal of noise or unwanted components through filtering of the collected data while data segmentation has to do with the splitting of the data into variable window lengths (where necessary) since the data is often in the form of a continuous stream. Feature extraction performs the selection of useful features from segments of the data. These features are discriminating enough to allow the models to distinguish the different activities. Depending on the nature of the data, features in different domains are extracted, such as in the time and frequency domains. Dimensionality reduction is sometimes performed to enable the selection of optional features from the extracted ones in order to improve the classification accuracy [20].

In [24], ADLs are recognised using pattern clustering to cluster the data collected using ambient sensors. Labels are assigned to the clusters representing each activity and then an Artificial Neural Network (ANN) is used as the classifier. The authors in [117] proposed an enhanced Fuzzy Finite State Machine (FFSM) model by combining classical FFSM, Long Short Term Memory (LSTM) and Convolutional Neural Network (CNN) to learn the relationship in the temporal human activity data for activity recognition. The binary data collected using ambient sensors first undergo a fuzzification process before the model training. For the LSTM-FFSM combination, an accuracy of 95.7% and 97.6% is achieved on the two evaluation datasets, while for the CNN-FFSM combination, an accuracy of 94.2% and 99.3% is obtained for the two datasets respectively. Meanwhile, in [161], both static and dynamic Bayesian Network (BN) are used to identify 13 activities from binary sensor data. The dynamic approach outperformed the static approach in both runtime and prediction accuracy. A portable smart home infrastructure is developed in

[33] known as Smart Home in a Box equipped with various sensing devices. The authors applied Support Vector Machine (SVM), Hidden Markov Model (HMM), Naive Bayes (NB) and Conditional Random Forest (CRF) for real-time activity recognition. The work is extended in [35] while an experimental approach based on temporal relation is proposed in [82, 84].

Different ADLs, such as having a meal, maintaining personal hygiene and ambulatory actions are detected using a smartwatch with an embedded accelerometer [32]. This is a follow-up work from the previous publication by the same authors in [31]. The authors used a series of handcrafted features to train 3 classifiers; Multilayer Perceptron (MLP), Radial Basis Function (RBF) Network and SVM. The result shows that the SVM achieved the highest accuracy of 90.23%. Similar research work is carried out in [13] using handcrafted features from a tri-axial accelerometer to identify eating and drinking gestures utilising a combination of K-Nearest Neighbor (KNN), Random Forest (RF), ANN and SVM. The authors aim to monitor self-neglect behaviour through monitoring of food and drink intake, achieving a per-class accuracy of over 99%. Deep learning models are used by the authors in [14] to eliminate the need for manual feature extraction achieving nearly the same result.

As highlighted in [13, 20], the wearable sensor placement has a significant effect on the model performance and it is highly dependent on the activities of interest. This assertion is also corroborated in [1, 19, 32, 95, 169]. Wearable devices are used to recognise activities with Conditional Hidden Markov Model (CHMM) in [168]. The obtained result shows an accuracy of over 95% is achieved. The work carried out in [21] utilised a smartphone with an accelerometer to recognise different ambulatory activities and was able to achieve an accuracy of 91% using an ensemble of classifiers. The RF classifier is used by the authors in [29] with data collected from a chest-mounted accelerometer device. The authors are able to extract over 20 features, thereby obtaining a classification accuracy of 94%.

From the reviewed literature, it is evident that computational models are widely applied for data modelling in the AAL context and the model selection is highly dependent on the type of activity and data collection modality. Temporal models such as RNN and HMM are largely utilised for time series data and are more suitable for recognising sequential activities (non-static activities) that occur over a time window such as movement, hand motion etc. This is due to the models' ability to analyse sequential data and make inferences based on proceeding and succeeding entries of the time series data. On the other hand, non-recurrent models such as KNN, SVM and NB are better applied to static data since the models distinguish data entries based on feature variations. The static models also have less training time compared to the sequential models since only one data entry is analysed at a time while in the recurrent approaches, multiple entries are analysed over a time window [130].

2.2.3 Gesture Recognition Intermediary

In line with the research overview presented in Figure 1.1, a robotic platform is considered as an intermediary with hand gestures as a communication modality. Over the years, research has been conducted with the aim of equipping robots with gesture recognition capabilities [118, 119, 176]. While some approaches require the use of specialised hardware such as sensor-equipped gloves or wrist-worn devices equipped with an accelerometer and a gyroscope [119], others only utilised image and video data stream [105]. Luo et al. [105] developed a robot with the capability of recognising hand gestures corresponding to sign languages using a combination of a 2D camera and a Kinect depth sensor. One of the aims, among other things, is that, by understanding sign language, the robot can help patients with hearing impairment in a hospital environment with enquiries, navigation, and can serve as an intermediary for communicating with people that do not understand sign languages. To provide better Human-Robot Interaction (HRI), the author in [118] implemented a gesture recognition system for a mobile robot with the ability to detect 6 hand gestures using images captured with the robot's 2D camera. The authors in [181] developed a hand gesture approach for complex background based on Fuzzy Gaussian Mixture Models (FGMMs). This is a two-step process that includes the estimation of the hand pose to locate key points, and the application of the FGMM to eliminate patterns that are termed as non-gesture while classifying the correct gestures based on the key points. Similarly, in [128], a dynamic pointing gesture recognition approach is proposed for HRI. The idea is to promote a naturalistic interaction by allowing humans to control the robot movement by pointing in the direction of interest. The mobile robot tracks the hand movement using a stereo camera as the users enter its line of sight. Unlike static gestures, where the recognition is performed on a single image frame, dynamic gestures require the processing of multiple image frames over time to be able to track the movements, thereby requiring the use of temporal models. Using RGB-Depth images collected with a Kinect sensor, the work in [90] proposed an approach for extracting features used for hand gesture recognition in real-time while the authors in [128] applied Hidden Markov Model (HMM) for hand movement tracking.

While there exist different models for the recognition of gestures from images, deep learning models such as CNN are often favoured due to their ability to learn features encoding directly from data without requiring explicit feature engineering [176]. This has been applied for static hand gesture recognition in [118, 176]. For dynamic gesture recognition, one of the proposed approaches in the literature involves using the combination of CNN and RNN for the prediction. The CNN learns the encoding of the various images while the RNN learns the sequential pattern of the ordered frames [56]. Another approach has to do with the utilisation of a 3D CNN. A 3D CNN model allows features to be extracted in both spatial and temporal dimensions by performing 3D convolution. This enables motion information encoded in various frames to be captured [79]. Traditional classification approaches have also been utilised. However, unlike with deep learning models, complex pre-processing of the data is required, such as image cropping and resizing, background removal, feature extraction and normalisation etc. [118].

The use of depth sensors for gesture recognition is more advantageous in terms of prediction accuracy than 2D cameras despite the higher cost of depth-sensing devices. This is due to the presence of a depth channel in addition to RGB channels in conventional visual sensors, thereby eliminating the problem of poor lighting conditions associated with 2D cameras. In terms of the computational models, CNN based models have proven to be better in this regard since the tedious tasks of feature extraction and selection are handled by the models. For the recognition of dynamic gestures, conventional CNN-based models are not very effective since they only apply to static images. A more feasible approach is to combine CNN with temporal models such as HMM and RNN since the temporal models can analyse images sequences over a specified window length.

2.3 Anomaly Detection in Human Activities

Anomaly detection has been a subject of interest over the years due to its application in numerous domains such as medical diagnosis [152], industrial processes [45, 93, 151], surveillance [98, 165], finance and fraud detection [4], sensor networks and Information security [162, 171]. The term "anomaly detection" is used interchangeably with "outlier detection" to refer to the means of identifying unusual or novel occurrences in a given dataset [4, 135]. This has the potential for utilisation in a smart home environment for monitoring older adults routines. In an AAL context, anomaly also referred to as an abnormality, is defined as any significant change or deviation in the usual behavioural routine of an individual which could be an early indication of health-related challenges [17, 18]. Several anomaly detection models exist based on different techniques. These can be classified into an unsupervised outlier-based approach, supervised approach, cognitive health assessment approach and a hybrid approach. The subsequent sections of this chapter review these approaches in the context of related research work and their applications in the AAL domain.

2.3.1 Unsupervised Outlier-based Approach

In this approach, unsupervised models are used to identify data entries that differ from the available training set and predicts those occurrences as outliers [135]. Since the anomalies are rare instances, modelling is usually carried out on datasets that are mostly outlier-free. One approach to achieving this involves estimating the probability density function of the training data such that data in the region of high density are considered normal while those in the low-density region as considered as outliers [135]. The major drawback of this approach is that the data is assumed to be of certain distribution which is not practical. To overcome this, non-parametric approaches that estimate the distribution from the training data are proposed [166, 175]. Another approach has to do with estimating the distance between a data point and its nearest neighbours. Data points with close neighbours are classified as inliers and those with far neighbours as outliers [9, 39, 135]. It is worth noting that this technique is computationally expensive in high dimensional spaces.

Clustering algorithms such as Density-Based Spatial Clustering of Applications with Noise (DBSCAN) have been applied in this context. The authors in [73] developed an anomaly detecting system termed "Holmes" using DBSCAN clustering. The idea is to model the daily activities based on selected features by clustering them based on the days of the week. Smaller clusters are then merged based on a defined threshold, and entries that do not belong to a cluster indicate a variability, therefore, are considered as outliers. Based on the author's evaluation, the false positive rate is reduced by over 40% while the false-negative rate is reduced by 27%. In [51], an approach for monitoring change in individuals' daily routine using longitudinal data analysis is proposed. The authors grouped activities by day and labelled newly detected activities incrementally. K-means clustering is then used to separate normal routines from unusual or abnormal routines. Similarly, in [52], the number of sensor events, time and duration of activities are extracted and clustered with DBSCAN. Instances with unusual duration or irregular events are classified as anomalous since they fall outside the generated clusters [52]. The authors extended the work using a similar approach in [53]. Activity instances are defined as anomalous if it consists of irregular sensor sub-events or has an unusual execution time. An Autoencoder is then trained to identify those anomalous instances. The obtained result for the accuracy varied among activities but a minimum of 90% accuracy is achieved [53].

Conventional anomaly detection models have been utilised for ADL anomaly detection [135]. The authors in [83] have applied One-Class Support Vector Machine (OC-SVM) for detecting anomalies in daily activities. This concept allows the model to be trained on outlier free data to build a baseline with subsequent data deviating from the baseline identified as outliers. OC-SVM has also been applied in other research areas such as in the detection of seizures in patients using Electroencephalogram (EEG) data [60], prediction of patients that are at risk after undergoing surgery [152], detection of cancerous mass in images [47], and in the diagnosis of faulty vehicles [156].

Utilising infrared motion sensors (PIR) installed in a home environment, the authors in [147] are able to model the human behavioural routine. The features used in training the Support Vector Data Description (SVDD) model are calculated as the activity level, mobility level and non-response interval, and are able to achieve an accuracy of 95.8% on a real dataset. It is worth noting that the abnormality is detected over a 24 hours window. Self Organising Map (SOM) is applied to binary data collected using ambient sensors to detect anomalies [121]. Since abnormalities can be in different forms for different scenarios, the authors only focussed on three distinct cases, namely; Long period of inactivity, lacking activity and unusual changes to the activities. The model is able to predict the segment of the binary data containing the defined anomalous cases. It should be noted however that outliers are artificially induced on the dataset during the evaluation of the methodology. The work is extended in [122] since the previous method considered one activity at a time. The authors achieved an accuracy between the range of 83.8% to 99.1%.

The use of unsupervised models for anomaly detection is effective since the models are trained based on the available training data with the assumption that the data represents the normal behavioural routine of the subject. This presents an advantage over supervised models since synthetic data representing normal or outlying activities are not required. Moreover, the synthetic data can be misleading due to behavioural variation among individuals. However, in a scenario where the dataset used for the unsupervised modelling is contaminated with outliers, the computational models can produce misleading results, thereby making the approach ineffective.

2.3.2 Supervised Classification Approach

In the classification approach, anomaly detection is treated as a binary classification problem. Binary classification task requires training data for the two classes to be available [86]. However, Due to the absence of training data representing normal and abnormal activities, the approach taken is to use the collected data as the training samples for the normal class while synthetic data is then generated for the abnormal class [17, 18]. In an AAL context, the synthetic data could be generated to reflect anomalies relating to MCI. For instance, since frequently interrupted sleep and confusion in the performance of daily activities are attributed to early symptoms of Dementia, the data is generated to mimic these symptoms. The authors in [18] used this approach to artificially insert anomalous events into the collected data such as random events in between sleeping activity to reflect a disturbed sleep, and altering the order of activity sequence to simulate confusion in carrying out daily activities. A classifier is then trained on the data to identify anomalous activities. The authors used a combination of CNN and LSTM to predict the anomalies. The ADL data is first converted into a 2D image map and then fed into a CNN model to learn the data encoding while an LSTM model learns the activity sequence. Evaluation of the methodology shows that the overall accuracy of over 90% is obtained.

In [71], an approach for early detection of anomalous behaviour in smart homes based on Causal Association Rule (CAR) is proposed. The identified causes are then used with Markov Logic Network (MLN) for identifying the risk of anomaly occurrence at a later time, and for recommending suitable actions to avoid future occurrences. In [180], the authors perform a comparative analysis of LSTM, CNN, CNN-LSTM and Autoencoder for accurate detection of ADL anomalies. The validation is performed on public datasets while employing an oversampling technique to alleviate the non-uniformity in the class data distribution.

This supervised approach has not received much attention in an AAL compared to the outlier-based approach. The major drawback of supervised approaches for anomaly detection is that, for the classifiers to detect any anomaly, synthetic data simulating that anomaly must be generated. However, generating every variation of an anomaly is nearly impossible due to the variability of the anomalies, especially in ADL, since behavioural routine varies among individuals.

2.3.3 Cognitive Health Assessment Approach

An emerging approach for anomaly detection in ADLs involves the assessment of older adults' functional health based on a clinical score. This concept relies on the clinical evaluation of the older adults in areas, such as cognitive health, mobility etc. for a given period of time (usually every six months) by a domain expert, and a score is assigned to the older adults based on their performance in the assessed areas [7, 16, 43]. Using the ambient sensor data collected over the assessment period, a regression model can be trained to map a relationship between the data and the functional evaluation score. Therefore, the logic behind this approach is that by supplying the ambient data for a subsequent time window, the model should be able to assign a functional health score for the monitored older adults.

Dawadi et al. [43] proposed Clinical Assessment using Activity Behaviour (CAAB) based on this concept and applied it to 18 smart home data collected over a period of 2 years. Statistical correlation is established between the CAAB predicted score and the clinically assigned score. Alberdi et al. [16] used this same clinical assessment approach on 18 smart home data utilising clinical score based on Instrumental Activities of Daily Living-Compensation (IADL-C), which consists of a larger subset of activities than CAAB. Activities covered in IADL-C are money and self-management, home daily living, travel and event memory, and social skills. The authors did not just predict the clinical score but also predicts if there is any reliable change in the older adults activity over the evaluation period. A regression model is used to determine the health score while a classifier is used to predict the change. The work is extended in [7] after oversampling the dataset with SmoteBOOST and wRACOG to cover for the class imbalance in the dataset. In both cases, the score assigned by the regression model shows a promising result. The classification result for predicting absolute change performed poorly, but a statistical evaluation shows a correlation between the activities and the assigned clinical score. The poor performance of the classifier may be connected to the unique nature of each human behaviour.

Based on the above hypothesis, The authors in [132] analysed smart home data to predict cognitive impairment in older adults. The authors utilised datasets from Washington State University[1] which contains data of people with and without MCI to train both supervised and unsupervised models achieving a promising result. This is the same approach taken by Dawadi et al. [40, 41] using data collected from 179 volunteers and was able to obtain an AUC value of 0.64. The authors used Principle component analysis (PCA) to reduce the dimension of the feature set and then an SVM to create an automated scoring of the activity based on previously assigned values. A longitudinal study can give a better insight into the functional health assessment over a time period for trend analysis. The studies in [42] and [144] attempted to look at longitudinal evaluation and its effect on the assessment score but the works are at a preliminary stage.

In general, this approach based on cognitive health assessment presents a different dimension to anomaly detection but a more user-centric approach is required and could be more effective since health decline varies among individuals. Another major drawback of this approach is the timing of the health assessments. The assumption is that the assessment reflects the health condition of the adults over a time period. However, the assessment may be performed at a time when the older adults are in a state of health that differs from their day to day health status. Given that the modelling is performed based on the health assessment score, the prediction of subsequent activities by the computational models may be inaccurate and misleading.

2.3.4 Hybrid Approaches

Any other approach not included in the three methods outlined earlier are classified as hybrid approaches and presented in this section. In this are approaches based on sequential or deep learning models as well as aggregation methods that combine different methodologies. A sequential modelling approach employing a HMM is used to learn the activity sequences over a time period and classify sequences that do not conform to the usual order as anomalies while adding a layer consisting of a Fuzzy Rule-Based System (FRBS) to infer if the identified sequences are actual outliers [59]. This is similar to the approach

[1] http://casas.wsu.edu/

proposed by Jakkula et al. [81]. However, in this case, the emphasis is more on the temporal relation between activities rather than the sequential order. The modelling of normal household activities to detect outliers using HMM is also performed in [137] with data collected using ambient sensors. The HMM is trained with the data considered to be normal while the data considered as anomalies are used for validation achieving an overall accuracy of 97%. In [158], HMM is used to analyse sleeping patterns with the Viterbi algorithm, although a detailed explanation on how the activity sequences are used is lacking as well as a comparative analysis with other models.

To distinguish fall from other similar ADL anomalies, using a wearable device, an approach based on HMM is proposed in [92] which is termed as 'X-Factor' Hidden Markov Model(XHMM). Similarly, a combination of LSTM and HMM is used for modelling sequential activities in a home environment to detect anomalies. The two evaluation datasets each consist of 10 activities, and the evaluation shows that in both scenarios, LSTM achieved a better result compared to HMM [136]. The authors highlighted that LSTM requires at least five (5) inputs representing distinct activities to achieve optimal performance. Approaches that rely on the sequential or temporal order of activities presents a good opportunity but more often, the same activity can be performed in a different sequence. An adaptive learning model could mitigate this shortcoming by learning all the possible sequences for the given activities. Howedi et al. [76] utilised ADL data collected with ambient sensors to detect abnormal activity instances using different entropy measures. Activities with entropy value exceeding a certain range are predicted as outliers. The authors first calculated the entropy values of the different activities and used a statistical method to estimate the optimal threshold. The entropy values of the test data are then compared to the estimated threshold to identify the anomalies. An accuracy of 100% is obtained in some cases after evaluation of the approach [76]. A drawback of this approach is that decision is made based on calculated threshold and no learning is performed. If the training data is contaminated with outliers, the performance cannot be guaranteed. Abnormality detection algorithms based on a Hidden State Conditional Random Field (HCRF) are proposed in [159]. The approach taken is to model the activity occurrence sequences such that if an activity is missing from the sequence, or newly added activities are observed, the distance of the feature vector is computed for the sub-activities to identify if the occurrence constitutes an abnormality. The evaluation of the method shows an achieved accuracy between the range of 70% - 95% [159].

The authors in [15] proposed a system for promoting the safety of autistic children in an indoor environment. This system is based on a 3D-CNN and LSTM applied to a video stream to predict activities with physical irregularities. The CNN model extracts spatio-temporal features from the data stream while the LSTM model calculates the temporal relationship to identify the irregularities. This approach is applicable for older adults monitoring since it allows deviating activity to be identified. To promote the safety of older adults, a context-aware approach is proposed in [38]. The authors used ambient sensors to identify occupancy of the home environment and performed a tree-based matching with the usual occupancy history of the residents for the given time. If the behaviour deviates from the usual resident's routine, an alarm is triggered to indicate possible intrusion (e.g. by burglars). Meanwhile, a Kinect depth sensor is used in [133] for identifying deviating activities that could constitute an abnormality in a smart home environment. To achieve this, a data-driven technique is used to define fuzzy sets over attributes of the occupant's behaviour and a fuzzy inference engine with a membership function is used to identify the abnormal patterns.

In [125], a system is developed to assist the disabled and older adults by learning their behaviours routine using Bayesian statistics. The authors introduced three probabilistic features termed 'sensor activation likelihood', 'sensor sequence likelihood', and 'sensor event duration likelihood'. These probabilistic features are used to detect activities that are likely to be anomalous or not. The evaluation of the approach using the three components achieved a specificity of 0.98 and 0.99 on real and synthetic data respectively. A similar system for behaviour modelling is developed by Eisa and Moreira [48] using HMM and CRF. The system learns the daily location to location transition of the older adult termed as 'daily mobility' using primarily motion sensors. A separate model for each location occupancy is built consisting of the usual arrival time and duration spent at the location daily. A related study on detecting anomaly using long-term data is performed in [179] by monitoring behavioural changes over a time period. The author did not use hourly or daily behaviour modelling but rather analysed the overall behavioural changes over a long period. A temporal similarity score is then defined for the activities such that once an anomaly is identified, the timeline is analysed to identify the outlying activities. The methodology is implemented based on an FRB approach, and achieved an accuracy of 96%. This work is built on top of the previous work by the same authors in [178] that attempts to analyse long-term data for anomaly detection using DBSCAN.

Research studies to identify abnormalities that could be early indicators of MCI have been conducted [18, 62, 87, 102, 142, 143]. In the work carried out in [142], sensors are attached to the appliances in the environment and the interaction are recorded and modelled along with other ambient sensors' data using statistical techniques. The proposed algorithm for detecting the anomalies is named Fine-grained Abnormal BEhavior Recognition (FABER) and is based on modelling activity sequences. The work is extended by proposing a SmartFABER in [143]. The difference between the two approaches is that SmartFABER incorporates a machine learning model and therefore can handle sequences with unaligned activities. Meanwhile, a rule-based approach is adopted in [87] using rules that are defined by a domain expert. Although precision and recall values of 0.97 and 0.85 are obtained, the approach is very subjective and may not generalise in a wider context. Lotfi et al. [102] use Echo State Network (ESN) for the detection of anomalies in ADL from the raw binary sensor data.

Grewal et al. [67] took a relatively different approach to detect anomalies in a smart home using an electricity usage pattern. Although this is not directly targeted at monitoring older adults, it can be utilised for the said purpose. The authors analysed the usage of electricity by dividing the days into time window segments and creating a cluster using Local Outlier Factor (LOF). If the power usage deviates from the usual usage of the given time slot, an abnormality is predicted. It is worth highlighting that these computational models for anomaly detection have also been applied in non-AAL domains. In [148], anomalies in human activity are detected from a video stream using a CNN with an adaptive compression technique. Similarly, a CNN based approach for anomaly detection in images is proposed in [63]. An approach based on RNN named Single-Tunnelled Gated Recurrent Unit (SiTGRU) for detecting anomalies in video streams is proposed [54] while an attempt at detecting anomalies in users' activities by finding deviations from their routine pattern has been carried out using a Random Forest classifier [104].

Ensemble of anomaly detection models often leads to enhanced performance compared to a single model since each model is good at certain characteristic features. Dib et al. [45] applied an ensemble of machine learning models to monitor structural health by detecting damages using guided waves generated by the building sensors. The authors utilised this approach to foresee a possible structural failure over a certain period and compared it with the recently recorded data. In an AAL context, the authors in [131] presented an ensemble approach based on similarity measures for novelty detection models, and a promising result is obtained after the model evaluation.

One of the major problems in the anomaly detection domain has to do with the availability of labelled normal (outlier-free) data. In some scenarios, the available data are partially labelled while a significant portion of the unlabelled data may contain outliers. To obtain a reliable model, the impurities in the unlabelled data must be eradicated or reduced to a bare minimum. The authors in [89] proposed Positive and Unlabeled Metric learning for Anomaly Detection (PUMAD) that can learn from unlabelled data as long as there exists a small number of partially labelled samples. Evaluation of the methodology against other models, such as Deep Autoencoder shows that the approach achieved a better result. This approach has the potential of being applied in other domains, however, a large feature set could affect the efficiency of the approach since distance-based filtering of the data is required. Another seeming challenge is that most systems for anomaly detection in ADL often operate in an offline manner. To address this limitation, an online learning approach could be explored. Meng et al. [115] proposed an Online Daily Habit Modelling and Anomaly Detection (ODHMAD) model capable of performing real-time activity recognition and anomaly detection from ambient sensor data.

The combination of different computational models based on their unique strengths makes the hybrid approach applicable for ADL anomaly detection. For example, The Online Daily Habit Modelling and Anomaly Detection (ODHMAD) framework proposed by Meng et al. [115] can be utilised in conjunction with both supervised and unsupervised models to achieve a real-time system for detecting anomalies. The limitation of the framework as well as the other existing approaches is the inability to support human intervention in the learning process for newly acquired data entries. The continuous adaption of the computational models to new data entries can opens room for errors since the models could adapt to outlying data. The Positive and Unlabeled Metric learning for Anomaly Detection (PUMAD) approach proposed in [89] suffers from the same limitation relating to outlying data entries in the dataset. This can be improved by eradicating outlying data entries from the datasets which is nearly impossible since outliers in human activities vary among individuals. The approach taken in this thesis to overcome these limitations is to make the anomaly detection approach user-centric and adaptive to new data entries. However, the adaptation approach incorporates human feedback as a filtering mechanism to eliminate outliers in the new data entries.

2.4 Discussion and Research Opportunity

Findings from this review of relevant literature indicate a research gap in current approaches for anomaly detection in the AAL context. The cognitive health assessment approach for detecting anomalies presents a meaningful way for monitoring the health of older adults and detecting early indicators of MCI. However, since the approach relies on an evaluation and scoring by domain experts over a time period, the assessment could be subjective. Moreover, the review period could have a significant effect on the methodology, for example, the initial health assessment may be performed when the older adults are in a good or bad state of health, and therefore, the assessment may not be an accurate representation of the individuals' cognitive capabilities.

The major downside of the classification based approach is the lack of sufficient training data representing the normal and abnormal behaviours since the classifiers require the training data for both classes. The lack of a standardise dataset is due to the variability in the behavioural routine among individuals. Although synthetic data are generated to simulate different anomalies, the approach cannot be generalised since abnormalities could be in different forms and user-dependent. Additionally, it is nearly impossible to generate data for every form of ADL anomalies. The unsupervised outlier-based approach has the potential of addressing the identified limitations since the model requires only one set of data representing the normal activity routine. It is worth noting that contamination of the training set, however, can have a significant effect on the model performance. Moreover, research studies carried out seem to focus more on building a generic model, leading to a high false alarm rate as a result of lack of adaptability of the models and behavioural variation [70]. Additionally, the inability of the approaches to provide a detailed description of the sources or reasons for the outlier prediction presents a bottleneck in the development of an adaptive system.

Considering the above-listed shortcomings, this thesis investigates a user-centric approach to anomaly detection. This is based on the outlier-based approach to model the usual behavioural routine on an individual basis. To counter the effect of behavioural changes due to seasonal or other factors, the approach proposed in this thesis is adaptive and data-driven, therefore, behavioural changes are easily incorporated. Additionally, this user-centric approach is environment invariant, therefore, learning is performed incrementally even when there is insufficient training data. To adapt to novel data and distinguish actual anomalies from normal behavioural change, a communication intermediary for transmitting user feedback is incorporated using a user-in-a-loop approach, hence the terminology "user-centric approach". While wearable sensors are generally adopted for motion-based activities, ambient sensors are generally used for in-home activity modelling. This work utilised data collected using non-visual ambient sensors since it is generally more acceptable compared to visual-based sensors [33, 138]

2.5 Conclusion

This chapter presents a comprehensive review of related research works on abnormality detection in ADL as well as activity recognition. The review also highlights the sensing modalities and computational methodologies for recognising daily activities from sensory data which serves as the basis for behaviour modelling. The research gap in the existing literature and shortcoming of the current approaches are explored to proffer an efficient solution. Finally, the approaches taken in this thesis to address the limitations of existing models are presented which are based on a user-centric approach to anomaly detection. Given the justifications provided for the selected approach, it has the potential of addressing the identified shortcomings and the realisation of a reliable in-home monitoring system for older adults. The next chapter presents a detailed description of some of the computational models for anomaly detection and a description of the datasets employed for the validation of the approach proposed in this thesis.

Chapter 3

Computational Models and Dataset

3.1 Introduction

Machine learning models create a mapping of input data to the desired output using mathematical functions [99, 153]. For a classification problem, the input data are mapped to the class labels that are known beforehand [109, 111, 153]. Unsupervised models for outlier detection undergo a similar training routine to identify a cluster boundary separating the normal data from the anomalous data (outliers) [131, 135, 153]. This requires a sufficient amount of training data to achieve an acceptable result. In a scenario with limited training data or where the data is overly dynamic and changes over time, these conventional models can be inefficient.

The task of learning the behavioural routine of an individual and detecting abnormalities in it is rather arduous. Moreover, the ADL data representing human behaviour vary from one individual to another. Novelty Detection algorithms can be used to model the ADL data representing the individual daily activity routine to serve as a baseline. Subsequent activities can be compared to the baseline model to detect deviation, which can be an indication of abnormality. The efficiency of the computational models is also dependent on the quality of the validation datasets. An understanding of the models and how it applies to the detection of ADL anomalies is required for the justification of their exploration throughout this thesis.

This chapter is structured as follows: Section 3.2 presents some of the utilised anomaly detection models while Section 3.3 presents the classification models for gesture recognition aspect of this research. Section 3.4 presents an overview of the data collection approach including a brief overview of the sensors, data collection scenarios, and a description of the datasets. The pre-processing techniques applied to the data are presented in Section 3.5 while the model evaluation metrics are discussed in Section 3.6. Lastly, a summary of the chapter is presented in Section 3.7.

3.2 Anomaly Detection Models

Anomaly Detection is a process of identifying the unusual pattern of new or unknown data. This is interchangeable with Novelty Detection, which involves detecting if a test data differs significantly from the data available during training [135]. Unlike in binary or multi-class classification approach where data for the different classes are available during training, in novelty detection, only one set of data is available. This is also referred to as One-Class Classification [135]. This methodology has proven useful in a scenario where there is sufficiently large training data representing the normal class and little or no training data for the anomalous (abnormal) class. This concept is relevant to ADL anomaly detection since abnormalities in ADL are rare, and data representing the anomalous cases are not readily available. Novelty detection enables a model to be fitted into the normal training data and subsequent data to be compared to the model to detect abnormalities that do not conform to the built model. Some machine learning models based on this concept are One-Class Support Vector Machine (OC-SVM), Local Outlier Factor (LOF) and Isolation Forest (iForest). These models are employed in the experimental chapters of this thesis.

The models above are selected due to their suitability for the ADL anomaly detection scenario after considering factors such as the size of the dataset, class label of the data entries, explainability of the models' output, and the feasibility for continuous data adaptation. Since the available training data is composed of only data representing normal activity routines, supervised learning approaches cannot be utilised, hence the most feasible approach is the use of novelty detection models [135]. Other machine models for outlier detection such as Self Organising Maps (SOM), Echo State Network (ESN) and Autoencoders require a large amount of training data, thereby making them unsuitable for ADL anomaly detection due to the limited training data size [102, 121, 141]. Additionally, the proposed approach in this thesis aims to promote the explainability of the model's prediction by identifying the sources of the abnormality. This cannot be achieved using models based on deep learning since features of the dataset that could be the anomaly sources are extracted by the models.

3.2.1 One-Class Support Vector Machine

The OC-SVM is an unsupervised novelty detection algorithm proposed by Scholkopf et al. [146]. To identify outliers, assuming the data has an underlying probability distribution P, the problem can be framed into a minimisation of a quadratic function.

Let $A = \{a_1, ..., a_n\}$ be an n set of d-dimensional data ($A \in \mathbb{R}^d$), and $\Phi : \mathbb{R}^d \to F$ be a non-linear mapping from a data space $\mathbb{R}^d$ to a feature space F, The support vector separating the data is computed by solving the quadratic problem:

$$\min_{w \in \mathbb{F}, \xi \in \mathbb{R}, \rho \in \mathbb{R}} \quad \frac{1}{2}||w||^2 + \frac{1}{vn}\sum_{i=1}^{n} \xi_i - \rho \tag{3.1}$$

subject to:

$$(w \cdot \Phi(a_i)) \geq \rho - \xi_i, \quad \xi_i \geq 0, \quad i = 1, ..., n$$

where n is the training sample size, $v \in (0, 1)$ is a trade-off parameter for the expected fraction of outliers, ξ is a slack variable, w and ρ are the hyper-plane parameters in the feature space F.

To solve Equation 3.1, the decision function $f(a)$ which determines if a_i is an outlier is obtained as:

$$f(a) = \sum_{i=1}^{n} \alpha_i k(a_i, a) - \rho \tag{3.2}$$

where α_i is a Lagrange multiplier for the vectors a_i and $k(a_i, a) = \Phi$ is the kernel function for the non-linear mapping.

$$k(a_i, a) = \Phi(a_i) \cdot \Phi(a) \tag{3.3}$$

The Radial Basis Function (RBF) kernel with spread parameter γ is used:

$$K(a_i, a) = exp\Big(- \gamma ||a_i - a||^2 \Big) \tag{3.4}$$

$$f(a) = \begin{cases} +1 & a \text{ is NOT an outlier} \\ -1 & a \text{ is an outlier} \end{cases} \tag{3.5}$$

The prediction of the model is determined from Equation 3.3 and Equation 3.5. A more details description of OC-SVM can be found in [146].

3.2.2 Isolation Forest

The Isolation Forest (iForest) is based on the premise that outliers are few and different, making them susceptible to isolation [101]. This makes the detection of anomalies based on the concept of isolation rather than using distance or density measures.

Let $A = \{a_1, ..., a_n\}$ be an n set of d-dimensional data ($A \in \mathbb{R}^d$). An isolation tree (iTree) is built from subsample instances ψ of $A' \subset A$ by recursively dividing A' with a randomly selected attribute q and a split value p, until either: (i) the node has only one instance or (ii) all data at the node have the same values. The iTree is a binary tree with each node having exactly zero or two child nodes. Assuming all instances are distinct, each instance is isolated to an external node when an iTree is fully grown, the number of external nodes is ψ, the number of internal nodes is $(\psi - 1)$, and the total number of nodes of an iTree is $(2\psi - 1)$. To detect outliers, a ranking of the data that reflects the degree of anomaly is performed through the sorting of the data points according to their path lengths or anomaly scores. The anomalies are points that are ranked at the top of the list. The different terminologies for the iForest are defined as follows:

- *Isolation Tree (Definition):* Let T be a node of an iTree. T is either an external node with no child, or an internal node with one test and exactly two child nodes (T_l, T_r). A test at node T consists of an attribute q and a split value p such that the test $q < p$ determines whether the data point belongs to either T_l or T_r.

- *Path Length (Definition):* The Path Length $h(a)$ of a point a is a measure of the number of edges a traverses an iTree from the root node to an external node such that a short path length indicates a high degree of susceptibility to isolation and a long path length indicating low susceptibility.

- *Anomaly Score (Definition):* The Anomaly Score $s(a, \psi)$ for a data point a of data set with n instances is given as:

$$s(a, \psi) = 2^{-\frac{E(h(a))}{c(\psi)}} \tag{3.6}$$

  $E(h(a))$ is the average path length $(h(a))$, $c(\psi)$ is the average path length of unsuccessful binary search, and ψ is the subsample size from $A = \{a_i\}_{i=1,\ldots,n}$.

$$c(\psi) = 2H(\psi - 1) - 2(\psi - 1)/\psi \tag{3.7}$$

  where $H(i)$ is a harmonic number estimated as $ln(i)+0.5772156649$ (Euler's constant is 0.5772156649).

The interpretation of the outcome of the anomaly score is summarised below with a more detailed discussion in [101].

- The instances are anomalies if the value of s is close to 1 $(E(h(a)) \rightarrow 0, s \rightarrow 1)$.

- The instances are not anomalies if the value of s is much smaller than 0.5 $(E(h(a)) \rightarrow (\psi - 1), s \rightarrow 0)$ and
- The entire instances have no distinct anomalies if the values of s is approximately 0.5 $(E(h(a)) \rightarrow c(\psi), s \rightarrow 0.5)$.

3.2.3 Local Outlier Factor

The Local Outlier Factor (LOF) is an unsupervised approach for outlier detection that is based on the density estimation of a data point relative to its nearest neighbours. Since the density around an outlier differs significantly from the density around its neighbours, data points with relatively lower density to its neighbours estimated in the form of a score are considered as outliers [26].

Let $A = \{a_1, ..., a_n\}$ be an n-sample dataset. Let $d(a, b)$ denote the distance between two objects a and b. Assuming $a = a_i$ is an observation in A, the k-distance of a denoted by $D_k(a)$ is the distance between a and its k-nearest neighbours while the k-distance neighbourhood of a denoted by $N_k(a)$ contains every element whose distance from a is not greater than k-distance $D_k(a)$ of a given as:

$$N_k(a) = \{x \in A | d(a, x) \leq D_k(a)\} \tag{3.8}$$

where x is an entry in A, which is a neighbour of a. The k-reachability distance $RD_k(a)$ of the object a with respect to its nearest neighbours a' is given as:

$$RD_k(a, a') = max\{D_k(a), d(a, a')\} \tag{3.9}$$

The local reachability density $LRD_k(a)$ of a is given as:

$$LRD_k(a) = \frac{|N_k(a)|}{\sum\limits_{a' \in N_k(a)} RD_k(a, a')} \tag{3.10}$$

The average of the ratio of the local reachability density of a and its k-nearest neighbours is presented as:

$$LOF_k(a) = \frac{\sum\limits_{a' \in N_k(a)} \frac{LRD_k(a')}{LRD_k(a)}}{|N_k(a)|} \tag{3.11}$$

The lower the local reachability density $LRD_k(a)$ of a the higher the local reachability density $LRD_k(a')$ of its nearest neighbours a', the higher the LOF $LOF_k(a)$, making the a an anomaly. A more detailed discussion of LOF can be found in [26].

