TriSense: RFID, Radar, and USRP-Based Hybrid Sensing System for Enhanced Sensing and Monitoring

Muhammad Zakir Khan

Submitted in fulfilment of the requirements for the
Degree of Doctor of Philosophy

School of Engineering
College of Science and Engineering
University of Glasgow

March 2024

Abstract

This thesis presents a comprehensive approach to contactless human activity recognition (HAR) using the capabilities of three distinct technologies: radio frequency identification (RFID), Radar, and universal software-defined radio peripheral (USRP) for capturing and processing Wi-Fi-based signals. These technologies are then fused to enhance smart healthcare systems. The study initially utilises USRP devices to analyse Wi-Fi channel state information (CSI), choosing this over received signal strength for more accurate activity recognition. It employs a combination of machine learning and a hybrid of deep learning algorithms, such as the super learner and LSTM-CNN, for precise activity localisation. Subsequently, the study progresses to incorporate a transparent RFID tag wall (*TRT-Wall*) that employs a passive ultra-high frequency (UHF) RFID tag array. This RFID system has proven highly accurate in distinguishing between various activities, including sitting, standing, leaning, falling, and walking in two directions. Its effectiveness and non-intrusiveness make it particularly suited for elderly care, achieved using a modified version of the *Transformer* model without the use of a decoder. Furthermore, a significant advancement within this study is the creation of a novel fusion (*RFiDARFusion*) system, which combines RFID and Radar technologies. This system employs a long short-term memory networks variational autoencoder (*LSTM-VAE*) fusion model, utilising RFID amplitude and Radar RSSI data. This fusion approach significantly improves accuracy in challenging scenarios, such as those involving long-range and non-line-of-sight conditions. The *RFiDARFusion* system notably improves the detection of complex activities, highlighting its potential to reduce healthcare costs and enhance the quality of life for elderly patients in assisted living facilities. Overall, this thesis highlights the significant potential of radio frequency technologies with artificial intelligence, along with their combined application, to develop robust, privacy-conscious, and cost-effective solutions for healthcare and assisted living monitoring systems.

Contents

List of Tables

List of Figures

List of Abbreviations

AAL	Ambient Assisted Living
ADLs	Activities of Daily Living
AI	Artificial Intelligence
ANN	Artificial Neural Network
BLE	Bluetooth Low Energy
CBID	Customer Behavior Identification
CNN	Convolutional Neural Network
CSI	Channel State Information
CTGAN	Conditional Tabular Generative Adversarial Network
CV	Computer Vision
DAE	Deep Auto-Encoder
DDAE	Deep Denoising Auto-Encoder
DL	Deep Learning
DNN	Deep Neural Network
DTW	Dynamic Time Warping
DWT	Discrete Wavelet Transform
EEG	Electroencephalogram
EKF	Extended Kalman Filter
FFT	Fast Fourier Transformer
FM	Frequency Modulation

FMCW Frequency-Modulated Continuous-Wave

FN False Negative

FP False Positive

GNSS Global Navigation Satellite System

GPU Graphic Processing Unit

GRU Gated Recurrent Unit

GSM Global System for Mobile Communications

HAR Human Activity Recognition

HARI Human Activity Recognition Indoor

HMM Hidden Markov Model

IMU Inertial Measurement Unit

IoT Internet of Things

IR Infrared

ISM Industrial, Scientific and Medical

KF Kalman Filter

KNN K-Nearest Neighbors

LDPL Log-Normal Distance Path Loss

LoS Line of Sight

LSTM Long Short-Term Memory

MLP Multi-Layer Perceptron

ML Machine Learning

MIMO Multiple Input Multiple Output

MMP Modified Matrix Pencil

MNL Mobile Node-Based Localisation

MUSIC Multiple Signal Classification

NAN	Not A Number
NHS	National Health Service
NIC	Network Interface Card
NLoS	Non-Line of Sight
NN	Neural Network
OFDM	Orthogonal Frequency Division Multiplexing
PCA	Principal Component Analysis
PC	Personal Computer
PF	Particle Filter
ReLU	Rectified Linear Unit
RF	Radio Frequency
RFID	Radio Frequency Identification
RGB	Red Green and Blue
RNN	Recurrent Neural Network
RTI	Radio Tomographic Imaging
Rx	Receiver
SDRs	Software-Defined Radios
sEMG	surface Electromyography
SL	Super Learner
STFT	Short-Time Fourier Transform
SVM	Support Vector Machine
TN	True Negative
TP	True Positive
ToF	Time of Flight
Tx	Transmitter

UHF Ultra High Frequency

UKF Unscented Kalman Filter

UN United Nations

USRP Universal Software Radio Peripheral

UWB Ultra Wide Band

VAE Variational Autoencoder

VLC Visible Light Communication

WiFi Wireless Fidelity

WHO World Health Organization

vRTI Variance-based RTI

List of Publications

In this section, I showcase a list of research papers that I have authored and co-authored throughout my Ph.D. journey, starting from December 2019. This list details about the authors, titles, publishing entities, and current status of each paper.

Journal:

1. **Khan, M. Z.,** Taha, A., Taylor, W., Imran, M. A., & Abbasi, Q. H. (2022). Non-invasive localization using software-defined radios. IEEE Sensors Journal, 22(9), 9018-9026. (Published)

2. **Khan, M. Z.,** Farooq, M., Taha, A., Qayyum, A., Alqahtani, F., Al Hassan, A. N., ... & Abbasi, Q. H. (2023). Indoor localization technologies for activity-assisted living: Opportunities, challenges, and future directions. Advances in Computers. (Published)

3. **Khan, M. Z.,** Usman, M., Ahmad, J., Ur Rahman, M. M., Abbas, H., Imran, M., & Abbasi, Q. H. (2024). Tag-Free Indoor Fall Detection by Fusion Using Transformer Network Encoders. *Journal: Nature Scientific Report* (Under Review)

4. **Khan, M. Z.,** Usman, M., Tahir, A., Qayyum, A., Ahmad, J., Abbas, H., Imran, M., & Abbasi, Q. H. (2024). TRT-Wall: Transparent RFID Tag Wall Enabled by Artificial Intelligence for Assisted Living. *Journal: Nature Scientific Report* (Under Review)

5. **Khan, M. Z.,** Ahmad, J., Abbas, H., Imran, M., & Abbasi, Q. H. (2024). RFiDAR: Augmenting Human Activity Recognition through Contactless Fusion of RFID and Radar Data. *Journal: Information Fusion* (Under Review)

6. **Khan, M. Z.,** Bilal, M., Abbas, H., Imran, M., & Abbasi, Q. H. (2024). Empowering the Blind: Contactless Activity Recognition with Commodity SDR and UHF RFID. *Journal: Computers, Materials & Continua* (Under Review)

Other Journals:

1. Taylor, W., **Khan, M. Z.,** Tahir, A., Taha, A., Abbasi, Q. H., & Imran, M. A. (2022). An Implementation of Real-Time Activity Sensing Using Wi-Fi: Identifying Optimal

Machine-Learning Techniques for Performance Evaluation. IEEE Sensors Journal, 22(21), 21127-21134. (Published)

2. Usman, M., Rains, J., Cui, T. J., **Khan, M. Z.,** Kazim, J. U. R., Imran, M. A., & Abbasi, Q. H. (2022). Intelligent wireless walls for contactless in-home monitoring. Light: Science & Applications, 11(1), 212. (Published)

3. Saeed, U., Shah, S. A., **Khan, M. Z.,** Alotaibi, A. A., Althobaiti, T., Ramzan, N., & Abbasi, Q. H. (2022). Intelligent reflecting surface-based non-LOS human activity recognition for next-generation 6G-enabled healthcare system. Sensors, 22(19), 7175. (Published)

Conference:

1. **Khan, M. Z.,** Bilal, M., Luke, A. M., Arshad, K., Assaleh, K., Shah. S. T., Abbas, H., Imran, M. A., & Abbasi, Q. H. (14-19 July 2024 • Florence, Italy). RFiDARFusion: Enhancing Contactless Activity Monitoring with Radar and RFID Fusion. In 2024 IEEE International Symposium on Antennas and Propagation and ITNC-USNC-URSI Radio Science Meeting.

2. **Khan, M. Z.,** Qayyum, A., Arshad, K., Assaleh, K., Abbas, H., Imran, M. A., & Abbasi, Q. H. (2023, July). Contactless Fall Detection Using RFID Wall and AI. In 2023 IEEE International Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting (USNC-URSI) (pp. 1491-1492). IEEE.

3. **Khan, M. Z.,** Farooq, M., Taha, A., Qayyum, A., Alqahtani, F., Al Hassan, A. N., ... & Abbasi, Q. H. (2023). Indoor localization technologies for activity-assisted living: Opportunities, challenges, and future directions. Advances in Computers.

4. **Khan, M. Z.,** Taha, A., Taylor, W., Alomainy, A., Imran, M. A., & Abbasi, Q. H. (2022, May). Indoor localization using software defined radio: A non-invasive approach. In 2022 3rd URSI Atlantic and Asia Pacific Radio Science Meeting (AT-AP-RASC) (pp. 1-4). IEEE.

5. **Khan, M. Z.,** Taha, A., Farooq, M., Shawky, M. A., Imran, M., & Abbasi, Q. H. (2022, July). Comparative Analysis of Artificial Intelligence on Contactless Human Activity localization. In 2022 International Telecommunications Conference (ITC-Egypt) (pp. 1-3). IEEE.

6. **Khan, M. Z.,** Farooq, M., Taha, A., Ramazan, N., Imran, M. A., & Abbasi, Q. H. (2022, October). Localization using wireless sensing for future healthcare. In 2022 29th IEEE International Conference on Electronics, Circuits and Systems (ICECS) (pp. 1-4). IEEE.

7. **Khan, M. Z.,** Ahmad, J., Boulila, W., Broadbent, M., Shah, S. A., Koubaa, A., & Abbasi, Q. H. (2023). Contactless Human Activity Recognition using Deep Learning with Flexible and Scalable Software Define Radio. arXiv preprint arXiv:2304.09756.

Dataset:

- Usman, M., Rains, J., Cui, T. J., Khan, M. Z., Kazim, J. U. R., Imran, M. A., & Abbasi, Q. H. (2022). Intelligent wireless walls for contactless in-home monitoring. Light: Science & Applications, 11(1), 212. [Data Collection].

Patent:

- Methods and System for Passive RFID-based Activity Determination (Submitted).

Acknowledgements

I would like to convey my deepest gratitude and acknowledge the extraordinary support and mentorship that have been essential throughout my doctoral journey. Spanning three years, this academic endeavor has been profoundly shaped and enriched by the guidance and contributions of several distinguished individuals.

First and foremost, I wish to extend my sincere appreciation to my supervisor, Prof. Qammer H. Abbasi. His unwavering support and invaluable guidance have been the bedrock of both my academic pursuits and personal growth. Prof. Abbasi's insightful advice and thoughtful mentorship have not only directed my research trajectory but have also been a source of continuous inspiration and encouragement.

I am equally indebted to my co-supervisors, Prof. Muhammad Imran, and to Dr. Hasan Abbas for their persistent support and invaluable insights throughout my doctoral journey. My experience was further enhanced by the collaborative environment at the Communication, Sensing, and Imaging Lab. Special thanks are due to Dr. Muhammad Usman and Dr. Jalil ur Rehman Kazim, whose engaging discussions, scholarly advice, and active involvement in my experiments have significantly contributed to my research.

I would also like to acknowledge the University of Glasgow for providing a supportive and enriching academic environment during times of difficulty. This support was crucial in facilitating my research and academic endeavors.

Lastly, but most importantly, my deepest appreciation goes to my parents and family. Their unwavering support, encouragement, and faith in my work have been a constant source of motivation and strength. Their belief in my capabilities has been unshakeable and has propelled me forward at every step of this journey.

This thesis stands as a tribute to the collective support, guidance, and encouragement provided by each of the individuals mentioned above. To them, I owe an immense debt of gratitude.

Declaration

Name: Muhammad Zakir Khan
Registration Number: xxxxxxx

I certify that the thesis presented here for examination for a PhD degree of the University of Glasgow is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it) and that the thesis has not been edited by a third party beyond what is permitted by the University's PGR Code of Practice.

I declare that the thesis does not include work forming part of a thesis presented successfully for another degree.

I declare that this thesis has been produced in accordance with the University of Glasgow's Code of Good Practice in Research.

I acknowledge that if any issues are raised regarding good research practice based on the review of the thesis, the examination may be postponed pending the outcome of any investigation of the issues.

Signature: Muhammad Zakir Khan

Date: 01-02-2024

Statement of Copyright

Chapter 1

Introduction

1.1 General Background

Human activity recognition (HAR) has significantly evolved over the last two decades, driven by its applicability in various sectors including healthcare, gaming, security, and surveillance [1]. This growth is attributable to technological advancements that have made it easier and more cost-effective to monitor daily activities such as cooking, eating, and sleeping. Initially, HAR primarily utilised vision-based methods involving cameras and computer vision algorithms. However, these methods raised privacy concerns and were dependent on consistent lighting conditions, leading to a shift towards more discreet and adaptable sensor-based methods [2].

Sensor-based HAR covers three primary categories: wearable sensors, object-tagged sensors, and device-free or dense sensing. Wearable sensors, which are worn or carried by individuals, have become increasingly popular in healthcare and fitness applications. However, their practicality is sometimes questioned, particularly among elderly users or those who find such devices obtrusive. Object-tagged sensors, on the other hand, are attached to daily-use objects and recognize activities based on user interactions with these objects. Though less intrusive than wearable sensors, their functionality is constrained by dependence on specific tagged objects.

The limitations of wearable and object-tagged sensors have led to the emergence of device-free sensing methods. This approach involves deploying sensors within an environment to passively monitor activities without requiring direct interaction or attachment to the user. Despite its advantages, device-free sensing faces challenges such as electromagnetic interference from electronics, signal disruptions by physical barriers, and environmental condition variations like temperature and humidity, which can introduce noise to data. The recent shift towards radio frequency (RF)-based technologies, such as FM-based, GSM-based, ultrawideband, Bluetooth-based, audio, light, zigbee, WiFi-based, Radar and radio frequency identification (RFID), for HAR represents a significant advancement. These technologies exploit the effects of human body movements on wireless signal propagation, such as reflection and scattering, to capture human activities. Wi-Fi, Radar, and RFID is referred as 'TriSense' selection for study due to

their complementary strengths in HAR. Wi-Fi is widespread, offering an economical option through existing networks. Radar excels in detecting precise movements and operates effectively in various visibility conditions, ensuring privacy. RFID is valued for its passive operation and minimal energy requirements, capable of tracking through obstacles, enhancing data richness. Collectively, these technologies provide a balanced of accuracy, privacy, and efficiency in diverse environments. Tri-Sense overcome many limitations of wearable and object-tagged methods by not requiring the target subject to wear or carry any device.

Tri-Sense technologies have a wide range of applications ranging from indoor navigation and health monitoring to human-computer interactions. They operate effectively without a line of sight and using the available signals in indoor environments. This allows for accurate capture of human body movements involved in daily activities. These technologies have also expanded beyond traditional uses, demonstrating versatility in applications in the indoor healthcare domain. Despite the advancements of Tri-Sense technologies, challenges remain, particularly in scenarios requiring cooperative and equipped subjects. Nevertheless, the development of device-free sensing methods marks a substantial evolution in the field of HAR. It offers more practical, privacy-conscious, and user-friendly solutions for monitoring and analysing human activities, paving the way for smarter, more responsive environments.

In conducting the research presented in this thesis, we adhered to the highest ethical standards. The ethical approval to conduct these experiments was obtained from the College of Science and Engineering at the University of Glasgow Research Ethics Committee (approval no.: 300200232, 300190109). This approval ensures that our research methods and interactions with subjects fully comply with ethical guidelines, emphasising our commitment to responsible and respectful research practices.

1.2 Problem Statement

Current HAR systems offer benefits but face several notable limitations. These limitations include privacy concerns associated with vision-based methods, dependency on environmental conditions, and practical challenges encountered with wearable sensors. Additionally, the functionality of object-tagged sensors is restricted. Moreover, device-free RF-based sensors, which are designed to improve privacy and convenience, face challenges often influenced by environmental factors.

This thesis aims to develop an accurate HAR system for monitoring human activities using radio frequency (RF) TriSense technologies. It focuses on two key aspects to improve accuracy: employing channel state information (CSI) for WiFi, received signal strength indicator (RSSI) for RFID, and amplitude for Radar; and developing algorithms that utilize raw CSI, RSSI, phase, and amplitude information over signal intervals. Traditional methods for indoor human activity identification and tracking often involve physical contact or invasive techniques, which can lead

to discomfort, inconvenience, and potential risks. To overcome these issues, this thesis proposes using Tri-Sense technology, a contactless and non-intrusive approach, to enhance monitoring in indoor environments. The goal is to create a more effective, efficient, and user-friendly HAR system by addressing existing limitations and harnessing the capabilities of RF-based technologies.

1.3 Aims and Objectives

The primary aim of this thesis is to develop a contactless, AI-enabled RF sensing system for indoor healthcare monitoring. This research pursues several key objectives to achieve this aim:

1. **Evaluation of Current Monitoring Systems**: To analyze current human activity monitoring systems, identifying their strengths and limitations to inform the development of a more effective system.

2. **Application of Advanced Machine Learning Techniques**: To implement advanced machine learning and deep learning algorithms for the accurate classification of human activities such as sitting, standing, leaning, walking, and falling.

3. **AI Integration in RF Sensing for Healthcare**: To augment RF sensing technologies with artificial intelligence algorithms, enhancing accuracy and performance for healthcare monitoring applications.

4. **Development of an Advanced RF Sensing System**: To design an RF sensing system capable of precise fall detection and activity localization without requiring physical contact with the individual.

5. **Utilization of Tri-Sense Technology for Non-Invasive Monitoring**: To leverage WiFi, Radar, and RFID technologies within the Tri-Sense framework to create a user-friendly and privacy-aware monitoring solution that surpasses conventional invasive monitoring methods.

6. **Enhancement of Activity Recognition Accuracy**: To refine activity recognition precision by exploiting specific technical parameters, including CSI for WiFi, RSSI for RFID, and amplitude measurements for Radar.

7. **Empirical Validation of Tri-Sense Technology**: To conduct empirical research to validate the practicality and effectiveness of Tri-Sense technology in real-world scenarios, demonstrating its benefits over traditional HAR systems.

1.4 Contributions

This section highlights our significant advancements in HAR through the TriSense technology and data fusion techniques for RFID and Radar technology. By leveraging machine learning and deep learning, we have achieved contactless activity recognition using a comprehensive set of sensing metrics including CSI, RSSI, phase, and amplitude, along with their synergistic fusion. Our study delivers key contributions as follows:

1. **Developed an Ensemble Machine Learning Algorithm:** Developed a super learner algorithm utilizing CSI for indoor activity recognition and localization, achieving superior accuracy over conventional models.

2. **Correlation Between Activity Position and Detection Accuracy:** Conducted a comprehensive analysis to establish the relationship between activity locations and detection accuracy, occupancy monitoring, and walking direction, providing insights into spatial influences on activity recognition systems.

3. **Enhanced Sequential and Simultaneous Activity Monitoring:** Leveraged UHF RFID tag arrays (TRT-Wall) to enhance the detection of both sequential and simultaneous activities, significantly improving monitoring granularity.

4. **Deep Learning-Enabled Contactless Fall Detection:** Utilized a modified deep learning Transformer model without a decoder, integrating TRT-Wall and fusing RSSI and phase data, to accurately detect falls in a contactless manner.

5. **Enhanced HAR through Radar and RFID Fusion:** Developed a LSTM-VAE fusion model that combines Radar and RFID data, markedly enhancing activity identification and localization capabilities. This approach introduces a novel method for comprehensive activity recognition leveraging the strengths of both technologies.

1.5 Thesis Organisation

This thesis is organised into the following chapters:

Chapter 2 provides an overview of current technologies and literature on human activity recognition, identification, and localisation using WiFi, RFID, and a fusion of RF sensing with deep learning algorithms for single and multi-modality applications. It also explores localisation methods aimed at enhancing the accuracy of identifying elderly individuals' activities in the home and healthcare environments, offering a comprehensive comparison of recent approaches, methods, and technologies for indoor localisation.

Chapter 3 details how RF sensing-based indoor localisation techniques can be applied for classifications of human activities and enhancing room occupancy in a contactless manner. The chapter makes use of traditional ML algorithms (super learner), an ensemble of multiple algorithms, showcasing superior accuracy, improving with distance from the transmitter.

Chapter 4 builds on the previous chapter by presenting a study comparing LSTM, CNN, and LSTM-CNN models for RF-based indoor activity recognition, focusing on seven activities in a single-room setting. The LSTM model demonstrated the highest accuracy in utilizing CSI for effective activity detection, highlighting RF sensing's potential in non-intrusive human activity recognition.

Chapter 5 explores the TRT-Wall, employed Impinj RFID readers, as an innovative, cost-effective, and private way to identify users by their activities without needing tags. It highlights the use of a data preprocessing technique and an LSTM for accurate activity detection, showcasing its superiority over existing methods with impressive identification accuracy rates across various data types, even in complex environments with multiple subjects.

Chapter 6 builds on the previous chapter by presenting a novel RFID-based method for contactless fall detection, utilizing passive UHF RFID tags and avoiding wearables. Employing a transformer model and early and late fusion technique for data analysis, the system achieved an impressive accuracy, showcasing its superiority to traditional methods like CNN, RNN, and LSTM.

Chapter 7 highlights the RFiDAR system's advancement in HAR, combining RFID and Radar through a LSTM-VAE model for improved accuracy. It showcases data fusion's role in overcoming sensor limitations, particularly in challenging environments. RFID's superior standalone accuracy is balanced by the fusion, underscoring the system's potential applications in elderly care.

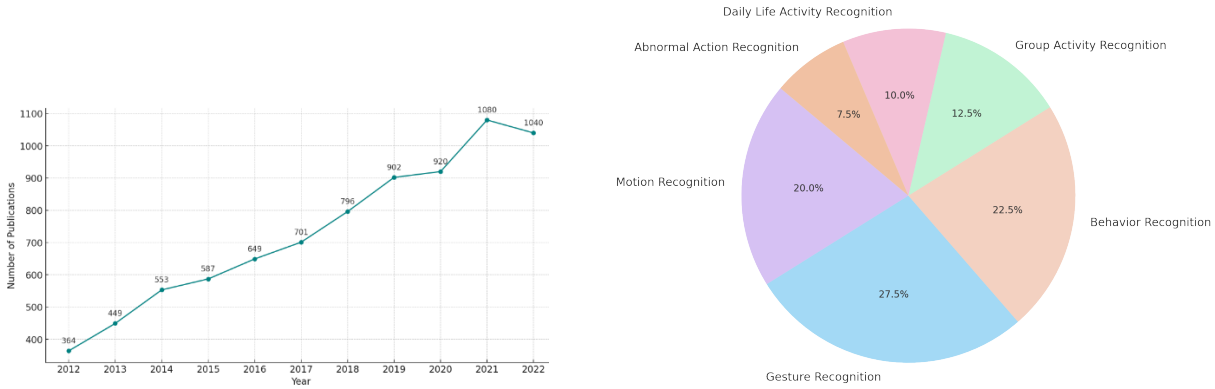
Chapter 8 concludes the thesis and details future work to be considered for expanding the work discussed throughout the thesis.

Chapter 2

Literature Review

2.1 Human Activity Recognition Technologies

The field of HAR has significant progress due to advancements in sensing technologies. This development offers an array of applications ranging from improving interactions in smart homes and providing care for the elderly, to monitoring health and enhancing security surveillance. This chapter primarily focuses on the precise recognition of human activity within indoor environment, which are essential for both daily life and controlled environments. At the heart of these advancements is the role of artificial intelligence (AI), which processes and classifies data from different devices. These include wearable sensors [3], smartphone inertial sensors [4], Kinect camera systems [5], and CCTV [6] devices. The fusion of AI with HAR devices has led to their application in diverse fields such as healthcare [3, 4], surveillance [7, 8], remote elderly care [7, 9], and the creation of intelligent environments in homes, offices, and urban areas [4, 7]. This has not only enhanced human safety but also significantly improved quality of life [10].
The employment of sensors, cameras, RFID, and Wi-Fi in HAR, while not entirely new, has been significantly pushed by rapid advancements in AI, enhancing the capabilities of these technologies in various domains [11]. This mutualism between AI models and HAR devices, as evidenced by recent research trends, suggests a move beyond the traditional reliance on single images or brief video sequences as shown in Figure 2.1. Instead, current advancements have expanded the potential for broader applications of HAR across different sectors, indicating a progressive integration of these technologies into daily and controlled environments. This progression in HAR, underpinned by AI, not only enhances the accuracy and efficiency of activity recognition within indoor environments but also push the field toward a future where smart environments are an integral part of enhancing human safety and improving quality of life. The mutual advancements in AI and sensing technologies continue to drive the HAR field forward, promising even greater integration and application possibilities in both daily life and controlled environments.



(a) Publication (b) Trends

Figure 2.1: Research publications in human activity recognition over ten-year trends [12].

2.1.1 Overview of WiFi-based Recognition

Recently, there has been a growing interest in utilizing RF technologies, particularly WiFi, to detect human activities. WiFi networks connect a variety of devices, including smart speakers, televisions, thermostats, and security systems, creating a comprehensive indoor network. As individuals move within these premises, WiFi signals, naturally reflecting off surfaces, undergo changes in strength and phase. These changes provide meaningful information into human presence and movement, enabling an innovative approach to monitoring activities without the need for direct visual observation or wearable sensors. This method leverages existing infrastructure to offer a non-intrusive, cost-effective means of collecting data on human activity. Furthermore, researchers investigating WiFi signal variations often focus on the physical properties of the wireless channel, such as RSSI and CSI. They use commercial network interface cards (NIC) with special software, like the Intel 5300 NIC [13] and the Atheros 9580 NIC [14], to access these signal properties. For more accurate detection, some researchers employ a platform known as the universal software radio peripheral (USRP) to fine-tune wireless signals, identifying frequency shifts caused by human motion [15]. They also explore the Doppler effect, where changes in frequency help track movement, using USRP platforms for precise signal handling [16].

This thesis explores the advanced applications of WiFi sensing to enhance quality of life by focusing on detecting human activities and locating people indoors. We categorize these applications into four main areas: intrusion detection and room occupancy monitoring, activity and gesture recognition, vital sign monitoring, and user identification and localisation. Intrusion detection aims to identify unauthorized entries, while occupancy monitoring assesses how many people are present in a space. Activity and gesture recognition involves identifying common activities or smaller movements. Vital sign monitoring leverages subtle signal changes to track health indicators such as breathing and heart rate. User identification and localisation employ

WiFi signals to determine people's indoor locations and identify them based on their activity patterns. Our research primarily focuses on HAR including monitoring occupancy, recognizing activities, and identifying users with localisation in an indoor environment.

A detailed review of HAR and localisation techniques is summarized in Table 2.1. Figure 2.2 illustrates a typical process in wireless human activity sensing systems. These systems detect signal changes caused by human activities using various techniques such as RSSI, CSI, frequency modulated continuous wave (FMCW), and doppler shift (DS) then clean up the signals to reduce errors, and finally extract features that machine learning models use to accurately identify human activities.

Table 2.1: Overview of WiFi sensing for activity recognition and localisation.

Application	Techniques			
	RSSI	CSI	FMCW	Doppler Shift
Daily Activity Recognition	[17, 18, 19, 20]	[21, 22, 23, 24, 25, 26, 27]	[15]	[28, 29]
Indoor Localisation	[30, 31, 32, 33, 34, 35, 36, 37, 38]	[39, 40, 41, 42, 43, 44]	-	-

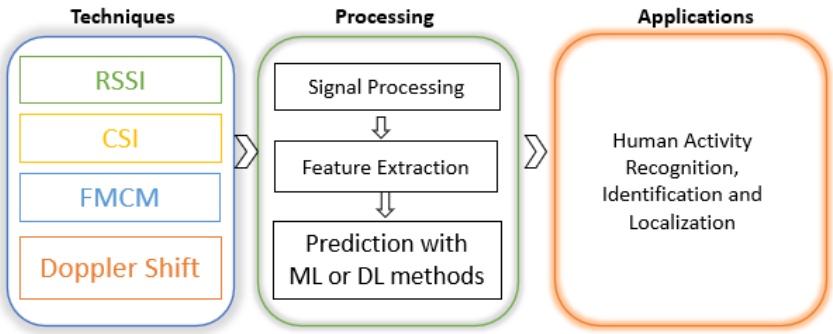


Figure 2.2: Workflow of sensing human activity through wireless signals.

2.1.2 WiFi Sensing Techniques

In this section, we explore various WiFi sensing techniques that utilise different physical layer properties of both commodity devices and customized hardware for human activity sensing. We focus on four key techniques: RSSI, CSI, FMCW, and Doppler shift. These techniques, leveraging distinct physical layer attributes, these techniques are essential for effective HAR, as detailed in Table 2.2.

Table 2.2: Comparison of WiFi sensing techniques for human activity recognition.

Technique	Literature	Derived Metrics	Additional Hardware
RSSI-based	[45, 17, 18, 19, 20]	Wireless signal strength	No
CSI-based	[21, 22, 23, 24, 25, 26, 27]	Channel condition	No
FMCW-based	[15, 28, 29, 16]	Frequency shift	Yes
DS-based	[15, 28, 29, 16]	Frequency shift with speed	Yes

Preliminary

1. Received Signal Strength Indicator (RSSI)

The RSSI is a common feature in WiFi devices that quantifies the signal attenuation, or path loss, of wireless signals over distance. This attenuation is modeled using the log-normal distance path loss equation:

$$P(d) = P(d_0) + 10\gamma \log\left(\frac{d}{d_0}\right) + X_\delta, \tag{2.1}$$

where $P(d)$ denotes the RSSI measurement representing path loss at distance d in decibels (dB), $P(d_0)$ is the path loss at the reference distance d_0, γ is the path loss exponent, and X_δ is the Gaussian noise due to flat fading [46].

RSSI has played an important role in applications like user positioning with portable WiFi devices, due to its ability to sense changes caused by human presence near wireless networks. This utility has extended RSSI's use to device-free indoor localisation [38, 47], room occupancy [45], and respiratory monitoring [48], leveraging its simplicity and availability on standard WiFi devices without extra hardware. However, its straightforward measure of signal attenuation over distance can limit reliability, particularly in static environments where the absence of movement challenges its effectiveness [49].

2. Channel State Information (CSI)

CSI plays an important role in precise human activity sensing within indoor environments. CSI captures the detailed effects of signal propagation, including scattering, fading, and power decay, which are influenced by the presence or movement of a human body. Unlike the RSSI, which provides a simple power metric, CSI offers a rich set of complex values that represent both amplitude and phase information across multiple orthogonal frequency-division multiplexing (OFDM) subcarriers. This fine-grained depiction of wireless channels allows for accurate tracking of human movements. The IEEE 802.11n standard enables CSI measurement over 52 or 128 subcarriers for 20MHz and 40MHz channel bandwidths, respectively, while the 802.11ac

standard expands this capacity even further. CSI is essential for detailed channel estimation and can be mathematically represented as:

$$H = [H_1, H_2, \ldots, H_i, \ldots, H_N]^T, \quad i \in [1, N], \tag{2.2}$$

where N is the number of subcarriers and H_i is defined as:

$$H_i = |H_i| e^{j\theta_i}, \quad \theta_i = \angle H_i, \tag{2.3}$$

with $|H_i|$ representing the amplitude and $\angle H_i$ the phase of the i^{th} subcarrier. Accessible through modified drivers on standard WiFi interfaces, such as the Intel 5300 NIC [13] and Atheros 9580 NIC [14], CSI has become increasingly utilised by researchers for a range of sensing applications, from intrusion detection to walking speed estimation and activity recognition [22].

3. Frequency Modulated Carrier Wave (FMCW)

FMCW technology is used to identify human activities by examining the reflections of radio waves. Given the fast speed of wireless signals, it's not feasible to measure the time it takes for these signals to travel (time of flight, or ToF) directly. FMCW overcomes this challenge by linking the time differences in signal travel to changes in frequency, thus making ToF measurements possible. In this process, the frequency of the signal sent by the transmitter changes across a specific range. When this signal bounces off a human body, it returns with a change in frequency and a slight delay. This change and delay help in determining the position and movement of the person, as the frequency shift directly relates to the time delay in the signal received relationship is given by:

$$\Delta t = \frac{\Delta f}{k}. \tag{2.4}$$

Here, Δt is the time shift, Δf is the frequency shift, and k is the frequency sweep rate. The round-trip distance is calculated as $d = c \cdot \Delta t$, where c is the speed of light. This technique, unlike standard WiFi OFDM, requires specialized equipment like USRP for signal generation. FMCW-based wireless sensing systems have been developed for various applications. These include tracking human figures through walls [50], 3D motion tracking [15], gait velocity, and stride length estimation [51], vital signs detection [52], sleep monitoring [53], and emotion recognition [54], among others.

4. Doppler Shift (DS)

The Doppler shift is an inherent property of wireless signals that plays an important role in sensing human activities. This phenomenon is observed as a change in the frequency of a wire-

less signal, attributed to movement. Specifically, when an individual move toward a receiver, the frequency of the signal perceptibly increases; conversely, it diminishes when the individual moves away from the receiver. This variation in frequency is directly related to the movement of an object, such as a hand, moving at a velocity v and at an angle θ with respect to the receiver, resulting in a DS. This shift is instrumental in accurately detecting and analyzing human movements, leveraging the dynamic nature of wireless signals to sense activities.

$$\Delta f = \frac{2v\cos(\theta)}{c} f, \tag{2.5}$$

where c represents the speed of light, and f is the central frequency of the wireless signal. utilising the DS, WiFi sensing systems built on software-defined radio devices like the USRP N210 have been developed to detect activities such as walking, running, and hand gestures.

2.1.3 Human Activity Recognition and Localisation

HAR serves as a cornerstone for various applications, such as human-computer interaction, elderly care, wellness management, and security surveillance. Traditional methods have depended on specific sensors like wearables and cameras. In contrast, RF-based methods offer a device-free approach to activity recognition, covering privacy issue by avoiding the capture of unnecessary or sensitive information. Human activities typically fall into two categories: regular (daily activities) and gesture (movements of the hands, fingers, and head).

Regular Activity Recognition and Localisation

Regular activities within a home environment including walking, sitting, cooking, and watching television offer valuable information into daily routines. Monitoring these activities can inform recommendations for healthier lifestyle. Additionally, this surveillance extends to applications such as elderly care and ensuring the safety of children left at home alone. In the subsequent sections, we discussed four WiFi-based methods for recognizing these regular activities, showcasing their potential in fostering a safer and healthier living environment.

1. Received Signal Strength Indicator

RSSI is an imporant metric of RF sensing technology in activity recognition and indoor tracking, exploiting the unique patterns of wireless signal fluctuations caused by human movements. Sigg et al. [17] demonstrated the capability of RSSI-based systems to detect movements and differentiate between various activities such as lying, standing, and walking without the need for carrying any device, through analyzing signal characteristics like peaks and employing a k-nearest neighbors (KNN) classifier. This approach is further refined to distinguish between static and dynamic activities using both ambient signals and dedicated transmitters, highlighting

the versatility of RSSI in capturing the essence of human motion [18]. Complementing this, advancements in radio tomographic imaging (RTI) and variance-based RTI (vRTI), as introduced by [20], significantly improve the accuracy of motion tracking and recognition by enabling the visualization of moving targets through sensor networks, thus underscoring the non-invasive nature of RSSI-based methods in activity recognitio without privacy concern.

Parallel to the development in activity recognition, RSSI has been extensively applied in indoor tracking and localisation, spearheaded by [30] RADAR system, which correlates signal strength with distance. The system's accuracy is further enhanced through the implementation of the kalman filter in the RSSI propagation model by Güvenc [31], and similar enhancements by [32]. Moreover, the integration of RSSI with other sensor data like GPS and infrared sensors by C[33], respectively, showcases the potential for more stable localisation solutions. The inherent sensitivity of RSSI to environmental factors necessitates robust approaches to localisation, leading to the development of differential RSSI methods using particle filters by [34], and dynamic techniques by [35]. to counteract environmental variability. To address the challenges posed by multipath attenuation, Xie et al. [36] employ a KNN scheme based on Spearman distance, enhancing localisation accuracy. Barsocchi et al. [55] propose a novel virtual calibration technique for signal propagation models, eliminating the need for manual training, and [56] introduce ArrayTrack, a system that achieves precise RSSI-based localisation without pre-calibration. Furthermore, the exploration of multi-user localisation through RF tomography of RSSI measurements by [57] expands the application horizon of RSSI technologies.

2. Channel State Information

CSI has been identified as an alternative to RSSI for HAR, due to its enhanced resolution and sensing capabilities. Pioneering research, such as Wang et al.'s introduction of E-eyes [21], used CSI's granularity to recognise daily activities. E-eyes ingeniously utilizes the correlation between location and activity traits to distinguish between stationary activities (cooking or sleeping) and locomotive activities (walking between rooms), all through a single WiFi access point. This distinction is made possible by observing the unique amplitude distributions of CSI, which vary significantly between different activities [21], and by noting that consistent CSI patterns emerge for the same walking trajectories, though these patterns change with different paths , as depicted in Figure 2.3. However, the initial applications of CSI in activity recognition faced limitations due to a lack of theoretical foundations linking CSI measurements with specific human activities. Addressing this, Wang et al. [23] furthered the field with the development of CARM. Such findings confirm the potential of CSI from WiFi signals to be a reliable indicator of specific in-place activities or unique walking paths, thereby serving as a feasible method for recognising routine daily activities.

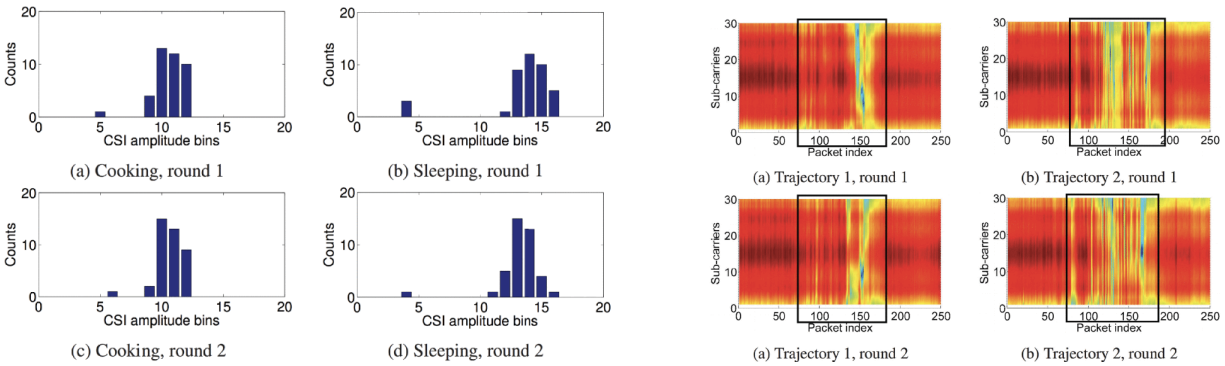
The importance of detecting abnormal movements, particularly falls, has led to the development of specialized systems using CSI for timely elder care. The WiFall system [24] was one

of the first to distinguish falls from other activities using CSI data, employing a support vector machine (SVM) for accurate identification. Enhancements to fall detection have been made with their Anti-Fall system [25], which considers both phase and amplitude of CSI, and by Wang et al. [26], who found that CSI phase differences between antennas effectively sense falls. Further advancing the field, Palipana et al. [27] introduced FallDeFi, applying short-time fourier transform (STFT) to extract time-frequency features from CSI and using feature selection techniques to adapt to environmental changes.

Beyond activity recognition and fall detection, CSI has been instrumental in user identification and indoor localisation. Techniques using CSI for user identification distinguish individuals through unique physiological and behavioral patterns, such as gait. We discuss these CSI-based approaches and summarize them in Table 2.3. Zhang et al. [58] and Wang et al. [59] have demonstrated methods to identify individuals by analyzing CSI variations resulting from walking, while WFID [60] and [61] have extended the application of CSI to device-free authentication and a wide range of activities through deep learning models. CSI-based indoor localisation techniques represent a robust improvement over RSSI-based methods. The foundational works by Wu et al. [39, 40] and Sen et al. [41, 42] have enhanced the precision of indoor localisation, with advancements like the SAIL system [43] achieving accurate localisation with a single WiFi access point. Further exploration by Wang et al. [44] into deep learning approaches has shown the potential for even greater accuracy by using neural networks to interpret CSI data for localisation purposes.

Table 2.3: Comparison of WiFi-based identification techniques using CSI.

Literature	Frequency	Accuracy	Activity	Distance
[58]	2.4-5GHz	93% for 2 Subjects, 77% for 6 Subjects	Walking	2m
[62]	2.4GHz	93% for 2 Subjects, 80% for 6 Subjects	Walking	2-3m
[59]	5GHz	79.28% for top-1, 89.52% for top-2, 93.05% for top-3	Movement speed of different body part	6m
[60]	Not mentioned	91.9% for 9 subjects, 93.1% for 6 subjects	standing and walking	3.6m
[61]	5GHz	94% for walking and 91% for static	walking and static activities	10m



(a) Cooking, round 1 (b) Sleeping, round 1 (a) Trajectory 1, round 1 (b) Trajectory 2, round 1

(c) Cooking, round 2 (d) Sleeping, round 2 (a) Trajectory 1, round 2 (b) Trajectory 2, round 2

(a) Cooking vs Sleeping (b) Walking Pattern

Figure 2.3: Comparative histogram analysis of specific subcarrier CSI amplitudes for cooking vs. sleeping activities and consistent patterns in identical walking paths [21].

3. Frequency-Modulated Continuous-Wave

FMCW recognition offers a refined approach to activity recognition, distinct from RSSI and CSI methodologies that utilise standard wireless devices. This technique employs the USRP platform, modulating transmitted wireless signals across a specific frequency band, like that of an FMCW radio. It then calculates time differences (Δt) from the reflected signals. The FMCW radio's architecture, based on the super-heterodyne principle, provides enhanced sensitivity and stability for activity recognition. WiTrack system [15] was the first in FMCW-based recognition which captures radio signals reflected off the human body to track three-dimensional movements. WiTrack employs a T-shaped directional antenna array to localise the human body's center in three dimensions and can approximately track body parts, such as estimating the direction of a pointing hand within a median error of 11.2 degrees. Moreover, WiTrack distinguishes between different activities, like falling or sitting, by observing the Z-axis value and changes in elevation, as depicted in Figure 2.4.

