Non-Invasive AI-Driven Human Activity Recognition

Fahad Ayaz

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

James Watt School of Engineering
College of Science and Engineering
University of Glasgow

July 2025

Abstract

The rapid proliferation of Internet of Things technologies, coupled with artificial intelligence-driven applications, has revolutionised human activity recognition, enabling pervasive real-time monitoring across smart homes, healthcare, security, and ambient-assisted living environments. This transformation holds particular significance for healthcare systems, as radar-based recognition of physical and physiological activities facilitates continuous remote monitoring through invasive and non-invasive technologies, supporting personalised care and early intervention at scale. Traditionally, activity recognition systems have relied primarily on invasive or contact-based devices, such as wearables and biosensors, which often lead to user discomfort, require frequent maintenance or charging, and risk non-compliance, especially among elderly individuals. Conversely, cameras, Wi-Fi, and radar are all treated as non-invasive sensing modalities; however, cameras raise serious privacy concerns and are constrained by lighting conditions, whereas Wi-Fi-based sensing suffers from multipath interference and spectrum-sharing challenges. Radar sensing emerges as a promising tool and privacy-preserving alternative with robustness to environmental variations. Despite these advantages, systems built on radar for activity recognition face significant challenges in real-world applications. This thesis addresses three critical challenges in radar-based human activity recognition: enabling non-intrusive recognition of both macro-level physical activities (e.g., falls, gait) and micro-level physiological signals (e.g., heart rate, respiration rate); data diversity and radar domain adaptation; and ensuring energy-efficient, privacy-aware edge deployment. The first contribution addresses the challenges of non-intrusive recognition of macro-level human activities and radar domain adaptation by developing a radar signal processing framework that transforms complex signals into four two-dimensional domain representations for robust activity recognition. By integrating domain-specific preprocessing with transfer learning, the framework improves adaptability across environments and reduces the complexity of raw signal data. Experimental results show up to 29.36% improvement in recognition accuracy compared to a baseline convolutional neural network, with transfer learning models achieving 96.03% on the primary dataset and demonstrating strong generalisation across two additional radar datasets. Building upon this, the second contribution focuses on finer-grained sensing, addressing the challenge of non-intrusive monitoring of micro-level physiological signals by extending the system to support radar-based extraction of vital signs such as heart rate and respiration rate. This is achieved using two radar modalities:

ultra-wideband and millimetre-wave frequency-modulated continuous wave radar. A comprehensive analysis was conducted to evaluate the impact of varying distances and radar positioning configurations on the accuracy of vital sign extraction. The third contribution addresses the challenges of domain adaptation and energy efficiency by optimising transfer learning models for lightweight, energy-efficient, and privacy-aware deployment on edge devices. Using post-training quantisation and selective domain-model pairing, the system significantly reduces computational costs while maintaining high recognition performance across radar domains. Results indicate energy consumption as low as 0.42 mWh and response times of 1.32 seconds for 5-second activities, confirming its suitability for real-time, on-device monitoring. Additionally, the framework incorporates differential privacy techniques to strengthen local inference privacy with minimal loss in accuracy. Collectively, these contributions enhance the scalability, robustness, and efficiency of activity recognition systems, paving the way for non-invasive, AI-driven applications in healthcare and real-world environments.

Keywords

Human Activity Recognition, Macro- and Micro-level Activity Monitoring, Radar Sensing, Transfer Learning, Energy Efficiency, Edge Computing, Radar Domain Adaptation.

Acknowledgements

I extend my deepest gratitude to all those who have supported and guided me throughout my doctoral journey. First and foremost, I am profoundly grateful to my supervisor, **Dr. Ahmed Zoha**, whose exceptional mentorship and unwavering support have been the cornerstone of my academic and personal development. His insightful advice and encouragement have continually inspired me to strive for excellence. His dedication to fostering a collaborative and intellectually stimulating environment has enhanced my research skills and broadened my perspective. His patience, wisdom, and willingness to guide me through challenges have been invaluable.

I am also deeply indebted to my co-supervisor, **Prof. Sajjad Hussain**, for his persistent support and invaluable insights. His visionary leadership has been instrumental in shaping the direction of my research. His ability to anticipate future trends and encourage innovative thinking has significantly enriched my academic experience. Under his guidance, I have learnt to approach problems with a forward-thinking mindset, which has not only enhanced the quality of my work but also prepared me for future challenges in the field. I am incredibly grateful to **Prof. Muhammad Ali Imran** for their persistent support and invaluable insights. Their expertise and guidance have significantly enriched my research experience.

A special thanks go to **Basim Alhumaily**, **Habib Ullah Manzoor**, and **Ahsan Raza khan**, whose collaborative spirit and scholarly discussions have greatly contributed to the progression of my work. I want to thank the University of Glasgow for providing a nurturing and stimulating academic environment, especially during challenging times. The resources and support offered have been instrumental in facilitating my research endeavours.

Above all, I am eternally grateful to my parents, whose unwavering love and support have been the foundation of all my achievements. Their endless encouragement and belief in my abilities have given me the strength and determination to pursue my dreams. To my siblings, thank you for your constant support and always being there when I needed you. To my loving wife, whose patience, understanding, and encouragement have been invaluable throughout this journey, thank you for your endless support and for being my pillar of strength.

Declaration

University of Glasgow

College of Science & Engineering

Statement of Originality

Name: Fahad Ayaz
Registration Number:

I certify that the thesis presented here for examination for a PhD degree in the University of Glasgow is solely my own work other than where I have 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 review of the thesis, the examination may be postponed pending the outcome of any investigation of the issues.

Signature:
Date: 23/07/2025

Dedication

To my beloved parents, whose endless prayers and support have been the foundation of my journey. To my siblings, for their constant encouragement and belief in my dreams.

And to my wife, my greatest companion, thank you for standing beside me through every high and low, for your unwavering patience, strength, and love.

This thesis is as much yours as it is mine.

Contents

List of Tables

List of Figures

List of Publications

Journal

Published Papers

1. **Fahad Ayaz**, B. Alhumaily, S. Hussain, M. A. Imran, K. Arshad and K. Assaleh, and A. Zoha, “Radar Signal Processing and Its Impact on Deep Learning-Driven Human Activity Recognition,” in *Sensors*, vol. 25, no. 3, p. 724, January 2025.
   Doi: https://doi.org/10.3390/s25030724

2. A. R. Khan, H. U. Manzoor, **F. Ayaz**, M. A. Imran, and A. Zoha, “A Privacy and Energy-Aware Federated Framework for Human Activity Recognition,” in *Sensors*, vol. 23, no. 23, p. 9339, Nov. 2023, doi:https://doi.org/10.3390/s23239339

Under Review

1. **Fahad Ayaz**, B. Alhumaily, S. Hussain, M. A. Imran, and A. Zoha, “Evaluating Radar Signal Preprocessing Techniques for Improved Transfer Learning in Human Activity Recognition,” in *Elsevier Pervasive and Mobile Computing*, 2025, ID: PMC-D-25-00359.

2. **Fahad Ayaz**, B. Alhumaily, S. Hussain, M. A. Imran, and A. Zoha, “Edge-optimised, Privacy-Preserving Human Activity Recognition Using FMCW Radar,” *Elsevier Internet of Things*, 2025.

Conference Proceedings

1. **F. Ayaz**, M. S. Khan, S. Hussain, W. Ahmad, F. Kawsar, M. A. Imran, and A. Zoha, "Contact-free vital sign estimation using ultra-wide band radar,” in Proc. 29th IEEE International Conference on Electronics, Circuits and Systems (ICECS), Glasgow, United Kingdom, Oct. 2022, pp. 1–4, doi: https://doi.org/10.1109/ICECS202256217.2022.9970856.

2. **F. Ayaz**, B. Alhumaily, S. Hussain, M. A. Imran, and A. Zoha, "VMD integrated FMCW radar system for non-invasive pediatric vital sign estimation,” in Proc. 28th IEEE Inter-

national Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD), Edinburgh, United Kingdom, Oct. 2023, pp. 264–269, doi: https://doi.org/10.1109/CAMAD59638.2023.10478425.

3. **F. Ayaz**, B. Alhumaily, S. Hussain, L. Mohjazi, M. A. Imran, and A. Zoha, "Integrating millimeter-wave FMCW radar for investigating multi-height vital sign monitoring," in Proc. IEEE Wireless Communications and Networking Conference (WCNC), Dubai, United Arab Emirates, Mar. 2024, pp. 1–6, doi: https://doi.org/10.1109/WCNC57260.2024.10570899.

4. N. AlQallaf, **F. Ayaz**, S. Bhatti, S. Hussain, A. Zoha, and R. Ghannam, "Solar energy systems design in 2D and 3D: A comparison of user vital signs," in Proc. 29th IEEE International Conference on Electronics, Circuits and Systems (ICECS), Glasgow, United Kingdom, Oct. 2022, pp. 1–4, doi: https://doi.org/10.1109/ICECS202256217.2022.9971065

5. H. U. Manzoor, M. S. Khan, A. R. Khan, **F. Ayaz**, D. Flynn, M. A. Imran, and A. Zoha, "Fedclamp: An algorithm for identification of anomalous client in federated learning," in Proc. 29th IEEE International Conference on Electronics, Circuits and Systems (ICECS), Glasgow, United Kingdom, Oct. 2022, pp. 1–4, doi: https://doi.org/10.1109/ICECS202256217.2022.9970909

Book Chapter

1. A. R. Khan, H. U. Manzoor, **F. Ayaz**, M. A. Imran, and A. Zoha, "Hybrid neuromorphic federated learning (HNFL): A spiking-LSTM for human activity recognition using wearable sensors," in *Multimodal Intelligent Sensing in Modern Applications*, 1st ed., Hoboken, NJ, USA: Wiley, 2024, ch. 6, doi: https://doi.org/10.1002/9781394257744.ch6.

List of Abbreviations

5G	Fifth Generation
ADAM	Adaptive Moment Estimation
AI	Artificial Intelligence
AAL	Ambient Assisted Living
ADC	Analog to Digital Converter
B5G	Beyond Fifth Generation
BPM	Beats Per Minute
BrPM	Breaths Per Minute
CSI	Channel State Information
CNN	Convolutional Neural Network
CAE	Convolutional Auto Encoder
CW	Continous Wave
DL	Deep Learning
DP	Differential Privacy
DNN	Deep Neural Network
DFT	Discrete Fourier Transform
EI	Edge Intelligence
EPR	Energy Precision Ratio
FFT	Fast Fourier Transform
FN	False Negative
FP	False Positive
FT	Fine Tuning
FMCW	Frequency Modulated Continuous Wave
GPU	Graphics Processing Unit
GHz	Giga Hertz
HAR	Human Activity Recognition
HR	Heart Rate
HPF	High Pass Filter
IoT	Internet of Things
IoMT	Internet of Medical Things

IF	Intermediate Frequency
KNN	K-Nearest Neighbor
LSTM	Long Short-Term Memory
LPF	Low Pass Filter
ML	Machine Learning
MAE	Mean Absolute Error
MCU	Microcontroller Unit
mD	Micro Doppler
MEC	Mobile Edge Computing
MIMO	Multiple-Input Multiple-Output
mmWave	Millimetre Wave
mWh	Milli Watt Hour
MTI	Moving Target Indicator
PSD	Power Spectral Density
PTQ	Post-Training Quantisation
QAT	Quantisation-Aware Training
RD	Range Doppler
R-HAR	Radar-Based Human Activity Recognition
RF	Radio Frequency
RSSI	Received Signal Strength Indicator
RR	Respiratory Rate
RMSE	Root Mean Squared Error
SBC	Single Board Computer
SGD	Stochastic Gradient Descent
SPWVD	Smoothed Pseudo Winger Ville Distribution
STD	Standard Deviation
SVM	Support Vector Machine
SNN	Spiking Neural Network
TF	Time Frequency
TL	Transfer Learning
THz	Terahertz
TF	Tensorflow
TP	True Positives
TN	True Negatives
TD	Time Doppler
UWB	Ultra-wideband
WHO	World Health Organisation

List of Symbols

B	Bandwidth
f_b	Beat Frequency
f_c	Center Frequency
β	Probability of class label
EPI	Energy per Item
$h(\tau)$	Frequency Smoothing Window
I	In-phase
$w(t)$	Instantaneous angular frequency
$\varphi(t)$	Instantaneous Phase
$\bar{t}_{\text{infer}}$	Inference Time of model
D	Latency of System
MAE	Mean Absolute Error
f_m	Modulation Frequency
ε	Privacy budget in differential privacy
Q	Quadrature
R_x	Receiver Antenna
σ_{infer}	Standard Deviation of Inference time
α	Slope of Transmitted Signal
c	Speed of Light
t_{cap}	Time required for image capture
d_{tx}	Time required for data preprocessing
t_{inf}	Time required for model inference
τ	Time Delay
$w(.)$	Time Window
T_x	Transmitter Antenna
λ	Wavelength of radar antenna

Chapter 1

Introduction

1.1 Background

The ageing population in the United Kingdom (UK) presents mounting healthcare challenges, with Office for National Statistics data showing the proportion of individuals aged 65 years and above increasing from 16.4% in 2011 to nearly 20% in 2024, projected to reach 25% by 2050 [1]. This demographic shift has intensified pressure on the National Health Service (NHS), particularly regarding hospital capacity, where the average bed occupancy reached 89.8% in 2023, significantly exceeding the 85% safety threshold [2]. A primary contributor to this strain is "bed blocking", where medically fit patients remain hospitalised due to delayed community or social care arrangements [3]. As illustrated in Fig. 1.1, bed occupancy has emerged as one of the most pressing concerns among NHS hospitals, showing an increasing trend over the past decade and often exceeding 90% in recent years, which highlights the sustained pressure on healthcare industry [4]. Age-related conditions exacerbate this problem, with falls being particularly significant; they cost the NHS over £2.3 billion annually, affect approximately one-third of individuals age over 65 years [5], and frequently result in extended hospitalisations with complex discharge requirements. These interconnected challenges of an ageing demographic, increased prevalence of chronic conditions, limited hospital capacity, and the specific burden of falls collectively underscore the urgent need for innovative approaches to healthcare delivery for older adults.

To address these challenges, the UK government has prioritised investment in technology-enabled care, allocating £30 million in 2023 to expand virtual wards and remote monitoring solutions [6]. Within this context, Human Activity Recognition (HAR) has emerged as a foundational field to support independent living for older and vulnerable individuals while reducing pressure on inpatient services. HAR automatically identifies and classifies human physical and physiological activities by analysing data from various sensing devices [7]. The systems or algorithms that perform these tasks are commonly referred to as HAR systems (HARS). Within the domain of HAR, human activities are broadly classified into two principal categories [8]:

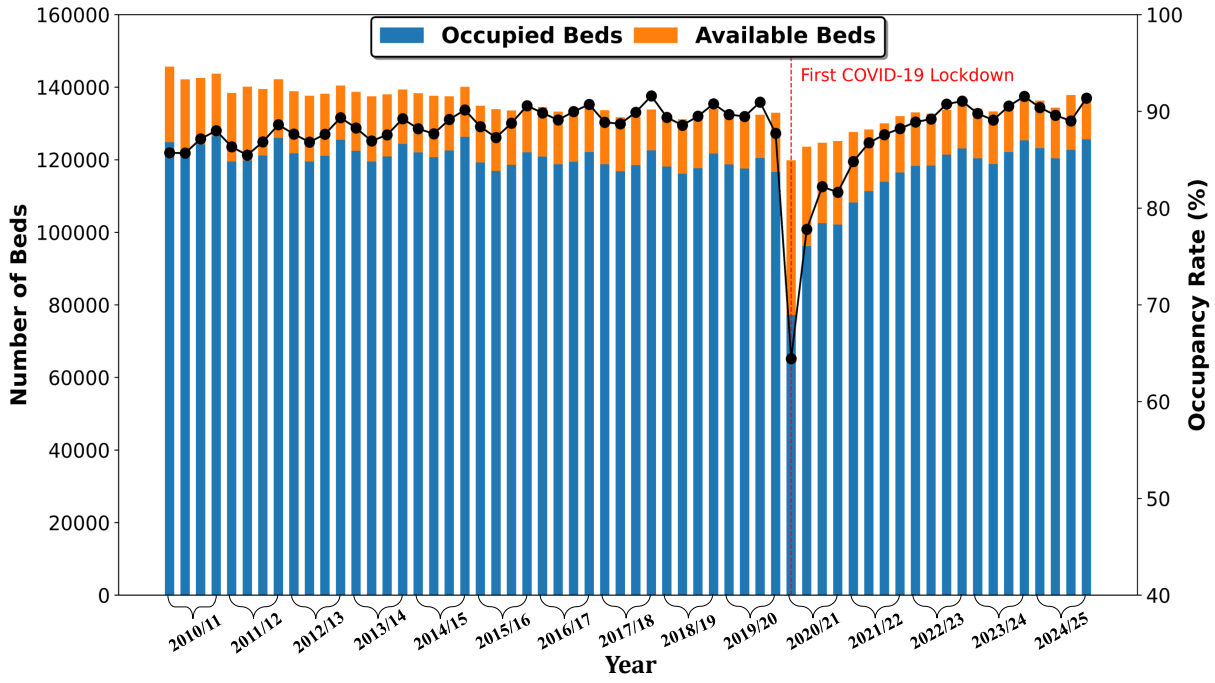


Figure 1.1: Most reported operational issues by NHS hospitals, with bed occupancy as the top concern [4].

1. **Macro-Activities:** These involve large, easily observable movements, such as walking, sitting, lying down, or falling. The detection of such actions is critical for applications such as fall detection and mobility monitoring.

2. **Micro-Activities:** These encompass subtle physiological and behavioural signals, including breathing patterns, chest wall displacement, and minor gestures. Continuous monitoring of these micro-activities, which represent vital signs such as heart rate (HR) and respiration rate (RR), enables the early detection of health deterioration and supports chronic disease management.

Several UK-based companies have developed commercial HAR solutions specifically designed for elderly care. For example, Tunstall healthcare provides telecare and telehealth systems with fall detection and emergency response capabilities [9]. Canary care offers smart home monitoring that analyses patterns of daily living to detect anomalies [10], while Howz uses energy consumption data to infer behavioural changes indicative of health risks [11]. These systems support older adults in maintaining independence by enabling continuous remote monitoring, thereby reducing hospital admissions and alleviating pressure on the UK healthcare system.

Despite these promising developments, conventional HARS, particularly those based on wearable and vision-based sensors, faces significant obstacles:

- **User Compliance and Comfort:** Despite their tracking capabilities, wearable devices face adoption barriers among older adults due to requirements for consistent interaction

that challenge individuals with cognitive or physical limitations. Long-term adherence is hindered by physical discomfort, battery maintenance requirements, and usability issues, including complex interfaces and small displays that collectively limit effectiveness in elderly populations [12].

- **Privacy and Ethical Concerns:** Vision-based HARS deliver rich contextual data but introduce privacy concerns through continuous visual capture in personal spaces, which proves particularly problematic in homes and care facilities [13]. Their performance varies with environmental conditions, such as lighting and physical obstructions, which further reduce reliability [14]. These privacy and performance limitations create significant barriers to acceptance among vulnerable users, despite the technical capabilities of the systems.