3.2.4 Covariance Matrix Estimation

In this approach, outliers are detected using the Maximum Likelihood Estimators (MLE) for the mean and the covariance matrix of the normal (non-outlying) data. Observations with abnormally large distances (e.g. Mahalanobis distances) are predicted as outliers [145].

Let $A = \{a_1, ..., a_n\}$ be an n-sample dataset of d-dimension and each observation $a_i = \{a_{i1}, ..., a_{id}\}$. Assuming the data follows a multivariate normal distribution, The mean ($\hat{\mu}$) and covariance matrix ($\hat{\Sigma}$) estimated using MLE are given as:

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} a_i \tag{3.12}$$

$$\hat{\Sigma} = \sum_{i=1}^{n} (a_i - \hat{\mu})(a_i - \hat{\mu})^T / (n-1) \tag{3.13}$$

The Mahalanobis distance of the observations are computed as:

$$MD(a, \hat{\mu}, \hat{\Sigma}) = \sqrt{(a - \hat{\mu})^T \hat{\Sigma}^{-1} (a - \hat{\mu})} \tag{3.14}$$

The estimations above assume that the observations are not contaminated as few outliers could lead to poor performance. Alternative robust approaches termed as Robust Covariance Estimation (RCE) for estimating the location (mean) and scatter (covariance) such as using Minimum Covariance Determinant (MCD) are proposed [145].

3.2.5 Ensemble of Detectors Based on Similarity Measure

Ensemble approaches allow for the aggregation of individual models for better performance. Ensemble models for outlier detectors based on similarity measures are proposed in [131]. The authors proposed two approaches, namely, Ensemble of Detectors with Correlated Votes (EDCV) and Ensemble of Detectors with Variability Votes (EDVV). The proposed approaches merely estimate the appropriate weight for each model in the ensemble using the individual model's results for a given dataset. The weight estimation in EDCV is based on a correlation coefficient while the weight estimation in EDVV is based on a measure of Mean Absolute Deviation (MAD) of the individual model's output. Using the estimated weights, a score for the data entries are calculated to determine if the data points are outliers.

Let $A = \{a_1, ..., a_n\}$ be an n set of d-dimensional data ($A \in \mathbb{R}^d$). Let $G = \{g_1, ..., g_m\}$ be a set of m outlier detection models. a_i is an observation in A and g_i is a model in G. The weight $W(g_i)$ for g_i is estimated using Equation 3.15 and Equation 3.16 for the EDCV and EDVV respectively.

$$W(g_i) = \frac{(\sum_{i=1}^{m} \mathbb{C}_i) - 1}{m - 1}, \text{ for EDCV} \tag{3.15}$$

where $\mathbb{C}$ is a matrix of the correlation coefficients of A for the corresponding model g_i and m is the number of models in the ensemble.

$$W(g_i) = \frac{(\sum_{i=1}^{m} \mathbb{D}_i)}{m - 1}, \text{ for EDVV} \tag{3.16}$$

and $\mathbb{D}$ is a matrix of the MAD of A for the corresponding model g_i and m is the number of models in the ensemble

The final score for a data entry a_i is calculated as:

$$S(a_i) = \frac{(\sum_{j=1}^{m} F(i,j) * V(i,j) * W(j))}{m} \tag{3.17}$$

$S(a_i)$ is the final outlier score, m the number of models, $F(i,j)$ is the outlier score, $V(i,j)$ is the label (output), and $W(j)$ is the weight of the j^{th} model respectively.

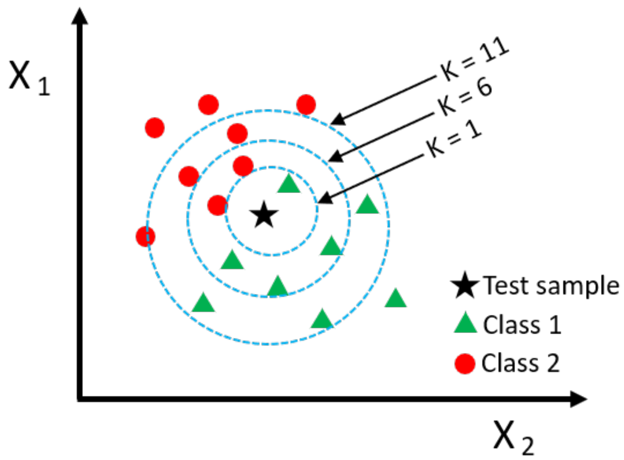


Figure 3.1: A binary KNN classification model with variable k sizes (i.e. k = 1, 3, 11).

3.3 Classification Models

The classification models are utilised for gesture recognition in this research. The task of identifying the gesture performed by the human agent involves the creation of a mapping of input variables X to a set of predefined outputs Y based on some defined mathematical functions. This supervised approach differs from anomaly detection earlier described since data for all the gestures (classes) are available during training. Once the models are trained with segments of the data, they are then used to predict subsequent gestures. This section presents two of the models used for gesture recognition, namely, KNN and a CNN implementation known as YOLO.

3.3.1 K-Nearest Neighbor

The KNN is a memory-based supervised model that predicts the class of a given data point x based on the labels of its nearest neighbours as shown in Figure 3.1 and therefore, does not require the model to be fit. The proximity of the k number of neighbouring data points is based on a distance function (usually Euclidean distance) [86].

Given a data point x to be classified, the k nearest data points $x_a, a = 1, ..., k$

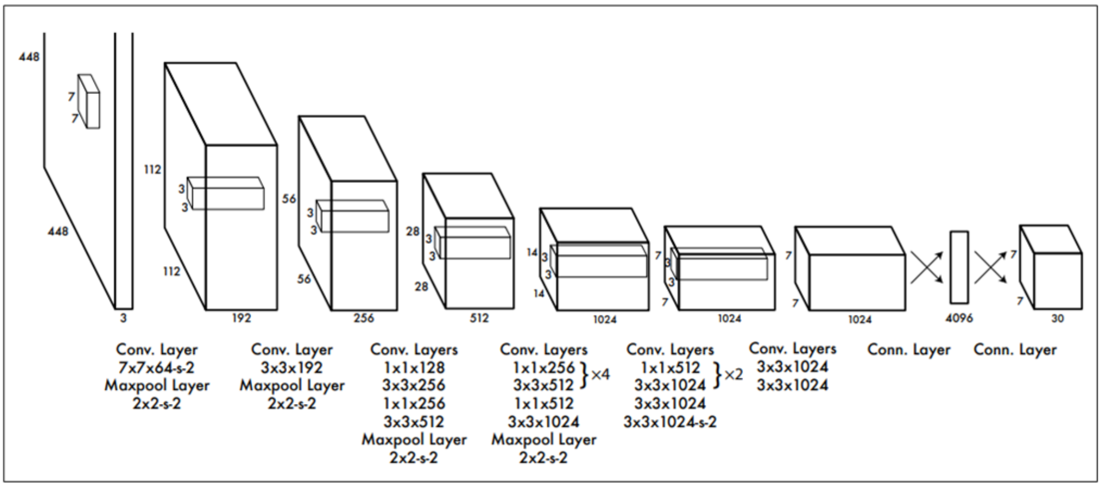


Figure 3.2: YOLO Architecture (Redmon et al. [139]).

are searched based on their distance to x. The class of x is determined based on the majority vote among the k neighbouring points.

3.3.2 YOLO

The You Only Look Once (YOLO) is an implementation approach of CNN for detection of objects in images proposed by Redmon et al. [139]. Over the years, different versions of YOLO has been developed and recently, the third version known as YOLOv3 is released with several performance improvements compared to its predecessors. In YOLO, a single CNN trained on full images is used to make predictions of the multiple objects specified simultaneously. The input image is divided into an M x M grid of equal sizes. Each of the grid cells predicts B bounding boxes around the detected objects of interest along with their respective confidence scores. The bounding boxes each consists of 5 values corresponding to the x and y coordinates representing the centre of the bounding box, the width and height of the bounding box relative to the image dimension, and the confidence score, respectively [139].

The YOLO architecture depicted in Figure 3.2 consists of 24 convolutional layers for extracting features from the image and 2 fully connected layer for making a prediction and generating the bounding boxes.

3.4 Dataset Description

To effectively utilise the presented computational models for anomaly detection, data representing the human behavioural routine is required. The data of interest for this research are collected from a smart living environment using non-visual low-cost ambient sensors installed in different locations of the living environment. As mention in the previous chapter, these sensors generate binary data and are generally more acceptable for monitoring purposes because they pose less privacy concern compared to visual-based ambient sensors [35]. Despite the high processing requirement and privacy-related concerns associated with the visual sensing devices, they provide the richest form of data for ADL monitoring and are particularly utilised for life-threatening scenarios such as fall detection [107, 127, 167]. This is because the data generated by visual-based sensors contain relevant information for posture identification that allows for easy identification of fall-related anomalies compared to non-visual sensors [177]. To address the privacy concerns, researchers have utilised depth-sensors [61, 108], image anonymisation approach [107], and a hybrid approach for combining both visual and non-visual sensors for fall detection [77, 127]. The hybrid approach presents an avenue for controlled monitoring such that visual sensors are only activated when an unidentified abnormal event is detected using the non-visual sensors. From the schematic diagram presented in Figure 1.1, it can be seen that the proposed methodology incorporates human feedback through a robotic intermediary. The feedback modality is via a hand gesture communication. Data for training the gesture recognition models is collected using a 2D camera and a wearable tri-axial sensor. A brief overview of the sensing devices utilised in this research are presented as follows:

- *Passive Infrared (PIR) Sensor:* It is also known as motion sensor, which detects the movement of a living embodiment in its field of view when an infrared light radiates from the object. It is used to identify the presence of individuals in a given location.

- *Pressure Sensor*: The pressure sensor is usually placed on the sofa or bed to detect the presence of a body mass. This is used to infer the activity performed by the resident such as sitting or sleeping.

- *Entry Door Contact Sensor*: This binary sensor comes as a pair and is placed on objects that support opening or closing such as doors and windows. It is triggered when the proximity of the pairs exceeds a certain distance threshold. This is used to identify if an object (such as a door, fridge or microwave) is opened or closed.

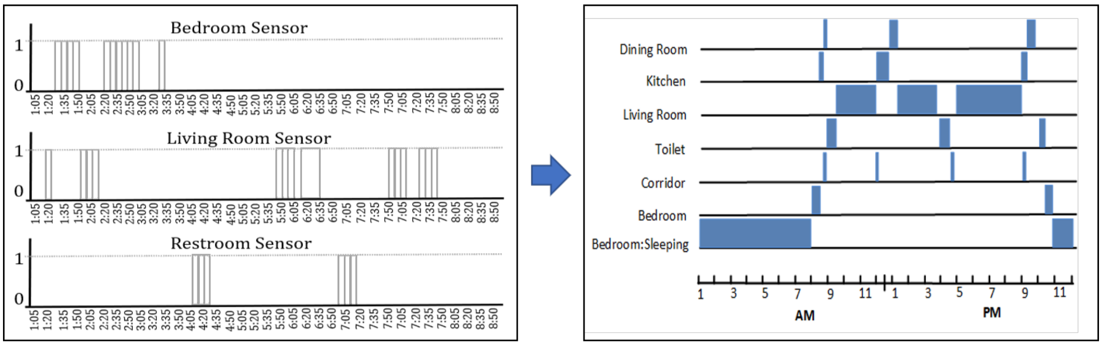
- *Visual 2D Camera*: The 2D camera is employed for the collection of image data for the hand gesture recognition task. A 2D camera is selected over depth sensors since it is the most common sensing device present on smart devices. This allows for the substitution of the device with other suitable modalities for any given use case.

- *Wearable Sensor*:The wearable device is used for gesture data collection. This is a wrist-worn tri-axial sensor equipped with an accelerometer and a gyroscope.

The collected activity data used for the validation of the computational models are grouped into distinct datasets and a detailed description is provided as follows:

3.4.1 Dataset 1: SmartNTU Data

The SmartNTU data is an ADL dataset collected from Nottingham Trent University (NTU) smart home facility as part of this research study. It contains the daily activity data of a single resident for the duration of 72 days. The ambient sensors presented earlier are utilised for the data collection. The recorded activities performed by the resident includes going to bed, using the restroom, eating, kitchen-related activities, entry and exit from the home etc. As highlighted earlier, data generated by these sensors are binary in nature with 1 and 0 signifying active and inactive states respectively. The activities performed by the residents are inferred from the binary data entries. For example, the firing of the PIR sensor in the restroom signifies that the resident is using the restroom while that of the pressure sensor on the bed is an indication that the resident is sleeping.



(a) Raw binary data (b) Interpreted data

Figure 3.3: A sample ADL data.

3.4.2 Dataset 2: CASAS HH111 Data

The CASAS HH111 data is a publicly available ADL dataset provided by the Center for Advanced Studies in Adaptive Systems (CASAS) [33]. The data is collected from a volunteer older adult living alone for the duration of 50 days. The sensing devices used for the data collected as well as the recorded activities are similar to that of the SmartNTU data which include activities such as eating, toileting, sleeping, bathing etc.

A sample plot of the binary data generated by the ambient sensors as well as the inferred data is shown in Figure 3.3. The attributes of the data include a timestamp representing the start and end time of an activity, a unique device identifier for identifying the sensor location, an event value etc. Different computational models are proposed in the literature for the data interpretation [34]. The interpreted activities have attributes such as start time, end time, activity type and the location of the performed activity. Table 3.1 shows a

Table 3.1: A sample of interpreted ADL data.

Start Time	End Time	Activity	Location
21:10:04	03:11:07	Sleeping	Bedroom
03:11:51	03:14:11	Toileting	Toilet
03:14:52	07:13:07	Sleeping	Bedroom
07:15:47	07:22:04	Preparing Meal	Kitchen
07:25:01	09:14:17	Watching TV	Living Room

sample of the interpreted activities from the collected sensor data.

3.4.3 Dataset 3: Over-sampled and Synthetic Data

To effectively evaluate the performance of the computational models and to reduce the effect of class imbalance in the evaluation metrics, the original dataset (SmartNTU and CASAS HH111) are over-sampled using the technique proposed in [30] known as Synthetic Minority Over-sampling Technique (SMOTE). Additionally, two separate synthetic datasets are generated each based on the distribution of the two datasets. Different anomalous activities relating to the sleeping pattern are artificially induced such as less sleeping, oversleeping, interrupted sleep etc. which are also based on the data distribution. The total of 100 anomalous instances are realised in the respective datasets after the oversampling and the artificial induction of outliers.

3.4.4 Dataset 4: Gesture Recognition Data

For the hand gesture recognition, data is collected from five (5) participants in a controlled lab environment. The data is captured using a 2D camera when the participants are performing the selected gestures corresponding to affirmation and denial as shown in Figure 3.4. The affirmation gestures indicate to the robotic intermediary that the model prediction is right, while the denial gestures signify wrong predictions. A 2D camera is used for the data collection since it is the most common camera found in smart device and robotic platforms, thereby giving room for the utilisation of other non-robotic intermediaries. The use of hand gestures as a communication modality is based on the assumption that the monitored individuals are in a position to certify the prediction of the anomaly detection system. While this may only apply to certain groups of individuals, this communication modality adds to the overall system flexibility in addition to other modalities such as speech and touch input. It also provides a means for the different modalities to be enabled or disabled based on the recommendation of domain experts for any given older adult. The data collection is performed in three different scenarios as follows:

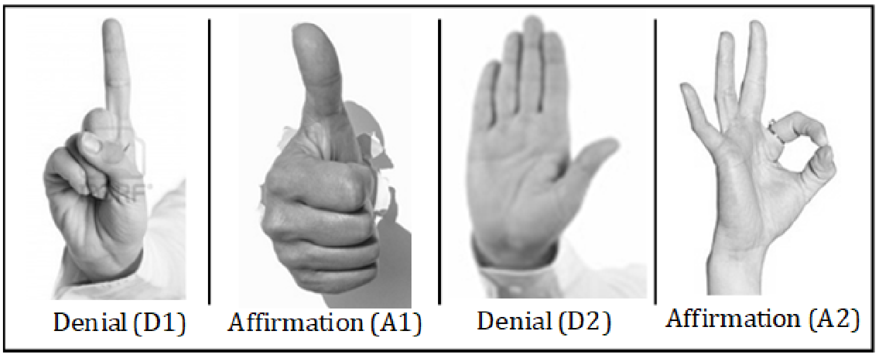


Figure 3.4: Hand gestures corresponding to affirmation and denial.

- **Scenario 1:** The gestures in Figure 3.4 are performed by each volunteer for approximately one (1) minute in a plain background setting under a good lighting condition while the videos of the volunteers performing the gestures are recorded. Frames of images are then extracted from the recorded video at the rate of 5 Frames Per Second (FPS), generating approximately 6,000 image frames.

- **Scenario 2:** In this scenario, the volunteers performed the defined gestures in a different background setting under a poor lighting condition. This scenario aims to test the robustness of the gesture recogniser since the system may be deployed in a home environment with a relatively different background setting. Each of the defined gestures is also performed for a duration of one (1) minute and the frames are extracted at the rate 5 FPS rate, resulting in over 6,000 images.

- **Scenario 3:** This scenario is meant to simulate the usability of the deployed system in a home environment. The volunteers performed the gestures only when instructed. The aim is to simulate a situation where an anomaly is detected and the intermediary administers a query to the user for confirmation. The user's response can be instantaneous or delayed. The video stream is recorded for a duration of ten (10) seconds from the time the query is issued. It is assumed that the gesture performed by the user is present in the recorded stream. The participants are instructed to perform each of the gestures ten (10) times, resulting in

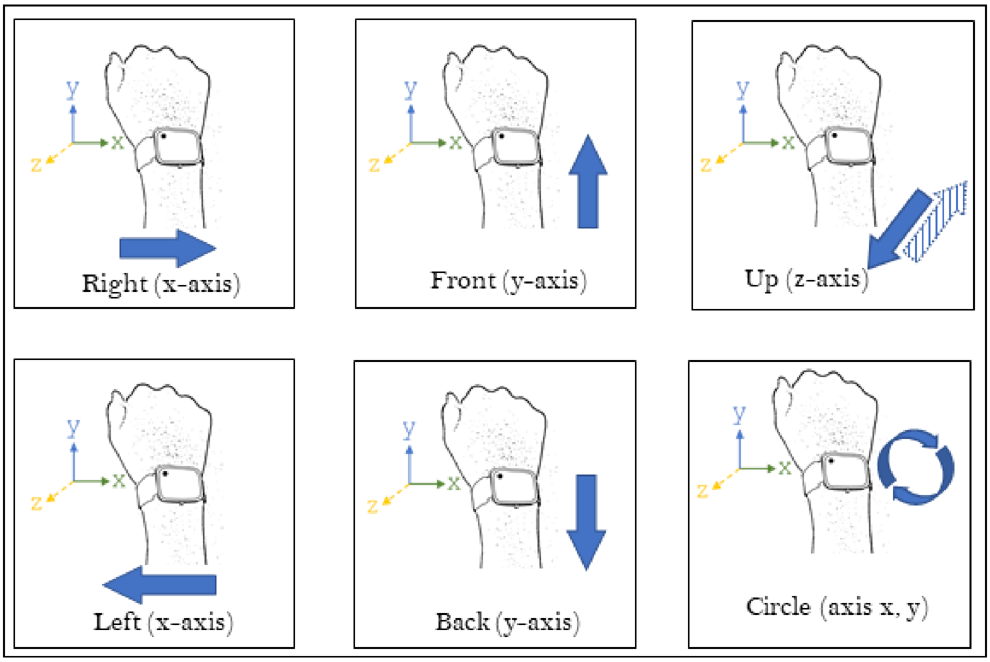


Figure 3.5: Motion related hand gestures.

a total of 200 video streams. The environmental setting for this scenario is similar to that of "Scenario 1".

An alternate approach for gesture recognition is considered using a wearable wrist-mounted sensor that is equipped with an accelerometer. This can be useful in recognising motion-based hand gestures. Six (6) motion-based hand gestures are defined to represent motions in different directions (axis), namely, Left, Right, Front, Back, Up and Circle as shown in Figure 3.5.

The data is collected from the same five participants while performing the listed gestures in Figure 3.5. The participants performed the gestures approximately twenty (20) times with a variation in the movement speed, leading to a total of 634 gestures. A wrist-mounted accelerometer sensor called "MetaMotionR (MMR)"[1] is utilised for the data collection using a set frequency sample of 50Hz.

[1] https://mbientlab.com/metamotionr/

3.5 Data Preprocessing

The collected data could contain unwanted noise, missing or invalid entries which could affect the model performance. The preprocessing step allows for the cleanup of the collected data. Conversion of the data into an acceptable format, filtering, feature extraction and normalisation is also performed.

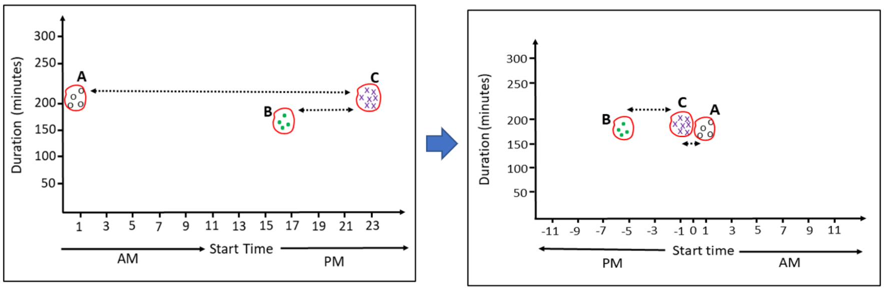
3.5.1 Activity of Daily Living Dataset

The Activity of Daily Living (ADL) datasets described above contains different daily activities of the monitored individuals. Since the anomaly detection approach is activity-dependent, sleeping activity is considered as the activity of interest. The dataset is filtered and the sleeping activity is selected for modelling and detection of abnormalities. Activities with missing entries are discarded while relevant features that can discriminate between the normal and anomalous cases are extracted as follows:

- Start time: This is the starting hour and minute of the activity. The start hour ranges from 0 to 23. It is then converted to a scale of -11 to +11 with 0 representing 12 midnight. This is because generally, people do go to bed at night time. An activity that starts at 11:50 pm is closer to that of 1:00 am with respect to the start time than an activity performed at 9:00 pm. However, without converting the start time to a scale of -11 to $+11$, the margin between 11:50 pm to 1:00 am will be larger as shown in Figure 3.6.

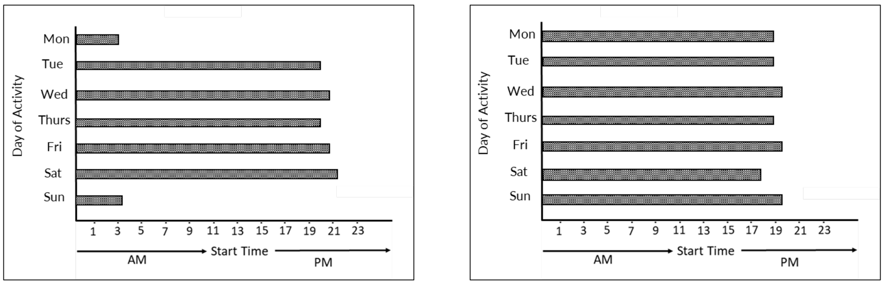
- Duration: This is the duration in minutes of the activity obtained by subtracting the start time from the end time.

- Number of interruptions: This is the number of times an individual leaves the bed and returns to it. For example, an individual may leave the bed in the middle of the night to use the restroom. If the interval (in minutes) between the time the individual leaves the bed and returns to it is less than an hour (60 minutes), it is considered an interruption, else it is assumed that the activity has ended.



(a) Unscaled start time (a) Scaled start time

Figure 3.6: A plot of activity start time scaling.



(a) User 1 bedtime pattern (b) User 2 bedtime pattern

Figure 3.7: A plot of activity start time patterns by day for two users.

- Duration of interruption: This is the total duration (in minutes) of all the interruptions within an activity.

- Day of activity: This represents the day in which the activity is performed, ranging from 0 to 6 representing the 7 days of a week. This is important because the individual may go to bed late on some days and early some other days due to his/her routine e.g. watching a specific late-night TV show every Monday and Sunday. Figure 3.7 depicts a simulated scenario of the bedtime pattern for two users in which "User 2" has a uniform start time pattern every day while the pattern for "User 1" differs across the days.

- Weekend or Weekday: This is to determine if the activity is performed on weekdays or weekends. Some individuals might have a different routine for weekdays and weekends while some might not. 0 and 1 represents weekdays and weekends respectively.

The extracted features are in different scales. Computational models sensitive to scaling may perform poorly on the dataset, therefore, the selected features are normalised.

3.5.2 Gesture Recognition Dataset

The datasets for the gesture recognition collected using a 2D camera and a wearable device is preprocessed prior to the application of the computational models. The recorded video streams described in the scenarios above are processed and images are extracted at the rate of 5 FPS. Segments of the image frames that do not contain the gestures of interest are removed. Regions of the images containing the gestures are marked with a bounding box, and the labelling of the dataset is performed thereafter. Table 3.2 shows a summary of the collected data highlighting the number of participants, gestures, image frames, and the size of the training and validation sets, respectively.

As for the data collected using the wearable sensor, the acceleration data is filtered using a digital high-pass filter with 0.5Hz cut-off frequency to filter out the gravity component of the data while a low-pass filter with a cut-off frequency of 20Hz is employed for noise reduction. The filtering process is performed along the three different axes (i.e. x, y and z) separately. A signal segmentation technique proposed in [12, 13] termed Crossings-based Adaptive Segmentation Technique (CAST) is utilised. According to the authors in [12], the CAST can detect segments of interest from the signal using the crossing of two moving

Table 3.2: A summary of the collected gesture data using 2D camera.

Feature Description	Scenario 1	Scenario 2	Scenario 3
Number of participants	5	5	5
Number of Gestures	4	4	4
Length of each gesture	60 sec	60 sec	100 sec
Length of all gestures per individual	240 sec	240 sec	400 sec
Length of all gestures for all individuals	1200 sec	1200 sec	2000 sec
Extracted images per gesture	300 images	300 images	0
Extracted images for all gestures (per individual)	1200 images	1200 images	0
Extracted images for all gestures (for all individuals)	6000 images	6000 images	0
Training data size	2000 images	0	0
Testing data size	4000 images	6000 images	200 vid

averages (i.e. a fast moving average and a slow-moving average) of different magnitude. A moving average is a signal smoothing technique for time series data that generates a new series based on the computed average of the old series using a specified sliding window length. The fast-moving average is a moving average obtained when the window length is short, leading to a rapid motion in the new time series data while a slow-moving average generated a series with a slower motion due to its large window length. The moving averages are calculated using Equation 3.18. After evaluating different values of the moving averages, an optimal value for the fast-moving average is obtained as 40 samples (i.e. 0.8 seconds), and 300 samples (i.e. 6 seconds) for the slow-moving average. Data segments with a duration of less than 0.7 seconds are ignored since this is too short for any of the specified gestures of interest [11].

$$|MA|_t = \frac{1}{n}(x_{t-\frac{n}{2}} + \ldots + x_t + \ldots + x_{t+\frac{n}{2}}) \tag{3.18}$$

where $|MA|_t$ is the moving average, x is the input data, t is the current point of the signal and n is the number of points over which the moving average is calculated.

The features set extracted from the collected time series data segments include Signal Magnitude Area (SMA), Tilt Angles, Standard Deviation, Mean, Root Mean Square (RMS), Correlation, Skewness, Kurtosis, Min, Max, MinMax, and Positive and Negative crossings of the zero-axis. All these outlined features are calculated along the three axes except for the SMA. Overall, a total of 37 features are realised after the preprocessing.

3.6 Evaluation Metrics

To evaluate the performance of the computational models, different metrics are measured in comparison to the ground truth. The ground truth for the anomaly detection is identified through the manual observation and annotation of the outliers in the dataset. For the artificially induced anomalies, this is easily obtained since the altered features of the data entries are known. The data entries with features that are altered to simulate an abnormally large variance are labelled as outliers. The four possible prediction outcomes for any given entry relative to the anomaly detection are given as follows:

- True Positive (TP): This indicates that the data annotated to be outliers are correctly predicted as outliers.
- False Positive (FP): This indicates that the data entries labelled as non-outlier are predicted as outliers.
- True Negative (TN): This indicates that the data entries annotated as non-outliers are correctly predicted as non-outliers.
- False Negative (FN): This indicates that the data entries labelled as outlier are predicted as non-outliers by the model.

The evaluation metrics measured are the prediction accuracy, precision and recall using the expression in Equation 3.19 - 3.21. The accuracy calculates the model ratio of the number of correctly predicted samples against the total number of samples in the evaluation dataset. The recall and precision measurement give a more accurate representation of the true positive rate of predicted data.

$$Precision = \frac{TP}{TP + FP} \tag{3.19}$$

$$Recall = \frac{TP}{TP + FN} \tag{3.20}$$

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{3.21}$$

It should be noted that these metrics are utilised throughout this thesis for the evaluation of the proposed approaches for anomaly detection as well as for the gesture recognition approaches. Given that the gesture recognition approaches are classification problems by default, a Cross-Validation (CV) strategy is adopted to evaluate the robustness of the approaches. A K-Fold CV strategy is utilised when required for the model evaluation. In a K-Fold CV, the dataset is shuffled and divided into 'K' number of folds or units. The model is then evaluated K times, but for each iteration, $K - 1$ folds are used for model training and the remaining '1' fold is used for the validation.

3.7 Conclusion

This chapter presents a description of the computational models employed for the detection of anomalies such as OC-SVM, iForest, LOF and approach based on RCE, as well as the supervised models for gesture recognition used for the communication intermediary. A detailed description of the datasets used for the validation of the approaches proposed in this thesis, including the choices of the sensing devices, data collection procedure and the preprocessing techniques are provided. The datasets include daily activity data collected using ambient sensors as part of this research study, a publicly available dataset, as well as the gesture recognition data gathered using a wearable device and a 2D camera. Lastly, the evaluation metrics utilised for measuring the performance of the models proposed in this thesis are described. The next chapter presents an ensemble model for aggregating the anomaly detection models presented in this chapter.

Chapter 4

Consensus Novelty Detection Ensemble

4.1 Introduction

In this chapter, an approach for creating an ensemble of novelty detection models is proposed. Anomaly detection approaches in ADL are too simplistic and therefore generate a high rate of false alarm [73], making it unsuitable for monitoring the wellbeing of older adults. To reduce the false alarm rate, the behaviour of the monitored individuals needs to be modelled accurately. This can be achieved by using an ensemble of novelty detection models since each model is good on certain characteristic features. For example, OC-SVM is sensitive to the presence of outliers in the training data, thereby resulting in poor performance, while iForest can perform well even when the training data is contaminated with outliers since it isolates the anomalies instead of profiling the normal data [91, 100]. An ensemble of machine learning models combines multiple models' predictions to achieve better accuracy. For anomaly detection, an ensemble of homogeneous models that produces the same output is not good enough. However, an ensemble of heterogeneous models is better since it will provide the much needed diversity and accuracy [131].

The proposed Consensus Novelty Detection Ensemble (CNDE) approach generates a score for an activity termed as "Normality Score" qualifying the activity as either normal (inlier) or abnormal (outlier). Ensembles of machine learning models are usually based on a voting approach and the resulting output is a label representing the class of the data point. In the context of this research, the output can be either “normal” or “abnormal”. This is not flexible for ADL anomaly detection since the threshold value for normal and abnormal activities cannot be explicitly adjusted. Human behavioural routine is complex and subject to changes due to seasonal or other factors. The normality score generated by this proposed approach allows the threshold to be dynamically adjusted to incorporate changes in the individual’s routine. Additionally, ensemble approaches for novelty detection models have not been given much attention, and according to our knowledge, none of the few proposed approaches takes the concept of normality score into consideration. The approach gives more flexibility since the threshold of the score signifying inliers and outliers can be altered dynamically.

The rest of this chapter is organised as follows: Section 4.2 presents the methodology of the ensemble approach while Section 4.3 presents the experimental results and discussion. Pertinent conclusions are drawn in Section 4.4.

4.2 Methodology

The proposed CNDE approach is based on the concept of internal and external consensus. This is inspired by the concept proposed by Mahmud et al. [110] in which a sensor node is certified based on its data and behavioural trust among the other nodes in the infrastructure. The internal consensus is an internal voting scheme within each model in the ensemble such that a number of child models are created for each model and their votes are aggregated, and a score is computed for the data points. The external consensus is a voting scheme among the parent models in the ensemble similar to the majority vote approach. Appropriate weights are estimated and assigned to the respective models in the ensemble based on their performance. The normality score generated by the CNDE enables the data to be classified as either normal or abnormal. A higher normality score indicates that the data point is an inlier (normal) while a lower

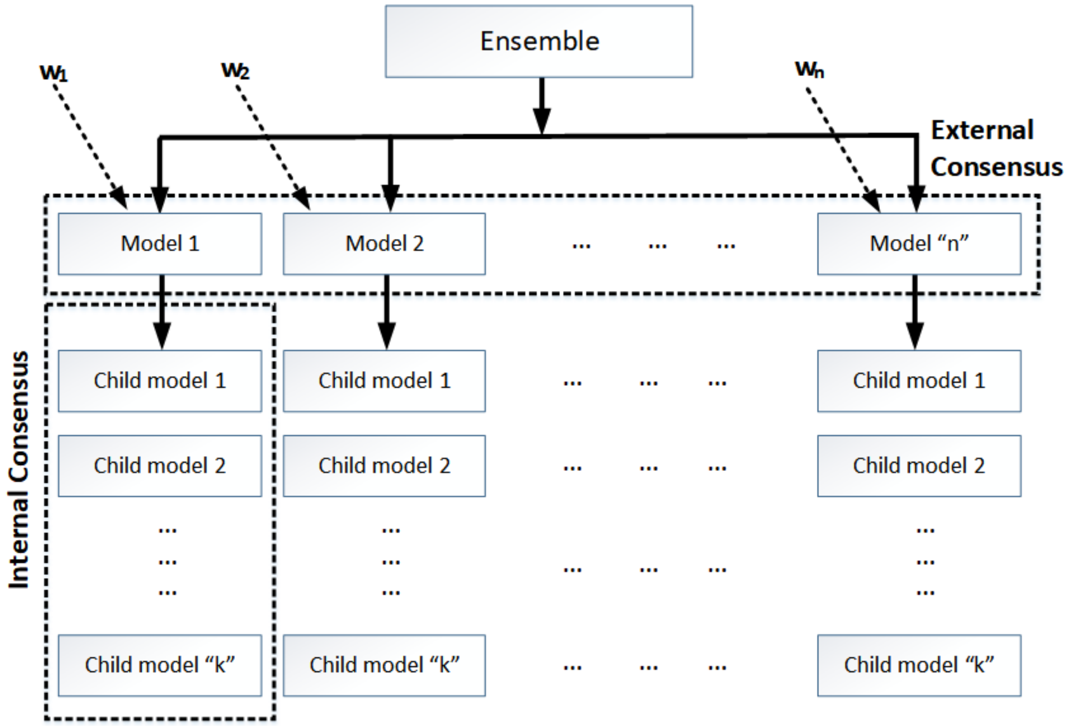


Figure 4.1: An schematic diagram of consensus voting scheme.

score signifies an outlier (anomaly). Figure 4.1 illustrates a schematic diagram of the consensus voting scheme of the CNDE approach.

Given a set of n training samples $X = \{x_1, x_2, ..., x_n\}$ of $d-$dimensional data (i.e., $X \in \mathbb{R}^d$), let $A = \{a_1, a_2, ..., a_m\}$ be a set of m models. A data point x can be classified as either an outlier or inlier by computing its Normality Score $\mathcal{N}_x$ using an ensemble of A as expressed in Equation 4.1 below.

$$\mathcal{N}_x = \frac{F_x^c + E_x^c}{2} \tag{4.1}$$

F_x^c is the Combined Internal Consensus Score (CICS) and E_x^c is the Combined External Consensus Score (CECS) of the data point x respectively.

$$E_x^c = \frac{1}{m} \sum v_x \tag{4.2}$$

v_x is the number of votes x received as an inlier from the models and m is the number of models in the ensemble.

The aggregate of the votes v_x is termed as the Combined External Consensus Vote (CECV). The class of a data point based on external consensus is determined by the majority vote of the CECV.

The CICS is the weighted average of the Internal Consensus Score (ICS) as expressed in Equation 4.3.

$$F_x^c = \frac{1}{m} \sum_{i=1}^{m} I_x^i * w_i \tag{4.3}$$

where m is the number of models in the ensemble, I_x^i is ICS of the i^{th} model for the data point x, and w_i is the weight of the i^{th} model. The w_i is estimated through model weight initialisation and penalisation using the expression in Equation 4.5.

The ICS expressed in Equation 4.4 is inspired by the Bagging approach in machine learning [86]. The training data is split randomly into k-folds. A k-child models are created for each model in the ensemble. The k-child models are trained each with one separate fold out of the k-fold training data as illustrated in Figure 4.1. The votes a data point x receives from the k-child models are termed as the Internal Consensus Vote (ICV). A data point is considered an inlier by a model if it has 1 or more ICV.

$$I_x = \frac{1}{k} \sum_{i=1}^{k} v \tag{4.4}$$

v is the number of votes x received as an inlier from the child models, and k is the number of child models (i.e. the number of folds).