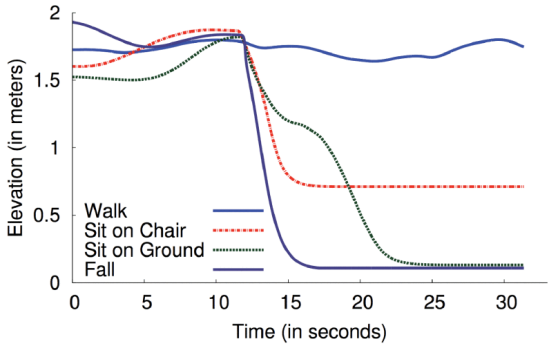


Figure 2.4: WiTrack: Monitoring falls through elevation and height changes [15].

4. Doppler Effect

The Doppler effect is utilised in HAR to detect subtle changes in WiFi signals induced by movements such as running, walking, or other daily activities [29, 28]. Chetty et al. [28] employed a passive WiFi radar on a USRP platform to monitor Doppler shifts through walls, indicating human activity. Subsequently, Adib and Katabi [29] enhanced this system by applying MIMO interference nulling, which mitigates the flash effect in Doppler shifts, thus improving recognition accuracy. Further advancements by Okamoto and Ohtsuki [63] included using the temporal phase shift from moving targets, along with MIMO interference, to ascertain the relative velocity between the target and the antenna. Building on this, Okamoto and Tomoaki [63] leveraged a bistatic radar model within a MIMO framework to categorize diverse human activities and track multiple subjects. Successive studies [64] extended these techniques by designing antenna arrays to measure Doppler shifts from everyday movements of the elderly, such as falling or sitting, providing valuable insights into their safety and well-being.

Related Work on WiFi-based Human Activity Recognition and Localisation

Recent studies in the field of HAR using RF signals have shown significant advancements, particularly in healthcare monitoring applications. Iqbal et al. [65] developed a deep-learning-based system using Wi-Fi sensor data to classify different user movement states, such as forward, backward, and no movement. Similarly, Nipu et al. [66], along with Al et al. [67], employed CSI from USRP devices to identify various participants, capturing CSI data as individuals walked across radio frequencies. This data was then subjected to machine learning algorithms like random forest and decision trees for accurate activity detection. The effectiveness of USRP devices in HAR has been further demonstrated through various studies. The use of USRP N210 devices by Wang et al. [68], Zhang et al. [69], and Pu et al. [16] achieved notable accuracy rates (91-95%) in activity detection. Bokhari et al. [70] utilised two USRPs to propose a deep-gated recurrent unit model for non-obtrusive HAR, achieving 95% accuracy. Additionally, Taylor et al. [71] reported 96% accuracy in differentiating sitting and standing activities using USRPs X300 and X310 and the random forest algorithm.

In the aspect of localisation and tracking using RF signals, the WiTrack2.0 system by Adib et al. [72] marked a significant development. This multi-person localisation system calculates the distance between users and antennas using the time taken for reflected wireless signals to travel. Other experiments in localisation, such as those reported by Shi et al. [73], have achieved substantial accuracy. The field has also seen progress in WiFi-based sensing for indoor activity, with Hoang et al. [74] focusing on user location and tracking using WiFi RSSI, which considers correlation among RSSI measurements in a trajectory. However, challenges remain in the field, particularly with intrusive methods that, despite their accuracy, are often cumbersome and unwelcome, especially among sensitive groups like the elderly or children [75]. The demand is

growing for contactless, long-term health monitoring systems that can efficiently detect activity in crowded environments [76]. While radar-based solutions have shown precision in detection and real-time monitoring [77], the high costs and energy consumption associated with these technologies call for more cost-effective and energy-efficient solutions for broader adoption.

2.1.4 Contactless RFID-based Activity Recognition

RFID technology, originally designed for military use to identify aircraft, has significantly evolved over recent years. It is now extensively utilised in tracking and supply chain management. The technology's range has expanded remarkably, with passive tags reaching up to 15 meters and active tags up to 100 meters. The RFID system comprises two primary components: the reader and the tags. The reader, equipped with an antenna, sends out radio waves that are modulated by RFID tags attached to objects. These tags transmit information, like their ID, back to the reader. There are two types of tags: active tags, which have their power source, and passive tags, which derive energy from the reader's radio waves. Active tags generally have a longer range than passive tags. Due to its simplicity, cost-effectiveness, and non-intrusive nature, RFID is increasingly utilised in various fields, including human activity recognition research. This involves posture, gesture recognition, tracking, localisation, and behavior recognition applications.

Contactless Activity Recognition

This section is focused on exploring device-free approaches that utilize RFID technology, consistent with the classification framework provided. We specifically focused on 'action-based' approaches, given their significant research importance in addressing the current issue. A schematic representation of the various research directions in activity recognition, according to a defined taxonomy, is presented in Figure 2.5. In the domain of device-free (contactless) activity monitoring, various innovative approaches leveraging RFID technology have been explored, each with its unique focus and application domain, ranging from gesture-based activity recognition to posture-based monitoring and tracking customer behavior in retail settings.

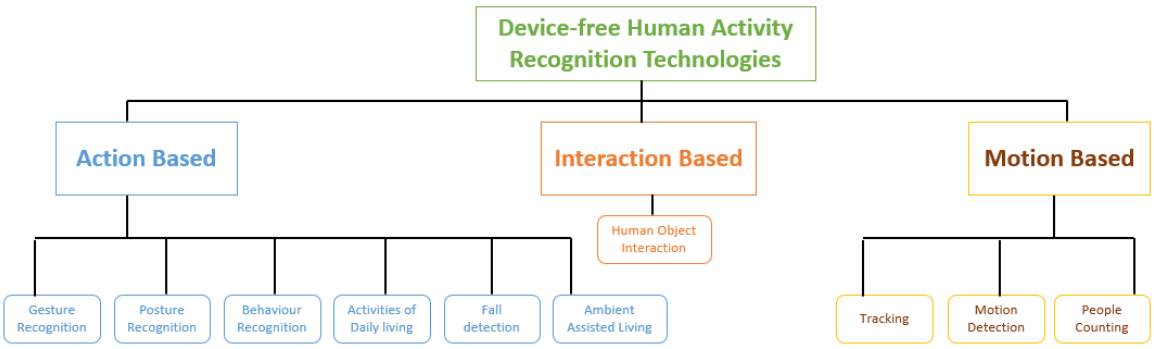


Figure 2.5: Taxonomy of human activity recognition.

Action-based Activity Recognition

Action-based activities are those that involve some action of the human body. This action can involve either the whole body or a specific portion of a body as shown in Figure 2.6. We provide an overview of the different solutions proposed for the recognition of action-based human activities.

Figure 2.6: RF-based action activities recognition .

Gesture-based activity recognition, a key component of human-machine interaction, has seen significant advancements through RFID technology. Ye et al. [78] introduced a method using passive RFID tag arrays for gesture recognition, known as the link state indicator (LSI), which calculates a gesture matrix to represent tag states during different gestures, achieving an average accuracy of 94% for twelve gestures. However, this approach struggles with similar gestures and is limited to offline use. Similarly, the Smart Surface technique [79] and Ding et al.'s work [80] furthered the application of passive RFID tags for recognizing hand motions and handwriting, respectively, albeit with limitations in gesture complexity and the need for close proximity to the RFID tag plate. Moreover, GRfid demonstrated the capability to detect six hand gestures with an accuracy of up to 96.5% through RFID signal phase changes, albeit with its own set of limitations, including basic gesture recognition and unspecified system latency [81].

Parallel to gesture recognition, posture-based activity recognition offers a more realistic alternative for monitoring without direct contact. RF-Care employs passive RFID tags to detect activity-related disturbances with high accuracy in steady posture recognition but faces challenges in latency and transition detection accuracy [82]. Subsequently, techniques such as Yao et al.'s [83] sparse dictionary-based approach and the R&P method that combines phase and RSSI values have shown significant accuracy in activity recognition across various scenarios, though they also face challenges with system complexity and latency [84]. Additionally, the exploration of passive RFID technology has extended to the retail sector for tracking shopping behavior. Han et al. developed the CBID system, which analyses RFID tag signals from store items to identify shopping patterns and item correlations, though its effectiveness with metallic products and latency issues remains unexplored [85]. Similarly, Zhou et al.'s system aims to mine customer shopping behaviors through the detection of actions such as item browsing

and picking, based on RFID tag phase changes. This approach, while successful in identifying trends, faces performance challenges in crowded shopping environments [86].

Related Work on RFID-based Fall Activity Recognition

Falls are a major health concern globally, ranking just behind traffic accidents as a leading cause of accidental injuries [87]. This issue is especially critical among the elderly, for whom falls are the main cause of various health problems, including death, disease, weakness, and disability [88]. It is noteworthy that about 30% of people over 65 fall at least once a year and these falls often have serious consequences: 90% result in broken hips and 60% in head injuries [87]. A significant number of these elderly individuals, nearly half, experience what is called a 'long-lie' after a fall, where they remain on the ground for an extended time. This situation increases the risk of serious health issues such as pressure ulcers, cold exposure, lung infections, dehydration, and even death [89]. Falls that go unnoticed can have particularly grave outcomes, especially if the person doesn't receive prompt medical attention. To address this, the implementation of automatic fall detection systems has been shown to greatly speed up the response time of medical professionals to these incidents [90].

Fall detection methods are primarily divided into two categories: wearable [91, 92, 93, 94] and non-wearable [95, 96, 97] systems. Wearable fall detection systems utilise sensors like gyroscopes and accelerometers, which are attached to the body to monitor changes in activity. Non-wearable systems, on the other hand, use different techniques such as visual detection [95, 98], environmental sensing, and RFID [99]. Wearable devices provide a detailed monitoring system, collecting data to represent daily activities. This is achieved through sensors like accelerometers, gyroscopes, and RFID technology. For example, Le and Pan [100] developed a system specifically for the elderly using wearable acceleration sensors. Vallejo et al. [101] used MLP for data classification from accelerometers. Micucci et al. [102] applied KNN and SVM for data categorization. Gjoreski et al. [103] attached accelerometers to sportswear, using Bluetooth for data transmission. Wang et al. [104] created a pendant with a three-axis accelerometer and barometric sensor. Similarly, Xiaoling et al. [105] utilised smartphone sensors for gesture recognition. These wearable technologies are highly accurate in fall detection but require users to wear specific equipment on certain parts of their bodies. This necessity could be seen as a limitation for some users, highlighting a trade-off between convenience and effectiveness in these systems.

Vision-based systems for detecting falls analyze images from cameras, using either standard RGB or depth-sensing technologies. These systems have evolved but still face challenges in accurately distinguishing falls from similar movements like lying down or leaning. To enhance fall detection, various research approaches have been explored. For instance, Rougier et al. [106] focused on how the human body's outline changes in images during a fall. Mirmahboub et al. [107] used cameras to observe falls among the elderly, analyzing behaviors captured in video

sequences to differentiate between normal activities and falls. Yun and Gu [108] developed a method to identify falls by examining the changing shape and movement of human figures in images. Khraief et al. [109] created a 4D-CNN network that processes different types of data through individual CNN streams for detecting falls. Another innovative approach by Abdo et al. [110] involved using RetinaNet for tracking movement in video frames, with MobileNets enhancing the categorization of movements as falls. These vision-based methods are advantageous as they don't require users to wear any devices and can offer a clear visual perspective of falls. However, they require high-performance computing due to the complexity of their algorithms and can raise privacy concerns. They are also sensitive to varying light conditions and need a clear line of sight between the camera and the subject, making their implementation in complex environments challenging.

Previous studies have primarily focused on tag-based approaches [111, 112, 113, 114] where RFID tags were attached to humans for activity capture, including object identification and tracking [115, 116, 75, 117], which can infer human activities from RFID-labeled objects [118, 119]. However, traditional tag-based systems require specific objects or users to have RFID tags attached to track their movement and infer various activities. This limitation arises when activities such as pushups, walking, and body rotation do not directly interact with RFID-tagged products or cause movement in the environment. In contrast, RF based systems, including TriSense, offer potential solutions to address the limitations of wearable devices and cameras. Among these, WiFi systems have shown promise in fall detection, leveraging CSI. For instance, Wang et al. [26] demonstrated real-time automatic data segmentation and fall detection using fine-grained CSI data from WiFi devices. Another notable system, WiFall [24], utilised WiFi-CSI for fall detection by measuring predefined motions and employing a one-class SVM and RF for classification. While WiFi signal detection effectively addresses challenges related to lighting and user privacy, its deployment costs can be prohibitive due to sensitivity and stability issues in complex monitoring environments, thus limiting its widespread implementation for the elderly [120]. On the other hand, Doppler radars have also been employed for fall detection based on human movement speed. However, they can be influenced by non-fall activities. To address this, Ma et al. [121] proposed the use of ultra-wideband (UWB) monostatic radar and an LSTM algorithm for fall detection. Nevertheless, the system's adaptability to new individuals and environments is constrained by residual environmental impacts. Tian et al. [122] presented a solution involving two perpendicular angle-range heat maps to differentiate between human daily activities and falls, leveraging a large-scale dataset under various scenarios to mitigate environmental impact.

Related Work on RFID-based Daily Activity Recognition

Recent research has been directed toward employing UHF passive RFID tags to enhance the quality of life for elderly individuals. These applications cover a broad spectrum, including tracking location and mobility, managing medications, and preventing falls. Researchers have

Table 2.4: Comparison of different technologies for fall recognition.

Approach	Technology	Advantages	Disadvantages
Wearable Device	Accelerometer + RFID, smartphone, barometer, magnetometer	Low cost	Constraint to wear the device
Device Free	Radio devices	Low cost	Customized hardware required
	RFID	Low cost, COTS available, passive	Environmental interference

investigated both 'tag-free' and 'tag-based' technologies for monitoring and analyzing daily activities of the elderly. In the 'tag-based' approach, various studies have been conducted. Raad et al. [123] introduced a prototype using passive RFID wearable anklets or bracelets. This system is designed to detect wandering elderly individuals within their homes. Similarly, Shuaieb et al. [124] proposed a cost-effective indoor location system using RFID tags, aimed at improving alarm systems in nursing homes and triggering emergency services when needed. Feng et al. [125] developed a posture-recognition system called 'SitR'. This system employs RF signals and lightweight RFID tags placed on the user's back to identify seven different sitting postures. The 'TagCare' system by [126], is a fall detection system for the elderly that utilises RSSI and Doppler frequency readings. Another approach for fall detection involves using passive RFID sensor tags in indoor footwear to monitor RSSI and pressure changes. Toda et al. [127] provided a comprehensive mechanism for fall detection by analyzing routine activities through shoe sole pressure data and RSSI fluctuations. Ruan et al. [128] introduced the 'TagFall' system, which detects falls from daily activities using abrupt changes in RSSI values. These examples illustrate the effectiveness of 'tag-based' solutions in recognising various activities and postures.

On the other hand, 'tag-free' sensing technologies offer benefits such as lower cost, non-intrusiveness, and simplicity in structure. Sigg et al. [18] introduced a 'tag-free' method using ambient FM radio signals for activity detection. Similarly, He et al. [129] proposed a technique to enhance the signal-to-noise ratio of RFID tags for detecting activities in a contactless manner. Moreover, Zou et al. [81] developed the 'GRfid' system for gesture recognition using multi-tag phase measurement and DTW distance normalization. The 'RFree-GR' system by Dian et al. [130] demonstrated the ability to recognise fine-grained gestures, including 16 American Sign Language words. Futhermore, Zhao et al. [131] presented the 'RF-Motion' system, which identifies six types of human motion with 87% accuracy, leveraging DTW, synthetic aperture algorithms, and data slicing techniques. RFID tag arrays have also been used to predict motion and pose-based activities, though challenges such as occlusions or physical obstructions exist in wall-mounted systems [83, 132]. The 'RF-HMS' system in [133], utilised an RFID tag array for tag-free human motion monitoring through walls, achieving an average accuracy of 90%. The 'IDSense' system integrates passive RFID tags into everyday objects to detect interactions

involving motion and touch, particularly near walls [134]. These 'tag-free' approaches typically depend on tags or tag arrays as reference points to detect motion or direction. They often require proximity to the tags for effective activity recognition, which can limit their applicability.

2.1.5 Human Activity Monitoring on Multi-modal Fusion

The recent advancements in technologies like artificial intelligence, 5G, and the Internet of Things have significantly altered how we live. These technologies have led to the creation of intelligent sensors that are now commonly found in wearable devices. These devices overcome some of the limitations found in computer vision, such as issues with space, time, privacy, and energy use. Instead, they provide a more effective way of using sensors to monitor human activities and emotions. These sensors are highly valued for their computational power, small size, and cost-effectiveness[135]. Focusing on the effectiveness of single sensors in recognising activities has been a key area in sensor technology research [136, 137]. Single sensors are user-friendly and durable, but they have limitations in accurately detecting complex movements due to their focus on localised motion. To address this, recent studies have been exploring ways to combine data, features, and classifiers to improve the monitoring of health and activities. The number of sensors used in these systems is an important consideration, balancing efficiency and user convenience. The decision to use multiple sensors often depends on the user's comfort and available space [138].

Data processing in systems that use multiple sensors can be divided into three categories: data-level, feature-level, and decision-level fusion [139]. Data-level fusion focuses on cleaning, extracting features, classifying, and compressing data from various sensors to improve the system's reliability and robustness. Feature-level fusion works by creating complex vectors from different sensor features and uses machine learning to effectively categorize data [140]. Decision-level fusion combines the outcomes from multiple sensors to produce a coherent final decision, often incorporating information processed in earlier stages [141]. For example, Guo et al. [142] combined multiple sensor nodes for activity and gesture recognition. This process of fusion minimizes errors and improves data categorization by using a variety of methods, ranging from statistical approaches to machine learning and deep learning algorithms. These methods are designed to manage both spatial and temporal data. Liu et al. developed a system for monitoring driving behaviour using motion capture and AI [143]. Similarly, [144] used a combination of machine learning and deep learning in a fusion framework for recognising human activities. These developments demonstrate the significant progress in sensor technology and its wide-ranging applications in our daily lives.

Related Work on Contactless Multi-modality Sensing Fusion

In the context of device-free HAR, there are two main methods of combining data from multiple modalities: score fusion and feature fusion. Score fusion involves combining the outcomes or decisions from different modalities, typically using techniques like weighted averaging [145], or through learning models [146]. Feature fusion, in contrast, merges distinctive characteristics from various data sources into a comprehensive feature set that is highly effective for HAR [147]. Furthermore, several studies have explored the integration of different modalities to improve activity recognition [148, 149, 150, 151], using decision-level fusion techniques that combine vision-based sensing with Wi-Fi -CSI or wearable sensors for a range of activities.

Recent advancements have shifted towards developing multi-modal fusion systems tailored for smart sensing applications. For instance, [152] introduced the RF-Focus system, combining RFID and computer vision (CV) to identify and locate tagged boxes on conveyor belts. However, this system's focus on recognising tagged boxes does not align with the broader context of HARI. Another study by Wu et al. [153] integrates RFID and CV for object detection using dynamic Bayesian networks but lacks user interaction and depends on potentially intrusive RFID reader bracelets. TagVision, proposed in [154], effectively combines RFID and CV for precise object identification and tracking, especially in motion target tracking using the CV subsystem. The DEEM system, as described in [155], leverages CV and RFID for assessing fitness effectiveness and user identification but does not focus on detailed gestural action recognition, facing challenges like low tag reading rates and unknown hand-tag offsets. The ShakeReader system in [156] connects tags with smartphones using unique reflector polarization models for indirect tag reading. Similarly, the RF-Grasp system, developed by [157], focuses on robotic tasks and object localisation using RFID in visual 3D models, but it primarily serves robotic applications and does not directly address the technical challenges in HAR.

The experimental setup of the *RFIDAR* system is centered on recognising activities such as sitting, standing, leaning, and walking within a specific area. This system is designed to demonstrate its effectiveness for patient monitoring in healthcare and accurate detection of daily activities. The *RFIDAR* system underscores the synergistic potential of combining RFID and radar sensing modalities, thereby boosting the reliability of HAR in indoor environment. This approach aligns with the recommendations of global public health agencies [158], highlighting its applicability in healthcare environments where precise indoor activity classification is crucial.

Multi-model Fusion Algorithms

This section focuses on the methods of feature fusion, which are crucial for combining data from different sensors. We will explore various algorithms used in this process. These include methods for manual feature extraction, which rely on probability and statistics, as well as techniques that utilise machine learning and deep learning for automatic feature representation. The structure and classification of these fusion algorithms are illustrated in Figure 2.7. Additionally,

there is a comprehensive comparison of the advantages and disadvantages of widely-used HAR methods, presented in Table 2.5.

Table 2.5: Pros and cons of information fusion algorithms.

Methods	Pros	Cons
HMM	Determine the most probable sequence of states in a model that results in a specific output sequence.	HMM are limited by their lack of memory, restricting them to use only the immediate previous state for context. To utilise more extensive contextual information, it's necessary to develop more advanced HMM models.
CNN	CNNs excel in handling local dependencies and are scale-invariant.	Requires a substantial dataset and extensive hyper-parameter tuning for optimal feature extraction
RNN	RNNs can learn complex nonlinear dynamical mappings, typical in nonlinear time series prediction.	Real-time activity prediction faces difficulties with substantial parameter updates.
LSTM	LSTM networks outperform other RNNs due to their design, which allows them to maintain aspects of their hidden state across extended time steps.	More parameters, easy to fit
GRU	The modified GRU in RNN resembles LSTM but is structurally simpler, leading to reduced learning time due to its fewer parameters. Additionally, it is capable of learning effectively from smaller datasets.	The LSTM outperforms the GRU in tasks requiring indefinite counting capabilities.

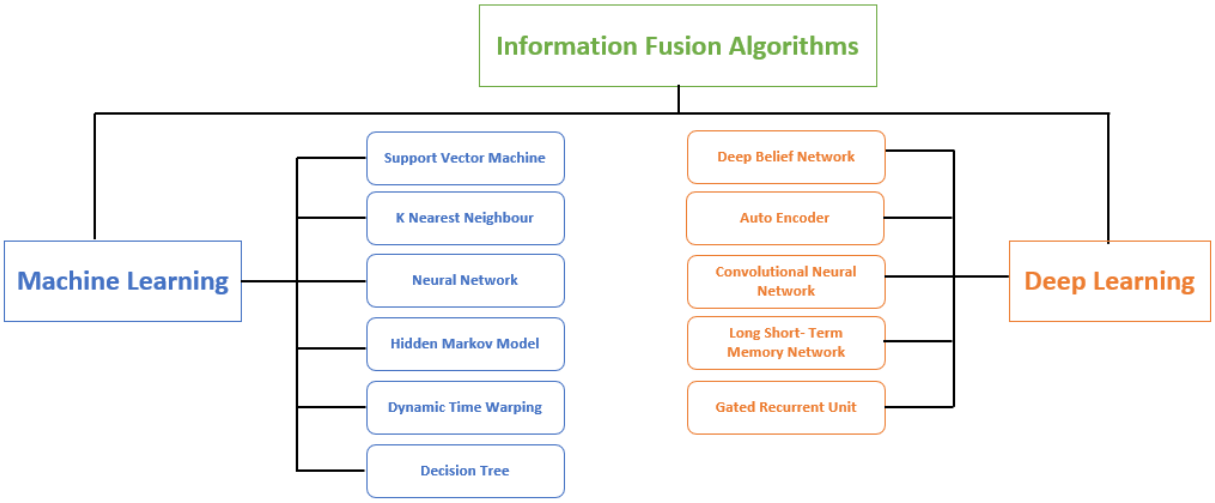


Figure 2.7: Multi-model sensor fusion information algorithm.

1. Machine Learning Algorithms

Machine learning techniques significantly improve the accuracy of activity detection. This improvement is achieved through a series of steps: data preprocessing, segmentation, feature extraction, dimensionality reduction, and the assessment of different learning algorithms [141]. These steps incorporate various methods including SVM [159, 160], KNN [160, 161], dictionary learning [161, 162], Naive Bayesian [159], decision trees [161, 163], ANN [164], hidden Markov models [165], linear discriminant classifiers [166], random forests [161, 71] and dynamic time regularization. For example, Guo et al. [162] created an algorithm for recognising aerobic exercises using multiple sensors and dictionary learning. Similarly, Spinsante et al. [167] utilised decision tree algorithms for monitoring activities to prevent sedentary behavior at workplaces. Dai et al. [168] introduced a wearable biosensor network for daily emotion recognition, employing technologies like EEG and pulse monitoring. They developed an unbalanced fuzzy SVM to effectively process datasets with various emotional states. Hidden Markov models, especially known for their application in speech recognition [169] and HAR [170], have been demonstrated to be effective in gait analysis. Additionally, Tahir et al. [171] showed that high accuracy could be achieved in signal variance measurement by extracting multifused models. Finally, Wu et al. [172] proposed an orientation-independent method for activity and gesture recognition, specifically tailored for low-power wearable sensors. This method employs DTW for segmentation and is notable for its high accuracy in real-life situations, irrespective of the sensor orientation. This approach illustrates the adaptability of machine learning to different sensor orientations, making it a viable option for practical applications. Experimental results indicated that this method achieved an average accuracy of 98.2% in daily life recognition and 95.6% in gesture recognition.

2. Deep Learning Algorithms (Relevant Literature)

Deep learning, an end-to-end learning method, begins with raw input data and independently manages both feature extraction and model learning with a multi-layered network structure. This approach bypasses the requirement for manually designed features, thus enabling the autonomous learning of effective features. It also facilitates the fitting of complex functions with fewer parameters and the creation of more intricate models. Deep learning's advantages lie in its capacity for feature expression, function modeling, and model generalization. These aspects highlight the differences and relationships between deep learning and conventional machine learning in human activity recognition.

The deep auto-encoder (DAE), focused on the unsupervised learning of efficient data encodings, consists of two parts: an encoder and a decoder. The encoder transforms input data into hidden features, and the decoder reconstructs these features back into approximate outputs, aiming to minimize error rates [173]. DAE has been applied in areas such as smartphone and health monitoring for feature representation [174]. Innovations in this field include a generative model for feature learning in motion detection [175], and a stacked auto-encoder for human

motion recognition, which surpasses traditional neural networks in accuracy [176]. Recent advancements in auto-encoder feature representation encompass stacked auto-encoders [177], deep denoising auto-encoders [178], and sparse auto-encoders, particularly in the context of human activity detection through feature fusion [173]. Sparse auto-encoders, which introduce sparse term loss functions, excel at representing high-dimensional feature vectors, providing high interpretability and implicit feature selection. DAEs also contribute to effective data representation, emphasising not only minimal reconstruction errors or sparsity but also robustness against partial data corruption. DDAEs improve coding robustness by incorporating noise and training with damaged input samples, thus capturing robust data invariants [179].

Convolutional neural networks (CNNs) [180] have been effective in fields like speech recognition [181], image classification [182], and video activity recognition [183]. Their superiority in HAR, particularly in classifying time series, arises from their local dependency and scale invariance. Novel applications include single-channel EEG-based drowsiness detection, real-time sEMG feature prediction, posture classification, human gesture recognition, fatigue level estimation, and EEG-based fatigue detection using a CNN model [184, 185]. Yang et al. [186] introduced a novel soft, wear-resistant smart shoe with sensors for learning sports intentions, employing a convolutional autoencoder for data fusion and feature extraction, combined with LSTM for comprehensive gait analysis.

Recurrent neural networks (RNNs), suitable for processing time series data like audio [187] and text [188], transfer information across hidden layer units to capture temporal relationships. Advances like LSTM networks and GRU networks address the problem of gradient vanishing in long sequences. LSTM, in particular, has shown leading results in various sequence processing tasks [189], with applications in gait analysis [190], anomaly detection [191], and more. The integration of LSTM with other models, such as Tran et al. [192] multi-modal network, enhances human activity recognition systems. Combining CNN and LSTM for feature extraction is increasingly popular, applied in multi-view settings, dual-channel LSTM and CNN models for various data types [193], and sensor-based activity recognition in sports [194, 195]. GRU networks, a simpler alternative to LSTM, have shown comparable efficacy in human activity recognition [196]. These networks, when combined with CNNs, are employed for automatic feature extraction and activity classification [197], gait type identification [198], and arm gesture recognition using deep convolution and recurrent networks [199].

2.2 Summary

The chapter starts by summarising HAR using contactless methods, with a special focus on TriSense and fusion technologies. WiFi-based HAR employs signal interaction techniques like the RSSI and CSI to monitor indoor activities, providing a non-intrusive way to understand human movements. It is noted that WiFi-based HAR prefers CSI over RSSI due to its higher

accuracy in detecting activities, due to CSI's capability to offer a detailed analysis of signal propagation. On the other hand, RFID-based HAR introduces contactless systems that do not require tagged objects, enhancing their usability across different settings. The combination of Radar and RFID technologies marks notable progress in HAR by integrating the strengths of both to improve detection accuracy and robustness in complex environments. However, this fusion brings forth challenges like system complexity, increased costs, and privacy issues that necessitate careful consideration for successful and ethical application.

Chapter 3

Advancing Localisation Techniques: The Role of Software-Defined Radios in Non-Invasive Approaches

This chapter explores the use of radio waves for non-invasive indoor human activity detection, a field with growing relevance in smart healthcare. It examines how activity localisation can enhance healthcare systems by accurately identifying patient locations. The research utilises CSI from radio frequencies, a method established as effective for non-invasively detecting human activities. Our experiments utilised an USRP devices to detect and localise the activities of a single individual, analysing CSI changes during various actions like sitting, standing, leaning, and walking forward and backward directions in a room. We also include scenarios with an empty monitored area for comprehensive analysis. An important aspect of this study is the application of artificial intelligence, particularly a Super Learner algorithm, which has shown a 96% accuracy rate in localising different activities, outperforming existing methods.

3.1 Introduction

Indoor localisation systems, aimed at estimating the location of entities within indoor environments, utilise technologies such as WiFi, UWB, Bluetooth, RFID, Infrared, inertial sensors, and cameras [200, 201, 202, 203]. Recently, these applications have broadened and extended to battlefield surveillance, disaster prediction, intelligent traffic, and indoor navigation [204, 205]. In healthcare monitoring, these systems are becoming essential and desirable, especially considering the growing elderly population [206, 207]. Indoor localisation is challenged by factors such as noise, signal fluctuation, and the presence of obstacles like furniture. Despite these challenges, significant advancements have been made due to technological developments in wireless communication, computing, and detection techniques. The field includes various approaches such as wearable device-based, context-aware, and contactless device-based systems. Wearable

devices, for instance, are known for their ability to recognise human activity without compromising privacy [208]. Context-aware systems, which utilise sensors like floor sensors and cameras, have limitations once the user leaves the monitored area and can raise privacy concerns, leading to legal restrictions in some countries [209].

The chapter further discusses how indoor localisation contrasts with outdoor environments, where advanced satellite systems like Beidou and GPS provide more precise location services. This difference is attributed to issues like weak satellite signals and low penetration indoors. In the field of indoor localisation, RF is often used due to the widespread application of low-power sensors. Researchers have explored technologies like UWB, WiFi, Bluetooth, audio, light, Zigbee, and RFID for effective indoor localisation [200, 201, 203]. This chapter specifically focuses on RF-based Wi-Fi sensing for indoor localisation, leveraging existing Wi-Fi infrastructures, and avoiding the need for additional sensing equipment. The discussion includes the use of CSI and RSSI for activity detection and localisation, with a preference for CSI due to its detailed data capture [210, 211]. The experiment conducted uses USRP devices to collect CSI data, aiming to localise human activities within a defined indoor environment.

The objective of this study is to use two USRP devices, one serving as a Tx and the other as the Rx to collect CSI data on a single human subject performing activities in different locations of a single room. The room is divided into three zones both horizontal and vertical directions (3x3 zones) shown in Figure 3.1. Each intersection point is referred to as a location. The six activities (empty room, no-activity, sitting, standing, leaning, walking forward and backward) are carried out in all nine locations. The amplitude changes observed in the CSI differentiate between activities performed in each location. This allows for CSI to be used in the localisation of a subject as the radio signals are affected differently in human movements occurring in different locations.

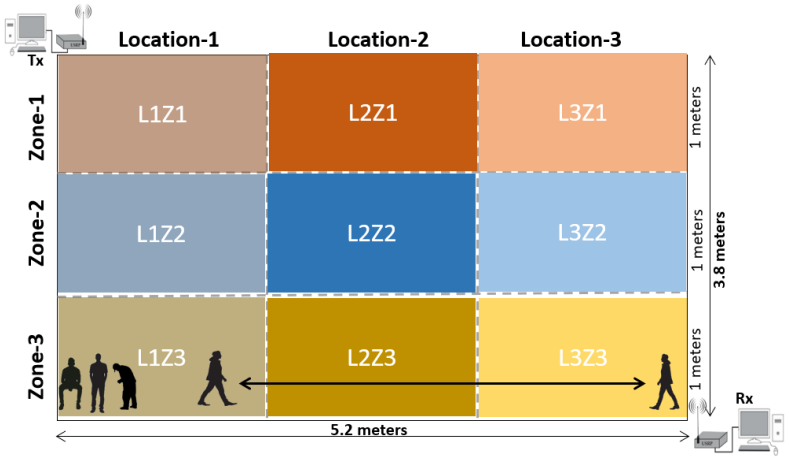


Figure 3.1: Experiment setup diagram showing the position of the horizontal and vertical zones.

3.2 Materials and Methods

This section covers materials and methods, as well as how data is collected using an experimental setup to produce various test cases before applying machine learning. The hardware and software subsection describes the components that are designed and used to collect CSI data depicting human activity from the sensing devices.

A. Hardware Specification

The hardware setup for data collection consists of two USRP devices, specifically national instrument (NI) $X310/X300$ models. These devices communicate with each other within their coverage area during the activity. They are connected to two PCs using 1G Ethernet cables. Each USRP is equipped with extended-bandwidth daughterboard slots, covering a range from DC to 6 GHz and supporting up to 120 MHz of baseband bandwidth. The PCs used in this setup are powered by Intel(R) Core(TM) i7-7700 3.60 GHz CPUs and equipped with 16 GB RAM. They run on an Ubuntu 16.04 virtual system as the operating system. For wireless communication, the USRPs are outfitted with VERT2450 omnidirectional antennas.

B. Software Specification

The USRP devices are configured with software to enable transmission from the transmitting USRP device to the receiving one. The GNU Radio software package, a free and open-source signal processing tool widely used in research, is utilised for configuring the software that facilitates USRP communication [212]. It offers examples of orthogonal frequency division multiplexing (OFDM) signal processing, which can be adapted for use with USRP devices, facilitating the extraction of CSI. GNU Radio allows the setting of various parameters for the USRP, such as the center frequency at 3.75 GHz, the use of 64 OFDM subcarriers, and gain levels set at 70dB for the transmitter (T_x) and 50dB for the receiver (R_x), chosen to optimise data transmission efficiency and signal robustness. The flow diagram created in GNU Radio is transformed into a Python script, which initiates OFDM communication on the USRP devices. This script outputs the CSI collected during the transmission, represented as complex numbers. To calculate the amplitude of these signals, the absolute value of these complex numbers is taken, prioritising amplitude data for its robustness and computational efficiency while phase is sensitive, which is essential for processing and machine learning applications. This CSI amplitude data is then converted into CSV files, which are compiled into datasets for training and testing machine-learning algorithms. The data flow diagram for data collection across the six classes of activities is illustrated in Figure 3.2.

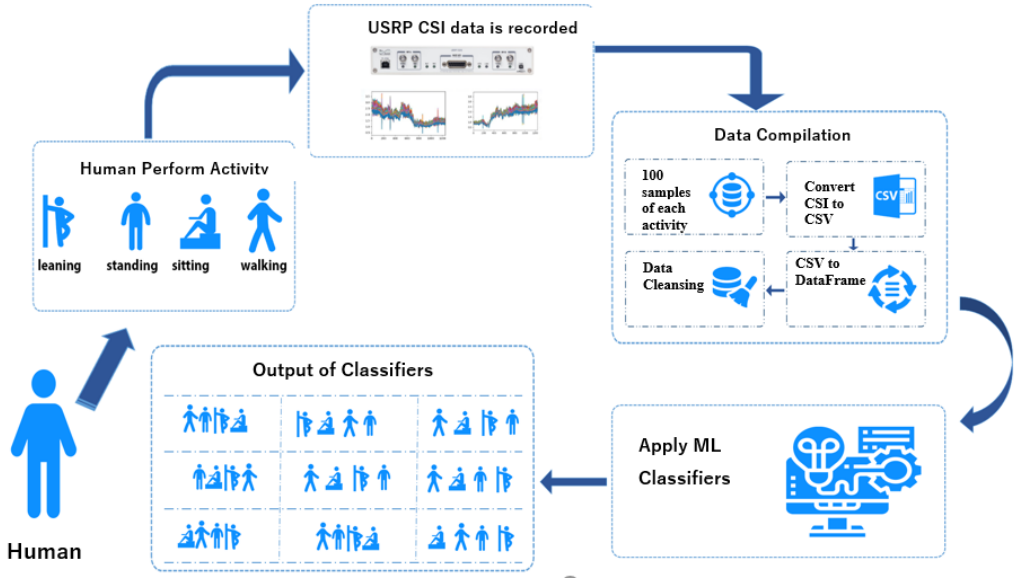


Figure 3.2: Data flow diagram showing the process of how the human movement is recorded as CSI and compiled into a dataset for ML classification

3.2.1 Experimental Setup

The experiments presented in this chapter are carried out in a $3.8 * 5.2\text{m}^2$ room in the James Watt South Building at the University of Glasgow, where an active and approved ethics application is in place. The experiment is conducted in an office setting where the room is divided into three zones and locations, both in the horizontal and vertical directions (see Figure 3.1). The three zones are separated by one meter, where all the activities take place. The T_x and R_x USRP devices are positioned in opposite corners, angled at 45° to each other, to capture all movement. The data is collected using the same single subject to perform the activities in each of the zones. This ensures that the only variables in the data collection are the activities and the location where the activity is performed.

3.2.2 Data Collection

This section explains the data collection process using the setup from section 3.2.1, essential for developing machine learning datasets. The experiment involves seven activities: sitting, standing, leaning, no activity, walking from T_x to R_x and vice versa, and an 'empty' classification. Data for the first four activities are collected from each zone shown in Figure 3.1, while the walking activities are observed as the subject travels diagonally across areas between the Tx and Rx devices. The 'empty' category involves data collection with no subject present.

Figure 3.3 illustrates CSI amplitude patterns for all six activities and the empty classification. In these visual representations, each color denotes a subcarrier during an activity, with the amplitude of the subcarrier on the y-axis and the packet count on the x-axis. Each data sample represents 3 seconds of OFDM communication, roughly equating to 1200 packets. In total, 100

samples are collected for each activity, culminating in 4300 data samples. This includes 100 samples each for sitting, standing, leaning, and no activity, collected from each area shown in Figure 3.1, amounting to 3600 samples (4 Activities x 100 samples x 9 areas). For the walking activities, 600 samples are collected (300 samples for each direction), considering the inclusion of 3 areas for each walking activity and the two different directions, resulting in 600 samples to represent 100 samples for each area in each direction. Additionally, 100 empty samples are gathered (1 activity x 100 samples). During the empty sample collection, the subject exits the room completely to ensure the integrity of these samples. This results in a total of 43 classifications, each with 100 samples. The samples are labelled according to their respective zones and locations, for instance, L2Z1 indicates a sample collected at location 2 and zone 1. Table 3.1 provides a comprehensive overview of the 43 classes and the total number of data samples collected for each location.

Table 3.1: Overview of data classes and their description.

S. No	Class	Class Description	No. of Classes	No. of Samples
1	Empty Activity	No human subject in the activity area	1	100
2	No Activity	No activity performed by human	9	900
3	Sitting	The action of "Sitting" at the designated location within Zone	9	900
4	Standing	The action of "Standing" at the designated location within Zone	9	900
5	Leaning	Leaning forward with the upper body at the designated location within Zone	9	900
6	Walking Rx-Tx and Tx-Rx	Walking from the USRP X310 Rx side to USRP X300 Tx side and vice versa	3*2	600

System Hypothesis

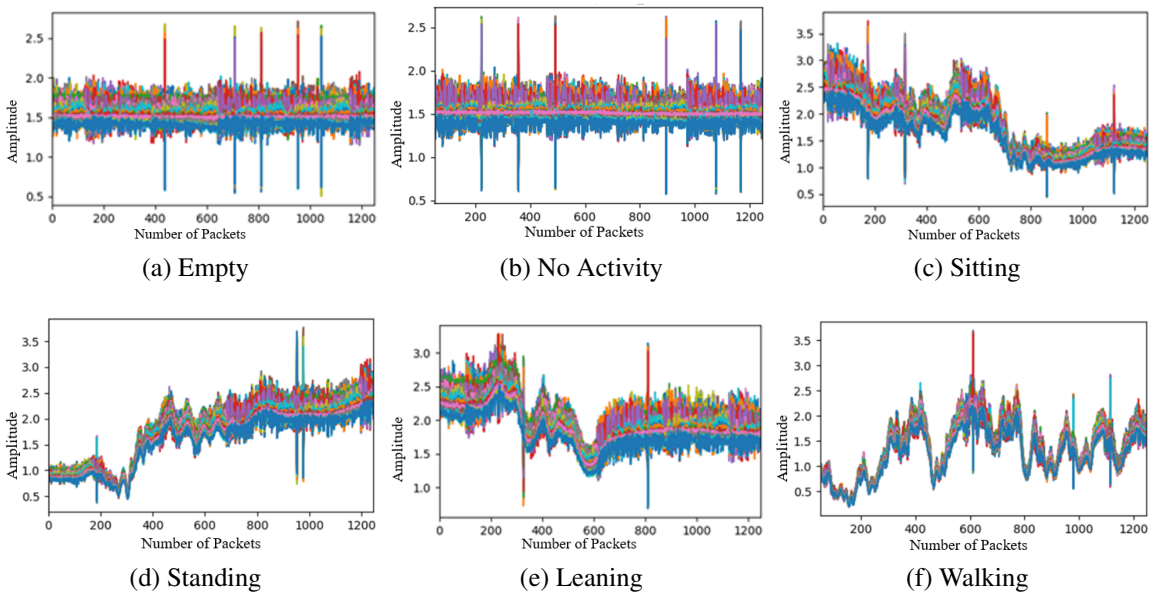
The hypothesis is summarised as follows:

1. The accuracy of determining the position of an activity within a room is expected to be 100% as movements are made either vertically or horizontally.

2. The hypothesis posits that the accuracy of detecting walking activity is higher compared to sitting or standing when movements occur vertically or horizontally.

3. It is anticipated that the horizontal and vertical distances between the transmitter (T_x) and receiver (R_x) will influence the accuracy of detecting the activity.