- **Activity Detection Accuracy and Reliability:** Recognising both macro and micro-activities in uncontrolled indoor environments presents significant technical challenges [8]. Differentiating between similar actions, filtering background noise, and capturing subtle physiological changes remains difficult. Wearables may fail to detect nuanced micro-movements and often require user input during emergencies, which is impractical when the individual is incapacitated [15].

- **System Integration and Scalability:** Deploying HARS at scale within existing healthcare infrastructures, such as the NHS, requires interoperability, standardisation, and extensive technical support. Adapting these technologies to support large and diverse populations involves substantial financial investment and long-term sustainability planning.

In response to these challenges, radio frequency (RF)-based HAR sensing has gained popularity owing to its advantages over vision and wearable systems [16]. RF-based approaches offer distinct advantages due to their non-invasive nature, eliminating the need for users to wear, charge, or interact with devices, while simultaneously preserving privacy by not capturing identifiable visual information. Among RF technologies, radar stands out as particularly promising, demonstrating performance advantages over Wi-Fi-based alternatives in both activity classification and vital sign monitoring [17]. Operating at higher frequencies, radar provides enhanced Doppler resolution and greater sensitivity to both macro and micro-motions. This capability enables more accurate detection of physiological changes that are crucial for health monitoring applications.

1.2 Scope and Motivation

Radar technology is increasingly recognised as a key enabler for next-generation HARS, offering contactless, privacy-preserving, and robust monitoring of human activities across diverse environments [18]. In contrast to other RF-based sensors such as WiFi or RFID [19], which often

suffer from lower spatial resolution, are more susceptible to environmental interference, and rely on requirements like the placement of RFID tags, radar operates non-intrusively and can sense through obstacles. This results in superior robustness and reliability, particularly for complex or cluttered indoor environments. Furthermore, radar systems provide high Doppler resolution and are highly sensitive to both macro- and micro-level human motions, enabling accurate detection across a wide range of activities without requiring user compliance or compromising privacy.

The integration of radar with artificial intelligence (AI) applications and Internet of Things (IoT) devices has created unprecedented opportunities for continuous, real-time activity monitoring in healthcare and ambient-assisted living (AAL) contexts. Radar-enabled systems directly support critical initiatives such as "ageing in place" [20], by facilitating early detection of adverse events and subtle behavioural changes, thereby reducing hospital admissions and alleviating pressure on strained healthcare systems. While traditional machine learning (ML) approaches applied to radar data have relied on handcrafted features [21], deep learning (DL) methodologies have demonstrated remarkable potential by automatically extracting discriminative spatio-temporal features from raw radar signals [22], substantially improving recognition accuracy while reducing dependence on manual feature engineering.

Despite the advantages of radar-based HAR (R-HAR), current frameworks face several challenges that hinder their widespread adoption and effectiveness in complex real-world environments. One key challenge is non-intrusive multi-scale monitoring, which in this context refers to the recognition of both macro-level physical activities (e.g. walking, falling), and micro-level physiological signals (e.g. heart rate (HR), respiration rate (RR)) within a unified system. Radar has been shown to be highly effective in capturing large-scale body movements due to its sensitivity to motion and robustness to environmental conditions [15, 23]. It has also demonstrated a strong capability in detecting subtle physiological signals through micro-Doppler analysis [24, 25]. However, most existing research focuses on either macro- or micro-activities in isolation, with few approaches attempting to integrate both levels within a comprehensive framework [19]. This separation complicates unified feature representation and cohesive system design, particularly in the presence of signal sparsity and environmental noise. Developing radar-based systems capable of robust multi-scale analysis, which here refers to the combined processing of macro- and micro-activity data, without relying on wearable devices or compromising user privacy remains a critical research direction for healthcare and AAL environments.

Another significant challenge is data diversity, as radar systems deployed across different settings often generate non-independent and identically distributed (non-IID) data due to variations in user behaviour, device specifications, sensor placement, and environmental conditions [26]. This heterogeneity complicates the training of DL models and often results in suboptimal generalisation when applied beyond the original training context. To achieve accurate recognition across such diverse conditions, complex models are often required, which increases the training cost and makes deployment on resource-constrained systems more difficult [27, 28]. In addi-

tion, most existing studies rely on a single-domain radar input, typically time-Doppler (TD) patterns [23, 29], which limits the model's ability to fully capture the diverse characteristics present in radar signals [30]. In this context, multiple radar domains refer to distinct signal representations derived from the same radar data, such as range, velocity, and phase. Evaluating each of these domains individually enables a clearer understanding of their respective strengths, limitations, and suitability for various HAR tasks.

For real-world deployment, R-HARS face a critical challenge in terms of achieving energy efficient systems, which involves both computational costs and communication overhead. The computation cost refers to the amount of energy required to process radar signals and perform inference of DL models on edge devices, whereas the communication cost is the amount of energy needed to transmit data or model updates during inference or training. Traditional HARS rely on cloud-based processing, which increases latency and raises privacy concerns due to the external transmission of sensitive data [31]. Edge computing offers a promising alternative by enabling local processing on resource-constrained devices [32], and has therefore become a preferred solution for real-time HAR in IoT applications. However, deploying large deep neural networks (DNNs) on these devices is limited by memory, computing capacity, and power consumption constraints [33]. These limitations are particularly critical for battery-powered systems, where continuous radar inference can significantly drain energy resources. In addition, maintaining data privacy during on-device processing remains a major concern, particularly in healthcare settings that involve sensitive physiological data. This thesis specifically considers privacy at the local inference stage, focusing on protecting personal health information through techniques such as differential privacy (DP). Addressing these challenges is essential to enable reliable, energy-efficient, and privacy-preserving radar-based activity recognition in real-world settings.

1.3 Problem Statement and Objectives

Despite the advancements in R-HAR, significant challenges persist in achieving optimal efficiency, security, and robustness across various intelligent applications. The scope of this thesis addresses three critical challenges that limit the deployment of R-HARS in real-world scenarios, which include:

C1 Non-Intrusive Multi-Scale Monitoring: Radar sensing, due to its non-intrusive, contactless nature and robustness to environmental conditions, is well-suited for capturing both macro-activities, and micro-physiological signals within indoor settings. For instance, detecting a fall followed by post-event HR variability could provide a critical context for emergency response [34]. However, most existing R-HARS target only one activity scale, limiting their effectiveness in continuous and holistic health monitoring. The divergent temporal and spectral characteristics of macro and micro-signals complicate unified sig-

nal processing and feature extraction, while environmental clutter further degrades the detection accuracy. Strong micro-Doppler (mD) signatures from limb movements can obscure subtle chest displacements associated with breathing, and conventional ML models often lack the complex structure needed to separate these mixed signals occurring at different scales [35]. Moreover, achieving balanced sensitivity and specificity across both macro- and micro-activities, while maintaining robust performance in dynamic and cluttered environments remains an open problem. These limitations underscore the need for integrated approaches capable of simultaneously and reliably interpreting human activity across scales without requiring user interaction or compromising privacy.

C2 Data Diversity: A core limitation of R-HARS lies in their vulnerability to data diversity, which significantly affects model generalisation across users and deployment environments. Radar signals vary due to differences in body morphology, individual movement patterns, sensor placement, and ambient conditions, leading to non-IID data distributions. These variations correspond to well-established heterogeneity types, including subject-level, spatial, and modality-based differences [36]. The high-dimensional and non-linear nature of radar signals further complicates the learning of robust and transferable features. Traditional ML models often rely on handcrafted features that fail to capture these complexities, whereas DL approaches require large volumes of annotated data that are expensive and difficult to obtain. Although using different radar domain representations may offer improved informativeness and efficiency, there is limited understanding of which domains are best suited for learning generalisable activity-specific features under data scarcity and heterogeneous conditions. Together, these factors constrain the scalability of R-HARS, particularly for applications involving diverse populations and real-world environments where frequent retraining is impractical.

C3 Energy Efficiency: Energy efficiency is another critical challenge in HAR research, particularly for resource-constrained edge devices, such as IoT sensors that coupled with radar systems. In R-HARS, energy consumption stems from two main sources: computational cost, which involves the energy required for both model training and inference, and communication cost, which is incurred when transmitting radar data, extracted features, or model updates between edge devices and remote servers. For battery-powered or standalone systems, these demands can lead to rapid energy depletion, increased latency, and reduced operational lifetime. Traditional DL models are typically large and computationally intensive, necessitating effective model compression techniques to reduce the memory footprint and processing requirements without sacrificing accuracy. In addition to these constraints, preserving user data privacy during inference adds further complexity, especially in sensitive settings within healthcare. While radar sensors avoid capturing identifiable visual information, they still process behavioural and physiological data that

require protection. These challenges highlight the need for lightweight, energy-efficient, and privacy-aware HAR frameworks that support real-time operation without compromising recognition performance.

1.3.1 Aims and Objectives

In response to the three challenges identified in the problem statement, the objectives of this thesis are outlined as follows:

1. Design and implement R-HAR frameworks capable of reliable multi-scale activity recognition, supporting both macro- and micro-activities. This involves developing signal processing and learning strategies that address overlapping motion patterns and environmental interference while ensuring non-intrusiveness and robustness in real-world healthcare environments.

2. Develop generalisable HAR algorithms that perform reliably across diverse users, settings, and radar signal characteristics. This includes evaluating multiple radar representations, applying transfer learning (TL) to mitigate data scarcity, and avoiding the limitations of handcrafted features or retraining during each deployment scenario.

3. Design and implement a scalable and energy-efficient R-HAR framework optimised for resource-constrained edge devices. The aim is to develop lightweight and compressed DL models that minimise computational and communication costs while supporting real-time responsiveness. This also includes integrating differential privacy (DP) techniques to protect sensitive physical data during inference, ensuring user privacy without sacrificing performance or sustainability during standalone deployment.

1.4 Research Contributions

The main contributions of this thesis are itemised as follows:

- The first contribution of this thesis is the development of a radar signal preprocessing and domain representation framework that addresses challenges **C1** and **C2** by recognising macro-activities and enhancing model generalisation across users and environments. The framework is built on a 5.8 GHz Frequency Modulated Continuous Wave (FMCW) radar, which offers reliable Doppler and range information critical for both robust motion capture and fine-resolution analysis. This contribution introduces a unified methodology to evaluate and compare four radar domain representations: range-time (RT), range-Doppler (RD), TD based short-time Fourier transform (STFT), and smoothed pseudo Wigner-Ville distribution (SPWVD), along with their associated preprocessing pipelines and computational costs. Each domain is encoded as a two-dimensional (2D) image matrix, which

is subsequently analysed using a baseline convolutional neural network (CNN) to benchmark the feature extraction capability of each representation. To overcome data scarcity and improve recognition accuracy, the framework integrates state-of-the-art TL models, which significantly reduce the misclassification of critical events such as falls. The analysis includes key performance metrics, such as classification accuracy, preprocessing time, training duration, and inference latency, providing a holistic view of the trade-offs between model performance and computational efficiency. A subject-wise data splitting strategy ensures that models are evaluated on unseen individuals, reducing the risk of data leakage and promoting realistic generalisation. The effectiveness of the selected radar domain and model combinations is validated using two additional publicly available FMCW radar datasets operating at 24 and 77 GHz, confirming their robustness across different sensing technologies. This contribution culminates in a decision framework that guides the selection of the optimal radar domain and model pairs for real-time deployment in R-HARS.

- The second contribution of this thesis is the development of a radar-based vital sign monitoring framework that addresses challenge **C1** by enabling non-invasive recognition of micro-activities, specifically HR and RR. This study extends R-HARS beyond macro-activity detection by supporting continuous physiological monitoring through subtle chest wall motion analysis. To extract cardiopulmonary parameters, we implemented a dedicated signal processing pipeline using spectral estimation and peak detection techniques, which also enabled breathing pattern classification to distinguish between normal, shallow, and held respiration states. Initially, the study utilised ultra-wideband (UWB) radar operating at a centre frequency of 6.5 GHz with a single-subject setup to establish the feasibility of radar-based micro-activity monitoring. However, limitations in bandwidth efficiency, system integration, and scalability prompted a transition to millimetre wave (mmWave) FMCW radar, operating at 60.25 GHz with a 3.75 GHz bandwidth. The FMCW radar offers significant advantages, including improved hardware integration, lower power consumption, and better proximity sensing, while providing high Doppler and range resolution for the precise detection of fine physiological motions and reliable differentiation of multiple targets. All radar measurements were benchmarked against medical-grade reference sensors to confirm their clinical relevance. This contribution demonstrates the feasibility of radar as a unified, privacy-preserving sensing modality for the continuous monitoring of both macro- and micro-activities.

- The third contribution of this thesis is the development of an edge-optimised, privacy-aware R-HAR framework that addresses challenges **C2** and **C3** by enabling efficient deployment on resource-constrained devices while preserving user privacy. Building on previously selected radar domains, this contribution adapts DL models for practical use on platforms such as Raspberry Pi and Jetson Nano. GPU-based energy profiling was con-

ducted during model training to measure power usage and estimate carbon emissions for each architecture, thus supporting environmentally sustainable model selection aligned with the green AI principles. To address deployment constraints, the framework applies model compression through post-training quantisation (PTQ), significantly reducing model size and computational cost without compromising accuracy. Radar signal preprocessing pipelines were integrated to support domain representations suitable for real-time edge inference. The system was evaluated on edge devices using key metrics, including preprocessing time, inference latency, and energy precision ratio (EPR), highlighting the trade-offs between accuracy and energy consumption. To ensure data privacy, a lightweight local differential privacy (LDP) mechanism was implemented directly on the device, securing prediction outputs without retraining or impacting performance. Altogether, this contribution delivers a deployable, energy-efficient, and privacy-preserving R-HARS optimised for real-world use in constrained and sensitive settings. Crucially, it culminates in an end-to-end decision framework that identifies the optimal combination of radar representation, TL model, and edge hardware, which enables informed trade-offs across performance, efficiency, and privacy for robust deployment in healthcare and AAL environments.

1.5 Thesis Organisation

This thesis is structured into six chapters, each addressing specific aspects of the research challenges, proposed solutions, and their applications. The detailed organisation of the remaining chapters is as follows:

Chapter 1 provides an overview of the research context, motivation, and objectives of this study. It highlights the contributions and outlines the challenges addressed, laying the foundation for the subsequent chapters.

Chapter 2 comprehensively reviews the existing literature on non-invasive AI-driven HAR and its applications. It begins by exploring state-of-the-art sensing technologies for HAR, their limitations, and their applications in healthcare domains. This chapter also analyses the critical challenges associated with different HAR motions such as macro and micro-level, data diversity, DL models, privacy considerations, and energy efficiency. This review identifies research gaps, particularly in R-HARS on the basis of two different HAR activities, edge computing, privacy-utility tradeoffs, and computational efficiency for real-world settings. The chapter concludes with a discussion of potential application areas for R-HAR, emphasising the need for innovative frameworks to address the identified three challenges discussed in **Section 1.3**.

Chapter 3 addresses challenges **C1** and **C2** by developing a radar preprocessing framework for macro-activity recognition and improving model generalisation across users and environments. It compares four radar-domain techniques using a 5.8 GHz FMCW radar and evaluates

DL models in terms of accuracy, training time, and inference latency. A decision framework is proposed to guide the selection of radar-domain and model combinations for practical R-HAR applications.

Chapter 4 explores non-invasive vital sign monitoring using radar technology, highlighting the non-intrusive capability of radar to detect micro-physiological movements that tackle challenge **C1**. This chapter investigates two different radar systems, a UWB and mmWave FMCW radar, for estimating HR and RR. It evaluates the effects of distance, radar height positioning, and signal processing approaches on the measurement accuracy, providing valuable insights for contactless health monitoring applications.

Chapter 5 presents an edge-optimised, privacy-aware R-HAR framework that addresses challenges **C2** and **C3** by enhancing model generalisation in real-world scenarios and enabling efficient deployment on resource-constrained devices. It applies model compression to reduce computational cost, evaluates energy-efficient inference on low-power edge devices, and integrates a lightweight DP mechanism at the output stage. The framework supports real-time, privacy, and efficient activity recognition suitable for real-world use.

Chapter 6 concludes the thesis by summarising the key contributions, discussing their broader impact, and reflecting on the advancements made in R-HAR frameworks in addressing the three core challenges outlined in **Section 1.3**. It also outlines future research directions, including advanced ML architectures, multimodal sensor fusion, and stronger privacy-preserving methods, with potential to further enhance radar-based sensing and extend its application to new domains.

Chapter 2

Literature Review

In the rapidly evolving landscape of HAR, radar technology augmented by ML has emerged as a powerful and privacy-preserving solution capable of enabling scalable, contactless monitoring in real-world environments. This chapter begins by outlining the fundamental aspects of HARS, with a focus on non-invasive, AI-driven solutions for healthcare and real-world environments. It presents foundational concepts, including sensing modalities in HAR, and the rationale behind radar selection in real-world settings. Radar offers both privacy and security advantages for users, this chapter also explores established methods, such as DP, to protect data confidentiality. Furthermore, this chapter specifically aligns the literature review with the three critical challenges identified in Section 1.3, and how these factors impact model performance and convergence in R-HAR settings. For instance, Section 2.2, covers radar signal processing, emphasising its importance in robust multi-scale feature extraction for both macro- and micro-activities, key aspect of challenge **C1**. AI-driven techniques for HAR, including conventional ML and DL models, have been explored in Section 2.3, and 2.5 for their limitations in generalisation across users and environments, which relate to challenge **C2**, while also prompting the development of edge AI solutions, energy-efficient architectures, and privacy-preserving inference methods to support real-world deployment, aligning with challenge **C3**. Further examination of challenge **C1** is presented in Section 2.4, with a review of radar for micro-activity monitoring in healthcare domains and AAL environments, identifying current limitations and research gaps. Finally, Section 2.6, summarises the research gap and links it with three core challenges identified in the thesis.

2.1 Human Activity Recognition: An Overview

In the context of AAL, HARS enable the real-time monitoring of human activities, offering critical insights into behavioral patterns, health, and well-being. By accurately identifying and classifying human activities, these systems can contribute to the development of context-aware AAL environments. Such systems are designed to provide timely assistance, detect anoma-

lies, and enhance daily support for individuals, particularly in healthcare and care homes [37]. Recent advancements in HAR research have led to significant improvements in the algorithm performance and system capabilities [38]. HARS have been applied across diverse domains, including automotive safety [39], human-machine interaction [40], smart home automation [41], and healthcare monitoring [42]. These systems play a vital role in understanding and interpreting human movements across various contexts, offering both practical and theoretical value.

Typically, the HAR process can be divided into four key stages [43], as illustrated in Fig. 2.1: (1) capturing activity signals, (2) data preprocessing, (3) AI-based activity recognition, and (4) a user interface for activity management. Each stage can be implemented using various techniques, resulting in various design choices. Factors such as the application domain, type of data acquisition device, and selection of AI algorithms for activity detection further complicate these decisions. This chapter delves into the technologies and methodologies underlying these stages, with a particular focus on sensing modalities, signal processing, and AI-driven approaches. The following subsection explores the sensing modalities in HAR and compare their strengths and limitations in the context of real-world applications.