The difference between the ICS and the CICS is that the ICS determines the score of a data point for an individual model in the ensemble, while CICS computes the score of the data point across all the models in the ensemble using the respective models' weights.

The weight of each model is a value ranging from 0 to 1. The models performing better receives larger weight and vice versa. This is estimated during training since it is impossible to manually assign appropriate weight values. The weight of each model is initialised to 1 and penalised by the percentage of wrong predictions made by the model. Wrong predictions are determined by comparing the CECV and the ICV. Variability of prediction

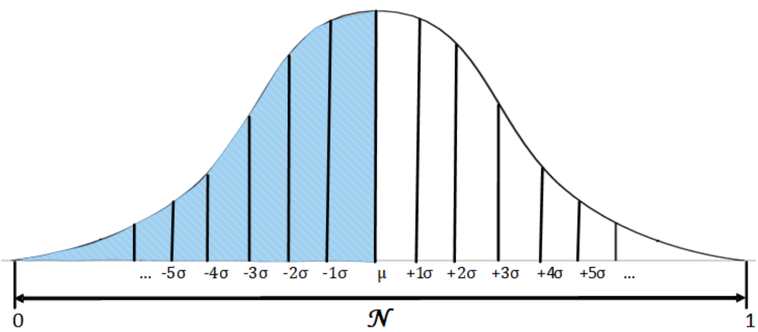


Figure 4.2: Normality threshold estimation.

between the CECV and ICV is considered a wrong prediction, therefore, warrants the penalisation of that model.

Let $W = \{w_1, w_2, ..., w_m\}$ be the weights of the m models in the ensemble, the final weight of each individual model after penalisation can be expressed as:

$$w_f = w_i - \frac{e}{n} w_i \tag{4.5}$$

where w_f is the final weight after penalisation, w_i is the initial weight before penalisation (i.e. 1), e is the number of wrong predictions made by the i^{th} model and n is the size of the training samples.

Introducing a threshold value ε termed as "Normality Threshold" to serve as cut-off point for the normality score, the function $f(x)$ that determines the class of x is expressed as:

$$f(x) = \begin{cases} \mathcal{N}_x \geq \varepsilon & \text{then x is an Inlier} \\ \mathcal{N}_x < \varepsilon & \text{then x is an Outlier} \end{cases} \tag{4.6}$$

The normality score is a value ranging from 0 to 1 (i.e. $0 \leq \mathcal{N}_x \leq 1$) with a higher score signifying inlier and a lower score signifying outlier. Certain standard deviation to the left of the normality score is considered as the threshold as shown in Figure 4.2. An ideal threshold is -3σ (i.e. any score below -3σ is considered as an outlier). This is determined using the standard statistical approach for outlier detection on a normal data distribution which considers entries beyond a certain standard deviation (usually -3σ or 3-sigma rule) as outliers [86, 96].

Algorithm 1, Algorithm 2, Algorithm 3 and Algorithm 4 show the procedures for computing the CECS, CICS, ICS and the weights of the models respectively as described above. The time complexities of the respective algorithms are tabulated in Table 4.1. The worst-case complexities (O) of the algorithms for the Weight Estimation and Combined Internal Consensus are dependent not just on the size of the input data (n) but also on the number of models (m) in the ensemble. It is highly unlikely that the number of models employed will exceed the size of the training data, and therefore, the complexity will most likely be of the form $O(n)$. However, if in any case, the number of models in the ensemble is equal to the size of the training data, then the complexity will be $O(n^2)$. Improved time complexity can be achieved by maintaining a fixed number of models in the ensemble and having the size of the training data larger than the number of models.

Algorithm 1 Combined External Consensus Algorithm (CECA)

Input: Dataset $X = \{x_1, x_2, ..., x_n\}$,
Models List $A = \{a_1, a_2, ..., a_m\}$
Output: Vote and Score $(V^e, E^c) = \{(v_{x_1}, e_{x_1}), (v_{x_2}, e_{x_2}), ..., (v_{x_n}, e_{x_n})\}$

procedure CECA
  for each $x \in X$ **do**
    $t = \sum v$ ▷ Aggregate votes of x as inlier from A
    $e_x = \frac{1}{m}t$ ▷ Computing the CECS of x
    if t is the majority **then**
      $v_x = 1$ ▷ x is an inlier by CECV
    else
      $v_x = 0$ ▷ x is an outlier by CECV
    end if
    end if
    $(V^e, E^c) \leftarrow (v_x, e_x)$ ▷ Append result to (V^e, E^c)
  end for
  end for
  return (V^e, E^c) ▷ Return the CECV and CECS of x
end procedure

Table 4.1: Time complexities of proposed algorithms.

Algorithm	Complexity
Combined External Consensus Algorithm (CECA)	$O(n)$
Combined Internal Consensus Algorithm (CICA)	$O(n * m)$
Internal Consensus Algorithm (ICA)	$O(n)$
Weight Estimation Algorithm(WEA)	$O(n * m)$

Algorithm 2 Combined Internal Consensus Algorithm (CICA)

Input: Dataset $X = \{x_1, x_2, ..., x_n\}$,
Models List $A = \{a_1, a_2, ..., a_m\}$,
Models Weights $W = \{w_1, w_2, ..., w_m\}$

Output: CICS $F^c = \{f^c_{x_1}, f^c_{x_2}, ..., f^c_{x_n}\}$

procedure CICA
  for each $x \in X$ **do**
    for each $a \in A$ **do**
      I_x = Compute the ICS of x
      $\rho \leftarrow I_x * w_a$ ▷ ICS and model's weight
    end for
    end for
    $f^c_x = \frac{1}{m}\sum_{i=1}^{m} \rho$ ▷ Compute CICS for x
    $F^c \leftarrow f^c_x$ ▷ Append result to F^c
  end for
  end for
  return F^c ▷ Return the CICS of x
end procedure

Algorithm 3 Internal Consensus Algorithm (ICA)

Input: Dataset $X = \{x_1, x_2, ..., x_n\}$,
Model M with k-child models $M = \{m_1, m_2, ..., m_k\}$

Output: Vote and Score $(V^i, I) = \{(v_{x_1}, i_{x_1}), (v_{x_2}, i_{x_2}), ..., (v_{x_n}, i_{x_n})\}$

procedure ICA
  for each $x \in X$ **do**
    $t = \sum v$ ▷ Aggregate votes of x as inlier from childrens of M
    $i_x = \frac{1}{k}t$ ▷ Computing the ICS of x
    if $t \geq 1$ **then**
      $v_x = 1$ ▷ x is an inlier
    else
      $v_x = 0$ ▷ x is an outlier
    end if
    end if
    $(V^i, I) \leftarrow (v_x, i_x)$ ▷ Append result to (V^i, I)
  end for
  end for
  return (V^i, I) ▷ Return the ICV and ICS of x
end procedure

4.2.1 Combined Concepts

Combining the concepts described, the diagram in Figure 4.3 shows the training and testing phases of the CNDE approach. The training phase involves the random splitting of the training data into k-folds, creation of k-child models for

Algorithm 4 Weight Estimation Algorithm (WEA)

Input: Dataset $X = \{x_1, x_2, ..., x_n\}$,
Models List $A = \{a_1, a_2, ..., a_m\}$
Output: Weights $W = \{w_1, w_2, ..., w_m\}$

procedure WEA
  Initialise weights $W = \{w_1, w_2, ..., w_m\}$ to 1
  Initialise errors $E = \{e_1, e_2, e_m\}$ to 0
  for each $x \in X$ **do**
    v_e = Get the CECV of x
    for each $a \in A$ **do**
      v_i = Get the ICV of x by model a
      if $v_e \neq v_i$ **then**
        $e_a = e_a + 1$ ▷ Increment error count of model a
      end if
      end if
    end for
    end for
  end for
  end for
  for each $a \in A$ **do**
    w_a = Get the weight of model a from W
    e_a = Get the errors of model a from E
    $w_a^* = w_a - \frac{e_a}{n} w_a$ ▷ Compute final weight by penalisation
    $W \leftarrow w_a^*$ ▷ Update weight of model a
  end for
  end for
  return W ▷ Return estimated weights of A
end procedure

each model in the ensemble, training and estimating an appropriate weight for each model, computing the normality score for the training data and estimating an optimal normality threshold. During the testing phase, the trained models along with their estimated weights and the normality threshold are applied to predict the class of the test data.

4.3 Experimentation

In this section, the proposed CNDE model is evaluated using the ADL datasets described in Section 3.4. This includes the two real datasets (SmartNTU and CASAS HH111 data) and the synthetic dataset for evaluation of different performance metrics. The extracted ADL features in Sections 3.5 are normalised prior to the model training.

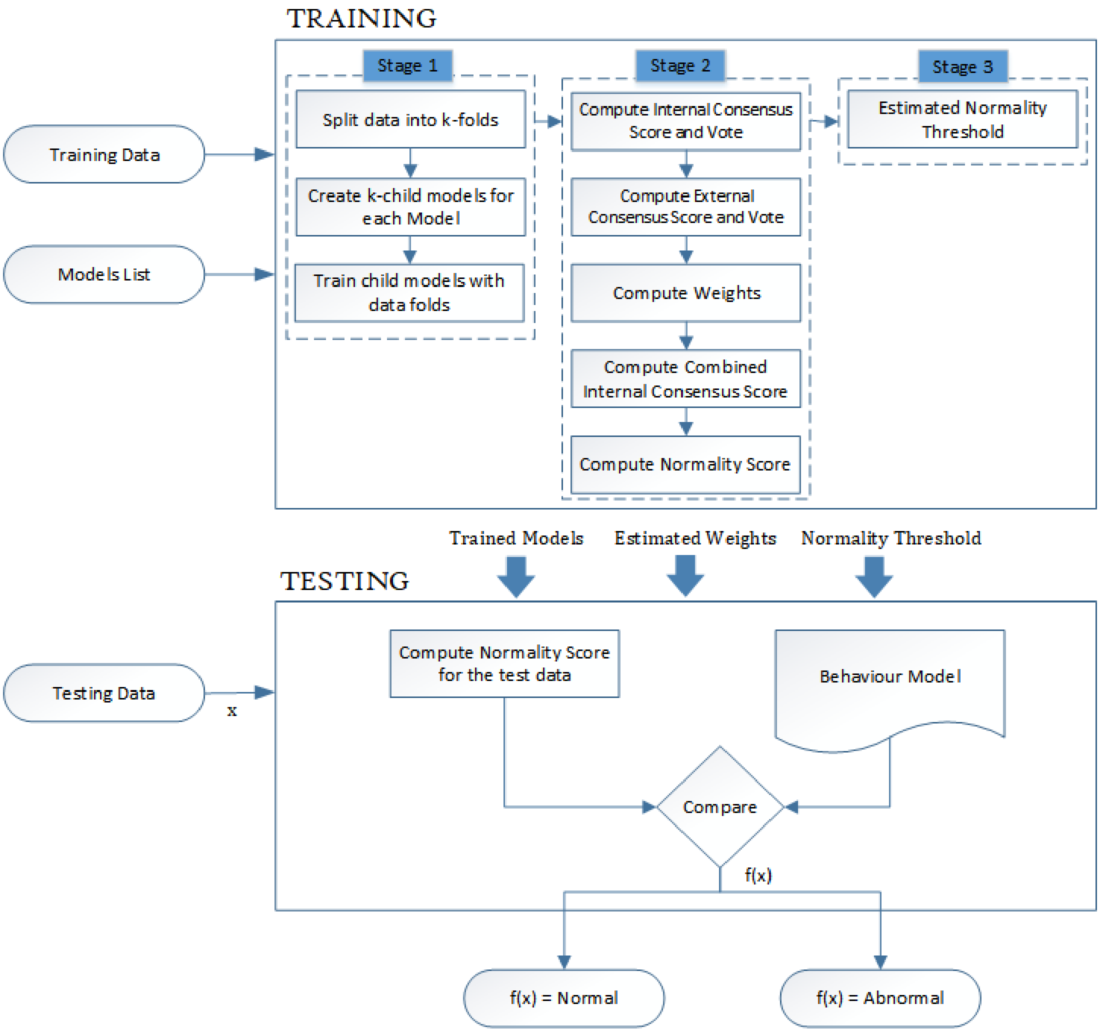


Figure 4.3: The training and testing phases of the proposed CNDE approach.

4.3.1 Model and Parameter Selection

The CNDE approach is generic and can be used with any number of novelty detection models. However, for this empirical evaluation, models employed are iForest, LOF, RCE, OC-SVM with a Radial Basis Function (RBF) kernel. The contamination rate (i.e. the rate of outliers in the training data) is set to 0.1 across all the models since the aim is to model the training data with minimal error. The number of folds for the internal consensus can be varied depending on the size of the training set. Three (3) folds are used for the utilised datasets and the size of the training data is set to 31 days. The weights of the respective models in the ensemble are initialised to 1 and the normality threshold is taken as -3σ as described in the methodology section.

To determine the minimum amount of data required for training the model and to verify if the size of the training data has any significant effect on the model's performance, a heuristic approach is used for the model training using different

Table 4.2: Accuracy of the ensemble model based on training size.

Data Size (days)	**SmartNTU Data**			**CASAS HH111 Data**		
	Iter. 1	Iter. 2	Iter. 3	Iter. 1	Iter. 2	Iter. 3
10	0.450	0.550	0.480	0.430	0.620	0.560
11	0.480	0.650	0.446	0.552	0.346	0.560
12	0.300	0.465	0.360	0.386	0.620	0.620
13	0.682	0.450	0.565	0.560	0.560	0.386
14	0.368	0.480	0.480	0.882	0.800	0.790
15	0.560	0.682	0.480	0.860	0.852	0.902
16	0.560	0.560	0.446	0.852	0.800	0.922
17	0.682	0.682	0.480	0.944	0.960	0.852
18	**0.775**	**0.776**	**0.775**	**0.972**	**0.972**	**0.972**
19	0.785	0.786	0.785	0.970	0.972	0.970
20	0.786	0.786	0.786	0.968	0.968	0.970
21	0.810	0.810	0.800	0.972	0.970	0.972
22	0.854	0.854	0.856	0.972	0.972	0.972
23	0.854	0.855	0.854	0.962	0.962	0.962
24	0.890	0.892	0.892	0.960	0.960	0.962
25	0.890	0.890	0.892	0.966	0.960	0.960
26	0.892	0.892	0.892	0.962	0.962	0.962
27	0.930	0.931	0.930	0.960	0.960	0.966
28	0.956	0.956	0.955	0.957	0.957	0.958
29	0.955	0.956	0.956	0.966	0.960	0.960
30	0.960	0.960	0.960	0.956	0.956	0.957
31	**0.986**	**0.986**	**0.987**	**0.958**	**0.958**	**0.958**
32	0.986	0.987	0.986	0.957	0.957	0.957
33	0.985	0.985	0.986	0.958	0.958	0.958
34	0.985	0.984	0.984	0.958	0.958	0.958
35	0.986	0.986	0.986	0.956	0.956	0.956

training sample sizes. The data used for this verification is the oversampled data described in Section 3.4. Table 4.2 shows the accuracy of the proposed ensemble approach for the two datasets tabulated across 3 iterations based on the size of the training data. It can be seen from the result that the accuracy across the 3 iterations remains consistent from day 18 (i.e. when 18 days training data is used). This shows that a minimum of 18 days of training data is required to achieve consistency since the accuracy across the 3 iterations normalises from day 18. It can also be seen that the accuracy of the ensemble model remains static from day 31 across the datasets. This is an indication that the performance peak of the model is reached using 31 days of training data. The two reference points (i.e. 18 and 31 days) are used for further evaluation of the proposed CNDE approach.

4.3.2 SmartNTU Dataset

The experimental evaluation carried out on the SmartNTU dataset utilises the first 31 days data for model training while the remaining 41 days data is used for testing. To ensure that the proposed ensemble approach generalises, multiple iterations of the experiment is executed since for each iteration, the data is split and shuffled randomly.

The result in Figure 4.4 shows a plot of the normality score for the SmartNTU dataset using 31 days training data across 3 iterations. Even though the normality score varies across the iterations for the different days, the days identified as anomalous are the same across all the iterations, thereby making the score difference insignificant. The dataset is examined for variations between the days identified as anomalous and those identified as normal and the findings shows that the entries are correctly predicted as outliers.

Additionally, a test is conducted with 18 days of training data to verify if the size of the training data has any significant effect on the model's performance. The obtained result is shown in Figure 4.5. It can be seen that the model performed poorly when trained with data for 18 days as compared to 31 days of training data.

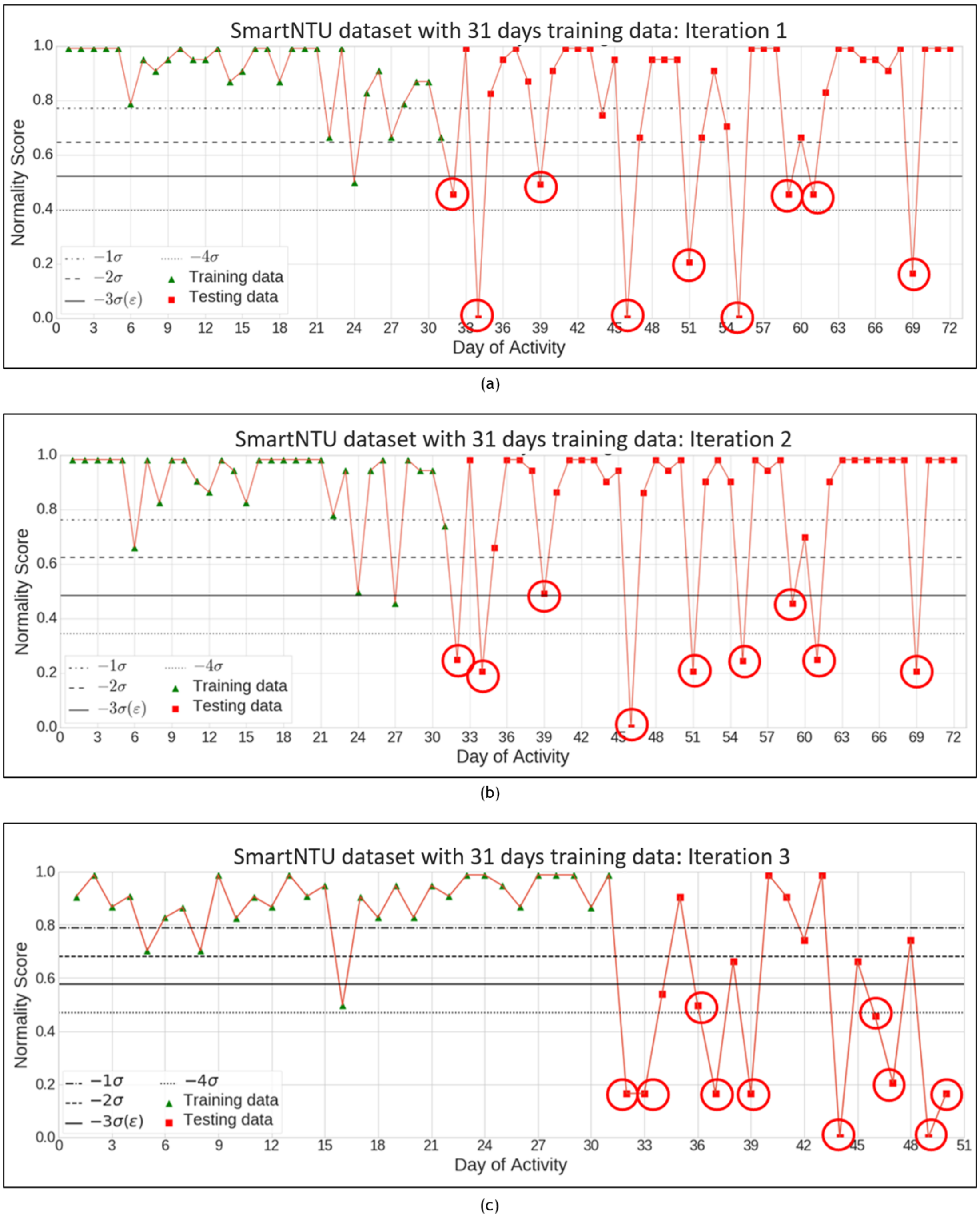


Figure 4.4: Normality score plot for SmartNTU dataset; a) Iteration 1, b) Iteration 2, c) Iteration 3

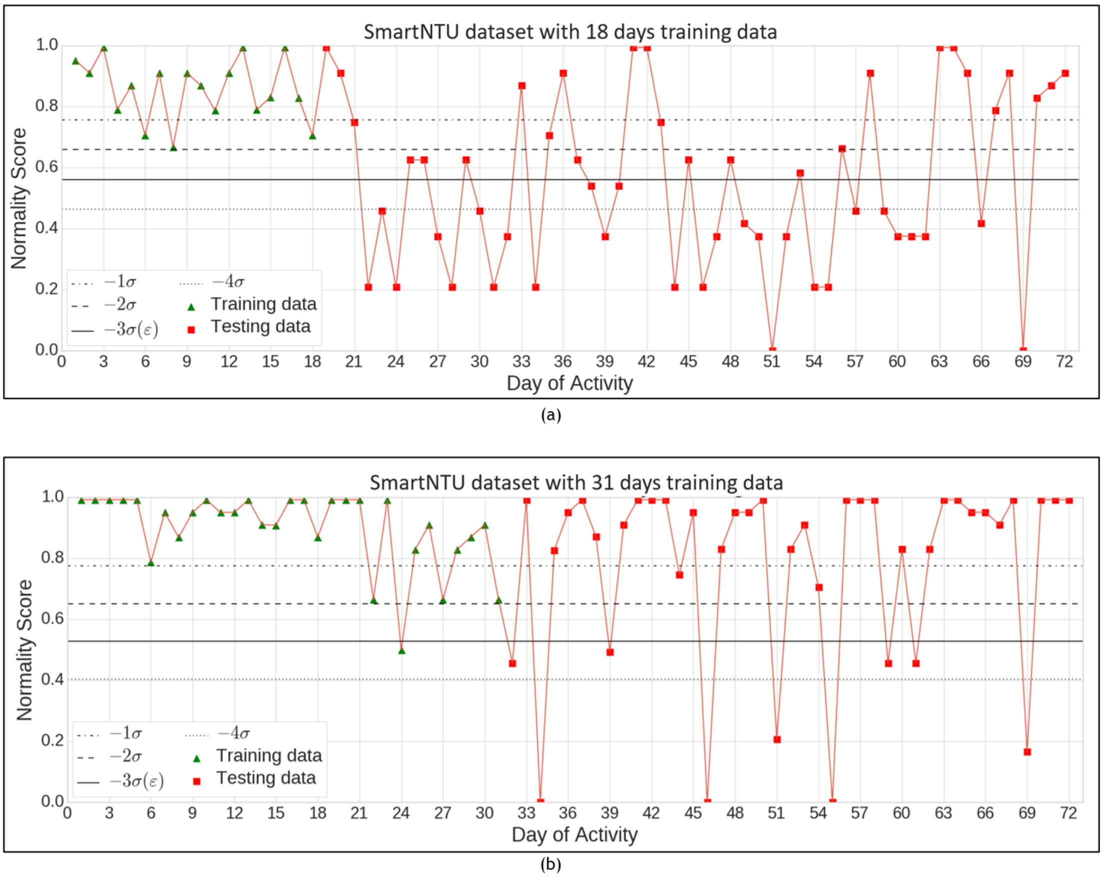
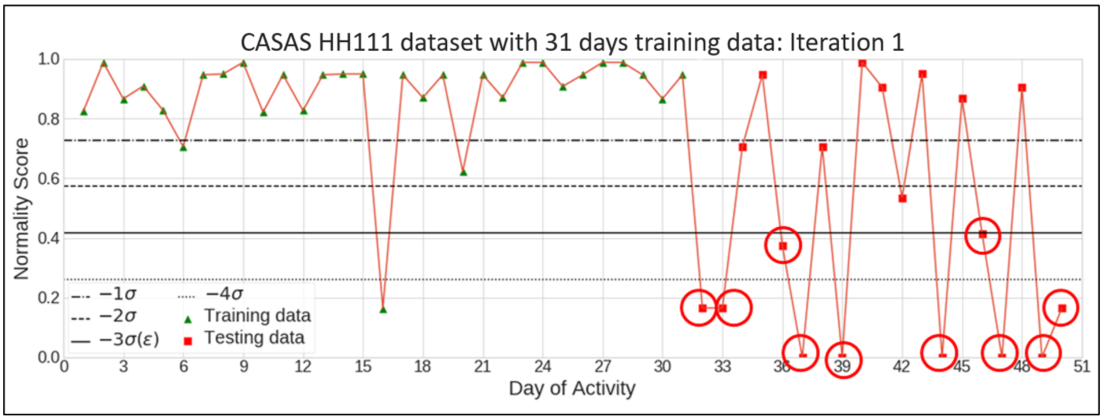


Figure 4.5: Normality score plot for SmartNTU dataset with different training data sizes; a) 18 days training data, b) 31 days training data.

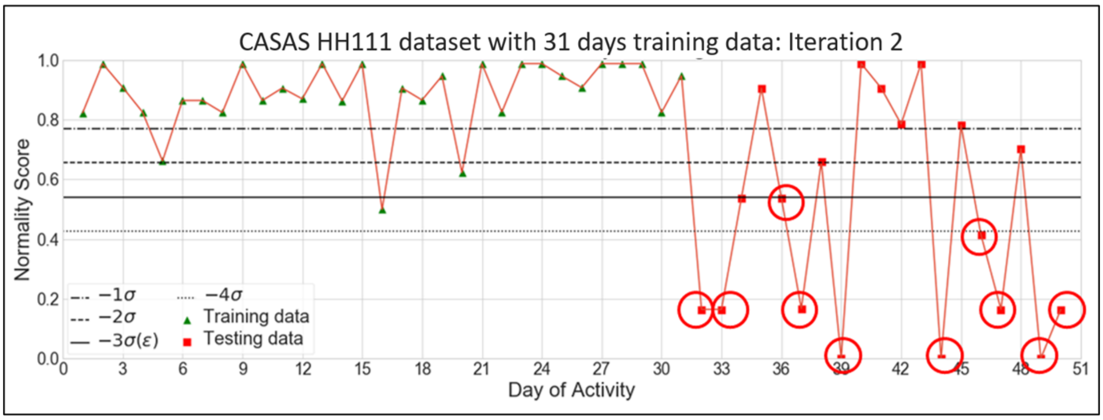
4.3.3 CASAS HH111 Dataset

Similar to the previous experiment, an evaluation of the ensemble approach is carried out using the CASAS HH111 data with 31 days data used for training and the remaining data for testing. Multiple iterations are also performed to verify the model generalisation. The normality score plot for the CASAS HH111 dataset for the 3 iterations is shown in Figure 4.6. The identified anomalous days are examined and verified from the dataset.

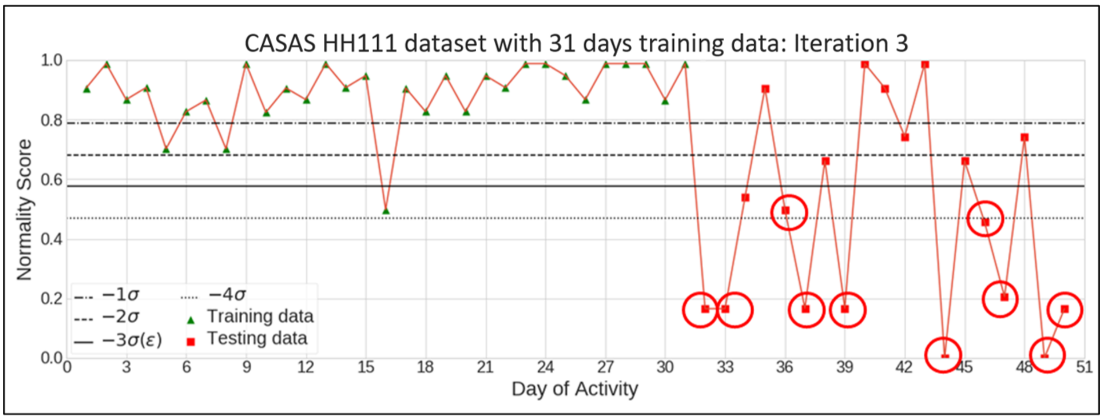
The model is able to identify data points that do not conform to the known resident's behavioural routine. Additional experiment to ascertain the model's robustness is conducted with 18 days training data and the result is shown in Figure 4.7. It is evident from the result that both the 18 and 31 days training data



(a)

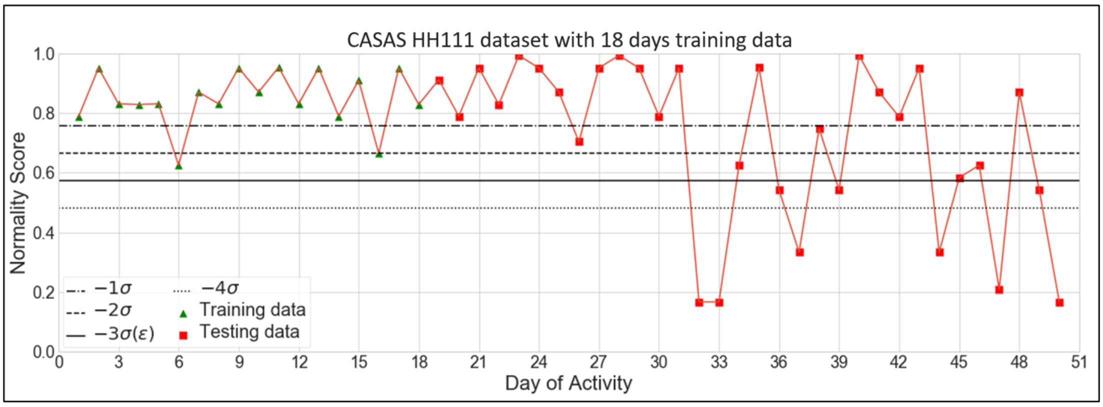


(b)

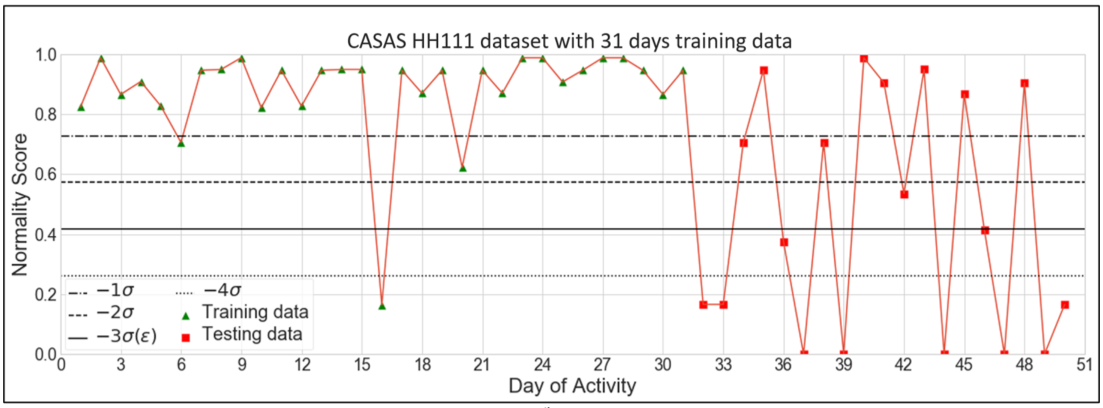


(c)

Figure 4.6: Normality score plot for CASAS HH111 dataset; a) Iteration 1, b) Iteration 2, c) Iteration 3



(a)



(b)

Figure 4.7: Normality score plot for CASAS HH111 dataset with different training data size; a) 18 days training data, b) 31 days training data.

produces a comparably similar result, unlike in the case of the SmartNTU data in Figure 4.5. This further proves the initial assertion that human behavioural routine varies among individuals and therefore, the model training should be done on an individual basis. Nonetheless, it can be established that the minimum number of days required for the modelling is 31 days.

4.3.4 Comparison with Other Models

To evaluate the proposed ensemble approach, a comparison is made with singular models, ensemble approach based on majority votes, as well as approaches proposed in [131], namely; Ensemble of Detectors with Correlated Votes (EDCV) and Ensemble of Detectors with Variability Votes (EDVV). In the ensemble approach based on a majority vote, the respective models in the ensemble vote a data entry as either inlier or outlier with the class having the majority of votes taken as the final prediction. Since the ensemble approach consists of 4 computational models, the majority vote approach is evaluated in two scenarios using different values of the majority vote threshold. For the first scenario, a threshold value of 2 is selected while for the second scenario, a threshold value of 3 is selected. The EDCV and EDVV are similar to the majority vote approach except that for the EDCV, weights of the ensemble model is estimated from the correlation coefficient of the models' predicted score, while for the EDVV, the weights are estimated from the Mean Absolute Deviation (MAD) of the prediction score as described in Section 3.2.

To measure the performance metrics, the oversampled data as well as synthetic anomalous data described in Section 3.4 are utilised. The synthetic data is generated to simulate different anomalous instances such as going to bed late, insufficient sleep, interrupted sleep etc. A total of 100 days anomalous activity data is realised. The four computational models (i.e. OC-SVM, iForest, LOF and RCE) are also used for the various ensemble models.

The results obtained from the comparison of the ensemble models is presented in Table 4.3. From the presented results, the proposed CNDE approach achieved

Table 4.3: Result of comparison with other ensemble methods based on accuracy.

	Training Data (Days = 31)		**Training Data (Days = 18)**	
Ensemble Approach	**Collected Data**	**CASAS HH111**	**Collected Data**	**CASAS HH111**
EDCV	0.92958	0.83099	0.88732	0.85915
EDVV	0.90141	0.81690	**0.91549**	0.83099
Majority Vote (v=2)	0.54930	0.77465	0.54930	0.78873
Majority Vote (v=3)	0.76056	0.92958	0.60563	0.94366
Our Approach (CNDE)	**0.98592**	**0.95775**	0.77465	**0.97183**

a better performance than the other ensemble approaches. The only exception is in the case of the SmartNTU dataset when 18 days data is used to train the model. This is not surprising since it is established earlier that 31 days of training data is the minimum requirement for the behaviour modelling. Similar to what is obtained in the Normality score plot in Figure 4.7, the results for the CASAS HH111 dataset when trained with data for 18 days and 31 days are comparably similar, while for the SmartNTU data, the results are significantly different. This is another confirmation of the variability in the behavioural routine of one individual to another. Overall, the ensemble approach based on majority votes with a 3 votes threshold outperformed that of a 2 votes threshold.

Having established that 31 days data is the minimum requirement for the behaviour modelling, general comparison is carried out with the singular models are well as the ensemble approaches using 31 days data for training. The result showing the model accuracy for the SmartNTU data is shown in Figure 4.8 while that of the CASAS HH111 data is shown in Figure 4.9. It can be seen that the CNDE outperformed all the other approaches on both datasets, thereby asserting the effectiveness of the proposed ensemble model.

4.3.5 Discussion

Based of the results presented in Figure 4.8, Figure 4.9 and Table 4.3, it can be seen that the proposed ensemble model outperformed the other existing approaches achieving an overall accuracy of 98% and 96% for the SmartNTU and CASAS HH111 datasets respectively, while at the same time having nearly a linear runtime complexity in all cases. Additionally, the weight estimation algorithm allows for the easy identification of better performing models for any given dataset. This is important since poor performing models for a given dataset can be identified and discarded from the ensemble. To confirm this assertion, an experimental evaluation is carried out by discarding the least performing models from the ensemble and the result is presented in Appendix A. Using the weight estimation approach in Algorithm 4, the best and least performing models are identified. For the SmartNTU dataset, the least performing model is identified as the Robust Covariance Estimation (RCE)

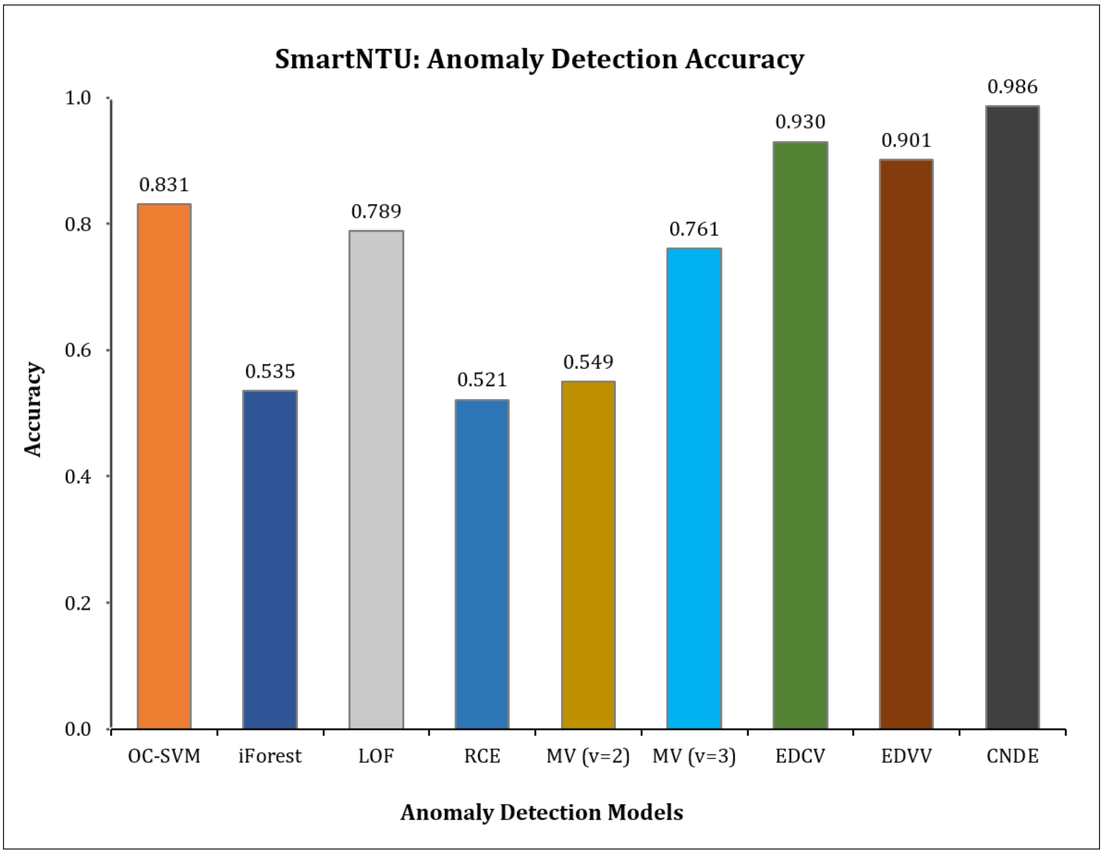


Figure 4.8: Result of the anomaly detection models for SmartNTU dataset.

while for the CASAS HH111 dataset, the least performing model is the Isolation Forest (iForest). The ensemble model is then evaluated by excluding the poor performing models and compared to the initial evaluation results (i.e. the results obtained when no model is discarded). From the presented result in Appendix A, it can be seen that discarding the poor performing models does not affect the overall accuracy of the ensemble approach. This is expected since the models contribute less to the overall prediction of the CNDE approach. However, a benefit derived from discarding the models is the improvement in the evaluation runtime of the model. This is because the time allocated for training the poor performing models is not required, thereby speeding up the evaluation process. The possibility of dynamically adjusting the normality threshold without explicitly retraining the models gives more flexibility to incorporate changes in the human behavioural routine.