This study aims to evaluate the capability of AI in precisely recognising different activities and pinpointing their precise locations in a room through the analysis of RF signal data. It explores the effectiveness of AI in detecting and localising activities by interpreting the subtle variations in RF signals.



(a) Empty (b) No Activity (c) Sitting

(d) Standing (e) Leaning (f) Walking

Figure 3.3: CSI examples for all six activities showing the amplitude values for all 64 subcarriers (represented in each colour) in the OFDM communication

3.2.3 Test Cases

In this section, several test cases are presented, which are used to apply different ML approaches for activity localisation. In this study, A total of two test cases are presented based on the data collected as shown in Table 3.2.

The test cases for the data collection are reported in Table 3.2. The dataset description is given below.

- **L1L2Zone1**: The dataset contains data from locations 1 and 2 for Zone 1, with a total of 11 classes of activities from both locations.
- **L1L3Zone1**: The dataset contains data from locations 1 and 3 for Zone 1, with a total of 11 classes of activities from both locations.
- **L2L3Zone1**: The dataset contains data from locations 2 and 3 for Zone 1, with a total of 11 classes of activities from both locations.
- **L1L2L3Zone1**: The dataset contains data from locations 1, 2, and 3 for Zone 1, with a total of 15 classes of activities from three locations.

Table 3.2: Test cases of (Figure 3.1)

Test case	Dataset	Test case description
The relationship between the activity's location and the detection accuracy		
Test-1.1 Test-1.2 Test-1.3 Test-1.4	**Zone 1** (L1L2Zone1, L1L3Zone1, L2L3Zone1) **Zone 2** (L1L2Zone2, L1L3Zone2, L2L3Zone2) **Zone 3** (L1L2Zone3, L1L3Zone3, L2L3Zone3) **Location wise combination** Location1-Z1Z2Z3, Location2-Z1Z2Z3, Location3-Z1Z2Z3)	To check the co-relation between zones and locations when we move horizontally or vertically from Tx towards Rx
To check localisation overall accuracy		
Test-2.1	Combined location and Zones ($L_1L_2L_3 - Zone_{123}$)	To check the localisation accuracy of all the activities within Room

- **Location1-Z1Z2Z3** The dataset contains data from Zones 1, 2, and 3 for location 1, with a total of 19 classes of activities from three locations.
- **L1L2L3-Zone123** The dataset contains data from all areas, with a total of 43 classes of activities from nine locations.

3.2.4 Data Pre-processing and Machine Learning

This section provides an overview of the data preprocessing and ML approaches that have been designed and implemented.

Data Preprocessing

In this chapter, we use Scikit, a popular Python data analysis toolkit [213], and Pandas, a Python library for parsing CSV files and converting them into data frames for analysis with Scikit-learn [73]. Data frames are labeled, and we address not a number (NaN) values, resulting from packet mismatches between USRP devices, by replacing them with the row mean using Scikit's 'SimpleImputer' function. This data cleansing step ensures the integrity of the overall data pattern before applying ML algorithms.

Machine Learning

The evaluation of the proposed Indoor localisation system for HAR is conducted using seven different ML algorithms, focusing on the accuracy of localising various human activities. In our experiment, we measure the performance of each algorithm using the accuracy metric, calculated independently for each test case dataset. To ensure a comprehensive analysis, we employ two approaches: (i) k-fold cross-validation and (ii) the train-test split method.

K-fold cross-validation, a widely used technique in ML, involves dividing the dataset into k groups for testing the efficacy of the ML approach [71]. In this experiment, we set k to 10, meaning the dataset is split into 10 groups. Each group serves once as the testing set while the remaining 9 groups are combined to form the training set. This process is repeated for each group, and the classification results from all groups are used to evaluate the performance across the entire dataset. The train-test split technique, on the other hand, divides the dataset into separate training and testing sets. Here, the ML model is trained using the training data and then makes predictions on the testing data. This method allows the algorithm to apply what it has learned from the training data to unseen data. In our study, we allocate 80% of each dataset for training purposes and reserve the remaining 20% for testing. The specific parameters used to configure the algorithms are detailed in Table 3.3.

Table 3.3: The parameters of machine learning algorithms.

Algorithm	**Hyper-parameters**	**N estimator**
Support Vector Machine	Kernel = rbf and sigmoid	gamma='scale'
K-Nearest Neighbors	Euclidean distance and K = 3,7	n-repeat = 3
Bagging	max-features, default= 1.0	n-estimators=20
Random Forest	max-features: ['auto', 'sqrt']	n-estimators=20
Extra Trees	max-features = auto, sqrt	n-estimators=20
Super Learner	multi-threading	n-estimators=20

3.3 Results and Discussion

This section presents and discusses the results of the test cases shown in Table 3.2.

3.3.1 Test-1.1

The results of Test-1.1 are shown in the below Table 3.4. These results show the relationship between localising and identifying activities across the horizontal zones of:

- L1Zone1 and L2Zone1
- L1Zone1 and L3Zone1
- L2Zone1 and L3Zone1

In Test-1.1, the super learner (SL) algorithm consistently demonstrated the highest accuracy across all three experiments. Notably, the experiments involving L1Zone1 and L3Zone1 showed the most impressive accuracy scores with all the algorithms, surpassing the results from L1Zone1 and L2Zone1, as well as L2Zone1 and L3Zone1. The SL algorithm, in particular, achieved a remarkable accuracy score of 95.90% in the L1Zone1 and L3Zone1 experiments.

Table 3.4: ML algorithms comparison using Cross-validation on test case 1.1 in (Table 3.2)

Algorithm	L1L2Zone1 Accuracy	L1L3Zone1 Accuracy	L2L3Zone1 Accuracy
Multilayer Perceptron	66.87%	73.42%	72.70%
Support Vector Machine	74.28%	84.71%	79.40%
K- Neighbors Classifier	78.70%	84.95%	82.25%
Bagging Classifier	79.56%	86.38%	82.68%
Random Forest	80.95%	87.60%	84.44%
Extra Trees	86.00%	92.93%	89.08%
Super Learner	**87.27%**	**95.90%**	**91.12%**

These findings suggest that the ability of the algorithms to distinguish between different locations improves when there is a greater spatial separation between them. This improvement is likely attributed to the more significant fluctuations in CSI as the distance from the transmitter increases. In contrast, locations situated closer to each other exhibit less variation in CSI fluctuations, yet the algorithms, especially the SL algorithm, still attain high accuracy levels. This indicates the effectiveness of the SL algorithm and others in accurately detecting and differentiating activities in various indoor locations, despite the proximity of these locations to each other.

3.3.2 Test-1.2

The results of Test-1.2 are shown in the below Table 3.5. These results show the relationship between localising and identifying activities across the horizontal zones of:

- L1Zone2 and L2Zone2
- L1Zone2 and L3Zone2
- L2Zone2 and L3Zone2

Table 3.5: ML algorithms comparison using cross-validation on test case 1.2 in (Table 3.2)

Algorithm	L1L2Zone2 Accuracy	L1L3Zone2 Accuracy	L2L3Zone2 Accuracy
Multilayer Perceptron	68.76%	87.59%	88.17%
Support Vector Machine	82.92%	82.22%	92.93%
K- Neighbors Classifier	85.98%	85.56%	90.84%
Bagging Classifier	85.58%	87.77%	90.47%
Random Forest	86.54%	87.81%	91.87%
Extra Trees	91.69 %	91.60%	95.07%
Super Learner	**95.54**	**94.23%**	**96.36**

The results from Test-1.2 demonstrated the most accurate outcomes in the experiment for L2Zone2 and L3Zone2. In these specific trials, the SL algorithm outperformed other methods, achieving the highest accuracy rate of 96.36% in both L2Zone2 and L3Zone2. These results suggest that as the subject moves farther from the transmitter, the fluctuations in CSI become more pronounced in the nearby locations. This pattern highlights the effectiveness of the SL algorithm in accurately detecting position changes based on CSI variations, especially in scenarios where the subject is at a greater distance from the transmitter.

3.3.3 Test-1.3

The results of Test-1.3 are shown in the below Table 3.6. These results show the relationship between localising and identifying activities across the horizontal zones of:

- L1Zone3 and L2Zone3
- L1Zone3 and L3Zone3
- L2Zone3 and L3Zone3

Table 3.6: ML algorithms comparison using cross-validation on test case 1.3 in (Table 3.2)

Algorithm	L1L2Zone3 Accuracy	L1L3Zone3 Accuracy	L2L3Zone3 Accuracy
Multilayer Perceptron	56.65%	71.88%	77.91%
Support Vector Machine	74.61%	82.58%	85.62%
K- Neighbors Classifier	77.61%	84.16%	84.64%
Bagging Classifier	80.37%	84.68%	85.65%
Random Forest	81.33%	86.21%	88.35%
Extra Trees	82.27 %	87.22%	90.45%
Super Learner	**85.00%**	**92.27%**	**93.18%**

The results from Test-1.3 align with those of Test-1.2, indicating that as the subject moves further from the transmitter, the CSI fluctuations become more pronounced in adjacent locations. This pattern is evident in the L2Zone3 and L3Zone3 experiments. In these tests, all algorithms demonstrated improved accuracy, with the SL algorithm achieving the highest accuracy of 93.18%, surpassing the performance of other algorithms.

3.3.4 Test-1.4

The cross-validation results of Test-1.4 are shown in the below Table 3.7. These results show the relationship between localising and identifying activities across the vertical zones of:

- Location1Z1, Location1Z2 and Location1Z3

- Location2Z1, Location2Z2 and Location2Z3
- Location3Z1, Location3Z2 and Location3Z3

Table 3.7: ML algorithms comparison using cross-validation on test case 1.4 in (Table 3.2)

Algorithm	Location1-Z1Z2Z3 Accuracy	Location2-Z1Z2Z3 Accuracy	Location3-Z1Z2Z3 Accuracy
Multilayer Perceptron	47.07%	64.46%	82.81%
Support Vector Machine	66.66%	75.45%	91.08 %
K- Neighbors Classifier	66.66%	69.00 %	91.08 %
Bagging Classifier	71.66 %	76.65%	89.46%
Random Forest	75.66 %	78.66%	91.14%
Extra Trees	77.00 %	83.36%	93.66%
Super Learner	**81.66 %**	**86.66%**	**96.66%**

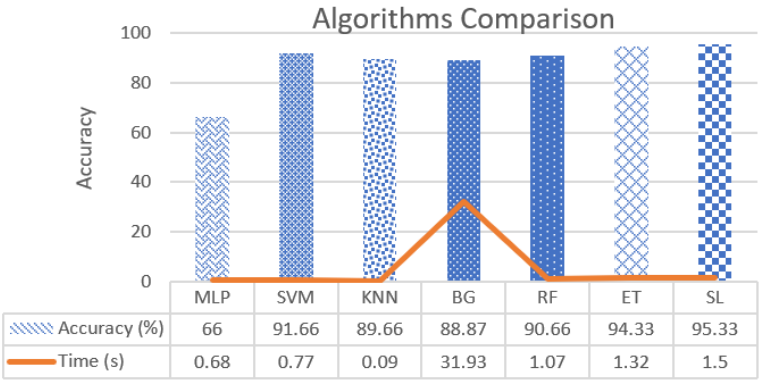
The results from Test-1.4 provide further evidence supporting increased accuracy as the distance between the subject and the transmitter grows, corroborating the results observed in Test-1.2 and Test-1.3. Notably, the experiments conducted at Location3Z1, Location3Z2, and Location3Z3 during Test-1.4, as well as in the previous tests of Test-1.2 and Test-1.3, demonstrated the highest accuracy levels. In these tests, the SL algorithm consistently outperformed other algorithms, achieving an accuracy of 96.66%. Consequently, the dataset from Location3Z1, Location3Z2, and Location3Z3 has been selected for further analysis using the train-test split technique. The results obtained through this technique are presented in Table 3.8.

Table 3.8: ML algorithms comparison using train test on test case 1.4 in Table 3.2

Algorithm	Location3-Z1Z2Z3 Accuracy	Precision	Recall	F1 score	Time (sec)
Multilayer Perceptron	66.00%	67.00%	67.00%	68.00 %	0.68
Support Vector Machine	91.66%	92.00%	91.00 %	91.00 %	0.77
K- Neighbors Classifier	89.66%	91.00 %	90.00 %	90.00 %	0.09
Bagging Classifier	88.87 %	89.00%	89.00%	88.00%	31.93
Random Forest	90.66 %	91.00 %	91.00 %	91.00 %	1.07
Extra Trees	94.33 %	95.00%	94.00 %	94.00 %	1.32
Super Learner	**95.33 %**	**95.00%**	**95.00 %**	**95.00 %**	1.50

In the train-test split results, the SL algorithm emerged as the superior performer among all algorithms, achieving an impressive accuracy of 95.33%. This was closely followed by the Extra Trees (ET) classifier, which secured an accuracy of 94.33%. The Support Vector Machine (SVM) and Random Forest algorithms also showed commendable results with accuracies of 91.66% and 90.66% respectively. As depicted in Figure 3.4, Bagged Trees and the K-Nearest

Neighbors (KNN) algorithm recorded the lowest accuracies, with scores of 88.87% and 89.66% respectively. Notably, the Multi-Layer Perceptron (MLP) demonstrated the least satisfactory performance, achieving an accuracy of only 66%.



	MLP	SVM	KNN	BG	RF	ET	SL
Accuracy (%)	66	91.66	89.66	88.87	90.66	94.33	95.33
Time (s)	0.68	0.77	0.09	31.93	1.07	1.32	1.5

Figure 3.4: Accuracy and time comparison of the tested machine learning algorithms on test case 1.4 in Table 3.2.

Figure 3.5 shows the confusion matrix of the results achieved using the train-test split technique for the SL algorithm.

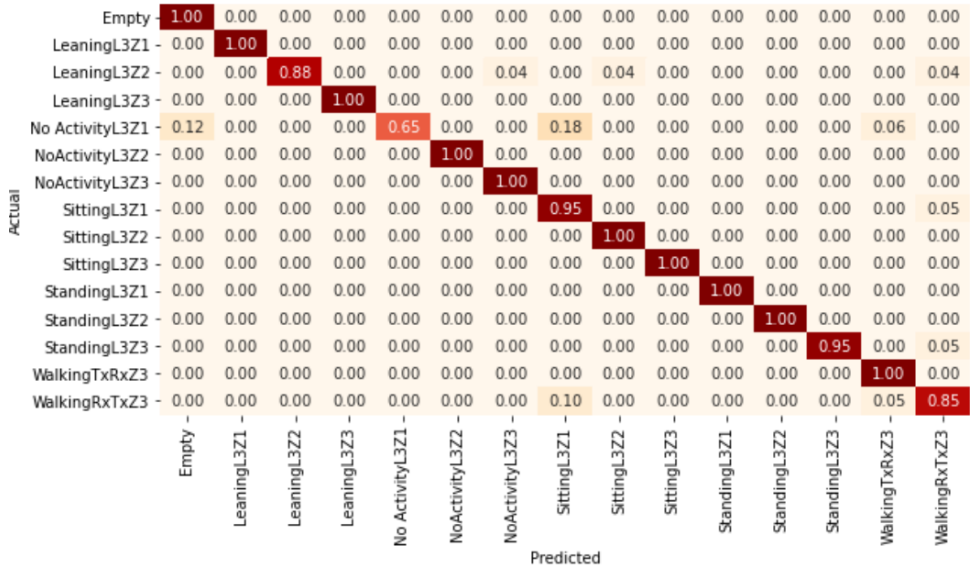


Figure 3.5: The confusion matrix of the SL algorithm on test case 1.4 Location3-Z1Z2Z3 (see Table 3.2) shows how the algorithm classified each sample of data.

3.3.5 Test-2.1

In this chapter, the main contributions are showcased in the results of Test-2.1. The initial experiment encompasses data from all locations and zones. To provide a comprehensive comparison, two additional experiments were conducted. The first of these focuses on all zones within locations 1 and 3, while the second examines all zones in locations 2 and 3. The accuracy of

each algorithm used in these experiments is detailed in Table 3.9, offering a clear view of their performance across different scenarios.

Table 3.9: ML algorithms comparison using cross-validation on test case 2.1 in (Table 3.2)

Algorithm	Accuracy $L_1L_2L_3 - Zone_{123}$	Accuracy $L_1L3 - Zone_{123}$	Accuracy $L_2L_3 - Zone_{123}$
Multilayer Perceptron	57.44%	68.00%	71.00%
Support Vector Machine	62.33%	70.15%	76.97 %
K- Neighbors Classifier	63.66%	71.66%	72.57%
Bagging Classifier	68.66%	75.90%	76.66%
Random Forest	68.15%	77.12%	75.60%
Extra Trees	76.11%	82.27 %	81.97%
Super Learner	**79.00%**	**85.60%**	**85.60%**

The SL algorithm demonstrated a high proficiency in identifying activities and locations with a 79% accuracy rate, outperforming other experiments. This superior performance is attributed to its integration of multiple classifiers, including SVM, KNN, Bagged Trees, Random Forest, and Extra Trees. This result aligns with the outcomes observed in previous test cases. Additionally, the SL algorithm maintained its status as the top performer in other experiments conducted in Test-2.1.

3.4 Summary

This chapter presents an indoor localization system that employs RF sensing technology to classify seven distinct activities within a room. The system is engineered to accurately locate activities, identify the nature of these activities, and determine room occupancy achieved through the use of non-contact RF sensing methods, thus eliminating the need for wearable devices. A significant highlight of this study is the exceptional performance of the super learner classifier, an ensemble of multiple algorithms, demonstrating a notable increase in activity detection accuracy with the distance from the transmitter. Specifically, the findings include a 3% improvement in horizontal accuracy and a 14% enhancement in vertical accuracy per meter. Further study linked to research that underscores the benefits of utilising Wi-Fi CSI over traditional vision-based systems, which are often invasive to privacy. The referenced study emphasizes Wi-Fi CSI's cost-effectiveness and seamless integration with existing infrastructures. Through experiments involving the collection of CSI data via USRP, the research evaluates different deep learning models such as CNN, LSTM networks, and a hybrid LSTM-CNN approach, in the context of accurate activity detection.

Chapter 4

Advances in Human Activity Recognition: A Deep Learning Approach with Flexible and Scalable Software-Defined Radio

This chapter explores the growing field of ambient computing, particularly its application in healthcare technology advancements. It focuses on the use of Wi-Fi CSI for non-intrusive human activity recognition in an indoor environment. The study highlights the advantages of Wi-Fi CSI over traditional, privacy-invasive vision-based systems, emphasizing its cost-effectiveness and ease of integration with existing infrastructure. Through experiments using universal software-defined radio to collect CSI data, the research assesses various deep learning approaches, including CNN, LSTM, and a hybrid model, for accurate activity detection. LSTM is identified as the most accurate model, promising future exploration in multi-user environments and dynamic settings.

4.1 Introduction

In recent years, there has been a growing interest in human activity recognition (HAR) due to the increasing need to monitor human behavior and activities indoors. This surge has led to various applications, including assisting individuals with health conditions, sports, augmented reality treatments, and a focus on preventing falls among the elderly and disabled. HAR systems can be categorized into three main types based on the sensors used to collect data on human behavior: radio frequency (RF)-based systems [23], wearable sensor-based systems [214], and vision-based systems [215]. RF-based approaches have gained attention for capturing changes caused by human activity. These systems work on the principle that RF signals bounce off human bodies and cause fluctuations in ambient RF signals. Unlike vision and wearable sensor-based systems, RF-based systems don't require users to wear sensors and address privacy concerns. They are especially useful for recognizing indoor activities in healthcare [23, 216]. Radar-based HAR

systems offer high bandwidth and precise spatial resolution, making them suitable for detecting fine-grained human activities [217]. However, their expensive hardware limits their widespread use. On the other hand, HAR systems using Wi-Fi don't require specialized hardware and can be easily integrated into existing Wi-Fi setups [218].

The widespread use of Wi-Fi routers at home has made it possible to use the received signal strength indicator (RSSI) and channel state information (CSI) for environmental sensing. This study explores the potential of using Wi-Fi CSI with deep learning (DL) for HAR. A Wi-Fi CSI and DL-based system could enhance security by detecting individuals in complete darkness, identifying falls in elderly individuals, recognizing suspicious activities, and triggering timely assistance. However, RSSI is unstable, varies between individuals [15], and can't detect signal fluctuations caused by human activities when users aren't precisely positioned between an access point and a Wi-Fi router. In contrast, CSI provides detailed data [219], including amplitude and phase distortions at different frequencies. It can identify human activity by analyzing amplitude fluctuations in RF signals within the Wi-Fi context [220]. This chapter aims to collect CSI data from a single person performing various activities in different parts of a room. Two universal software radio peripheral (USRP) devices, one acting as a transmitter (T_x) and the other as a receiver (R_x), are used for this purpose. Changes in CSI amplitude can distinguish between activities, enabling the recognition of human activity based on RF signal changes. DL algorithms, including long short-term memory (LSTM), convolutional neural networks (CNN), and LSTM-CNN hybrids, are employed to classify six different activities in a single room. An additional class is introduced to detect an empty room. This study's contributions lie in the effective use of DL algorithms to accurately identify six distinct activities in an indoor environment based on CSI data collected from USRP devices.

4.2 Data and Methods

This section presents the experimental setup used to collect CSI data from contactless sensing devices for HAR. Additionally, it includes information on the techniques employed for data collection and pre-processing.

4.2.1 Data Collection

The data collection in this experiment involved two USRP devices positioned at opposite corners of a room, as shown in Figure 4.1, to transmit Wi-Fi signals and receive CSI. A participant performed various activities for data capture. The transmitting device (T_x) operated at 3.75 GHz with 64 OFDM sub-carriers at a 70dB gain and the receiver (R_x) at 50 dB. Data capture was managed through a GNU radio flow diagram converted to a Python script. Each data packet contained T_x, R_x, and CSI information. Figure 3.3 depicts six activities and an empty class, analyzed through CSI amplitude, with each sub-carrier color-coded for activity identification.

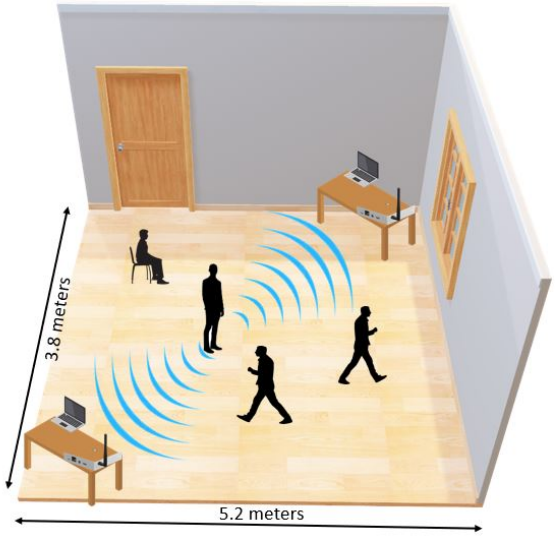


Figure 4.1: Experiment setup diagram.

The study collected 700 samples, each a three-second OFDM transmission with 1200 packets, distributing 100 samples across each activity.

In the preprocessing step of our study, we organised data to facilitate algorithmic processing, replacing missing values 'NaN' with the mean to maintain data integrity. Using Python's 'scikit' library and 'pandas' package, we processed raw data from CSV files and analysed them with 'Scikit-learn'. Dimensionality reduction was achieved using PCA's 'fit-transform' method, complemented by feature selection through 'selectKBest' and 'ANOVA f-test'. We applied a Butterworth filter using the 'butter(1, 0.05)' function for signal processing. Label encoding was implemented on the categorical values in the first column of data frames for efficient integration in Deep Learning algorithms. The output, CSI data represented as complex numbers, was processed to extract amplitude information and converted into CSV files for algorithm training and testing. This entire process, including the data flow for the seven activity classes, is depicted in Figure 4.2.

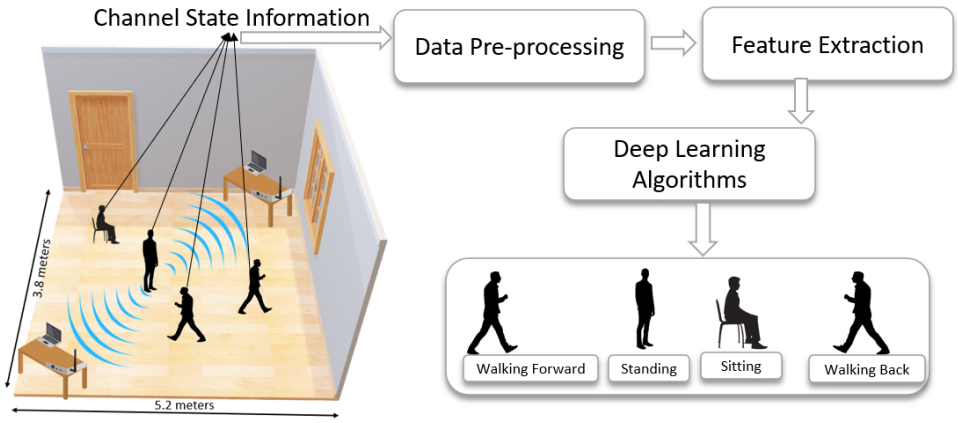


Figure 4.2: Data flow diagram depicting human activity captured as CSI & compiled into a dataset for classification.

4.2.2 Human Activity using Deep Neural Networks

Deep learning methods are increasingly used to automatically extract features from CSI impacted by various activities. Although multi-layer neural networks enhance classification accuracy, they risk overfitting and performance decline with limited data. Traditional solutions like adjusting learning rates, reducing batch sizes, and weight decay often fall short in addressing these issues. Consequently, WiFi-based systems require a tailored number of neural layers for optimal performance. Our study employed DL models optimised for small datasets, offering both effectiveness and reduced processing time.

Human Activity using LSTM

This chapter employed LSTM networks, a type of artificial RNN, recognized for their efficacy in handling time-series data, as suitable for our context per [221, 222]. The LSTM's primary role was to extract CSI values and reduce noise. We fed the raw CSI amplitude from each of the 64 subcarriers into a 64-dimensional feature vector. The network's hidden layer was configured with dimensions $(20, 50)$ and utilized the 'tanh' activation function. Optimisation was achieved using the Adam optimizer, with parameters set at a batch size of 64, a learning rate of 0.01, and a decay rate of $1e^{-6}$. A key advantage of using LSTM in our classification task was its ability to directly learn from raw series data, obviating the need for manual feature engineering and providing greater flexibility for domain experts.

Human Activity using CNN

CNN is one of the most popular DL architectures due to its ability to automatically extract deep, high-dimensional features as compared to just a few shallow ones [223, 224]. We contend that CNN and LSTM are compatible DL techniques based on the work [225]. Although LSTM is better suited for time-domain analysis and responds better to short-duration movement, CNN focuses primarily on changes in the frequency domain and has a greater reaction to long-duration movement. The size of the hidden layers was maintained at $(20, 32)$, the maximum pooling size was set at $(3, 1)$, and the activation is 'tanh'. At this stage, we can make the final classification prediction in the output layer by adopting the general CNN design principle.

Human Activity using LSTM-CNN

The LSTM, a variant of RNNs, excels in handling continuous temporal relationships, making it suitable for time series processing. Meanwhile, CNNs are adept at reducing frequency domain changes and extracting spatial features, as detailed in [226]. In our approach, we integrate the advantages of both LSTM and CNN into a single LSTM-CNN model. This study highlights that various LSTM configurations yield differing results with CSI data as input. We compare the LSTM and LSTM-CNN models specifically for human activity recognition. The LSTM

Table 4.1: Hyper-parameters of deep learning algorithms

S.No	Algorithms	Hyper Parameters
1	LSTM	optimizer='adam', Activation = tanh, lr=(0.1,0.01), decay= $1e^{-6}$, epochs=(20,50), hidden-layer-size=(20,50), Dropout=0.2 , batch-size=32, connected-layer-activation:'softmax'
2	CNN	optimizer='adam', Activation = tanh, lr=(0.1,0.01), loss = 'binary-crossentropy', hidden-layer-size=(20,32), epochs=(20,50), batch-size=32, Max-pooling-size = (3,1), connected-layer-activation:'softmax'
3	LSTM-CNN	optimizer='adam', Activation = tanh, lr=(0.1,0.01), loss = 'binary-crossentropy', epochs=(20,50), LSTM hidden-layer-size= (20,50), CNN hidden-layer-size= (50,32), Max-pooling-size = (3,1)

layer first processes the input, using signal fluctuation changes to gather data and extract time-domain characteristics from the initial signal. This is followed by the application of a 1D-CNN layer to the LSTM's output. Through convolution, this layer extracts high-dimensional implicit features. These features are then refined by a maximum pooling layer, which helps in forming an optimal feature sequence that serves as the basis for final classification. The LSTM-CNN model's specifications, including LSTM hidden layer sizes $(20, 50)$, CNN hidden layer sizes $(50, 32)$, 1D max-pooling sizes $(3, 1)$, and activation functions 'tanh' and 'softmax', are detailed in Table 4.1.

4.3 Result

4.3.1 Deep Learning Parameters and Evaluation Metrics

In this section, we evaluate the model's performance for multi-classifications using several metrics, including accuracy, loss, and confusion matrices. In the experiment, 700 samples of data were collected for training and testing, comprising two groups of static data (Empty and No activity) and 500 samples of dynamic data (sitting, standing, leaning forward, and walking in two directions).

The three important parameters in DL are the number of *epochs*, which dictates how many times the model processes the entire training dataset; the *batch size*, which determines how much training data to process before updating the model's internal parameters; and the *learning rate*, which governs the amount of change in the model weights following an estimated error update. Our experimental results indicate that modifying the batch size to 32 and reducing the vector learning rate to 0.01 gives a noteworthy enhancement in the network's classification performance.

Our results show that vector selection is important to network performance and that hyper-parameter optimization can increase classification accuracy. Table 4.2 shows that after 50 iterations of varying experimental conditions, the rate of change shows minimal variation while keeping a training size of 0.80. The Keras DL library supports the 'scikit-learn' module's *make_multilabel_classification*() function and utilises the 'ReLU' and 'softmax' activation functions in the hidden layer.

Table 4.2: Classification accuracy of deep learning algorithms

S.No	**Algorithms**	**Accuracy**	**F1 Score**	**Precision**	**Recall**
1	LSTM	95.3%	0.78	0.82	0.74
2	LSTM-CNN	91.0%	0.60	0.85	0.46
2	CNN	86.5%	0.20	0.40	0.14

4.3.2 Comparison with other Models

To evaluate the classification performance, we compared the LSTM network with that of the LSTM-CNN and CNN networks. Both of them use the TensorFlow library in the Python environment, and the number of LSTM hidden layers they employ matches the input data. The LSTM network has a better classification result, as demonstrated in Figure 4.3. The average loss value for multi-classification is 12.5%, the average accuracy rate is 95%, and all indices outperform LSTM-CNN which has 91%, and CNN which has 86%. The average loss values for LSTM-CNN and CNN are 20.3% and 27.3% respectively. The learning rate and epochs are slightly changed, leading to accuracy degradation, as shown in Table 4.3. In conclusion, compared to other algorithms in the same family, the suggested LSTM network provides a straightforward, effective, and highly accurate approach for recognizing seven different activities.

Table 4.3: Learning rate vs. epochs.

epochs= 50	**learning rate = 0.01**		**learning rate = 0.1**	
Algorithms	Accuracy	Loss	Accuracy	Loss
LSTM	95.3%	12.5	93.4%	14.7
LSTM-CNN	91.0%	20.3	75%	1.85
CNN	86.5%	27.3	77%	1.59
epochs= 20	**learning rate = 0.01**		**learning rate = 0.1**	
Algorithms	Accuracy	Loss	Accuracy	Loss
LSTM	93.5%	11	89.4%	23
LSTM-CNN	88%	23	74%	1.38
CNN	84%	31	74%	1.61

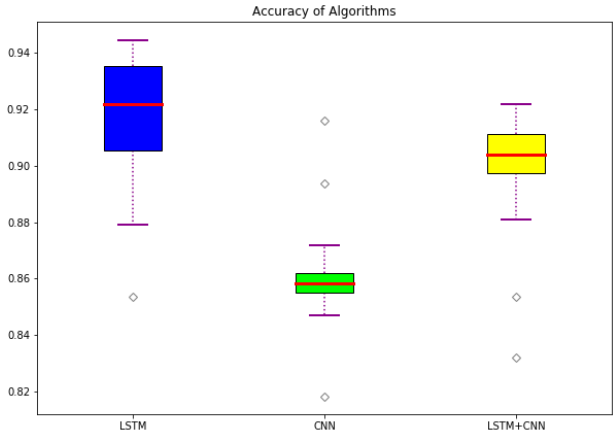


Figure 4.3: Comparison of deep learning Algorithms.

4.3.3 Discussion

The confusion matrix of the final results of the LSTM, CNN, and LSTM-CNN models on multi-classifications is shown in Figure 4.4. The classification performance generated by the "empty" and "standing" activities is optimal. However, the classification performance of "walking" is poor. It is evident that certain "walking" actions are subdivided into "forward and backward," most likely because the two signals don't have clear classification criteria and have more comparable amplitude and frequency. Overall, the classification accuracy rate remains steady at above 95%, indicating a positive classification performance.

4.4 Summary

In this chapter, we have conducted a comprehensive study comparing the effectiveness of LSTM, CNN, and LSTM-CNN network models in radio frequency-based indoor human activity recognition using channel state information data. The focus of our research was to recognize seven distinct activities within a single-room environment. Our results demonstrate that radio frequency sensing serves as a reliable and contactless method for human activity recognition. Key steps in our methodology included preprocessing the data, extracting relevant features, and implementing a robust system for CSI data collection. Among the models tested, the LSTM model exhibited superior classification accuracy. It was particularly effective in extracting hidden features from the CSI data, thereby enhancing the performance of activity recognition. These results underscore the significance of our approach in advancing the field of human activity recognition, particularly highlighting the potential of radio frequency sensing as a tool for indoor activity recognition.

(a) Normalised confusion matrix of LSTM (b) Normalised confusion matrix of CNN

(c) Normalised confusion matrix of LSTM+CNN

Figure 4.4: The deep learning model's normalized confusion matrix for multi-classification.

Chapter 5

AI-Enhanced Transparent RFID Tag (TRT-Wall) for Assisted Living

This chapter discusses the challenges faced by current assisted living technologies, which are often seen as complex, expensive, and intrusive. These characteristics can lead to a lack of practicality and acceptance among users. However, with the development of artificial intelligence and advances in wireless technology, there's a great chance to make assisted living systems better. Such improvements could help reduce healthcare costs and the need for hospital stays by more efficiently spotting, tracking, and pinpointing dangerous activities, allowing for quick action in emergencies. To address these issues, we present an innovative system called the Transparent RFID Tag Wall (TRT-Wall). This new system uses a passive ultra-high-frequency (UHF) RFID tag array along with deep learning methods to monitor human activities in a contactless manner. The TRT-Wall has undergone thorough testing for five different activities: sitting, standing, walking in both directions and being inactive. The experiments have shown very promising results, with the TRT-Wall being able to identify these activities with an impressive average accuracy of 95.6%. These findings suggest that the TRT-Wall, as a non-intrusive assisted living system, offers significant promise for improving assisted living conditions for the elderly.

5.1 Introduction

Human activity recognition plays an important role in enabling remote health monitoring for elders desiring to live independently at home. With an aging population, there's a growing need for assistance among individuals seeking to preserve their autonomy, as highlighted by United Nations estimates [227]. These estimates predict a decrease in the ratio of individuals aged 15 to 64 to those over 65 from 7:1 in 2020 to 4:1 by 2050, indicating a global elderly population of approximately 2 billion by 2050 [228]. This demographic trend suggests a looming workforce shortage in elderly care, emphasizing the importance of AAL research [229]. AAL incorporates various technologies to support caregivers, addressing challenges like limited mobility, chronic

disease monitoring, reducing social isolation, and managing medication services [230]. The collected medical data from these technologies is essential in the context of the growing demand for technology-driven healthcare solutions, particularly for disabled patients in indoor environments [231].

In recent years, human activity has been recognized using camera systems or on-body sensors such as infrared, accelerometers/gyroscopes, and FMCW radar [232, 233]. However, these systems face challenges [234, 235]. For example, camera-based monitoring can encounter issues like occlusion, restricted perspective, low lighting and frame resolution, and high computational demands for video processing. Privacy is also a major concern with these systems, although studies indicate a willingness among the elderly to trade some privacy for increased autonomy [234]. Moreover, wearable sensors can be burdensome during activities like sleep or physical exercise, and there's a risk of users forgetting or losing interest in wearing them [113]. As an alternative, contactless (tag-free) sensors using TriSense technologies offer benefits such as enhanced privacy and better performance in complex indoor environments with various obstacles and moving objects, which create multiple signal paths [210]. While radar-based solutions with large antenna systems and wide bandwidths have been successful in accurate, real-time activity monitoring, they are often expensive, power-hungry, and not widely accessible [236]. A more economical option is to use UHF RFID readers in combination with battery-less and compact RFID tags.

RFID technology is a practical and cost-effective solution for remotely monitoring elderly healthcare [237]. Its benefits include low cost, compact size, scalability, shareability, and battery-free operation [238]. Recent advances have produced inexpensive, highly sensitive passive tags with read ranges exceeding 10m, supporting their use as an economically viable solution for pervasive healthcare [231]. The emergence of 'tag-free' sensing, which employs contactless technology instead of attaching tags directly to the human body [239], presents a potential solution for AAL challenges [240]. This approach is less cumbersome and invasive for recognizing fundamental activities like standing, sitting, running, and walking, which are essential for well-being. These systems retrodict target objects or events by analyzing signal characteristic changes, such as RSSI and phase shift. Commonly, coarse-grained measurements of RSSI and phase value are used for sensing [111], but accurate results for indoor AAL and positioning require consideration of factors such as obstacles causing non-line-of-sight (NLOS) conditions, signal weakening due to rapid fading, and multipath effects from indoor construction materials, along with climate changes impacting signal propagation speed. These factors collectively affect the accuracy and reliability of AAL and positioning systems in indoor environments.

This chapter introduces a monitoring system that uses COTS UHF RFID technology, operating within the frequency range of 860 to 950 MHz. The main goal is to develop intelligent walls equipped with RFID tag arrays on each side to distinguish five distinct human activities:

sitting, standing, walking in two directions, and no-activity. Our proposed *TRT-Wall* system hypothesizes that the presence and movement of the human body within the radio field will result in recognizable RFID signal patterns due to attenuation, diffraction, reflection, and multi-path effects. Notably, the *TRT-Wall* approach enables monitoring the daily activities of elderly patients using pseudo-localization, reusing low-cost printed RFIDs and existing RFID readers for indoor activity recognition. Additionally, it ensures simple deployment using COTS RFID readers, requiring only a single UHF reader with a single antenna.

Specific contributions to this chapter are:

1. We propose *TRT-Wall* that uses contactless UHF RFID tags for sequential and simultaneous activity detection. Specifically, we collected a dataset for four different activities i.e., including sitting, standing, and walking (forward and backward).

2. The propose *TRT-Wall* leverages RSSI and phase data fluctuations for activity localization.

3. We perform an extensive evaluation of the collected dataset to determine the walking direction (i.e., forward or backward).

4. We calculate the speed of the moving object in order to establish a relationship between detection and activity location.

5.2 Data and Methods

This section presents the methodologies and materials used for collecting information through an experimental setup, which involved creating multiple test scenarios before applying machine learning (ML) and deep learning (DL) techniques for predictive analytics. The hardware and software components that were meticulously organised and utilised to collect RSSI and Phase information from the RFID UHF passive tags array using reader sensing devices that indicate human activity are expounded upon in subsections 5.2.1 and 5.2.1. Our proposed methodology is depicted in Figure 5.1, which comprises four major components that are elaborated upon below.

5.2.1 Experimental Setup

The experiments presented in this paper were conducted in accordance with ethical approval, within a $10 \times 10 m^2$ room located in the James Watts South building at the University of Glasgow. The experimental setup implemented a *TRT-Wall* structure, typical of a room with dimensions of $1.5 \times 1.5\ m^2$, which included several metal storage boxes and writing tables, thus offering rich multipath characteristics and a strong NLoS environment. The *TRT-Wall* is segregated into five columns, each consisting of three tags, and three rows, with a total of fifteen tags being utilised.

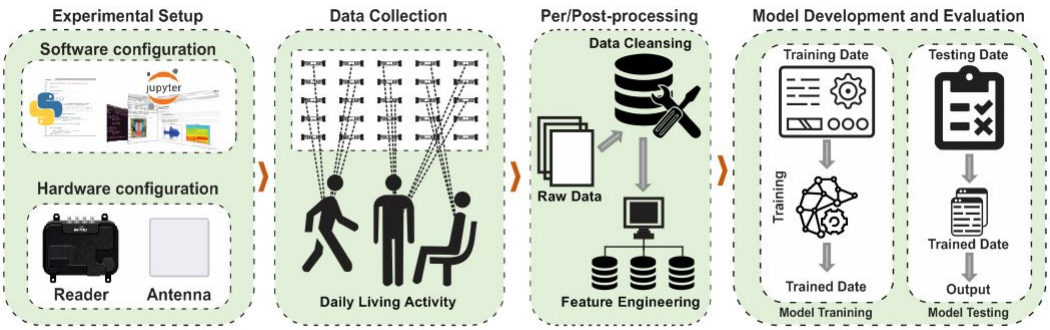
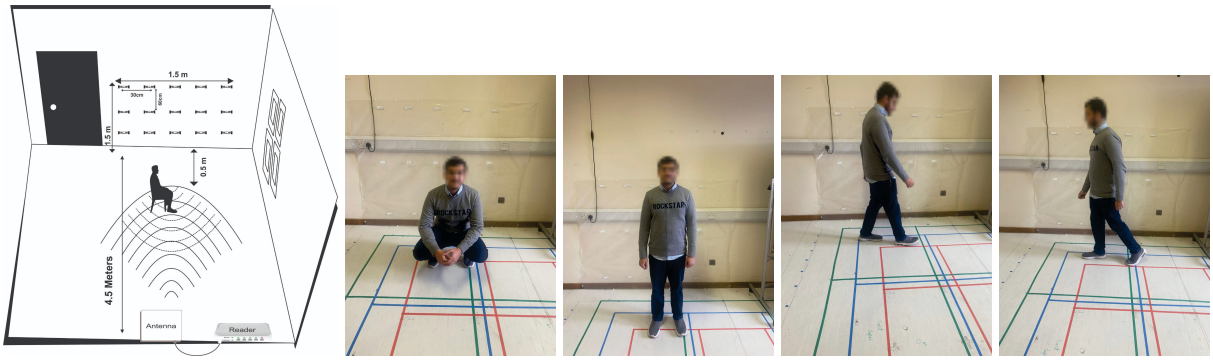


Figure 5.1: Data flow diagram: RSSI and phase capture for human activity, dataset compilation for ML/DL classification.