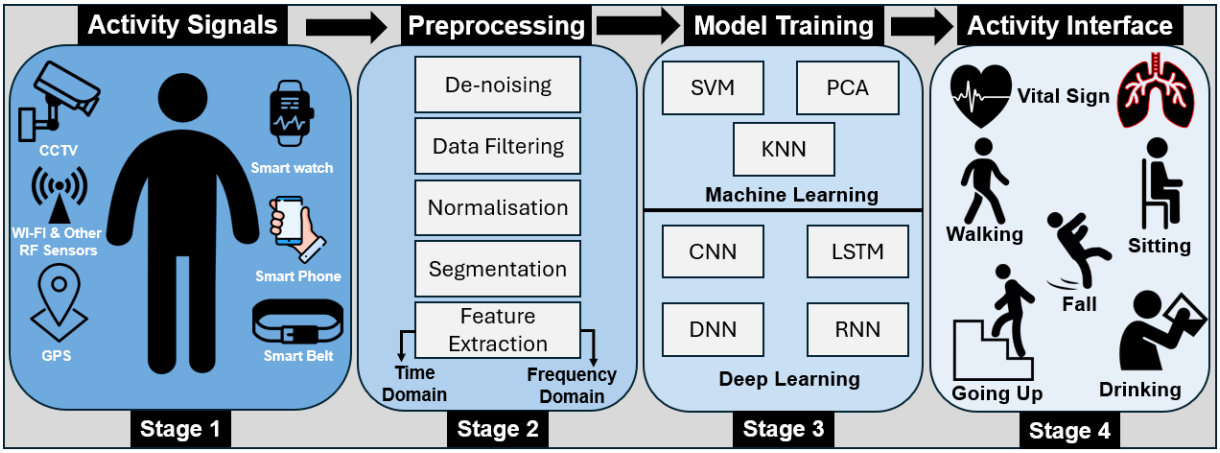


Figure 2.1: End-to-end HARS architecture for activity monitoring and classification

2.1.1 Sensing Modalities in Human Activity Recognition

Sensing technologies play a key role in healthcare, providing innovative solutions for monitoring, and diagnostics in AAL settings. At the core of HARS are algorithms that automatically identify activities based on the data collected from different sensors, as shown in Fig. 2.1 (stage 1). Over the years, various sensing modalities have been explored, including vision-based sensors [44], acoustic sensors [45], inertial measurement units (IMUs) [46], and radio frequency (RF) [47] sensors, as illustrated in Table 2.1, and broadly explained in detail below. The choice of sensing modality directly affects the accuracy, robustness, and applicability of the HARS.

Vision-Based Sensors

Vision-based sensors utilise visual data to interpret human actions and environmental context. This category includes standard red, green, and blue (RGB) cameras as well as advanced imaging technologies such as infrared cameras, thermal cameras, and depth sensors such as Microsoft Kinect [48]. These sensors capture rich spatial and temporal information, making them ideal for detailed motion analyses and posture recognition. However, their reliance on line-of-sight and susceptibility to varying lighting conditions can limit their performance in uncontrolled environments. Despite these challenges, vision-based sensing has been extensively adopted in domains such as healthcare, surveillance, and smart-home automation [49].

Acoustic Sensors

Acoustic sensing leverages audio signals to identify human activity and environmental events. Microphones or microphone arrays are typically used to capture sound patterns, which are then processed to detect specific activities, such as speech, footsteps, or ambient sounds [50]. Acoustic sensors are particularly useful in scenarios where visual data are unavailable or where privacy concerns restrict the use of cameras. However, these sensors are highly sensitive to noise interference, which can degrade their accuracy in dynamic and noisy environments. Advanced audio signal processing techniques such as noise suppression and feature extraction are often employed to address these limitations.

Inertial Sensors

Wearable inertial sensors, including accelerometers, gyroscopes, and magnetometers, are widely used because of their ability to capture motion and orientation data [51]. These sensors are often embedded in devices such as smartwatches, fitness trackers, or smartphones, making them convenient for everyday use. Inertial sensing is particularly effective for detecting fine-grained motion patterns such as walking, running, or hand gestures. However, these systems require consistent user compliance because improper device placement or usage can affect the performance. Recent advancements in ML algorithms have significantly improved the robustness of inertial sensing for HAR applications.

Radio Frequency Based Sensors

RF sensors, particularly those employing Wi-Fi [52] and radar [53], have gained more interest for various applications including HAR, due to their contactless nature. These sensors influence the unique characteristics of radio waves in the detection and classification of human activities. Wi-Fi-based systems use the fluctuations in wireless signal strength, such as received signal strength indicators (RSSI) and channel state information (CSI), caused by human movement

[52]. On the other hand, radar systems rely on Doppler shifts, range profiles, and mD signatures to analyse motion patterns, as detailed in Section 2.2.

Beyond Wi-Fi and radar, other RF-based technologies, such as Bluetooth, universal software radio peripherals (USRP), and channel sounders, are also relevant for HAR [54]. RF sensors exploit variations in signal propagation characteristics, including amplitude, phase, or frequency, caused by body movements. By analysing these changes, RF systems can capture the unique signatures of moving body segments, enabling precise and reliable activity recognition.

Multi-Modal Sensing

To enhance accuracy and robustness, HARS often integrate multiple sensing modalities, a practice known as multi-modal sensing [55]. By combining complementary data sources, such as visual and inertial sensors or acoustic and RF sensors, these systems can achieve higher reliability and contextual awareness. Advanced strategies such as feature-level and decision-level fusion enable the integration of diverse sensor outputs into a unified framework [56]. Additionally, context-aware algorithms and ensemble learning techniques further enhance system performance by continuously adjusting their models in response to shifts in environmental conditions and user behaviour patterns.

As summarised in Table 2.1, a wide range of sensing modalities has been explored for HARS, each offering unique advantages and presenting specific challenges. Among these, vision, inertial, and RF sensors are the most widely used because of their ability to capture diverse activity patterns. For simplicity, these sensors can be broadly classified into invasive and non-invasive systems.

2.1.2 Comparison of Invasive vs. Non-Invasive Human Activity Recognition System

HARS are broadly categorised into invasive and non-invasive systems, each presenting distinct strengths and limitations. Invasive systems typically require direct contact with the human body, such as wearables or implantables to capture physiological or motion-related data, offering high-fidelity signals but at the cost of user compliance and comfort. In contrast, non-invasive systems operate remotely through modalities such as cameras and RF sensors. Each category has its unique advantages and limitations, as detailed below:

- **Invasive HARS:** Invasive or contact-based HARS use physical sensors worn by users, including smartwatches, smart belts, and in-ear sensors. These devices capture precise physiological and kinematic signals, excelling in fitness monitoring, rehabilitation, and clinical applications [51]. However, they require consistent user compliance, creating barriers for continuous monitoring, especially among older adults or those with cognitive impairments [12]. Practical limitations such as discomfort, charging requirements,

Table 2.1: Overview of HARS implemented using various sensing modalities and associated techniques with their strength and limitations.

Sensing Type	Techniques and Sensors	References	Strengths and Weaknesses
Vision Sensing	Infrared, Thermal and wearable Cameras	[44, 57, 58]	High spatial resolution and interpretability; affected by lighting, occlusion, and privacy concerns.
	Depth sensors	[59]	
	Pose estimation	[60]	
Acoustic Sensing	Microphone	[61]	Low-cost and passive sensing; highly sensitive to ambient noise and lacks spatial detail.
	Audio signal processing	[50]	
	Acoustic event detection	[45]	
Inertial Sensing	Accelerometers	[51, 62]	Lightweight, energy-efficient, and widely used; prone to signal drift, calibration issues, and user compliance.
	Gyroscopes	[51, 63]	
	Magnetometers	[46]	
	Wearable sensors (e.g., smartwatches)	[64]	
RF Sensing	SISO and MIMO radars	[53, 65–67]	Non-intrusive and robust to lighting; requires complex signal processing and can be affected by multipath interference.
	Wi-Fi systems	[52, 68]	
	Pattern recognition technique	[69, 70]	
Multi-modal Sensing	Multi-modal sensor framework	[56]	Combines complementary strengths for enhanced accuracy; adds computational complexity and data synchronisation challenges.
	Sensor Fusion techniques	[55]	

and susceptibility to damage or calibration drift restrict their effectiveness for multi-scale monitoring scenarios, where tracking both macro- and micro-activities is necessary. Furthermore, these user-specific systems often face generalisation problems across populations or contexts, corresponding to the data heterogeneity issue [71].

- **Non-invasive HARS:** Such systems rely on remote sensing technologies such as vision-based sensors and RF sensors. These systems eliminate the need for physical contact, enabling more user-friendly and continuous monitoring. However, not all non-invasive

modalities are equally privacy-preserving; vision-based systems often raise significant privacy concerns in sensitive environments [13]. In contrast, RF-based methods like radar offer a contactless and privacy alternative, making them particularly suitable for long-term ambient monitoring [66]. Unlike Wi-Fi-based systems, radar sensors do not suffer from carrier frequency offsets, which can lead to noisy channel-frequency responses. Radar systems demonstrate robust performance in capturing both macro-activities (e.g., walking, falling) and micro-physiological signals (e.g., RR, HR), even under poor lighting or occlusion conditions [19]. This capability makes radar especially valuable in addressing the challenge of non-intrusive multi-scale monitoring, as it supports comprehensive and continuous observation without user intervention. Nevertheless, radar systems face a significant challenge regarding data diversity and generalisation due to the high variability in user posture, movement styles, environmental layout, and sensor positioning. This variability leads to non-independent and identically distributed (non-IID) data, where each instance reflects localised patterns and biases that hinder the generalisation of learned models across users and settings [26].

R-HARS is gaining popularity due to its robustness across diverse environmental conditions, including fog, rain, poor lighting, and darkness [72]. While it complements other sensing modalities such as visual, infrared, and wearable systems, two major challenges arises. First, multi-scale monitoring demands accurate tracking of both macro-activities and subtle physiological micro-signals to estimate vital sign accurately, requiring advanced signal processing. Second, radar data is highly heterogeneous and non-IID due to variability in user posture, movement, and environmental settings, limiting generalisation. Cluttered environment and overlapping micro-Doppler (mD) features further complicate classification, necessitating domain-adaptive ML models. Addressing these challenges forms the core of this thesis, as elaborated in Section 2.2 and Chapters 3 and 5. Table 2.2 provides a detailed comparison of invasive and non-invasive devices, summarising their sensor types, clinical usage, advantages, and limitations, which highlights the complementary role of these technologies in HAR.

2.1.3 Overview of Non-Invasive Macro-Activity Monitoring using Radar Technology

The comprehensive range of sensing technologies described in Section 2.1.2 and in Table 2.2, prompt the question: What is the rationale for adopting radar technology for HAR? Its main advantage is its contactless sensing capability, allowing it to recognise human posture, movements, and specific activities without the need to wear, carry, or interact with any sensors, which is especially useful for quickly identifying critical events, such as falls [15]. This feature significantly improves user compliance, especially for the elderly, who may forget to carry the sensor or recharge their batteries. Additionally, radar does not produce traditional visual images

Table 2.2: Comparison of invasive and non-invasive sensor technologies for HARS, including clinical applications, advantages, and limitations.

Category	Sensor Type	Clinical Applications	Advantages	Limitations	Ref.
Invasive Devices	Smartwatch	HR, RR Blood pressure, Physical activity	Notifications, Individualised, health tracking	Limited battery life, Inaccurate information, expensive	[73]
	Smart Belt	HR, RR	Precise, Reliable, Easy to connect	Required extra strap, uncomfortable	[74]
	Temperature Sensor	Body Temperature	Affordable, accurate	Limited utility under extreme conditions (e.g., freezing point of mercury)	[75]
	EEG Sensor	Brain disorder, Emotions, Sensitivity	High temporal precision	Limited spatial resolution, lacks precision in locating the source of brain activity	[76]
	In-ear sensor	HR, RR, Temperature	Stable placement Environmental isolation	Miniaturisation, Power consumption Signal artifacts, standardisation	[77]
Non-Invasive Devices	Other RF sensors	HR, RR, Gait analysis Human activity	Real-Time data, Scalability, Ease of integration	Interference, Security risks, Reliability Bandwidth sharing decrease the speed, Power consumption	[78, 79]
	Camera	Fall Detection, Physical activity, Target detection	High image resolution, Versatile for object and activity detection	Fails to preserve privacy, Unsuitable for certain environments, Poor performance in low-light conditions	[44]
	Radar sensor	Gait Analysis, Vital sign Fall detection	Privacy-preserving ,Non-intrusive, robust to noise, Capable of penetrating barriers, Range, velocity and angle measurements	Limited spatial resolution, Background noise sensitivity, need for advanced signal processing	[15, 80]

or videos, thereby protecting the privacy of individuals and their surroundings. These privacy-preserving features make radar sensor more acceptable to end-users, alleviating concerns about intrusion and data security, especially in the context of AAL.

In recent years, radar-based technologies have gained significant attention, with numerous researchers exploring their potential in specific healthcare domains [81,82]. As illustrated in Fig. 2.2, radar sensors have evolved into distinct healthcare research areas. This thesis focuses on two key applications: radar-based macro-activity recognition, which is discussed in this section and detailed further in Sections 2.3, 2.5, and Chapters 3 and 5, and radar-based micro-activity monitoring, which is detailed in Sections 2.4 and Chapter 4 for vital sign monitoring. Together, these applications demonstrate the versatility of radar as a comprehensive non-intrusive monitoring technology capable of detecting both physical activity and subtle physiological signals, representing the main contribution of this thesis. R-HARS consists of two main technological components: the radar system and the ML network, as illustrated in Fig. 2.3. The radar system captures human activity data, whereas the ML models process these data to identify and classify various activities, as shown in Fig. 2.1 and described in detail in Section 2.3.

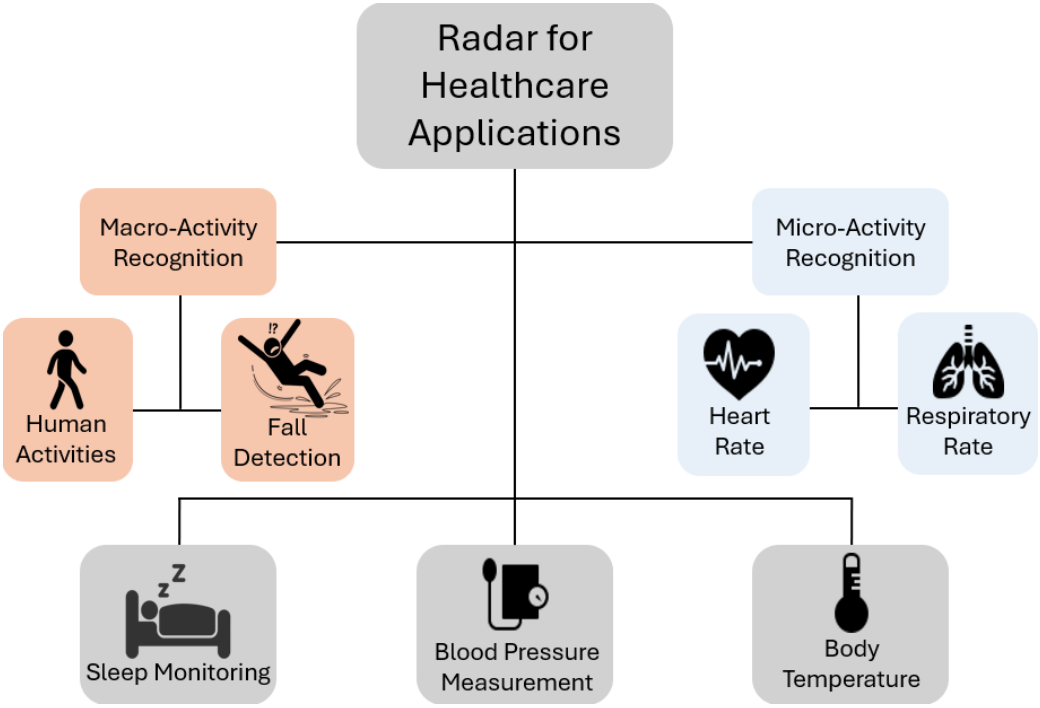


Figure 2.2: Categorisation of radar-based healthcare applications.

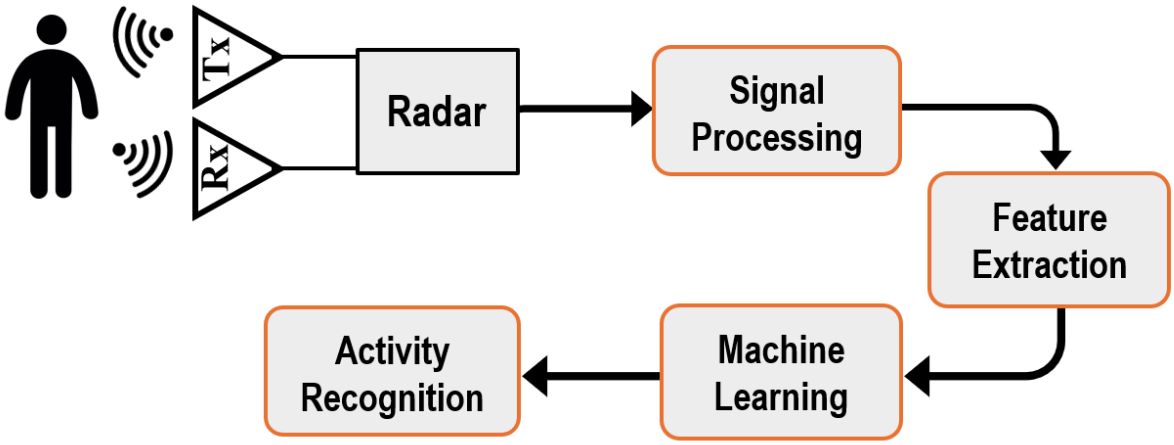


Figure 2.3: Overview of R-HARS architecture and data flow.

Radar Types for Non-Invasive Activity Sensing

Radar systems are categorised based on their signal transmission methodology, functioning either actively or passively. Active radar systems generate and transmit their own RF signals via an antenna toward targets and analyse the backscattered signals to determine crucial parameters, including range, velocity, and physical dimensions [83]. The reflected signals are then processed using a dedicated signal processor to extract meaningful information [84]. In contrast, passive radar systems operate without internal transmitters, instead utilising external illuminators of opportunity, such as television or radio broadcasts, for target detection and tracking purposes [83].

Table 2.3: The comparison of active radars configurations.

Feature	**Monostatic Radar**	**Bistatic Radar**	**Multistatic Radar**	**MIMO Radar**
Antenna Configuration	Single or co-located antenna for Tx and Rx	Separate antennas for Tx and Rx	Multiple antennas for Tx and/or Rx	Multiple antennas for Tx and Rx
Hardware Complexity	Lower	Higher	Higher (due to number of units)	Highest
Cost	Generally lower	Generally higher	Higher	Highest
Signal Processing	Simplified	More complex	More complex	Most complex
Synchronisation	Easier (co-located Tx/Rx)	More complex	More complex	Most complex
Latency	Generally lower	May vary	May vary	May vary
Coverage Area	Limited to line-of-sight	Potentially broader depending on geometry	Wide-area due to distributed units	Wide, but depends on antenna layout
Resilience to Interference	Lower	Moderate	High (spatial diversity)	Very high (waveform and spatial diversity)

Active radar systems can be further classified into four principal configurations based on the spatial arrangement of their transmitting and receiving elements: monostatic, bistatic, multistatic, and Multiple-Input Multiple-Output (MIMO) radar systems [83]. To facilitate comparison among these configurations, Table 2.3 summarised their key features, highlighting relative advantages and operational trade-offs.