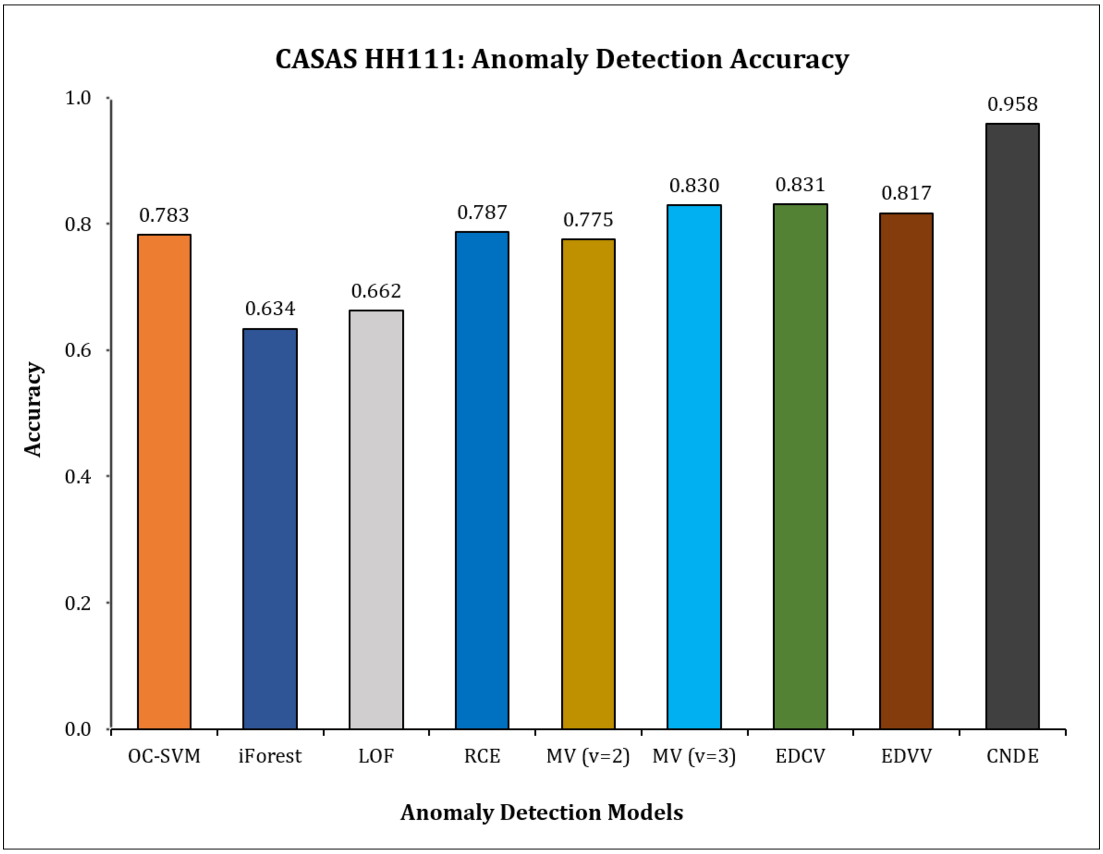


Figure 4.9: Result of the anomaly detection models for CASAS HH111 dataset.

Novelty detection models are created on the promise that there is only one set of available training data. This means that the training data contains none or a negligible amount of outliers [135]. A significant amount of outliers (noise) in the training data can drastically affect the performance of the models, and therefore, the proposed ensemble approach. In a scenario where a significant amount of outliers exists in the training set, the most feasible approach is to reduce the class imbalance problem by undersampling the majority class or oversampling the minority class. Supervised learning algorithms can then be utilised to classify the data as applied in [18]. While the proposed ensemble approach is applied in a batch manner where all the needed training data are available, it has the potential of being utilised in an incremental learning scenario where the models are required to adapt to new data when they become available, as discussed in Chapter 6.

4.4 Conclusion

This chapter presents an ensemble approach for novelty detection models based on the concept of internal and external consensus for the detection of ADL anomalies. This approach allows the weights of the models in the ensemble to be estimated during training based on the models' performance, thereby making the identification of suitable models to be incorporated in the ensemble possible. The resulting output of the ensemble approach is a score termed as "Normality Score" qualifying the data as either inliers or outliers. This offers more flexibility since the threshold of the normality score can be dynamically adjusted to incorporate changes in human activities, thereby allowing new or unknown data entries to be learned incrementally. The experimental evaluation of the CNDE on ADL datasets outputs an excellent result. While the CNDE is able to detect ADL anomalies, the features of the dataset that are likely to be the source or reason for the outlier prediction are not identifiable. The next chapter presents a novel similarity measure approach for identifying the sources of anomalies in ADL data based on the prediction result of the ensemble model presented in this chapter.

Chapter 5

Similarity Measure for Anomaly Source Identification

5.1 Introduction

In the preceding chapter, an ensemble approach for anomaly detection models was proposed. The limitation of the proposed approach as well as existing anomaly detection models is their inability to identify the anomaly sources. The term "Source(s) of Anomaly" is used to refer to the feature of the data entry responsible for the classification of the entry as an outlier. For example, for an activity-based anomaly detection system for sleeping routine, different features of the sleeping activity such as the time, duration, day or week, interruptions, activity transitions, proceeding or succeeding activities can be the reason for the abnormality. As an illustration, Figure 5.1 shows a sample plot representing a sleeping activity pattern with two features, sleeping start time and duration, selected from several features for easy visualisation. The clusters "X1", "X2" and "X3" represent the normal activity routines while A, B, and C indicate the outliers. Existing computational models can identify the anomalies but not the sources of the anomalies. The sources can be observed from Figure 5.1 as follows: I) the anomaly source for "A" is the duration feature indicating insufficient sleep or oversleeping, II) the anomaly source for "B" is the start time feature indicating going to bed early or late while III) the anomaly source

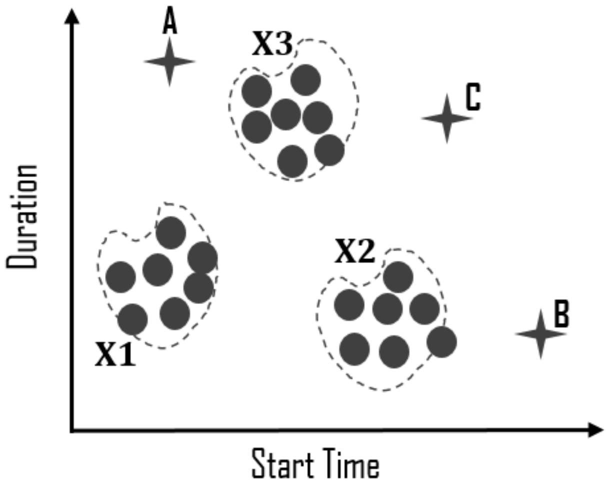


Figure 5.1: A sample plot representing sleeping activity with three outliers A, B and C.

for “C” is both the start time and duration features.

The main contribution of this chapter is the proposal of a similarity measure approach for identifying the sources of anomalies in human activities. The proposed exploratory approach is based on a pairwise distance measure of the features extracted from the activity data. Two approaches for measuring the pairwise distance are presented and termed as One vs One Similarity Measure (OOSM) and One vs All Similarity Measure (OASM). Statistical measures are then applied to estimate the threshold of the extracted features with features exceeding the threshold predicted as the anomaly sources. The CNDE model introduced in Chapter 4 is applied for the detection of outliers in the data while the proposed similarity measure approach is applied to the predicted outliers in order to identify their sources.

The rest of this chapter is organised as follows: Section 5.2 presents the proposed similarity measure approach. Section 5.3 contains experimental evaluation of the approach while a conclusion of the work is provided in Section 5.4.

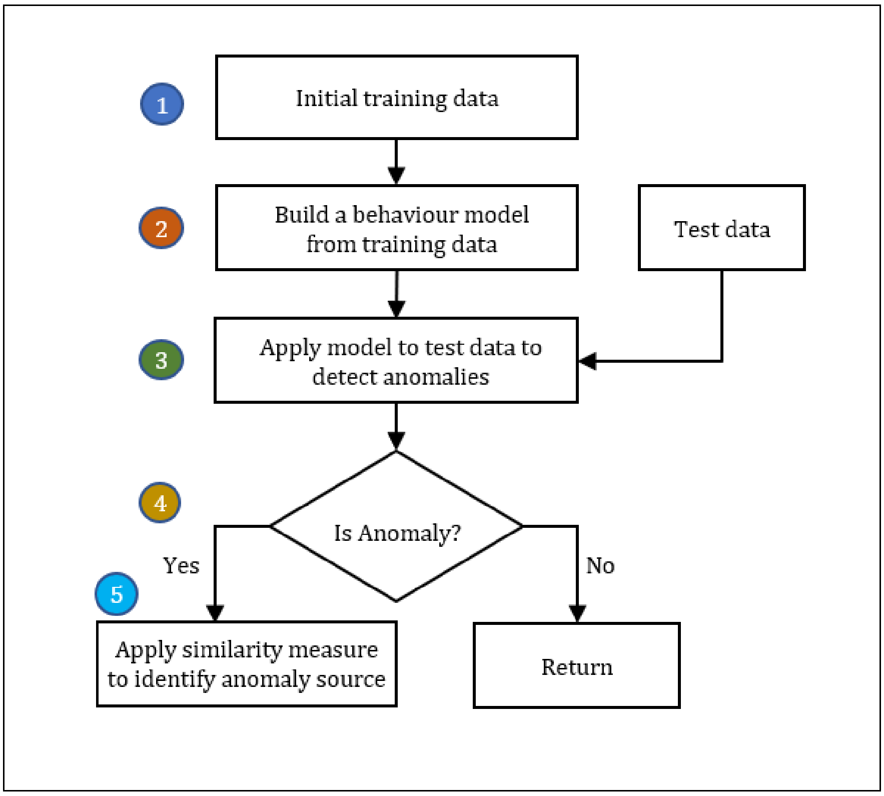


Figure 5.2: A workflow of the anomaly detection and source identification process.

5.2 Methodology

The proposed workflow for identifying the sources of abnormalities in human activities presented in 5 steps is shown in Figure 5.2. To detect the anomaly sources, the underlining ADL data must contain outliers identifiable using computational models. Detecting abnormalities is a complex task due to the variability of activities from one individual to another, thereby leading to the absence of standardised datasets representing normal and abnormal activities. Additionally, the existence of anomalies in different activities makes the anomaly detection task rather arduous. To address this, a data-driven activity-based approach is adopted whereby the computational models for anomaly detection are trained on data representing the activities of interest.

The initial set of data collected from the home environment representing the usual behavioural routine of the monitored individual is used to train the

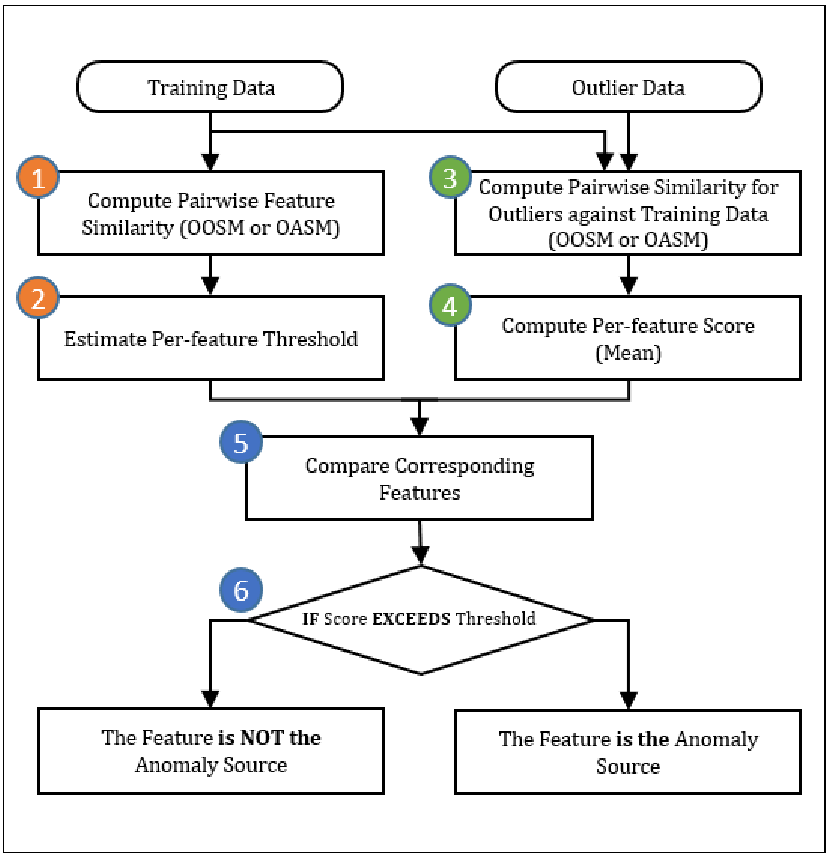


Figure 5.3: A schematic diagram of the proposed similarly measure approach.

novelty detection models to serve as a baseline. The baseline model is then utilised on subsequent activity data to detect deviating routines that constitute an abnormality. In this context, the result of the model for each activity is considered as either *normal* or *anomaly*. The observations predicted as anomalies are then explored further to identify the actual reason for the abnormality. In the remaining part of this section, the exploratory similarity measure approach for identifying the anomaly sources (stage 5 of the workflow) is presented.

Computational models identify outliers by measuring the similarity between data points such that entries with a significant variance in the measured similarities are predicted as anomalies. The process of identifying the sources of the anomaly as shown in the schematic diagram in Figure 5.3 involves:

- Estimating the similarity matrix of the training data.
- Estimating the threshold of the similarity matrix.
- Computing the similarity matrix of the anomalous observation and,
- Calculating the similarity score of the anomalous observation to identify the source.
- Perform a feature by feature comparison of the similarity measure entry for the anomalous and training observation.
- Predict a feature as the outlier source if it exceeds the defined threshold.

Different similarity measures exist in the literature for numerical and categorical data with distance functions such as Euclidean, Minkowski, Manhattan distances commonly used as a measure of similarity for numerical data [171]. In general, machine learning models for both supervised and unsupervised learning utilise distance measures. For example, KNN uses distance functions such as Euclidean and Manhattan to measure the distances between a data point and its neighbours, with instances having relatively close neighbours classified as observations of the same class [78, 171]. Clustering algorithms such as K-Means uses distance measure to estimate the proximity of the data entries to their assigned cluster centroids [171].

Two approaches for the distance measurement are proposed and termed as; "One vs One Similarity Measure (OOSM)" and "One vs All Similarity Measure (OASM)". The approaches involve the pairwise distance measurement of the corresponding features of the dataset. We utilised three distance functions namely; Euclidean, Chebyshev and Canberra distance satisfying the fundamental distance measure requirements as follows [171]:

- Non-negativity rule: $D(x, y) \geq 0$
- Symmetry rule: $D(x, y) = D(y, x)$
- Identity of indiscernible rule: $D(x, y) = 0$ iff $x = y$

In the OOSM, the distance is measured between corresponding features of the data entries one at a time. For example, the distance of the i^{th} feature of the first entry a_1 is measured against the i^{th} feature of the second entry a_2, then with that of the third entry a_3 and vice-versa across all the features as given in Equation 5.7. In the OASM, the distance of the corresponding features is measured all together in one single expression. For example, the distance of the i^{th} feature of the first entry a_1 is measured against the i^{th} feature of all the remaining entries as presented in Equation 5.11.

To express the baseline formula for the selected distance functions, let D be a distance function between two n-dimensional data entries x and y. The Euclidean, Chebyshev and Canberra distance are defined as follows:

$$\text{Euclidean distance: } D(x,y) = \sqrt{\sum_{i=1}^{n} \Big(x_i - y_i\Big)^2} \tag{5.1}$$

$$\text{Chebyshev distance: } D(x,y) = \max_{i}^{n} |x_i - y_i| \tag{5.2}$$

$$\text{Canberra distance: } D(x,y) = \sum_{i=1}^{n} \frac{|x_i - y_i|}{|x_i| + |y_i|} \tag{5.3}$$

5.2.1 One vs One Similarity Measure (OOSM)

Let $A = \{a_1, a_2, ..., a_n\}$ be a d-dimensional set of n normal activity data used for training the anomaly detection model. Let $\mathbb{S}(A, A)$ be the similarity matrix for A calculated as the distance of the pairwise features of the observations. Expressing the data A in a matrix form:

$$A = \begin{bmatrix} a_1^1 & - & - & a_1^d \\ - & - & - & - \\ - & - & - & - \\ a_n^1 & - & - & a_n^d \end{bmatrix} \tag{5.4}$$

Given the similarity matrix $\mathbb{S}(A, A)$ for A:

$$\mathbb{S}(A, A) = \begin{bmatrix} \mathbb{S}(a_1, A) \\ - \\ - \\ \mathbb{S}(a_n, A) \end{bmatrix} \tag{5.5}$$

The similarity $\mathbb{S}(a_i, A)$ for the i^{th} observation of A is presented in Equation 5.6 with the $\ominus$ notation indicating a pairwise distance measure.

$$\mathbb{S}(a_i, A) = \begin{bmatrix} a_i^1 & - & - & a_i^d \end{bmatrix} \ominus \begin{bmatrix} a_1^1 & - & - & a_1^d \\ - & - & - & - \\ - & - & - & - \\ a_n^1 & - & - & a_n^d \end{bmatrix} \tag{5.6}$$

using OOSM approach, $\mathbb{S}(a_i, A)$ is calculated as:

$$\mathbb{S}(a_i, A) = \begin{bmatrix} \partial(a_i^1, a_1^1) & - & - & \partial(a_i^d, a_1^d) \\ - & - & - & - \\ - & - & - & - \\ \partial(a_i^1, a_n^1) & - & - & \partial(a_i^d, a_n^d) \end{bmatrix}, \mathbb{S}(a_i, A) \in \mathbb{R}^{n\text{x}d} \tag{5.7}$$

where $\partial(a_i^k, a_j^k)$ is a 1-dimensional pairwise distance whose expression is obtained by transforming the three distance functions in Equation 5.1 - 5.3. The $\partial(a_i^k, a_j^k)$ can be expressed as:

$$\text{Euclidean: } \partial(a_i^k, a_j^k) = \sqrt{(a_i^k - a_j^k)^2} = |a_i^k - a_j^k| \tag{5.8}$$

$$\text{Chebyshev: } \partial(a_i^k, a_j^k) = \max |a_i^k - a_j^k| = |a_i^k - a_j^k| \tag{5.9}$$

$$\text{Canberra: } \partial(a_i^k, a_j^k) = \frac{|a_i^k - a_j^k|}{|a_i^k| + |a_j^k|} \tag{5.10}$$

5.2.2 One vs All Similarity Measure (OASM)

The similarity matrix $\mathbb{S}(a_i, A)$ in Equation 5.6 is calculated as:

$$\mathbb{S}(a_i, A) = \left[\partial(a_i^1, a_{j=1,..,n}^1) \quad - \quad - \quad \partial(a_i^d, a_{j=1,..,n}^d)\right], \mathbb{S}(a_i, A) \in \mathbb{R}^{1xd} \tag{5.11}$$

The 1-dimensional pairwise distance $\partial(a_i^k, a_{j=1,..,n}^k)$ obtained by transforming Equation 5.1 - 5.3 is expressed as:

$$\text{Euclidean: } \partial(a_i^k, a_{j=1,..,n}^k) = \sqrt{\sum_{j=1}^{n}(a_i^k - a_j^k)^2} \tag{5.12}$$

$$\text{Chebyshev: } \partial(a_i^k, a_{j=1,..,n}^k) = \max_{j=1}^{n} |a_i^k - a_j^k| \tag{5.13}$$

$$\text{Canberra: } \partial(a_i^k, a_{j=1,..,n}^k) = \sum_{j=1}^{n} \frac{|a_i^k - a_j^k|}{|a_i^k| + |a_j^k|} \tag{5.14}$$

The threshold $\overrightarrow{\mathbb{S}}(A)$ of the features (columns) of $\mathbb{S}(A, A)$ in Equation 5.5 is a row vector estimated from the positively (right) skewed distribution of the entries using median and interquartile range rule for outlier detection [157].

$$\overrightarrow{\mathbb{S}}(A) = \left[\delta^1 \quad - \quad - \quad \delta^d\right], \overrightarrow{\mathbb{S}}(A) \in \mathbb{R}^{1xd} \tag{5.15}$$

$$\delta^k = \alpha^k + e\beta^k \tag{5.16}$$

where α, β are the median and interquartile range of the k^{th} feature while e is a constant (usually set to 1.5).

Given a set of d-dimensional outliers $B = \{b_1, b_2, ..., a_m\}$, the anomaly source for each observation b_i is determined by estimating its similarity with the non-outlying data using Equation 5.17. The $\ominus$ notation in the expression indicates a pairwise distance measure.

$$\mathbb{S}(b_i, A) = \begin{bmatrix} b_i^1 & - & - & b_i^d \end{bmatrix} \ominus \begin{bmatrix} a_1^1 & - & - & a_1^d \\ - & - & - & - \\ - & - & - & - \\ a_n^1 & - & - & a_n^d \end{bmatrix} \tag{5.17}$$

$\mathbb{S}(b_i, A)$ is calculated using either OOSM in Equation 5.7 or using OASM in Equation 5.11. The similarity score $\tilde{\mathbb{S}}(b_i, A)$ for b_i is a row vector obtained by calculating the feature mean of $\mathbb{S}(b_i, A)$ from Equation 5.17.

$$\tilde{\mathbb{S}}(b_i, A) = \begin{bmatrix} \rho^1 & - & - & \rho^d \end{bmatrix}, \tilde{\mathbb{S}}(b_i, A) \in \mathbb{R}^{1\text{x}d} \tag{5.18}$$

$$\rho^k = \begin{cases} \text{mean}(\mathbb{S}(b_i^k, A)) & \text{for the OOSM} \\ \mathbb{S}(b_i^k, A) & \text{for the OASM} \end{cases} \tag{5.19}$$

where ρ^k is the score for the k^{th} feature of the observation b_i.
The function $f(\rho^k)$ that determines if the k feature is the anomaly source is expressed as:

$$f(\rho^k) = \begin{cases} 0 \leq \rho^k \leq \delta^k & \text{The } k^{\text{th}} \text{ feature is NOT the anomaly source} \\ \text{otherwise} & \text{The } k^{\text{th}} \text{ feature is the anomaly source} \end{cases} \tag{5.20}$$

The anomaly source is identified in Equation 5.20. If the similarity score of the feature ρ^k has an abnormally large variance to the feature threshold δ^k, the k^{th} feature of the observation is the anomaly source.

The procedure to implement the proposed similarity measure is provided in Algorithm 5. In the next section, an empirical evaluation of the approaches is carried to identify the source(s) of the ADL anomalies in the datasets.

5.3 Experimentation

The ADL datasets describe in Section 3.4 are used for the evaluation of the proposed similarity measure approach. This approach for identifying the sources

Algorithm 5 Similarity Measure

Input: Training set $A = \{a_i\}, i = 1, ..., n$
Outlier observation b_j
Output: Features list (sources) $f = \{\}$

procedure Anomaly Source
  $\mathbb{S}(A, A) \leftarrow$ Compute similarity of A
  $\overrightarrow{\mathbb{S}}(A) = \{\delta^k\}, k = 1, ..., d \leftarrow$ Estimate threshold of $\mathbb{S}(A, A)$
  $\mathbb{S}(b_j, A) \leftarrow$ Compute similarity of outlier b_j
  $\tilde{\mathbb{S}}(b_j, A) = \{\rho^k\}, k = 1, ..., d \leftarrow$ Compute score of $\mathbb{S}(b_j, A)$
  for $k = 1 : d$ **do**
    if $\rho^k \geq 0$ and $\rho^k \leq \delta^k$ **then**
      $f \leftarrow \text{index}(k)$ ▷ k^{th} feature is a source
    else
      $f \leftarrow f$ ▷ k^{th} feature is NOT a source
    end if
    end if
  end for
  end for
  return f
end procedure

of anomaly applies to data already classified as outliers by the anomaly detection models. Therefore, the anomaly detection results in the previous chapter are utilised. The results indicate that the models are able to identify abnormalities in the sleeping activity, which may be due to the variability in the extracted features.

The observations from the datasets (SmartNTU and CASAS HH111) that are predicted as outliers are further explored using the proposed similarity measure to identify the anomaly sources. After the oversampling of the datasets with Synthetic Minority Oversampling Technique (SMOTE) prior to the model training, the respective datasets contains approximately 100 anomalous observations with a relatively equal number of anomaly sources in four of the most discriminating features that are identified using Principal Component Analysis (PCA). The four features are the start time, duration, interruption, and interruption length.

The outlying observations are explored, and the features of the data that are the likely anomaly sources are identified. Since multiple features can be identified as the source, A condition is set such that the identification of at least one anomaly source per activity is sufficient. In a scenario where multiples anomaly sources (features) are identified, a true positive prediction is assigned if one of the predicted sources matches the ground truth.

To evaluate the performance of the similarity measure approaches, the accuracy, precision and recall are measured using Equations 3.19 - 3.21. The ground truth of the anomaly sources is identified through the manual observation and annotation of the outliers on a feature by feature basis and from the annotated data used for anomaly detection. The features that are altered to simulate an abnormally large variance for a given data entry are labelled as the sources of the anomaly for that entry. The precision and recall are computed on a feature by feature basis while the accuracy is computed across all the features for the given dataset. In this context of anomaly source identification, the following four parameters for measuring the accuracy, precision and recall are defined as follows:

- *True Positive (TP):* when a feature of the data entry labelled as the anomaly source is correctly predicted as the source by the model.

- *False Positive (FP):* when a feature of the data entry that is not labelled as the anomaly source is wrongly predicted as the source by the model.

- *True Negative (TN):* when a feature of the data entry labelled as not the anomaly source is correctly predicted as not the source by the model.

- *False Negative (FN):* when a feature of the data entry labelled as the anomaly source is wrongly predicted as not the source by the model.

For a more robust analysis, additional synthetic data is generated with different anomalous entries for the duration of 100 days. The features of the generated data are based on the distribution of the midpoints for the first and second quartile of the SmartNTU dataset. The simulated outliers sources are set as the third quartile of the data distribution of the features. This data, along with the two real datasets, are then used to validate the similarity measure approaches. The result obtained for the similarity measure approaches based on Euclidean distance is presented in Figure 5.4 showing the precision and recall for the different features. Similarly, Figure 5.5 and Figure 5.6 shows the results for the Chebyshev and Canberra distances respectively. It can be seen that similarity measure approaches are able to identify the anomaly sources across the validation datasets. A summary of the performance metrics is tabulated in Table 5.1 and Table 5.2.

5.3.1 Discussion

From Table 5.1, it can be seen that the obtained result is slightly better for the SmartNTU dataset compared to that of the CASAS HH111 dataset. However, the result for the Synthetic data surpasses that of the real datasets. This is expected since the synthetic data is carefully crafted to fit the model's requirement and therefore, should not be used as a measure of accuracy given that the models can be biased towards the data. The average accuracy in Table 5.2 is grouped based on the distance function and based on the similarity measure approach across the two real datasets (SmartNTU and CASAS HH111). The result of the synthetic data is excluded in the computation due to its biased nature. It is evident that the Euclidean distance function has the best performance while the Canberra distance has the least. Additionally, the OASM approach outperforms the OOSM approach across the datasets. This gives the OASM approach an additional advantage in addition to having a linear runtime complexity $O(n)$ over the OOSM approach that has a quadratic runtime complexity $O(n^2)$.

The identified anomaly sources (e.g. start time, duration etc.) may not present a significant meaning in an AAL context. To better understand the anomalies, a mapping is created from the extracted feature space to the real-life space. This means that the extracted features are translated and attributed to real-life causes of the anomalies. For example, if the anomaly source is the start time, then the real-life source could be an indication of going to bed early or later than usual. Similarly, if the anomaly source is the duration, then this can be attributed to having an insufficient sleep or oversleeping. The mapping may vary depending on the extracted features and their ground truth representation. A

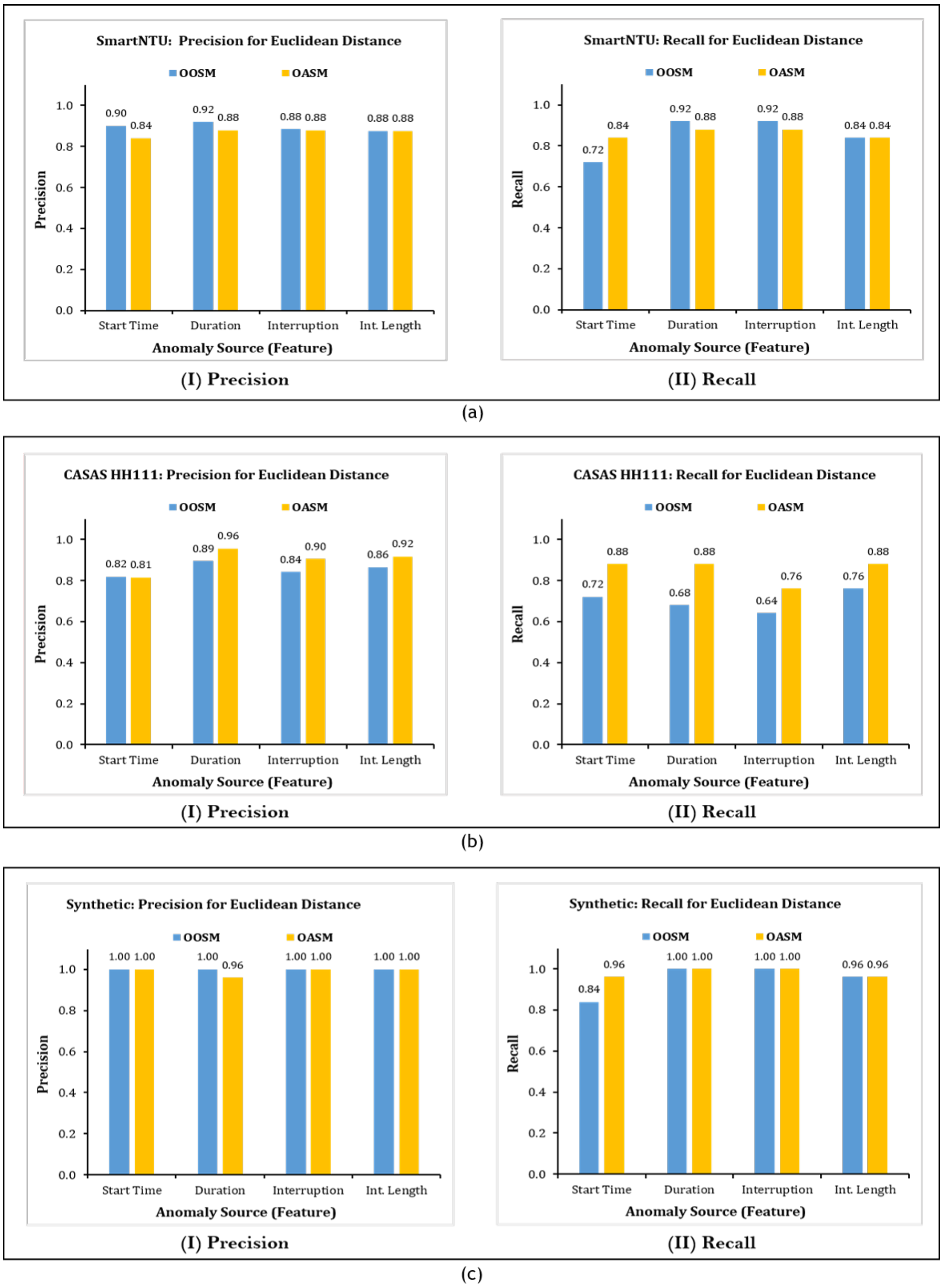
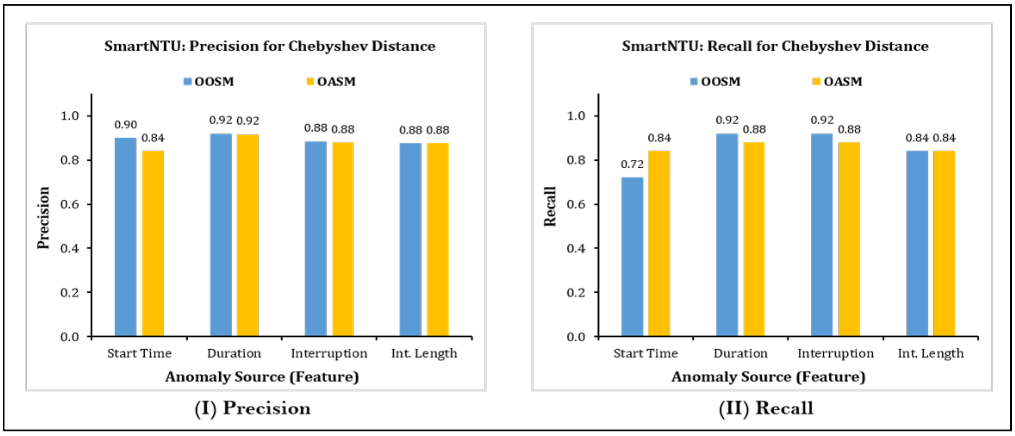
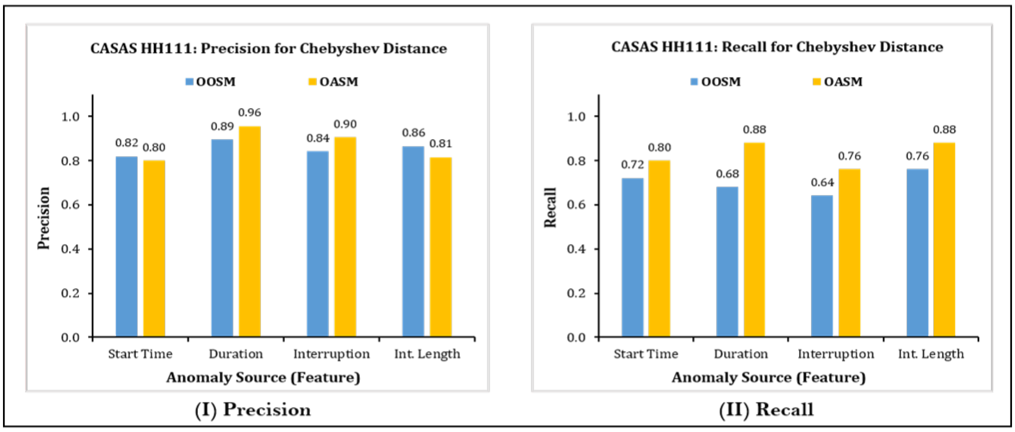


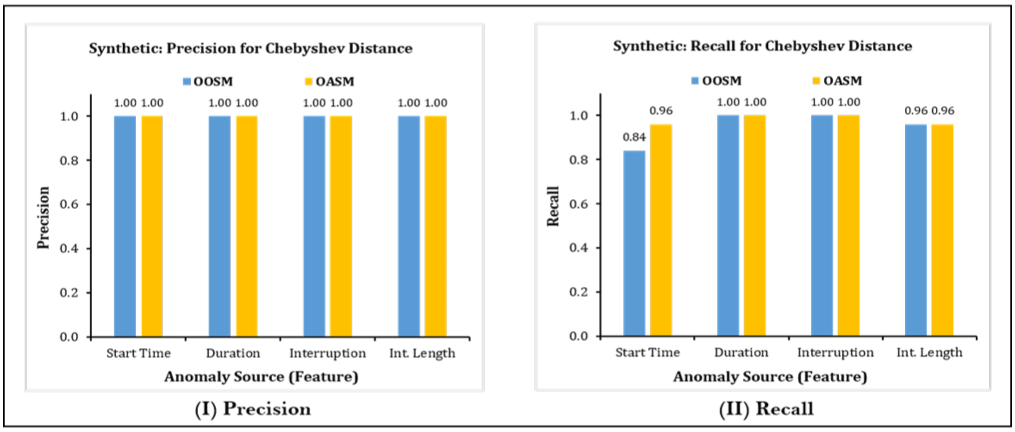
Figure 5.4: Result for the similarity measures based on Euclidean distance on; a) SmartNTU dataset, b) CASAS HH111 dataset, c) Synthetic dataset.