The circularly polarized antenna was placed at horizontal distances of $2, 2.5, 3.5$, and 4.5 meters from the center of the *TRT-Wall*, while the subject was positioned 0.5 meter away from the *TRT-Wall*. The height of the antenna was maintained at 0.75 meters above the floor surface while the subject was instructed to perform activities at designated locations. The only factors involved in the data collection were the activities and the surrounding environment. The data collection setup has two components, i.e., hardware and software setup, which are described below.



(a) Experimental Setup (b) Sitting (c) Standing (d) Walking L->R (e) Walking R->L

Figure 5.2: Experimental setup in a mocked room for activity recognition and localisation using TRT-Wall.

Hardware Setup

The proposed *TRT-Wall* for the AAL system utilises COTS UHF Gen-2 RFID devices without any hardware or firmware modifications. The system comprises a UHF passive RFID tag array and an RFID reader (*Impinj R*700). The reader operates between $865 - 868$ MHz using time-division multiplexing mode, is capable of reading up to 1100 tags per second, and is compliant with the EPC Class 1 and Gen 2 standard tags. These tags are attached to a board with a size of 3×5 grid points, where each tag is placed uniformly at $30cm$ apart, numbered from 1 to 15, arranged from left to right, and top to bottom. A circularly polarised antenna with dimensions

of 250 mm $\times$ 250 mm $\times$ 14 mm and an $8.0dBi$ gain is connected to the reader. The wavelength λ is set at 0.34 m, and the RF transmitter power is set to $30dBm$. The model training backend module runs on a laptop with an Intel® Core $i7-10850H$ CPU running at 2.7 GHz, dual-core, and 16 GB of RAM.

Software Setup

To collect data, the reader's data collection program is run on a laptop using the Impinj ItemTest Software[1]. The process involves the reader interrogating tags repeatedly and capturing RSSI and phase information from the back-scattered signals. The transmitter then transmits the received measurements of RSSI and phase information from the tag array through a laptop's RS232 serial port in a continuous manner. The streamed readings are received by the backend module and processed there.

5.2.2 Data Collection and Preprocessing

This section delineates the methodology employed for data collection. Firstly, we expound upon the various scenarios that were taken into consideration for data collection. This study has considered four distinctive test scenarios, which are explained below.

Test Scenario 1: One subject performing activities (reader & antenna 2m from subject, subject-*TRT-Wall* distance 0.5m).

Test Scenario 2: One subject performing activities (reader & antenna 2.5m from subject, subject-*TRT-Wall* distance 0.5m).

Test Scenario 3: Three subjects performing activities (reader & antenna 3.5m from subject, subject-*TRT-Wall* distance 0.5m).

Test Scenario 4: Three subjects performing activities (reader & antenna 4.5m from subject, subject-*TRT-Wall* distance 0.5m).

Data Collection

This chapter considered three subjects to conduct experiments with varying ages, height, and weight. To maintain consistency in both training and testing data, the subjects were instructed to perform four distinct activities, including sitting, standing, and walking (forward and backward) at their natural pace between the antenna and the *TRT-Wall*, as depicted in Figure 5.2. Every subject in the study gave their consent through the signing of an ethical approval form, which was authorized by the University of Glasgow's institutional review board. It is important to note

[1] https://support.impinj.com

that the data collection was carried out for each scenario, with the subject completing all activities while being mindful of their proximity to the *TRT-Wall* and antenna. Each activity having 20 samples, including RSSI and phase information, was collected for each subject. Only one subject was allowed to perform each activity as simultaneous recognition of multiple subjects was not the intention of this study. As a result, the data matrix contains information from a total of 15 tags. The data was collected when the subject was at a distance of $2, 2.5, 3.5$, and 4.5 m from the antenna. It is emphasized that each subject's role in collecting data was equally valuable while the inclusion of three subjects aimed to enhance diversity in the collected dataset. A total of 1200 valid training and testing activity samples were collected in four distinct scenarios, and each tag was read approximately $30-36$ times during a 3-second interval. These collected raw RSSI and radian are parsed using a Python script to extract relevant information for further preprocessing before they can be utilised for the training and testing of various ML/DL algorithms. A summary of the collected dataset is presented in Table 5.1.

Table 5.1: Dataset summary using *TRT-Wall*: scenarios, subjects, and activities performed.

Activity	**4.5 meters**		**3.5 meters**		**2.5 meters**		**2 meters**	
	RSSI	**Phase**	**RSSI**	**Phase**	**RSSI**	**Phase**	**RSSI**	**Phase**
Empty Room	20	20	20	20	20	20	20	20
Sitting	20	20	20	20	20	20	20	20
Standing	20	20	20	20	20	20	20	20
Walking Forward	20	20	20	20	20	20	20	20
Walking Backward	20	20	20	20	20	20	20	20

Data Preprocessing

Data preprocessing is an essential step in analysing raw RSSI data, as it involves cleaning, formatting, and transforming the data into a structured format that can be used for further analysis. As we collected raw data, we used essential mathematical/statistical techniques such as moving average window and signal processing methods including bandpass, low-pass, and high-pass filters to concentrate on a certain pattern. Initially, we processed the data using the following mathematical expression.

$$T_f \frac{y(k)-y(k-1)}{T} + y(k) = x(k), \tag{5.1}$$

$$y(k) = \frac{T}{T_f+T} x(k) \frac{T_f}{T_f+T} y(k-1) = ax(k) + (1-a)y(k-1). \tag{5.2}$$

The collected data for each activity was formatted in the form of a 2D matrix, where each row contained an observation and the columns represented the corresponding RSSI and phase tags.

To refine and ensure the quality of the data, we applied standard preprocessing functions from well-known libraries (i.e., *Scikit* and *Pandas*). The data for each activity was then stored in a data matrix with 540 columns of RSSI data (15 tags x 36 columns) when the activity was performed without any blockages. To ensure robust and unbiased training of ML/DL models, we consider generating synthetic data using generative adversarial network (GAN) and conditional tabular generative adversarial network (CTGAN) that helped in standardizing the number of samples for each activity class [241]. To handle the data, we saved the phase and RSSI data for activity as separate data files with a *frame size* of 36 (representing roughly 36 times in three seconds). During various activities such as sitting or standing, the corresponding tags were either partially or fully read. To ensure a 36-time tag reading, any missing data for each tag was replaced with zeros and included in the data matrix. The analysis of collected data revealed that no tag was read more than 36 times and *NaN* values were imputed with the mean of each row using the *SciKit* built-in *SimpleImputer* function. Then, we applied the pandas *unique* functions to divide the timestamp into seconds and monitored the correct reading of each tag for three seconds.

5.2.3 Activity recognition using RSSI

The use of passive UHF RFID tags in indoor activity localisation for AAL is activated by a reader employing an air interface protocol such as EPC class1 Gen-2 and ISO-18000-6c for data transmission and reception [242]. Within the realm of passive tag-based AAL, the RSSI RF capability can be effectively harnessed with COTS readers. In practical applications, passive RFID tags furnish the reader with raw data in a 5-tuple format, encompassing RSSI, timestamp, EPC, TID, and frequency. This raw data is collected from the tags. Nevertheless, the process of creating an RSSI dataset entails several steps outlined in Algorithm 1. The transformation required for solving the indoor propagation path loss model, which is utilized in various studies about RSSI distance transformation, can be derived through a simplified derivation.

$$P_{\text{ower}}\ (d_{\text{istance}}\ )_{dBm} = P\,(d_0)_{dBm} - 10n\log\left(\frac{d}{d_0}\right) + X_{dBm}, \tag{5.3}$$

$$d = 10^{\frac{P(d_0)_{dBm} - P(d)_{dBm}}{10n}}, \quad \forall \quad d_0 = 1 \text{ and } X_{dBm} = 0 \tag{5.4}$$

where $P(d_0)_{dBm}$ is received power along the propagation path of relative distance d, and $P(d_0)_{dBm}$ along the propagation path of reference distance $d_0(1m)$.

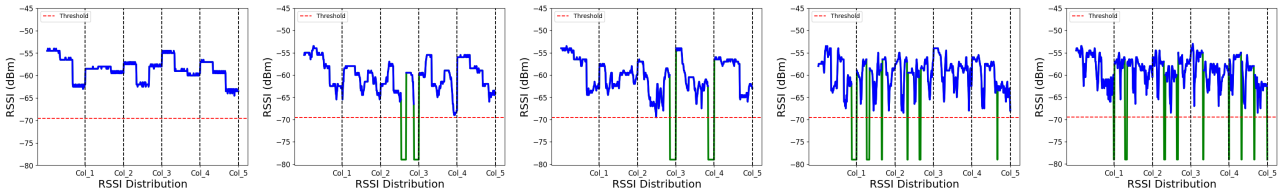
The recognition of AAL was made possible through a series of carefully planned activities, as illustrated in Figure 5.2. Specifically, five distinct activities were performed in the designated area in front of the *TRT-Wall*. For instance, the walking activity was performed from the first column to the fifth, with the main focus being on columns 3^{rd} and 4^{th} for sitting and standing activity. Figure 5.3 presents the results of the performed AAL, showing that the RSSI variations can be used to easily recognise each activity in the same location, with tags being blocked caus-

Algorithm 1 Pseudo code for RSSI Dataset Creation

Input : filePath, columns, groupCol, valueCol, keys, defaultValue, outputPath
Output: flattened

```
data ← Read CSV(filePath)
selected ← Select Columns(data, columns)
groups ← Group Values(selected, groupCol, valueCol)
dict ← Create Dictionary(keys, defaultValue)
for key in groups.keys() do
    if key in dict then
        dict[key] ← groups[key]
    end
end
for values in dict.values() do
    flattened ← Flatten(dict)
end
Save Data(flattened, outputPath)
```

ing the drop in the RSSI values. Notably, the threshold for RSSI strengths was determined by observing the maximum and minimum values, which were recorded at $-55dbm$ and $-69dbm$ respectively. The latter value, $-69dbm$, was selected as the threshold, taking into account potential instances of non-reading or blocking of tag activity detection. Instances where RSSI data wasn't read or an activity wasn't recognized are indicated by the color green. Specifically, in Figure 5.3 (a), the RSSI values for an empty room are displayed. Following this, Figures 5.3 (b) and (c) showcase sitting and standing activities in front of columns 3^{rd} and 4^{th}. Walking patterns from right to left and left to right are depicted in Figures 5.3 (d) and (e), respectively.

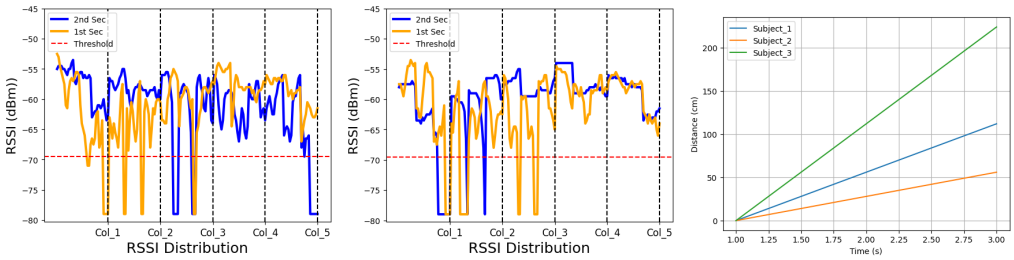


(a) No Activity (b) Sitting (c) Standing (d) Walking L->R (e) Walking R->L

Figure 5.3: Illustration showcasing the diversity of activity recognition data through RSSI distribution and magnitude analysis.

RSSI-based Walking Direction Analysis

To provide clarity regarding the walking direction, we partitioned the RSSI data for the walking patterns into one-second intervals and distinguished the data for each second using two distinct colors, as shown in Figure 5.4. In Figure 5.4 (a), during the 1^{st} second (represented in orange), the subject began walking from column-1 towards column-5, as indicated by the orange color, while during the 2^{nd} second (represented in blue), moved from column-2 towards column-5.



(a) Walking Forward (b) Walking Backward (c) Walking Speed

Figure 5.4: Illustration of walking speed and direction recognition via RSSI method with time split.

Similarly, in Figure 5.4 (b), during the 1^{st} second, the subject walked from column-5 towards column-1, with a clear indication from column-5 to column-3, whereas during the 2^{nd} second, they walked from column-2 to column-1. This partitioning and color-coding approach effectively illustrates the walking direction in the RSSI data.

Walking Speed Estimation

To accurately estimate walking speed, we acquired RSSI measurements at 3-second intervals and collected data from three subjects for preprocessing to address noise issues and enhance the precision of speed calculations. This preprocessing step encompasses the removal of outliers, filtering out unreliable readings, and applying smoothing techniques to mitigate the impact of measurement fluctuations caused by environmental factors. The calculation of walking speed involves analyzing changes in the distance over these 3-second intervals, as depicted in Figure 5.4 (c). This speed calculation is performed using a basic geometric formula: $\mathrm{v} = \frac{\Delta \mathrm{d}}{\Delta \mathrm{t}}$.

5.2.4 Activity Recognition using Phase

The utilisation of RF backscatter enables the signal to traverse a distance of $2d$ in dual directions, thereby facilitating the monitoring of human activity through the analysis of RF features' phase differences using cross co-relationship. The subsequent formula elucidates the correlation between distance, antenna phase rotation, and tag phase rotation:

$$\theta = (2\pi\frac{2d}{\lambda} + \theta_{Ant} + \theta_{Tag})\mathrm{mod}(2\pi)\lambda\,. \tag{5.5}$$

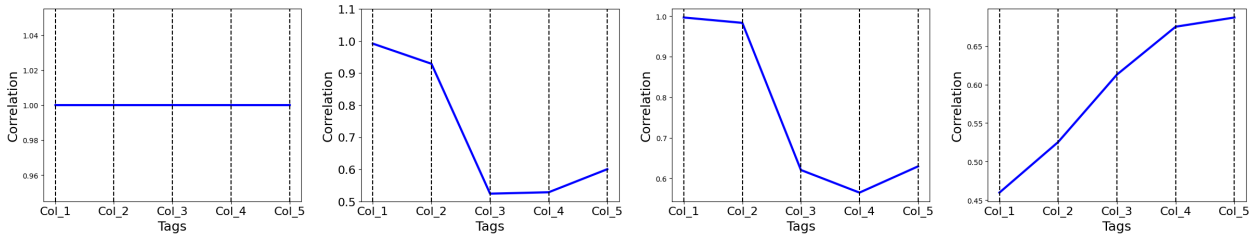
The phase is a periodic function of 2π radians occurring every $\lambda/2$ in the RF communication distance. The rotations of the antenna and tag phases are described by θ_{Ant} and θ_{Tag}, respectively.

It is essential to assess the precision and discriminatory nature of phase difference calculations during activity. The significance of this is demonstrated in Figure 5.5, which displays phase difference patterns during sitting and empty activities in front of columns 3^{rd} and 4^{th} using

numpy *np.corrcoef(x,y)*. The cross co-relationship function is explained in Algorithm 2. The co-relationship difference pattern suggests an effective method for modeling activities. The smooth variation of phase differences across blocking tags during sitting highlights their accuracy and reliability. The visualisation results indicated that the calculated phase differences are reliable and sensitive to AAL activity. To quantify the strength of the relationship between two different activities performed against the same tag (each tag has 36 phase reading values), the following formula can be used to calculate the correlation coefficient:

$$r_{xy} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}}, \tag{5.6}$$

where r_{xy} represent the correlation coefficient of the linear relationship between the tag value of empty activity and sitting activity tags, x_i, y_i the values of the empty and sitting activity tags values whereas $\bar{x}$ and $\bar{y}$ denotes the mean of the values respectively.



(a) No Activity (b) Empty-Sitting (c) Empty-Standing (d) Empty-Walking

Figure 5.5: Illustration of activity recognition using phase difference with co-relationship

Algorithm 2 Pseudo Code for Phase Correlation

Input : file paths 1 and 2, header=None, rowIndex, start, end, step, size

Output: correlations list

10 $data1 \leftarrow$ Read CSV($file\ path\ 1$, header = None) $data2 \leftarrow$ Read CSV($file\ path\ 2$, header = None) $row1 \leftarrow$ extractRow($data1$, rowIndex) $row2 \leftarrow$ extractRow($data2$, rowIndex) $correlations \leftarrow []$

11 **for** $i \leftarrow start\ to\ end\ by\ step$ **do**

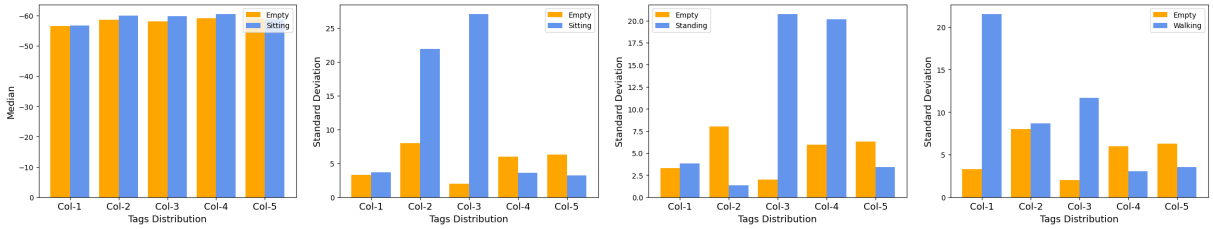
12 $subrow1 \leftarrow$ extractSubArray($row1, i, i+$size) $subrow2 \leftarrow$ extractSubArray($row2, i, i+$size) $corr \leftarrow$ calcCorr($subrow1, subrow2$) $correlations \leftarrow$ append($corr$)

13 **end**

5.2.5 Activity Recognition using FeatureSet

The performance of AI models may be improved significantly by feature engineering. Flexible features allow for the development of models that are less complicated and easier to maintain, which leads to better performance. Moreover, feature engineering reduces the amount of time

required for variable extraction, enabling the extraction of several variables. This research focuses on processed matrix data, which is made up of rows that indicate activities or occurrences. There are 540 columns for each activity, with each row including 36 RSSI/phase samples (frame sizes) that were collected using 15 tags during a 3-second timeframe. The objective is to extract high-order features from the CSV file that may significantly reduce data dimensions and boost the robustness and classification accuracy of the system. By combining frequently used statistical features such as mean, median, mean absolute value, standard deviation, variance, minimum, maximum, skewness, kurtosis, count, entropy, trimmed mean, trimmed variance, trimmed minimum, trimmed maximum, trimmed standard deviation, trimmed standard error, variation, score at percentile, and correlation coefficient, the study determines the optimal feature subset for AAL classification. Figure 5.6 demonstrated activity recognition by highlighting the standard deviation of empty activity with sitting, empty with standing, and empty with walking.



(a) Empty-Sitting Mean (b) Empty-Sitting Std Dev (c) Empty-Standing Std Dev (d) Empty-Walking Std Dev

Figure 5.6: Illustration of activity recognition using standard deviation on the feature set.

5.3 Evaluation and Results

This section presents the results of four distinct experimental scenarios using four distinct approaches, each involving different subjects performing various activities, as depicted in Figure 5.2. The evaluation of the experimental results comprises an assessment of both the overall performance and the impact of reader-subject and subject-tag distances, as well as the type of activity on the accuracy of the system. Tables 5.2 and 5.3 provide a comprehensive evaluation that is utilised to assess the system's robustness.

5.3.1 Artificial Intelligence Model Development

The development of an artificial intelligence (AI) model to provide assistance and support to individuals with limited mobility or disabilities, allowing them to perform daily activities in a contactless manner. One of the key challenges in developing an AI model for contactless AAL is to ensure the privacy and security of the user's data.

Table 5.2: Deep learning classification accuracy in multi-distance scenarios.

Subjects	4.5 meters			3.5 meters			2.5 meters			2 meters		
	RSSI	Phase	FeatureSet	RSSI	Phase	FeatureSet	RSSI	Phase	FeatureSet	RSSI	Phase	FeatureSet
LSTM Accuracy												
1	**81.2%**	68.3%	75.2%	**87.4%**	81.2%	62.5%	60.6%	62.7%	**74.1%**	53.5%	62.9%	**69.5%**
2	**91.6%**	75.2%	88.6%	**91.8%**	87.7%	85.7%	72.4%	68.5%	**79.8%**	66.8%	66.3%	**73.4%**
3	**93.7%**	92.1%	90.3%	**95.6%**	94.3%	91.6%	76.4%	73.5%	**82.4%**	70.2%	70.3%	**76.4%**
CNN Accuracy												
1	**78.7%**	75.3%	78.1%	68.2%	**86.6%**	81.4%	63.4%	**66.5%**	61.3%	60.3%	66.5%	**70.4%**
2	**82.3%**	81.3%	79.5%	**89.5%**	88.4%	87.6%	**69.3%**	67.5%	64.3%	64.8%	70.4%	70.4%
3	89.3%	**93.2%**	85.7%	**93.5%**	92.6%	87.5%	76.5%	76.6%	**84.4%**	73.7%	74.5%	**76.3%**
LSTM+CNN Accuracy												
1	**79.6%**	75.7%	78.4%	**86.6%**	68.7%	81.5%	63.4%	**66.9%**	61.6%	60.5%	66.8%	**70.5%**
2	81.6%	**82.7%**	79.4%	**88.5%**	87.7%	87.6%	67.5%	**69.8%**	64.3%	64.5%	70.3%	70.7%
3	**92.3%**	89.5%	85.6%	**92.6%**	91.7%	87.6%	76.5%	76.5%	84.6%	73.6%	74.6%	**76.4%**

Table 5.3: Machine learning classification accuracy in multi-distance and multi-subject.

Subjects	4.5 meters			3.5 meters			2.5 meters			2 meters		
	RSSI	Phase	FeatureSet	RSSI	Phase	FeatureSet	RSSI	Phase	FeatureSet	RSSI	Phase	FeatureSet
SVM Accuracy												
1	**80.2%**	78.4%	70.5%	**85.4%**	84.5%	75.6%	**76.5%**	74.8%	70.6%	66.7%	**67.6%**	65.5%
2	81.6%	**82.7%**	73.7%	**90.5%**	85.7%	81.6%	**76.5%**	73.6%	72.7%	**66.6%**	65.5%	63.8%
3	**89.4%**	89.6%	82.6%	**91.6%**	89.6%	84.5%	81.5%	**83.6%**	77.6%	70.4%	**72.9%**	68.4%
RF Accuracy												
1	**73.7%**	71.6%	71.0%	**76.6%**	74.2%	73.3%	**69.3%**	66.9%	65.5%	61.4%	59.4%	**63.4%**
2	**81.2%**	73.2%	74.5%	79.3%	**81.4%**	73.4%	**71.8%**	66.4%	70.3%	**65.5%**	63.5%	63.9%
3	**81.7%**	78.7%	76.3%	**90.8%**	85.4%	81.4%	**78.6%**	78.5%	73.5%	**67.7%**	66.5%	66.9%
DT Accuracy												
1	73.5%	70.5%	**77.4%**	74.7%	75.6%	**78.4%**	70.5%	63.2%	**68.6%**	**66.3%**	60.3%	63.5%
2	77.4%	75.5%	**81.4%**	79.5%	77.4%	**80.5%**	**73.4%**	67.5%	71.6%	**65.5%**	62.6%	64.6%
3	85.6%	81.3%	**85.4%**	79.2%	**82.3%**	82.6%	**79.8%**	74.7%	76.4%	68.6%	**70.8%**	66.4%

Data Analysis using DL Models

To assess the efficacy of the system, experiments were conducted on datasets collected from a scenario-based environment. The front layers of the RSSI and Phase profiles were exclusively used for all activities. The performance of the system was evaluated through the application of three distinct DL models, including long short-term memory (LSTM), convolutional neural network (CNN), and a combination of the two (LSTM+CNN). Before training an LSTM-CNN network, a series of raw RSSI data was transformed into a stack of matrices with dimensions of 5×3, corresponding to the size of the tag array. The experiment data was simplified by focusing solely on the 1D layer rather than converting to 3 levels using the timestamp dimension. The instantaneous RSSI value for each RFID tag was entered into the i^{th} row and j^{th} column of a matrix. If certain tags failed to provide RSSI readings due to blockage caused by human activity, the corresponding value defaulted to 0. The processed data was normalized with a mean of 0 and a standard deviation of 1 before being passed into the network as input. The proposed hybrid

DL models incorporated a single LSTM layer with a single dropout and flatten layer, while the CNN model utilised a $1D$ convolution layer due to the linear data structure. Two equally sized $1D$ convolutional layers and two identically sized *max pooling* layers were utilised, with a dense layer employed between the $flatten$ and output layers. For the third model, fully connected layers were used to merge the LSTM and CNN models. The optimizer utilised was *adam*, with a decay of $1e^{-6}$ and a learning rate of 0.01. The activation function was set to *tanh*. The models were trained across 50 epochs, with user recognition being addressed as a problem of multi-class classification. The parameters tracked during the training included accuracy and loss.

Data Analysis using ML Models

In addition to DL algorithms, three classical ML algorithms, namely SVM, random forest, and decision tree classifier were applied in the current experiments to evaluate the collected dataset. Accuracy was assessed using a train-test split technique, where predictions were generated using data that was not used for model training. The data was divided into training and testing subsets, with a train-test ratio of 0.8, which indicates that 80% of the data was utilised for training, and the remaining 20% was utilised for testing as mentioned in Table 5.4.

Table 5.4: Hyper-parameters of ML/DL algorithms

S.No	Algorithms	Hyper Parameters
1	LSTM	optimizer= adam, hidden-layers-activation = tanh, lr= 0.01, loss = binary-crossentropy, batch-size = default, hidden-layer-size(15,50), dropout = 0.2, out-layer-activation = softmax, epochs= 50
2	CNN	optimizer= adam, hidden-layers-activation = tanh, lr= 0.01, loss = binary-crossentropy, batch-size = default, hidden-layer-size(15,32), max-pooling-layer-size = (3,1), out-layer-activation = softmax, epochs= 50
3	LSTM+CNN	optimizer= adam, hidden-layers-activation = tanh, lr= 0.01, loss = binary-crossentropy, batch-size = default, LSTM-hidden-layer-size(15,50), CNN-hidden-layer-size(15,50), dropout = 0.2,max-pooling-layer-size = (3,1), out-layer-activation = softmax, epochs= 50
4	SVM	degree = 3, gamma = auto, kernel = linear, tol = 0.001, shrinking = true, C = 1.0
5	RF	n-estimator = 10, criterion = gini, max-features = auto, min-samples-leaf = 1, $\min_i mpurity_d ecrease = 1e-7, n-jobs = 1$
6	ET	min-samples-leaf = 1, splitter = best, $\min_i mpurity_d ecrease = none, criterion = gini$

5.3.2 User Recognition's Overall Performance

To evaluate the effectiveness of user recognition with our collected dataset, we have employed the k-fold cross-validation technique. This method involves dividing the dataset into k groups of equal size and randomly shuffling them. In each fold, k-1 groups are used for training the

model, while the remaining group is reserved for validation. The process is repeated k times, and the average of the outcomes serves as the final estimate. In this study, we have chosen k=5, resulting in each fold consisting of eight samples from each user for testing (240 samples total; 20%), while the remaining samples (960 samples total; 80%) are used for training. As a result, 960 samples are used precisely once as validation data after all five folds. The overall results and normalized confusion matrix are depicted in Figures. 5.7 and 5.8.

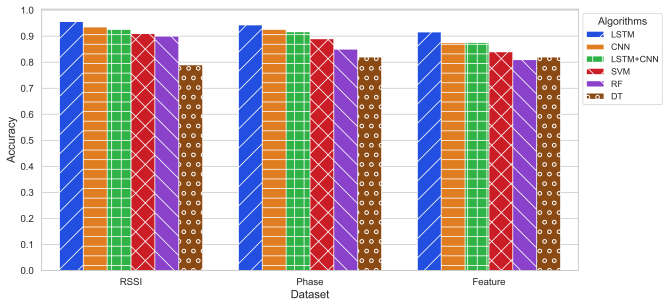


Figure 5.7: Comparison of DL and ML algorithms across various scenarios and approaches.

(a) SVM (b) LSTM

Figure 5.8: A normalized confusion matrix of various activities recognition using SVM and LSTM.

5.4 Discussion

This study assessed six algorithms, namely LSTM, CNN, LSTM+CNN, SVM, RF, and DT, for contactless RFID human activity recognition across four distinct scenarios, employing three different approaches as detailed in section 5.3. The results revealed that in scenario 1, where the reader and antenna were 2 meters from the subject and the subject-TRT-Wall distance was 0.5 meters, deep learning algorithms (LSTM, CNN, and LSTM+CNN) outperformed the machine learning algorithms (SVM, RF, and DT). In contrast, in scenario 2, characterised by noise and a weak line of sight (LoS) situation, the machine learning algorithms demonstrated performed better, with SVM achieving the highest accuracy of 83.6%. Moving on to scenario 3, which entailed a strong NLoS environment, the deep learning algorithms, particularly LSTM, outperformed the machine algorithms, boasting an impressive accuracy of 95.6%. In scenario 4, where

the reader and antenna were positioned 4.5 meters from the subject, the deep learning algorithms again exhibited better performance, with LSTM achieving the highest accuracy of 93.7%. Cross-validation further affirmed the reliability of both SVM and LSTM algorithms, yielding average accuracies of 91.6% and 95.6% respectively, as demonstrated in Figure 5.8. The normalised confusion matrix underscored the LSTM model's consistent recognition accuracy, surpassing 91% for all subjects. This observation reinforces the potential of the proposed approach for practical implementation in user recognition applications, as depicted in the Figures 5.7, 5.9, and 5.11. Overall, the performance of deep learning and machine learning algorithms for contactless RFID human activity recognition is dependent on the distance from *TRT-Wall* to the reader antenna. Deep learning algorithms are better suited for scenarios where there is a need to capture the temporal dynamics of human activities, while machine learning algorithms are better suited for scenarios where the data is noisy or the distance is limited.

5.4.1 Ablation Studies

The section systematically investigated key factors influencing RFID-based human activity recognition. It concluded that maintaining a 0.5 meter proximity between *TRT-Wall* and antenna settings is optimal. Additionally, it highlighted the significance of reducing the number of tags for more efficient activity data transmission. The study also explored the influence of subject quantity on detection accuracy, revealing a non-linear relationship. Lastly, it emphasized the critical role of antenna height in optimizing system performance, concluding that a default height of 0.75 meters ensures robust outcomes.

Impact of Distance from *TRT-Wall* to Antenna Settings

The study adopts the *TRT-Wall* approach, which employs an array of tags to decouple activity recognition of subjects. Throughout the experiments, accuracy is measured at different distances between the antennas and tags, spanning from 2 to 4 meters. The results reveal a statistically significant correlation between the *TRT-Wall* distance and the antenna. Notably, a *TRT-Wall* distance of 1.0 and 2 meters proximally causes severe distortion in RSSI and Phase waves, resulting in false positives during recognition. The initial tag distance between the subject and the *TRT-Wall* is set at 0.5 meters. Subsequently, the accuracy of activity recognition is further tested at various subject-to-antenna distances, ranging from 2 to 4.5 meters. The results indicate that the accuracy decreases with decreasing distance for unobstructed readings and beyond 3.5 m, various factors such as weak signals and the lower reading rate of RFID tags affect accuracy. To address the impracticality at shorter distances, the proposed study maintains a default setting of 0.5 meters between tags and subjects. Meanwhile, the distance between the subject and the antenna is tested at $2, 2.5, 3.5$, and 4.5 meters.

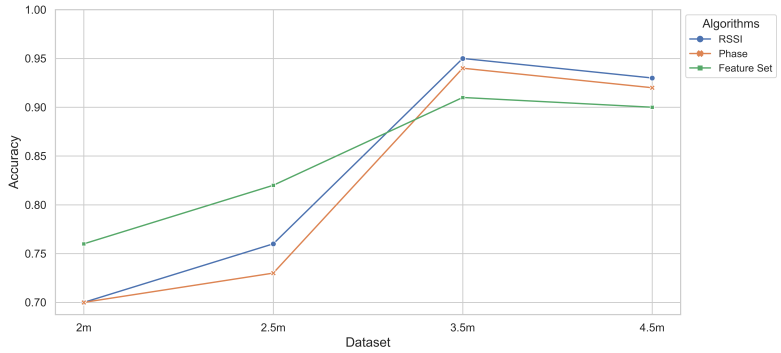
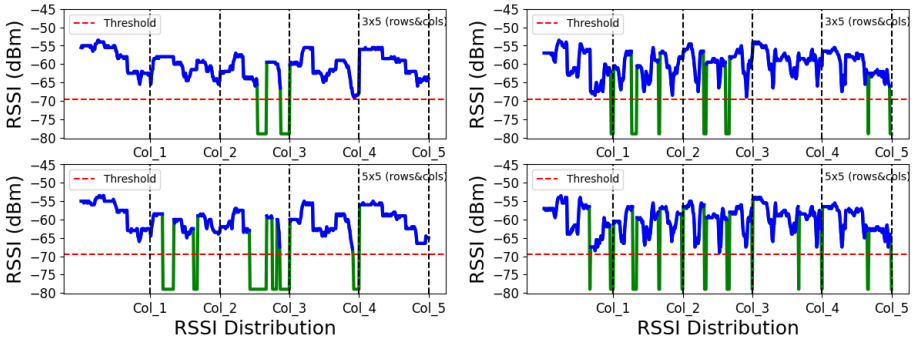


Figure 5.9: Assessing LSTM accuracy through four different scenarios and approaches.

Impact of Number of Tags

Increasing the number of tags leads to a greater number of reflected signals, broader pathway coverage, and enhanced data collection for activity detection. Our experiments reveal that a reduced tag count facilitates efficient coupling and transmission of activity data. By eliminating two rows containing five tags each, we bring the total tag count down to 15 (arranged in 3 rows and 5 columns) from the initial 25 (arranged in 5 rows and 5 columns). This reduction not only provides supplementary information but also enhances activity prediction performance, as depicted in Figure 5.10 (a) illustrating sitting activity and Figure 5.10 (b) showcasing walking activity. If the goal is to augment the environmental path complexity, it is advisable to augment the number of tags in the columns. Additionally, within indoor environments, the incremental cost of adding more tags is marginal compared to the expense of incorporating additional readers with antennas [243]. To implement this, our study employs a default configuration comprising a single circularly polarized antenna and 15 passive UHF tags.



(a) Sitting 3×5 vs 5×5 (b) Walking 3×5 vs 5×5

Figure 5.10: Impact of a number of tags.

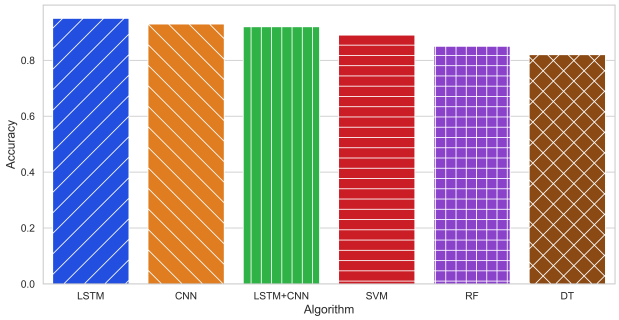


Figure 5.11: A comparison of DL and ML algorithms on Scenario 3.

Impact of Number of Subjects

This section examines how the number of subjects affects performance and concludes that detection accuracy does not increase linearly with the number of subjects, but rather gradually. Since fewer subjects indicate less variation in user characteristics, this result is consistent with our expectations. We assess a number of parameters in our study, including RSSI, Phase, and feature values. When there are three subjects involved, our suggested method has an accuracy of 96%, but when there is just one subject, it only has an accuracy of 81.20%. Nevertheless, WiFi-ID [58] and WiWho [62] outperform our method when detecting six subjects at a single distance using Wi-Fi signals, achieving accuracy levels of 77% and 80%, respectively. As shown in Table 5.5, RFree-ID [244], TagFall [128], and [245, 246] respectively, achieved high accuracy levels of 93%, 94%, 94%, and 95% for detecting human activities such walking or falling using RFID. It needs to be noted that it may not be practical or convenient to enable subjects to do activities at a distance of 4.5 m for proper recognition.

Table 5.5: Analysis of different user activity recognition approaches.

Approach	Technology	Group Size	Antenna	Accuracy
WiFi-ID [58]	WiFi	6	1	77%
WiWho [62]	WiFi	6	3	80%
RFree-ID [244]	RFID	5	1	93%
TagFall [128]	RFID	1	4	94%
Ruan et al. [245]	RFID	1	1	94%
Wenjie et al. [246]	RFID	1	4	95%
TRT-Wall	*RFID*	*3*	*1*	**95.6%**

Impact of Antenna Height

This study examines how the performance of RFID tag reading systems is affected by antenna height. The system's range and precision are significantly influenced by the antenna height. Raising the antenna height to ensure visibility of all tags enhanced system performance. Nonetheless, we ran an experiment to examine three different antenna placement scenarios in

our RFID tag reading system. The first case included mounting the antenna against a $1.5 \times 1.5m^2$ wall on the ground level. Our research showed that the tags on the top row were not properly read can reduce the recognition accuracy. Secondly, we then adjusted the antenna height to improve the read range of the system to 0.75 m, which was the centre of the wall and had an LoS. This was created as it was thought that a small change in antenna height would have essentially little impact on accuracy. Nevertheless, we noticed a decrease in the tag signal strength when the antenna height was raised further from 0.75 m to 1.5 m. The reason for this was that the reader's signal was insufficiently strong to reach the lower row of tags. According to the findings of our experiments, maintaining the default antenna height of 0.75 m, or mid-height of the wall, is necessary to achieve optimal system performance. After experimental verification, it is clear that the system is robust.

5.5 Limitations and Future Directions

The *TRT-Wall* approach is a key step toward making it possible to accurately detect indoor activities without requiring users to wear or carry any RFID tags. However, there are a lot of opportunities to strengthen the basic TagFree concept in the future.

5.5.1 System Flexibility

The study investigated the precision with which *TRT-Wall* can detect each subject's activity. Nonetheless, it is expected that the system's scope could expand to monitor multiple subjects simultaneously owing to the complex subject interactions. This would need more extensive data pre-processing and analysis. According to our preliminary study, DL algorithms are better at detecting individual activities than conventional methods. Moreover, the *TRT-Wall* range is limited to 12 meters when using a single antenna. This reading range can be extended by installing larger antenna arrays using an Impinj antenna hub and using many RFID tags as references for wider coverage.

5.5.2 Model Generalisation

The implementation of our proficient DL model is limited to non-uniform antenna configurations or tag placements. In consequence, the model necessitates retraining to accommodate heterogeneous environments. An alternative approach for enhancing the predictive outcome involves tuning the model through Federated learning. Federated learning facilitates the training of the model on various room samples, presuming that certain activities transpire in a particular sequential order, resulting in a generalized and uniform trained model, adaptable to any heterogeneous environment

5.5.3 User Authentication

The ability of the current approach to distinguish between a limited number of activities and user detection is limited. In the future, it could be possible to get around this limitation by developing a user identification and recognition model that makes use of strong and complete user attributes rather than an activity-centric one.

5.6 Summary

This chapter presents a study on using commercial Impinj RFID readers to create TRT-Wall, a cost-effective, non-invasive, and privacy-preserving user identification mechanism leveraging multipath signals from various activities. This tag-free model uses a data preprocessing technique for detailed activity recognition and employs a long short-term memory network to effectively detect activities without tagging targets. Extensive experiments with commercial RFID devices demonstrate that TRT-Wall surpasses existing methods, achieving average activity identification accuracy rates of 95.6% with RSSI data, 94.3% with phase difference, and 91.6% with features, even in multipath-rich environments with up to three subjects. Further extending the applications of RFID technology, the study introduces an innovative method for fall detection in elderly individuals using RFID technology, employing passive RFID tags that operate by querying these tags and analysing the RSSI and phase data. This data is then processed using a sophisticated transformer model combined with data fusion techniques before comparing the results with CCNN, RNN, and LSTM networks.

Chapter 6

TFree-FD: Tag-Free Fall Detection in Indoor Environments using Fusion Approach with Transformer Encoders

This chapter introduces an innovative method for fall detection in elderly individuals using radio frequency identification (RFID) technology. This technique offers a convenient and effective alternative to traditional wearable devices, which often cause discomfort and negatively impact user experience. The proposed system employs passive RFID tags placed in strategic locations to facilitate unobtrusive monitoring. It operates by querying these tags and analyzing the signal strength (RSSI) and phase data. The data is then processed using a sophisticated transformer model combined with data fusion techniques, resulting in enhanced accuracy in both activity recognition and non-contact fall detection. Remarkably, this method achieves an impressive average accuracy rate of over 96%, surpassing the effectiveness of other prevalent technologies like convolutional neural networks (CNN), recurrent neural networks (RNN), and long short-term memory (LSTM) networks. This high level of accuracy demonstrates the method's reliability and its significant potential for real-world application.

6.1 Introduction

Falls are a significant global health concern, particularly among the elderly population. As the second most common cause of accidental injury after traffic accidents, falls pose a serious threat to the elderly, leading to injuries, mortality, morbidity, frailty, and disability in individuals over 65 [87, 88]. Statistically, around 30% of the elderly experience at least one fall annually, with 90% of these falls resulting in hip fractures and 60% leading to head injuries [87]. Moreover, the phenomenon of 'long-lie' elderly individuals remaining on the ground for extended periods post-fall exacerbates risks like pressure sores and hypothermia [89]. In light of these concerns, the implementation of automatic fall detection systems is crucial, as they significantly reduce

response times for medical assistance, thereby mitigating the impacts of falls [90].

Fall detection methods are broadly categorized into wearable and non-wearable approaches. Wearable systems integrate sensors like gyroscopes and accelerometers, attached to the body, to monitor activity changes [91, 92, 93, 94]. Non-wearable methods include visual detection using cameras, environmental sensor placement, and RFID-based detection [95, 96, 97]. RFID technology, in particular, has been noted for its affordability and non-intrusive nature. Traditional RFID systems are tag-based, where activities are tracked by attaching RFID tags to objects or individuals [111, 112, 113, 114, 115, 116, 75, 117]. However, these systems face limitations in tracking untagged activities like walking or body rotation. To overcome the constraints of 'tag-based' systems, recent advancements have led to the development of 'tag-free' solutions. These systems utilize multiple stationary RFID tags as reference points, with their signals influenced by human activities nearby [247, 248, 249]. Deep learning approaches have been employed to analyze the signal fluctuations caused by activities, offering a more nuanced understanding of the environment [250, 251, 249]. However, these methods are not without challenges; they are sensitive to environmental changes and require extensive training data for neural networks.