Radar configuration plays a vital role in the design and performance of radar-based systems, particularly in applications such as HAR, where the quality and nature of the acquired signals have a direct influence on system effectiveness [85]. The key configuration parameters include the operating frequency band, pulse repetition frequency, and antenna geometry. These factors collectively determine the spatial resolution, signal penetration, and robustness of the radar [85]. Radar systems can be classified in several ways; however, one of the most widely used approaches is based on the nature of transmitted signals. Within this framework, radar technologies are typically divided into continuous wave (CW) radar and pulse radar.

- The CW radar operates by continuously transmitting an RF signal and analysing the Doppler shift in the reflected signal to estimate the target velocity. This type of radar is characterised by its simplicity, low power consumption, and high resolution in velocity

Table 2.4: Comparison of radar types based on the transmitted signal characteristics for non-invasive activity sensing.

Feature	**CW Radar**	**Conventional Pulse Radar**	**UWB Radar**	**FMCW Radar**
Signal Type	Unmodulated continuous wave	Pulsed with modulations	Ultra-wide band, ultra-short pulses	Frequency-modulated continuous
Range Measurement	No	Yes	Yes	Yes
Short Range Detection	Better	Moderate	Best (very high resolution)	Better
Velocity Measurement	Yes	Yes	No (typically not Doppler-based)	Yes
Complexity	Low	High	Moderate to High	High
Power Consumption	Low	High	Low (often low average power)	High

estimation, although it does not provide range information [86]. When the CW is modulated in frequency, typically using a linear or sinusoidal pattern, the system is referred to as frequency modulated CW (FMCW) radar [87], with which we can measure both range and velocity, respectively.

- In contrast, pulse radar transmits discrete bursts of RF energy, calculates the target range based on the time delay of the returning echo, and can also estimate velocity using Doppler analysis [88]. A specific form of pulse radar is ultra-wideband (UWB) radar, which transmits extremely short pulses with a wide bandwidth [89]. Due to its high temporal resolution, UWB radar is particularly effective in short-range applications and is well suited for tasks such as micro-motion detection and through-wall sensing.

A detailed comparison of these radar types, focusing on their relevance to macro- and micro-activity monitoring, is provided in Table 2.4. In this thesis, we employed an FMCW radar for both macro- and micro-activity monitoring, as it is particularly suitable for HAR research in indoor environments. The application of the FMCW radar for macro-activity recognition is discussed in chapter 3, and micro-activity monitoring is addressed in chapter 4. The frequency modulation capability of the FMCW radar enables representation of time-domain information in the frequency domain, allowing simultaneous estimation of range and velocity through efficient signal processing techniques such as the Fourier transform [90]. This approach reduces the computational burden while maintaining a high temporal and range resolution [87]. The high-resolution output generated by the FMCW radar proves particularly effective for ML-based HAR models, enhancing recognition accuracy while maintaining system efficiency. Addition-

ally, in chapter 4, we initially investigated micro-activity monitoring under controlled experimental conditions using UWB radar before transitioning to FMCW radar. Both radar systems offer advantages, including low power consumption and moderate hardware complexity, making them ideal for continuous monitoring applications in home or clinical environments.

Working Principle of FMCW Radar

Recent advances in computing and processing have made it possible to implement new modulation techniques that can estimate range and velocity information. FMCW radar is one such system that continuously transmits a frequency-modulated signal. A synoptic of FMCW radar system is illustrated in Fig. 2.4. The transmitted signal is reflected by the target and the radar receiver recognises the reflected signal. The range (distance) of the target was estimated by comparing the frequencies of the transmitted and received signals. The velocity (Doppler) information between radar and target is determined by the Doppler shift caused by human movement [91].

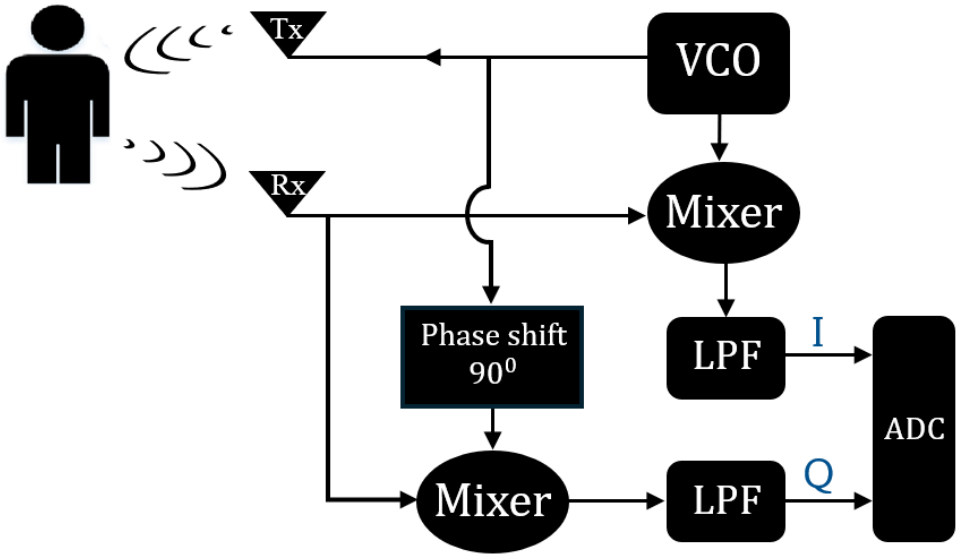


Figure 2.4: Block diagram of FMCW radar system.

The operation of an FMCW radar involves transmitting a CW signal modulated by a linear frequency sweep over time with a fixed duration, commonly referred to as a "chirp" signal, as shown in Fig. 2.5. A chirp signal is characterised by three key parameters [92]: center frequency (f_c), bandwidth (B) and duration of the chirp (T_c). In addition, it has a slope (α), which defines the rate of frequency change over time (B/T_c). These parameters can be configured to optimise the transmitted and received radar signals to satisfy specific application requirements. When f_c is used to sweep through the linear frequency chirp sequence, the instantaneous frequency at any

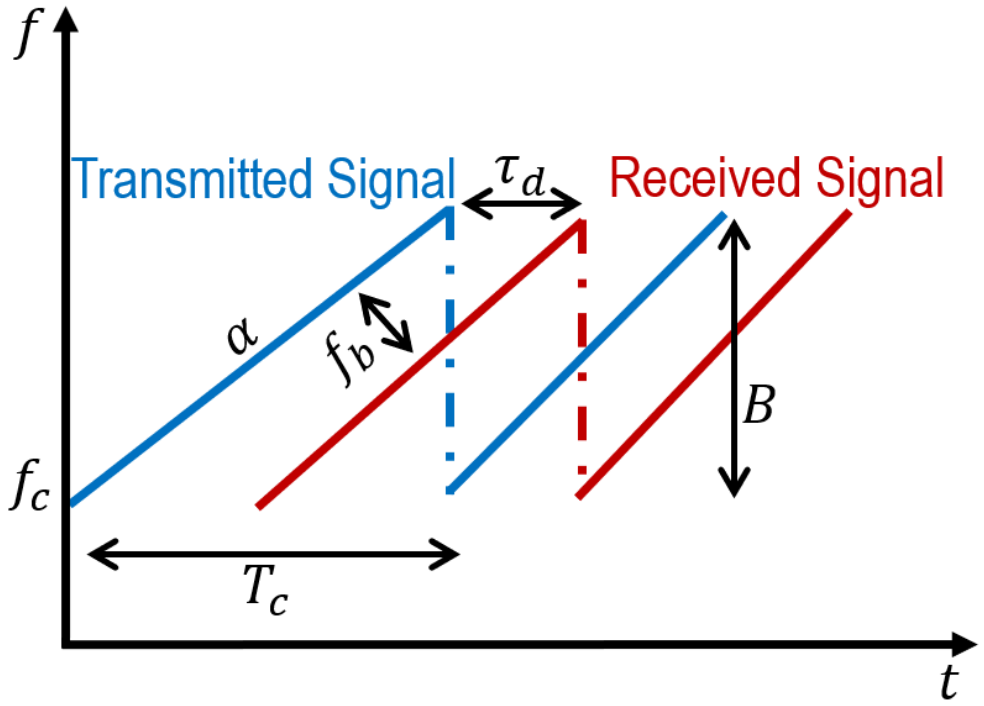


Figure 2.5: A sample of chirp pattern.

given time $f(t)$, can be expressed using the following Eq. (2.1) [84]:

$$f(t) = f_c + \frac{B}{T_c}t = f_c + \alpha t \tag{2.1}$$

The instantaneous angular frequency $\omega(t)$ that correspond to $f(t)$ by a factor of 2π and can be derived using Eq. (2.2):

$$\omega(t) = \frac{d\varphi(t)}{dt} = 2\pi f(t) \tag{2.2}$$

Here $\varphi(t)$ shows the instantaneous phase of the signal. Substitute Eq. (2.1) into Eq. (2.2), we can define $\omega(t)$ as:

$$\omega(t) = 2\pi(f_c + \alpha t) \tag{2.3}$$

Thus, we can acquire the equation of $\varphi(t)$ as:

$$\varphi(t) = \int_0^t 2\pi f(t)dt = 2\pi\left(f_c t + \frac{\alpha}{2}t^2\right) + \varphi_0 \tag{2.4}$$

Where φ_0 denotes the initial phase at $t = 0$. The transmitted signal from radar T_x antenna now expressed as:

$$s(t) = Acos(\varphi(t)) = Acos\left(2\pi\left(f_c t + \frac{\alpha}{2}t^2\right) + \varphi_0\right) \tag{2.5}$$

Here A shows the amplitude of T_x signal, which is a constant term and related to signal power. In the signal processing stage [84], the frequency difference between the transmitted and received

signals is called the beat frequency (f_b), as shown in Fig. 2.5. This frequency difference is proportional to the time delay (τ_d) between the radar and target, allowing the range (R) information to be extracted directly through the f_b and frequency offset analysis, as expressed in Eq. (2.6) and (2.8)

$$\frac{\tau_d}{T_c} = \frac{f_\mathrm{b}}{B} \tag{2.6}$$

From Eq. (2.6) we can define the beat frequency (f_b) as:

$$f_d = \frac{B\tau_d}{T_c} \tag{2.7}$$

$$\mathrm{R} = \frac{\tau_d c}{2} \tag{2.8}$$

Where c is the speed of light. For a target at a range R, the τ_d of received signal can be define as:

$$\tau_d = \frac{2R}{c} \tag{2.9}$$

According to Eq. (2.5), the received signal $r(t)$ can be expressed as follows:

$$r(t) = B\cos\left(2\pi\left(f_c(t-\tau_d) + \frac{\alpha}{2}(t-\tau_d)^2\right) + \varphi_0\right) \tag{2.10}$$

Where B denotes the amplitude of R_x signal, which is also a constant term and related to received signal power. Assuming that both the transmitted and received chirp signals have normalised amplitudes, excluding the initial phase, we derive two general expressions for the transmitted signal $s(t)$ and received signal $r(t)$ [84]:

$$s(t) = cos\left(2\pi(f_c t + \frac{\alpha}{2}t^2)\right) \tag{2.11}$$

$$r(t) = \cos\left(2\pi\left(f_c(t-\tau_d) + \frac{\alpha}{2}(t-\tau_d)^2\right)\right) \tag{2.12}$$

The demodulator mixes the transmitted and received signals to generate I/Q components. Therefore, as shown in Fig. 2.4, the output of the in-phase mixer or component I is the product of two functions, which can be written as follows:

$$I(t) = s(t)r(t) = \cos\left(2\pi\left(f_c t + \frac{\alpha}{2}t^2\right)\right)\cos\left(2\pi\left(f_c(t-\tau) + \frac{\alpha}{2}(t-\tau)^2\right)\right) \tag{2.13}$$

By employing trigonometric equation $cos(x).cos(y) = \frac{1}{2}(cos(x+y)+cos(x-y))$, the Eq. (2.13) can be rewritten as:

$$I(t) = \frac{1}{2}\left(\cos\left(2\pi t\left(2f_c - \alpha\tau\right) + 2\pi\alpha t^2 - 2\pi f_c\tau\right) + \cos\left(\pi\alpha\tau^2 - 2\pi\alpha t\tau - 2\pi f_c\tau\right)\right) \tag{2.14}$$

From Fig. 2.4, a low-pass filter (LPF) is used to eliminate high-frequency components after the

mixing stage. The final equation for the I signal is given as:

$$I(t) = \frac{1}{2}\cos\left(\pi\alpha\tau^2 - 2\pi\alpha t\tau - 2\pi f_c\tau\right) \tag{2.15}$$

On the other hand, the Q component is generated by mixing a 90°-shifted replica of the transmitted signal, as shown in Fig. 2.4, with the received signal, as shown in Eq. (2.16):

$$Q(t) = \frac{1}{2}\left(\sin\left(2\pi t\left(2f_c - \alpha\tau\right) + 2\pi\alpha t^2 - 2\pi f_c\tau\right) + \sin\left(\pi\alpha\tau^2 - 2\pi\alpha t\tau - 2\pi f_c\tau\right)\right) \tag{2.16}$$

After using the LPF again to remove high-frequency components, the final Q signal is expressed as:

$$Q(t) = -\frac{1}{2}\sin\left(\pi\alpha\tau^2 - 2\pi\alpha t\tau - 2\pi f_c\tau\right) \tag{2.17}$$

The capability of I/Q components to represent radar signals in a complex form is one of their most important advantages. Therefore, the general FMCW radar signal $s[n]$ can be expressed in complex form as Eq. (2.18), and the phase $\varphi(t)$ can be expressed in Eq. (2.19):

$$s[n] = I[n] + jQ[n] = e^{-\pi\alpha\tau^2 + 2\pi\alpha t\tau + 2\pi f_c\tau} \tag{2.18}$$

$$\varphi(t) = -\pi\alpha\tau^2 + 2\pi\alpha t\tau + 2\pi f_c\tau \tag{2.19}$$

Although the f_b is related to the phase derivative of the mixed signal, it is typically calculated using frequency domain analysis. The R (range) can be expressed as a function of τ_d, so that the angular frequency ω_{beat}, f_b, and the estimated range R_e can be expressed in Eq. (2.20) and in Eq. (2.22):

$$\omega_{beat} = \frac{d(-\pi\alpha\tau^2 + 2\pi\alpha t\tau + 2\pi f_c\tau)}{dt} = 2\pi\alpha\tau \tag{2.20}$$

$$f_{beat} = \frac{\omega_{beat}}{2\pi} = \alpha\tau \tag{2.21}$$

$$R_e = \frac{c f_{beat}}{2\alpha} \tag{2.22}$$

In the previous equations, the focus was on a simple, ideal scenario: a single chirp interacting with a stationary target. In a real-world scenario, an FMCW radar typically transmits a series of continuous chirps as the target moves back and forth from the radar at radial velocity v. Consequently, the time delay τ_r can be restated as:

$$\tau_r = \frac{2\left(R + v(t_s + nT)\right)}{c} = \frac{2R + 2v(t_s + nT)}{c} \tag{2.23}$$

Here n shows number of chirps that we analysed, t_s is the start time to n^{th} chirp. By substituting

τ with τ_r in Eq. (2.21), the f'_{beat} can be obtained as:

$$f'_{\text{beat}} = \frac{(2R\alpha + 2vf_c + 2nBv)}{c} = \frac{2R\alpha}{c} + \frac{2vf_c}{c} + \frac{2nBv}{c} \tag{2.24}$$

The beat frequency f'_{beat} includes not only the range information as shown in Eq. (2.24), but also the Doppler shift produced by the radial velocity of the target. The velocity produces an additional phase shift that appears as an additional bound on f'_{beat}. Since $c \gg 2nBv$, this term can usually be ignored in normal cases. However, in cases involving a large number of chirp signals meaning that the transmitting time is long, this limitation may become important and should not be ignored.

2.2 Radar Signal Representations for Human Activity Recognition

R-HARS rely heavily on how raw radar signals are transformed and represented before being fed into ML models for classification. These transformed views, referred to as radar data domains or signal representations, determine how effectively spatial, temporal, and motion-related patterns are extracted. Building on the FMCW radar principles, the signal processing pipeline typically begins with complex baseband I/Q signals derived from the IF signal. These are processed to extract features such as range, Doppler, and mD signatures, which form the basis for various signal representations. In the context of FMCW radar, common domains include spatial-frequency profiles such as the range-fast Fourier transform (R-FFT) and time-frequency (TF) maps based on the short-time Fourier transform (STFT) [90]. Each representation captures different aspects of human macro-motion and directly affects the model's ability to learn discriminative features.

Despite the availability of diverse transformation techniques, most existing R-HAR studies adopt a single fixed representation [23, 93], typically STFT, without assessing how domain choice influences generalisation across users, settings, or radar devices. However, real-world radar signals vary significantly due to factors such as body morphology, movement style, clothing, sensor placement, and environmental interference. These variations result in non-IID data, making it challenging for models trained under one condition to perform reliably in others [94]. In this context, the choice of signal representation is critical for achieving adaptive and generalisable HAR. However, limited insight exists into which radar domain representations are most effective under diverse or unseen conditions.

2.2.1 Range-FFT Processing for Radar-based Human Activity Recognition System

An important component of FMCW radar systems, particularly for HAR applications, is the generation of a range profile. The range profile provides a quick overview of the surrounding environment by showing the presence of targets at different distances from the radar [95]. Similar to the previous formulations, the Fourier transform can be used to convert a beat-note signal from the time domain to the frequency domain in order to obtain its range information. By substituting Eq. (2.24) into Eq. (2.19), one can get the radar signal mathematical formulations by neglecting $\frac{2nBv}{c}$, discussed before, so it is not considered:

$$\varphi_R(t) = \frac{4\pi f_c v}{c} nT + 2\pi \left(\frac{2v f_c}{c} + \alpha\tau \right) \tag{2.25}$$

As from the Eq.(2.18), we can write as:

$$s(t) = e^{-\varphi_R(t)} \tag{2.26}$$

It is noteworthy that the Fourier transform is traditionally defined as a continuous integral over infinite time. However, computing continuous integrals is impractical and the necessity of an infinite duration cannot be achieved in practical applications. To address these limitations, it is important to apply a discrete Fourier transform (DFT) [95].

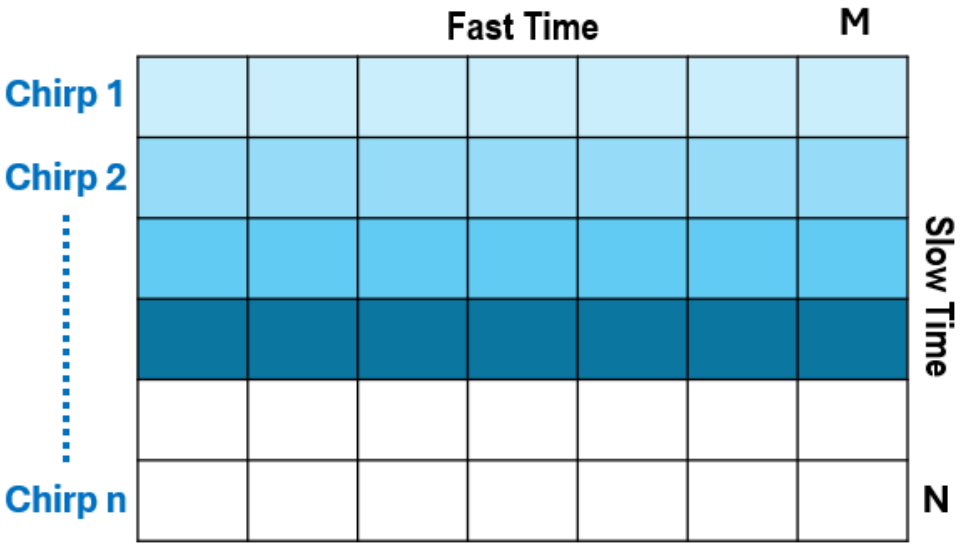


Figure 2.6: FMCW radar chirp information.