(a)



(b)



(c)

Figure 5.5: Result for the similarity measures based on Chebyshev distance on; a) SmartNTU dataset, b) CASAS HH111 dataset, c) Synthetic dataset.

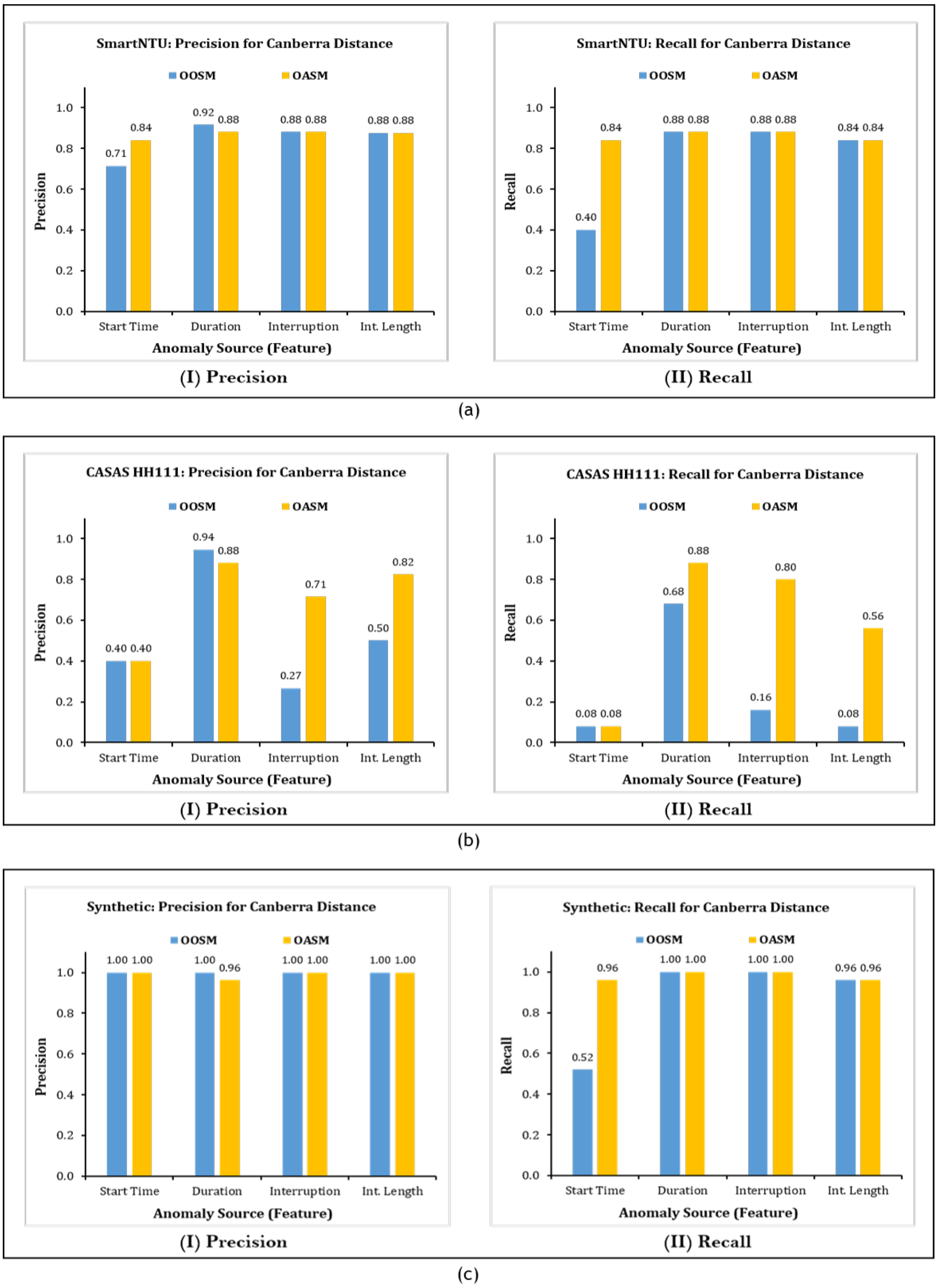


Figure 5.6: Result for the similarity measures based on Canberra distance on; a) SmartNTU dataset, b) CASAS HH111 dataset, c) Synthetic dataset.

sample of the translated anomaly sources is summarised in Table 5.3 indicating the identified source and its real-life mapping observed from the data. To the best of our knowledge, no other exploratory approach exists in the literature for identifying anomaly sources in this context, thereby making a comparative analysis of the proposed similarity measure nearly impossible.

Additionally, the datasets used for this validation is made up of small data samples with fewer features. In a scenario where the dataset is significantly large or contains a large feature set, the proposed approach for estimating the similarity matrices may be time-consuming. This is because the similarity is measured for each outlying observation against the entire training samples across all the features. A heuristic approach for selecting fewer samples of the training data to be used for the similarity estimation can improve the efficiency of the methodology.

Table 5.1: Overall accuracy of the similarity measure approaches.

Dataset	Distance Function	OOSM	OASM
SmartNTU	Euclidean	0.85	0.86
	Chebyshev	0.85	0.86
	Canberra	0.75	0.86
CASAS HH111	Euclidean	0.70	0.85
	Chebyshev	0.70	0.83
	Canberra	0.25	0.58
Synthetic	Euclidean	0.95	0.98
	Chebyshev	0.95	0.98
	Canberra	0.87	0.98

Table 5.2: Average accuracy of the similarity measures grouped by distance function and by similarity measure (SmartNTU and CASAS HH111 datasets).

Description	Measure	Average Acc.
Based on distance function	Euclidean	**0.82**
	Chebyshev	0.81
	Canberra	0.61
Based on similarity measure	OOSM	0.68
	OASM	**0.81**

Table 5.3: Identified anomaly sources for SmartNTU and CASAS HH111 dataset.

Dataset	Day	Data Feature	Interpreted Source	Detailed Description
SmartNTU	34, 46 & 55	Duration	Less sleeping	The individual has insufficient sleep compared to his/her usual sleep duration
	51	Duration	Over sleeping	The individual sleeps for a longer duration than usual
	39 & 69	Interruption	Interrupted sleep	The individual experienced several unusual interruptions while sleeping
	59	Start time	Going to bed late	The individual goes to bed late compared to his/her usual bedtime
CASAS HH111	50	Duration	Less sleeping	The individual has insufficient sleep compared to his/her usual sleep duration
	39 & 49	Interruption	Interrupted sleep	The individual experienced several unusual interruption
	37, 44 & 46	Interruption Length	Interrupted sleep	The individual experienced a long transition during sleeping time. This may be an indication of performing an activity in the middle of the night
	32, 33 & 47	Start time	Going to bed Early	The individual goes to bed earlier than usual

5.4 Conclusion

This chapter proposes a novel exploratory approach for identifying the sources of abnormalities in human activities based on a similarity measure approach. This is relevant as existing methodologies are unable to identify the sources of the abnormalities, thereby hindering the development of adaptive monitoring systems with reduced false prediction rate. The proposed approach involves the pairwise distance measurement of the extracted features of the activity dataset in the form of a similarity matrix. A statistical approach is then applied to estimate the optimal threshold of the features and for the identification of the anomaly sources. Experimental evaluation of this approach achieved an overall average accuracy of 0.82, showing the credibility of the proposed approach for utilisation in an in-home monitoring system. In the next chapter, a continuation of this work is carried out by making the anomaly detection approach adaptive to the changes in human behaviour over time.

Chapter 6

Adaptive Anomaly Detection Approach

6.1 Introduction

This chapter presents an adaptive model-based approach to anomaly detection system in ADL. The approach adapts to new data corresponding to changes in the human behavioural routines over time. A data-driven filtering approach termed as "Forgetting Factor" is presented. It allows the system to identify outdated activity data to be discarded while incorporating newly identified data representing the changes in human behaviour for adaptation. This is a practical approach since the goal is for the system to adapt to the current behavioural routine of an individual while discarding irrelevant characteristics feature of the old routines. Two forgetting factor approaches are proposed, namely, Forgetting Factor based on Data Ageing (FFDA) and Forgetting Factor Based on Data Dissimilarity (FFDD) while using the CNDE approach proposed in Chapter 3 for the data modelling.

The existing systems for abnormality detection are not often adaptive to changes in human behavioural routine, thereby generating high false alarm rate [70, 73]. Changes in daily routines are often predicted as abnormalities due to this lack of adaptation. A self-supervised approach called Abnormal Event Detection Network (AED-Net) for detecting anomalies in crowded scenes is proposed in [165] while another approach based on Conditional Random Field (CRF) and Mixtures of Dynamic Textures (MDT) is proposed in [98]. However, these two approaches utilised video stream data which differs from our use case where binary ADL data collected using ambient sensors are utilised. Moreover, a sufficient amount of training data is required for the modelling to achieve an acceptable result. In a limited data scenario, or where the data is overly dynamic and changes over time, a model capable of learning continuously and adapting to new data is desirable. This can be deployed in a dynamic environment and can play a significant role in ADL anomaly detection.

The rest of this chapter is organised as follows: Section 6.2 presents the formulation of the proposed adaptive approach. In Section 6.3, the methodology is validated and the obtained results along with a discussion are provided. A conclusion of the work is provided in Section 6.4.

6.2 Methodology

The adaptive model pipeline allows the system to adapt to novel data representing changes in human activity routine. Novel activity instances may be a reflection of actual abnormality or a behavioural change requiring confirmation from an external agent (such as humans). Figure 6.1 shows an overview of the proposed adaptive approach with support for a human confirmation of the detected abnormalities. To realise this, the following assumptions and constraints in line with the research objectives are defined.

- *Assumption 1*: Anomalous activities of interest occur consecutively over a time period. Isolated outlier instances may be due to erroneous data, misprediction or rarely a true anomaly.

- *Assumption 2*: A consecutive occurrence of anomalous activity could be an indication of behavioural change or abnormality due to health-related challenges. An external communication intermediary can be incorporated to verify the status of the detected outliers as shown in Figure 6.1.

- *Assumption 3*: Anomalous instances verified as behavioural changes

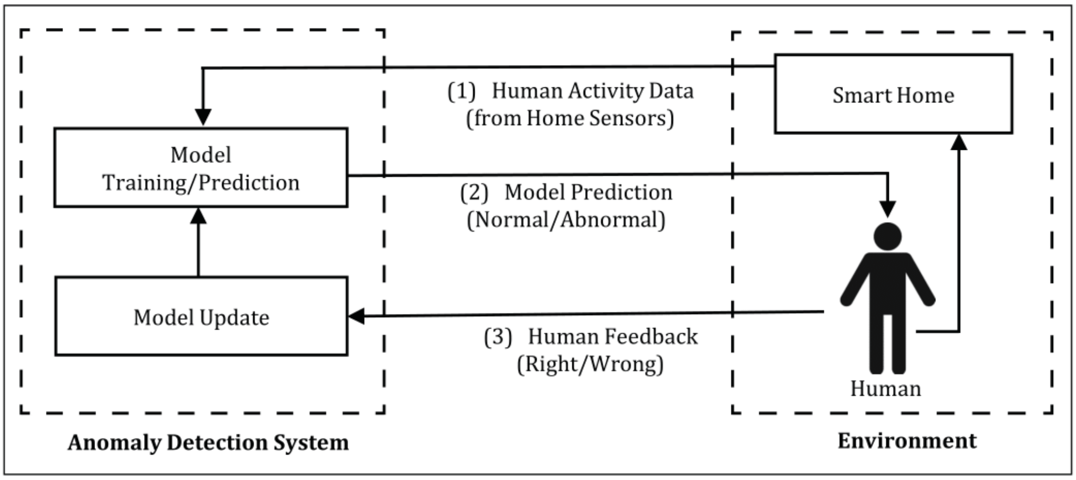


Figure 6.1: An overview of the adaptive anomaly detection system.

should be incorporated into the system for adaptation. A forgetting factor is introduced to discard outdated behavioural pattern while incorporating the currently verified routines.

6.2.1 Forgetting Factor

TThe forgetting factor is introduced to enable the filtering of the activity data used for updating the model. The characteristic features of the newly verified data are incorporated into the system by updating the anomaly detection model while the outdated features are discarded. This is achieved by filtering both the existing and new activity dataset to remove entries that do not conform to the current behavioural routine of the individual based on the two forgetting factor approaches, i.e. FFDA and FFDD.

6.2.1.1 Forgetting Factor based on Data Ageing (FFDA)

In this approach, the existing training data entries are discarded based on the age of the data. This is based on the hypothesis that the oldest activity data represents an obsolete behavioural routine, while the recent data represents the current routine of an individual. Since the activity data contains a timestamp, the entries are sorted using the timestamp and the oldest entries are replaced

Algorithm 6 Forgetting Factor based on Data Ageing

D_{Old} is the existing data
D_{New} is the newly verified data ▷ (n = dataset size)
D_{Final} is the filtered data
procedure FORGETOLDESTDATA(D_{Old}, D_{New})
    Sort (D_{Old}) ▷ sort by data age
    RemoveOldest(D_{Old}, n) ▷ discard n-top entries
    $D_{\text{Final}} \leftarrow$ Append($D_{\text{Old}}, D_{\text{New}}$)
    return D_{Final}
end procedure

Algorithm 7 Forgetting Factor based on Data Dissimilarity

D_{Old} is the existing data
D_{New} is the newly verified data ▷ (n = dataset size)
D_{Final} is the filtered data
procedure FORGETDISSIMILARDATA(D_{Old}, D_{New})
    SimMatrix $\leftarrow$ EstimateSimilarity(D_{Old}, D_{New})
    Rank(SimMatrix) ▷ using Linear Regression
    RemoveDissimilarData(D_{Old}, n)
    $D_{\text{Final}} \leftarrow$ Append($D_{\text{Old}}, D_{\text{New}}$)
    return D_{Final}
end procedure

with the newly verified data. A practical limitation of this approach is that the oldest activity data may have more similarity to the current routine than some of the recent entries. Algorithm 6 contains the FFDA implementation procedures.

6.2.1.2 Forgetting Factor based on Data Dissimilarity (FFDD)

In this approach, the discarded activity data are identified based on a similarity measure. Data entries in the existing training data are measured for similarity against the newly verified data. A ranking of the data entries is performed based on linear regression to identify the most dissimilar entries. Unlike in the FFDA approach where the oldest data are assumed to be less representative of the current behavioural routine, the FFDD is more generic since it relies on the data similarity measure. The FFDD implementation procedures are given in Algorithm 7.

6.2.2 Data Similarity Measure

The proposed FFDD performs a similarity measure between the existing and the newly verified activity data entries. The similarity is measured using a Euclidean distance function. Computational models for classification and outlier detection utilises distance as a measure of similarity [171]. For example, K-Mean clustering uses distance measure to estimate the proximity between data entries and the centroids while KNN uses distance measure such as Euclidean distance to identify the neighbours of data entries for classification [78, 171]. The choice of Euclidean distance measure over other distance functions is supported by the experimental evaluation carried out in Chapter 5 in which Euclidean distance achieved a better result compared to Chebyshev and Canberra distance.

Let $A = \{a_1, a_2, ..., a_n\}$ be a d-dimensional set of n activity data used for the initial training of the anomaly detection model and $B = \{b_1, b_2, ..., b_m\}$ be a d-dimensional set of m verified data for adaptation. The similarity $\mathbb{S}(A, B)$ of A and B is a matrix estimated from the Euclidean distance between the pairwise features of the data. Expressing the data A and B in a matrix form:

$$A = \begin{bmatrix} a_1^1 & - & - & a_1^d \\ - & - & - & - \\ - & - & - & - \\ a_n^1 & - & - & a_n^d \end{bmatrix}, B = \begin{bmatrix} b_1^1 & - & - & b_1^d \\ - & - & - & - \\ - & - & - & - \\ b_m^1 & - & - & b_m^d \end{bmatrix} \tag{6.1}$$

Given the similarity measure matrix $\mathbb{S}(A, B)$ for A and B:

$$\mathbb{S}(A, B) = \begin{bmatrix} \overrightarrow{\mathbb{S}}(a_1, B) \\ - \\ - \\ \overrightarrow{\mathbb{S}}(a_n, B) \end{bmatrix}, \mathbb{S}(A, B) \in \mathbb{R}^{nxd} \tag{6.2}$$

The similarity $\mathbb{S}(a_i, B)$ for the i^{th} entry of A is calculated as the per feature distance after normalisation of the entry using Equation 6.3, with the $\ominus$ notation indicating a pairwise distance measure:

$$\mathbb{S}(a_i, B) = \begin{bmatrix} a_i^1 & - & - & a_i^d \end{bmatrix} \ominus \begin{bmatrix} b_1^1 & - & - & b_1^d \\ - & - & - & - \\ - & - & - & - \\ b_m^1 & - & - & b_m^d \end{bmatrix} \tag{6.3}$$

$$\mathbb{S}(a_i, B) = \begin{bmatrix} \partial(a_i^1, b_1^1) & - & - & \partial(a_i^d, b_1^d) \\ - & - & - & - \\ - & - & - & - \\ \partial(a_i^1, b_m^1) & - & - & \partial(a_i^d, b_m^d) \end{bmatrix}, \mathbb{S}(a_i, B) \in \mathbb{R}^{mxd} \tag{6.4}$$

where $\partial(a_i^k, b_j^k)$ is a 1-dimensional pairwise Euclidean distance expressed as:

$$\partial(a_i^k, b_j^k) = \sqrt{(a_i^k - b_j^k)^2} = |a_i^k - b_j^k| \tag{6.5}$$

The matrix $\mathbb{S}(a_i, B)$ is converted into a row vector $\overrightarrow{\mathbb{S}}(a_i, B)$ by calculating its per-feature (column) mean.

$$\overrightarrow{\mathbb{S}}(a_i, B) = mean\Big(\mathbb{S}(a_i, B)\Big), \quad \overrightarrow{\mathbb{S}}(a_i, B) \in \mathbb{R}^{1xd} \tag{6.6}$$

$$\overrightarrow{\mathbb{S}}(a_i, B) = \Big[mean\Big(\partial_{i,1}^1, \partial_{i,m}^1\Big) \quad - \quad - \quad mean\Big(\partial_{i,1}^d, \partial_{i,m}^d\Big)\Big] \tag{6.7}$$

Since $\mathbb{S}(A, B)$ consist of row vectors, the vectors are converted into a scalar (ranked) for easy sorting as mentioned in Algorithm 7. The ranking $\mathbb{G}(A, B)$ of $\mathbb{S}(A, B)$ is determined using a linear regression model. Linear regression is expressed as the linear combination of the attributes(features) with their corresponding weights as expressed below:

$$\mathbb{G} = w^0 + w^1\mathbb{S}^1 + w^2\mathbb{S}^2 + ... + w^d\mathbb{S}^d = \sum_{k=1}^{d} w^k\mathbb{S}^k \tag{6.8}$$

where $[w^1, ..., w^d]$ are the feature weights, w^0 is the error term set to 0, $\mathbb{S}^k$ is a scalar representing the value of the k^{th} column for the given row, and d is the number of features in the dataset.

Algorithm 8 Procedure for computing Similarity Measure.

$A = \{a_i\}, i = 1, ..., n$ is the initial training data
$B = \{b_j\}, j = 1, ..., m$ is newly verified data
$\mathbb{S}(A, B)$ is an $[n, d]$ similarity matrix for A & B
procedure ESTIMATESIMILARITY(A, B)
    for $a_i \in A$ **do**
        $\mathbb{S}(a_i, B)$ is an $[m, d]$ similarity matrix
        for $b_j \in B$ **do**
            $\overrightarrow{V_r}$ is a $[1, d]$ row vector
            for $[k := 1 : d]$ **do** ▷ for-each feature
                $\partial := distance(a_i^k, b_j^k)$ ▷ euclidean dist.
                $\overrightarrow{V_r} \leftarrow \partial$ ▷ insert column to $\overrightarrow{V_r}$
            end for
            $\mathbb{S}(a_i, B) \leftarrow \overrightarrow{V_r}$ ▷ insert row to matrix
        end for
        $\overrightarrow{\mathbb{S}}(a_i, B)$ is a $[1, d]$ row vector
        for $[k := 1 : d]$ **do** ▷ for-each feature
            $\alpha^k := mean\Big[\mathbb{S}\Big(a_i^k, B^k_{\{b=1,...,m\}}\Big)\Big]$
            $\overrightarrow{\mathbb{S}}(a_i, B) \leftarrow \alpha^k$ ▷ insert as column
        end for
        $\mathbb{S}(A, B) \leftarrow \overrightarrow{\mathbb{S}}(a_i, B)$ ▷ insert row to matrix
    end for
    return $\mathbb{S}(A, B)$
end procedure

The PCA is applied during the training of the anomaly detection model to determine the most discriminating features of the dataset. The weights of the features (ranging from $[0, 1]$) is taken as the percentage fraction of the ranking of the features based on PCA using the expression $(1 - \frac{\text{Feature Rank (r)}}{\text{Number of Features (d)}})$. The weight of the most discriminating feature is taken as $(1 - \frac{1}{d})$, while that of the second most discriminating feature is given as $(1 - \frac{2}{d})$ and vice versa. Algorithm 8 outlines the procedure for computing the similarity measure for the given sample datasets A and B.

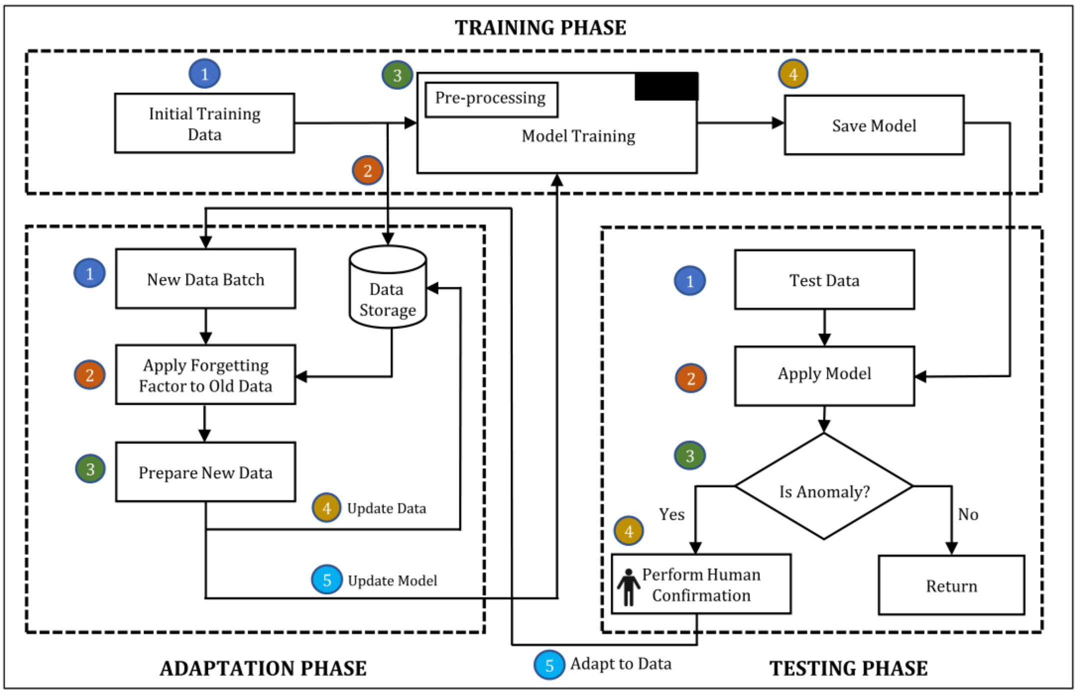


Figure 6.2: A schematic diagram of the data-driven adaptive system pipeline.

6.2.3 Adaptive System Pipeline

The approaches described earlier are combined to form the adaptive system pipeline as shown in the schematic diagram in Figure 6.2. This proposed adaptive approach is activity-dependent and data-driven. Its goal is to adapt to the newly verified data while forgetting the characteristic features of the outdated data identified using the proposed forgetting factor approaches (i.e. FFDA and FFDD). The forgetting factor feature makes the system environment invariant and allows for easy adaptation to a different environment and datasets.

Considering that the proposed approach may be similar to a brute-force method of appending (merging) the newly verified data to the existing one, a significant difference is that this approach allows the system to forget the old behavioural routine data completely without retaining its characteristic features while a brute-force approach still retains features of the old behavioural data.

From the illustration in Figure 6.2, the training of the anomaly detection models (such as OC-SVM) is performed with the initial set of collected data. A copy of the data and the trained model are saved on a storage component of the system for future use. Subsequent data collected after the system deployment is validated by the model to detect outliers (anomalies). The result of the model prediction can either be normal or anomaly for any given activity. In an ideal scenario, a single abnormality prediction may not constitute a major concern as this may be due to erroneous sensor reading or model misprediction, while consecutive repeated anomalies will undergo human verification as highlighted in the defined constraints. In a situation where the prediction is confirmed to be wrong and the activities are identified as normal behavioural changes, the stored data is retrieved from the storage component. The proposed forgetting factor approaches are then applied to both the existing and newly verified activity data. A model update is performed and the storage component is updated with the consolidated data entries to enable future occurrences of similar activity pattern to be identified as normal behavioural routine.

While this approach puts more emphasis on adapting to changes in behavioural routine that are wrongly identified as outliers, the adaptive system can benefit from periodic model updates at regular intervals (e.g., weekly, monthly, quarterly, etc. based on deployment requirements and data size) with data of normal behavioural routines (i.e. data that are not identified as outliers) to ensure that the model reflects the current behaviour of the monitored individuals.

This chapter only aims to show how the model is able to adapt to new data representing the changes in the behavioural routine of an individual. The reporting and confirmation aspect of the detected abnormalities, which can be achieved through various modalities, is discussed in Chapter 7. The methodology proposed in this thesis as presented in Chapter 1 incorporate human in the loop and highlights the role of a human agent in the learning process of the system while utilising an assistive robot as an intermediary. Additionally, the detected anomalies are meant to be confirmed by the human agent before the data entries are transmitted to the computational model for adaptation. Therefore, the role of the human agent is to confirm whether detected activities are abnormal or not with the help of a communication intermediary. To facilitate the Human-Robot Interaction (HRI), different communication modalities are considered, such as touch, speech and gesture. A detailed discussion of the proposed gesture recognition model is presented in the next chapter.

6.3 Experimentation

The two datasets utilised for this evaluation are the SmartNTU dataset and CASAS HH111 dataset described in Section 3.4. The evaluation is conducted after the data preprocessing which include the removal of data instances with missing entries and erroneous readings, filtering and selection of activity, features extraction and normalisation. The proposed adaptive system is activity-dependent (i.e. for optimal performance, activities of interest must be selected and modelled independently) in which the sleeping activity is selected.

The baseline anomaly detection model in the adaptive system is the CNDE ensemble approach proposed in Chapter 4 with the recommended 31 days data used for training and a model contamination rate of 0.1 across all the four models in the ensemble. This baseline ensemble model is not adaptive by default, therefore, novel or new data entries are always predicted as outliers.

The system capability of forgetting outdated data while adapting to the newly verified ADL data is evaluated. To achieve this, a cross-dataset evaluation strategy is performed such that the anomaly detection model is trained on one dataset and evaluated for adaptation on the other dataset. For example, the model is trained on the SmartNTU dataset and adapted to the CASAS HH111 dataset and vice versa. This validation strategy is adopted to ensure that the system is environment invariant, i.e. the system can be migrated to a different environment with easy adaptation.

In the cross dataset validation technique, to evaluate the adaptive approach on the SmartNTU dataset, the first 31 days data is used for training while the remaining data (referred to as a segment of the training dataset) is used for validating the error rate of the model. The CASAS HH111 data is then used to validate the model adaptation rate, and the data is split into an adaptation set and a validation set with a duration of 30 and 35 days data, respectively.

The adaptation set is injected into the system incrementally using different batch sizes (such as 5, 10 and 15 days data). After each incremental injection of the adaptation data, the two validation sets (i.e. a segment of the training data for evaluating the model error rate and a segment of the cross-dataset for evaluating the adaptation rate) are used. The system is expected to adapt to the new dataset after a certain number of incremental injection of the adaptation data, therefore, the adaptation rate is expected to increase over time while the error rate is expected to decrease. The oversample versions of the datasets are used in order to achieve uniformity in the sample sizes.

The performance metrics of the proposed approaches with forgetting factor are computed and compared with other anomaly detection models, namely, OC-SVM, iForest, and the aforementioned CNDE ensemble approach (using the same model parameters). Additionally, a Deep Q-Network (DQN) is also implemented with two actions (normal and outlier) with the rewards of $+1$ and -1. For the DQN, the datasets are sufficiently over-sampled prior to training and the weights are saved when the training is complete. During the adaptation phase, the adaptation data batch is used to retrain the saved model weights of the DQN instead of training the model from scratch. For the baseline anomaly detection models, since the models are not adaptive by default, a brute-force injection of the adaptation data is performed, i.e. the adaptation data is appended to the original training data incrementally.

6.3.1 Discussion

The results obtained for the evaluation of the proposed adaptive approaches are shown in Figure 6.3 and Figure 6.4. The performance metrics measured are the adaptation rate and the error rate. The adaptation rate is the rate at which the system adapts to the newly verified data. This is achieved by calculating the prediction accuracy of the validation set. The adaptation rate is expected to increase over time as more data representing the new behavioural routine are injected into the model. The error rate (forgetting inverse) is the rate at which the system forgets the initial (old) training data. This is obtained by calculating the error rate of the model prediction on the segment of old training data. As

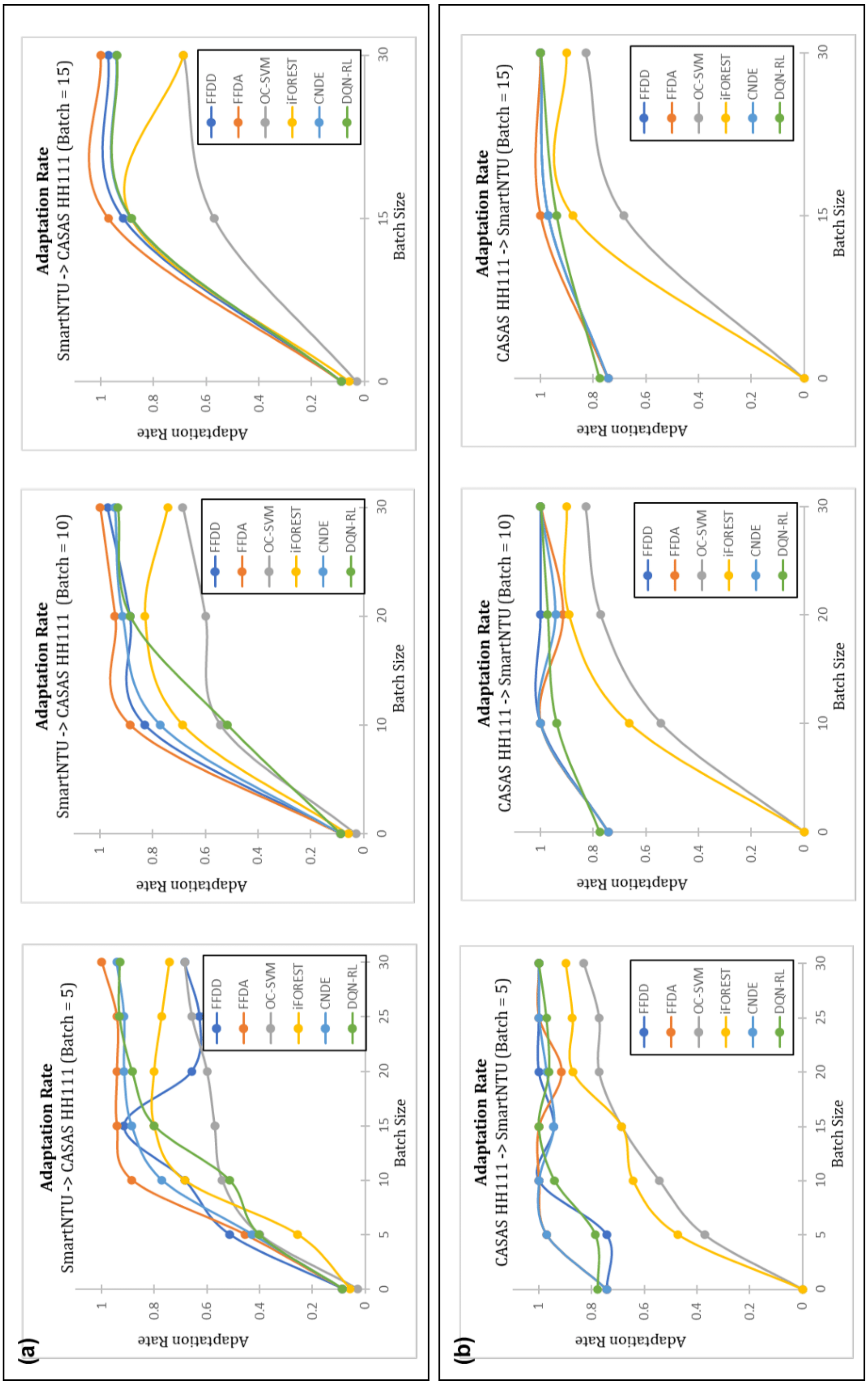


Figure 6.3: Adaptation rate plots; a) SmartNTU to CASAS HH111, b) CASAS HH111 to SmartNTU.

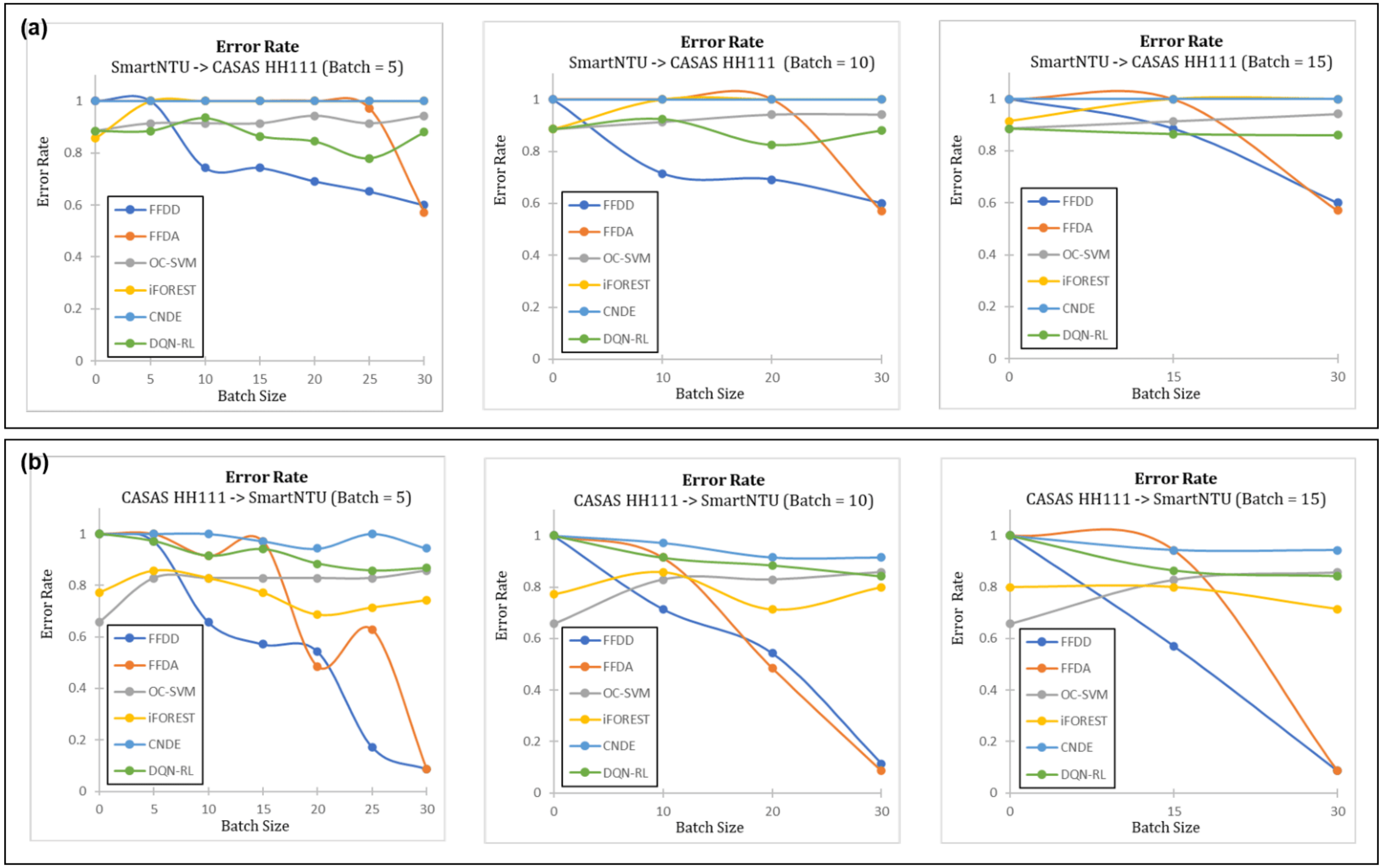


Figure 6.4: Error Rate (Forgetting Inverse) plots; a) SmartNTU to CASAS HH111, b) CASAS HH111 to SmartNTU.

highlighted in the previous section, the initial training data is split into a training set and a validation set prior to the model training. This remaining segment (validation set from the initial training data), is used to compute the error rate. It should be noted that this data segment is not used for the model training. The model is expected to predict this data segment as outliers, and an inlier prediction is considered an error. The error rate is expected to decline over time since the newly injected data (adaptation data) is from a different dataset and differ from the training data segment, thereby causing the model to forget the characteristic features of the initial training data gradually. Obtaining a lower error rate is the final goal since this is an indication that the model predicts the segment of the training as outliers after adaptation.