Introducing our tag-free fall detection *TFree-FD* system marks a significant advancement in fall detection technology. This system leverages a passive RFID tag array and a neural network to detect falls more precisely and reliably. By processing raw radian and RSSI data into spectrum frames, *TFree-FD* utilizes supervised learning for accurate activity classification, including various fall scenarios and normal activities like standing or leaning. The system is built on a transformer architecture, notable for its efficiency and reduced training time compared to LSTM and CNN models [252, 253]. This approach not only improves fall detection accuracy but also holds potential for smart home applications, such as temperature control and gym equipment assistance. Importantly, *TFree-FD* is compatible with standard RFID readers, making it a cost-effective and practical solution for indoor activity monitoring.

6.2 Data and Methods

This section presents an extensive overview of the methodologies and materials utilized in the experimental setup for data collection, specifically aimed at predictive analytics using deep learning techniques. Before applying these techniques, two test scenarios were designed to facilitate the data collection process. Subsections 'hardware setup' and 'software setup' provide detailed information on the hardware and software components meticulously organized and employed to capture RSSI and phase information from the RFID UHF passive tags array. Our proposed methodology, comprising five major components, is illustrated in Figure 6.1, with each component thoroughly explained in the subsequent sections.

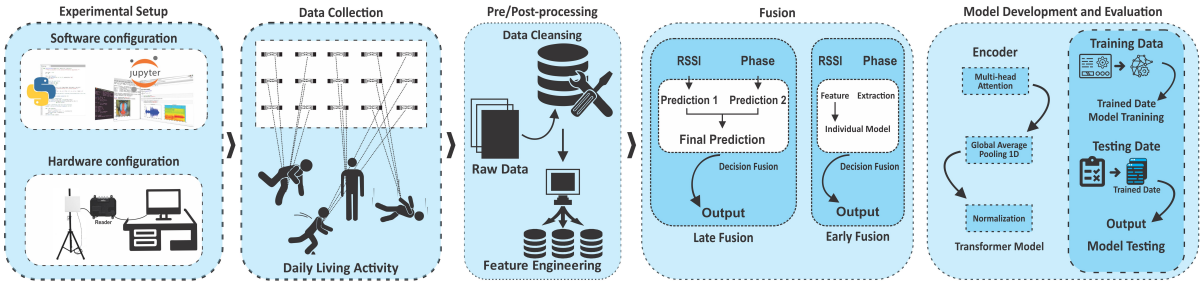


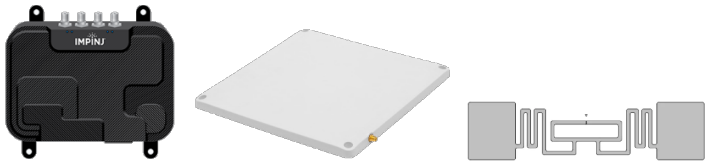
Figure 6.1: Integrated workflow for intelligent experimentation from experimental setup to transformer model development.

6.2.1 Experimental Setup

The experimental setup for this study is identical to the one described in Chapter 5, Section 5.2.1, with a focus exclusively on the last two scenarios, which are at distances of 3.5 and 4.5 meters.

Hardware Setup

The hardware setup for this study is identical to the one described in Chapter 5, Section 5.2.1, are outlined in Figure 6.2.



(a) Reader (b) A5010 Antenna (c) UHF Tag

Figure 6.2: Hardware used in experimental setup

Software Setup

In this study, we processed the collected data and trained our fall detection transformer model using the TensorFlow 2.0 development platform and the Python programming language on a laptop. To facilitate this process, we utilized the Impinj ItemTest Software[1] to continuously transmit the collected measurements of RSSI and phase information from the tag array through the laptop's RS232 serial port by the transmitter. Figure 6.3 provides a visual representation of the experimental scene, offering a contextual view of the testing environment.

[1]https://support.impinj.com/hc/en-us/articles/204059593-Impinj-ItemTest-Software

(a) Experimental Setup (b) Standing (c) Leaning (d) Normal Fall (e) Walk Fall

Figure 6.3: Experimental setup and fall-related activities scenario.

6.2.2 Data Collection and Preprocessing

In this section, we discuss the methodology applied for data collection, which involved two distinct test scenarios:

Test Scenario 1: The subject performed activities with the reader and antenna positioned 3.5 meters away from the subject, while the subject-*TRG-Wall* distance was maintained at 0.5 meters.

Test Scenario 2: The subject performed activities with the reader and antenna placed 4.5 meters away from the subject, and the subject-*TRG-Wall* distance was also kept at 0.5 meters. These two carefully selected test scenarios captured different configurations and distances between the subject, reader, antenna, and the *TRG-Wall*, aiming to collect data that represented real-world situations and variations in the experimental setup.

Data Collection

In this study, we conducted an experiment involving three subjects who varied in terms of age, weight, and height. In the absence of a publicly available RFID-based dataset, we proactively create our own dataset. To ensure the reliability of our results, each subject performed five distinct activities: no-activity, standing, leaning, normal fall, and walking fall. These activities were performed at a natural pace within the designated area between the antenna and the *TRG-Wall*, as illustrated in Figure 6.3. During data collection, we took meticulous care to ensure subjects consistently maintained their proximity to both the *TRG-Wall* and the antenna. To maintain experimental control and focus on individual recognition, only one subject performed the activities at a time, resulting in data from a total of 15 RFID tags. All subjects in the experiment provided their consent by signing an ethical approval form sanctioned by the institutional review board of the University of Glasgow. The data collection process resulted in a total of 2600 valid training and testing samples across two distinct scenarios. Each RFID tag was read approximately $32-36$ times within a 3-second interval. Subsequently, we utilized a Python script to parse the 50 collected samples of each activity, extracting pertinent information for further pre-processing.

Table 6.1: Dataset Summary using *TRG-Wall*: scenarios, subjects, and activities performed.

Activity	**Senario-1 (3.5 m)**		**Senario-2 (4.5 m)**	
	RSSI	**Phase**	**RSSI**	**Phase**
Empty Room	50	50	50	50
Standing	50	50	50	50
Leaning	50	50	50	50
Normal Fall	50	50	50	50
Walk Fall	50	50	50	50

The processed dataset was then employed to train and test DL algorithms. A summary of the collected dataset, including its composition and structure, is provided in Table 6.1.

Data Preprocessing

To ensure accurate fall detection, the reflected signal from the RFID reader undergoes several essential preprocessing steps. This encompasses raw RSSI and phase values, meticulously calibrated to address environmental and system-specific factors. Advanced filtering methods, including adaptive filtering and wavelet denoising, are then employed to eliminate initial phase and RSSI noise. A signal segmentation process, utilizing techniques like adaptive thresholding and dynamic time warping, is applied to isolate fall-relevant segments. Subsequently, comprehensive normalization standardizes the data, reducing potential biases. These steps form a robust foundation for reliable fall detection algorithms, significantly enhancing accuracy and efficiency. This preprocessing ensures subsequent stages of the fall detection algorithm operate on a refined and optimized dataset, resulting in a more dependable system.

1) Phase Normalisation: We employ neighboring phase averaging to mitigate hardware-induced thermal noise, aligning and refining phase values for improved accuracy (Eq. 6.1).

$$\alpha = \left(\frac{4\pi d}{\lambda} + k \cdot 2\pi \right) \bmod (2\pi). \tag{6.1}$$

a) Noise Reduction: Gaussian smoothing complements phase normalization, reducing high-frequency noise for enhanced clarity in phase-distance relationships.

b) Temporal Segmentation: Data is partitioned into discrete time intervals to isolate and analyze specific activities, including falls.

c) Quality Control: Checks identify and rectify anomalies in RFID data, ensuring data integrity through validation for missing or erroneous readings.

d) Signal Alignment: Dynamic alignment synchronizes data from multiple RFID tags for temporal consistency, critical for accurate fall detection.

For data consistency, we assume $\Delta d \leq \lambda = 4$ as the tag's distance difference between consecutive sampling points, given the short time delay between rounds.

$$\begin{aligned} \alpha &= \left(\frac{4\pi d}{\lambda} + k \cdot 2\pi + \Delta\alpha \right) \bmod (2\pi) \\ \Delta\alpha &= \left(\frac{4\pi \Delta d}{\lambda} - \Delta\phi \right) \bmod (2\pi). \end{aligned} \tag{6.2}$$

2) RSSI Signal Noise Reduction:

Wavelet filtering comprises three main steps: decomposition, thresholding, and reconstruction, aimed at analyzing patterns in RSSI data. To achieve this, we utilized the discrete wavelet transform (DWT) to decompose the signal into wavelet coefficients, which capture frequency content at different scales. The DWT convolves the signal with wavelet basis functions, denoted by $\psi_{j,k}$ and $\phi_{j,k}$, operating at various scales and positions. Specifically, we employed the Coiflet-5 (coif5) wavelet to divide the raw data into five layers during the decomposition process. The decomposition equation is provided as follows:

$$\begin{aligned} X &= \sum_{j=0}^{J-1} \sum_{k} \langle X, \psi_{j,k} \rangle \phi_{j,k} + \langle X, \phi_{J,k} \rangle \phi_{J,k} \\ X &= \sum_{j=0}^{J-1} \sum_{k} \langle X, \psi_{j,k}^{coif5} \rangle \phi_{j,k}^{coif5} + \langle X, \phi_{J,k}^{coif5} \rangle \phi_{J,k}^{coif5} \\ X &= \sum_{j=0}^{J-1} \sum_{k} \langle X, \psi_{jk} \rangle \phi_{jk}. \end{aligned} \tag{6.3}$$

During the DWT procedure, the initial signal X is subjected to decomposition, generating wavelet and scaling components across various levels, limited by a maximum level J. The inner products $\langle X, \psi_{j,k} \rangle$ and $\langle X, \phi_{J,k} \rangle$ represent correlations between the signal and the wavelet and scaling functions, respectively. This decomposition facilitates the examination of RSSI data across diverse scales and a thresholding phase contrasts each data point in the RSSI signal with a predefined threshold of -70 dBm. Crossing the threshold indicates noteworthy activity detection, whereas points below the threshold imply the absence of detection.

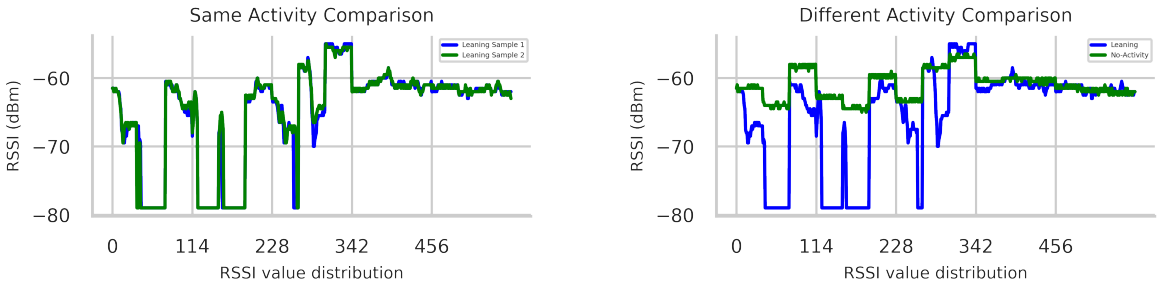
$$f(X(t)) = \begin{cases} X(t) & \text{if } X(t) \geq T \\ \text{Not detect} & \text{if } X(t) < T \end{cases}$$

Assessing the Feasibility for Fall Detection

A feasibility study used an RFID UHF tag array to detect various fall postures and prepare data for DL models as shown in Figure 6.4. The study focused on how different body movements affect RFID signal waveforms, particularly RSSI. The graphical representation in Figure 6.4a showed strong correlations in RSSI waveforms for repeated instances of the same leaning activity. Figure 6.4b demonstrated how various daily activities influence RSSI waveforms, distinguishing between leaning and no-activity. Clear fluctuations due to human activities were vis-

ible, while consistent patterns were evident without interference before or after activity. These results highlight the potential of using the RFID phase and RSSI waveforms to classify human activity attributes for detection.

The feasibility analysis confirmed that RFID signals can effectively capture and differentiate various fall and fall-related activities. However, to optimize system performance, addressing signal noise and implementing precise action segmentation methods during data preprocessing are critical challenges.



(a) Leaning Activities

(b) Leaning and No Activity

Figure 6.4: Data curves: Distinguishing between the same and different activities.

6.3 System Methodology

In this section, we employed the attention-based transformer model, which entirely avoids the use of decoder, recurrence, and convolutions. The proposed transformer architecture follows an encoder structure, effectively capturing their behavioral characteristics. To elaborate, the encoder, situated on the left side of the transformer architecture, maps an input sequence into continuous representations. Figure 6.6 illustrates the components of the transformer, showcasing stacked self-attention and point-wise fully connected layers in the encoder section, as depicted in the left half of Figure 6.5.

6.3.1 Model Architecture

Below is a detailed description of the model architecture:

a) Input Layer: The feature matrix (X) is fed into the model with a shape of (*number_of_features*, 1), where *number_of_features* represent the number of columns in the feature matrix.

b) Transformer Blocks: The core of the model consists of four identical transformer blocks, each comprising the following components:

- LayerNormalization followed by a *MultiHeadAttention* layer and a *Dropout* layer.

- The output of the Dropout layer is combined with the input of the transformer block using a residual connection.
- Another LayerNormalization layer follows.
- Two Conv1D layers, where the first one employs *ReLU* activation. The output of the second Conv1D layer is merged with the output of the first LayerNormalization layer using another residual connection.

c) GlobalAveragePooling1D Layer: This layer is implemented to reduce the model's output dimensions.

d) Dense Layers: The standard fully connected layers with *gelu* activation function are used. The number and sizes of these layers are determined by the *mlp_units* parameter.

e) Output Layer: The final layer of the model is another Dense layer, containing the same number of neurons as the number of classes. The activation function used is *softmax*, making this model suitable for multiclass classification tasks.

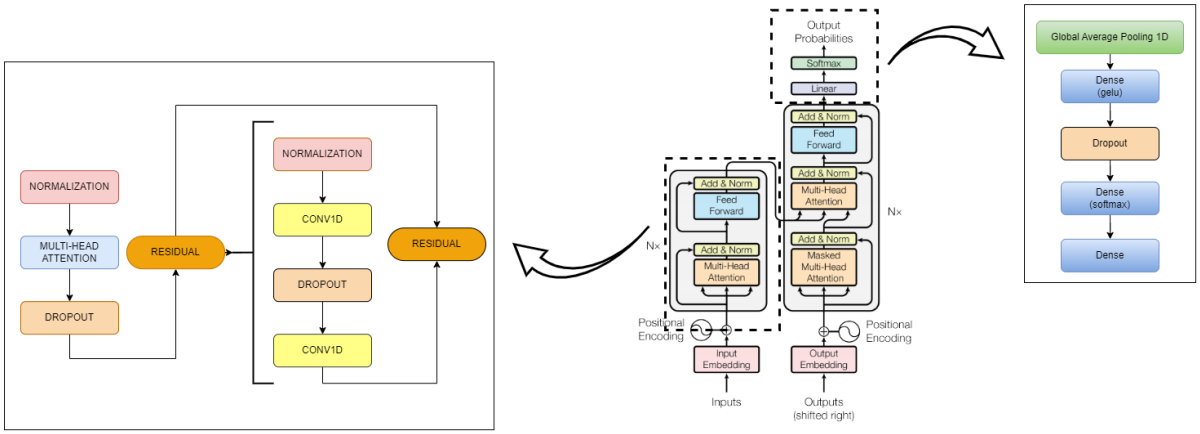


Figure 6.5: Comparative rrchitecture: Original Encoder-Decoder in the center with modified encoder-only versions on both sides

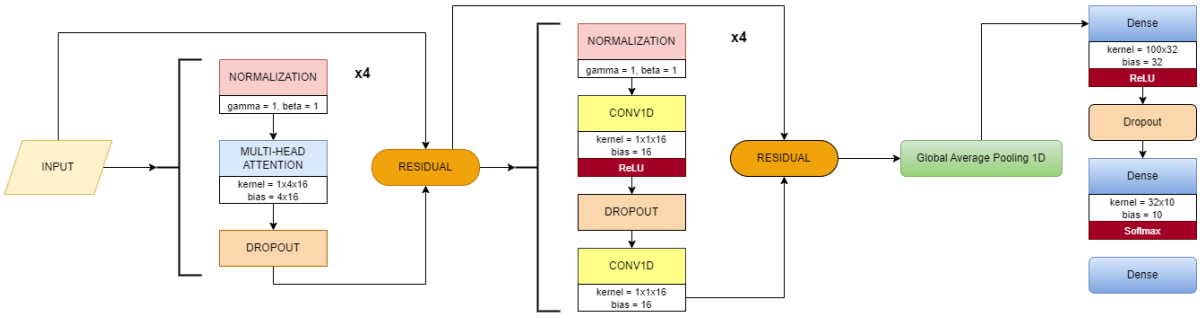


Figure 6.6: Proposed transformer encoder structure

6.3.2 Transformer Architecture Encoder Structure

The encoder consists of $N = 4$ identical layers, each containing two sublayers:

a) The first sublayer uses a 4-headed multi-head self-attention mechanism, with each head processing unique linear projections of queries, keys, and values, contributing to the final output.

b) The second sublayer includes two *Conv1D* layers, with the first layer utilizing a *ReLU* activation function.

$$\mathrm{FFN}(\mathrm{x}) = \mathrm{ReLU}\left(\mathrm{W}_1\mathrm{x} + \mathrm{b}_1\right)\mathrm{W}_2 + \mathrm{b}_2 \tag{6.4}$$

In the transformer architecture, each sublayer is augmented with a residual connection and followed by a *LayerNormalization* layer. This ensures proper normalization of the sublayer's input, denoted as 'x', and its output, *sublayer(x)*. However, the transformer architecture lacks inherent positional awareness due to its non-recurrent nature. To address this, positional encodings are introduced by adding them to the input embeddings, providing essential positional information.

$$\text{LayerNormalization}\left(\text{input} + \text{sublayer}(\text{output})\right) \tag{6.5}$$

6.3.3 The Transformer Multi-Head Attention

The attention function operates on query and key-value pairs, generating an output represented as vectors. The output is obtained through a weighted summation of values, with the weights determined by a compatibility function applied to the query and key. To retain the spatial information of RSSI and phase data, we avoid the use of positional embedding [254].

The feature extraction module of the transformer comprises two sub-layers: multi-head self-attention and multiscale residual CNN with adaptive scale attention. These sub-layers utilize a residual connection (Add) [255] and layer normalization (LayerNorm) [256]. Self-attention effectively handles long-term dependencies in sequences, surpassing the limitations of RNN and LSTM models. It captures global information from the entire sequence and overcomes the constraints of CNN's perception field and reliance on time-domain information. In the transformer, the input contains (*queries, keys, and values*), each having dimensions d_k and d_v respectively. These inputs undergo a *softmax* function to derive attention weights, which are then used to scale the values through weighted multiplication. The *multi-head* attention blocks in the transformer execute a scaled dot-product attention operation. This process can be summarized as follows:

$$\mathrm{Attention}(\mathrm{Q},\mathrm{K},\mathrm{V}) = \mathrm{softmax}\left(\frac{\mathrm{QK}^{\mathrm{T}}}{\sqrt{\mathrm{d_k}}}\right)\mathrm{V} \tag{6.6}$$

The attention mechanism in the transformer enables capturing dependencies between data sequence elements and extracting pertinent features for subsequent processing.

$$\begin{aligned} \text{MultiHead}(Q,K,V) &= \text{Concat}(\text{head}_1,\ldots,\text{head}_\text{h})W^O \\ \text{where head}_\text{i} &= \text{Attention}(QW_i^Q, KW_i^K, VW_i^V) \end{aligned} \tag{6.7}$$

In this study, we employ 8 parallel attention layers (or heads) with $h = 8$, each having 64 dimensions. This reduction in dimensionality ensures a computational cost similar to that of single-head attention with full dimensionality.

6.3.4 Architecture Comparison

The proposed architecture is adapted from Vaswani et al.[252] transformer model, as depicted in Figure 6.6. However, it differs from the comprehensive transformer model shown in Figure 6.5 as follows:

1. Encoder Only vs. Encoder-Decoder: The proposed model exclusively uses transformer *Encoder* blocks, similar to BERT [254]. In contrast, the original transformer incorporates both *Encoder* and *Decoder* blocks, primarily for sequence-to-sequence tasks like translation.

2. Global Average Pooling: The proposed architecture includes a *GlobalAveragePooling1D* layer after the transformer blocks to reduce output dimensionality for classification tasks.

3. Fully Connected Layers: The proposed architecture introduces *Dense* layers with dropouts after the transformer blocks, setting it apart from the original transformer.

4. Output Layer: For multi-class classification, the proposed architecture utilizes a Dense layer with *softmax* activation, while the original transformer uses a *final linear* layer followed by softmax for sequence-to-sequence tasks, predicting the next word in a sequence.

5. Positional Encoding: Unlike the original transformer model, the proposed architecture does not employ positional encoding.

6.4 Performance Evaluation

This section aims to assess the fall detection system's performance by employing diverse features, including RSSI, phase, and fusion of both (RSSI + phase). The effectiveness and accuracy of the proposed contactless RFID-based fall detection approach, utilizing an early fusion transformer model, will be evaluated through a set of comparative experiments.

6.4.1 Activity Recognition Methods

Activity recognition methods play an important role in precisely detecting and classifying human daily living activities, especially in indoor environments. In this subsection, we explore three essential approaches for activity recognition: *Activity Recognition using RSSI*, *Activity Recognition using Phase*, and *Activity Recognition using Fusion*. These approaches greatly enhance the accuracy and comprehension of various human activities within an indoor environment.

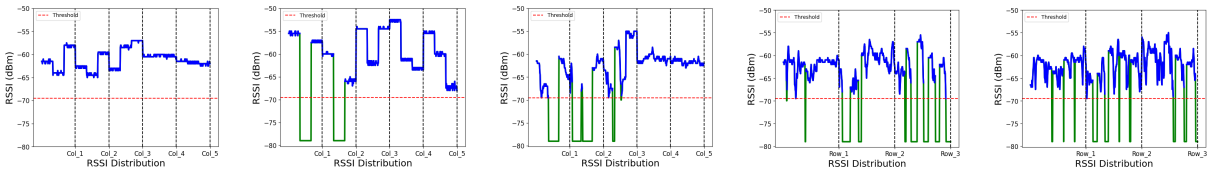
Activity Recognition using RSSI

The literature extensively discusses the use of passive UHF RFID tags for indoor activity recognition, particularly in fall detection. These RFID tags are activated by readers using air interface protocols (EPC Class 1 Gen-2 and ISO-18000-6c) for data transmission and reception [242]. During practical scenarios, passive RFID tags provide raw data to the reader in a 5-tuple format, consisting of RSSI, timestamp, EPC, TID, and frequency. Nevertheless, the process of creating an RSSI dataset entails several steps outlined in Algorithm 1. To generate an RSSI dataset, several preprocessing steps are performed as outlined in subsection 6.2.2.

In this study, the recognition of falls' activities is conducted using a meticulously designed experimental setup as depicted in Figure 6.3. A designated area in front of the *TRG-Wall* was used to perform five distinct activities, and the data collected for each activity is presented in a tabular format. To provide a comprehensive analysis of the collected data, we employed both column-wise and row-wise presentations of the tag arrangement data. Column-wise visualization proved effective in illustrating standing and leaning activities, while the row-wise representation is better suited for depicting falls and fall-related activities. Specifically, this was observed about normal falls and walking falls, with a particular emphasis on rows 2^{nd} and 3^{rd}. In contrast, columns 1^{st} and 2^{nd} primarily depicted leaning and standing activities, exerting minimal influence on the remaining columns.

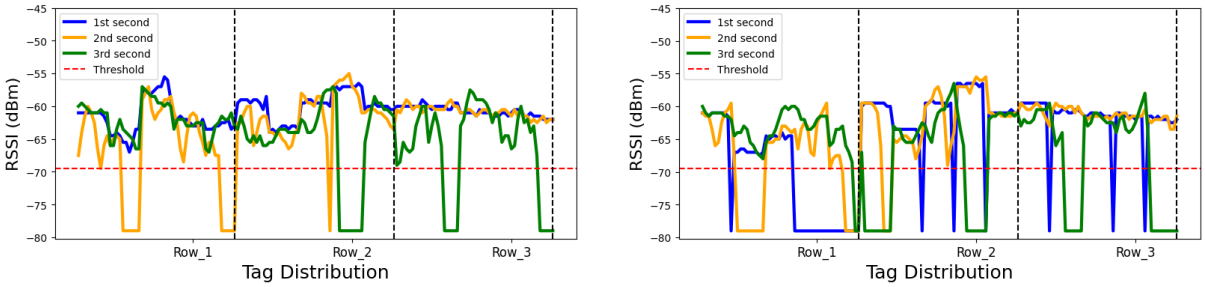
Building upon these insights, the compelling evidence supporting accurate fall and fall-related detection in shared spatial environments is presented in Figure 6.7. Recorded RSSI strengths ranged from $-50dBm$ to $-69dBm$, with obstructed tags leading to decreased RSSI readings. The red dotted line serves as an activity recognition threshold, meticulously set to address scenarios involving non-reading or tag-blocking. The use of green highlights instances of RSSI data reading obstruction or activity recognition challenges. Specifically, Figures 6.7b and 6.7c depict RSSI values for standing and leaning activities, respectively, in rows 2^{nd} and 3^{rd}. Meanwhile, Figures 6.7d and 6.7e showcase normal and walking fall activities performed in the same location.

To enhance the representation of falling events, we split the RSSI data linked with three-second falling patterns into one-second intervals and employ distinct color codes for visual clarity. The color scheme is as follows: blue for the 1^{st} second, green for the 2^{nd} second, and black for the 3^{rd} second of the falling pattern data. This visual representation effectively demonstrates



(a) No Activity (b) Standing (c) Leaning (d) Normal Fall (e) Walk Fall

Figure 6.7: Analysis of falling activity recognition: Distribution and magnitude of RSSI (row-wise and column-wise)



(a) Normal Fall (b) Walk Fall

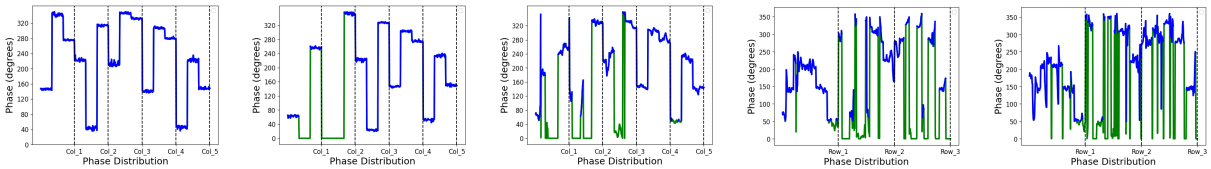
Figure 6.8: Visual representation: Analysis of falling per second.

the sequential progression of the fall activity. For instance, in Figure 6.8, the initiation of the fall by the subject is evident during the 1^{st} second (blue color) from row-1 to row-3. This is followed by movement observed during the 2^{nd} second (green color) and the 3^{rd} second (black color) from row-2 to row-3, signifying the progression of the fall activity in a row-wise manner.

Activity recognition using Phase

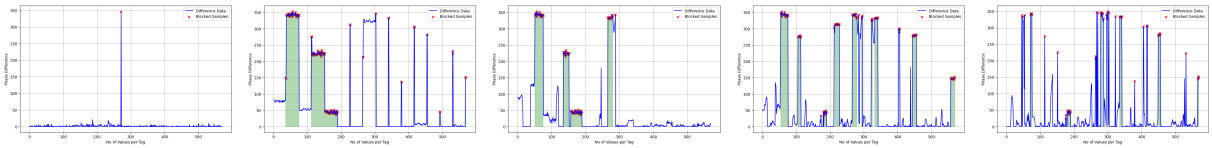
RF backscatter technology enables bidirectional signal transmission over a distance of $2d$, making it possible to monitor human activity by analyzing phase differences in RF features using cross-correlation. The relationship between distance, antenna phase rotation, and tag phase rotation can be mathematically described in Equations 5.5 and 5.6.

Figure 6.9 showcases the analysis of falling activity recognition, both row-wise and column-wise, utilizing phase information. Figure 6.10 illustrates the significance of phase difference patterns in fall activities. These patterns were obtained using the *numpy* function $np.corrcoef(x,y)$, which quantifies the cross-correlation between two sets of data. For instance, in Figure 6.10b, the standing activity demonstrates the phase difference between an empty state and a standing position. The red cross indicates the detection or blockage of phase, signifying the occurrence of activity detection. The resulting cross-correlation difference pattern proves to be an effective method for modeling activities. Notably, the smooth variation of phase differences observed when blocking tags were present during sitting activities emphasizes the accuracy and reliability of the measurements.



(a) No Activity (b) Standing (c) Leaning (d) Normal Fall (e) Walk Fall

Figure 6.9: Phase-based analysis of falling activity recognition: Row-wise vs Column-wise



(a) No Activity (b) Standing (c) Leaning (d) Normal Fall (e) Walk Fall

Figure 6.10: Phase difference-based analysis of falling activity recognition: Row-wise vs Column-wise

Activity recognition using Fusion

Activity recognition through fusion involves integrating data from multiple modalities to enhance accuracy, specifically by leveraging both RSSI and phase-based information to optimize activity recognition algorithms. This fusion process can be accomplished using two primary concepts: early fusion and late fusion.

1) Early Fusion in RFID-based Activity Recognition: Early fusion, or feature-level fusion, combines unprocessed RSSI and phase data from RFID tags before classification. This integration reduces redundancy and complexity in features, resulting in more efficient models for quicker inference and reduced computational overhead. It enhances accuracy, robustness, and adaptability in RFID-based human activity detection, enabling effective recognition of various activities in indoor environments.

The algorithm for early fusion in RFID-based activity recognition involves the following steps:

1. Initialize empty datasets for RSSI and phase data: $\mathbf{D}_{\text{RSSI}}$ and $\mathbf{D}_{\text{Phase}}$.
2. Preprocess RSSI and phase data separately to create $\mathbf{D}_{\text{RSSI}}$ and $\mathbf{D}_{\text{Phase}}$.
3. Combine $\mathbf{D}_{\text{RSSI}}$ and $\mathbf{D}_{\text{Phase}}$ to create $\mathbf{D}_{\text{Combined}}$ containing integrated information.
4. Extract features from $\mathbf{D}_{\text{Combined}}$ using the *feature_extraction* method, resulting in $\mathbf{F}_{\text{Combined}}$.
5. Perform classification on $\mathbf{F}_{\text{Combined}}$ to predict activity labels, storing them in $\mathbf{Y}_{\text{Pred}}$.

The algorithm for early fusion in RFID-based activity recognition can be concisely summarized as follows:

Algorithm 3 : Pseudo code for early fusion

Initialize empty datasets: D_{RSSI}, D_{Phase}, and $D_{Combined}$
$D_{RSSI} \leftarrow$ RSSI_data_preprocessing(D)
$D_{Phase} \leftarrow$ Phase_data_preprocessing(D)
$D_{Combined} \leftarrow$ combine_data(D_{RSSI}, D_{Phase})
$F_{Combined} \leftarrow$ feature_extraction$(D_{Combined})$
Perform classification: $Y_{Pred} \leftarrow$ classification$(F_{Combined})$
return Y_{Pred}

2) Late Fusion with merit in RFID-based Activity Recognition: Late fusion, or decision-level fusion combines independent decisions from RSSI and phase data for falling and daily living activity prediction. Separate classifiers are trained for each modality (RSSI and phase). One classifier recognizes falling patterns using RSSI, and the other relies on phase data. These classifiers make independent predictions. Late fusion employs two methods: voting-based and merit-based. In the voting-based approach, the final decision is based on majority votes from individual classifiers while merit-based fusion optimizes modality weights based on performance and reliability measures. We use merit-based to avoid bias in classifiers. This approach leverages the complementary information from RSSI and phase data, enhancing robust falling event detection, especially in complex indoor environments. The pseudo-code of late fusion with a merit-based algorithm is summarized as follows:

Algorithm 4 : Pseudo code of merit-based late fusion

$D_{\text{RSSI}} \leftarrow$ RSSI_data_preprocessing(D)
$D_{\text{Phase}} \leftarrow$ Phase_data_preprocessing(D)
$F_{\text{RSSI}} \leftarrow$ feature_extraction(D_{RSSI})
$F_{\text{Phase}} \leftarrow$ feature_extraction(D_{Phase})
$Y_{Pred\_RSSI} \leftarrow$ classification(F_{RSSI})
$Y_{Pred\_Phase} \leftarrow$ classification(F_{Phase})
$M_{\text{RSSI}} \leftarrow$ calculate_merit(Y_{Pred_RSSI})
$M_{\text{Phase}} \leftarrow$ calculate_merit(Y_{Pred_Phase})
$Y_{Pred\_Combined} \leftarrow$ fusion_with_merit$(Y_{Pred_RSSI}, Y_{Pred_Phase},$
$M_{\text{RSSI}}, M_{\text{Phase}})$
return $Y_{Pred\_Combined}$

6.4.2 Experimental Results

We evaluated our contactless *TFree-FD* method using an 80 : 20 train-test split (random state: 42), with 32 batch size, 50 epochs, and a dropout rate of 0.01 during training. For a comprehensive view of network hyperparameters, discussed in Table 6.2. Our approach was compared against four prominent deep learning models: RF-finger [257], LiteHAR [258], Tagfree [251],

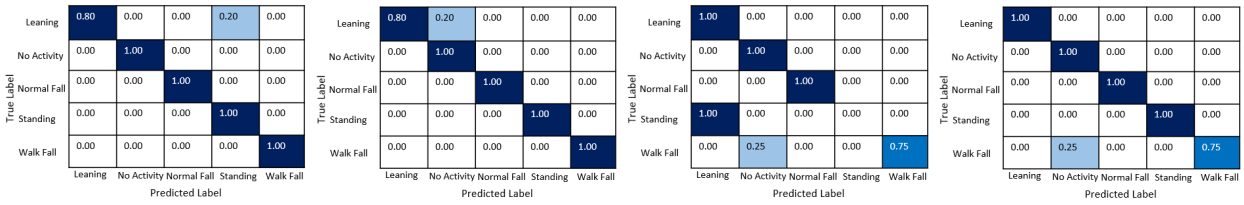
and Dense-LSTM [259]. These models represent benchmarks for device-free RF-based activity detection using traditional deep-learning methods

Table 6.2: Proposed transformer model hyperparameters.

S.No	Components	Parameters
1	LayerNormalization	gamma (1) and beta (1)
2	Multi-Head Attention	Kernel $(1 \times 4 \times 64)$, bias (4×64)
3	Dense-gelu	Kernel (100×32), bias (32)
4	Dense-Softmax	Kernel (32×10), bias (32)
5	Conv1D	Kernel $(1 \times 1 \times 64)$, bias (4×64), ReLU

Overall Performance of TFree-FD

The effectiveness of the *TFree-FD* technique, which employs a transformer, is demonstrated in Figure 6.11. The confusion matrix in the figure reveals a distinct diagonal pattern for RSSI, phase, early fusion, and late fusion results. The Figure 6.11a achieves an accuracy of 93.3%. Similarly, Figure 6.11b achieves 94% accuracy for RSSI, while the transformer model fused with late fusion in Figure 6.11c shows 86.2%. Remarkably, the late fusion in Figure 6.11d achieves the highest accuracy at 96.5%. Comparing these results to the baselines in Table 6.3, the transformer outperforms significantly, consistently achieving the highest accuracy and F1 score across all categories. Specifically, it establishes new benchmarks on our experimental dataset, surpassing RF-finger, Dense-LSTM, LiteHAR, TagFall, and Tagfree by 11.1%, 6.5%, 5.5%, 4.8%, and 3.7% in average accuracy, respectively. Moreover, our proposed F1-score outperforms RF-finger, TagFall, Tagfree, Dense-LSTM, and LiteHAR by 12.1%, 8.7%, 5.6%, 4.0%, and 2.9%, respectively. These results underscore the exceptional performance of our proposed encoder-based transformer architecture in fall detection.



(a) Phase (b) RSSI (c) Late Fusion (d) Early Fusion

Figure 6.11: Visualization of the normalized confusion matrix pre and post fusion.

Comparative Analysis of Existing Methods

We compared our fusion-based transformer model with common models (CNN, RNN, LSTM) using accuracy and prediction time. Results in Table 6.4 and 6.5 show our model's superior accuracy and F1-scores (epochs = 50). Unlike CNN, RNN, and LSTM using convolutional

Table 6.3: Comparing performance of the proposed model with baselines.

S.No	Appraoch	Accuracy	F1 Score
1	RF-finger [257]	85.4 %	83.6
2	Dense-LSTM [259]	90 %	87.0
3	LiteHAR [258]	91 %	90.1
4	TagFall [128]	91.7 %	91.7
5	Tagfree [251]	92.8 %	92.8
6	**Proposed**	**96.5 %**	**95.7**

layers, our transformer model stands out with self-attention mechanisms for better accuracy. However, due to its extensive self-attention use and parameters, the transformer requires more computational resources, leading to less efficient prediction times compared to other methods.

Table 6.4: Comparing accuracies among different algorithms

Algorithms	Scenario-1 Moderate SNR (3.5)		Scenario-2 High SNR (4.5)	
	RSSI	Phase	RSSI	Phase
CNN	82.3%	79.5%	73.7%	71.9%
RNN	85.1%	78.2%	77.5%	70.7%
LSTM	88.8%	75.1%	80.6%	74.8%
Transformers	94%	93.3%	84.1%	77%
Transformers (Late Fusion)	86.2%		81.4%	
Transformers (Early Fusion)	**96.5%**		**87.9%**	

Table 6.5: Comparing F1-scores among different algorithms

Algorithms	Scenario-1 Moderate SNR (3.5)		Scenario-2 High SNR (4.5)	
	RSSI	Phase	RSSI	Phase
CNN	0.64	0.58	0.46	0.43
RNN	0.70	0.56	0.54	0.41
LSTM	0.78	0.50	0.60	0.48
Transformers	0.88	0.86	0.68	0.54
Transformers (Late Fusion)	0.72		0.62	
Transformers (Early Fusion)	**0.93**		**0.74**	

Comparing Head Count's Influence on Transformer Model Accuracy

The effectiveness of transformer models is notably impacted by the number of attention heads within their architecture. This effect is demonstrated in Table 6.6, offering a comparative accuracy analysis for different head counts (epoch = 10). The accuracy of the model fluctuates with varying head counts, highlighting the influential role of headcount in determining model effectiveness. The results reveal a discernable pattern where increasing headcount generally improves accuracy, although this relationship may not be strictly linear. For example, as headcount

increases from 2 to 6, accuracy consistently rises, peaking at 100%. However, further increasing the headcount to 8 leads to a slight accuracy reduction. This underscores the complex interplay between model complexity, attention mechanisms, and accuracy. Therefore, consideration of headcount is important when optimizing transformer models for optimal performance.

Table 6.6: Comparing number of heads (epoch =10)

Number of heads	2	4	6	8	10
Accuracy	76.3%	92.5%	100%	86.0%	81.2%

Comparing Convolution Utilization and Analyzing 1D and 2D Techniques' Impact on Accuracy

We conducted an extensive study to compare the impact of utilizing 1D and 2D convolutions on model accuracy. The analysis spanned ten epochs and employed four attention heads. The results, presented in Tables 6.7 and 6.8, reveal a significant accuracy difference between the two convolution types across epochs. Specifically, 1D convolution demonstrated robust accuracy at 92.5%, surpassing the comparatively modest 68.1% accuracy achieved by 2D convolution. This contrast underscores the intricate interplay of convolution methodologies and underscores the pivotal role of selecting an appropriate technique to enhance transformer model precision.

Table 6.7: Comparing 1D and 2D convolution (epochs = 10 vs 50 and heads = 4)

Convolution Configuration	**Epochs = 10**		**Epochs = 50**	
Convolution Type	1D	2D	1D	2D
Accuracy	92.5%	68.1%	96.5%	75.4%

Table 6.8: Comparing epochs vs accuracy's

Epochs	10	30	40	50	70	100
Accuracy	91.9%	90.7%	93.7%	96.5%	100%	100%

6.5 Discussion

This study undertook a comprehensive evaluation of four deep learning-based models—CNN, RNN, LSTM, and the attention-based transformer—leveraging both early and late fusion methodologies for RFID-based fall activity recognition across two distinct scenarios. The investigation encompassed three well-defined approaches (RSSI, Phase, and fusion), expounded upon in Section 6.4.1. The findings illuminate the transformer model's exceptional performance when coupled with the early fusion technique, surpassing the capabilities of CNN, RNN, and LSTM

across both scenarios. These scenarios entailed the strategic placement of a reader and antenna at a 3.5-meter distance from the subject, with an additional subject-TRG-Wall distance of 0.5 meters. Analysis of the normalized confusion matrix unveils the transformer model's capacity to achieve recognition accuracy exceeding 99% for the majority of activities, with the exception of walk fall, which demonstrated confusion with leaning activity. This observation underscores the potential of the proposed approach for practical user recognition applications, as highlighted in Table 6.4. Significantly, the efficacy of the transformer model, particularly when coupled with early fusion, for contactless RFID human activity recognition, remains contingent on the distance between the TRG-Wall and the reader antenna.

6.5.1 Ablation Studies

The ablation study systematically analyzes components affecting system performance. It assesses user diversity and location impacts, explores antenna height effects on RFID tag reading, and examines the advantages of combining RSSI and phase data. These insights contribute to robust fall detection.

Impact of User Diversity and Location on Performance

To assess the system's stability and generalization ability, we conducted experiments with multiple subjects, utilizing distinct fall-related activity data without any training or validation sets. Under the same system deployment mode, there are five distinct activities, including falls, with predefined distances of 3.5 and 4.5 meters between the antenna and the tag. Figure 6.12 illustrates the model's adaptability for fall detection across different activities. We further examined the impact of the target subject's position within the fall perception system on detection accuracy. At 3.5 meters, accuracy exceeded 98%, except WalkFall achieving 94%. This variation indicates that RFID signals are influenced by the multipath effect in the physical environment, leading to a slight decline in recognition accuracy for non-preset positions. Nevertheless, the system maintains satisfactory recognition performance. This underscores the robustness of the proposed fall detection method, ensuring optimal performance in indoor environments.

Impact of Antenna Height

This study explores the impact of antenna height on RFID tag reading performance. Antenna placement has a significant effect on range and precision. Two placements were tested: The initial setup mounted the antenna on a ground-level wall measuring $1.5 \times 1.5\ m^2$. This affected the reading of top-row tag RSSI/phase data. Adjusting the height to 0.75m, aligning with the wall's center and LoS, aimed to extend the range with minimal accuracy impact. However, raising it further to 1.5m weakened the signal to the lower tag row due to insufficient reader strength. Results suggest maintaining the default 0.75m height for optimal system performance.