The radar beat frequency signal was sampled over time, creating a structured data matrix, as illustrated in Fig. 2.6. This matrix has dimensions $N \times M$, where: N represents the number of successive chirps (slow time) and M corresponds to the number of samples per chirp (fast time). Each row in the matrix corresponds to a single radar chirp and each column represents a range

bin. The sampling process creates a structured grid, where the radar captures the range and time information of the signal for each transmitted chirp. The number of bins, M, is determined by the sampling rate f_s and the chirp duration T_c as shown in Fig. 2.5, and expressed as:

$$M = f_s T_c. \tag{2.27}$$

To extract range information, an FFT is applied to the beat frequency signal $s(p,q)$ along the fast-time dimension (chirp-wise), as illustrated in Fig. 2.6. The transformed signal $S(p,q)$ in the frequency domain is given by:

$$S(p,q) = \sum_{q=1}^{M} s(p,q) e^{-j\frac{2\pi}{M}qt}, \tag{2.28}$$

where p represents the chirp index (slow time), q corresponds to the range bin index in the frequency domain, and $S(p,q)$ encapsulates the range profile information of the target. The frequency profiling stage, obtained by applying FFT along the fast time dimension, serves as a fundamental component in many R-HARS. This range profile identifies the spatial position of the subject, which is often used to define a region of interest (ROI) for further analysis [96]. In more advanced pipelines, such as TF processing, STFT is applied specifically to the range bins corresponding to the detected target location. This targeted application helps reduce background clutter and computational costs while focusing on human motion dynamics.

However, despite its central role in radar signal processing, the use of R-FFT as a standalone domain representation for macro-activity recognition has received limited attention. Most R-HAR studies have focused almost exclusively on STFT spectrograms, overlooking the potential of range profiles to convey motion-related energy patterns that can serve as informative input features [97]. This highlights a critical gap in the literature regarding the independent evaluation of radar frequency profiling as a valid and potentially complementary domain. We address this gap by systematically evaluating R-FFT maps alongside other domain representations and assessing their ability to capture discriminative features under data diversity conditions. This includes their integration into ML models for macro-activity classification, as discussed in chapter 3, as well as their application for precise chest region localisation in micro-activity estimation, as explored in chapter 4.

2.2.2 TF-Domain Analysis in Radar-based Human Activity Recognition System

FMCW radar systems leverage the Doppler effect to detect macro- and micro-motion patterns, enabling the identification of human activities ranging from dynamic actions such as running to subtle events such as falls [98]. The Doppler effect is a phenomenon in which radar signals reflected from a moving target experience a frequency shift, providing valuable information re-

garding the velocity and direction of the target. By analysing these velocity data, radar systems can effectively monitor human motion and facilitate HAR techniques [99]. For R-HAR, TF representations play a crucial role in capturing the non-stationary characteristics of human motion. Among these, STFT [90] is widely used as a robust method for analysing such signals. Unlike traditional fourier transform, which operates purely in the time or frequency domain, STFT divides the signal into segments and applies the transform to each, yielding a localised analysis of frequency content over time. The STFT representation for the signal $s(n)$, as derived in Eq. (2.18), is mathematically expressed as [90]:

$$\text{STFT}(n,\omega) = \sum_{m=-\infty}^{\infty} s(n)h(n-mR)e^{-j\omega n} \tag{2.29}$$

For $n = 0,....,N-1$, N represents the total number of time samples, $h(n)$ is the window function applied to each segment, R is the hop size or step size in samples, which controls the overlap between adjacent windows, and m is the length of the window. A spectrogram is a widely utilised technique for visualising the time-varying spectral density of a radar signal. It serves as a spectro-temporal representation that captures the dynamic changes in the Doppler frequency components of a signal over time [100]. The spectrogram is computed using STFT and is represented by the squared magnitude of the STFT, excluding the phase information of the signal, mathematically expressed as,

$$\text{Spectrogram}(n,\omega) = |\,\text{STFT}(n,\omega)|.^2 \tag{2.30}$$

STFT is regarded as one of the most computationally efficient techniques for generating TF spectral representations and remains the dominant method in R-HAR research. Its ability to highlight time-localised Doppler shifts makes it particularly effective for capturing dynamic motion events. However, STFT operates with a fixed window size, leading to an inherent trade-off between time and frequency resolution. This limitation constrains its ability to accurately resolve fine-grained or overlapping micro-motions, which are especially relevant in complex activity scenarios [100]. Despite its widespread adoption, most R-HAR studies apply STFT as a default representation without critically evaluating its robustness under data diversity or its ability to generalise across subjects, sensor placements, or environments. Furthermore, there has been little effort to benchmark STFT against alternative signal representations within a unified experimental framework [101]. This study presents two key gaps in the literature: First, comparative studies evaluating multiple radar domain representations under consistent preprocessing and evaluation conditions are lacking. Second, existing methods largely ignore the challenge of domain adaptation under non-IID data, such as subject-to-subject or frequency-to-frequency generalisation. These limitations restrict the development of scalable adaptive HARS suitable for deployment in real-world unconstrained settings.

To address these challenges, this thesis proposes a unified radar signal processing and repre-

sentation framework that systematically evaluates multiple-domain representations derived from a single radar signal. In addition to STFT, this includes the implementation of the Smoothed Pseudo Wigner-Ville distribution (SPWVD), a subset of the Wigner-Ville distribution (WVD) [102], which provides high-resolution quadratic TF features with reduced cross-term interference. While STFT remains the most commonly adopted method in prior R-HAR literature, its performance under non-IID conditions and in comparison to other domain representations like R-FFT has not been extensively studied. This thesis addresses that gap by experimentally evaluating the generalisability of each domain representation under cross-subject and cross-frequency conditions. The framework measures trade-offs between recognition accuracy, computational efficiency, and robustness. By integrating these representations into AI-driven models, this thesis contributes new insights into domain adaptation, signal encoding, and feature generalisation for scalable and real-world R-HARS.

2.3 AI-Driven Techniques in Human Activity Recognition

As illustrated in Fig. 2.1 (stage 3), represents a critical phase in R-HARS, where the processed sensor data are fed into AI models to recognise and classify human activities. Importantly, the ability to handle data diversity and perform domain adaptation, particularly under non-IID conditions, significantly depends on the capabilities of these AI-driven models. Without robust AI methods, it becomes difficult to extract generalisable features from varied domain representations or to adapt models across subjects, scenes, or sensing conditions. The integration of AI has been transformative, significantly expanding the capabilities of HAR by enabling more accurate and efficient activity recognition. This synergy between AI and radar-based human activity not only enhances system performance but also broadens its application across diverse domains [18]. However, the advent of advanced AI methodologies, particularly DL, has enabled HARS to process intricate motion patterns and handle diverse datasets effectively [103]. As a result, AI continues to drive the evolution of R-HARS, serving as the core mechanism through which representational diversity is harnessed and generalisation challenges are addressed.

AI-driven HAR techniques can be broadly categorised into two approaches [104]: conventional ML and DL based methods. Conventional ML techniques are built on well-established theoretical frameworks, making them highly interpretable and computationally efficient. In contrast, DL approaches utilise large datasets and sophisticated neural network architectures to uncover complex patterns, thereby offering superior performance in recognising intricate activities.

2.3.1 Machine Learning for Human Activity Recognition

ML is a subfield of AI that focuses on developing algorithms and statistical models to enable computers to perform tasks without explicit programming [105]. The primary aim of ML is to allow machines to learn from data to make accurate predictions or decisions. In HAR, sensors

like radar collect raw signals that are then processed into feature representations used to train ML algorithms for activity classification, as illustrated in Figures 2.1 and 2.3. A number of conventional ML algorithms, such as Decision Trees [106], K-Nearest Neighbor (KNN) [107], and Support Vector Machines (SVM) [107], have been applied to R-HAR data for activity classification. These methods rely on handcrafted features, which are extracted through a manual process based on human visual judgment and domain expert knowledge rather than being automatically learned by the system [104]. R-HAR has widely used conventional ML algorithms, but their robustness and generalisation capabilities suffer from several limitation:

- Conventional ML-based classification methods depend heavily on heuristic and manual feature extraction, which are time consuming and require significant human expertise.
- Manual features often capture only low-level statistical information, such as the mean, variance, frequency, and amplitude, which have weak transferability and generalisation [108]. Consequently, applying a trained ML model to a new scenario or dataset often results in significant performance degradation.
- Achieving classification using conventional ML techniques involves a complex, multi-step pipeline, including preprocessing, feature extraction, feature selection, learning, and classification [108], as shown in Fig. 2.7. Each of these steps requires careful tuning and domain expertise. For instance, feature selection has a significant impact on the performance of ML models, and biased or suboptimal feature selection can lead to incorrect discrimination between classes [109]. This multi-step process not only increases the complexity of the system but also introduces potential sources of error at each stage, making ML methods less robust and scalable for real-time R-HARS.

In contrast, DL enables learning and classification to be achieved in a single shot, as shown in Fig. 2.7. DL models excel in R-HAR due to their ability to automatically learn hierarchical features directly from raw radar data, unlike conventional ML methods [109]. This eliminates the need for manual feature extraction and reduces dependency on human intervention and domain expertise. Furthermore, the emergence of graphics processing units (GPUs) has enabled DL algorithms to leverage massive datasets and perform fast data processing and computation through parallel computing [110]. This significantly accelerated the training and deployment of DL models, making them more practical for real-world applications.

Given these advantages, conventional ML models are beyond the scope of the R-HAR in this thesis. Instead, we focused on DL methods, which are discussed in detail in the following subsections.

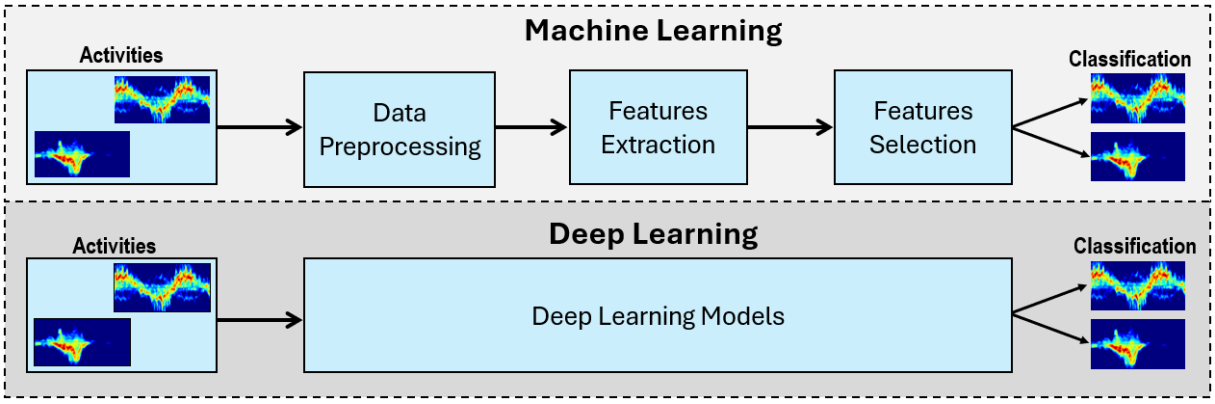


Figure 2.7: Comparison between conventional ML and DL algorithm pipelines.

2.3.2 Deep Learning for Human Activity Recognition

In recent years, DL has emerged as an exceptionally popular subfield of AI-driven by the rapid expansion and evolution of big data [111]. Its ability to process large-scale data and uncover intricate patterns has enabled significant progress across various domains, including image super-resolution [112], object detection [113], and image recognition [114]. The architecture of DL models typically consist of three primary components: an input layer, hidden layers, and an output layer [111]. The input layer receives training data, which are processed through a series of hidden layers before generating predictions at the output layer. The term "deep learning" stems from the depth of these hidden layers, which can range from a few to hundreds or even thousands depending on the complexity of the model and task [115]. Each hidden layer performs mathematical operations such as multiplying input values by learned weights, followed by the application of activation functions to extract and learn patterns within the data [116].

The network's predictions were compared to the ground truth labels to calculate the loss using categorical cross-entropy, which is appropriate for the multi-classification task of R-HAR. This loss is then used to adjust the network weights through backpropagation [109]. The backpropagation algorithm iteratively fine-tunes the weights by propagating the error backward through layers [111]. As this process is repeated, small adjustments to the weight values refine the network's ability to identify and learn patterns in the data, thereby improving its predictive accuracy [115]. This iterative optimisation is performed over multiple cycles, known as "epochs," during which the network becomes progressively better at capturing complex relationships in the data [116]. By leveraging these principles, DL models can achieve a high performance across a wide range of applications, including R-HAR.

Radar echo signals carry rich information about the targets in the environment, including the range and Doppler frequencies. A critical challenge in R-HAR is the design of DL based signal processing algorithms to effectively extract target related information from these echoes. This thesis focuses on exploring and applying DL approaches to radar-based macro-activity signals, specifically leveraging the dimensions of radar returns. Radar signals can be transformed into

two dimensional (2D) representations, same as camera-based images. Consequently, compared to other 2D radar echoes, TF maps are most commonly used for macro-activity recognition [93]. In this work, we primarily focused on 2D radar maps, generating four distinct types of representations, as discussed in Section 2.2, and further elaborated in chapter 3.

To fully exploit information within radar echoes, DL methods must be tailored to the unique characteristics of these representations. Although 3D human backscattering echoes provide extensive activity-related data, their complexity makes them challenging to process [22]. In contrast, 2D radar echoes strike a balance, offering sufficient activity information, while being more computationally manageable [22]. These 2D echoes are often treated as images, aligned with methodologies in computer vision, and used for activity classification tasks. Consequently, HAR based on 2D radar echoes often resembles an image classification problem, where DL architectures, particularly those rooted in vision-based models, excel in extracting meaningful patterns for accurate activity recognition. For HAR, DL techniques can be broadly categorised into three main types: deep neural networks (DNNs) [117], hybrid DL (HDL) models [118], and transfer learning (TL) based models [119].

DNNs include models such as convolutional neural networks (CNNs) [120], recurrent neural networks (RNNs) [121], and RNN variants such as long short-term memory (LSTM) and gated recurrent units (GRUs) [122]. Hybrid HAR models integrate CNN and RNN architectures to effectively process spatio-temporal data. Notable hybrid models proposed recently include DeepSense [123] and DeepConvLSTM [124]. These approaches have demonstrated significant advancements in HAR by combining the spatial feature extraction capabilities of CNNs with the temporal analysis strengths of RNNs. The various DL-based models applied to radar data domains are summarised in Table 2.5. Each row presents one representative example from the literature, however several other studies also exist for each radar type and configuration. Although CNNs, RNNs, and their hybrid variants have significantly advanced R-HAR by effectively utilising spatial and temporal information, these methods often require large, well annotated datasets and computational resources for training. This presents challenges in healthcare applications, where datasets can be scarce and computational efficiency is crucial. To address these limitations, TL has emerged as a promising alternative, enabling the reuse of knowledge from pre-trained models for R-HAR tasks.

Transfer Learning (TL) Approach

DL is highly data dependent and requires large-scale labelled datasets to achieve optimal performance [137]. However, acquiring such datasets for macro-activity monitoring is challenging, given the difficulties in collecting and annotating radar data. TL offers a powerful solution by enabling models to reuse the knowledge learned from one task and adapt it to another, thus reducing data dependency and improving generalisation [138]. Instead of training a DNN from scratch, TL allows models to leverage pre-trained feature representations from large scale

Table 2.5: Summary of existing R-HAR methods utilising DL models.

Radar Type	Carrier Freq.	Radar Domains	Classifier/Models	Example Ref.
CW radar	4 GHz	time-Doppler maps	CNN	[125]
	24 GHz		CNN	[93]
	8 GHz		LSTM	[122]
Doppler radar	2.4 GHz		CNN	[126]
	24 GHz		1D CNN+LSTM	[127]
	5.8 GHz		CNN	[23]
	25 GHz		LSTM	[128]
UWB radar	4 GHz		CNN	[129]
	4.3 GHz		CNN	[130]
	7.25 GHz		CNN	[131]
	4 GHz		CNN	[132]
FMCW radar	24 GHz		CNN	[133]
	60 GHz		CNN+LSTM	[134]
	77 GHz		1D CNN+LSTM	[29]
	5.8 GHz		CNN	[120]
UWB radar	3.9 GHz	time-range maps	CNN	[135]
FMCW radar	24 GHz		3D CNN+LSTM	[136]

datasets, such as ImageNet, which has 1.2 million images distributed over 1000 classes, or from simulated radar data [129]. This significantly accelerates training convergence and improves model performance, even with limited radar-based datasets.

The TL mechanism involves training a deep CNN model on a large dataset, thereby allowing the network to learn generalisable feature representations. During training, the model optimises the key parameters, including weights and biases, to improve classification accuracy [137]. In the next phase, these learned parameters are transferred to a new network that is fine-tuned using a smaller radar dataset. This transfer of knowledge enables the new model to use pre-trained weights, eliminating the need to train from scratch, while accelerating convergence and enhancing accuracy [138].

Given that radar signals can be converted into 2D dimension, TL has been successfully applied by adapting CNNs pre-trained on optical image data to analyse radar TF maps, as demonstrated in recent studies [139]. In R-HAR task, TL is applied using two primary approaches: transferring models trained on large scale natural image datasets [140], and transferring of models trained on simulated radar datasets to real-world HAR applications [141]. To implement TL effectively, two techniques are commonly employed [138]:

- **Feature Extraction:** The last layers of the pre-trained network were removed and replaced with new task-specific layers, which were then randomly initialised and trained on

the target dataset.

- **Fine-tuning (FT):** A pre-trained model is trained further using a smaller radar dataset, optimising the final layers while keeping the earlier layers frozen to retain the previously learned features.

By FT CNNs pre-trained on large-scale datasets, radar-based DL models benefit from faster convergence, decrease computational cost, and reduced overfitting, even with limited training data [142]. For instance, one study introduced a DCNN model pre-trained on ImageNet, and subsequently FT the network using measured radar TF maps for human aquatic activity classification [141]. Similarly, another study [143] proposed a residual learning-based model, DivNet, that was initially trained on a simulated radar spectrogram dataset and later FT with real measured radar data to classify seven distinct human activities. Additionally, a convolutional autoencoder (CAE) model [144] was first pre-trained in an unsupervised manner and then FT using a limited number of labelled radar maps.

The FT strategy adopted in these methods leverages a small amount of target data to adjust pre-trained DL models, thereby transferring knowledge from the source domain to mitigate the limitations of target-domain data insufficiency. Various pre-trained DL architectures, including ResNet-50 [145], MobileNet-v2 [146], VGGNet [147], have been extensively employed in TL. These architectures have been effectively FT for R-HAR, capitalising on their robust feature-extraction capabilities. In this thesis, TL plays a central role in recognition of macro-activities, enabling the integration of pre-trained CNN models to effectively train on limited radar datasets. By leveraging TL, we enhance classification accuracy while ensuring that the model adapts well to diverse scenarios. The detailed implementation and fine-tuning strategies adopted in this thesis are discussed in chapters 3 and 5.