From the result in Figure 6.3, it can be seen that proposed approaches (FFDA and FFDD) outperformed the Reinforcement Learning (RL) model and the classical novelty detection models. The adaptation rate of the CNDE and DQN-RL is comparably close to that of the proposed approaches. The FFDD reaches the peak of its performance faster than the FFDA. This is because the FFDD removes dissimilar data at each iteration while FFDA only discards data based on the data age. A performance drop can be seen for the FFDD approach from Figure 6.3(a) when the batch size is set to 5. This is as a result of the continuous training of the model despite reaching its convergence, thereby causing the model to overfit. This is also evident from the result in Figure 6.3 since the drop in accuracy for the proposed approaches (FFDD and FFDA) appears only after the models reached the peak of the performance. The error rate plots in Figure 6.4 shows that only the proposed approaches in this chapter can forget (discard) the outdated data while the other models (i.e. CNDE, OC-SVM and iForest) performed poorly. This is expected since the models are not adaptive and the new adaptation data is appended using a brute-force approach. The RL model adapts to the adaptation data without retraining on the entire dataset by using the pre-trained weights but still retains the properties of the old data. Although the error rate can be improved by fine-tuning the model parameters, the properties of the old data may still be retained. This deviates from the objective of forgetting the data in its entirety. Additionally, the FFDD approach converges faster than the FFDA since

Table 6.1: A summary of the highest achieved adaptation rates.

Approach	Training: SmartNTU	Training: CASAS HH111
iForest	0.886	0.890
OC-SVM	0.686	0.829
CNDE	0.943	0.982
DQN-RL	0.940	0.982
FFDA	**0.992**	**0.982**
FFDD	**0.971**	**0.983**

Table 6.2: A summary of the lowest achieved error rates.

Approach	Training: SmartNTU	Training: CASAS HH111
iForest	0.857	0.686
OC-SVM	0.886	0.657
CNDE	0.989	0.914
DQN-RL	0.780	0.841
FFDA	**0.571**	**0.086**
FFDD	**0.600**	**0.086**

dissimilar entries are removed on each iteration. The FFDA will eventually converge when the old data entries are completely discarded based on their age but with a longer convergence runtime for large datasets. The best results obtained for the adaptation and error rates are summarised in Table 6.1 and Table 6.2 respectively.

The batch size for the incremental adaptation has a significant effect on the convergence rate as it can be seen from Figure 6.3 and Figure 6.4. A small batch size (e.g. 5 days) will result in a regular system update with a slow convergence rate (more prediction errors) and vice versa. From the obtained results, it is evident that the performance deteriorates after a peak is reached. This may be due to model over-fitting as a result of the continuous model update on similar data entries. Therefore, the model update can be terminated when the characteristic features of the adaptation data are similar to the existing ones. The termination of the model update can be realised in a real-life scenario through the human confirmation mechanism.

6.4 Conclusion

This chapter presents a data-driven adaptive system pipeline for detecting abnormalities in human activities. The adaptive approach is based on data ageing and data dissimilarity forgetting factor approaches referred to as FFDA and FFDD. These approaches enable the anomaly detection model in the system to adapt to changes in human behavioural routine while discarding the characteristic features of old routines. Experimental evaluation of these approaches results in an overall average adaptation accuracy of 98% on both validation datasets, thereby outperforming the conventional anomaly detection models. The FFDD discard dissimilar data by measuring the similarity of the existing data with the new activity data, while the FFDA discard data based on its age. This results in a faster convergence time for the FFDD since dissimilar data are discarded at each iteration. The batch size of the adaptation data also have an effect on the forgetting factor approaches as small batch size results in a regular model update but leads to a high rate of false prediction, while a large batch size results in less error rate with a slower model update.

Chapter 7

The Role of an Assistive Robot Intermediary

7.1 Introduction

Assistive robots have the potential of being utilised in a home environment for various purposes such as domestic services, companionship and monitoring. A companion robot capable of helping older adults suffering from MCI to keep track of their reminders (such as medication time), establishing communication with carers and family, as well as administering exercises to improve their cognitive abilities is developed in [68]. Assistive robots have also been utilised to promote engagement and to serve as a means of teaching, for example, in [28], a robotic platform is used to serve as a teaching tool to help people in managing their health condition, while in [173], a similar platform is used to facilitate engagement. Related research works are also carried out to promote older adults' wellbeing using a robotic platform [25, 88, 103].

The previous chapter presents an adaptive approach for anomaly detection based on forgetting factors. The adaptation data is expected to undergo a filtering process where human approval is required through a communication intermediary. This chapter explores the means of utilising an assistive robot platform as the communication intermediary for the abnormality detection system as depicted in Figure 1.1. This is in line with the research question on exploring the feasibility of realising an adaptive system for ADL anomaly detection with support for human input. Based on the workflow of the adaptive system, the activities identified as anomalous by the computational models are communicated to humans through the intermediary while the human feedback in the form of a gesture can be transmitted back to the computational model for incremental adaptation. Both vocal and non-vocal communication approaches are considered as feedback mechanisms for the intermediary. The choice of an assistive robot over screen-based interfaces is due to the robot's support for multi-modal interaction such as through voice, touch and gestures, as well as the presence of physical embodiment which facilitates interaction according to some studies [27, 164].

The structure of this chapter is as follows: Section 7.2 presents the gesture recognition model using a 2D camera while Section 7.3 present an alternate approach for the gesture recognition using a wearable sensor for communicating with the intermediary. Section 7.4 explored a means of utilising the robotic intermediary as a physical activity trainer to promote the health and wellbeing of older adults. A conclusion of the chapter is provided in Section 7.5.

7.2 Gesture Recognition with 2D Camera

This section presents the methodology for the hand gesture recognition from 2D images to be utilised on the robotic intermediary. The interpreted hand gestures corresponding to affirmations or denials will serve as an input to the anomaly detection model to confirm or deny the model's predictions. The hand gestures corresponding to a denial signify that the prediction of the computational model is incorrect, thereby prompting the model to learn the characteristic features of the activity data and vice-versa. The four gestures corresponding to affirmations and denials are shown in Figure 3.4.

7.2.1 Methodology

The detection of any of the gestures defined in Figure 3.4 during the HRI interaction represents the human response to the query administered by the

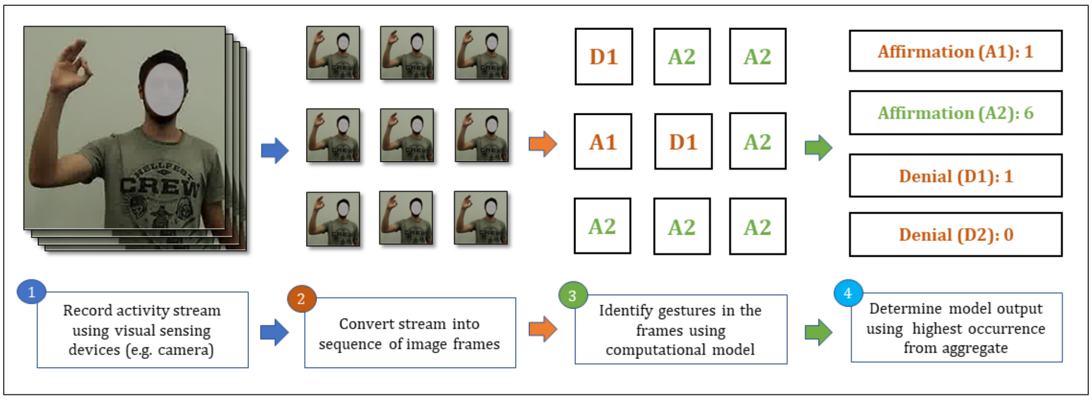


Figure 7.1: Hand gesture recognition procedure.

robotic intermediary. Since the gestures are static in nature, the most feasible approach is to identify them from image frames. However, it is an arduous task to ascertain the exact interval between the gesture feedback from the user and the time the query is administered by the intermediary. To address this, the approach in the schematic diagram in Figure 7.1 is proposed.

This involves obtaining a video stream of the performed gesture using a visual sensing device (i.e. 2D camera) for a given period (e.g. 5-15 secs long) from the moment the query is administered by the robotic intermediary. The recorded stream is then converted into frames of images at a given frame rate (e.g. 5 FPS). A computational model trained to detect the gestures of interest can then be applied to the extracted image frames. Uniquely identified gestures from the frames are aggregated. Frames containing no gesture are discarded and are assumed to be the recorded stream before and after the gestures were performed. The aggregated gestures detected from the image frames are used to predict the human's response, with the gesture having the highest number of occurrences taken as the final prediction.

A CNN based model for object detection known as YOLOv3 is utilised for detection of the hand gestures in Figure 3.4. The choice of YOLOv3 over other object detection methods such as Single Shot Detector (SDD), Region-based CNN (R-CNN), Fast R-CNN and Faster R-CNN is due to YOLOv3 superior processing speed [8, 112, 150]. A fast processing speed in the gesture recognition component is necessary since it enables the realisation of a near real-time anomaly detection system. Despite having less prediction accuracy compared to the aforementioned models, the performance difference between YOLOv3 and the other models is negligible. A more detailed description of YOLO is provided in Section 3.3 and in [139].

7.2.2 Experimentation

The dataset used for the validation of the proposed gesture recognition approach is described in Section 3.4. The data is collected using a 2D camera from five participants in a controlled lab environment performing the defined gestures in three different scenarios. The splitting of the data into training and testing sets is carried out at an individual level to achieve a generalised uniform set, i.e. the training and testing data are split for each participant separately before merging the data. A split ratio of “1:2” is adopted for the training and testing sets respectively, resulting in 33.33% of the data used for training. Overall, approximately 2,000 images are used for training with over 500 samples per gesture while the remaining 4,000 images are used for the validation. While acknowledging that the training set is usually selected to be larger than the validation set, for this experiment, the small split ratio for the training data is selected to minimise the labelling time and because the gestures across the data are expected to be similar. The labelled data is then used to train the YOLOv3 model to ascertain if the trained model can identify the gestures in the validation data for the remaining two scenarios. A summary of the collected data is tabulated in Table 3.2.

The YOLOv3 model confidence threshold reflecting the likelihood of an image containing one of the gestures is set to 0.5 across all the 3 scenarios i.e. a gesture is only predicted if the confidence value of the prediction is above 0.5. The threshold for the Non-Maximum Suppression responsible for the removal of duplicate prediction for the same class is also set to 0.5. In a situation where 2 or more unique gestures are predicted for the same image, the gesture with the highest confidence value is selected. The model is implemented on a computer equipped with an Intel Core i7 processor and NVIDIA GTX 1070 graphics card running Ubuntu 18.04. The YOLO implementation in Darknet is used. Darknet

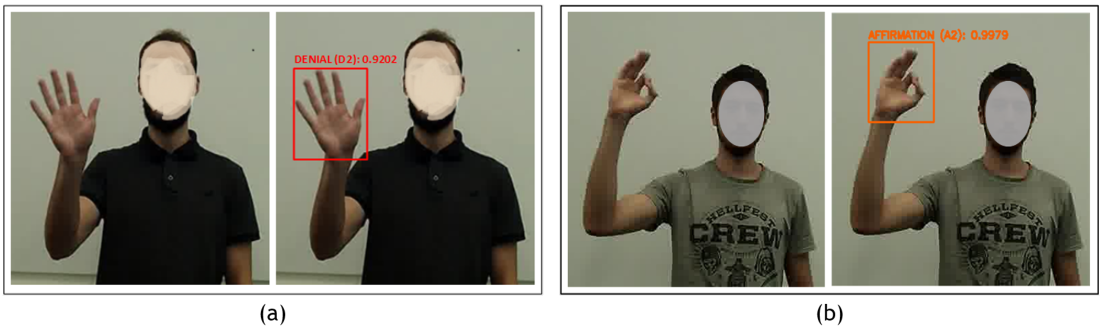


Figure 7.2: Sample output of the gesture recognition model (a) Denial (D2); (b) Affirmation (A2).

which is a C based Neural Network (NN) framework is compiled with CUDA and OpenCV enabled. Over 20,000 iterations are performed during the training and the model weight with the best result is selected for each experimental scenario.

Figure 7.2 shows a sample output of the gesture recognition model when applied to an image frame with the bounding box representing the identified gesture along with its confidence score. The confusion matrices of the prediction result for the three scenarios are used to calculate the precision, recall and F1-Score of the model in the respective scenarios as tabulated in Table 7.1. This result only shows the performance metrics of the model for the individual gestures across the three scenarios. It is worth mentioning that the training data is taken from the data collected in Scenario 1. To test for model generalisation, a K-Fold CV is performed with $K = 5$ on the Scenario 1 dataset. The overall average accuracy of the gesture recognition model across all the different gestures for the 3 scenarios is shown in Figure 7.3.

7.2.3 Discussion

The evaluation of the gesture recognition approach shows that the approach is successful in identifying the gestures as shown from the obtained results in Figure 7.3 and Table 7.1. It can be seen that the approach achieved a better performance in Scenario 1 compared to Scenario 2. This is expected since the model training is performed with data samples taken from Scenario 1. The aim

Table 7.1: Result of the gesture recognition model for the individual gestures across the three different scenarios.

Experimental Scenario		Precision	Recall	F1 Score
Scenario 1	Denial (D1)	0.99	0.91	0.95
	Denial (D2)	1.00	0.92	0.96
	Affirmation (A1)	0.96	0.84	0.89
	Affirmation (A2)	1.00	0.99	0.99
Scenario 2	Denial (D1)	0.94	0.52	0.67
	Denial (D2)	0.75	0.70	0.73
	Affirmation (A1)	0.89	0.64	0.74
	Affirmation (A2)	0.89	0.89	0.89
Scenario 3	Denial (D1)	1.00	1.00	1.00
	Denial (D2)	1.00	1.00	1.00
	Affirmation (A1)	1.00	1.00	1.00
	Affirmation (A2)	1.00	1.00	1.00

of the validation on the Scenario 2 dataset is to test the model's robustness in a different environmental setting. Observations from the confusion matrix for Scenario 2 shows that the number of unidentified gestures is high, indicating that the model is unable to predict a portion of the dataset. The portion of the Scenario 2 dataset could be incorporated in the training set for improved accuracy and reduced bias.

The model accuracy for Scenario 3 outperformed that of the other scenarios. This may be because the Scenario 3 experiment is carried out in the same environmental setting as that of Scenario 1, and as such, the high accuracy could be attributed to the similarity of the scenarios. Furthermore, the gesture recognition approach as presented in Figure 7.1 aggregates the model predictions on the image frames and selects the gesture with the highest number of occurrence as the final prediction. The likelihood of this approach making a wrong prediction is low since the number of correctly identified gestures is expected to be higher in the video stream. The accuracy of the 5-Fold CV on the Scenario 1 data is slightly lower than the baseline accuracy of the same scenario. The evaluation of the model shows the potential of utilising gesture as a non-vocal feedback mechanism on the communication intermediary.

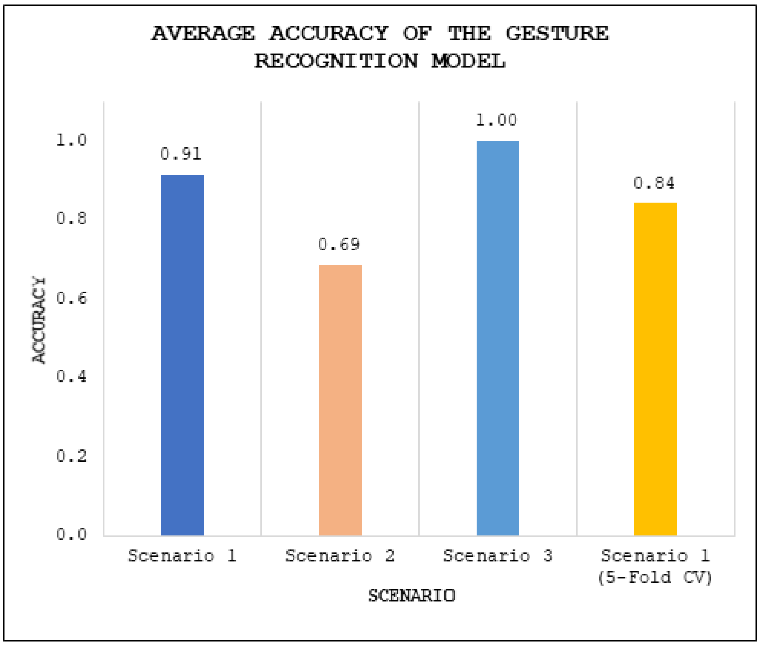


Figure 7.3: Average accuracy of the gesture recognition model.

7.3 Gesture Recognition with Wearable Sensor

An alternative approach for gesture recognition involves the utilisation of a wearable wrist-mounted tri-axial sensor that is equipped with an accelerometer. This can be useful in recognising motion-based hand gestures. The six motion-based hand gestures representing motions in different directions are shown in Figure 3.5, therefore, this section presents the methodology for their identification.

7.3.1 Methodology

As described in Section 3.4, data for the motion-based gesture recognition is collected using a device equipped with an accelerometer from five participants using a set frequency sample of 50Hz. The data undergoes a filtering process with a high-pass filter with 0.5Hz cut-off frequency and a low-pass filter with a cut-off frequency of 20Hz for removing the gravity component and noise respectively along the three different axes.

To evaluate the gesture recognition approach, five classifiers are utilised,

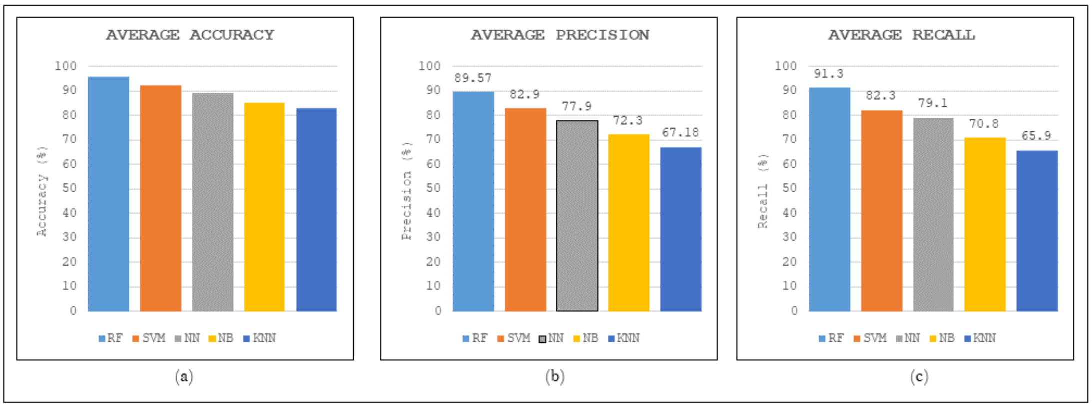


Figure 7.4: A comparative chart of results obtained for gesture recognition using a wrist-mounted accelerometer; (a) Accuracy, (b) Precision, (c) Recall.

Table 7.2: Result of hand gesture recognition with wrist-mounted accelerometer using Random Forest (RF) classifier.

Hand Gesture	Precision	Recall	Accuracy
Right	0.88	0.83	0.94
Left	0.87	0.94	0.96
Front	0.87	0.97	0.97
Back	0.92	0.88	0.95
Up	0.94	0.95	0.98
Circle	0.90	0.84	0.94

namely, Random Forest (RF), Support Vector Machine (SVM), Naive Bayes (NB), K-Nearest Neighbours (KNN) and a Multi-layer Perceptron Neural Network (MLP-NN). A 5-fold CV approach is employed. After the models training and validation, the comparison of the obtained results is shown in Figure 7.4 for the five different classifiers. It can be seen that the RF classifier achieved the best result, thereby outperforming the four remaining models. Table 7.2 shows the breakdown of the performance metrics such as the accuracy, precision and recall for each of the defined motion-based gestures obtained using the RF classifier. Overall, the RF classifier has the best result with percentage average accuracy, precision and recall of 95.9%, 89.6% and 91.3% respectively. Meanwhile, KNN has the lowest performance with an overall average of 82.9%, 67.2% and 65.9% for the accuracy, precision and recall, respectively.

7.3.2 Discussion

The recognition of hand gestures using a wearable device presents an opportunity for incorporating this sensing modality in the anomaly detection system as a means of communication. Although these are motion-based gestures that differ from the defined hand gestures recognised with a 2D camera in Figure 3.4, the results of the evaluation show that wearable devices have the potential of being utilised as a communication modality and the possibility of combining the two modalities. The fusion of the two modalities may present an avenue for the recognition of more complex and dynamic gestures.

7.4 Assistive Robot as an Exercise Trainer

In this section, the possibility of using the assistive robot intermediary as an exercise coach is explored. As mention earlier, other communication intermediaries such as a screen-based interface and push buttons could be utilised but the support for multi-modal interaction by robotic platforms present numerous opportunities. In addition to promoting physical activity and serving as a communication intermediary, the robotic platform can assist older adults in related cognitive tasks such as keeping track of meals or medication intake, serving as a task reminder and promoting social engagement.

7.4.1 Background

The benefits of physical activities cannot be over-emphasised, especially for older adults. Health-related issues such as diabetes, dementia, heart disease, high blood pressure and obesity are improved with physical activities. Moreover, as people become older, their body tissue takes longer to repair and exercise is a key counteracting factor [149]. According to a World Health Organisation (WHO) report, premature deaths attributed to lack of physical activity have risen to 3.2 million annually worldwide [154]. Given the increase in the global ageing population and the need to improve the quality of life of older adults, regular physical activities should be incorporated as part of their daily routines. Some additional benefits of physical activities include improving the body immune system, good blood circulation, and removal of skin toxins. Research shows that a significant number of older adults fail to meet their required level of exercise. For instance, in the US, over 60% of people above the age of 50 fail to meet the expected level of physical activities while in the UK, only 18% of people between the age of 65 - 74 meet the required level [154]. For older adults above the age of 75, only 7–8% are able to meet their required exercise level [154].

Engaging a personal exercise instructor may add to the demand on the societal workforce and may not be cost-effective. A possible solution is to utilise the robotic platform to engage the older adults in the needed physical activities while at the same time assessing their performance. Researchers have developed robotic platforms to promote physical activities. Bickmore and Picard [22] developed a virtual exercise advisor robot while in [134], a similar exercise robot is developed for post-stroke patients' therapy to administer arm exercises to help with the grasping and transportation of objects. A humanoid robot called "Taizo" is developed in [113]. Researchers in Singapore developed an exercise robot called Xuan to carry out training sessions for older adults [172]. Similarly, Fasola and Mataric [55] developed a robot that engages the older adults in a seated arm exercise and motivates the participants by giving praises and feedback after evaluating their performance. Findings by the authors showed that people preferred to be trained by a physical object (such as a robotic platform) than a virtual training software. Therefore, this section aims to establish the role of the robotic intermediary in administering different physical activities in addition to serving as a communication intermediary for the anomaly detection system.

7.4.2 Methodology

To develop an assistive robot that coaches, monitor and evaluate the older adults in performing physical activities and providing suitable feedback to encourage further participation, several factors are taken into consideration such as the choice of exercises, assessment criteria, and the feedback mechanisms. The following subsections provide an in-depth review of the outlined factors.

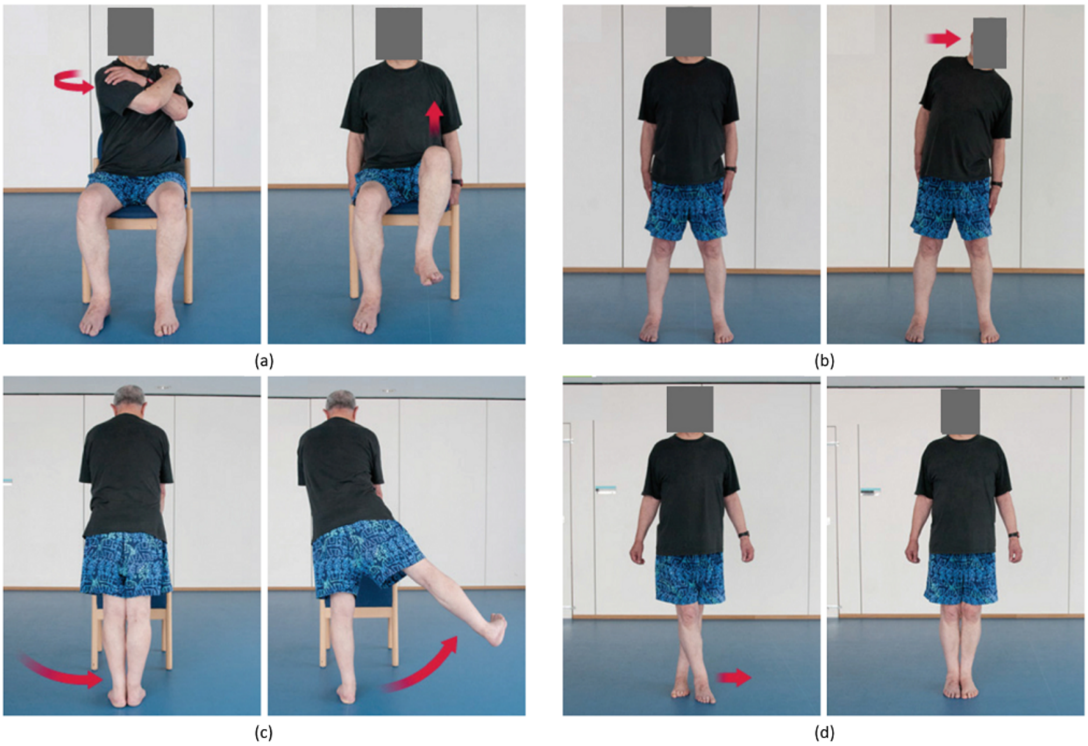


Figure 7.5: A sample of exercises recommended by the NHS for older adults: (a) sitting exercises; (b) flexibility exercises; (c) strength exercises; and (d) balance exercises.

7.4.2.1 Choice of Exercises

An exercise can be defined as any physical activity that is repetitive and structured to maintain the fitness or condition of the body which is proven to improve health and resistance to diseases. This is broadly classified into aerobic, anaerobic and flexibility exercises [50]. According to the UK Medical Officers' guideline, the benefits of physical activities to older adults include improved sleep, reduced stress, improved quality of life, maintaining a healthy weight, and an overall increase in health [65, 140]. According to the same guideline, regular exercises reduces the chances of:

- Type II Diabetes by 40%
- Cardiovascular disease by 35%

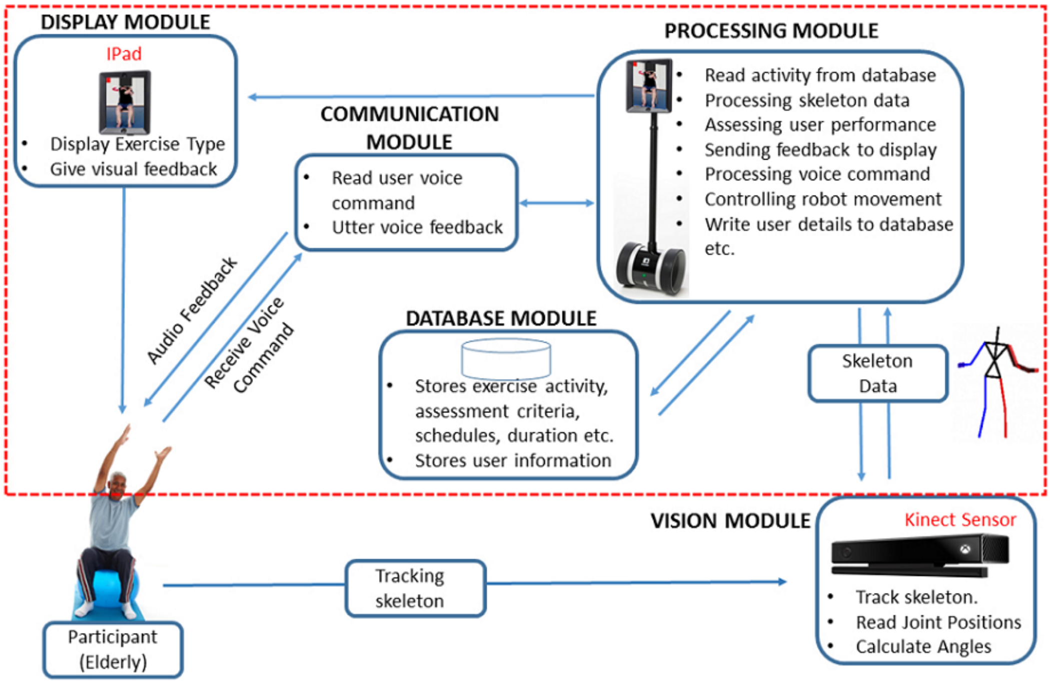


Figure 7.6: Architecture of the robot exercise system.

- Joint and back pain by 25%
- Fall, Dementia, and Depression by over 30% and
- Colon and breast cancer by 20%

According to LaVona Traywick, an assistant professor of gerontology at the University of Arkansas [160], the recommended level of exercise for older adults is 150 minutes weekly at a moderate level. This is in line with the NHS guideline and can be supplemented with a 75 minutes vigorous weekly exercise [64, 65, 140]. Figure 7.5 shows the recommended exercises according to the NHS published guidelines divided into sitting, strength, flexibility and balance exercises. These exercises are selected for implementation on the assistive robot intermediary.

7.4.2.2 System Design and Pose Matching

The schematic diagram of the system architecture is shown in Figure 7.6 consisting of five modules. The Vision module is responsible for tracking the

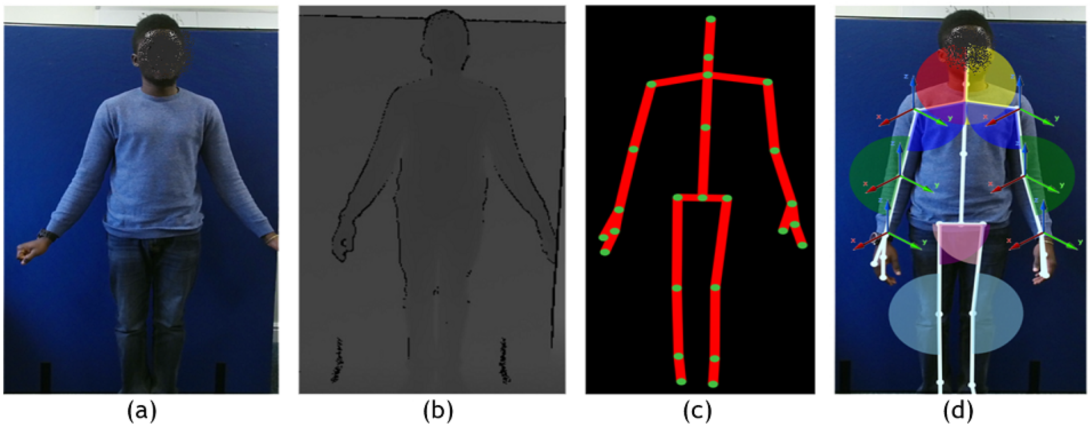


Figure 7.7: Illustration of depth image transformation: (a) colour stream; (b) depth stream; (c) skeleton (joint are shown in green dots); and (d) tracked skeleton and joints (similar joints are presented with the same colour).

user's activity to ascertain the performed activity. Microsoft Kinect sensor[1] (v2) is used for the skeleton tracking. The Display module presents visual feedback in the form of a facial expression, as well as displays the exercise poses to be performed. The display screen of the robotic platform is utilised for this task. The Communication (Sound) module provides speech recognition and audio feedback. The built-in speaker and microphone of the platform serve the purpose of receiving voice commands from the participants and providing vocal feedback. The Processing module is the core unit of the system where the data obtained from the vision module is processed. Activity recognition is performed, and the desired data for comparison is retrieved from the database. The output from the comparison process is then routed to the display and communication modules. The processing module also processes commands from the user received through the communication module. Lastly, the Database module serves as a repository for the predefined exercise data, participant information, as well as assessment information.

The skeleton data collected using the Kinect sensor is used in determining the joint coordinates and relative angles. Since the sensor provides 3D coordinates of all the tracked joints, the angles between the joints can be calculated using

[1]https://developer.microsoft.com/en-us/windows/kinect/

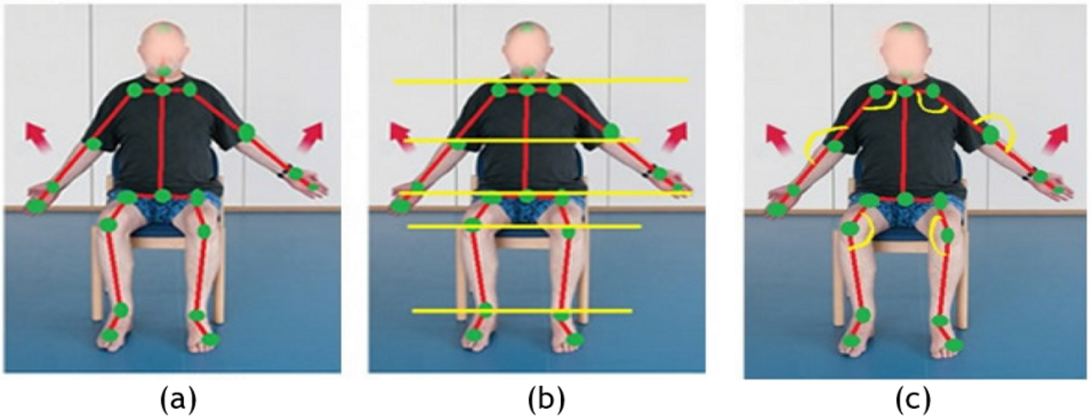


Figure 7.8: Exercise pose matching: (a) tracked skeleton; (b) comparing joint positions; and (c) comparing joint angles.

an Angular Kinematics model. Figure 7.7 shows the transformation of the depth data from the Kinect sensor. To ascertain if the participants' performed the right physical activity, pose matching is performed between the collected skeleton data and the predefined exercise poses stored in the database module. Once the joint angles and coordinates are determined from the skeletal data, the resulting values are compared to that of the stored pose information with a degree of tolerance. Figure 7.8 illustrates the pose matching approach by first tracking the skeletal and joint positions. A rule-based approach is adopted for the evaluation of the performed activities. The joints are then mapped to the required exercise pose, for instance, the wrist of the left hand must be at the same height level as that of the right hand. The joint angles are compared to the predefined angles of the given exercise, for example, the right and left shoulder angles must be approximately 120 degrees while both ankle angles must be approximately 180 degrees.

The participants are assessed and assigned a score based on the level of correctness of the performed exercises using the pose matching approach described above. The rule-based approach requires a set of conditions (rules) to be defined for each of the predefined exercises along with the tolerance (error) threshold. The feedback given to the participants is in the form of a facial expression (using emojis) and a list of predefined vocal responses signifying if the physical activities are performed correctly or not.

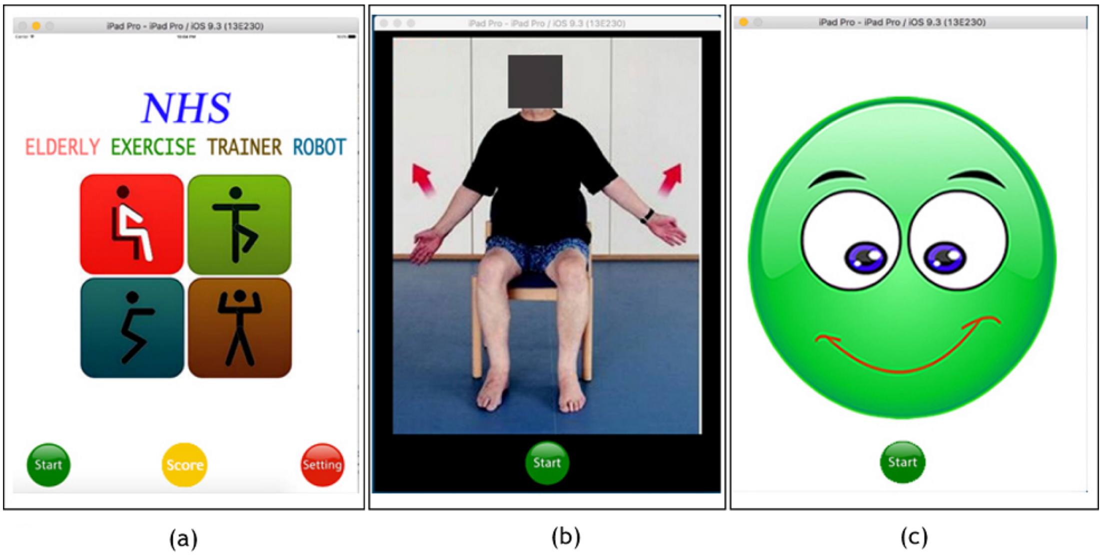


Figure 7.9: A sample of the implemented exercise application: (a) startup page; (b) exercise scenario; and (c) feedback page.