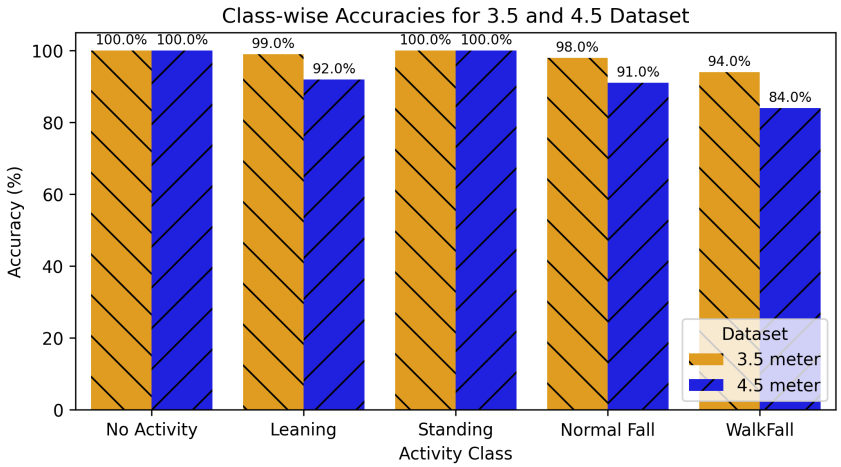
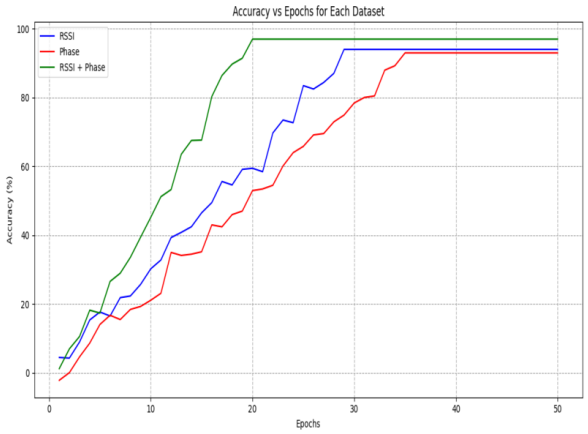


Figure 6.12: Comparison of class-wise accuracies between 3.5-meter and 4.5-meter datasets.

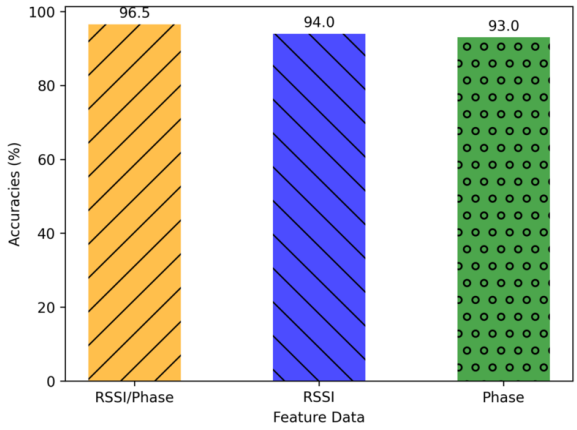
Impact of Multimodal Analysis

In this section, we explore the impact of multimodal features on recognition accuracy. To evaluate the system's performance, we initially select RSSI, phase, and fusion signals as sample data. The incorporation of multiple signal types aims to harness a broader spectrum of features, potentially leading to enhanced recognition accuracy. We proceed by training and validating the model using three datasets. The accuracy curve of the test set during the training process is depicted in Figure 6.13a. The results indicate that, under identical training conditions, the combination of RSSI and phase data as feature inputs facilitates faster convergence and yields higher accuracy during the training process. Our experimental findings demonstrate that the fused feature data outperforms the individual phase or RSSI features in terms of fitting speed and final accuracy, as illustrated in Figure 6.13b. This superiority arises from the fact that the phase signal is more sensitive to environmental factors, while the RSSI resolution diminishes with increasing distance from the antenna. Therefore, employing fused data features becomes preferable for the fall detection system.

Figure 6.13b demonstrates that combining RSSI and phase data yields higher accuracy compared to using either RSSI or phase data alone. Building on this finding, we investigate the impact of distance and the number of tags on the system's recognition performance. The accuracy results are presented in Table 6.4. The experiment reveals that increasing the distance from 3.5 to 4.5 adversely affects the model's recognition effectiveness. As the fall detection system prioritizes swift recognition, incorporating additional RFID tags in close proximity to the activity area significantly reduces the system's sampling rate, while also imposing a higher computational burden on the model. Therefore, to maintain a reasonable computational cost and meet the fall detection requirements in indoor scenarios, we have adopted a layout with fifteen tags, as confirmed in this study. This configuration satisfies the system's needs and ensures efficient fall detection.



(a) Impact of Phase, RSSI, and Fused (Phase + RSSI) Data on Model Performance



(b) Assessment of accuracy across multimodal data features.

Figure 6.13: Comparison of model performance based on different data representations.

6.6 Conclusion, Limitations and Future Directions

This paper investigates the impact of RFID tag signals on monitoring the activity of the elderly in a contactless manner. The study focuses on using phase and RSSI features as inputs for a deep neural network, specifically a transformer model, to develop a fall detection method. We collected a dataset comprising multimodal data, including phase and RSSI, from five different daily human activities. The data underwent preprocessing steps, such as phase calibration, noise reduction, action segmentation, and normalization, before training the transformer model. The trained model achieved a high accuracy of 96.5% in two distinct scenarios. The experimental results demonstrate the robustness of the contactless fall detection system against variations in users and locations. However, the proposed fall detection method in this paper still has certain limitations that require further improvement in future research.

1. **Multi-subject Fall Detection:** The proposed method detects falls in one person only and doesn't address detecting falls in multiple individuals due to challenges with contact-less deployment and signal separation. Research is ongoing to explore contactless real-time fall detection for multiple users using human body feature signals.

2. **Collecting Authentic Human Fall Data:** The collected fall data may not fully represent real falls due to challenges in collecting authentic human fall data. Falls can be categorized as object-related or fainting-induced, with real falls being sudden and unpredictable. Controlled environments in experiments capture diverse fall actions but differ from real falls. Training the system on simulated data may lead to deviations when detecting sudden falls in real bodies. Collecting more real fall data is vital for enhancing the system's accuracy in detecting unforeseen falls.

3. **Scalability and Application Potential:** The proposed system, initially tested in a 10×10

m^2, excels at detecting falling-related activities within designated area. The contactless RFID-based scheme, known for its remarkable scalability, can easily extend its coverage to larger spaces by adding more tags and antennas. This scalability not only maintains a low average cost but also positions it as an ideal solution for healthcare facilities, nursing homes, smart homes, and other environments where cost-effectiveness and non-intrusiveness are essential, making RFID tags a viable and adaptable solution for widespread adoption.

4. **Environment Dependency and Training Costs:** The phase and RSSI signals depend on the deployment environment of the system, which can vary in practical applications (e.g., hospitals, homes, and nursing homes). Different objects and environments can reduce accuracy. Adding sample data or training models in new environments incurs high learning and training costs for RFID. Future work should prioritize passive sensing across diverse environments and devices.

6.7 Summary

This chapter summarised an innovative RFID-based method for contactless fall detection in the elderly, employing strategically placed passive UHF RFID tags. The system operates without the need for uncomfortable wearables. Through tag querying and data processing using a transformer model and early fusion technique, we achieved a remarkable average accuracy rate surpassing 96.5%, thus demonstrating the efficacy and practicality of our approach over conventional methods like CNN, RNN, and LSTM. Furthermore, the study introduces RFIDAR, an innovative fusion system that marks a critical advancement in HARI within indoor environments, by combining RFID with Radar technology to address the privacy and cost limitations of single-sensor systems. This fusion capitalizes on the unique sensing capabilities of both systems, employing a sophisticated algorithm to accurately identify complex human activities, especially in long-range and non-line-of-sight scenarios.

Chapter 7

Enhancing Human Activity Recognition with RFiDAR: Integrating RFID and Radar Data for Contactless Detection

This chapter introduces the innovative fusion system RFIDAR, a critical advancement in human activity recognition HAR within indoor environments. Addressing the privacy and cost limitations of single-sensor systems, RFIDAR combines radio frequency identification (RFID) with radar technology. This fusion leverages the unique sensing capabilities of both systems, using a sophisticated algorithm to accurately identify complex human activities, especially in long-range and non-line-of-sight scenarios. Our research involved testing RFIDAR on individuals performing a range of activities, including sitting, standing and walking direction. We employed an advanced LSTM-based variational autoencoder (LSTM-VAE) fusion model by its superior capability in capturing long-term dependencies and temporal features from sequential sensor data. Additionally, we explored two different data fusion algorithms, finding significant improvements in HAR accuracy, with increases of 5.8%, and 7.9%, for data-level and feature-level fusion algorithms, respectively. The non-intrusive, passive nature of RFIDAR, combined with its precision and robustness, demonstrates its immense potential to enhance the quality of life for elderly patients in assisted living environments.

7.1 Introduction

Human activity recognition in an indoor environment is imperative for advancing modern living standards and preemptive safety measures. The demand for an accurate, cost-effective, and unobtrusive monitoring solution has highlighted the importance of contactless RF sensing technologies. Leveraging RFID and radar-based methodologies presents a promising pathway to elevate monitoring capabilities while prioritising individual privacy and comfort. These methodologies offer versatile applications across various medical domains, including elderly monitor-

ing [260, 261], smart living [262], and health care [263]. Existing HAR methodologies can be classified into vision-based and sensor-based approaches. Sensor-based methods encompass wearable, object-tagged, and device-free sensors [264]. Vision-based methods, while offering accurate results by analysing recorded videos, face challenges related to occlusion, viewpoint restrictions, lighting dependencies, and privacy concerns [234, 265]. Wearable sensors (gyroscopes, accelerometers, and heart rate monitors), widely utilised due to their diversity and affordability, encounter issues of discomfort and reliability, especially among neurodegenerative patients and the elderly [266]. Object-tagged sensors, which involve tags attached to daily-use objects for activity recognition, address the obtrusiveness of wearable sensors but rely on user interactions [267]. In contrast, device-free sensors employing technologies such as Wi-Fi, radar, and RFID offer an appealing alternative offering improved privacy and performance in complex indoor settings with multiple signal paths [268, 269]. However, wearable sensor-based approaches, although precise, lack practicality in real-life situations, particularly for elderly individuals or those with dementia. Notably, radar-based solutions exhibit accurate real-time monitoring in scenarios involving multiple targets [236]. A cost-effective approach combines UHF RFID readers with battery-less and compact RFID tags, making it advantageous for HAR applications. Table 7.1 summarises the pros and cons of various HAR methods.

Table 7.1: Comparison of sensors used for human activity recognition.

Methods	**Examples**	**Pros**	**Cons**
Wearable	Accelerometer, gyroscope, microphone, smartphone, etc	preserve privacy, low cost, usable for vital sign acquisition and good accuracy	battery-powered, intrusive, power consumption and sensor specified to activity
object-attached	RFID, Proximity sensors	unobtrusive, preserve privacy, low sensitive to static objects, good accuracy	limited to tag objects, limited to activities and relatively costly
Vision-based	RGB-D camera, Kinetic, etc	unobtrusive, good accuracy	sensitive to occlusion, illumination, require LoS, high cost and computational resources
Device-free	pyroelectric infrared sensors, RFID, WiFi, Radar, etc	better performance in complex environments, privacy preserve, unobtrusive, low cost	sensitive to static objects, low accuracy, and sensitivity to noise.

Device-free human sensing encompasses various sensing technologies such as motion sensors, cameras, thermal imagers, light sensors, and others. However, certain sensors like motion detectors, thermal imagers, and cameras, require specific deployment for human sensing, making their ubiquitous availability limited. Furthermore, privacy concerns arise with sensors like

microphones and cameras. In contrast, radio signals offer unique advantages; they are widely available, including WiFi signals at home, and are non-intrusive regarding privacy compared to cameras and microphones. Radio signals possess unique capabilities, allowing penetration through walls and functioning in darkness. RF-based device-free human sensing has emerged as a significant area of research with considerable advancements, with emerging commercial solutions by various startups [270, 271, 272] applications such as sleep monitoring, vital sign monitoring, fall detection, localisation, and tracking, activity monitoring, and people counting.

In recent years, RFID and radar have gained significant attention due to their cost-effectiveness and non-intrusive nature in human sensing and monitoring. RFID technology, widely used in remote elderly healthcare monitoring [237], is known for its cost-effectiveness, compactness, scalability, shareability, and battery-free operation [238]. Recent advancements include highly sensitive passive tags with a read range exceeding 10 meters, making it viable for pervasive healthcare solutions [231]. The innovation of 'tag-free' sensing in RFID, eliminating the need for direct tag attachment, addresses challenges in ambient assisted living (AAL) [240], facilitating the recognition of activities like standing, sitting, running, and walking without the cumbersome nature of attaching tags [264]. However, RFID sensors have limitations such as limited information, space constraints, and deployment complexity due to managing various communication protocols. Nonetheless, accurate results for indoor AAL and positioning necessitate accounting for various factors such as non-line-of-sight (NLOS) conditions, signal weakening, indoor material effects inducing signal reflection and refraction, and climate-induced changes affecting signal propagation speed. Meanwhile, radar, an active sensing technology, operates by transmitting and capturing electromagnetic waves from objects. Continuous-wave radars like Doppler [273] and frequency-modulated continuous wave (FMCW) [274] radars are widely utilised in HAR. Doppler radars determine body part velocities, generating micro-doppler signatures essential for HAR, while FMCW radars measure target distances efficiently. Radar spectrograms offer resilience in various environments, ensure visual privacy, and enable through-wall detection. Challenges like lack of precise range information due to its narrowband nature in continuous-wave (CW) radar [275] and indoor performance issues in FMCW radars due to multipath interference [276] are addressed by alternative solutions like ultra-wideband impulse radar (UWB), offering higher resolution information through high-frequency pulse signals [275].

Earlier studies in RF-based human sensing made extensive use of conventional machine-learning (ML) techniques to classify human actions and contexts by extracting manually designed features from radio signals [277]. While these methods showed promise in smaller-scale experiments, their accuracy struggled when deployed on a larger scale. Presently, researchers are increasingly turning to advancements in deep learning (DL) to enhance the accuracy, scalability, and universality of RF sensing. This transition is evident from the growing number of publications in prominent conferences and journals (as illustrated in Figure 7.1), which explore various deep neural network architectures and algorithms to advance RF-based human sensing. Recent

observations highlight the superior suitability of DL intricate network structures for supervised and incremental learning, outperforming traditional machine learning approaches [278]. Deep learning, being instrumental in comprehending complex information embedded in raw RFID and radar data, assumes significance due to the complex variations inherent in the data [279, 280]. Specifically in HAR applications, DL's capacity to automatically extract intricate features surpasses that of traditional methods [281]. Notably, recurrent neural network (RNN) models have demonstrated efficacy in processing time-series data and recognizing diverse human actions, encompassing gestures, gaits, and language [282, 283]. Incorporating long short-term memory (LSTM) units into RNNs has addressed challenges associated with long-term learning [284]. Given the success of deep learning models and considering the time-dependent nature of RFID and radar data, the proposed LSTM-based variational autoencode (LSTM-VAE) model augmented with LSTM units holds promise for precise HAR classification within the RFiDAR system.

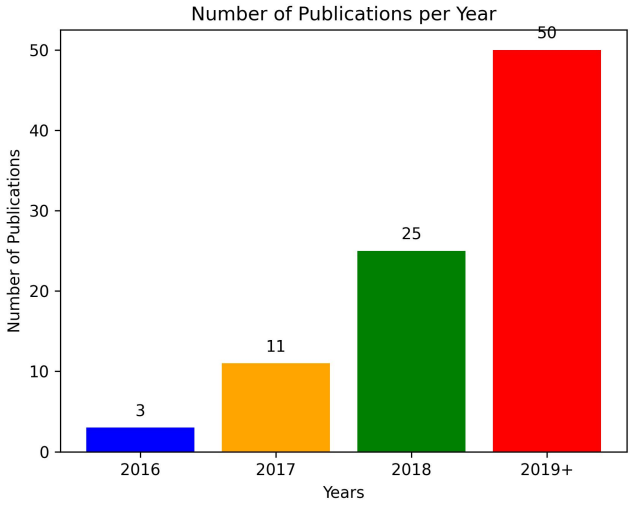


Figure 7.1: Recent growth in deep learning application for RF-based human sensing publications. [285]

Several single-modality and single-type ambient sensing solutions rely on DL algorithms for accurate HAR. However, these solutions have limitations when using only one sensor or relying on a specific sensor [286]. For instance, low recognition accuracy may persist for similar activities, and specific sensor solutions are susceptible to external interference such as temperature variations, multipath interference, LOS issues, and long detection ranges. Each sensor type also faces its challenges; for instance, RFID systems are influenced by LOS dependence and environmental factors while radar systems may face challenges in precisely detecting human movement over long distances. Addressing these challenges and aiming to provide a highly accurate HAR system suitable for long-term human monitoring, we propose RFiDAR, a signal fusion system leveraging information from RFID devices and radar radiation. *RFiDAR* adopts an LSTM DL model with VAE units to comprehend complex scenarios in multiple data streams. This fusion

approach was validated using data from ten subjects performing five different activities. This system underscores the effectiveness of the sensor fusion framework for independent, long-term human monitoring solutions at home.

To ensure interpretability and transparency in predictions made by the initial black-box system. For interpretability, which refers to connecting data input to AI predictions, and transparency, enabling human users to understand the algorithm's process. The contributions of this work are outlined below:

- Introduction of RFiDAR - an interpretable, passive, and non-intrusive fusion system for accurate and transparent HAR, designed to be inexpensive.
- Mitigation of limitations seen in single modality solutions, such as RFID LOS and ambient dependence, and radar susceptibility to electronic interference. RFiDAR offers a robust and reliable classification system for HAR.
- Utilisation of LSTM-VAE Fusion to demonstrate the system's ability to reduce the impact of LOS and detection range, ambient dependence, and electronic interference, thereby enhancing visual application prospects.

7.2 System Model and Preliminaries

This section introduces the conceptual framework and fundamental components of the RFiDAR system, setting the stage for a detailed exploration of its innovative multi-modal approach to HAR in challenging environments.

7.2.1 System Model

This chapter presents a novel multi-modal system, RFiDAR, tailored for HAR, addressing limitations in RFID and radar-based systems, especially in long-range and LOS environments. The RFiDAR system combines an RFID reader and antenna, RFID tags on a wall, and a UWB Xethru radar connected to a laptop. Subjects engage in distinct activities while RFID readers and radar collect data in a contactless manner, capturing tag ID, phase angle, RSSI, Doppler frequency, timestamps, and raw IQ data, all stored in CSV and .dat formats. This time-series data enables the RFiDAR system to identify subject-tag interactions and activities within the designated area. Integrating UWB radar and RFID data, our system employs data-level and feature-level fusion techniques to heighten accuracy. Fundamental to our approach is sensor calibration through data and signal level fusion, coupled with the collection and preprocessing of diverse human activity data from multiple sensors to prepare inputs for our model. For precise HAR, we propose an LSTM-VAE model complemented by fusion algorithms for data classification, yielding defini-

tive outcomes. A visual representation of our multi-sensor system is depicted in Figure 7.2.

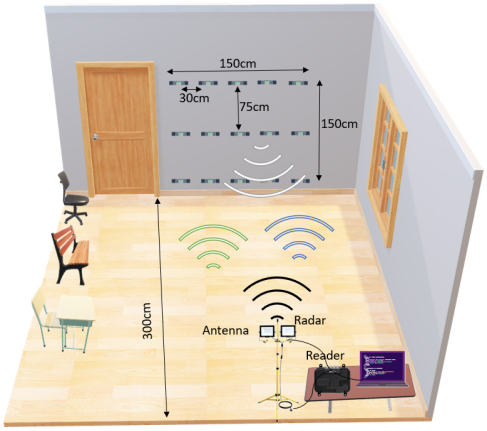


Figure 7.2: Experimental setup of RFiDAR system.

7.2.2 Preliminaries

In this section, we will describe the foundational aspects of RFID technology and Xethru radar.

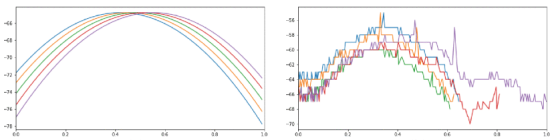
RFID The reader uses backscatter communication to interrogate tags. In this process, electromagnetic waves emitted from the reader antenna interact with a tag, causing the tag's tiny antenna to modulate its resonant properties and embed ID information into the backscattered signal. When multiple tags simultaneously backscatter signals, signal collisions occur on the reader side, resulting in the reader receiving no meaningful data. To address this issue, various anti-collision protocols, such as framed slotted Aloha protocols or tree-walking protocols [287], were proposed. Subsequently, following each successful tag reading, the reader can report not only the tag ID but also certain low-level data related to tag signals, including phase, RSSI, doppler frequency, and timestamp. The RSSI value, measured in dBm, ranges from negative values up to 0, with 0 representing ideal conditions where a tag receives the full signal from the antenna. In an ideal scenario, when an antenna scans through tags, its RSSI values exhibit periodic and hierarchical patterns. Figure 7.3 [288] illustrates this with five differently colored lines, each representing a tag. The left-right relationship between any two tags is distinctly visible in the figure. RSSI and Phase are influenced by factors like distance, medium, and others, commonly following the log-distance path loss model [289].

$$\mathrm{PL}(d)(\mathrm{dB}) = \mathrm{PL}(d_0) + 10n\lg\left(\frac{d}{d_0}\right) + X_\sigma, \quad X_\sigma \sim \mathcal{N}(0,\sigma) \tag{7.1}$$

$$\theta = \left(2\pi\frac{2d}{\lambda} + \theta_T + \theta_R + \theta_{Tag}\right) \bmod 2\pi. \tag{7.2}$$

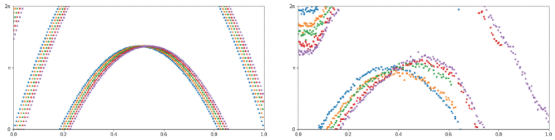
where $PL(d)$ represents path loss at distance d from transmitter to receiver, while $\mathrm{PL}(d_0)$ is

the reference path loss at a close distance established through actual testing. 'n' signifies the path loss factor influenced by the environment. Additionally, X_σ is a normal random variable with a standard deviation of σ. The right side equation represents the distance between the reader antenna and the tag, denoted as d, with a wavelength of λ. A signal's round-trip distance from transmitter to receiver is $2d$. Phase shifts $\theta_T + \theta_R + \theta_{Tag}$ are respectively introduced by the reader's transmitter circuit, receiver circuit, and tag's reflective traits [290]. Similar to RSSI, actual phase values significantly differ from theoretical ones, leading to interleaved tag data and occasional data absence.



(a) (b)

(a) Comparison between predicted and observed RSSI variations



(a) (b)

(b) Comparison between predicted and observed phase variations

Figure 7.3: Illustrating the dynamics of RSSI and Phase in response to Antenna movement

Radar The ultra-wideband radar functions by emitting ultra-short impulses lasting nanoseconds or picoseconds, enabling high pulse repetition frequencies for excellent range resolution. These rapid pulses are transmitted by a UWB impulse radar, propagating via a transmitter [276]. Upon reaching the target, the reflected pulse returns to the receiver, forming an echo composed of varied time delays from human scattering centers (e.g., limbs, torso) and environmental elements (e.g., ground, windows). This leads to diverse scattering coefficients [291]. The Novelda Xethru X4M03 UWB impulse radar module initially used a direct-RF synthesizer to generate Gaussian pulses with an analytic signal shape [275]. These pulses were then broadcasted in an analytic signal form into the operational area from the transmitter side, as stated in Eq 7.3.

$$g(t) = Aexp\,(j2\pi f_o t)\exp\left(\frac{-t}{T_p}\right)^2 x[n] = exp\,(j2\pi f_o n_\tau)\exp\frac{-(n-n_\tau)^2}{{T_p}^2}. \tag{7.3}$$

The RF data can be read directly or processed through on-chip digital down-conversion to obtain the baseband analytic signal. This signal, illustrated in Eq 7.3, consists of in-phase (I) component, $xl[n] = rex[n]$, and quadrature (Q) component, $xQ[n] = imx[n]$. The raw UWB baseband signal represents the data collected during various human activities after collection.

7.3 Detailed Design of RFiDAR System

This section will begin with an overview of the proposed RFiDAR system, followed by a detailed, sequential explanation of its core components.

7.3.1 System Overview

Figure 7.4 illustrates the proposed RFiDAR system comprising five primary building blocks: experimental setup, data collection, pre/post-processing, fusion model, and model deployment and evaluation.

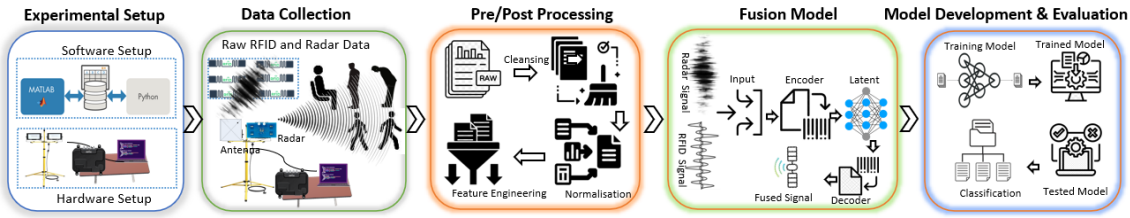


Figure 7.4: Overview of the RFiDAR system.

- **Experimental Setup:** This block outlines the experimental framework for the RFiDAR system. It details the specific hardware and software configurations employed, providing a foundation for the subsequent data collection and analysis processes.
- **Data Collection:** This block is dedicated to establishing the data structures necessary for storing radar and RFID data. It involves setting up tag arrays for RFID data and defining the parameters for radar data collection, ensuring all relevant data is captured effectively.
- **Pre/post-processing:** The focus here is on the steps involved in processing the data, both before and after the main analysis. This process includes the initial handling of raw data, its normalisation, and the preparation for input into the fusion model.
- **Fusion model:** In this part, the intricacies of the LSTM-VAE fusion model are discussed. The discussion revolves around how the model integrates and processes the data from both radar and RFID sources, highlighting its role in the effective fusion of these diverse data streams.
- **Model deployment and evaluation:** This final block discusses the deployment of the system model and a detailed examination of the results obtained. It includes an evaluation of the model's performance, assessing its efficacy in the context of the objectives set forth for the RFiDAR system.

7.3.2 Experimental Setup

The experiments conducted in this study followed ethical approval within a controlled environment, a room measuring $10 \times 10m^2$ in the James Watts South building at the University of Glasgow. To mimic typical room conditions, we built a structure called the *RFiDAR-Wall*, sized at $1.5 \times 1.5m^2$. This setup intentionally included various items, such as metal storage boxes and

tables, to create different signal paths and simulate a strong NLOS environment. The *RFiDAR-Wall* had five columns, each containing three tags, totaling fifteen tags used for our experiments. The hardware for our *RFiDAR-Wall* system (shown in Figure 7.2) comprised an Impinj R700 series reader, an Impinj Times-7 A5010 slim circular polarized UHF antenna, UHF RFID tags, and a Novelda Xethru X4M03 UWB radar. This system cost around £3,500, broken down as follows: Impinj reader (£2000), antenna (£150), tag (£0.03), X4M03 (£350), and laptop server (£1000). We implemented the *RFiDAR-Wall* system using MATLAB and Python frameworks, incorporating software components like Impinj for RFID reader control, X4M03 for HAR, and our algorithms and models. The RFID reader operated between 865MHz and 868MHz and has a transmission power range of 10dBm to 32.5dBm. We set the reader to work at 865.5MHz frequency with a default transmission power of 32.5dBm. In a controlled lab setting, we placed fifteen Impinj tags strategically for our RFiDAR system. The circularly polarized antenna, Novelda Xethru X4M03, was positioned at distances of 2 and 3.0 meters from the center of the "RFiDAR-Wall". Meanwhile, the participants involved in activities stayed 0.5 meters away from the wall. Positioned 0.75 meters above the floor, the antenna captured the activities performed by the participant. Data collection focused solely on these activities and their immediate surroundings. Our setup comprised two main components: the hardware, including the circularly polarised antenna and Novelda Xethru X4M03, and the software for collecting and analysing data.

Hardware Setup

The AAL's *RFiDAR* system integrates the Impinj *R*700 reader, functioning at $865-868$ MHz, and the Novelda Xethru X4M03 UWB radar within the $6-8.5$GHz range. A UHF passive RFID tag array, arranged in a 3×5 grid with 30 cm spacing, supports swift reading of up to 1100 tags per second. The radar operates at 50 frames per second alongside a circularly polarized antenna boasting an 8.0 dBi gain and a 30 dBm RF transmitter. Data collected by the reader seamlessly transmits to a connected laptop, where it interfaces with the radar system for further processing. The Impinj ItemTest Software[1] manages continuous RSSI data capture, while a developed script captures IQ data from back-scattered signals emitted by tags and subjects. This information is relayed via the laptop's RS232 serial port to a dedicated backend module for thorough analysis and processing.

Software Setup

The system's programming, written in Python 3.6, utilizes essential packages including tensorflow-gpu 1.1.0, keras 2.1.2, numpy 1.12.1, pandas 0.23.0, and matplotlib 2.2.2. This programming framework operates efficiently on a robust Dell PC equipped with an Intel® Core $i7-10850H$

[1] https://support.impinj.com

CPU (2.70 GHz, 4 cores), and 16 GB of RAM. These specifications ensure high-performance execution for complex machine-learning algorithms and data analysis tasks.

7.3.3 Data Collection

Following this, we delineate the RFID data collection procedures and radar data collection procedures within the RFiDAR system separately.

RFID Data Collection

The system's RFID reader continuously scans tags within the monitoring area, relaying tag ID, RSSI, phase, and timestamp data. This communication occurs through a narrow RFID channel, allowing numerous tags to coexist but providing limited reading chances as the number of tags increases. As a consequence, the reading rate of a target tag significantly decreases with an increase in tag numbers. The sparsity of tag RSSI data has a notable impact on accurate RSSI unwrapping, crucial for methods such as localization or activity recognition. Improving reading rates for target tags in such scenarios is a non-trivial challenge. For experimental purposes, the study involved five subjects of varying ages, heights, and weights. These subjects were directed to perform sitting, standing, leaning, and walking activities at their natural pace between the antenna and the *RFiDAR-Wall*, depicted in Figure7.2. Each subject completed these activities near the *RFiDAR-Wall* and antenna, resulting in 50 samples for each activity, containing RSSI information. The study's focus was on one subject per activity rather than simultaneous recognition of multiple subjects, resulting in data from a total of 15 tags. Including five subjects aimed to diversify the dataset, producing 1000 valid samples across four scenarios. Within a 3-second interval, each tag was read approximately 30-36 times. Following data collection, raw RSSI data underwent processing using a Python script to extract relevant information necessary for subsequent preprocessing. The goal was to prepare the data for training and testing algorithms effectively.

Radar Data Collection

The HAR data collection employed the Novelda Xethru X4M03 UWB radar operating in the 6-8.5GHz range. It captured various activities (sitting, standing, leaning, walking in two directions) using integrated scripts, recording 50 basebands per second. These basebands depicted reflection amplitude across distances ahead of the radar, with each containing 184 values within the range of [0.4–9.8 m]. The fast-time rate of 23.328 GHz sampled these values, defining each as a range bin and totaling 184 range bins per scan. A hundred radar frames accumulated each second, resulting in a slow-time range from 0.0 to 2.0 s. The data underwent processing through programmed scripts, saved as NumPy arrays based on a guided algorithm described in Table 7.2. Participants completed multiple rounds of activities, producing six distinct output files (one for

each activity), adhering to the data collection protocol. This process ensured a unified dataset, effectively documenting all activities as per the collection policy.

Table 7.2: Dataset summary using *RFiDAR-Wall*: scenarios, subjects, and activities performed.

Activity	**2 meters**		**3 meters**	
	RSSI	**Amplitude**	**RSSI**	**Amplitude**
Empty Room	50	50	50	50
Sitting	50	50	50	50
Standing	50	50	50	50
Walking Forward	50	50	50	50
Walking Backward	50	50	50	50

iii. System Calibration

System calibration serves as an essential step in fusing diverse modality data. An important aspect of this integration is time calibration, especially in multi-sensor data fusion scenarios where inconsistent exist in data collection frequencies across sensors. For instance, the continuous-wave radar records data at a rate of 50 frames per second, while the RFID operates at a speed of 12 samples per second. As the collection sample rate of the radar is much higher than that of the RFID, downsampling the radar data is necessary to match its time stamp with the RFID data collection frequency. This time calibration process is demonstrated in detail in Figure 7.5.

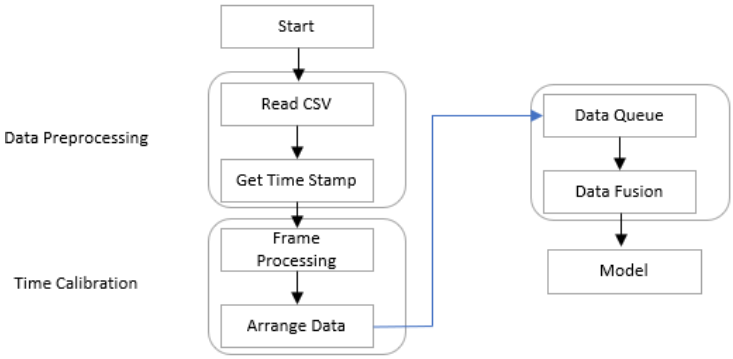


Figure 7.5: Calibrating Time for Data Alignment.

7.3.4 Data Preprocessing

RFID Data Preprocessing The data preprocessing phase is important for analyzing raw RSSI data, involving cleaning, formatting, and transforming it into a structured format suitable for further analysis. Various mathematical and signal processing techniques such as moving average window, bandpass, low-pass, and high-pass filters were applied to focus on specific patterns

within the collected raw data. Initially, the data underwent processing using mathematical expressions defined by Equations 7.4.

$$T_f \frac{y(k) - y(k-1)}{T} + y(k) = x(k), \tag{7.4}$$

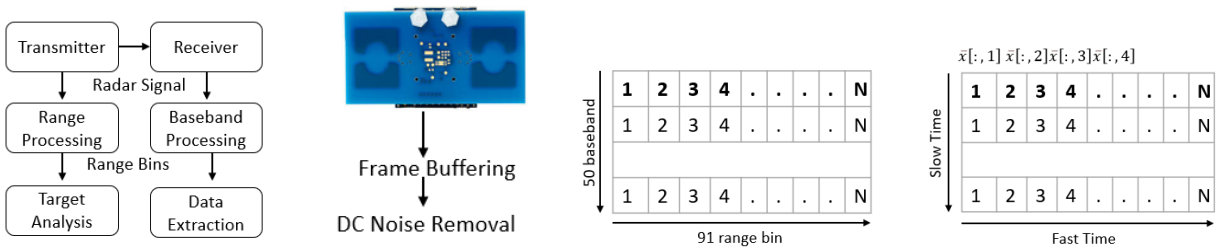
$$y(k) = \frac{T}{T_f + T} x(k) \frac{T_f}{T_f + T} y(k-1) = ax(k) + (1-a)y(k-1). \tag{7.5}$$

The raw data for each activity is organized into a 2D matrix format, where rows represent activity samples and columns represent RSSI data. Standard preprocessing functions and libraries such as (*Scikit* and *Pandas*) were employed to refine and ensure data quality. Each activity's data was formatted into a matrix with 540 columns of RSSI data (15 tags x 36 columns) when the activity occurred contactless.

To maintain uniformity in sample numbers for each activity class and enable robust training of ML/DL models, synthetic data generation using generative adversarial network (GAN) and conditional tabular generative adversarial network (CTGAN) methods is considered [241]. The RSSI data for each activity were stored separately in data files with a *frame size* of 36, representing approximately 36 readings within three seconds. During different activities (e.g., sitting or standing), some tags were partially or fully read. To ensure a consistent 36-time tag reading, any missing tag data was replaced with zeros within the data matrix. Analysis indicated that no tag was read more than 36 times, and missing values (NaN) were imputed with the mean of each row using the built-in *SimpleImputer* function from *SciKit*. Subsequently, the pandas *unique* function was used to segment the timestamp into seconds, enabling the monitoring of tag readings for a precise three-second duration.

***Radar Data Preprocessing*:** The process of Radar data processing involves transforming collected amplitude data into range-time radar frames to depict human activity and position within specific time frames. The method includes frame buffering and clutter removal techniques, as depicted in Figure 7.6b. For signal observation across various activities, baseband data from all range bins was organized per second, utilizing the X4 radar with a range resolution of approximately 5.35 centimeters. Data was recorded at distances of 2.0 and 3.0 meters, forming a 1500 column data matrix. The frame rate was set at 50 frames per second, creating 50 rows for each radar range-time frame, as illustrated in Figure 7.6c.

The xethru X4M03 UWB module captures reflection signals from the target, including receiver noise, atmospheric interference, and surrounding environmental factors within the detection area [292]. These reflections significantly impact HAR's performance. Initially, utilising raw UWB radar data for HAR led to a 19% lower performance compared to reported results, necessitating data processing for enhanced accuracy. Clutter removal, crucial for eliminating non-focusing target signals, involves DC noise removal and background subtraction. DC noise,



(a) Building Blocks (b) Preprocessing (c) Range-time matrix (d) Mean values over time.

Figure 7.6: Visual representation of radar data processing: Blocks, preprocessing, range-time matrix, time-averaged range bins.

a constant signal, is eliminated by computing the mean of each time step in every range bin and subtracting it from the original signal using Eq 7.6,

$$X[:,n] = X[:,n] - \bar{X}[:,n] \tag{7.6}$$

where $X[:,n]$ represents baseband data for all time steps in each range bin, and $\bar{X}[:,n]$ denotes the mean of all time steps in each range bin.

7.4 MultiModality Sensing Fusion System

Multi-sensor data fusion involves combining information from multiple sensors to improve an HAR system's accuracy, reliability, and overall effectiveness. This combination addresses limitations inherent in single-sensor data, including environmental constraints. Our study explored two specific data fusion algorithms: data level fusion and feature level fusion fusion, and conducted experiments to compare their performance.

The LSTM-VAE fusion architecture, as illustrated in Figures 7.7 and 7.8, represents a state-of-the-art approach for the combination of heterogeneous sensor data, specifically from RFID and radar modalities. The system begins with distinct LSTM encoding layers, each fine-tuned to handle the temporal sequences of $4,301$ time steps from XeThru radar and $4,525$ time steps from RFID sensors. In these initial layers, a descending neuron arrangement (100, 80, and 60) precisely encapsulates the high-dimensional time series into a compact set of features. Following the initial transformation, these features are fused into a single 120-dimensional feature vector with a concatenation layer. This vector is then processed through a dense layer, important to the fusion mechanism, which computes the mean and log variance—crucial for the latent distribution that is central to the VAE framework. A sophisticated sampling function then reparameterizes this integrated feature space, resulting in a latent representation secure for decoding.

The model's symmetrical design is reflected in the decoding phase, where another suite of dense layers and LSTM units with 60 neurons each is employed to reconstruct the modality data. The model ensures the reliability of the original data by replicating the dense outputs to the time

series dimensions specific to each modality, followed by an LSTM processing stage that crafts the final output to closely resemble the original data. This complex process is managed by the LSTM-VAE model's forward method, which seamlessly transitions from encoding the input data through reparameterization of the latent space to decoding and reconstructing the input data. The result is a balanced fusion and accurate reconstruction of data from the radar and RFID sensors. By leveraging the sequential data handling capabilities of LSTM and the generative properties of VAE, the model adeptly fuses and restores complex temporal data, demonstrating the robustness of the LSTM-VAE framework in multi-sensor data integration.

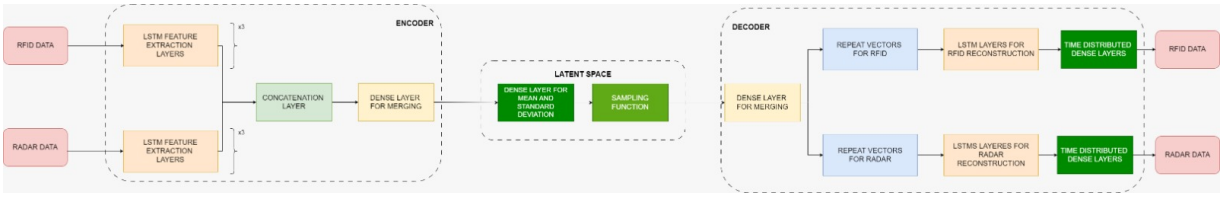


Figure 7.7: High level diagram of LSTM-VAE.

Data Level Fusion:

The data level fusion method manages raw data at the system's foundation, ensuring minimal data loss and optimal reliability. However, its effectiveness greatly depends on the sensor type in use. Inconsistencies in sensor information can complicate the fusion process. By employing timestamp validation and channel stacking, data level fusion integrates varied sensor data. The resulting fused data is then fed into the model for training and classification, as shown in the block diagram (Figure 7.9a) illustrating the data level fusion process.

Feature Level Fusion:

Feature-level fusion does not directly combine original data from diverse sensors. Rather, it entails extracting distinct features from the data processed by individual sensors and combinations of these features for recognition during the fusion process. Our methodology employed two separate LSTM networks: one specialized in extracting radar features, and the other focused on extracting RFID features. These extracted feature maps were then amalgamated using an addition operation. After adjusting the feature map sizes for consistency, they were effectively merged within a VAE model as shown in Figure 7.9b.