2.3.3 Limitations and Research Gap

Despite the widespread used of TL in HAR applications, most TL-based research has been predominantly applied to wearable sensors [64], vision-based approaches [58], and Wi-Fi-based HARS [68]. In wearable sensor-based HAR, TL has facilitated cross-subject and cross-device adaptation, enabling models trained on one user group to effectively generalise to others. In vision-based HAR, TL has been extensively utilised to FT pre-trained CNNs, such as ResNet, VGG, and Inception, achieving notable improvements in accuracy across diverse datasets [148]. Similarly, Wi-Fi-based HAR has benefited from domain adaptation techniques, such as Decision-boundary Iterative Refinement Training with a Teacher (DIRT-T) and Virtual Adversarial Domain Adaptation (VADA) models, to handle cross-environment variability [149]. However, R-HAR remains largely under explored in the context of TL, as highlighted by Ray and Kolekar [150]. This lack of focused research presents a major gap in this field. To address this challenge,

our work systematically evaluates TL models across multiple radar domains, demonstrating the feasibility and effectiveness of radar-based macro-activity recognition.

One significant gap in R-HAR is the lack of a systematic evaluation of radar representations for TL. Radar data can be represented in various forms, unlike traditional image or wearable sensor data, radar signals encode complex motion-related features that require specialised processing techniques. However, current research has not adequately investigated the impact of different radar domains on TL effectiveness. In contrast, vision-based HAR studies have shown that different modalities such as RGB, optical flow, and depth can significantly influence TL performance [151]. A similar comparative analysis is notably absent in R-HAR, which represents a critical gap in optimising TL-based models for radar domains.

Recent studies have shown that TL can effectively address the challenges in R-HAR, such as limited labeled datasets and computational inefficiency. Models such as AlexNet, ResNet, VGG, GoogLeNet, and DenseNet have been adapted for TF-based spectrograms classification, achieving high accuracy while reducing the training complexity [152]. Some studies have further enhanced the TL performance through domain adaptation, adversarial learning, and data augmentation, making models more robust in diverse environments [153].

Although previous studies have primarily focused on TF-based preprocessing with advanced CNN models in TL settings, as illustrated in Table 2.5, this work presents a comprehensive evaluation of multiple radar domain representations and their integration with state-of-the-art TL models. By exploring the strengths and limitations of each method, we aim to identify radar preprocessing techniques that optimise both recognition accuracy and computational efficiency. To achieve this, we compare the performance of various TL architectures, including DenseNet-201 [154], ResNet-34 [155], VGG-16 [154], VGG-19 [154], and EfficientNet-B0 [156], across different radar domain representations. This thesis also investigates how TL models address the limitations of training CNNs from scratch, particularly in low-data scenarios, and their effectiveness in reducing misclassification rates, especially for fall detection in applications for elderly care.

Furthermore, there is limited comparative research between TL models and CNNs trained from scratch in R-HAR. DCNNs trained from scratch are resource-intensive and require large-scale labelled datasets and substantial computational power [150]. In other HAR domains like wearable and Wi-Fi-based systems, TL has proven to reduce training costs and enhance model robustness [157]. However, no comprehensive study has been conducted to quantify the benefits of TL over traditional CNNs for macro-activity recognition, specifically in terms of the following:

- Improvement in classification accuracy and reduction in computational cost.
- Enhanced training efficiency and faster convergence rates.
- Performance in unseen radar environments and real-time applications, particularly for

edge devices.

In summary, while TL presents promising opportunities for enhancing radar-based macro-activity recognition, significant gaps persist in the systematic evaluation of its effectiveness across various radar domains, optimisation of domain adaptation techniques, and its application to complex activity recognition. This thesis addresses these challenges by systematically evaluating TL models, benchmarking their performance against CNNs trained from scratch, and exploring innovative approaches to improving R-HAR capabilities, as detailed in chapter 3.

2.4 Micro-Activity Monitoring using Radar

Micro-activities such as heart rate (HR) and respiration rate (RR) are critical indicators of an individual's health and are routinely monitored in clinical environments to detect early signs of conditions such as congestive heart failure or cardiac arrest [158, 159]. These physiological parameters are highly dynamic, with patterns that vary considerably during sleep, physical exertion, or psychological stress. A healthy adult typically breathes 12–20 times per minute at rest, while the HR ranges between 60 and 100 beats per minute [160]. RR is commonly inferred from chest or abdominal movements during the respiratory cycle. Conventional measurement tools, such as chest straps or contact-based sensors, although effective, suffer from limitations related to user comfort, long-term wearability, and privacy. These limitations render them unsuitable for continuous and unobtrusive monitoring. Within the broader context of R-HARS, accurately capturing micro-activity signals in a non-intrusive manner remains a major challenge. Unlike macro-activities, micro-physiological signals involve extremely subtle motions that are often masked by environmental clutter or stronger overlapping movements. This thesis addresses this gap by exploring radar technologies capable of reliably extracting such signals without requiring direct contact or compromising user privacy.

Recent advancements in the Internet of Medical Things (IoMT) have accelerated the development of remote, non-contact sensing systems. Radar-based solutions have emerged as promising alternatives, offering unique benefits, including comfort, unobtrusiveness, privacy preservation, and real-time continuous monitoring. Radar sensors can estimate HR and RR by detecting minute chest wall displacements and micro-movements associated with cardiovascular and respiratory activity without any physical attachment to the body [25]. In Section 2.1.3, we discuss two radar types explored for micro-activity monitoring: UWB and mmWave FMCW radars. While the core principles of the FMCW radar were covered previously, additional technical details and implementation results are provided in chapter 4.

Since the Federal Communications Commission (FCC) approved UWB usage in 2002 [161], it has been widely studied for contactless RR monitoring owing to its ability to penetrate walls and deliver accurate time-domain measurements [162]. Among UWB-based solutions, the Walabot device has gained attention for its affordability and capability to detect respiratory motion.

While commonly used for through-wall imaging and construction applications, Walabot's potential for remote health monitoring has made it a notable candidate for low-cost, non-invasive RR detection. A detailed discussion of the use and adaptation of the Walabot system is provided in chapter 4.

2.4.1 Significant Studies and Findings in Radar-Based Micro-Activity Estimation

This section provides an overview of state-of-the-art literature on non-invasive micro-activity monitoring. The discussion highlights key studies, methods, and advances in this field, emphasizing their contributions and limitations. In one study, short-range radar technologies were used for contactless respiration, HR, and stress monitoring [163]. They found that while automotive radars can estimate vital signs, they are power hungry and optimised for long-range applications rather than healthcare. Studies also highlight the potential of frequency bands such as the 60 GHz and 122 GHz ISM bands for FMCW and UWB radar-based monitoring, while high-frequency CW refractometers (75–110 GHz) present challenges in interference and harmonics, necessitating mmWave FMCW radar and ML techniques for improved estimation [164].

Other studies, such as those by [165], emphasize the challenges of separating the RR from HR signals and reducing motion artifacts, which require advanced signal processing techniques. One study reviewed non-contact sensors, including microwave radar and Doppler-based systems, which enhance robustness and efficiency, but still struggle with motion artifacts [166]. FMCW radar and wireless sensing show promising results but highlight classification and interference challenges. Beyond radar, alternative technologies such as thermal imaging, lasers, and ultrasound have been studied, but they often suffer from sensitivity to movement or inefficiency in HR detection. Emerging methods, such as airflow-based acoustic sensing, provide novel approaches for sleep monitoring, but require further validation [167]. A comparative overview of micro-activity monitoring studies is presented in Table 2.6, highlighting key differences based on radar type, experimental setup, and limitations.

2.4.2 Research Gap

Although significant progress has been made in micro-activity monitoring using radars, there are still some research gaps. Most existing studies focus either on UWB or FMCW radars, with limited comparative analysis under standardised conditions [25, 169]. This thesis bridges this gap by using and comparing these two types of radar and taking advantage of their strengths in vital sign monitoring. Moreover, previous studies have mainly focused on sitting or lying positions and ignored standing positions, which are essential for practical applications, as illustrated in Table 2.6. This thesis addressed this limitation by incorporating experiments with standing subjects. Furthermore, this thesis provides a detailed comparison of signal processing tech-

Table 2.6: Comparison between contactless micro-activity monitoring systems based on radar type.

Ref.	Radar Type		Subject Position		Distance	Radar Position	Key Outcomes	Limitations
	UWB	FMCW	Sit	Lying				
[168]	✓	–	✓	✓	0.5-1.5 m	Towards chest, Beside the bed	Achieved high accuracy with an average MAE of 0.73 BrPM across three studies.	Sensitive to body movements, resolution/granularity, and irregular breathing.
[169]	✓	–	✓	–	1.7 m	3 radars at equilateral triangle vertices	Average HR accuracy of 94.3% across six body orientations.	Orientation dependence, static postures, setup complexity.
[170]	✓	–	✓	–	1.1 m	Positioned aimed towards upper body	CNN-based method achieved an average RMSE of 20.6 ms for HR.	Synchronization dependency, static scenario, fixed radar position.
[171]	✓	–	✓	–	0.5-3 m	Directly towards subject chest	Eigenvalue improved the SNR: +12 dB for RR and +18 dB for HR. HR achieved an error of 2.56% compared to reference.	Synchronization dependency, static scenario, noise removal dependency, fixed radar position.
[172]	✓	–	✓	–	0.6-1 m	In front of subject Positioned towards chest	Average HR detection accuracy of 89%, 100% accuracy for RR detection in 3 out of 4 subjects	Short Distances, Static scenario, fixed radar position
[173]	✓	–	✓	–	1-1.5 m	In front & back of human torso	Average HR detection accuracy of 98.5%, accurate 2D localization and vital sign monitoring	Short Distances, Static scenario, fixed radar position, Complex Setup
[174]	✓	–	✓	✓	20 cm	Front of bed positioned towards abdomen	RR: 95.02%, MAE of 0.2297 brpm	Short Distances, Static scenario, fixed radar position
[175]	–	✓	–	–	2-4 m	single radar in front of subject	RR: 12 and 16 Brpm, HR: 70 and 80 Bpm compared to ground truth	One scenario, one radar position resolution limits
[176]	–	✓	–	✓	2 m	Radar mounted on the ceiling above the bed	Achieved an average of 96.30% accuracy with 1.03 MAE for HR, 1.41 for RR	One radar position, only one distance, limited study group
[25]	–	✓	–	✓	1.7 m	Radar mounted on the ceiling above the bed	HR: Achieved an average of 80% accuracy, RR: 94% accuracy correlation with reference	One radar position, single setup, only one individual
[177]	–	✓	✓	–	1 m	In front of subject	RMSE: 1.5 Bpm for HR, and 1 Brpm for RR	One radar position, single setup, subject variability
[178]	–	✓	✓	–	1 m	In front of seated subject	Using CS-OMP and RA-DWT algorithms obtained 93% accuracy	One radar position, single setup, limited validation
[179]	–	✓	✓	–	30 cm	In front of seated subject towards chest	21.3 Brpm for RR and 72.6 Bpm for HR, with minimal distortion using DC offset calibration	One radar position, single setup, short distance
[180]	–	✓	✓	–	0.5-1.5 m	In front of seated subject towards chest	HR: 96% (0.5 m), 94% (1 m),92% (1.5 m) and for RR: 97% (0.5 m), 95% (1 m), 94% (1.5 m)	Single setup, controlled environment

niques for UWB and FMCW radar data, and compares these methods with reference sensors. Research progresses from traditional signal processing for UWB radar demonstration to more robust and high-data rate and real-time applications using mmWave FMCW radar, providing a comprehensive and scalable approach for non-invasive R-HARS.

2.5 Radar-based Human Activity Recognition for Edge Devices

In the previous sections, we discussed the rationale for using radar sensing and DL models for macro-motion classification. While DL has significantly improved classification accuracy, real-time inference remains a critical requirement for real-world applications. In scenarios such as elderly fall detection, industrial safety, and human-computer interaction, activities must be recognised within milliseconds (ms) to enable timely responses and prevent hazardous outcomes. Traditional cloud-based HARS, however, introduce challenges such as network latency, privacy concerns, and dependency on internet connectivity, making them unsuitable for time-sensitive healthcare applications [181].

To address these limitations, edge computing has emerged as a promising solution. By enabling on-device processing, edge computing reduces inference delays and supports fast, autonomous decision-making [181]. However, deploying DL models on low-power edge devices presents significant challenges, including computational constraints, memory limitations, and energy efficiency trade-offs. These challenges are particularly relevant in the context of the IoT, where the rapid expansion of connected devices has driven the need for efficient ML implementations on resource-constrained hardware. While many IoT-based HAR applications leverage ML models for intelligent decision-making, small embedded devices often lack the computational resources required to run complex DL models effectively.

A key advancement in this domain is TinyML [182], an innovative technology that enables the deployment of DL models on microcontrollers and other edge devices with stringent resource constraints. TinyML facilitates real-time data analysis and inference directly on the device, eliminating the need to transmit raw data to the cloud. This approach enhances privacy and efficiency by localising data processing, reducing communication costs, and minimizing security risks. The integration of edge computing with IoT has opened new possibilities for implementing intelligent HARS at the network edge [181]. Unlike traditional cloud-based ML, which relies on significant computational power and incurs network delays, edge-based ML processes data locally, thereby reducing latency and energy consumption. This is particularly advantageous for applications such as smart home automation, healthcare monitoring, and industrial automation, where real-time responses and uninterrupted operations are critical.

Several studies have explored the deployment of DL models for real-time R-HAR on edge devices, addressing challenges related to computational efficiency, energy consumption, and inference speed. One study proposed Mobile-RadarNet [183], which is a lightweight CNN optimised for low-power edge devices. Using depth-wise separable convolutions significantly reduces computational complexity while achieving 97.31% accuracy and 91% accuracy for unseen subjects. Deployed on a Hisilicon Kirin 710F ARM processor with TensorFlow Lite, it delivers inference 3× faster than 1D-CNNs and 23× faster than 2D-CNNs, with a 15× reduction

in parameters and 40× lower computational cost. These results confirmed that Mobile-RadarNet is an efficient and scalable solution for R-HAR on edge devices. Similarly, another study developed an edge deployable DNN model that processes range-doppler maps from FMCW radar signals in real time [184]. Their study demonstrated the feasibility of recognising an on-device five human activity, achieving 93.2% accuracy with an inference time of 2.95 seconds on a Raspberry Pi 4.

In addition to CNN-based models, hybrid architectures that combine CNNs and recurrent networks have demonstrated improved performance for sequential radar data. One study introduced tinyRadar [185], a system that integrates an LSTM network with CNN-based feature extraction to classify multi-target human activities in real time. This approach effectively captures the spatial and temporal dependencies in radar-based spectrograms, achieving 93% accuracy while operating within 10 ms per frame on a Raspberry Pi 4. Similarly, another study extended the tinyRadar framework for gesture recognition by leveraging hardware accelerators and quantised DL models to optimise the power consumption below 80 mW, making it ideal for HAR applications [186].

Despite these advancements, transitioning DL models from high-performance cloud environments or server-based training to low-power edge devices presents several challenges:

- Preserving classification accuracy while reducing model size and computational complexity.
- Optimising DL models for inference to meet real-time requirements on embedded radar devices.
- Minimising energy consumption to ensure efficient deployment in portable or battery-powered systems.

To overcome these challenges, researchers have focused on developing lightweight DL architectures or employing model compression techniques [187]. These methods aim to maintain classification accuracy while minimising model size, computational complexity, and energy consumption, thereby enabling efficient deployment on resource-constrained edge devices. Advancements in this area are crucial for ensuring that DL models can operate effectively in real-time applications. Addressing these challenges is a major contribution of this thesis, and is further explored in chapter 5.

2.5.1 Model Complexity in Deep Learning

The model complexity in R-HARS is determined by factors such as model size, number of parameters, and computational requirements, all of which directly impact the feasibility of deploying models on resource-constrained edge devices [188]. Traditional DL models have grown significantly in size over the years. For instance, AlexNet [189], introduced in 2012, contained

60 million parameters, whereas VGG-Net [147], published in 2013, expanded to 133 million parameters, achieving 71.1% top-1 accuracy. These models were developed for the ImageNet Large Scale Visual Recognition challenge (ILSVRC) [190], where the primary evaluation metric was the top-1 absolute accuracy, with little consideration of execution time or computational efficiency. This focus on accuracy alone led to neural network architectures with significant redundancy in their designs.

As DL models continue to evolve, modern architectures have reached unprecedented scales, with some models exceeding 175 billion parameters as of 2020 [191]. While such large-scale models are well suited for cloud-based data centers, they are impractical for real-time edge computing applications, where computational resources, power, and memory are severely constrained. To address these limitations, researchers have developed reduced-parameter models that maintain competitive accuracy, while being computationally efficient. For example, GoogLeNet [192] achieves an accuracy comparable to that of VGG-16 (69.78% top-1 accuracy), but with only seven million parameters. Similarly, MobileNet [183], designed specifically for mobile and edge computing, achieves 70% top-1 accuracy while containing just 4.2 million parameters and requiring only 1.14 Giga Floating-points (GFLOPs). These developments highlight the growing need for efficient model architectures that balance the accuracy and computational feasibility, particularly for edge-based HAR applications.

However, while MobileNet and similar lightweight models are optimised for mobile and edge computing [183], they become significantly larger and computationally expensive when trained on radar data representations, such as TF maps. These models consume more memory and energy during inference, making them less suitable for direct deployment on low-power edge devices. Addressing this challenge is the key focus of this thesis, as detailed in chapter 5. To address this issue, various model compression techniques have been explored to reduce the complexity of DL models, facilitating their efficient deployment on resource-constrained devices, while preserving the classification accuracy:

- **Model Pruning:** Model Pruning is a technique that reduces model complexity by removing redundant parameters, thereby minimizing memory usage and computational load. It achieves this by selectively eliminating weights or neurons that contribute minimally to the model predictions. Two primary types of pruning exist: Structured pruning removes all neurons, channels, and layers while preserving the overall architecture of the model, while Unstructured pruning removes individual weights based on their magnitude, resulting in a sparse model [187].

- **Knowledge Distillation:** Knowledge Distillation is a method used to compress DL models. It works by transferring knowledge from a large, complex model (called the "teacher") to a smaller, simpler model (called the "student"). This allows the student model to perform almost as well as the teacher model while using much less computational power and time. Unlike regular training, which uses only hard labels (like "class A" or "class B"),

knowledge distillation uses the teacher's "soft probabilities". These probabilities provide more detailed information regarding how the teacher model views the relationships between classes. This extra detail helps the student model to learn better and becomes more accurate and robust. However, achieving high accuracy when compressing the model is still difficult, particularly if knowledge distillation is used alone without other techniques [187].

- **Quantisation:** Quantisation is a popular technique for making DL models faster and requires less memory. It converts high-precision numbers (such as floating-point values) into simpler low-bit integer formats [193]. There are two types of quantisation: quantisation aware training (QAT) and post training quantisation (PTQ).
  - QAT includes quantisation during the Training process. This helps the model adapt to lower precision while it learns, thereby reducing the accuracy loss that usually occurs with quantisation. QAT is especially useful for models that require high precision, such as those that use less than 8-bit representations. However, QAT requires retraining of the model with simulated quantisation, which can be slow and computationally expensive [194].
  - PTQ applies quantisation after the model is already trained without changing the model's weights. This makes PTQ faster and easier to use, particularly for deploying models on edge devices or in real-time applications. Some models may lose their accuracy when moving from high-precision floating-point numbers to low-bit integers [194].