7.4.3 Experimentation

In order to evaluate the proposed system, implementation of the architecture presented in Figure 7.6 is carried using a robotic platform called Double[1]. Developed by Double Robotics, Double is a video conferencing robot equipped with two wheels and an iPad capable of self-navigation. The implemented solution relating to the display of the administered physical activities and providing feedback runs on the robotic platform while the pose processing and matching algorithm run on a PC connected to the Kinect sensor for skeleton tracking. Figure 7.9 presents the implemented application that runs on the robotic platform showing the exercise scenarios and the sample feedback in the form of a facial expression.

The system has been evaluated with a group of volunteers. Figure 7.10 shows a scenario where the volunteers interact with the implemented system by performing the instructed exercises. The assessment and feedback are provided in real-time so that the participants can have an immediate indication of whether they were doing the exercises correctly. Ethical approval had been

[1]https://www.doublerobotics.com/

(a) (b)

Figure 7.10: A demo of the participants during evaluation (a) participant 1; (b) participant 2.

obtained for the validation and all the subjects gave their informed consent for inclusion before they participated in the study. The protocol was approved by the Non-Invasive Human Ethics Committee at the School of Science and Technology, Nottingham Trent University. The evaluation of the process is conducted in two parts; Initially, some students and staff in the university campus were contacted to take part in the testing of the system. Although this group of volunteers were not in the right age group, the main purpose of the experiment is to test and evaluate the hardware and software developed. The initial evaluation scenario and the participants' feedback are presented in Appendix B. After completion of the preliminary evaluation, in the second phase of the evaluation, seventeen (17) participants in three age groups of 50s, 60s and 70s volunteered to take part in interacting with the developed system to conduct the recommended physical activities. The participants used the system for a 10-minute session, and thereafter, evaluate the effectiveness of the system using a questionnaire. Both female and male participants from a relatively wide age group were involved in this study. Figure 7.11 shows the distribution of the participants based on their age group, gender and level of physical activity.

7.4.4 Discussion

After an initial evaluation of the system with a group of students and staff from the university as well as the concrete evaluation by the external older adults

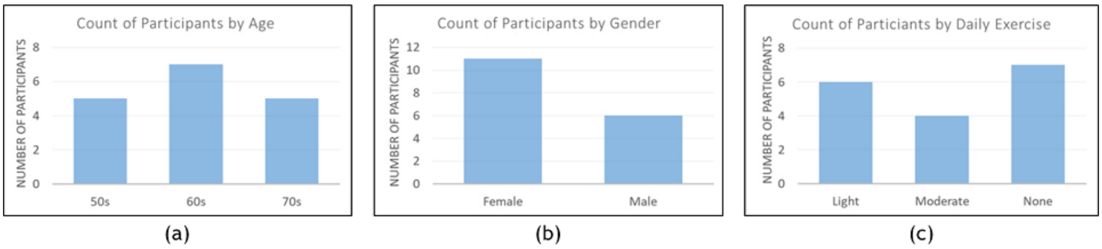


Figure 7.11: A Distribution of participants based on: (a)age; (b) gender; and (c) daily exercise level.

participants, it is established that the system is effective in administering and evaluating the physical activities. The participants also found that the system engaged them and that the feedback was appropriate and timely. Additionally, the participants indicated that they would recommend the system to others. It was observed that most of the female participants were more engaged with the system than the male participants, however, both groups found that the interaction process has encouraged them to be engaged with the exercises. A leaflet is provided to the participants with the same physical activities. Comparing the developed system with the instruction provided on a leaflet, all the participants reached a consensus that performing the physical activity with the help of the robotic platform is more fun, and it is more likely that they will take part in the daily exercises.

The only criticism received from some of the participants is with regards to the set-up configuration of the Double Robot platform and the tripod stand for mounting the Kinect sensor. Their suggestion is to integrate the robotic platform and the depth sensor to make it less cumbersome. A number of participants who had the experience of following exercise instruction from the TV found the system to be a better means of interaction. Additionally, two of the participants from the trial suggested equipping the robotic platform with the capability of recording the exercise sessions. They believed that this would be useful for therapy and evaluation such that physicians or therapists can perform an extensive review of the participant's performance and shortcomings. Some of the issues identified are attributed to the evaluation of sideways exercises. This may be due to the placement of the Kinect sensor relative to the positioning of the participants. The evaluation, in general, proves that the robotic intermediary has the potential of being utilised in administering physical activity.

7.5 Conclusion

This chapter explores the role of the robotic intermediary in the anomaly detection system. The intermediary is capable of identifying hand gestures using a 2D camera and a wearable device. The defined gestures corresponding to affirmation and denial are used to validate the prediction of the anomaly detection model. A CNN implementation for object detection known as YOLO is applied on the 2D image frames to identify the hand gestures and achieved an accuracy of 85% using a 5-fold CV approach. The classification models applied for the gesture recognition on the wearable sensor's data achieved an excellent result with a Random Forest classifier having an overall accuracy, precision and recall of 95.9%, 89.6% and 91.3% respectively. The ability of the robotic intermediary to serve as an exercise trainer is also explored. This resulted in the development of a personalised application for administering physical activities for older adults that are recommended by domain experts. An evaluation of the system by different participants shows that the system development is a success and the robotic intermediary can be used for administering physical activities.

Chapter 8

Conclusion and Future Work

8.1 Conclusion

The work in this thesis is a novel user-centric approach to anomaly detection in activities of daily living. The conducted research and findings have demonstrated the plausibility of realising a dynamic system capable of adapting to the changing human behavioural routine and detecting abnormalities that could be an early indication of health decline. This is motivated by the need to support independent living and improve the quality of life for older adults in a cost-effective manner through automation. This thesis attempted to answer the research questions identified in Chapter 1 outlined as follows:

- Can human activities be modelled in a user-centric manner to identify abnormalities in their daily activities despite the variability in the individual behavioural routine?
- Can the anomaly detection model be adaptive to the changes in human activities by incorporating human feedback through a communication intermediary?

The research questions have been addressed successfully in the respective chapters of this thesis. Chapter 4 presents an ensemble model for accurate identification of sleeping abnormalities, while Chapter 5 presents an approach for identifying the sources or reasons for the predicted anomalies. Section 6 presents a novel approach for an adaptive anomaly detection system capable of adapting to the changes in human behaviour. Lastly, the work in Chapter 7 explores the role of a robotic platform as a communication intermediary. Gesture recognition approach based on data collected with a 2D camera and a wearable device is developed as a communication modality. Additionally, an approach for utilising the robotic intermediary as an exercise trainer is explored with the implementation of an application for coaching and assessing the physical activities of older adults.

8.2 Major Contributions

This section presents a descriptive summary of the work carried out in this thesis for addressing the research questions.

8.2.1 A Novel Ensemble Anomaly Detection Model

A novel ensemble approach for novelty detection models is proposed based on the concept of consensus vote and score. The approach termed as Consensus Novelty Detection Ensemble (CNDE) is based on the principle of internal and external consensus majority vote representing the votes of the submodels, and the combined majority votes of the models, thereby allowing for the aggregation of heterogeneous models. Additionally, the approach allows for the estimation of the models' weight based on their performance on the given dataset. The outcome of the ensemble model is a normality score indicating whether the data entry is an outlier or inlier based on an estimated threshold such that a higher score indicates normal activity while a lower normality score signifies an abnormality. Experimental evaluation of the CNDE approach achieved an overall accuracy of 96% and 99% on the two validation datasets. Additionally, the ensemble approach offers the flexibility of incorporating novel data entries corresponding to the changes in human behavioural routine through the adjustment of the normality threshold.

8.2.2 A Novel Similarity Measure for Identifying Anomaly Sources

A novel similarity measure approach is proposed for identifying the sources (causes) of anomalies in human activities. This exploratory approach is based on a pairwise distance measure of the extracted features of the dataset. The pairwise distance measure can be computed using a One vs One Similarity Measure (OOSM) or One vs All Similarity Measure (OASM). In the OOSM approach, the distance is measured between corresponding features of the data entries one at a time, for example, the distance of the i^{th} feature of the first data entry is measured against the i^{th} feature of the second entry, then with that of the third entry and vice-versa across all the features. Meanwhile, in the OASM approach, the distance of the corresponding features is measured all together in one single expression, for example, the distance of the i^{th} feature of the first entry is measured against the i^{th} feature of all the other remaining entries. Three distance functions are used for the similarity measure, namely, Euclidean, Chebyshev and Canberra distance. A statistical measure is then applied to estimate the threshold of the extracted features with features exceeding a defined threshold predicted as the sources of the anomaly. The evaluation of the methodology shows that the OASM outperformed the OOSM, achieving an overall average accuracy of 81% against 68%, while the average accuracy based on the distance functions is obtained as 82%, 81%, and 61% for the Euclidean, Chebyshev and Canberra distances respectively. Additionally, a one-to-one mapping is carried out for interpreting the anomalous features into a human readable source for the ADL anomaly.

8.2.3 A Novel Adaptive Anomaly Detection Model

A novel data-driven adaptive anomaly detection approach is proposed based on the concept of the forgetting factor. The approach allows the computational models to discard old behavioural routine while incorporating new activity data representing the current behavioural routine of an individual. This is based on two forgetting factor approaches namely; Forgetting Factor based on Data Ageing (FFDA) and Forgetting Factor based on Data Dissimilarity (FFDD). The FFDA allows the model to discard old activity data based on the data age using the hypothesis that the new data entries represent the current behavioural routine while the FFDD discards the data based on a similarity measure using a distance function. The adaptive model pipeline is validated on two datasets and an overall minimum adaptation rate of 97% is obtained with a negligible error rate, thereby outperforming other computational models. Since the approach is data-driven, the model can be utilised in the absence of large training data making the model capable of learning incrementally and reducing the false prediction rate.

8.2.4 Addressing other Research Questions

To address the research question on the feasibility of realising an adaptive anomaly detection system with support for human input, modalities for incorporating human feedback are considered with an assistive robot selected as a platform of choice. Although the novelty of this approach is minor compared to the outlined contributions above, it provides the means of exploring the feasibility of incorporating human input in the modelling process and the role the robotic intermediary can play in the anomaly detection system. Hand gesture recognition models are implemented for utilisation on the robotic intermediary as a means of communication using two sensing modalities namely; a 2D camera and a wearable tri-axial sensor. The result of the gesture recognition models shows that hand gestures can be utilised as a communication modality. To demonstrate other use cases and additional benefits of utilising a robotic intermediary, a physical activity training application is developed. This allows the platform to administer and assess the older adults in performing physical activities capable of improving their overall health and wellbeing.

8.3 Future Work

The work undertaken in this thesis addressed all the identified research question successfully, but improvements to the proposed methodologies are necessary. This section outlines the possible improvements and recommendation on future research direction.

- The proposed ensemble model achieved an excellent result in the detection of anomalies. However, extensive evaluation of the methodology on longitudinal ADL data can be performed for trend analysis and to determine behavioural changes over time. For instance, behavioural variation can be identified on a dataset collected over a one year period on a monthly, quarterly or semiannual basis. The trend analysis can also provide an insight into functional health fluctuation and can be used for health evaluation by domain experts. Other means of utilising the proposed model in a non-AAL domain will also be explored.

- The proposed similarity measure for identifying anomaly sources can be improved. The datasets used for the validation consists of small data samples and fewer feature sets. In a scenario where the dataset is significantly large with a large feature set, the distance computation for the similarity matrices can be time-consuming because the similarity is measured for each outlying observation against the entire training samples across all the features. A heuristic approach for selecting fewer samples of the training data to be used for the similarity estimation can improve the efficiency of the methodology. Additionally, the current approach identifies anomaly sources in one feature of the dataset. The model can be enhanced to identify multi-feature anomaly sources. Apart from the identification of anomaly sources, the means for utilising the proposed approach for anomaly detection will be explored.

- The adaptive approach presented for anomaly detection tend to overfit as a result of the continuous training of the models whenever new adaptation data is acquired. This can affect the model performance and its ability to generalise. To address this, a tolerance threshold can be introduced to serve as a guard for preventing further model training once a performance peak is reached. The tolerance threshold can be in the form of a maximum acceptable error rate or minimum acceptable accuracy rate. Therefore, if the prediction error is within the acceptable error limit or if the accuracy is within the acceptable accuracy limit, the re-training of the model can be suspended even if new adaptation data is acquired.

- A line of research to be explored as part of the future work plan has to do with the incorporation of knowledge from domain experts in the monitoring system. The approach taken in this thesis is to build the behaviour model based on data representing the usual behavioural routine of the subjects. A knowledge-based system that incorporates domain experts' knowledge can enhance the modelling process and help in identifying appropriate normality thresholds based on relevant metadata of the older adults such as age, fragility, medical history etc.

- The gesture recognition model utilised two sensing modalities, with the camera device being more efficient in the recognition of static gestures and the wearable device more suited for the recognition of motion-based gestures. An implementation of a sensor fusion strategy for combining the two modalities can be carried out. Due to the heterogeneous nature of the sensing devices, our initial findings show that fusion at the data and decision layer is more feasible and likely to improve the accuracy of the gesture recognition approach.

- The application implemented on the robotic platform for administering physical activities is based on a rule-based approach. This approach is not efficient since each activity requires a pre-defined set of rules and an error threshold. Further work is required to incorporate a supervised model for learning the recommended physical activities and for performing pose matching in real-time. To address the participants criticism of discarding the externally connected sensor, utilising the robotic platform's in-build vision sensor can provide the needed solution.

- A significant contribution towards the field of AAL is the development and a large scale deployment of the realised system. This can promote the adoption of an automated in-home monitoring system that is cost-effective and capable of promoting ageing in place for the older adults.

Appendix A - More Evaluation of the CNDE Approach

Table 1: Accuracy of the CNDE approach after discarding poor performing models.

Dataset and Model Selection		Accuracy
SmartNTU	Without discarding a model	0.98592
	After discarding a model (RCE)	0.98571
CASAS HH111	Without discarding a model	0.95775
	After discarding a model (iForest)	0.95775

Appendix B - Initial Evaluation of the Assistive Robot

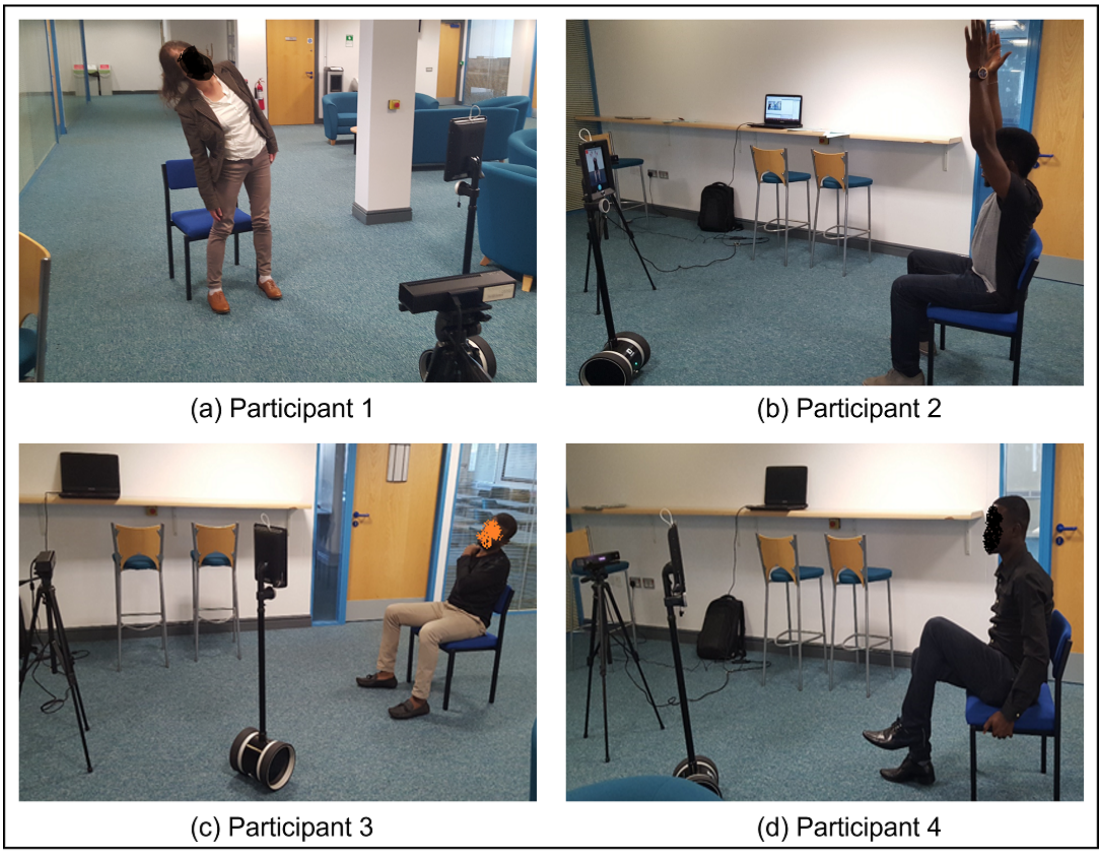
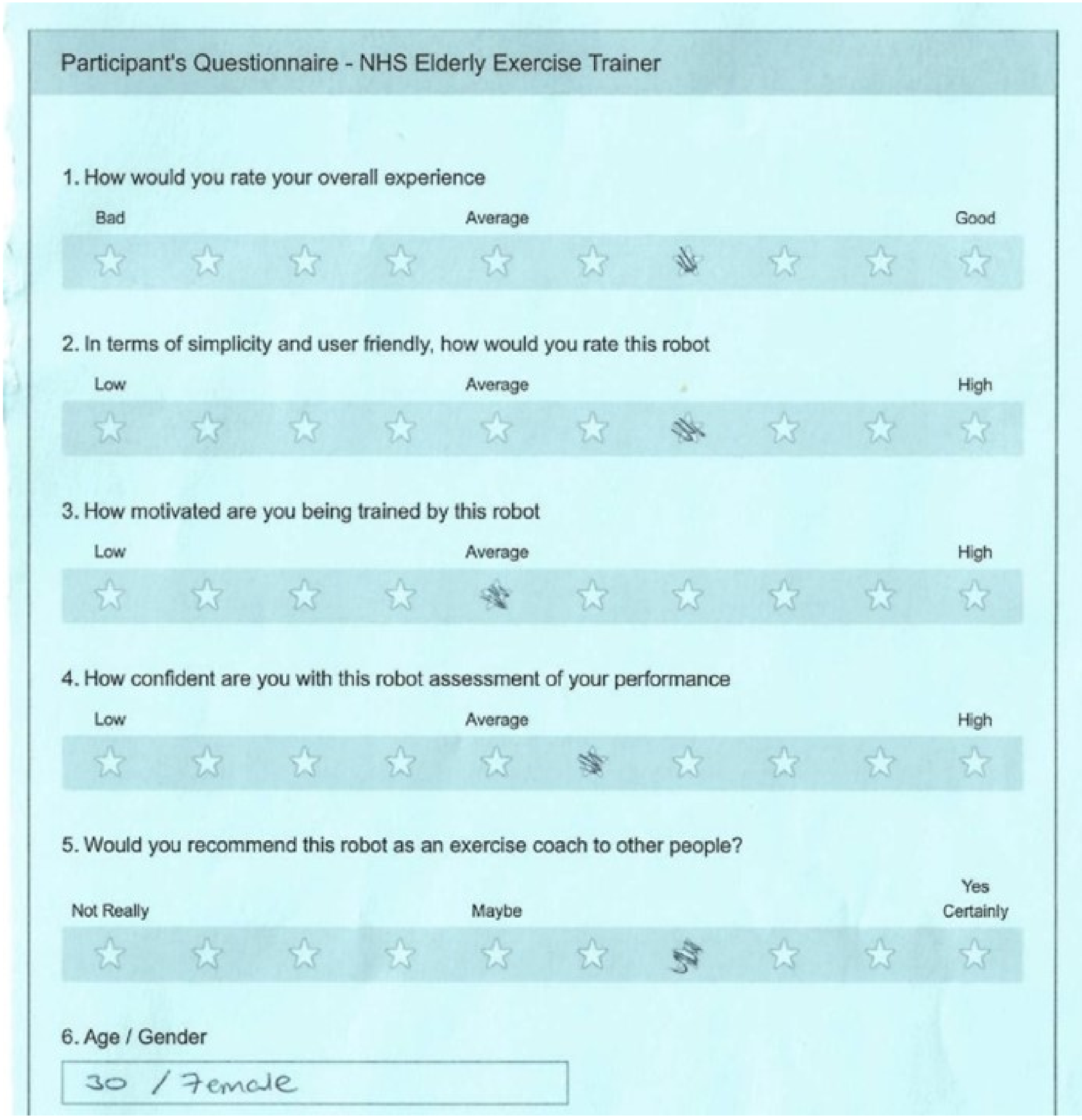


Figure 1: A scenario for the initial evaluation of the assistive robot by some of the participants.



Participant's Questionnaire - NHS Elderly Exercise Trainer

1. How would you rate your overall experience

Bad | Average | Good

2. In terms of simplicity and user friendly, how would you rate this robot

Low | Average | High

3. How motivated are you being trained by this robot

Low | Average | High

4. How confident are you with this robot assessment of your performance

Low | Average | High

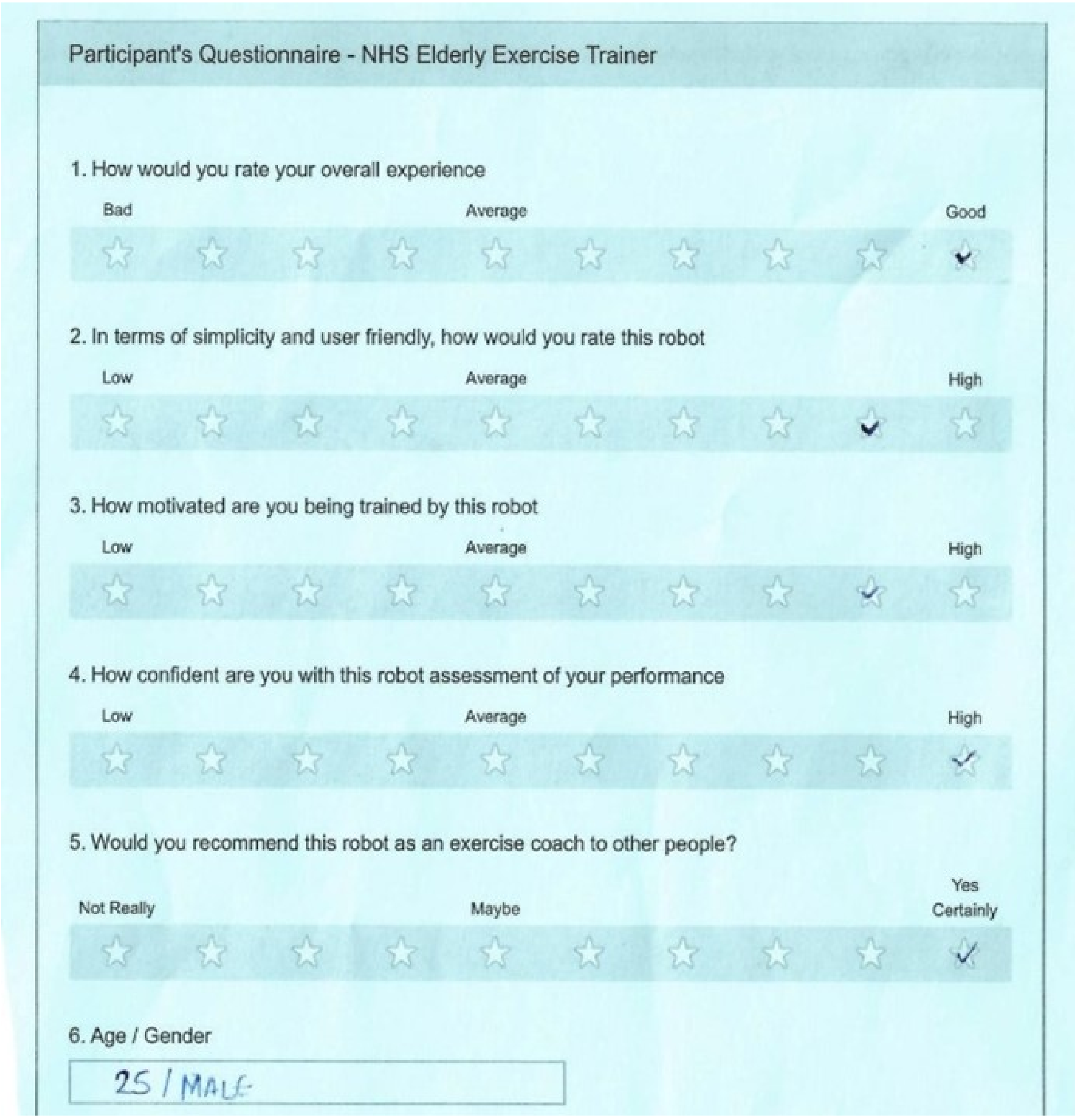
5. Would you recommend this robot as an exercise coach to other people?

Not Really | Maybe | Yes Certainly

6. Age / Gender

30 / female

Figure 2: A sample questionnaire for the assistive robot evaluation (i).



Participant's Questionnaire - NHS Elderly Exercise Trainer

1. How would you rate your overall experience

Bad — Average — Good

2. In terms of simplicity and user friendly, how would you rate this robot

Low — Average — High

3. How motivated are you being trained by this robot

Low — Average — High

4. How confident are you with this robot assessment of your performance

Low — Average — High

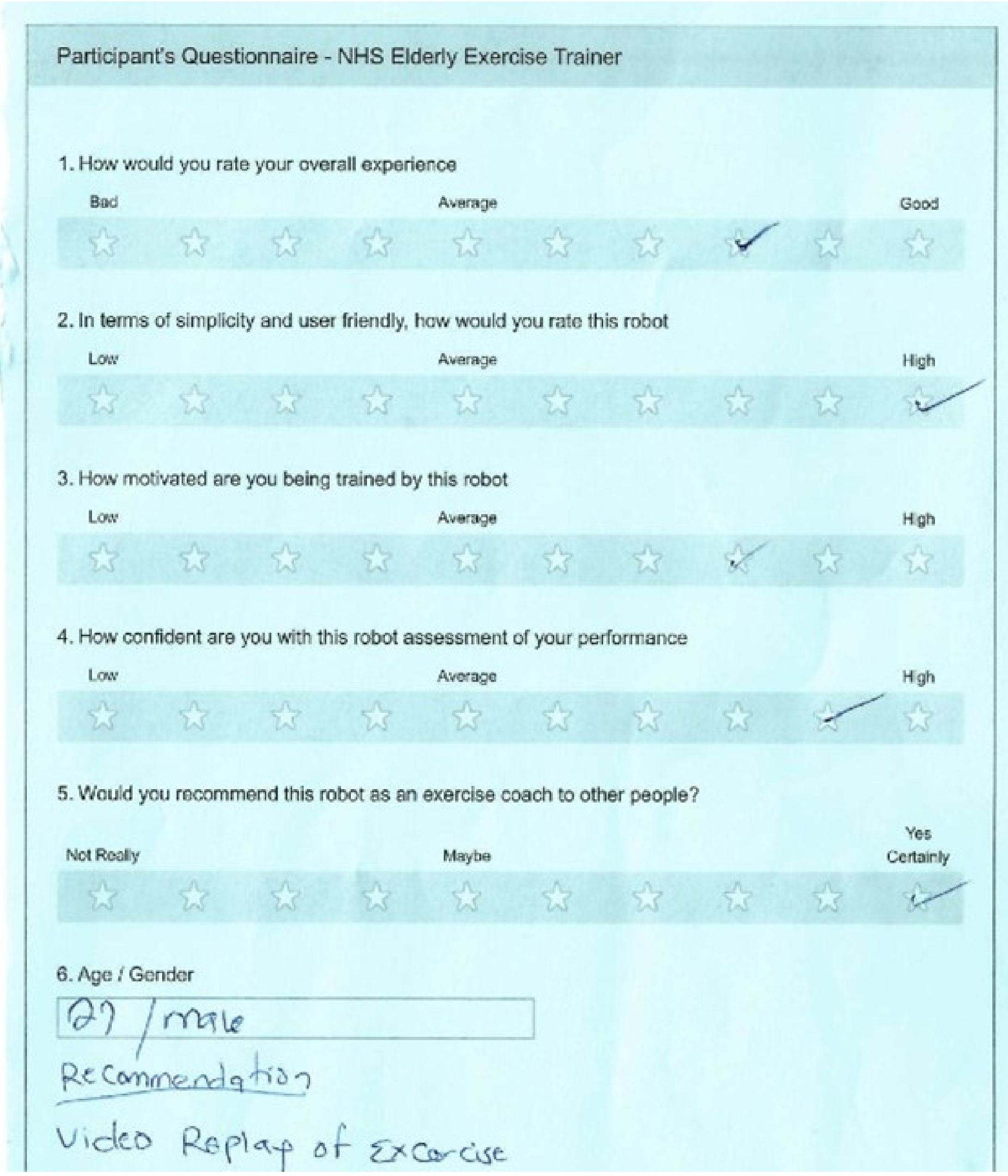
5. Would you recommend this robot as an exercise coach to other people?

Not Really — Maybe — Yes Certainly

6. Age / Gender

25 / MALE

Figure 3: A sample questionnaire for the assistive robot evaluation (ii).


Participant's Questionnaire - NHS Elderly Exercise Trainer

1. How would you rate your overall experience

Bad   Average   Good

2. In terms of simplicity and user friendly, how would you rate this robot

Low   Average   High

3. How motivated are you being trained by this robot

Low   Average   High

4. How confident are you with this robot assessment of your performance

Low   Average   High

5. Would you recommend this robot as an exercise coach to other people?

Not Really   Maybe   Yes Certainly

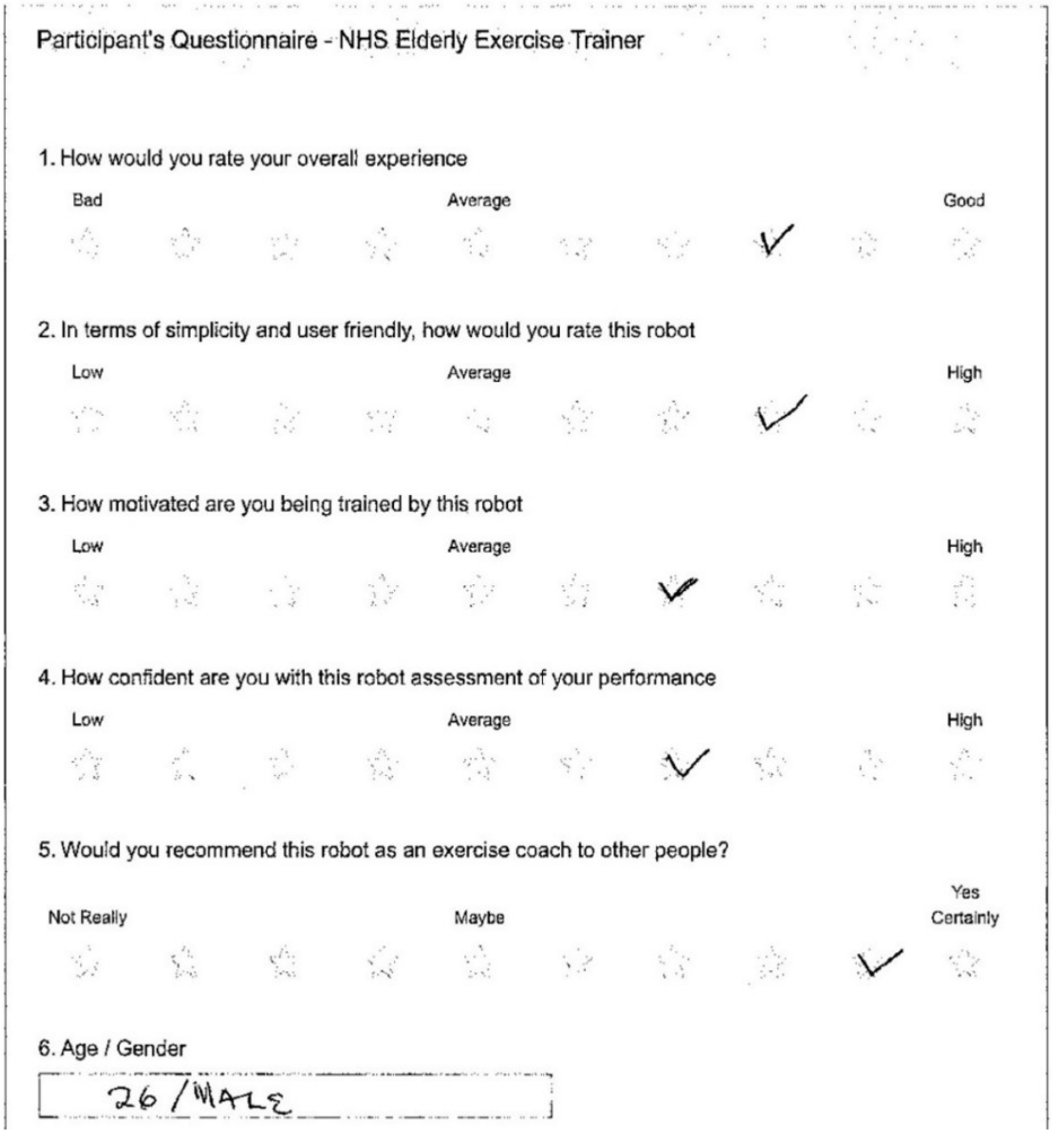
6. Age / Gender

27 / male

Recommendation

Video Replap of Excercise

Figure 4: A sample questionnaire for the assistive robot evaluation (iii).



Participant's Questionnaire - NHS Elderly Exercise Trainer

1. How would you rate your overall experience

Bad | Average | Good

2. In terms of simplicity and user friendly, how would you rate this robot

Low | Average | High

3. How motivated are you being trained by this robot

Low | Average | High

4. How confident are you with this robot assessment of your performance

Low | Average | High

5. Would you recommend this robot as an exercise coach to other people?

Not Really | Maybe | Yes Certainly

6. Age / Gender

26 / MALE

Figure 5: A sample questionnaire for the assistive robot evaluation (iv).