7.5 Results and Discussion

This section presents the performance evaluation of radar and RFID data fusion using LSTM-VAE models. We conducted experiments that entailed comparing the Fusion model against single-domain models. The results of our study on the fusion of radar and RFID data utilising

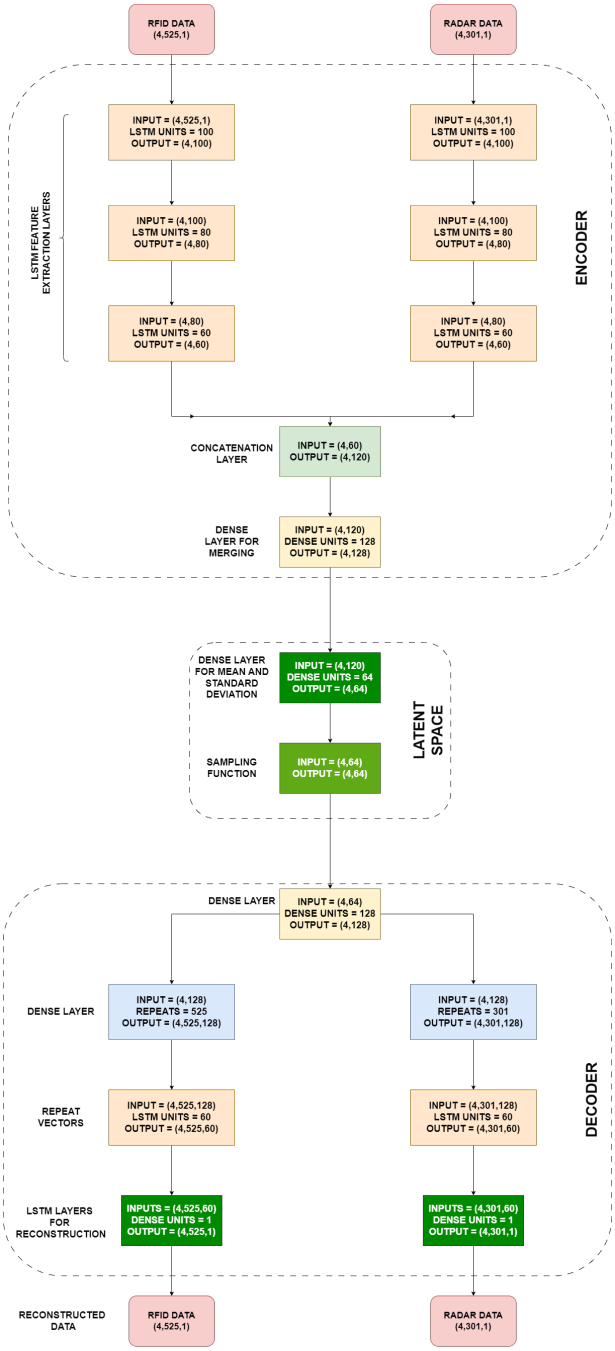
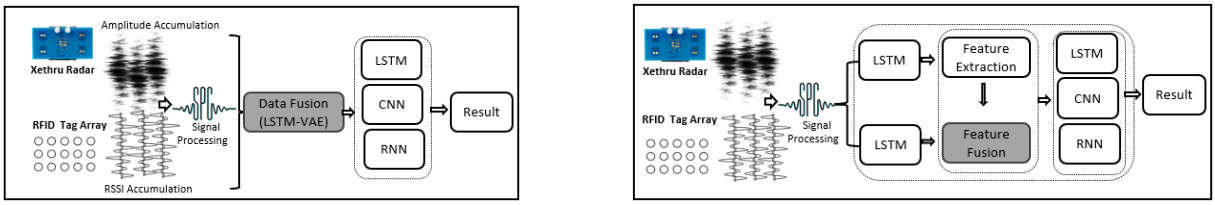


Figure 7.8: Structure of the implemented LSTM-VAE architecture diagram.



(a) Data Level

(b) Feature Level

Figure 7.9: Illustration of data-level and feature-level fusion algorithms.

LSTM-VAE models. Additionally, we aim to demonstrate the efficacy of various algorithms, namely CNN, LSTM, and RNN. Each was applied to the datasets obtained from the radar and RFID sensors to evaluate their individual and combined performance in activity recognition between five distinct activities at a proximity of 2 and 3 meters from the setup.

7.5.1 Assessment Indicators

The assessment of the proposed model's performance in this chapter utilised two metrics: accuracy and confusion matrix. These metrics are applied and performed on the two distinct scenarios. In addition, the assessment indicators of individual activities are shown using the confusion matrix, commonly referred to as the error matrix. Presenting the confusion matrix as a graphical representation offers a more distinct perspective on the model's classification outcomes for each activity. Within the confusion matrix, the column elements signify the actual activity type, while the row elements represent the predicted activity type. Specifically, activities such as no-activity, sitting, standing, leaning, and walking forward and backward correspond to labels.

$$\text{Accuracy } = \frac{TP+TN}{TP+FN+TN+FP}, \tag{7.7}$$

where TP and TN represent the count of accurately and inaccurately predicted samples within positive cases, whereas FP and FN signify the count of accurately and inaccurately predicted samples within negative cases.

7.5.2 Evaluation of Model Performance in Human Activity Recognition

In this study, we investigate the effectiveness of deep learning models in identifying human activities using data from RFID tags and radar under two scenarios: at distances of 2 meters and 3 meters. We systematically compare the accuracies achieved by CNN, RNN, and LSTM models across distances of 2 meters and 3 meters for individual sensing technologies. We also examine the benefits of combining this data through a method known as LSTM-VAE fusion to improve the accuracy of activity recognition.

Individual Sensing Accuracy

Our analysis starts by assessing the accuracy of activity recognition using data from either RFID or radar sensors alone, processed through CNN, RNN, and LSTM models as detailed in Tables 7.3 and 7.4. The LSTM model emerged as the top performer, achieving an impressive average accuracy of 91.5% at 2 meters (91% for RFID and 92% for radar) and 90% at 3 meters (87% for RFID and 93% for radar). This indicates the LSTM model's strong capability of capturing the temporal patterns of human activities.

LSTM-VAE Fusion Model Performance

Our next step is to explore the impact of merging RFID and radar data using the LSTM-VAE fusion approach. This approach led to a significant increase in accuracy as shown in Tables 7.6 and 7.7. At 2 meters, simple data fusion raised the accuracy to 96%, while a more intricate feature-level fusion boosted it to 98%. At a distance of 3 meters, we observed similar improvements, with accuracy reaching 95.8% for basic data fusion and 97.9% for feature-level fusion. These findings underscore the value of integrating data from both sensors to achieve more precise activity recognition.

Moreover, it is important to highlight the key results of the data-level and feature-level fusion strategies:

- Data fusion at 2 meters resulted in a 4.5% increase in accuracy over the baseline, whereas feature-level fusion showed a 6.5% improvement.
- Feature-level fusion notably enhances the accuracy for specific activities, achieving nearly reasonable accuracy scores in some cases.
- At 3 metres, the enhancements are even more pronounced, with feature-level fusion showing a 7.9% improvement over the baseline accuracy.

The comparison reveals that the fusion models, particularly those employing feature-level integration, significantly outperform the single-sensor models. This demonstrates the substantial benefits of synthesising sensor data for improved accuracy in activity recognition.

Overall Performance Comparison

A detailed examination highlights that LSTM models process radar data slightly more effectively than RFID data, particularly in dynamic activities, likely due to radar's superior motion detection capabilities. This advantage is more evident at the 3-meter distance, where radar data processed by LSTM models achieves an accuracy of 93%.

Overall, our results confirm that LSTM models are exceptionally effective in interpreting data for activity recognition, both when using data from individual sensors and when employing a fusion-based approach. By integrating data from RFID and radar sensors, especially through feature-level fusion, we significantly enhance our ability to accurately recognize human activities, showcasing the potential of advanced AI models in pushing the boundaries of technology in this area.

Individual Sensing Accuracy

The individual accuracies for both Radar and RFID, as assessed by CNN, LSTM, and RNN models, are presented in Tables 7.3 and 7.4 respectively. These accuracies are important for understanding the baseline performance of each sensor before fusion.

Table 7.3: Individual sensor accuracy for 2 Meters Scenario

Sensor	Input Data	Model	Overall Accuracy	Empty	Sitting	Standing	Walking Forward	Walking Back
RFID	RSSI	CNN	82%	100%	74%	71%	83%	82%
		LSTM	91%	100%	90%	90%	90%	95%
		RNN	86%	100%	87%	78%	85%	90%
Radar	Amplitude	CNN	76%	100%	60%	60%	75%	65%
		LSTM	92%	100%	90%	90%	98%	92%
		RNN	88%	100%	88%	82%	89%	81%

Table 7.4: Individual sensor accuracy for 3 Meters Scenario

Sensor	Input Data	Model	Overall Accuracy	Empty	Sitting	Standing	Walking Forward	Walking Back
RFID	RSSI	CNN	78%	100%	70%	67%	79%	78%
		LSTM	87%	100%	85%	85%	85%	90%
		RNN	82%	100%	82%	73%	80%	85%
Radar	Amplitude	CNN	77%	100%	62%	62%	77%	67%
		LSTM	93%	100%	91%	91%	97%	93%
		RNN	87%	100%	87%	80%	88%	80%

Fusion Accuracies

Table 7.5 delineates the accuracies obtained through data-level and feature-level fusion techniques implemented using the LSTM-VAE model. These techniques enhance the model's capacity to leverage the strengths of each sensor modality.

Table 7.5: Fusion accuracy's using LSTM-VAE

Fusion Techniques	Fusion Accuracy (2 meters)	Fusion Accuracy (3 meters)
Data-level Fusion	94%	95.1%
Feature-level Fusion	97%	97.8%

Fusion Classification Accuracy

Post-fusion, the classification accuracies for each activity, as discerned by the fusion of radar and RFID data, are summarized in the Tables 7.4 below. The fusion methodologies amplify the robustness of activity recognition.

Table 7.6: Classification accuracies after fusion (2 meters)

Fusion	**Model**	**Overall Accuracy**	**Activities**				
			Empty	**Sitting**	**Standing**	**Walking**	**Walking Back**
Data level	CNN	87%	100%	89%	76%	85%	80%
	LSTM	96%	100%	97%	93%	99%	94%
	RNN	82%	100%	70%	60%	70%	70%
Feature Level	CNN	81%	100%	80%	75%	80%	80%
	LSTM	98%	100%	98%	98%	99%	99%
	RNN	93%	100%	92%	92%	94%	94%

Table 7.7: Classification accuracies after fusion (3 meters)

Fusion	**Model**	**Overall Accuracy**	**Activities**				
			Empty	**Sitting**	**Standing**	**Walking**	**Walking Back**
Data level	CNN	87.2%	100%	89.2%	76.1%	85.4%	80.3%
	LSTM	95.8%	100%	97.2%	93.2%	98.8%	94.1%
	RNN	82.4%	100%	70.5%	60.8%	70.7%	70.3%
Feature Level	CNN	81.3%	100%	80.5%	75.6%	80.2%	80.7%
	LSTM	97.9%	100%	98.1%	97.7%	98.9%	98.6%
	RNN	92.7%	100%	91.8%	92.3%	93.6%	93.9%

7.6 Limitation and Practical Issues

7.6.1 Active Tags vs. Passive Tags

Generally, there are two types of RFID tags: active tags, which incorporate internal batteries, and passive tags, capable of harvesting energy from reader radio waves. Active tags typically offer longer communication ranges but require battery replacement or recharging, posing manpower challenges in large-scale RFID systems. In contrast, passive tags, as thin as paper, are easier to attach to objects than active tags. Additionally, passive tags are generally more cost-effective.

7.6.2 Optimizing RFID and Radar Integration for Large-Scale Monitoring

In large monitoring areas with numerous tags, multiple antennas, multiple readers, and radar, a common issue arises with the shared narrow communication channel among a large number of tags, leading to reduced tag reading rates. This paper proposes a Radar-assisted RFID scheduling mechanism to ensure high reading rates for targeted tags of interest. Moreover, due to the limited communication range of RFID, deploying multiple readers with overlapping coverage becomes necessary. However, adjacent readers may interfere with each other. Some existing works have proposed strategies to optimize reader deployment and mitigate reader collisions [293]. When

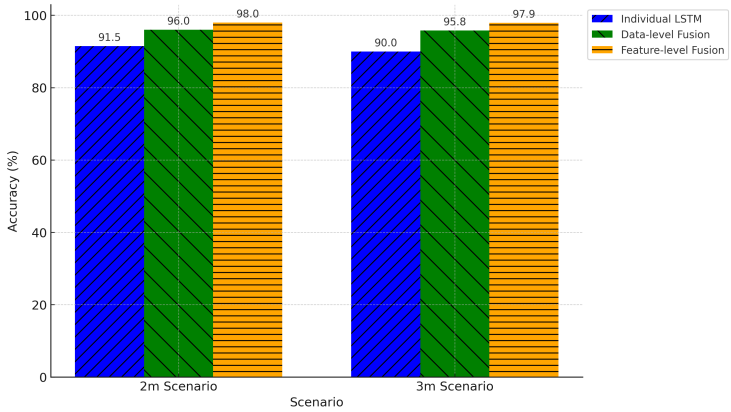


Figure 7.10: Comaprison of LSTM-based accuracy by scenario and fusion technique.

deploying multiple radars, the locations and radiation directions can be predetermined. For subjects within the monitoring region, their relative positions and orientations concerning the radar can be known. Integrating data from multiple radars enables effective subject tracking, possibly consolidating data into a single super radar with adequate monitoring range.

7.6.3 Optimizing RFID and Radar Systems for Efficient Monitoring in Large-Scale Environments

For long-term monitoring applications, energy conservation holds significance for environmental and economic reasons. To reduce energy consumption to some extent, a simple approach is proposed. Initially, when no subject is in the monitoring region, the RFID reader operates in a low-duty cycle mode to save energy. Similarly, the radar can be configured initially in a sleeping status. Upon a subject entering the monitoring region, changes in phase data of certain tags can serve as a signal to awaken the radar, thus further reducing the energy consumption of the RFID-Radar system.

7.7 Summary

This chapter summarises the significant advancement of the RFiDAR system in HAR, smoothly merging RFID and radar technologies to address the challenges inherent in single-sensor solutions. Central to this system is the LSTM-based variational autoencoder (LSTM-VAE) model, adept at extracting and analysing temporal features from both RFID and radar data. This model significantly improves HAR accuracy through data-level and feature-level fusion techniques. RFiDAR's fusion approach not only upholds the high accuracy of individual sensors but also enhances it, especially in complex non-line-of-sight scenarios. Notably, RFID demonstrates superior accuracy over radar in standalone evaluations, a disparity that diminishes with fusion. The effectiveness of the LSTM-VAE model in harmonizing diverse data sets, combined with

RFiDAR's unobtrusive and passive nature, positions it as an invaluable tool for applications such as elderly care in assisted living environments. Overall, RFiDAR emerges as a reliable, precise, and cost-effective system for long-term, autonomous human monitoring, showcasing the transformative impact of integrating multi-sensor data fusion in elevating safety and quality of life.

Chapter 8

Conclusion and Future Work

The Ph.D. thesis titled “TriSense: USRP, RFID, and Radar-Based Hybrid Sensing System for Enhanced Sensing and Monitoring" explores the integration of RFID, Radar, and USRP technologies to refine indoor human activity monitoring systems. This research aims to solve issues like enhancing system accuracy, reliability, signal interference, and non-intrusiveness, thereby marking a significant advancement in the detection and classification of indoor activities.

8.1 Comprehensive Summary

This thesis contributes to the field of human activity monitoring by developing an RF signal-based system that integrates machine learning and deep learning techniques. Initially, it contrasts contact-based methods, like wearables, which are accurate but often impractical due to discomfort and compliance issues, especially in elderly care, with contactless methods for their ubiquity and privacy advantages. Specifically, it explores the use of Wi-Fi’s CSI for detecting activities through signal disruption, advancing RFID from traditional tag-based systems to ‘tag-free’ monitoring through signal variation interpretation, and exploring radar’s potential for movement monitoring with privacy concerns. The thesis ‘TriSense’ then presents a novel integration of RFID, and radar to overcome their drawbacks. This fusion achieves enhanced accuracy in detecting a wide range of activities by leveraging the unique strengths of each technology: RFID’s cost-effectiveness, radar’s broad-area motion detection, and the precision of USRP. Challenges like hardware limitations and signal processing complexities were tackled through innovative solutions. The thesis underscores the need for further machine learning advancements to refine this integrated approach, aiming for broader deployment and greater accuracy in real-time applications.

8.1.1 Advancements in Non-Invasive Human Activity Recognition using USRP and Deep Learning Technologies

Our research has significantly contributed to the advancement of HAR by utilising USRP devices to capture Wi-Fi CSI, introducing a non-intrusive method for accurately monitoring and identifying human activities within indoor environments. This enhancing approach leverages the ubiquitous availability of Wi-Fi, providing a scalable and privacy-conscious alternative to conventional HAR techniques that are often perceived as intrusive. The result of our study highlights the potential of combining AI technologies with RF signal analysis to enhance the accuracy of activity detection, offering a promising solution for the development of intelligent healthcare systems.

The first part of our study successfully leveraged CSI captured through USRP devices, showcasing its efficacy in the precise localisation and identification of various indoor activities, including sitting, standing, and walking. By applying machine learning techniques, specifically a super learner algorithm, we achieved an accuracy rate of 96% in activity localisation, significantly outperforming existing methods. This illustrates the robust capability of AI in enhancing HAR systems, making it a viable solution for smart healthcare applications that require precise monitoring and assistance without compromising patient privacy. Furthermore, our subsequent research explored the application of deep learning models, including LSTM, CNN, and a hybrid LSTM-CNN model, to further enhance the accuracy and reliability of HAR systems. The LSTM model, in particular, proved exceptionally adept at processing the complex patterns present in CSI data, achieving a classification accuracy of 95.3% for a range of indoor activities. This highlights the strength of deep learning techniques in deciphering the subtle nuances of human movement and the flexibility of USRP devices in capturing these dynamics for real-time activity recognition.

Moreover, our research highlights the revolutionary potential of integrating software-defined radio technology with AI in the domain of HAR. The combination of CSI data and the analytical power of machine learning and deep learning algorithms offers a comprehensive solution that markedly improves the accuracy, reliability, and applicability of HAR systems. Moving forward, the exploration of multi-user environments, dynamic settings, and the inclusion of a broader spectrum of activities presents exciting opportunities to further enhance the sophistication and responsiveness of HAR systems. Such advancements hold significant promise for the development of smarter, more responsive healthcare solutions, ultimately contributing to the enhanced welfare and independence of individuals across various environments.

8.1.2 Enhancing Assisted Living with AI-Enhanced RFID Technologies

Our research has made significant contributions to the recent developments in technologies for assisted living, focusing on the integration of AI with RFID technology and transformer-based

models. This study focuses on elderly care with non-intrusive, accurate activity and fall detection systems. By using the AI-powered transparent RFID tag wall (TRT-Wall) and the TFree-FD system employ a passive UHF RFID tag array and deep learning techniques to accurately monitor human activities like sitting, standing, and walking in a contactless manner, achieving an impressive 95.6% accuracy. This is a significant step in enhancing non-intrusive assisted living technologies, promising to improve elderly care by combining simplicity with a non-intrusive approach. In addition, the TFree-FD system presents an innovative approach to tag-free fall detection by integrating a transformer model with RFID tag data, including RSSI and phase information, to achieve high precision in fall detection. With an accuracy exceeding 96%, it establishes a new standard in the field, surpassing traditional models through effective analysis of signal variations caused by human movements. This advancement is particularly significant for elderly care, offering swift and accurate fall detection. The system's innovative fusion strategy, combining both early and late fusion techniques, supports its reliability and versatility across different indoor environments, enabling it to accurately identify a broad spectrum of activities and fall incidents.

These developments signify a transformative approach to activity monitoring and fall detection within assisted living environments. By harnessing the power of AI and RFID technology, these systems not only promise to reduce the healthcare burden associated with elderly falls but also to improve the quality of life for the elderly, allowing them to live independently and safely in their homes. The integration of these advanced technologies into everyday care practices is secure to redefine the standards of elderly care, emphasizing the importance of innovation, accuracy, and non-intrusiveness in assisted living solutions.

8.1.3 Contactless Human Activity Recognition Through RFID and Radar Fusion

The fusion of RFID and radar (RFiDAR) system represents a significant advancement in indoor human activity recognition to overcome traditional monitoring systems' privacy and efficiency issues. This method enhances privacy by avoiding intrusive devices and reduces monitoring costs. RFiDAR excels in detecting diverse activities accurately, even in challenging conditions where conventional methods struggle due to long-range and non-line-of-sight scenarios. The RFiDAR system's effectiveness is driven by a sophisticated LSTM-VAE model that extracts temporal features from different modality data to improve activity recognition. This fusion model employes deep learning to improve recognition accuracy through sophisticated data fusion algorithms, including both data-level and feature-level fusion. Moreover, the integration of RFID and radar data through RFiDAR effectively addresses the individual limitations of these technologies, such as difficulties in non-line-of-sight conditions and signal degradation due to environmental factors. This fusion facilitates the detection of complex human activities with

high precision, paving the way for its application across various domains. It holds particular promise for enhancing the quality of life for elderly individuals in assisted living situations, offering a new paradigm in indoor monitoring technology.

8.2 Limitation and Future Direction

This thesis presents an RF signal-based system for detecting, recognising, identifying, and localising human activities, highlighting the potential for further enhancement. It meticulously documents experiments using CSI from WiFi, RSSI from RFID, and radar amplitude across predetermined, controlled environments. To elevate the system's efficacy, future endeavors will focus on diversifying the dataset with a broader array of positions and orientations. Enhancing signal processing techniques will be essential, aiming to diminish noise within CSI/RSSI readings and achieve precise activity classification. Future initiatives will explore deployment in dynamic environments like shopping malls and integrate technologies to harness each modality's strengths. While acknowledging the high costs and complexity of SDRs and the limitations posed by RFID's line-of-sight requirements, it proposes employing more cost-effective solutions, such as Raspberry Pis, for RF signal capture and cloud server integration for enhanced RSSI data management.

Our research analysis explores the fine roles of CSI and RSSI in indoor monitoring, specifically for detecting occupancy, recognising activities, and localisation. While RSSI is widely supported across a diverse range of devices, from IoT to WiFi, offering a straightforward approach to sensing, CSI processing stands out in WiFi-enabled devices equipped for multi-antenna operations. This investigation aims to delineate the comparative strengths and limitations inherent in both methods.

CSI Advantages

1. **Precision:** CSI enables precise detection of people's presence and activities by providing detailed information about signal propagation characteristics, including phase shifts and attenuation.

2. **Non-Intrusive Monitoring:** CSI processing is non-intrusive and can be implemented using readily available WiFi infrastructure, which is widespread in indoor environments.

3. **Environmental Resilience:** CSI relies on the signal's physical properties, making it resilient to environmental variations like temperature and humidity.

CSI Limitations

1. **Coverage Constraints:** The limited coverage of WiFi signals within indoor environments restricts CSI processing.

2. **Interference Sensitivity:** The accuracy of CSI may be compromised by electronic disturbances such as microwaves and Bluetooth devices, which can degrade the accuracy of occupancy detection.

3. **Scalability Challenges:** The extensive monitoring setups may face limitations in applying CSI due to its heavy data processing demands.

RSSI Advantages

1. **Implementation Ease:** RSSI processing is a straightforward and readily implementable method for occupancy detection, as it solely entails measuring the total strength of the received signal. In contrast, CSI processing, even when phase information is disregarded (as is typical), necessitates consideration of the signal strength across all sub-carriers.

2. **Energy Efficiency:** Processing RSSI is a low-power technique, as it demands minimal processing power and can be executed using inexpensive hardware.

3. **Extensive Reach:** RSSI processing enables the detection of signals from multiple access points, allowing a single device to cover a large area.

RSSI Limitations

1. **Low Accuracy:** RSSI processing is less precise compared to CSI processing as it solely indicates signal strength while disregarding other signal attributes.

2. **Resolution Limitations:** RSSI processing is limited in resolution as it merely offers a coarse estimation of the signal source's location.

3. **Density Dependence:** To attain precise outcomes, RSSI processing might require a high node density, potentially escalating the system's expenses and intricacy.

The choice between CSI and RSSI depends on balancing accuracy, system complexity, and deployment costs. However, other factors such as network topology and the potential to integrate signal amplitude with phase information also play a role. This comparison highlights the need for a strategic approach to integrating these technologies into indoor monitoring systems, aiming for a balanced solution that meets project requirements and constraints effectively.

8.2.1 Future Directions in USRP-Based Human Activity Detection

1. **Advanced Signal Processing for Enhanced Differentiation:** Improve signal processing algorithms to accurately distinguish between multiple subjects using deep learning and pattern recognition techniques, particularly in crowded environments.

2. **Improved Penetration and Range with Frequency and Antenna Innovations:** Explore a wider range of frequencies and innovative antenna designs to enhance the USRP's ability to detect activities over long ranges and through obstacles. Investigate sub-terahertz frequencies and metamaterials for antennas to address these challenges.

3. **Real-Time Processing Enhancements:** Focus on hardware and software optimizations to achieve near real-time data processing. This involves utilizing more powerful computing resources, efficient data handling algorithms, and implementing edge computing strategies to minimize latency.

8.2.2 Future Directions in RFID-Based Monitoring

1. **Enhanced Localisation Precision:** Improve RFID-based localisation and tracking accuracy through the integration of machine learning models capable of interpreting fluctuating RSSI values more accurately. This may entail developing algorithms that adapt to environmental changes and distinguish between human-induced signal variations and those caused by non-human factors.

2. **Integrated Multi-Sensor Monitoring:** Augment RFID systems by integrating data from additional sensing modalities (e.g., WiFi-RSSI or radar amplitude) of subjects. This multimodal approach offers richer contextual information, enhancing the system's ability to monitor complex activities and interactions.

3. **Energy-Efficient Sustainability:** Innovate in RFID tag and reader design to minimise energy consumption and prolong operational lifespan. Exploring energy harvesting technologies capable of powering RFID systems from ambient sources (e.g., solar, RF energy) could revolutionize long-term monitoring applications.

8.2.3 Future Directions in Radar-Based Monitoring

1. **Enhanced Radar Resolution:** Invest in research to improve the resolution of radar-based monitoring systems, enabling better detection of fine-grained human activities and subtle movements in complex scenarios.

2. **Refined Through-Wall Imaging:** Develop and enhance through-wall radar imaging techniques, possibly utilising UWB radar systems for improved penetration and resolution in

detecting activities obstructed by barriers.

3. **AI-Enabled Activity Recognition:** Integrate artificial intelligence and machine learning algorithms with radar data to enhance the classification and recognition of human activities, leveraging diverse datasets to improve real-time activity differentiation.

8.2.4 Future Directions in Fusion of USRP, RFID, and Radar Technologies

1. **Unified Sensing Framework Development:** Design a single framework to integrate data from USRP, RFID, and radar sensors. This framework should synchronise data collection, process inputs from different modalities, and extract meaningful insights using fusion algorithms.

2. **Advanced Cross-Modal Fusion Algorithms:** Develop sophisticated algorithms to fuse data from USRP, RFID, and radar sensors effectively. These algorithms should leverage the strengths of each modality, enhancing accuracy and reliability in human activity detection and monitoring.

3. **Privacy-Preserving Data Analysis Implementation:** Integrate privacy-preserving mechanisms into the sensing framework. This is important for maintaining individuals' privacy and trust, particularly in sensitive environments. Techniques like anonymization, encryption, and secure multi-party computation can safeguard individual privacy while enabling comprehensive activity monitoring.

Advancing these future directions necessitates a multidisciplinary strategy that integrates expertise in signal processing, hardware design, artificial intelligence, and privacy technologies. By concentrating on these domains, substantial progress can be made in enhancing the state-of-the-art in human activity detection and monitoring using USRP, RFID, radar, and their integration, thus surpassing existing constraints and unlocking novel opportunities for applications in healthcare, security, and smart environments.

Bibliography

[1] Paola Patricia Ariza-Colpas et al. “human activity recognition data analysis: History, evolutions, and new trends”. In: *Sensors* 22.9 (2022), p. 3401.

[2] Ehab El-Adawi et al. “Wireless body area sensor networks based human activity recognition using deep learning”. In: *Scientific Reports* 14.1 (2024), p. 2702.

[3] Cuong Pham et al. “SensCapsNet: deep neural network for non-obtrusive sensing based human activity recognition”. In: *IEEE Access* 8 (2020), pp. 86934–86946.

[4] Ran Zhu et al. “Efficient human activity recognition solving the confusing activities via deep ensemble learning”. In: *Ieee Access* 7 (2019), pp. 75490–75499.

[5] Cho Nilar Phyo, Thi Thi Zin, and Pyke Tin. “Deep learning for recognizing human activities using motions of skeletal joints”. In: *IEEE Transactions on Consumer Electronics* 65.2 (2019), pp. 243–252.

[6] Yegang Du, Yuto Lim, and Yasuo Tan. “A novel human activity recognition and prediction in smart home based on interaction”. In: *Sensors* 19.20 (2019), p. 4474.

[7] Samundra Deep and Xi Zheng. “Leveraging CNN and transfer learning for vision-based human activity recognition”. In: *2019 29th International Telecommunication Networks and Applications Conference (ITNAC)*. IEEE. 2019, pp. 1–4.

[8] Djamila Romaissa Beddiar et al. “Vision-based human activity recognition: a survey”. In: *Multimedia Tools and Applications* 79.41-42 (2020), pp. 30509–30555.

[9] Ahmed Abobakr, Mohammed Hossny, and Saeid Nahavandi. “A skeleton-free fall detection system from depth images using random decision forest”. In: *IEEE Systems Journal* 12.3 (2017), pp. 2994–3005.

[10] Kaixuan Chen et al. “Deep learning for sensor-based human activity recognition: Overview, challenges, and opportunities”. In: *ACM Computing Surveys (CSUR)* 54.4 (2021), pp. 1–40.

[11] Binjal Suthar and Bijal Gadhia. “Human activity recognition using deep learning: a survey”. In: *Data Science and Intelligent Applications: Proceedings of ICDSIA 2020*. Springer. 2021, pp. 217–223.

[12] Md Golam Morshed et al. “Human Action Recognition: A Taxonomy-Based Survey, Updates, and Opportunities”. In: *Sensors* 23.4 (2023), p. 2182.

[13] Daniel Halperin et al. “Tool release: Gathering 802.11 n traces with channel state information”. In: *ACM SIGCOMM computer communication review* 41.1 (2011), pp. 53–53.

[14] Yaxiong Xie, Zhenjiang Li, and Mo Li. “Precise power delay profiling with commodity WiFi”. In: *Proceedings of the 21st Annual international conference on Mobile Computing and Networking*. 2015, pp. 53–64.

[15] Fadel Adib et al. “3D tracking via body radio reflections”. In: *11th USENIX Symposium on Networked Systems Design and Implementation (NSDI 14)*. 2014, pp. 317–329.

[16] Qifan Pu et al. “Whole-home gesture recognition using wireless signals”. In: *Proceedings of the 19th annual international conference on Mobile computing & networking*. 2013, pp. 27–38.

[17] Stephan Sigg et al. “Leveraging RF-channel fluctuation for activity recognition: Active and passive systems, continuous and RSSI-based signal features”. In: *Proceedings of International Conference on Advances in Mobile Computing & Multimedia*. 2013, pp. 43–52.

[18] Stephan Sigg et al. “RF-sensing of activities from non-cooperative subjects in device-free recognition systems using ambient and local signals”. In: *IEEE Transactions on Mobile Computing* 13.4 (2013), pp. 907–920.

[19] Joey Wilson and Neal Patwari. “Radio tomographic imaging with wireless networks”. In: *IEEE Transactions on Mobile Computing* 9.5 (2010), pp. 621–632.

[20] Joey Wilson and Neal Patwari. “See-through walls: Motion tracking using variance-based radio tomography networks”. In: *IEEE Transactions on Mobile Computing* 10.5 (2010), pp. 612–621.

[21] Yan Wang et al. “E-eyes: device-free location-oriented activity identification using fine-grained wifi signatures”. In: *Proceedings of the 20th annual international conference on Mobile computing and networking*. 2014, pp. 617–628.

[22] Wei Wang et al. “Understanding and modeling of wifi signal based human activity recognition”. In: *Proceedings of the 21st annual international conference on mobile computing and networking*. 2015, pp. 65–76.

[23] Wei Wang et al. “Device-free human activity recognition using commercial WiFi devices”. In: *IEEE Journal on Selected Areas in Communications* 35.5 (2017), pp. 1118–1131.

[24] Yuxi Wang, Kaishun Wu, and Lionel M Ni. “Wifall: Device-free fall detection by wireless networks”. In: *IEEE Transactions on Mobile Computing* 16.2 (2016), pp. 581–594.

[25] Daqing Zhang et al. “Anti-fall: A non-intrusive and real-time fall detector leveraging CSI from commodity WiFi devices”. In: *Inclusive Smart Cities and e-Health: 13th International Conference on Smart Homes and Health Telematics, ICOST 2015, Geneva, Switzerland, June 10-12, 2015, Proceedings 13*. Springer. 2015, pp. 181–193.

[26] Hao Wang et al. “RT-Fall: A real-time and contactless fall detection system with commodity WiFi devices”. In: *IEEE Transactions on Mobile Computing* 16.2 (2016), pp. 511–526.

[27] Sameera Palipana et al. “FallDeFi: Ubiquitous fall detection using commodity Wi-Fi devices”. In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1.4 (2018), pp. 1–25.

[28] Kevin Chetty, Graeme E Smith, and Karl Woodbridge. “Through-the-wall sensing of personnel using passive bistatic wifi radar at standoff distances”. In: *IEEE Transactions on Geoscience and Remote Sensing* 50.4 (2011), pp. 1218–1226.

[29] Fadel Adib and Dina Katabi. “See through walls with WiFi!” In: *Proceedings of the ACM SIGCOMM 2013 conference on SIGCOMM*. 2013, pp. 75–86.

[30] Paramvir Bahl and Venkata N Padmanabhan. “RADAR: An in-building RF-based user location and tracking system”. In: *Proceedings IEEE INFOCOM 2000. Conference on computer communications. Nineteenth annual joint conference of the IEEE computer and communications societies (Cat. No. 00CH37064)*. Vol. 2. Ieee. 2000, pp. 775–784.

[31] İsmail Güvenc. “Enhancements to RSS based indoor tracking systems using Kalman filters”. PhD thesis. University of New Mexico, 2003.

[32] Anindya S Paul and Eric A Wan. “RSSI-based indoor localization and tracking using sigma-point Kalman smoothers”. In: *IEEE Journal of selected topics in signal processing* 3.5 (2009), pp. 860–873.

[33] Krishna Chintalapudi, Anand Padmanabha Iyer, and Venkata N Padmanabhan. “Indoor localization without the pain”. In: *Proceedings of the sixteenth annual international conference on Mobile computing and networking*. 2010, pp. 173–184.

[34] Jie Wang et al. “Robust device-free wireless localization based on differential RSS measurements”. In: *IEEE transactions on industrial electronics* 60.12 (2012), pp. 5943–5952.

[35] Ugur Bekcibasi and Mahmut Tenruh. “Increasing RSSI localization accuracy with distance reference anchor in wireless sensor networks”. In: *Acta Polytechnica Hungarica* 11.8 (2014), pp. 103–120.

[36] Yaqin Xie et al. “An improved K-nearest-neighbor indoor localization method based on spearman distance”. In: *IEEE signal processing letters* 23.3 (2016), pp. 351–355.

[37] Paolo Barsocchi et al. “A novel approach to indoor RSSI localization by automatic calibration of the wireless propagation model”. In: *VTC Spring 2009-IEEE 69th Vehicular Technology Conference*. IEEE. 2009, pp. 1–5.

[38] Maurizio Bocca et al. “Multiple target tracking with RF sensor networks”. In: *IEEE Transactions on Mobile Computing* 13.8 (2013), pp. 1787–1800.

[39] Kaishun Wu et al. “FILA: Fine-grained indoor localization”. In: *2012 Proceedings IEEE INFOCOM*. IEEE. 2012, pp. 2210–2218.

[40] Kaishun Wu et al. “CSI-based indoor localization”. In: *IEEE Transactions on Parallel and Distributed Systems* 24.7 (2012), pp. 1300–1309.

[41] Souvik Sen et al. “You are facing the Mona Lisa: Spot localization using PHY layer information”. In: *Proceedings of the 10th international conference on Mobile systems, applications, and services*. 2012, pp. 183–196.

[42] Souvik Sen et al. “Avoiding multipath to revive inbuilding WiFi localization”. In: *Proceeding of the 11th annual international conference on Mobile systems, applications, and services*. 2013, pp. 249–262.

[43] Alex T Mariakakis et al. “Sail: Single access point-based indoor localization”. In: *Proceedings of the 12th annual international conference on Mobile systems, applications, and services*. 2014, pp. 315–328.

[44] Xuyu Wang et al. “CSI-based fingerprinting for indoor localization: A deep learning approach”. In: *IEEE transactions on vehicular technology* 66.1 (2016), pp. 763–776.

[45] Chenren Xu et al. “SCPL: Indoor device-free multi-subject counting and localization using radio signal strength”. In: *Proceedings of the 12th international conference on Information Processing in Sensor Networks*. 2013, pp. 79–90.

[46] Scott Y Seidel and Theodore S Rappaport. “914 MHz path loss prediction models for indoor wireless communications in multifloored buildings”. In: *IEEE transactions on Antennas and Propagation* 40.2 (1992), pp. 207–217.

[47] Dian Zhang et al. “Rass: A real-time, accurate, and scalable system for tracking transceiver-free objects”. In: *IEEE Transactions on Parallel and Distributed Systems* 24.5 (2012), pp. 996–1008.

[48] Ossi Kaltiokallio et al. “Non-invasive respiration rate monitoring using a single COTS TX-RX pair”. In: *IPSN-14 proceedings of the 13th international symposium on information processing in sensor networks*. IEEE. 2014, pp. 59–69.

[49] Ambili Thottam Parameswaran, Mohammad Iftekhar Husain, Shambhu Upadhyaya, et al. “Is rssi a reliable parameter in sensor localization algorithms: An experimental study”. In: *Field failure data analysis workshop (F2DA09)*. Vol. 5. IEEE Niagara Falls, NY, USA. 2009.

[50] Fadel Adib et al. “Capturing the human figure through a wall”. In: *ACM Transactions on Graphics (TOG)* 34.6 (2015), pp. 1–13.

[51] Chen-Yu Hsu et al. “Extracting gait velocity and stride length from surrounding radio signals”. In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 2017, pp. 2116–2126.

[52] Fadel Adib et al. “Smart homes that monitor breathing and heart rate”. In: *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. 2015, pp. 837–846.

[53] Chen-Yu Hsu et al. “Zero-effort in-home sleep and insomnia monitoring using radio signals”. In: *Proceedings of the ACM on Interactive, mobile, wearable and ubiquitous technologies* 1.3 (2017), pp. 1–18.

[54] Mingmin Zhao, Fadel Adib, and Dina Katabi. “Emotion recognition using wireless signals”. In: *Proceedings of the 22nd annual international conference on mobile computing and networking*. 2016, pp. 95–108.

[55] Jihoon Hong and Tomoaki Ohtsuki. “Signal eigenvector-based device-free passive localization using array sensor”. In: *IEEE Transactions on Vehicular Technology* 64.4 (2015), pp. 1354–1363.

[56] Jie Xiong and Kyle Jamieson. “{ArrayTrack}: A {Fine-Grained} indoor location system”. In: *10th USENIX Symposium on Networked Systems Design and Implementation (NSDI 13)*. 2013, pp. 71–84.

[57] Santosh Nannuru et al. “Radio-frequency tomography for passive indoor multitarget tracking”. In: *IEEE Transactions on Mobile Computing* 12.12 (2012), pp. 2322–2333.

[58] Jin Zhang et al. “Wifi-id: Human identification using wifi signal”. In: *2016 International Conference on Distributed Computing in Sensor Systems (DCOSS)*. IEEE. 2016, pp. 75–82.

[59] Wei Wang, Alex X Liu, and Muhammad Shahzad. “Gait recognition using wifi signals”. In: *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 2016, pp. 363–373.

[60] Feng Hong et al. “WFID: Passive device-free human identification using WiFi signal”. In: *Proceedings of the 13th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services*. 2016, pp. 47–56.

[61] Cong Shi et al. “Smart user authentication through actuation of daily activities leveraging WiFi-enabled IoT”. In: *Proceedings of the 18th ACM international symposium on mobile ad hoc networking and computing*. 2017, pp. 1–10.

[62] Yunze Zeng, Parth H Pathak, and Prasant Mohapatra. "WiWho: WiFi-based person identification in smart spaces". In: *2016 15th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*. IEEE. 2016, pp. 1–12.

[63] Yoshihisa Okamoto and Tomoaki Ohtsuki. "Human activity classification and localization algorithm based on temporal-spatial virtual array". In: *2013 IEEE International Conference on Communications (ICC)*. IEEE. 2013, pp. 1512–1516.

[64] Yugo Agata, Tomoaki Ohtsuki, and Kentaroh Toyoda. "Doppler analysis based fall detection using array antenna". In: *2018 IEEE International Conference on Communications (ICC)*. IEEE. 2018, pp. 1–6.

[65] Saad Iqbal et al. "Indoor motion classification using passive RF sensing incorporating deep learning". In: *2018 IEEE 87th Vehicular Technology Conference (VTC Spring)*. IEEE. 2018, pp. 1–5.

[66] Md Nafiul Alam Nipu et al. "Human identification using wifi signal". In: *2018 Joint 7th International Conference on Informatics, Electronics & Vision (ICIEV) and 2018 2nd International Conference on Imaging, Vision & Pattern Recognition (icIVPR)*. IEEE. 2018, pp. 300–304.

[67] Mohammed Al-Khafajiy et al. "Intelligent control and security of fog resources in healthcare systems via a cognitive fog model". In: *ACM Transactions on Internet Technology (TOIT)* 21.3 (2021), pp. 1–23.

[68] Guanhua Wang et al. "We can hear you with Wi-Fi!" In: *IEEE Transactions on Mobile Computing* 15.11 (2016), pp. 2907–2920.

[69] Tao Zhang et al. "WiGrus: A WiFi-based gesture recognition system using software-defined radio". In: *IEEE Access* 7 (2019), pp. 131102–131113.

[70] Syed Mohsin Bokhari et al. "DGRU based human activity recognition using channel state information". In: *Measurement* 167 (2021), p. 108245.

[71] William Taylor et al. "An intelligent non-invasive real-time human activity recognition system for next-generation healthcare". In: *Sensors* 20.9 (2020), p. 2653.

[72] Fadel Adib, Zachary Kabelac, and Dina Katabi. "Multi-person localization via {RF} body reflections". In: *12th {USENIX} Symposium on Networked Systems Design and Implementation ({NSDI} 15)*. 2015, pp. 279–292.

[73] Shuyu Shi et al. "Accurate location tracking from CSI-based passive device-free probabilistic fingerprinting". In: *IEEE Transactions on Vehicular Technology* 67.6 (2018), pp. 5217–5230.

[74] Minh Tu Hoang et al. "Recurrent neural networks for accurate RSSI indoor localization". In: *IEEE Internet of Things Journal* 6.6 (2019), pp. 10639–10651.

[75] Yuxiao Hou, Yanwen Wang, and Yuanqing Zheng. "TagBreathe: Monitor breathing with commodity RFID systems". In: *2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS)*. IEEE. 2017, pp. 404–413.

[76] Zhanjun Hao et al. "Wi-SL: contactless fine-grained gesture recognition uses channel state information". In: *Sensors* 20.14 (2020), p. 4025.