In this thesis, we used PTQ because it allows us to deploy models efficiently on edge devices without re-training. It also strikes a good balance between computational efficiency and accuracy. A detailed explanation of this approach and its implementation are provided in chapter 5.

2.5.2 Limitations and Research Gaps

R-HARS has shown great promise for real-time applications; however, several challenges remain, particularly in terms of energy efficiency, privacy preservation, and efficient edge deployment. Although DL models have been widely used for macro-activity recognition, they are often too computationally intensive for resource-constrained, battery-operated edge devices, discussed in Section 2.4. This section outlines the key limitations of the current study and identifies the gaps addressed in this thesis.

Most R-HAR studies rely on DL architectures, such as CNNs and hybrid CNN-LSTM models, which achieve high classification accuracy but require millions of parameters. This makes them impractical for real-time, low-power edge deployments. For instance, a CAE achieved

94.2% accuracy but was not implemented on edge devices [195]. Similarly, a compact 1D-CNN reached 95.8% accuracy but did not address real-time embedded inference [196]. More recent models, such as a CNN-LSTM hybrid (94.75% accuracy, 70M parameters) [127] and a deep CNN with inception blocks (96.1% accuracy, 365 K parameters) [197], demonstrate strong classification performance, but remain too large for efficient deployment. Finally, Table 2.7 summarises the contributions and limitations of related work.

Radar sensing offers inherent privacy advantages over camera-based systems, as it does not capture identifiable images, detailed in Section 2.2. However, deploying radar-based DL models on edge devices introduces new privacy concerns such as data security and user privacy. Current studies often overlook privacy-preserving techniques like differential privacy (DP) [198] or secure Federated Learning (FL) [199]. Moreover, the processing of radar data on edge devices without robust security mechanisms increases the risk of data leakage. To address these challenges, this thesis implemented local DP (LDP), ensuring secure and private radar data processing directly on edge devices. LDP minimises the risk of sensitive data exposure while maintaining model accuracy, making R-HARS a privacy-aware and practical solution for real-world deployment. In addition to privacy preservation, this thesis leverages PTQ to reduce the model size and computational complexity without sacrificing classification accuracy. The quantised models are deployed on edge devices, where their energy consumption and power efficiency are evaluated to ensure their suitability for battery-operated applications.

Despite the potential of combining radar signal representations with model compression techniques to enhance the efficiency and accuracy in real-time implementations, this area remains underexplored. By integrating PTQ and LDP, this thesis bridges these gaps, enabling an efficient, privacy-preserving, energy-optimised R-HAR framework tailored for edge devices.

2.6 Summary of Literature Review, Research Gaps, and Link with challenges

This chapter has examined the evolution of HARS, beginning with an overview of different sensor modalities, including vision-based, acoustic, inertial, and RF sensors, and progressing toward the emerging prominence of radar-based approaches. Radar technology has increasingly become a preferred modality owing to its non-intrusive, privacy-preserving nature and robustness under occlusion, lighting variations, and complex indoor environments. Compared with vision-based or wearable systems, radar offers the unique advantage of supporting both macro-activity recognition and micro-physiological monitoring within a single sensing framework. Despite these advantages, the literature reveals several limitations that align with the three core challenges identified in **Section 1.3**. Non-intrusive multi-scale monitoring **C1** remains a critical challenge in R-HARS, where the effective integration of macro-activity recognition and micro-physiological monitoring is significantly hindered by their divergent temporal-

Table 2.7: Comparative analysis of recent advances in HARS with contributions and limitations.

Ref	Contributions	Limitations
[200]	Introduces a lightweight DL and parameter estimation hybrid model for multi-directional HAR. TF maps were used for activity representation, achieving 96.67% accuracy with reduced complexity.	Lacks edge deployment, real-time inference. No quantisation or privacy techniques are applied. Focuses solely on TF maps from the target's range bin, ignoring the extended target features and boundary scenarios.
[201]	A lightweight framework for HAR in wearable devices using edge computing was developed. Evaluate DL models on a microcontroller, optimising with quantisation and model compression.	It uses wearable sensor data (not radar) and focuses on acceleration data. Lacks PTQ, LDP, and evaluation of metrics on edge device.
[202]	Developed a cloud-based HAR and gait monitoring system using radar. The employed RD maps and GRU-based DL for real-time classification achieved 93% accuracy in controlled settings.	Lack of standalone edge inference relying on cloud computing. No PTQ and LDP explored for real-time systems.
[203]	Developed TWR-FMSN, a lightweight multi-scale neural network for indoor HAR using through-the-wall radar, achieving 94% accuracy with 0.6s inference time.	Despite being lightweight, it lacks edge deployment and real-time inference on embedded hardware.
[204]	Fall detection system using FMCW radar with RD maps, a Bi-LSTM model, and a two-stage CNN-based detection process, achieving 96% accuracy.	Focuses on a single human activity class with only one radar representation.
[205]	Integrates the feasibility of quantizing radar-based HAR models for edge deployment with minimal accuracy loss for efficient, low-power real-time recognition.	Requires QAT for the best accuracy, but is time consuming. The accuracy drops with lower-bit quantisation. Scalability and real-world edge performance were not evaluated in this study.

spectral characteristics and environmental clutter. Despite radar advantages in contactless and privacy-preserving monitoring, existing systems predominantly focus on either macro- or micro-activity recognition, lacking unified frameworks that effectively bridge these scales for continuous health tracking. The coexistence of these divergent features complicates signal processing because strong macro-Doppler components often obscure the subtle variations required for micro-activity detection. Current approaches rely heavily on a single radar representation such as STFT-based maps [90], without systematically evaluating their suitability under overlapping

motion conditions. This limitation has restricted the exploration of alternative or complementary domains such as range-Doppler (RD) and SPWVD that might offer advantages in cluttered environments. Addressing these gaps requires more robust signal processing techniques, feature extraction strategies, and comprehensive evaluation frameworks that can effectively handle multi-scale monitoring under realistic conditions.

Data diversity and generalisation **C2** presents a fundamental bottleneck in R-HARS development, where non-IID radar data across subjects, device configurations, and ambient environments significantly degrades model generalisation in real-world, unconstrained settings [26]. Conventional ML methods, such as SVM and KNN, have shown limited success owing to their heavy dependence on manual feature engineering, resulting in brittle performance under dynamic contexts [107]. Although DL approaches, particularly CNNs, have improved classification accuracy, they remain prohibitively data-intensive in radar settings where large, labelled datasets are difficult to acquire. The TL approach has emerged as a promising solution to mitigate data scarcity; however, there is a critical lack of systematic comparative analysis on how different radar domain representations influence TL effectiveness [119]. This field suffers from insufficient evaluation across multiple radar signal representation types, carrier frequencies, and feature encodings, which has impeded the development of scalable R-HARS solutions. This research gap restricts progress toward building truly generalisable systems capable of adapting across domains, frequencies, and user populations without extensive retraining, ultimately limiting widespread practical deployment.

Energy efficiency and edge deployment **C3** present a substantial barrier to the practical implementation of R-HARS in resource-constrained environments. While DL models achieve high accuracy for macro-activity recognition, they impose prohibitive computational, memory constraint, and power demands that limit deployment on low-cost battery-operated platforms, such as IoT nodes, Raspberry Pi, or Jetson Nano. Recent advances in TinyML, model compression techniques, and efficient architectures such as MobileNet [183], offer promising pathways forward, yet their integration into radar-specific processing pipelines remains severely limited. Although radar systems provide inherent privacy advantages over vision-based alternatives, the on-device processing of sensitive behavioural and physiological data still necessitates robust privacy-preserving mechanisms such as LDP, which are rarely considered in current implementations. This critical gap in lightweight, quantised, and privacy-aware DL frameworks specifically designed for radar data significantly hampers the feasibility of deploying R-HARS at scale in real-world healthcare and AAL scenarios, despite their potential to transform continuous monitoring applications. Furthermore, the fundamental trade-off between model complexity and communication efficiency presents an ongoing challenge: complex models enhance recognition accuracy but correspondingly increase communication overhead, whereas simplified models may compromise performance. Consequently, innovative optimisation strategies are imperative to effectively balance these competing demands, particularly for resource-constrained edge de-

vices operating in dynamic monitoring environments.

In summary, addressing the core challenges in HAR requires innovative frameworks that effectively handle radar data diversity and domain adaptation without increasing computational overhead while simultaneously enhancing privacy preservation and optimising model complexity for deployment on resource-constrained devices. This thesis addresses these critical gaps through contributions that collectively advance the real-world viability of macro- and micro-activity monitoring systems by enhancing their accuracy, adaptability, efficiency, and privacy-preservation capabilities. The proposed solutions make R-HARS more practical for widespread implementation in healthcare, AAL settings, and intelligent environments without compromising performance or user privacy, ultimately bridging the gap between laboratory demonstrations and practical sustainable deployment in everyday settings.

Chapter 3

Macro-Activity Recognition Using FMCW Radar Signals

This chapter presents the first core contribution of this thesis, focusing on improving radar-based macro-activity recognition by addressing challenges **C1** and **C2**, outlined in **Section 1.3**. This study develops a framework that transforms a single radar signal into multiple domain representations, each capturing different aspects of range, velocity, and frequency information. This approach leverages the inherent diversity and complexity of radar data to enhance model stability and performance. By integrating transfer learning models, the framework aims to improve classification accuracy and generalisation across subjects under non-IID conditions. The evaluation uses subject-wise validation and performance metrics, such as classification accuracy, inference time, and preprocessing cost, to guide the selection of effective domain-model combinations for robust, non-intrusive activity monitoring.

3.1 Introduction

Radar-based macro-activity has become an important technology for ensuring human safety and well-being, especially in scenarios that require continuous monitoring and assessment [80]. One of the main advantages of radar systems is their ease of use, as they require little preparation or cooperation from the monitored person [206]. Radar sensors are non-intrusive devices that can penetrate opaque objects such as walls and furniture, making them effective even in obstructed environments [207].

The effectiveness of R-HARS depends heavily on preprocessing radar echoes into 2D signal representations such as range-time (RT), range-Doppler (RD), and time-frequency (TF) maps [204]. These representations are crucial for extracting meaningful features to improve recognition accuracy and reliability. Conventional preprocessing methods, such as applying range fast Fourier transform (R-FFT) to generate RT maps, provide a spatial representation of the target's movement over time and range. However, these RT maps lack Doppler information,

making them less effective in capturing fine motion details such as slight limb movements or Doppler changes. To overcome this limitation, we perform additional Doppler FFT on slow-time samples (across multiple chirps) after the R-FFT, resulting in RD maps. These RD maps encode both range and Doppler (velocity) information, providing richer insights into motion dynamics and enabling better discrimination of activities based on velocity profiles.

TF representation further enhances R-HAR by capturing the Doppler variations over time. In this regard, two widely used Doppler analysis techniques are linear and quadratic methods. Linear TF analysis, such as STFT, provides a clear and interpretable representation of the time-varying frequency content but is limited by its fixed window length, resulting in a trade-off between time and frequency resolution. On the other hand, quadratic methods, such as Wigner-Ville distribution (WVD) [102] and its subset SPWVD provide higher resolution in both time and frequency domains. While WVD provides powerful TF aggregation, it suffers from cross-term interference, making it less suitable for complex multi-component signals. SPWVD is an improved version of WVD that effectively suppresses cross-term interference while maintaining a high-resolution TF representation. In this study, both STFT and SPWVD are used to generate a TF representation, providing insight into the time-varying Doppler shift. While STFT provides a simpler and more computationally efficient representation, SPWVD can capture finer motion details but is computationally more expensive, potentially limiting real-time feasibility.

Furthermore, seamless integration of radar with ML models enables recognition, detection, and classification of human activities. R-HARS have transitioned from traditional ML to DL, significantly improving feature extraction and generalisation capabilities. CNNs, known for their hierarchical feature extraction capabilities, have become the foundation of modern HARS, effectively processing radar images to capture complex activity details that were previously difficult to extract manually [108]. While early studies explored traditional ML methods, recent advances have shifted to DL applications on radar datasets [103].

However, a major challenge facing R-HAR is the limited availability of labelled training data, which can hinder the effectiveness of DL models. To address this issue, we utilise transfer learning (TL) method [138] to leverage pre-trained models on large-scale datasets to improve recognition performance and reduce the need for large amounts of labelled radar data. For macro-activity monitoring, it is critical to address the challenges associated with different radar domain representations and optimise TL models for activity classification.

A notable example of DL is the fall detection system proposed in [208], which employs a contactless impulse radio ultra-wideband (IR-UWB) radar sensor and a CNN model. Their framework effectively distinguishes between "fall" and "non-fall" activities of daily living (ADL), protecting the privacy and comfort of the user, while achieving an impressive 96% test accuracy. However, their approach converts the radar signal into a single-channel grayscale image before feeding it into the CNN model for classification. We argue that ignoring the multi-domain radar information and reducing it to a single grayscale image may limit the model's ability to

capture essential activity features. This study explores alternative preprocessing strategies, including leveraging multi-domain radar representations, to enhance classification performance and robustness in HAR applications. This study not only benchmarks the performance of traditional radar preprocessing methods but also explores their integration with advanced TL models, providing a unique perspective for optimising the accuracy and computational efficiency of radar-based macro-activity recognition.

3.2 Contributions

This study aims to comprehensively evaluate different radar domains using FFT and its corresponding preprocessing techniques to improve HAR performance. Unlike previous studies that focus on a single radar domain or preprocessing method [23, 93], our work systematically compares the extraction and communication of essential features from raw radar data. To this end, the radar domain representation is converted into image format and analysed using different CNNs models trained specifically on radar data. This work addresses a significant gap in the literature by integrating four distinct radar domain representations in a single study and evaluating them using various state-of-the-art DNN models. We focus on understanding how these techniques convey features from raw radar data and influence the model's ability to classify human activities.

The primary contributions of this study are summarised and elaborated as follows:

- A detailed computational evaluation of four radar domain representations, such as RT, RD, TF using STFT, and SPWVD is performed, along with their preprocessing techniques and computational costs. This evaluation provides insights into the computational requirements and feature extraction capabilities of each domain, enabling the selection of the best representation for a specific HAR application.

- The feature extraction capabilities of each radar domain are benchmarked by analysing how they convey features from raw radar data using a baseline CNN trained from scratch. This baseline serves as a reference for evaluating the effectiveness of TL models and highlights the strengths and limitations of each preprocessing method in capturing essential activity features.

- A key contribution of this study is the application of advanced TL models to minimise the misclassification of falls, which is particularly important indicator for elderly care. By leveraging the advantages of TL models, we improve the reliability of fall detection and reduce false alarms, thereby improving the overall safety and effectiveness of the system.

- A comprehensive comparative analysis of the computational cost and recognition performance of various combinations of radar preprocessing techniques and TL models is

performed. The analysis identifies the most effective and efficient methods for real-time HAR applications, striking a balance between accuracy and computational feasibility.

- The cross-frequency generalisability of the optimal radar domain model pairs was validated using two additional publicly available FMCW radar datasets operating at 24 and 77 GHz. This validation demonstrates the robustness of the proposed approach across diverse radar-sensing technologies and confirms its applicability to various radar-based macro-activity systems, regardless of the operating frequency.

3.3 System Model for Macro-Activity Monitoring

This section introduces the system model underlying the proposed approach for R-HAR. The model comprises three core components: radar data acquisition, signal preprocessing through multiple domain transformations, and DL-based classification. It was specifically designed to support robust macro-activity recognition under indoor conditions using FMCW radar signals. Fig. 3.1 presents the data processing pipeline for macro-activity recognition, outlining the key stages from data acquisition to radar signal preprocessing and dataset preparation. The figure illustrates how raw FMCW radar signals are transformed into four distinct domain representations: RT, RD, TF-based STFT, and SPWVD maps, each capturing unique aspects of human activities. These processed radar maps form the input to several CNN models, including a baseline CNN and several TL models, which are evaluated for their effectiveness in classifying macro-activities. The following subsections describe each component of the processing pipeline in detail.

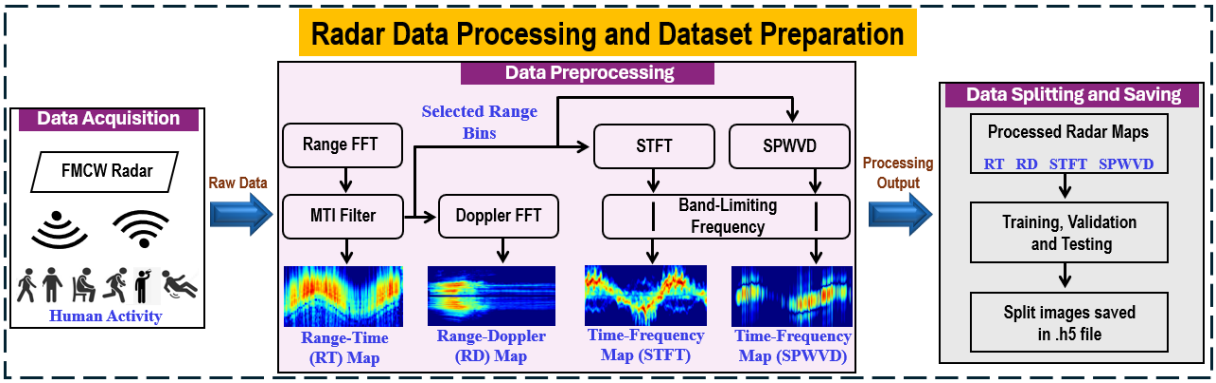


Figure 3.1: The workflow illustrates the radar data preprocessing flow, resulting in four distinct maps.

3.3.1 Radar Data Acquisition

This section outlines the radar data acquisition process, first focusing on the rationale for dataset selection, followed by the radar system used to collect the data. As the dataset forms the basis for

subsequent preprocessing and classification tasks, it is essential to understand both the context in which it was collected and the raw signal structure, which concludes with an intermediate frequency (IF) signal.