References

[1] G. Acampora, D. J. Cook, P. Rashidi, and A. V. Vasilakos. A survey on ambient intelligence in healthcare. *Proceedings of the IEEE*, 101(12):2470–2494, 2013. 1, 12, 13, 14, 15, 17

[2] David Ada Adama, Ahmad Lotfi, Caroline Langensiepen, Kevin Lee, and Pedro Trindade. Human activity learning for assistive robotics using a classifier ensemble. *Soft Computing*, 22(21):7027–7039, 2018. 14

[3] David Ada Adama, Ahmad Lotfi, and Robert Ranson. Adaptive segmentation and sequence learning of human activities from skeleton data. *Expert Systems with Applications*, 164:113836, 2021. 14

[4] Mohiuddin Ahmed, Abdun Naser Mahmood, and Md. Rafiqul Islam. A survey of anomaly detection techniques in financial domain. *Future Generation Computer Systems*, 55:278–288, 2016. 20

[5] M.A. Akkaş, R. Sokullu, and H. Ertürk Çetin. Healthcare and patient monitoring using iot. *Internet of Things*, 11:100173, 2020. 1

[6] Suad Albawendi, Kofi Appiah, Heather Powell, and Ahmad Lotfi. Video based fall detection with enhanced motion history images. In *Proceedings of the 9th ACM International Conference on PErvasive Technologies Related to Assistive Environments*, 2016. 14

[7] A. Alberdi, A. Weakley, M. Schmitter-Edgecombe, D. J. Cook, A. Aztiria, A. Basarab, and M. Barrenechea. Smart home-based prediction of multidomain symptoms related to alzheimer's disease. *IEEE Journal of Biomedical and Health Informatics*, 22(6):1720–1731, 2018. 4, 24

[8] Ugur Alganci, Mehmet Soydas, and Elif Sertel. Comparative research on deep learning approaches for airplane detection from very high-resolution satellite images. *Remote Sensing*, 12(3), 2020. 111

[9] I. Ali and G. Saha. A distance metric based outliers detection for robust automatic speaker recognition applications. In *2011 Annual IEEE India Conference*, pages 1–4, 2011. 21

[10] Michele Amoretti, Sergio Copelli, Folker Wientapper, Francesco Furfari, Stefano Lenzi, and Stefano Chessa. Sensor data fusion for activity monitoring in the persona ambient assisted living project. *Journal of Ambient Intelligence and Humanized Computing*, 4(1):67–84, 2013. 15

[11] D. O. Anderez, L. P. Dos Santos, A. Lotfi, and S. W. Yahaya. Accelerometer-based hand gesture recognition for human-robot interaction. In *2019 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 1402–1406, 2019. 15, 52

[12] Dario Ortega Anderez, Ahmad Lotfi, and Caroline Langensiepen. A novel crossings-based segmentation approach for gesture recognition. In *Advances in Computational Intelligence Systems*, pages 383–391, 2019. 51

[13] Dario Ortega Anderez, Ahmad Lotfi, and Amir Pourabdollah. Eating and drinking gesture spotting and recognition using a novel adaptive segmentation technique and a gesture discrepancy measure. *Expert Systems with Applications*, 140:112888, 2020. 15, 17, 51

[14] Dario Ortega Anderez, Ahmad Lotfi, and Amir Pourabdollah. A deep learning based wearable system for food and drink intake recognition. *Journal of Ambient Intelligence and Humanized Computing*, pages 1–13, 2020. 15, 17

[15] Manocha Ankush and Singh Ramandeep. An intelligent monitoring system for indoor safety of individuals suffering from autism spectrum disorder (asd). *Journal of Ambient Intelligence and Humanized Computing*, 2019, 2019. 27

[16] Ane Alberdi Aramendi, Alyssa Weakley, Asier Aztiria Goenaga, Maureen Schmitter-Edgecombe, and Diane J. Cook. Automatic assessment of functional health decline in older adults based on smart home data. *Journal of Biomedical Informatics*, 81:119 – 130, 2018. 4, 24

[17] Damla Arifoglu and Abdelhamid Bouchachia. Activity recognition and abnormal behaviour detection with recurrent neural networks. *Procedia Computer Science*, 110:86 – 93, 2017. 1, 4, 20, 23

[18] Damla Arifoglu and Abdelhamid Bouchachia. Detection of abnormal behaviour for dementia sufferers using Convolutional Neural Networks. *Artificial Intelligence in Medicine*, 94:88–95, 2019. 4, 7, 20, 23, 28, 73

[19] 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. 15, 17

[20] A. Avci, S. Bosch, M. Marin-Perianu, R. Marin-Perianu, and P. Havinga. Activity recognition using inertial sensing for healthcare, wellbeing and sports applications: A survey. In *23th International Conference on Architecture of Computing Systems 2010*, pages 1–10, 2010. 16, 17

[21] Akram Bayat, Marc Pomplun, and Duc A. Tran. A study on human activity recognition using accelerometer data from smartphones. *Procedia Computer Science*, 34:450–457, 2014. 17

[22] Timothy Bickmore and Rosalind W Picard. Future of caring machines. *Studies in health technology and informatics*, 118:132–145, 2005. 118

[23] M. Borazio, E. Berlin, N. Kücükyildiz, P. Scholl, and K. V. Laerhoven. Towards benchmarked sleep detection with wrist-worn sensing units. In *2014 IEEE International Conference on Healthcare Informatics*, pages 125–134, 2014. 1, 4, 7

[24] Serge Thomas Mickala Bourobou and Younghwan Yoo. User activity recognition in smart homes using pattern clustering applied to temporal ann algorithm. *Sensors*, 15(5):11953–11971, 2015. 16

[25] Hannah Louise Bradwell, Katie Jane Edwards, Rhona Winnington, Serge Thill, and Ray B Jones. Companion robots for older people: importance of user-centred design demonstrated through observations and focus groups comparing preferences of older people and roboticists in south west england. *BMJ Open*, 9(9), 2019. 109

[26] Markus M. Breunig, Hans-Peter Kriegel, Raymond T. Ng, and Jörg Sander. Lof: Identifying density-based local outliers. In *Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data*, page 93–104, 2000. 38, 39

[27] John-John Cabibihan, Hifza Javed, Marcelo Ang, and Sharifah Mariam Aljunied. Why robots? a survey on the roles and benefits of social robots in the therapy of children with autism. *International Journal of Social Robotics*, 5(4):593–618, 2013. 110

[28] Lola Cañamero and Matthew Lewis. Making new "new ai" friends: Designing a social robot for diabetic children from an embodied ai perspective. *International Journal of Social Robotics*, 8(4):523–537, 2016. 109

[29] Pierluigi Casale, Oriol Pujol, and Petia Radeva. Human activity recognition from accelerometer data using a wearable device. In *5th Iberian Conference on Pattern Recognition and Image Analysis*, page 289–296, 2011. 17

[30] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer. Smote: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16:321–357, 2002. 46

[31] S. Chernbumroong, A. S. Atkins, and H. Yu. Activity classification using a single wrist-worn accelerometer. In *5th International Conference on Software, Knowledge Information, Industrial Management and Applications (SKIMA) Proceedings*, pages 1–6, 2011. 17

[32] Saisakul Chernbumroong, Shuang Cang, Anthony Atkins, and Hongnian Yu. Elderly activities recognition and classification for applications in assisted living. *Expert Systems with Applications*, 40(5):1662 – 1674, 2013. 2, 3, 12, 15, 17

[33] D. J. Cook, A. S. Crandall, B. L. Thomas, and N. C. Krishnan. Casas: A smart home in a box. *Computer*, 46(7):62–69, 2013. 17, 31, 45

[34] D. J. Cook, M. Schmitter-Edgecombe, and P. Dawadi. Analyzing activity behavior and movement in a naturalistic environment using smart home techniques. *IEEE Journal of Biomedical and Health Informatics*, 19(6): 1882–1892, 2015. 45

[35] Diane J. Cook and Narayanan Krishnan. Mining the home environment. *Journal of Intelligent Information Systems*, 43(3):503–519, 2014. 1, 13, 14, 15, 17, 43

[36] Diane J. Cook, Juan C. Augusto, and Vikramaditya R. Jakkula. Ambient intelligence: Technologies, applications, and opportunities. *Pervasive and Mobile Computing*, 5(4):277–298, 2009. 12, 13

[37] A. S. Crandall, D. J. Cook, and V. R. Jakkula. Knowledge discovery in entity based smart environment resident data using temporal relation based data mining. In *7th IEEE International Conference on Data Mining Workshops*, pages 625–630, 2007. 15

[38] Jessamyn Dahmen, Brian L. Thomas, Diane J. Cook, and Xiaobo Wang. Activity learning as a foundation for security monitoring in smart homes. *Sensors*, 17(4), 2017. 27

[39] X. Dai and M. Bikdash. Distance-based outliers method for detecting disease outbreaks using social media. In *SoutheastCon 2016*, pages 1–8, 2016. 21

[40] P. N. Dawadi, D. J. Cook, and M. Schmitter-Edgecombe. Automated cognitive health assessment using smart home monitoring of complex tasks. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 43(6): 1302–1313, 2013. 25

[41] Prafulla Dawadi, Diane Cook, Carolyn Parsey, Maureen Schmitter-Edgecombe, and Miya Schneider. An approach to cognitive assessment in smart home. In *Proceedings of the 2011 Workshop on Data Mining for Medicine and Healthcare*, page 56–59, 2011. 25

[42] Prafulla Dawadi, Diane J. Cook, and Maureen Schmitter-Edgecombe. Smart home-based longitudinal functional assessment. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*, page 1217–1224, 2014. 25

[43] Prafulla Nath Dawadi, Diane Joyce Cook, and Maureen Schmitter-Edgecombe. Automated Cognitive Health Assessment from Smart Home-Based Behavior Data. *IEEE Journal of Biomedical and Health Informatics*, 20(4):1188–1194, 2016. 4, 24

[44] Francisco de Arriba-Pérez, Manuel Caeiro-Rodríguez, and Juan M. Santos-Gago. How do you sleep? using off the shelf wrist wearables to estimate sleep quality, sleepiness level, chronotype and sleep regularity indicators. *Journal of Ambient Intelligence and Humanized Computing*, 9(4):897–917, 2018. 4, 7, 15

[45] Gerges Dib, Oleksii Karpenko, Ermias Koricho, Anton Khomenko, Mahmoodul Haq, and Lalita Udpa. Ensembles of novelty detection classifiers for structural health monitoring using guided waves. *Smart Materials and Structures*, 27(1):15003, 2018. 20, 29

[46] NHS Digital. Social care spending up £556m as costs also rise, 2017. URL `https://tinyurl.com/SocialCareSpending`. [Online; accessed 01-February-2021]. 3

[47] Stephan Dreiseitl, Melanie Osl, Christian Scheibbock, and Michael Binder. Outlier Detection with One-Class SVMs: An Application to Melanoma Prognosis. In *Proceedings of the AMIA Annual Fall Symposium*, pages 172–178, 2010. 22

[48] Samih Eisa and Adriano Moreira. A behaviour monitoring system (bms) for ambient assisted living. *Sensors*, 17(9), 2017. 27

[49] Abubaker Elbayoudi, Ahmad Lotfi, and Caroline Langensiepen. The human behaviour indicator: A measure of behavioural evolution. *Expert Systems with Applications*, 118:493–505, 2019. 15

[50] The European Food Information Council (EUFIC). Types of exercise, 2008. URL http://www.eufic.org/article/en/expid/types-of-exercise/. [Online; accessed 17-June-2016]. 119

[51] L. Fahad, A. Ali, and M. Rajarajan. Long term analysis of daily activities in smart home. In *European Symposium on Artificial Neural Networks (ESANN)*, pages 419–424, 2013. 21

[52] L. G. Fahad and M. Rajarajan. Anomalies detection in smart-home activities. In *IEEE 14th International Conference on Machine Learning and Applications (ICMLA)*, pages 419–422, 2015. 21

[53] Labiba Gillani Fahad and Syed Fahad Tahir. Activity recognition and anomaly detection in smart homes. *Neurocomputing*, 423:362–372, 2021. 21

[54] Habtamu Fanta, Zhiwen Shao, and Lizhuang Ma. Sitgru: Single-tunnelled gated recurrent unit for abnormality detection. *Information Sciences*, 524: 15–32, 2020. 29

[55] J. Fasola and M. J. Mataric. Using socially assistive human–robot interaction to motivate physical exercise for older adults. *Proceedings of the IEEE*, 100(8):2512–2526, 2012. 118

[56] Christoph Feichtenhofer, Axel Pinz, and Andrew Zisserman. Convolutional two-stream network fusion for video action recognition. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1933–1941, 2016. 19

[57] The National Institute for Health Research. Multi-morbidity predicted to increase in the uk over the next 20 years, 2018. URL https://tinyurl.com/Multimorbiditypredicted. [Online; accessed 01-February-2021]. 3

[58] Office for National Statistics. Overview of the uk population: January 2021, 2021. URL https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/articles/overviewoftheukpopulation/january2021. [Online; accessed 15-August-2021]. 2

[59] Abdur Rahim Mohammad Forkan, Ibrahim Khalil, Zahir Tari, Sebti Foufou, and Abdelaziz Bouras. A context-aware approach for long-term behavioural change detection and abnormality prediction in ambient assisted living. *Pattern Recognition*, 48(3):628 – 641, 2015. 25

[60] Andrew B. Gardner, Abba M. Krieger, George Vachtsevanos, and Brian Litt. One-class novelty detection for seizure analysis from intracranial eeg. *Journal of Machine Learning Research*, 7:1025–1044, 2006. 21

[61] Samuele Gasparrini, Enea Cippitelli, Susanna Spinsante, and Ennio Gambi. A depth-based fall detection system using a kinect® sensor. *Sensors*, 14 (2):2756–2775, 2014. 43

[62] D. Gold. An examination of instrumental activities of daily living assessment in older adults and mild cognitive impairment. *Journal of Clinical and Experimental Neuropsychology*, 34:11 – 34, 2012. 28

[63] Maoguo Gong, Huimin Zeng, Yu Xie, Hao Li, and Zedong Tang. Local distinguishability aggrandizing network for human anomaly detection. *Neural Networks*, 122:364 – 373, 2020. 29

[64] GOV.UK. Exercises for older people, 2014. URL `http://www.nhs.uk/Tools/Pages/Exercises-for-older-people.aspx`. [Online; accessed 15-August-2016]. 120

[65] GOV.UK. Physical activity benefits for adults and older adults, 2015. URL `https://www.health-ni.gov.uk/publications/physical-activity-benefits-adults-and-older-adults`. [Online; accessed 01-February-2021]. 119, 120

[66] Kalpana Goyal and Jyoti Singhai. Review of background subtraction methods using gaussian mixture model for video surveillance systems. *Artificial Intelligence Review*, 50(2):241–259, 2018. 14

[67] Anupam Grewal, Maninder Kaur, Jong Hyuk Park, and Davide Mattera. A unified framework for behaviour monitoring and abnormality detection for smart home. *Wireless Communications and Mobile Computing*, 2019: 1–16, 2019. 28

[68] H. . Gross, C. Schroeter, S. Mueller, M. Volkhardt, E. Einhorn, A. Bley, T. Langner, C. Martin, and M. Merten. I'll keep an eye on you: Home robot companion for elderly people with cognitive impairment. In *2011 IEEE International Conference on Systems, Man, and Cybernetics*, pages 2481–2488, 2011. 109

[69] H. P. Gupta, H. S. Chudgar, S. Mukherjee, T. Dutta, and K. Sharma. A continuous hand gestures recognition technique for human-machine interaction using accelerometer and gyroscope sensors. *IEEE Sensors Journal*, 16(16):6425–6432, 2016. 15

[70] Karen Zita Haigh, Liana M. Kiff, and Geoffrey Ho. The independent lifestyle assistant: Lessons learned. *Assistive Technology*, 18(1):87–106, 2006. 3, 4, 31, 93

[71] Sfar Hela, Bouzeghoub Amel, and Raddaoui Badran. Early anomaly detection in smart home: A causal association rule-based approach. *Artificial Intelligence in Medicine*, 91:57 – 71, 2018. 23

[72] Netzahualcoyotl Hernandez, Jens Lundström, Jesus Favela, Ian McChesney, and Bert Arnrich. Literature review on transfer learning for human activity recognition using mobile and wearable devices with environmental technology. *SN Computer Science*, 1(66), 2020. 13

[73] E. Hoque, R. F. Dickerson, S. M. Preum, M. Hanson, A. Barth, and J. A. Stankovic. Holmes: A comprehensive anomaly detection system for daily in-home activities. In *International Conference on Distributed Computing in Sensor Systems*, pages 40–51, 2015. 4, 21, 55, 93

[74] A. Howedi, A. Lotfi, and A. Pourabdollah. Accelerometer-based human fall detection using fuzzy entropy. In *2020 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pages 1–7, 2020. 15

[75] Aadel Howedi, Ahmad Lotfi, and Amir Pourabdollah. Exploring entropy measurements to identify multi-occupancy in activities of daily living. *Entropy*, 21(4), 2019. 15

[76] Aadel Howedi, Ahmad Lotfi, and Amir Pourabdollah. An entropy-based approach for anomaly detection in activities of daily living in the presence of a visitor. *Entropy*, 22(8), 2020. 26

[77] Feng-Shuo Hsu, Tang-Chen Chang, Zi-Jun Su, Shin-Jhe Huang, and Chien-Chang Chen. Smart fall detection framework using hybridized video and ultrasonic sensors. *Micromachines*, 12(5), 2021. 43

[78] Li-Yu Hu, Min-Wei Huang, Shih-Wen Ke, and Chih-Fong Tsai. The distance function effect on k-nearest neighbor classification for medical datasets. *SpringerPlus*, 5, 2016. 79, 97

[79] Zhongxu Hu, Youmin Hu, Jie Liu, Bo Wu, Dongmin Han, and Thomas Kurfess. 3d separable convolutional neural network for dynamic hand gesture recognition. *Neurocomputing*, 318:151 – 161, 2018. 19

[80] W. A. Jabbar, T. K. Kian, R. M. Ramli, S. N. Zubir, N. S. M. Zamrizaman, M. Balfaqih, V. Shepelev, and S. Alharbi. Design and fabrication of smart home with internet of things enabled automation system. *IEEE Access*, 7: 144059–144074, 2019. 1

[81] V. Jakkula, D. J. Cook, and A. S. Crandall. Temporal pattern discovery for anomaly detection in a smart home. In *The 3rd IET International Conference on Intelligent Environments*, pages 339–345, 2007. 26

[82] V. R. Jakkula, A. S. Crandall, and D. J. Cook. Knowledge discovery in entity based smart environment resident data using temporal relation based data mining. In *Seventh IEEE International Conference on Data Mining Workshops*, pages 625–630, 2007. 16, 17

[83] Vikramaditya Jakkula and Diane J. Cook. Detecting anomalous sensor events in smart home data for enhancing the living experience. In *7th AAAI Conference on Artificial Intelligence and Smarter Living: The Conquest of Complexity*, pages 33–37, 2011. 4, 21

[84] Vikramaditya R. Jakkula and Diane J. Cook. Using temporal relations in smart environment data for activity prediction. In *24th International Conference on Machine Learning*, 2007. 16, 17

[85] Ahmad Jalal and S. Kamal. Real-time life logging via a depth silhouette-based human activity recognition system for smart home services. In *11th IEEE International Conference on Advanced Video and Signal-Based Surveillance, AVSS 2014*, pages 74–80. IEEE, 2014. 13

[86] Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. *An Introduction to Statistical Learning: With Applications in R*. Springer Publishing Company, Incorporated, 2014. 16, 22, 41, 58, 59

[87] Zaffar Haider Janjua, Daniele Riboni, and Claudio Bettini. Towards automatic induction of abnormal behavioral patterns for recognizing mild cognitive impairment. In *Proceedings of the 31st Annual ACM Symposium on Applied Computing*, page 143–148, 2016. 28

[88] Nancy S. Jecker. You've got a friend in me: sociable robots for older adults in an age of global pandemics. *Ethics and Information Technology*, 2020: 1–9, 2020. 109

[89] Hyunjun Ju, Dongha Lee, Junyoung Hwang, Junghyun Namkung, and Hwanjo Yu. Pumad: Pu metric learning for anomaly detection. *Information Sciences*, 523:167–183, 2020. 29, 30

[90] Zhaojie Ju, Dongxu Gao, Jiangtao Cao, and Honghai Liu. A novel approach to extract hand gesture feature in depth images. *Multimedia Tools and Applications*, 75(19):11929 – 11943, 2016. 19

[91] Shehroz S. Khan and Michael G. Madden. One-class classification: taxonomy of study and review of techniques. *The Knowledge Engineering Review*, 29(3):345–374, 2014. 55

[92] Shehroz S. Khan, Michelle E. Karg, Dana Kulić, and Jesse Hoey. Detecting falls with x-factor hidden markov models. *Applied Soft Computing*, 55: 168–177, 2017. 26

[93] Dongil Kim, Pilsung Kang, Sungzoon Cho, Hyoung joo Lee, and Seungyong Doh. Machine learning-based novelty detection for faulty wafer detection in semiconductor manufacturing. *Expert Systems with Applications*, 39(4): 4075 – 4083, 2012. 20

[94] Andrew Kingston, Louise Robinson, Heather Booth, Martin Knapp, Carol Jagger, and for the MODEM project. Projections of multi-morbidity in the older population in England to 2035: estimates from the Population Ageing and Care Simulation (PACSim) model. *Age and Ageing*, 47(3): 374–380, 2018. 3

[95] O. D. Lara and M. A. Labrador. A survey on human activity recognition using wearable sensors. *IEEE Communications Surveys Tutorials*, 15(3): 1192–1209, 2013. 13, 15, 17

[96] Daniel T. Larose and Chantal D. Larose. *Data Mining and Predictive Analytics*. Wiley Publishing, 2nd edition, 2015. 59

[97] 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. 12

[98] Weixin Li, Vijay Mahadevan, and Nuno Vasconcelos. Anomaly detection and localization in crowded scenes. *IEEE Trans. Pattern Anal. Mach. Intell.*, 36(1):18–32, 2013. 20, 94

[99] Yinghao Li, Qiumei Pu, Shuheng Li, Hong Zhang, Xiaofeng Wang, Haodong Yao, and Lina Zhao. Machine learning methods for research highlight prediction in biomedical effects of nanomaterial application. *Pattern Recognition Letters*, 117:111 – 118, 2019. 33

[100] F. T. Liu, K. M. Ting, and Z. Zhou. Isolation forest. In *2008 Eighth IEEE International Conference on Data Mining*, pages 413–422, 2008. 55

[101] Fei Tony Liu, Kai Ming Ting, and Zhi-Hua Zhou. Isolation-based anomaly detection. *ACM Transactions on Knowledge Discovery from Data*, 6(1), 2012. 36, 37

[102] Ahmad Lotfi, Caroline Langensiepen, Sawsan M. Mahmoud, and M. J. Akhlaghinia. Smart homes for the elderly dementia sufferers: identification and prediction of abnormal behaviour. *Journal of Ambient Intelligence and Humanized Computing*, 3(3):205–218, 2012. 3, 4, 28, 35

[103] Sylwia Łukasik, Sławomir Tobis, Sylwia Kropińska, and Aleksandra Suwalska. Role of assistive robots in the care of older people: Survey study among medical and nursing students. *Journal of Medical Internet Research*, 22(8):e18003, 2020. 109

[104] Jens Lundström, Eric Järpe, and Antanas Verikas. Detecting and exploring deviating behaviour of smart home residents. *Expert Systems with Applications*, 55:429–440, 2016. 29

[105] R. C. Luo, Y. C. Wu, and P. H. Lin. Multimodal information fusion for human-robot interaction. In *2015 IEEE 10th Jubilee International Symposium on Applied Computational Intelligence and Informatics*, pages 535–540, 2015. 18

[106] A. L'Heureux, K. Grolinger, H. F. Elyamany, and M. A. M. Capretz. Machine learning with big data: Challenges and approaches. *IEEE Access*, 5:7776–7797, 2017. 16

[107] Chao Ma, Atsushi Shimada, Hideaki Uchiyama, Hajime Nagahara, and Rin ichiro Taniguchi. Fall detection using optical level anonymous image sensing system. *Optics and Laser Technology*, 110:44–61, 2019. 43

[108] Xin Ma, Haibo Wang, Bingxia Xue, Mingang Zhou, Bing Ji, and Yibin Li. Depth-based human fall detection via shape features and improved extreme learning machine. *IEEE Journal of Biomedical and Health Informatics*, 18 (6):1915–1922, 2014. 43

[109] M. Mahmud, M. S. Kaiser, A. Hussain, and S. Vassanelli. Applications of deep learning and reinforcement learning to biological data. *IEEE Trans. Neural Netw. Learn. Syst.*, 29(6):2063–2079, 2018. 33

[110] Mufti Mahmud, M. Shamim Kaiser, M. Mostafizur Rahman, M. Arifur Rahman, Antesar Shabut, Shamim Al-Mamun, and Amir Hussain. A brain-inspired trust management model to assure security in a cloud based iot framework for neuroscience applications. *Cognitive Computation*, 10(5): 864–873, 2018. 56

[111] Mufti Mahmud, M Shamim Kaiser, T Martin McGinnity, and Amir Hussain. Deep learning in mining biological data. *Cognitive Computation*, 13(1):1–33, 2021. 33

[112] Priyanka Malhotra and Ekansh Garg. Object detection techniques: A comparison. In *7th International Conference on Smart Structures and Systems (ICSSS)*, pages 1–4, 2020. 111

[113] Y. Matsusaka, H. Fujii, T. Okano, and I. Hara. Health exercise demonstration robot taizo and effects of using voice command in robot-human collaborative demonstration. In *The 18th IEEE International Symposium on Robot and Human Interactive Communication*, pages 472–477, 2009. 118

[114] D. McIlwraith, J. Pansiot, and G. Yang. Wearable and ambient sensor fusion for the characterisation of human motion. In *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 5505–5510, 2010. 15

[115] Lei Meng, Chunyan Miao, and Cyril Leung. Towards online and personalized daily activity recognition, habit modeling, and anomaly detection for the solitary elderly through unobtrusive sensing. *Multimedia Tools and Applications*, 76(8), 2017. 29, 30

[116] Gadelhag Mohmed, Ahmad Lotfi, and Amir Pourabdollah. Human activities recognition based on neuro-fuzzy finite state machine. *Technologies*, 6(4), 2018. 15

[117] Gadelhag Mohmed, A. Lotfi, and A. Pourabdollah. Enhanced fuzzy finite state machine for human activity modelling and recognition. *Journal of Ambient Intelligence and Humanized Computing*, 11:6077–6091, 2020. 16

[118] J. Nagi, F. Ducatelle, G. A. Di Caro, D. Cireşan, U. Meier, A. Giusti, F. Nagi, J. Schmidhuber, and L. M. Gambardella. Max-pooling convolutional neural networks for vision-based hand gesture recognition. In *2011 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*, pages 342–347, 2011. 18, 19

[119] Davi Hirafuji Neiva and Cleber Zanchettin. Gesture recognition: A review focusing on sign language in a mobile context. *Expert Systems with Applications*, 103:159 – 183, 2018. 18

[120] Qin Ni, Ana Belén García Hernando, and Iván Pau 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. 3

[121] M. Novák, M. Biňas, and F. Jakab. Unobtrusive anomaly detection in presence of elderly in a smart-home environment. In *2012 ELEKTRO*, pages 341–344, 2012. 22, 35

[122] Marek Novák, Frantisek Jakab, and Luis Lain. Anomaly detection in user daily patterns in smart-home environment. *Journal of Selected Areas in Health Informatics (JSHI)*, 3:1–11, 2013. 22

[123] A. Nurhadiyatna, W. Jatmiko, B. Hardjono, A. Wibisono, I. Sina, and P. Mursanto. Background subtraction using gaussian mixture model enhanced by hole filling algorithm (gmmhf). In *IEEE International Conference on Systems, Man, and Cybernetics*, pages 4006–4011, 2013. 14

[124] Government of the United Kingdom. Chief medical officer annual report 2020, 2020. URL https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/945929/Chief_Medical_Officer_s_annual_report_2020_-_health_trends_and_variation_in_England.pdf. [Online; accessed 15-August-2021]. 3

[125] Francisco Javier Ordóñez, P. Toledo, and A. Sanchis. Sensor-based bayesian detection of anomalous living patterns in a home setting. *Personal and Ubiquitous Computing*, 19:259–270, 2014. 27

[126] Simon Ouellet and Franois Michaud. Enhanced automated body feature extraction from a 2d image using anthropomorphic measures for silhouette analysis. *Expert Systems with Applications*, 91(C):270–276, 2018. 14

[127] Koray Ozcan and Senem Velipasalar. Wearable camera- and accelerometer-based fall detection on portable devices. *IEEE Embedded Systems Letters*, 8(1):6–9, 2016. 43

[128] Chang-Beom Park and Seong-Whan Lee. Real-time 3d pointing gesture recognition for mobile robots with cascade hmm and particle filter. *Image and Vision Computing*, 29(1):51 – 63, 2011. 18, 19

[129] S.U. Park, J.H. Park, M.A. Al-masni, M.A. Al-antari, Md.Z. Uddin, and T.-S. Kim. A depth camera-based human activity recognition via deep learning recurrent neural network for health and social care services. *Procedia Computer Science*, 100:78 – 84, 2016. 14

[130] Razvan Pascanu, Tomas Mikolov, and Yoshua Bengio. On the difficulty of training recurrent neural networks. In *30th International Conference on Machine Learning*, page III–1310–III–1318, 2013. 18

[131] José Ramón Pasillas-Díaz and Sylvie Ratté. An unsupervised approach for combining scores of outlier detection techniques, based on similarity measures. *Electronic Notes in Theoretical Computer Science*, 329:61 – 77, 2016. 29, 33, 40, 55, 69

[132] Ramesh Paudel, Kimberlyn Dunn, W. Eberle, and Danielle Chaung. Cognitive health prediction on the elderly using sensor data in smart homes. In *FLAIRS Conference*, 2018. 24

[133] Hossein Pazhoumand-Dar, Leisa J. Armstrong, and Amiya Kumar Tripathy. Detecting deviations from activities of daily living routines using kinect depth maps and power consumption data. *Journal of Ambient Intelligence and Humanized Computing*, 11(4), 2020. 27

[134] Orsolya Peter, Ibolya Tavaszi, Andras Toth, and Gabor Fazekas. Exercising daily living activities in robot-mediated therapy. *Journal of physical therapy science*, 29(5):854–858, 2017. 118

[135] Marco A.F. Pimentel, David A. Clifton, Lei Clifton, and Lionel Tarassenko. A review of novelty detection. *Signal Processing*, 99:215 – 249, 2014. 20, 21, 33, 34, 35, 73

[136] S. Poh, Y. Tan, X. Guo, S. Cheong, C. Ooi, and W. Tan. Lstm and hmm comparison for home activity anomaly detection. In *IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, pages 1564–1568, 2019. 26

[137] S. Ramapatruni, S. N. Narayanan, S. Mittal, A. Joshi, and K. Joshi. Anomaly detection models for smart home security. In *IEEE Intl Conference on High Performance and Smart Computing*, pages 19–24, 2019. 26

[138] P. Rashidi and A. Mihailidis. A survey on ambient-assisted living tools for older adults. *IEEE Journal of Biomedical and Health Informatics*, 17(3): 579–590, 2013. 1, 2, 3, 13, 31

[139] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi. You only look once: Unified, real-time object detection. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 779–788, 2016. xv, 42, 112

[140] Hamish Reid and Charlie Foster. Infographic. physical activity benefits for adults and older adults. *British Journal of Sports Medicine*, 51(19): 1441–1442, 2017. 119, 120

[141] Manassés Ribeiro, André Eugênio Lazzaretti, and Heitor Silvério Lopes. A study of deep convolutional auto-encoders for anomaly detection in videos. *Pattern Recognition Letters*, 105:13–22, 2018. 35

[142] D. Riboni, C. Bettini, G. Civitarese, Z. H. Janjua, and R. Helaoui. Fine-grained recognition of abnormal behaviors for early detection of mild cognitive impairment. In *2015 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pages 149–154, 2015. 28

[143] Daniele Riboni, Claudio Bettini, Gabriele Civitarese, Zaffar Haider Janjua, and Rim Helaoui. Smartfaber: Recognizing fine-grained abnormal behaviors for early detection of mild cognitive impairment. *Artificial Intelligence in Medicine*, 67:57–74, 2016. 28

[144] Saskia Robben, Margriet Pol, and Ben Kröse. Longitudinal ambient sensor monitoring for functional health assessments: A case study. In

Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication, page 1209–1216, 2014. 25

[145] Peter J. Rousseeuw and Mia Hubert. Anomaly detection by robust statistics. *WIREs Data Mining and Knowledge Discovery*, 8(2):e1236, 2018. 39

[146] Bernhard Schölkopf, John C. Platt, John C. Shawe-Taylor, Alex J. Smola, and Robert C. Williamson. Estimating the support of a high-dimensional distribution. *Neural Computation*, 13(7):1443–1471, 2001. 35, 36

[147] J. H. Shin, B. Lee, and K. Suk Park. Detection of abnormal living patterns for elderly living alone using support vector data description. *IEEE Transactions on Information Technology in Biomedicine*, 15(3):438–448, 2011. 22

[148] D. G. Shreyas, S. Raksha, and B. G. Prasad. Implementation of an anomalous human activity recognition system. *SN Computer Science*, 1 (3), 2020. 1, 28

[149] Marlo Sollitto. How seniors can benefit from adopting an exercise regimen, 2017. URL `https://www.agingcare.com/Articles/Exercise-benefits-for-the-Elderly-95383.htm`. [Online; accessed 01-February-2021]. 117

[150] Shrey Srivastava, Amit Vishvas Divekar, Chandu Anilkumar, Ishika Naik, Ved Kulkarni, and Pattabiraman V. Comparative analysis of deep learning image detection algorithms. *Journal of Big Data*, 8(66), 2020. 111

[151] G. A. Susto, A. Beghi, and S. McLoone. Anomaly detection through on-line isolation forest: An application to plasma etching. In *28th Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC)*, pages 89–94, 2017. 20

[152] Zeeshan Syed, Mohammed Saeed, and Ilan Rubinfeld. Identifying high-risk patients without labeled training data: Anomaly detection methodologies to predict adverse outcomes. In *AMIA Annual Symposium proceedings*, pages 772–776, 2010. 20, 22

[153] Mark Ryan M. Talabis, Robert McPherson, I. Miyamoto, Jason L. Martin, and D. Kaye. Analytics defined. In *Information Security Analytics*, pages 1–12. Syngress, 2015. 33

[154] Denise Taylor. Physical activity is medicine for older adults. *Postgraduate Medical Journal*, 90(1059):26–32, 2014. 117, 118

[155] H. Thapliyal, R. Kumar Nath, and S. P. Mohanty. Smart home environment for mild cognitive impairment population: Solutions to improve care and quality of life. *IEEE Consumer Electronics Magazine*, 7(1):68–76, 2018. 1, 12, 13

[156] Andreas Theissler. Multi-class Novelty Detection in Diagnostic Trouble Codes from Repair Shops. In *IEEE 15th International Conference on Industrial Informatics (INDIN)*, pages 750–763, 2017. 22

[157] Nick T. Thomopoulos. *Statistical Distributions*. Springer International Publishing, 2017. ISBN 978-3-319-65112-5. 82

[158] K. Tonchev, P. Koleva, A. Manolova, G. Tsenov, and V. Poulkov. Non-intrusive sleep analyzer for real time detection of sleep anomalies. In *39th International Conference on Telecommunications and Signal Processing (TSP)*, pages 400–404, 2016. 26

[159] Yu Tong, Rong Chen, and Jian Gao. Hidden state conditional random field for abnormal activity recognition in smart homes. *Entropy*, 17(3): 1358–1378, 2015. 26, 27

[160] LaVona S. Traywick. Exercise recommendations for older adults, 2016. URL `http://www.todaysgeriatricmedicine.com/news/ex_092210_03.shtml`. [Online; accessed 01-February-2021]. 120

[161] T. v. Kasteren and B. Krose. Bayesian activity recognition in residence for elders. In *3rd IET International Conference on Intelligent Environments*, pages 209–212, 2007. 15, 16

[162] A. M. Vartouni, S. S. Kashi, and M. Teshnehlab. An anomaly detection method to detect web attacks using stacked auto-encoder. In *6th Iranian*

Joint Congress on Fuzzy and Intelligent Systems (CFIS), pages 131–134, 2018. 20

[163] Michalis Vrigkas, Christophoros Nikou, and Ioannis A. Kakadiaris. A review of human activity recognition methods. *Frontiers in Robotics and AI*, 2:28, 2015. 13

[164] J. Wainer, D. J. Feil-seifer, D. A. Shell, and M. J. Mataric. The role of physical embodiment in human-robot interaction. In *15th IEEE International Symposium on Robot and Human Interactive Communication*, pages 117–122, 2006. 110

[165] Tian Wang, Zichen Miao, Yuxin Chen, Yi Zhou, Guangcun Shan, and Hichem Snoussi. Aed-net: An abnormal event detection network. *Engineering*, 5(5):930–939, 2019. 20, 94

[166] Xiaoxia Wang, Peter Tiňo, Mark A. Fardal, Somak Raychaudhury, and Arif Babul. Fast parzen window density estimator. In *Proceedings of the 2009 International Joint Conference on Neural Networks*, page 410–417, 2009. 21

[167] Xueyi Wang, Joshua Ellul, and George Azzopardi. Elderly fall detection systems: A literature survey. *Frontiers in Robotics and AI*, 7:71, 2020. 43

[168] Zhelong Wang, Cong Zhao, and Sen Qiu. A system of human vital signs monitoring and activity recognition based on body sensor network. *Sens. Rev.*, 34(1):42–50, 2014. 17

[169] Zhihua Wang, Zhaochu Yang, and Tao Dong. A review of wearable technologies for elderly care that can accurately track indoor position, recognize physical activities and monitor vital signs in real time. *Sensors*, 17(2), 2017. 15, 17

[170] Mark Weiser. The computer for the 21¡sup¿st¡/sup¿ century. *SIGMOBILE Mob. Comput. Commun. Rev.*, 3(3):3–11, 1999. 1

[171] D. J. Weller-Fahy, B. J. Borghetti, and A. A. Sodemann. A survey of distance and similarity measures used within network intrusion anomaly

detection. *IEEE Communications Surveys Tutorials*, 17(1):70–91, 2015. 20, 79, 97

[172] Tessa Wong. Singapore's robot exercise coach for the elderly, 2015. URL https://www.bbc.co.uk/news/av/world-asia-35149344. [Online; accessed 01-February-2021]. 118

[173] Luke Wood, Ben Robins, Gabriella Lakatos, Dag Sverre Syrdal, Abolfazl Zaraki, and Kerstin Dautenhahn. Utilising humanoid robots to assist children with autism learn about visual perspective taking. In *1st UK-RAS Conference on Robotics and Autonomous Systems*, 12 2017. 109

[174] L. D. Xu, W. He, and S. Li. Internet of things in industries: A survey. *IEEE Transactions on Industrial Informatics*, 10(4), 2014. 1

[175] Dit-Yan Yeung and C. Chow. Parzen-window network intrusion detectors. In *Object recognition supported by user interaction for service robots*, pages 385–388, 2002. 21

[176] X. Yingxin, L. Jinghua, W. Lichun, and K. Dehui. A robust hand gesture recognition method via convolutional neural network. In *6th International Conference on Digital Home*, pages 64–67, 2016. 18, 19

[177] Miao Yu, Adel Rhuma, Syed Mohsen Naqvi, Liang Wang, and Jonathon Chambers. A posture recognition-based fall detection system for monitoring an elderly person in a smart home environment. *IEEE Transactions on Information Technology in Biomedicine*, 16(6):1274–1286, 2012. 43

[178] Dorsaf Zekri, Thierry Delot, Mikael Desertot, Sylvain Lecomte, and Marie Thilliez. Using learning techniques to observe elderly's behavior changes over time in smart home. In *The Impact of Digital Technologies on Public Health in Developed and Developing Countries*, pages 129–141, 2020. 28

[179] Dorsaf Zekri, Thierry Delot, Marie Thilliez, Sylvain Lecomte, and Mikael Desertot. A framework for detecting and analyzing behavior changes of elderly people over time using learning techniques. *Sensors*, 20(24), 2020. 28

[180] Meriem Zerkouk and Belkacem Chikhaoui. Spatio-temporal abnormal behavior prediction in elderly persons using deep learning models. *Sensors*, 20(8), 2020. 23

[181] Tong Zhang, Huifeng Lin, Zhaojie Ju, and Chenguang Yang. Hand gesture recognition in complex background based on convolutional pose machine and fuzzy gaussian mixture models. *International Journal of Fuzzy Systems*, 22(4):1330 – 1341, 2020. 18

[182] Zhenyu Chen, M. Lin, Fanglin Chen, N. D. Lane, G. Cardone, Rui Wang, Tianxing Li, Yiqiang Chen, T. Choudhury, and A. T. Campbell. Unobtrusive sleep monitoring using smartphones. In *7th International Conference on Pervasive Computing Technologies for Healthcare and Workshops*, pages 145–152, 2013. 4, 7