[77] Anjie Zhu et al. "Indoor localization for passive moving objects based on a redundant SIMO radar sensor". In: *IEEE Journal on Emerging and Selected Topics in Circuits and Systems* 8.2 (2018), pp. 271–279.

[78] Sheng Ye et al. "Lsi-rec: A link state indicator based gesture recognition scheme in a rfid system". In: *2014 9th IEEE Conference on Industrial Electronics and Applications*. IEEE. 2014, pp. 406–411.

[79] Raúl Parada et al. "Smart surface: RFID-based gesture recognition using k-means algorithm". In: *2016 12th International Conference on Intelligent Environments (IE)*. IEEE. 2016, pp. 111–118.

[80] Han Ding et al. "Rfipad: Enabling cost-efficient and device-free in-air handwriting using passive tags". In: *2017 IEEE 37Th international conference on distributed computing systems (ICDCS)*. IEEE. 2017, pp. 447–457.

[81] Yongpan Zou et al. "Grfid: A device-free rfid-based gesture recognition system". In: *IEEE Transactions on Mobile Computing* 16.2 (2016), pp. 381–393.

[82] Lina Yao et al. "Rf-care: Device-free posture recognition for elderly people using a passive rfid tag array". In: (2015).

[83] Lina Yao et al. "Compressive representation for device-free activity recognition with passive RFID signal strength". In: *IEEE Transactions on Mobile Computing* 17.2 (2017), pp. 293–306.

[84] Liyao Li et al. "R&P: an low-cost device-free activity recognition for E-health". In: *IEEE Access* 6 (2017), pp. 81–90.

[85] Jinsong Han et al. "Cbid: A customer behavior identification system using passive tags". In: *IEEE/ACM Transactions on Networking* 24.5 (2015), pp. 2885–2898.

[86] Zimu Zhou et al. "Design and implementation of an RFID-based customer shopping behavior mining system". In: *IEEE/ACM transactions on networking* 25.4 (2017), pp. 2405–2418.

[87] Chris Todd and Dawn Skelton. *What are the main risk factors for falls amongst older people and what are the most effective interventions to prevent these falls?* World Health Organization. Regional Office for Europe, 2004.

[88] Jian He, Chen Hu, and Yang Li. "An autonomous fall detection and alerting system based on mobile and ubiquitous computing". In: *2013 IEEE 10th International Conference on Ubiquitous Intelligence and Computing and 2013 IEEE 10th International Conference on Autonomic and Trusted Computing*. IEEE. 2013, pp. 539–543.

[89] Mary E Tinetti, Wen-Liang Liu, and Elizabeth B Claus. "Predictors and prognosis of inability to get up after falls among elderly persons". In: *Jama* 269.1 (1993), pp. 65–70.

[90] AK Bourke et al. "Evaluation of waist-mounted tri-axial accelerometer based fall-detection algorithms during scripted and continuous unscripted activities". In: *Journal of biomechanics* 43.15 (2010), pp. 3051–3057.

[91] Mustafa Şahin Turan and Billur Barshan. "Classification of fall directions via wearable motion sensors". In: *Digital Signal Processing* 125 (2022), p. 103129.

[92] Bor-Shing Lin et al. "Fall detection system with artificial intelligence-based edge computing". In: *IEEE Access* 10 (2022), pp. 4328–4339.

[93] Dariusz Mrozek, Anna Koczur, and Bożena Małysiak-Mrozek. "Fall detection in older adults with mobile IoT devices and machine learning in the cloud and on the edge". In: *Information Sciences* 537 (2020), pp. 132–147.

[94] Vinod Kumar, Neelendra Badal, and Rajesh Mishra. "Elderly fall due to drowsiness: Detection and prevention using machine learning and IoT". In: *Modern Physics Letters B* 35.07 (2021), p. 2150120.

[95] Evelien E Geertsema et al. "Automated remote fall detection using impact features from video and audio". In: *Journal of biomechanics* 88 (2019), pp. 25–32.

[96] Moeness G Amin et al. "Radar signal processing for elderly fall detection: The future for in-home monitoring". In: *IEEE Signal Processing Magazine* 33.2 (2016), pp. 71–80.

[97] Nathaniel Faulkner et al. "CapLoc: Capacitive sensing floor for device-free localization and fall detection". In: *IEEE Access* 8 (2020), pp. 187353–187364.

[98] Anurag De et al. "Fall detection approach based on combined two-channel body activity classification for innovative indoor environment". In: *Journal of Ambient Intelligence and Humanized Computing* (2022), pp. 1–12.

[99] Guodong Feng et al. "Floor pressure imaging for fall detection with fiber-optic sensors". In: *IEEE Pervasive Computing* 15.2 (2016), pp. 40–47.

[100] Thinh M Le and Rui Pan. "Accelerometer-based sensor network for fall detection". In: *2009 IEEE Biomedical Circuits and Systems Conference*. IEEE. 2009, pp. 265–268.

[101] Marcela Vallejo, Claudia V Isaza, and Jose D Lopez. "Artificial neural networks as an alternative to traditional fall detection methods". In: *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE. 2013, pp. 1648–1651.

[102] Daniela Micucci et al. "Falls as anomalies? An experimental evaluation using smartphone accelerometer data". In: *Journal of Ambient Intelligence and Humanized Computing* 8 (2017), pp. 87–99.

[103] Hristijan Gjoreski et al. "RAReFall—Real-time activity recognition and fall detection system". In: *2014 IEEE International Conference on Pervasive Computing and Communication Workshops (PERCOM WORKSHOPS)*. IEEE. 2014, pp. 145–147.

[104] Changhong Wang et al. "Low-power fall detector using triaxial accelerometry and barometric pressure sensing". In: *IEEE Transactions on Industrial Informatics* 12.6 (2016), pp. 2302–2311.

[105] Yin Xiaoling et al. "Human motion state recognition based on smart phone built-in sensors [J]". In: *Journal of Communications* 40.03 (2019), pp. 157–169.

[106] Caroline Rougier et al. "Robust video surveillance for fall detection based on human shape deformation". In: *IEEE Transactions on circuits and systems for video Technology* 21.5 (2011), pp. 611–622.

[107] Behzad Mirmahboub et al. "Automatic monocular system for human fall detection based on variations in silhouette area". In: *IEEE transactions on biomedical engineering* 60.2 (2012), pp. 427–436.

[108] Yixiao Yun and Irene Yu-Hua Gu. "Human fall detection in videos by fusing statistical features of shape and motion dynamics on riemannian manifolds". In: *Neurocomputing* 207 (2016), pp. 726–734.

[109] Chadia Khraief, Faouzi Benzarti, and Hamid Amiri. "Elderly fall detection based on multi-stream deep convolutional networks". In: *Multimedia Tools and Applications* 79 (2020), pp. 19537–19560.

[110] Hadir Abdo, Khaled M Amin, and Ahmad M Hamad. "Fall detection based on RetinaNet and MobileNet convolutional neural networks". In: *2020 15th International Conference on Computer Engineering and Systems (ICCES)*. IEEE. 2020, pp. 1–7.

[111] Ju Wang et al. "Are RFID sensing systems ready for the real world?" In: *Proceedings of the 17th Annual International Conference on Mobile Systems, Applications, and Services*. 2019, pp. 366–377.

[112] Han He et al. "Clothing-integrated passive RFID strain sensor platform for body movement-based controlling". In: *2019 IEEE International Conference on RFID Technology and Applications (RFID-TA)*. IEEE. 2019, pp. 236–239.

[113] Yanwen Wang and Yuanqing Zheng. "TagBreathe: Monitor breathing with commodity RFID systems". In: *IEEE Transactions on Mobile Computing* 19.4 (2019), pp. 969–981.

[114] Jintao Zhang et al. "RF-RES: Respiration monitoring with COTS RFID tags by doppler-shift". In: *IEEE Sensors Journal* 21.21 (2021), pp. 24844–24854.

[115] Meng Jin et al. "Fliptracer: Practical parallel decoding for backscatter communication". In: *Proceedings of the 23rd Annual International Conference on Mobile Computing and Networking*. 2017, pp. 275–287.

[116] Xiulong Liu et al. "Fast and accurate detection of unknown tags for RFID systems–hash collisions are desirable". In: *IEEE/ACM Transactions on Networking* 28.1 (2020), pp. 126–139.

[117] Yunfei Ma, Nicholas Selby, and Fadel Adib. "Minding the billions: Ultra-wideband localization for deployed rfid tags". In: *Proceedings of the 23rd annual international conference on mobile computing and networking*. 2017, pp. 248–260.

[118] Jue Wang, Deepak Vasisht, and Dina Katabi. "RF-IDraw: Virtual touch screen in the air using RF signals". In: *ACM SIGCOMM Computer Communication Review* 44.4 (2014), pp. 235–246.

[119] Chao Yang, Xuyu Wang, and Shiwen Mao. "RFID-pose: Vision-aided three-dimensional human pose estimation with radio-frequency identification". In: *IEEE transactions on reliability* 70.3 (2020), pp. 1218–1231.

[120] Syed Aziz Shah et al. "Privacy-preserving non-wearable occupancy monitoring system exploiting Wi-Fi imaging for next-generation body centric communication". In: *Micromachines* 11.4 (2020), p. 379.

[121] Liang Ma et al. "Room-level fall detection based on ultra-wideband (UWB) monostatic radar and convolutional long short-term memory (LSTM)". In: *Sensors* 20.4 (2020), p. 1105.

[122] Yonglong Tian et al. "RF-based fall monitoring using convolutional neural networks". In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2.3 (2018), pp. 1–24.

[123] Muhammad Wasim Raad, Mohamed Deriche, and Olfa Kanoun. "An RFID-based monitoring and localization system for dementia patients". In: *2021 18th International Multi-Conference on Systems, Signals & Devices (SSD)*. IEEE. 2021, pp. 1–7.

[124] Wafa Shuaieb et al. "RFID RSS fingerprinting system for wearable human activity recognition". In: *Future Internet* 12.2 (2020), p. 33.

[125] Lin Feng et al. "SitR: Sitting posture recognition using RF signals". In: *IEEE Internet of Things Journal* 7.12 (2020), pp. 11492–11504.

[126] Licai Zhu et al. “TagCare: Using RFIDs to monitor the status of the elderly living alone”. In: *IEEE Access* 5 (2017), pp. 11364–11373.

[127] Koichi Toda, Banno Wattana, and Norihiko Shinomiya. “Machine learning-based fall detection system using smart passive UHF sensor tag on the shoes”. In: (2020).

[128] Wenjie Ruan et al. “Tagfall: Towards unobstructive fine-grained fall detection based on uhf passive rfid tags”. In: *proceedings of the 12th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services on 12th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services*. 2015, pp. 140–149.

[129] Xuanke He et al. “RFID based non-contact human activity detection exploiting cross polarization”. In: *IEEE Access* 8 (2020), pp. 46585–46595.

[130] Cao Dian et al. “Towards domain-independent complex and fine-grained gesture recognition with RFID”. In: *Proceedings of the ACM on Human-Computer Interaction* 4.ISS (2020), pp. 1–22.

[131] Jumin Zhao et al. “RF-motion: a device-free RF-based human motion recognition system”. In: *Wireless Communications and Mobile Computing* 2021 (2021), pp. 1–9.

[132] George A Oguntala et al. “SmartWall: Novel RFID-enabled ambient human activity recognition using machine learning for unobtrusive health monitoring”. In: *IEEE Access* 7 (2019), pp. 68022–68033.

[133] Zhongqin Wang et al. “A see-through-wall system for device-free human motion sensing based on battery-free RFID”. In: *ACM Transactions on Embedded Computing Systems (TECS)* 17.1 (2017), pp. 1–21.

[134] Hanchuan Li, Can Ye, and Alanson P Sample. “IDSense: A human object interaction detection system based on passive UHF RFID”. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 2015, pp. 2555–2564.

[135] Mohammad Mehedi Hassan et al. “Human emotion recognition using deep belief network architecture”. In: *Information Fusion* 51 (2019), pp. 10–18.

[136] Alvin Raj et al. “Rao-Blackwellized particle filters for recognizing activities and spatial context from wearable sensors”. In: *Experimental Robotics: The 10th International Symposium on Experimental Robotics*. Springer. 2008, pp. 211–221.

[137] Yi-Liang Kuo et al. “Measuring distance walked and step count in children with cerebral palsy: an evaluation of two portable activity monitors”. In: *Gait & posture* 29.2 (2009), pp. 304–310.

[138] Anindya Das Antar, Masud Ahmed, and Md Atiqur Rahman Ahad. "Challenges in sensor-based human activity recognition and a comparative analysis of benchmark datasets: A review". In: *2019 Joint 8th International Conference on Informatics, Electronics & Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision & Pattern Recognition (icIVPR)*. IEEE. 2019, pp. 134–139.

[139] Martin Liggins II, David Hall, and James Llinas. *Handbook of multisensor data fusion: theory and practice*. CRC press, 2017.

[140] Chen Chen, Roozbeh Jafari, and Nasser Kehtarnavaz. "A survey of depth and inertial sensor fusion for human action recognition". In: *Multimedia Tools and Applications* 76 (2017), pp. 4405–4425.

[141] Andreas Bulling, Ulf Blanke, and Bernt Schiele. "A tutorial on human activity recognition using body-worn inertial sensors". In: *ACM Computing Surveys (CSUR)* 46.3 (2014), pp. 1–33.

[142] Yongcai Guo, Weihua He, and Chao Gao. "Human activity recognition by fusing multiple sensor nodes in the wearable sensor systems". In: *Journal of Mechanics in Medicine and Biology* 12.05 (2012), p. 1250084.

[143] Long Liu, Zhelong Wang, and Sen Qiu. "Driving behavior tracking and recognition based on multisensors data fusion". In: *IEEE Sensors Journal* 20.18 (2020), pp. 10811–10823.

[144] Thien Huynh-The et al. "Physical activity recognition with statistical-deep fusion model using multiple sensory data for smart health". In: *IEEE Internet of Things Journal* 8.3 (2020), pp. 1533–1543.

[145] S Sandhya Rani, G Apparao Naidu, and V Usha Shree. "Kinematic joint descriptor and depth motion descriptor with convolutional neural networks for human action recognition". In: *Materials Today: Proceedings* 37 (2021), pp. 3164–3173.

[146] Lichen Wang et al. "Generative multi-view human action recognition". In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2019, pp. 6212–6221.

[147] Hossein Rahmani and Mohammed Bennamoun. "Learning action recognition model from depth and skeleton videos". In: *Proceedings of the IEEE international conference on computer vision*. 2017, pp. 5832–5841.

[148] Han Zou et al. "WiFi and vision multimodal learning for accurate and robust device-free human activity recognition". In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*. 2019, pp. 0–0.

[149] Seungeun Chung et al. "Sensor data acquisition and multimodal sensor fusion for human activity recognition using deep learning". In: *Sensors* 19.7 (2019), p. 1716.

[150] Yao Shu et al. “Research on human motion recognition based on Wi-Fi and inertial sensor signal fusion”. In: *2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/AT* IEEE. 2018, pp. 496–504.

[151] Ramin Ramezani, Yubin Xiao, and Arash Naeim. “Sensing-Fi: Wi-Fi CSI and accelerometer fusion system for fall detection”. In: *2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*. IEEE. 2018, pp. 402–405.

[152] Wang Zhongqin et al. “Computer vision-assisted region-of-interest rfid tag recognition and localization in multipath-prevalent environments”. In: *Proc. ACM Interact. Mobile Wear. Ubiq. Technol* 3.29 (2019), pp. 1–30.

[153] Jianxin Wu et al. “A scalable approach to activity recognition based on object use”. In: *2007 IEEE 11th international conference on computer vision*. IEEE. 2007, pp. 1–8.

[154] Xinxin Mei et al. “Energy efficient real-time task scheduling on CPU-GPU hybrid clusters”. In: *IEEE INFOCOM 2017-IEEE Conference on Computer Communications*. IEEE. 2017, pp. 1–9.

[155] Zijuan Liu, Xiulong Liu, and Keqiu Li. “Deeper exercise monitoring for smart gym using fused rfid and cv data”. In: *IEEE INFOCOM 2020-IEEE Conference on Computer Communications*. IEEE. 2020, pp. 11–19.

[156] Kaiyan Cui et al. “ShakeReader:’Read’UHF RFID Using Smartphone”. In: *IEEE Transactions on Mobile Computing* (2021).

[157] Tara Boroushaki et al. “Robotic grasping of fully-occluded objects using rf perception”. In: *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 2021, pp. 923–929.

[158] Clinical Trials Transformation Initiative et al. “CTTI Recommendations: developing novel endpoints generated by mobile technology for use in clinical trials”. In: *Clinical Trials Transformation Initiative* (2017).

[159] Riley A Bloomfield et al. “Machine learning and wearable sensors at preoperative assessments: functional recovery prediction to set realistic expectations for knee replacements”. In: *Medical Engineering & Physics* 89 (2021), pp. 14–21.

[160] Sen Qiu et al. “Sensor combination selection strategy for kayak cycle phase segmentation based on body sensor networks”. In: *IEEE Internet of Things Journal* 9.6 (2021), pp. 4190–4201.

[161] Sibasankar Padhy. “A tensor-based approach using multilinear SVD for hand gesture recognition from SEMG signals”. In: *IEEE Sensors Journal* 21.5 (2020), pp. 6634–6642.

[162] Ming Guo, Zhelong Wang, and Ning Yang. "Aerobic exercise recognition through sparse representation over learned dictionary by using wearable inertial sensors". In: *Journal of Medical and Biological Engineering* 38 (2018), pp. 544–555.

[163] CD Katsis et al. "A wearable system for the affective monitoring of car racing drivers during simulated conditions". In: *Transportation research part C: emerging technologies* 19.3 (2011), pp. 541–551.

[164] Xiyang Peng et al. "Experimental analysis of artificial neural networks performance for accessing physical activity recognition in daily life". In: *2020 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCloud/SocialCom/Sus* IEEE. 2020, pp. 1348–1353.

[165] Hongyu Zhao et al. "Adaptive gait detection based on foot-mounted inertial sensors and multi-sensor fusion". In: *Information Fusion* 52 (2019), pp. 157–166.

[166] Kaitlin G Rabe et al. "Ultrasound sensing can improve continuous classification of discrete ambulation modes compared to surface electromyography". In: *IEEE Transactions on Biomedical Engineering* 68.4 (2020), pp. 1379–1388.

[167] Susanna Spinsante et al. "A mobile application for easy design and testing of algorithms to monitor physical activity in the workplace". In: *Mobile Information Systems* 2016 (2016).

[168] Yixiang Dai et al. "Wearable biosensor network enabled multimodal daily-life emotion recognition employing reputation-driven imbalanced fuzzy classification". In: *Measurement* 109 (2017), pp. 408–424.

[169] Lalit Bahl et al. "Maximum mutual information estimation of hidden Markov model parameters for speech recognition". In: *ICASSP'86. IEEE international conference on acoustics, speech, and signal processing*. Vol. 11. IEEE. 1986, pp. 49–52.

[170] Junji Yamato, Jun Ohya, and Kenichiro Ishii. "Recognizing human action in time-sequential images using hidden Markov model." In: *CVPR*. Vol. 92. 1992, pp. 379–385.

[171] Sheikh Badar Ud Din Tahir, Ahmad Jalal, and Kibum Kim. "Wearable inertial sensors for daily activity analysis based on Adam optimization and the maximum entropy Markov model". In: *Entropy* 22.5 (2020), p. 579.

[172] Jian Wu and Roozbeh Jafari. "Orientation independent activity/gesture recognition using wearable motion sensors". In: *IEEE Internet of Things Journal* 6.2 (2018), pp. 1427–1437.

[173] Lukun Wang. "Recognition of human activities using continuous autoencoders with wearable sensors". In: *Sensors* 16.2 (2016), p. 189.

[174] Nicholas C Jacobson et al. "Deep learning paired with wearable passive sensing data predicts deterioration in anxiety disorder symptoms across 17–18 years". In: *Journal of affective disorders* 282 (2021), pp. 104–111.

[175] Mario Munoz-Organero and Ramona Ruiz-Blazquez. "Time-elastic generative model for acceleration time series in human activity recognition". In: *Sensors* 17.2 (2017), p. 319.

[176] Xi Zhou, Junqi Guo, and Shenling Wang. "Motion recognition by using a stacked autoencoder-based deep learning algorithm with smart phones". In: *Wireless Algorithms, Systems, and Applications: 10th International Conference, WASA 2015, Qufu, China, August 10-12, 2015, Proceedings 10*. Springer. 2015, pp. 778–787.

[177] Belkacem Chikhaoui and Frank Gouineau. "Towards automatic feature extraction for activity recognition from wearable sensors: a deep learning approach". In: *2017 IEEE international conference on data mining workshops (ICDMW)*. IEEE. 2017, pp. 693–702.

[178] Qin Song et al. "An evolutionary deep neural network for predicting morbidity of gastrointestinal infections by food contamination". In: *Neurocomputing* 226 (2017), pp. 16–22.

[179] Pascal Vincent et al. "Extracting and composing robust features with denoising autoencoders". In: *Proceedings of the 25th international conference on Machine learning*. 2008, pp. 1096–1103.

[180] Laith Alzubaidi et al. "Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions". In: *Journal of big Data* 8 (2021), pp. 1–74.

[181] Shuo-Yiin Chang and Nelson Morgan. "Robust CNN-based speech recognition with Gabor filter kernels". In: *Fifteenth annual conference of the international speech communication association*. 2014.

[182] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. "Imagenet classification with deep convolutional neural networks". In: *Advances in neural information processing systems* 25 (2012).

[183] Karen Simonyan and Andrew Zisserman. "Two-stream convolutional networks for action recognition in videos". In: *Advances in neural information processing systems* 27 (2014).

[184] Hyeong Kyu Jang, Hobeom Han, and Sang Won Yoon. "Comprehensive monitoring of bad head and shoulder postures by wearable magnetic sensors and deep learning". In: *IEEE Sensors Journal* 20.22 (2020), pp. 13768–13775.

[185] Weiwei Zhang et al. "Partial directed coherence based graph convolutional neural networks for driving fatigue detection". In: *Review of Scientific Instruments* 91.7 (2020).

[186] Jiantao Yang and Yuehong Yin. "Novel soft smart shoes for motion intent learning of lower limbs using LSTM with a convolutional autoencoder". In: *IEEE Sensors Journal* 21.2 (2020), pp. 1906–1917.

[187] Nicolas Boulanger-Lewandowski, Yoshua Bengio, and Pascal Vincent. "Modeling temporal dependencies in high-dimensional sequences: Application to polyphonic music generation and transcription". In: *arXiv preprint arXiv:1206.6392* (2012).

[188] Ilya Sutskever, James Martens, and Geoffrey E Hinton. "Generating text with recurrent neural networks". In: *Proceedings of the 28th international conference on machine learning (ICML-11)*. 2011, pp. 1017–1024.

[189] Long Short-Term Memory. "Long short-term memory". In: *Neural computation* 9.8 (2010), pp. 1735–1780.

[190] Javier Conte Alcaraz, Sanam Moghaddamnia, and Jürgen Peissig. "Efficiency of deep neural networks for joint angle modeling in digital gait assessment". In: *EURASIP Journal on Advances in Signal Processing* 2021.1 (2021), pp. 1–20.

[191] Xiaoyu Duan et al. "A Generative Adversarial Networks for Log Anomaly Detection." In: *Comput. Syst. Sci. Eng.* 37.1 (2021), pp. 135–148.

[192] Lam Tran et al. "Multi-model long short-term memory network for gait recognition using window-based data segment". In: *IEEE Access* 9 (2021), pp. 23826–23839.

[193] Murtadha D Hssayeni et al. "Ensemble deep model for continuous estimation of Unified Parkinson's Disease Rating Scale III". In: *Biomedical engineering online* 20 (2021), pp. 1–20.

[194] S Ai, A Chakravorty, and C Rong. "2019 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)". In: (2019).

[195] Sakorn Mekruksavanich and Anuchit Jitpattanakul. "Lstm networks using smartphone data for sensor-based human activity recognition in smart homes". In: *Sensors* 21.5 (2021), p. 1636.

[196] Junyoung Chung et al. "Empirical evaluation of gated recurrent neural networks on sequence modeling". In: *arXiv preprint arXiv:1412.3555* (2014).

[197] Nidhi Dua, Shiva Nand Singh, and Vijay Bhaskar Semwal. "Multi-input CNN-GRU based human activity recognition using wearable sensors". In: *Computing* 103 (2021), pp. 1461–1478.

[198] Shuqin Yang et al. "Robust navigation method for wearable human–machine interaction system based on deep learning". In: *IEEE Sensors Journal* 20.24 (2020), pp. 14950–14957.

[199] Ji-Hae Kim et al. “deepGesture: Deep learning-based gesture recognition scheme using motion sensors”. In: *Displays* 55 (2018), pp. 38–45.

[200] Zheng Yang et al. “Mobility increases localizability: A survey on wireless indoor localization using inertial sensors”. In: *ACM Computing Surveys (Csur)* 47.3 (2015), pp. 1–34.

[201] Jialin Liu et al. “A research on CSI-based human motion detection in complex scenarios”. In: *2017 IEEE 19th International Conference on e-Health Networking, Applications and Services (Healthcom)*. IEEE. 2017, pp. 1–6.

[202] Germán Martín Mendoza-Silva, Joaquín Torres-Sospedra, and Joaquín Huerta. “A meta-review of indoor positioning systems”. In: *Sensors* 19.20 (2019), p. 4507.

[203] S Sophia et al. “Bluetooth Low Energy based Indoor Positioning System using ESP32”. In: *2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA)*. IEEE. 2021, pp. 1698–1702.

[204] Xinyue Yang et al. “Human posture recognition in intelligent healthcare”. In: *Journal of Physics: Conference Series*. Vol. 1437. 1. IOP Publishing. 2020, p. 012014.

[205] Qammer H Abbasi et al. *Advances in body-centric wireless communication: Applications and state-of-the-art*. 65. Institution of Engineering and Technology, 2016.

[206] Daniyal Haider et al. “An efficient monitoring of eclamptic seizures in wireless sensors networks”. In: *Computers & Electrical Engineering* 75 (2019), pp. 16–30.

[207] Sarah Naja, MMED Makhlouf, and Mohamad Abdul Halim Chehab. “An ageing world of the 21st century: a literature review”. In: *Int J Community Med Public Health* 4.12 (2017), pp. 4363–9.

[208] José Antonio Santoyo-Ramón, Eduardo Casilari, and José Manuel Cano-García. “Analysis of a smartphone-based architecture with multiple mobility sensors for fall detection with supervised learning”. In: *Sensors* 18.4 (2018), p. 1155.

[209] Juris Klonovs et al. *Distributed computing and monitoring technologies for older patients*. Springer, 2015.

[210] Zheng Yang, Zimu Zhou, and Yunhao Liu. “From RSSI to CSI: Indoor localization via channel response”. In: *ACM Computing Surveys (CSUR)* 46.2 (2013), pp. 1–32.

[211] Nishtha Chopra et al. “THz time-domain spectroscopy of human skin tissue for in-body nanonetworks”. In: *IEEE Transactions on Terahertz Science and Technology* 6.6 (2016), pp. 803–809.

[212] J Adarsh, P Vishak, and R Gandhiraj. “Adaptive noise cancellation using NLMS algorithm in GNU radio”. In: *2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS)*. IEEE. 2017, pp. 1–4.

[213] Jiangang Hao and Tin Kam Ho. “Machine learning made easy: a review of scikit-learn package in python programming language”. In: *Journal of Educational and Behavioral Statistics* 44.3 (2019), pp. 348–361.

[214] Ahsen Tahir et al. “A novel functional link network stacking ensemble with fractal features for multichannel fall detection”. In: *Cognitive Computation* 12.5 (2020), pp. 1024–1042.

[215] Isaac Cohen and Hongxia Li. “Inference of human postures by classification of 3D human body shape”. In: *2003 IEEE international SOI conference. Proceedings (cat. No. 03CH37443)*. IEEE. 2003, pp. 74–81.

[216] Wadii Boulila et al. “Noninvasive detection of respiratory disorder due to covid-19 at the early stages in saudi arabia”. In: *Electronics* 10.21 (2021), p. 2701.

[217] Fadel Adib, Zachary Kabelac, and Dina Katabi. “Multi-person motion tracking via RF body reflections”. In: (2014).

[218] Bo Wei et al. “From real to complex: Enhancing radio-based activity recognition using complex-valued CSI”. In: *ACM Transactions on Sensor Networks (TOSN)* 15.3 (2019), pp. 1–32.

[219] Siamak Yousefi et al. “A survey on behavior recognition using WiFi channel state information”. In: *IEEE Communications Magazine* 55.10 (2017), pp. 98–104.

[220] Yongsen Ma, Gang Zhou, and Shuangquan Wang. “WiFi sensing with channel state information: A survey”. In: *ACM Computing Surveys (CSUR)* 52.3 (2019), pp. 1–36.

[221] Iria Santos et al. “Artificial neural networks and deep learning in the visual arts: A review”. In: *Neural Computing and Applications* 33.1 (2021), pp. 121–157.

[222] Mohammed Al-Sarem et al. “A novel hybrid deep learning model for detecting COVID-19-related rumors on social media based on LSTM and concatenated parallel CNNs”. In: *Applied Sciences* 11.17 (2021), p. 7940.

[223] Safa Ben Atitallah et al. “Randomly initialized convolutional neural network for the recognition of COVID-19 using X-ray images”. In: *International journal of imaging systems and technology* 32.1 (2022), pp. 55–73.

[224] Safa Ben Atitallah et al. “Fusion of convolutional neural networks based on Dempster–Shafer theory for automatic pneumonia detection from chest X-ray images”. In: *International Journal of Imaging Systems and Technology* 32.2 (2022), pp. 658–672.

[225] Klaus Greff et al. “LSTM: A search space odyssey”. In: *IEEE transactions on neural networks and learning systems* 28.10 (2016), pp. 2222–2232.

[226] Nils Y Hammerla, Shane Halloran, and Thomas Plötz. "Deep, convolutional, and recurrent models for human activity recognition using wearables". In: *arXiv preprint arXiv:1604.08880* (2016).

[227] DP Leong et al. "World Population Prospects 2019. Department of Economic and Social Affairs Population Dynamics. New York (NY): United Nations; 2019 (https://population. un. org/wpp/Download/, accessed 20 September 2020). The decade of healthy ageing. Geneva: World Health Organization". In: *World* 73.7 (2018), 362k2469.

[228] Unated Nations. "Department of Economic and Social Affairs, population division". In: *Int Migr Rep* (2015).

[229] Davide Calvaresi et al. "Exploring the ambient assisted living domain: a systematic review". In: *Journal of Ambient Intelligence and Humanized Computing* 8 (2017), pp. 239–257.

[230] Rytis Maskeliūnas, Robertas Damaševičius, and Sagiv Segal. "A review of internet of things technologies for ambient assisted living environments". In: *Future Internet* 11.12 (2019), p. 259.

[231] George Oguntala et al. "Indoor location identification technologies for real-time IoT-based applications: An inclusive survey". In: *Computer Science Review* 30 (2018), pp. 55–79.

[232] Yuchu He et al. "Miniaturized circularly polarized Doppler radar for human vital sign detection". In: *IEEE Transactions on Antennas and Propagation* 67.11 (2019), pp. 7022–7030.

[233] Zhou Ren, Jingjing Meng, and Junsong Yuan. "Depth camera based hand gesture recognition and its applications in human-computer-interaction". In: *2011 8th international conference on information, communications & signal processing*. IEEE. 2011, pp. 1–5.

[234] Daphne Townsend, Frank Knoefel, and Rafik Goubran. "Privacy versus autonomy: a tradeoff model for smart home monitoring technologies". In: *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE. 2011, pp. 4749–4752.

[235] Dimitrios N Serpanos and Andreas Papalambrou. "Security and privacy in distributed smart cameras". In: *Proceedings of the IEEE* 96.10 (2008), pp. 1678–1687.

[236] Dongheng Zhang, Yang Hu, and Yan Chen. "Mtrack: Tracking multiperson moving trajectories and vital signs with radio signals". In: *IEEE Internet of Things Journal* 8.5 (2020), pp. 3904–3914.

[237] Abubakar Sharif et al. "Low-cost inkjet-printed RFID tag antenna design for remote healthcare applications". In: *IEEE Journal of Electromagnetics, RF and Microwaves in Medicine and Biology* 3.4 (2019), pp. 261–268.

[238] George Roussos and Vassilis Kostakos. “RFID in pervasive computing: state-of-the-art and outlook”. In: *Pervasive and Mobile Computing* 5.1 (2009), pp. 110–131.

[239] Daqiang Zhang et al. “TASA: Tag-free activity sensing using RFID tag arrays”. In: *IEEE Transactions on Parallel and Distributed Systems* 22.4 (2010), pp. 558–570.

[240] Jie Wang et al. “Device-free wireless sensing: Challenges, opportunities, and applications”. In: *IEEE network* 32.2 (2018), pp. 132–137.

[241] Ramyar Saeedi et al. “Personalized human activity recognition using wearables: A manifold learning-based knowledge transfer”. In: *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE. 2018, pp. 1193–1196.

[242] Chen Ying and Zhang Fu-Hong. “A system design for UHF RFID reader”. In: *2008 11th IEEE International Conference on Communication Technology*. IEEE. 2008, pp. 301–304.

[243] Ju Wang et al. “D-watch: Embracing" bad" multipaths for device-free localization with COTS RFID devices”. In: *Proceedings of the 12th International on Conference on emerging Networking EXperiments and Technologies*. 2016, pp. 253–266.

[244] Qian Zhang et al. “RFree-ID: An unobtrusive human identification system irrespective of walking cofactors using cots RFID”. In: *2018 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. IEEE. 2018, pp. 1–10.

[245] Wenjie Ruan et al. “Device-free human localization and tracking with UHF passive RFID tags: A data-driven approach”. In: *Journal of Network and Computer Applications* 104 (2018), pp. 78–96.

[246] Wenjie Ruan et al. “Device-free indoor localization and tracking through human-object interactions”. In: *2016 IEEE 17th international symposium on a world of wireless, mobile and multimedia networks (WoWMoM)*. IEEE. 2016, pp. 1–9.

[247] Run Zhao et al. “Gesture recognition with RFID: an experimental study”. In: *CCF Transactions on Pervasive Computing and Interaction* 3 (2021), pp. 397–412.

[248] Shang Jiang et al. “RF-Gait: Gait-Based Person Identification with COTS RFID”. In: *Wireless Communications and Mobile Computing* 2022 (2022).

[249] Yi Liu et al. “TransTM: A device-free method based on time-streaming multiscale transformer for human activity recognition”. In: *Defence Technology* (2023).

[250] Xinyu Li et al. “Deep learning for rfid-based activity recognition”. In: *Proceedings of the 14th ACM Conference on Embedded Network Sensor Systems CD-ROM*. 2016, pp. 164–175.

[251] Xiaoyi Fan, Wei Gong, and Jiangchuan Liu. "Tagfree activity identification with rfids". In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2.1 (2018), pp. 1–23.

[252] Ashish Vaswani et al. "Attention is all you need". In: *Advances in neural information processing systems* 30 (2017).

[253] Sepp Hochreiter and Jürgen Schmidhuber. "Long Short-Term Memory". In: *Neural Computation* 9.8 (1997), pp. 1735–1780. DOI: 10.1162/neco.1997.9.8.1735.

[254] Jacob Devlin et al. "Bert: Pre-training of deep bidirectional transformers for language understanding". In: *arXiv preprint arXiv:1810.04805* (2018).

[255] Kaiming He et al. "Deep residual learning for image recognition". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, pp. 770–778.

[256] Jimmy Lei Ba, Jamie Ryan Kiros, and Geoffrey E Hinton. "Layer normalization". In: *arXiv preprint arXiv:1607.06450* (2016).

[257] Chuyu Wang et al. "Multi-touch in the air: Device-free finger tracking and gesture recognition via cots rfid". In: *IEEE INFOCOM 2018-IEEE Conference on Computer Communications*. IEEE. 2018, pp. 1691–1699.

[258] Hojjat Salehinejad and Shahrokh Valaee. "LiteHAR: lightweight human activity recognition from WIFI signals with random convolution kernels". In: *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE. 2022, pp. 4068–4072.

[259] Jin Zhang et al. "Data augmentation and dense-LSTM for human activity recognition using WiFi signal". In: *IEEE Internet of Things Journal* 8.6 (2020), pp. 4628–4641.

[260] Kevin Chapron et al. "Highly accurate bathroom activity recognition using infrared proximity sensors". In: *IEEE Journal of Biomedical and Health Informatics* 24.8 (2019), pp. 2368–2377.

[261] Ming Tao et al. "Jointly optimization for activity recognition in secure IoT-enabled elderly care applications". In: *Applied Soft Computing* 99 (2021), p. 106788.

[262] Liangqi Yuan and Jia Li. "Smart cushion based on pressure sensor array for human sitting posture recognition". In: *2021 IEEE Sensors*. IEEE. 2021, pp. 1–4.

[263] Jay-Shian Tan et al. "Human activity recognition for people with knee osteoarthritis—a proof-of-concept". In: *Sensors* 21.10 (2021), p. 3381.

[264] Zawar Hussain, Quan Z Sheng, and Wei Emma Zhang. "A review and categorization of techniques on device-free human activity recognition". In: *Journal of Network and Computer Applications* 167 (2020), p. 102738.

[265] Mayank Lovanshi and Vivek Tiwari. "Human skeleton pose and spatio-temporal feature-based activity recognition using ST-GCN". In: *Multimedia Tools and Applications* (2023), pp. 1–26.

[266] Daniel Garcia-Gonzalez et al. "New machine learning approaches for real-life human activity recognition using smartphone sensor-based data". In: *Knowledge-Based Systems* 262 (2023), p. 110260.

[267] Bin Hu et al. "BioTag: Robust RFID-based continuous user verification using physiological features from respiration". In: *Proceedings of the Twenty-Third International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing*. 2022, pp. 191–200.

[268] Anubhav Natani, Abhishek Sharma, and Thinagaran Perumal. "Sequential neural networks for multi-resident activity recognition in ambient sensing smart homes". In: *Applied Intelligence* 51 (2021), pp. 6014–6028.

[269] Reza Shahbazian and Irina Trubitsyna. "Human Sensing by using Radio Frequency Signals: A Survey on Occupancy and Activity Detection". In: *IEEE Access* (2023).

[270] *EmeraldInno*. `https://www.emeraldinno.com/`. Accessed: 2024-02-15.

[271] *Walabot Fall Alert System: Detect Falls With No Wearables*. (Visited on 02/15/2024).

[272] K. Thompson et al. *xkcorp*. Accessed: 2024-02-15. URL: `https://xkcorp.com/`.

[273] Adrian Lin and Hao Ling. "Doppler and direction-of-arrival (DDOA) radar for multiple-mover sensing". In: *IEEE transactions on aerospace and electronic systems* 43.4 (2007), pp. 1496–1509.

[274] AG Stove. "Modern FMCW radar-techniques and applications". In: *First European Radar Conference, 2004. EURAD.* IEEE. 2004, pp. 149–152.

[275] Sruthy Skaria, Akram Al-Hourani, and Robin J Evans. "Deep-learning methods for hand-gesture recognition using ultra-wideband radar". In: *IEEE Access* 8 (2020), pp. 203580–203590.

[276] Maytus Piriyajitakonkij et al. "SleepPoseNet: Multi-view learning for sleep postural transition recognition using UWB". In: *IEEE Journal of Biomedical and Health Informatics* 25.4 (2020), pp. 1305–1314.

[277] Muhammad Attique Khan et al. "Hand-crafted and deep convolutional neural network features fusion and selection strategy: an application to intelligent human action recognition". In: *Applied Soft Computing* 87 (2020), p. 105986.

[278] Md Zia Uddin et al. "A body sensor data fusion and deep recurrent neural network-based behavior recognition approach for robust healthcare". In: *Information Fusion* 55 (2020), pp. 105–115.

[279] Jack Andrews and Jia Li. “Human Detection and Biometric Authentication with Ambient Sensors”. In: *Biomedical Sensing and Analysis: Signal Processing in Medicine and Biology*. Springer, 2022, pp. 55–98.

[280] Jenny Liu et al. “Human presence detection via deep learning of passive radio frequency data”. In: *2019 IEEE National Aerospace and Electronics Conference (NAECON)*. IEEE. 2019, pp. 296–301.

[281] Jindong Wang et al. “Deep learning for sensor-based activity recognition: A survey”. In: *Pattern recognition letters* 119 (2019), pp. 3–11.

[282] Chrsitine F Martindale et al. “Wearables-based multi-task gait and activity segmentation using recurrent neural networks”. In: *Neurocomputing* 432 (2021), pp. 250–261.

[283] Francis R Willett et al. “High-performance brain-to-text communication via handwriting”. In: *Nature* 593.7858 (2021), pp. 249–254.

[284] Sepp Hochreiter. “The vanishing gradient problem during learning recurrent neural nets and problem solutions”. In: *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 6.02 (1998), pp. 107–116.

[285] Isura Nirmal et al. “Deep learning for radio-based human sensing: Recent advances and future directions”. In: *IEEE Communications Surveys & Tutorials* 23.2 (2021), pp. 995–1019.

[286] Sen Qiu et al. “Multi-sensor information fusion based on machine learning for real applications in human activity recognition: State-of-the-art and research challenges”. In: *Information Fusion* 80 (2022), pp. 241–265.

[287] Muhammad Shahzad and Alex X Liu. “Probabilistic optimal tree hopping for RFID identification”. In: *IEEE/ACM Transactions on Networking* 23.3 (2014), pp. 796–809.

[288] Leixian Shen et al. “PRDL: Relative localization method of RFID tags via phase and RSSI based on deep learning”. In: *IEEE access* 7 (2019), pp. 20249–20261.

[289] Damian E Grzechca, Piotr Pelczar, and Lukas Chruszczyk. “Analysis of object location accuracy for iBeacon technology based on the RSSI path loss model and fingerprint map”. In: *International Journal of Electronics and Telecommunications* (2016).

[290] Longfei Shangguan et al. “STPP: Spatial-temporal phase profiling-based method for relative RFID tag localization”. In: *IEEE/ACM Transactions on Networking* 25.1 (2016), pp. 596–609.

[291] Zhengliang Zhu et al. “A Dataset of Human Motion Status Using IR-UWB Through-wall Radar”. In: *arXiv preprint arXiv:2008.13598* (2020).

[292] William Taylor et al. “Radar sensing for activity classification in elderly people exploiting micro-doppler signatures using machine learning”. In: *Sensors* 21.11 (2021), p. 3881.

[293] Bin Cao et al. “RFID reader anticollision based on distributed parallel particle swarm optimization”. In: *IEEE Internet of Things Journal* 8.5 (2020), pp. 3099–3107.