The Dataset

Despite the significant progress in R-HAR, a major limitation is the lack of large, standardised dataset for benchmarking classification algorithms. Most studies rely on proprietary datasets, which are typically collected using different radar sensors and involve a small, homogeneous group of participants, typically students. Such datasets limit the generalisability of DL models, as they may not represent realistic variations in human motion, environmental conditions, or population diversity. To address this issue, the radar community is increasingly advocating for open datasets, similar to what is done in image and audio processing. Efforts such as the Open Radar Initiative and classification challenges at major radar conferences (e.g., IEEE Radar Conference 2022 and IET Radar Conference 2020) encourage the public release of R-HAR datasets [209]. Some of the most widely used datasets include the following:

1. The Open Radar Initiative was presented at the 2021 IEEE Radar Conference to promote the sharing of datasets for algorithm benchmarking [209].
   Available: https://doi.org/10.1109/RadarConf2147009.2021.9455239

2. The "Radar Signatures of Human Activity" published by the Communication Sensing and Imaging Group at the University of Glasgow (UoG) is one of the first publicly shared datasets in R-HAR, with over 100 downloads [210].
   Available: https://researchdata.gla.ac.uk/848/

3. The DopNet dataset published by UCL Radar Group focuses on human gestures for human-computer interaction [211]. Available:https://github.com/UCLRadarGroup/DopNet

4. The PARrad dataset released by the University of Ghent focuses on recording patient activities in hospital environments [212]. Available: https://sumo.intec.ugent.be/parrad

5. The Radar Computational Intelligence Laboratory at the University of Alabama provides several datasets for research purposes [213]. Available: https://github.com/ci4r

For this study, we selected the radar dataset from the UOG (item 2 in the above list). This dataset forms the basis of all experimental work on macro-activity recognition presented in this chapter and is also used in Chapter 5 for edge-optimised inference. The following sections describe how the dataset was preprocessed, partitioned using subject-based splits, and used to train and evaluate a range of DL models across multiple radar domain representations. Table 3.1, provides an overview of the activity classes in the dataset, including the number of samples and their respective durations. Despite the limited number of activity types, the dataset offers several unique advantages, which are discussed in detail below:

Table 3.1: Description of macro-activity classes in the UoG radar dataset.

Short Name	Activity Description	Samples	Duration
A1	Walking back and forth	312	10 s
A2	Sitting on a chair	312	5 s
A3	Standing up	311	5 s
A4	Bend to pick an object	311	5 s
A5	Drinking water	310	5 s
A6	Fall	198	5 s

- First, the dataset included recordings from actual nursing home residents collected in their natural living environments and additional data from students and staff collected in laboratory settings. Specifically, nursing home residents were recorded at the Glasgow NG Nursing Home (three rooms) and the West Cumbria Elderly Centre (two rooms), while student and staff data were collected at the UoG laboratories, common rooms, and MAST labs. This environmental diversity introduces variations in background clutter (e.g. furniture and walls) and radar-to-subject distance, which are critical factors affecting radar signal quality. Unlike conventional datasets collected under strictly controlled laboratory settings [213], this setup captured more realistic variability. Consequently, it provides a valuable test bed for evaluating the model robustness under non-IID conditions, which is one of the key challenges addressed in this study.

- Second, the dataset comprised recordings from 105 subjects, including 26 female participants, with ages ranging from 21 to 98 years. This demographic diversity allows for the analysis of variations in gait and activity patterns across different age and sex groups, which are often under-represented in datasets composed primarily of younger individuals [22]. Although the dataset contains 1,754 samples, which is relatively small compared to those used in wearable sensor research, it still provides a valuable contribution to R-HAR by capturing the inter-subject variability that is critical for developing generalisable models.

Radar Configuration and Signal Characteristics

The dataset was collected using a commercial off-the-shelf (COTS) FMCW radar model 580-B from Ancortek. The system operates at a center frequency of 5.8 GHz with a chirp bandwidth of 400 MHz, providing high-resolution range and Doppler information. The experimental setup uses two Yagi antennas, one for transmission and one for reception, with a gain of approximately +18 dBi for each antenna. Both antennas are identical −17 dBi Yagi models, ensuring consistent signal transmission and reception characteristics [210]. The raw radar data consists of de-chirped complex beat frequency samples with 128 in-phase (I) and quadrature (Q) samples per chirp. These I and Q samples are essential for capturing amplitude and phase information,

allowing for detailed analysis of the radar signal. At this stage, the data is considered as "raw", because no preprocessing has been applied. Detailed radar parameters for the dataset are provided in the Table 3.2.

Table 3.2: FMCW radar parameters set during data collection.

Radar Parameters	**Values**
Duration of 1 chirp	1 ms
ADC samples per chirp	128
Pulse-repetition Factor (PRF)	1 KHz
Range Resolution	37.5 cm
Doppler Resolution	1.25 Hz
Velocity Resolution	3.2 cm/s^2
Max Doppler frequency $(f_d\text{max})$	500 Hz
Max Range (R_{max})	24 m

The working principle of the FMCW radar system is explained in Section 2.1.3. Briefly, the IF signal is obtained by mixing the transmitted signal with the conjugate of the received signal, followed by low-pass filtering and digitisation via the radar's onboard ADC. The resulting complex IF signal contains beat frequency information and is expressed as,

$$IF(t) = e^{j(2\pi f_d t + \phi_b)}, \tag{3.1}$$

where $\phi_b = 2\pi f_c \tau_d$ is the phase term, which represents the phase difference between the transmitted and received signals. This IF signal is then demodulated to obtain the I/Q components, from which the range and Doppler information are extracted. Thus, the dataset consists of radar data matrices rich in spatio-temporal information, forming the input for the subsequent preprocessing and classification stages.

3.3.2 Radar Data Preprocessing

Human activities exhibit complex spatiotemporal signatures that cannot be fully captured by a single radar domain. To address this, raw radar signals are transformed into three complementary 2D representations: RT, RD, and TF domains. Each of these domains captures a different aspect of human motion. The RT maps provided spatial positioning and coarse movement patterns over time. RD maps emphasise velocity characteristics through Doppler shifts, whereas TF maps highlight periodic or transient motion through time-varying spectral analysis. As illustrated in Fig. 3.1, this multi-domain approach enables CNNs to extract hierarchical and cross-domain features that are otherwise inaccessible through a single representation. By integrating these perspectives, the system becomes more robust to variations in viewpoints, partial occlusions, and ambiguous motion patterns, which are common challenges in real-world deployment scenarios. The preprocessing pipeline ensures that each domain preserves essential motion dynamics while

effectively reducing background noise, thereby producing optimal inputs for DL based macro-activity recognition.

R-FFT for RT Domain

The IQ components of the IF signals undergo several processing steps to enhance their quality and accurately represent human activity. Initially, these signals are combined to form a complex digital signal $y[n] = I[n] + jQ[n]$, where j is the imaginary unit, detailed in Section 2.1.3. The complex signal is then reshaped into a 2D matrix to align the data for subsequent processing steps.

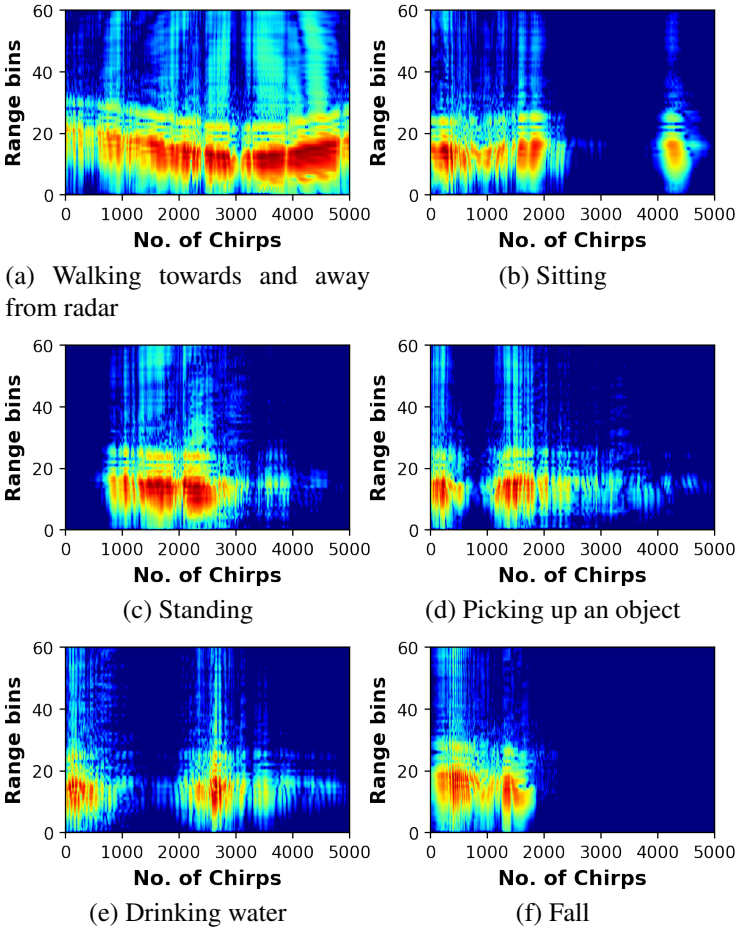
The process begins by applying the Hamming window to minimize spectral leakage, followed by conducting an FFT on the fast-time axis (also referred to as the number of ADC samples per chirp) to extract range information over time, known as R-FFT or range profile. The range profile $X_n[k]$, for chirps $n = 1, 2, ..., N_{tot}$ (N_{tot} is the total number of chirps), is defined as the DFT of $y[i]$ [95]. For the presence of a target, we examine the peaks in the magnitude of $X_n[k]$, which is acquired via DFT as:

$$X_n[k] = \sum_{i=0}^{N} w[i] y_n[i] e^{-j2\pi \frac{ki}{N}}, \tag{3.2}$$

where $N = 128$ is the number of ADC samples of the received IF signal y_n for chirp n, k shows the frequency bin, and w shows the window function, which is hamming in our case.

The RT domain reflects the time-varying range information between the radar and the target. A filter called a moving target indicator (MTI) detects only moving targets and effectively removes any clutter or stationary objects from the radar signal. The MTI filter was designed as a fourth-order Butterworth high-pass filter with a cut-off frequency of 0.0075 Hz. After applying the MTI filter, relevant range bins corresponding to the target's expected range are extracted, and the data are further processed to generate the RT map. The filtered RT map now provides a clearer representation, as shown in Fig. 3.2, by eliminating background noise and focusing on target movement. This graphical representation shows how the signal amplitude varies over time and range, effectively visualising the target's movement.

Once the RT map is generated, it is used as one of the preprocessed input for the DL models. By visually inspecting these maps, we identified the range bins of interest k^*, where the targets are predominantly detected. Additionally, only the range bins of interest, specifically from 5 to 25 (1.8 m to 9.3 m), where the subjects are detected, are selected. These selected bins form the basis for subsequent RD and TF analysis using STFT and SPWVD techniques.



(a) Walking towards and away from radar

(b) Sitting

(c) Standing

(d) Picking up an object

(e) Drinking water

(f) Fall

Figure 3.2: The RT maps of a young adult performing different activities.

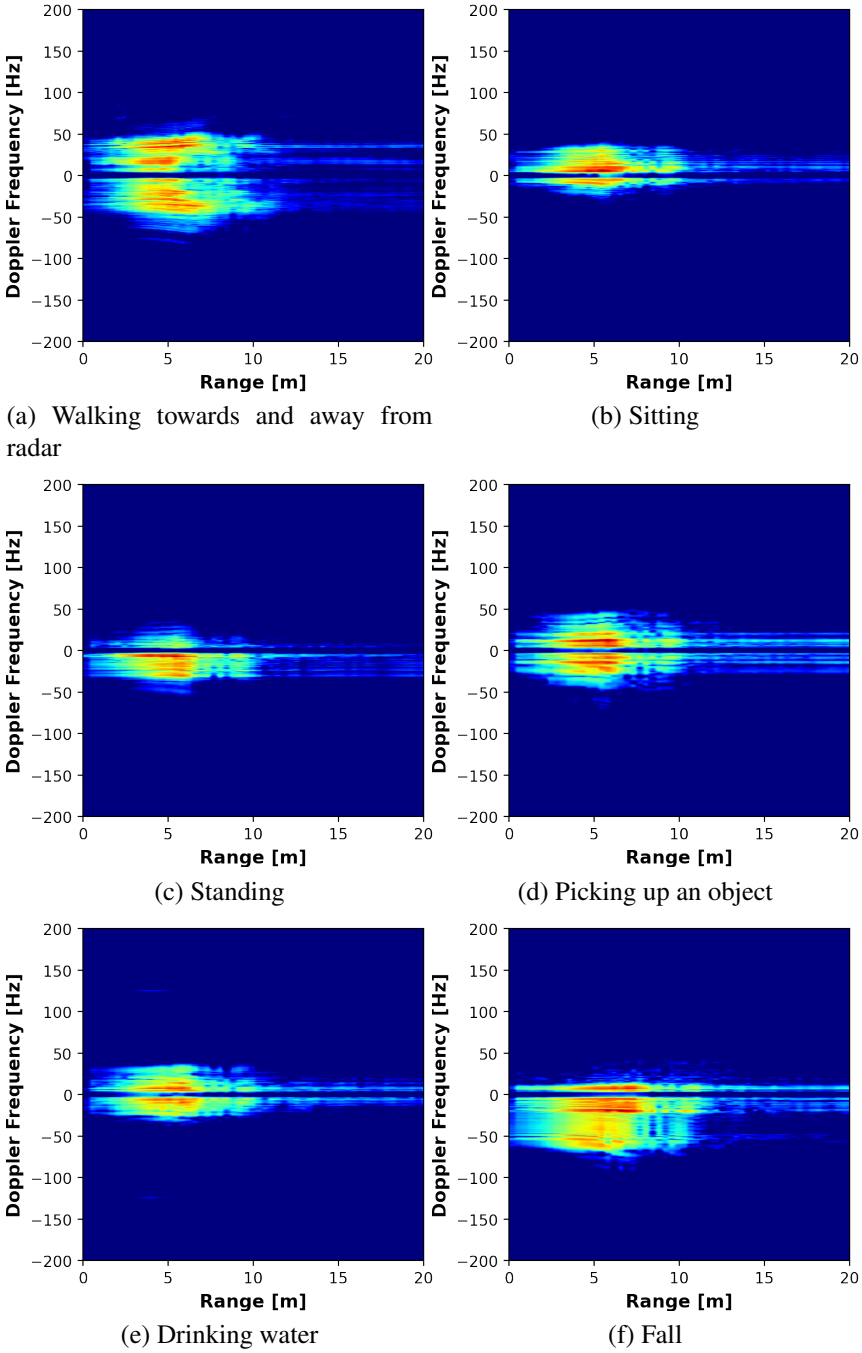
RD Domain

After performing the FFT and obtaining the RT Map, the next step was to extract Doppler information. This is achieved by applying a second FFT across the chirp dimension (slow-time axis) for each range bin, which is also known as Doppler-FFT (D-FFT). The result reveals the Doppler shift, a frequency shift caused by the relative motion of the target with respect to the radar. The Doppler shift is proportional to the target's velocity. The D-FFT is performed independently for each range bin, resulting in a 2D matrix called the RD Map. The mathematical representation of D-FFT is [214]:

$$D[m,l] = \sum_{n=0}^{M-1} X_{\text{MTI}}[m,n] e^{-j\frac{2\pi}{M}ln} \tag{3.3}$$

Hence, $D[m,l]$ represents the RD domain, m is the range bin, and l is the Doppler frequency bin. $X_{\text{MTI}}[m,n]$ is the output of the R-FFT and MTI filtering for the m^{th} range bin and n^{th} chirp, and M is the number of chirps. Each column in the resulting RD matrix corresponds to a single Doppler bin for all chirps, and each row represents a specific range bin. The RD maps for

different macro-activities are shown in Fig. 3.3, and is used as input to the CNN, which serves as the second radar domain representation method after RT map, as illustrated in Fig. 3.1.



(a) Walking towards and away from radar

(b) Sitting

(c) Standing

(d) Picking up an object

(e) Drinking water

(f) Fall

Figure 3.3: The RD maps of a young adult performing different activities.

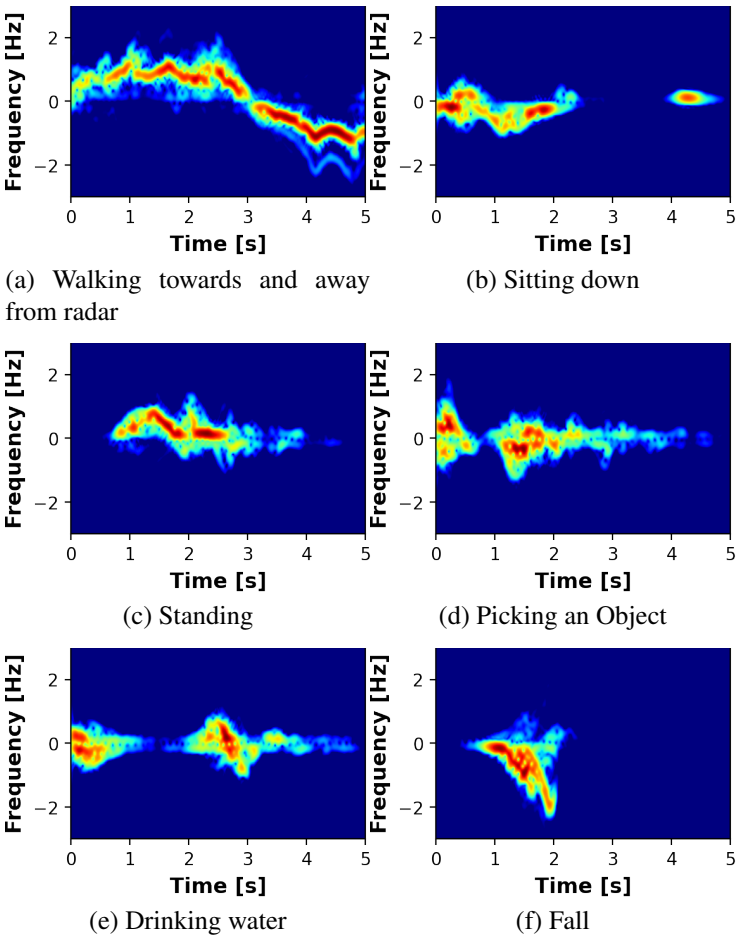
STFT-Based TF Domain

To obtain a TF-based radar representation of the macro-activity, we employed the STFT technique. In particular, the STFT employs a Hann window of size 200, with the FFT incorporating 800 sampling points, a zero padding factor of 4, and a 95% overlap between consecutive frames

(i.e., 190 samples). This approach balances frequency and time information, resulting in a 2D image, as shown in Fig. 3.4. The mathematical equation of the STFT applied to the selected range bins of interest k^* from the range profile $X_n[k^*]$ is given by:

$$\Theta[m,f] = \sum_{n=-\infty}^{\infty} X_n[k^*] g_s[n-mR] e^{-j2\pi fn}. \tag{10}$$

Where k^* denotes the range bins of interest (5 to 25), and g_s represents the window function, which is a Hann window with a segment length of $s = 200$ samples and a hope size of $R = 10$ samples. The TF spectrograms which were generated by taking the magnitude of $\Theta[m,f]$, are shown in Fig. 3.4. The TF resolution of the STFT domain is contingent upon the selection of



(a) Walking towards and away from radar

(b) Sitting down

(c) Standing

(d) Picking an Object

(e) Drinking water

(f) Fall

Figure 3.4: The TF maps based on STFT of a young adult performing different activities.

the window function. An extended window increases the frequency resolution, but a reduced window maximises the time resolution. As a result, the STFT experiences a trade-off, rendering it unable to achieve high resolution in both the time and frequency domains concurrently.

SPWVD-Based TF Domain

Since STFT is limited by non-independent time and frequency windows, SPWVD provides a more sophisticated approach to TF analysis by utilising independent windows. SPWVD is a subset of the WVD, which offers a high TF resolution but suffers from cross-term artifacts. SPWVD addresses this limitation by applying smoothing operations in both the time and frequency domains, effectively eliminating the cross terms in both dimensions. This dual smoothing operation makes SPWVD advantageous for WVD by providing a cleaner TF representation. SPWVD is also applied to the extracted range bins of interest, $X_n[k^*]$, and is mathematically expressed as follows [215]:

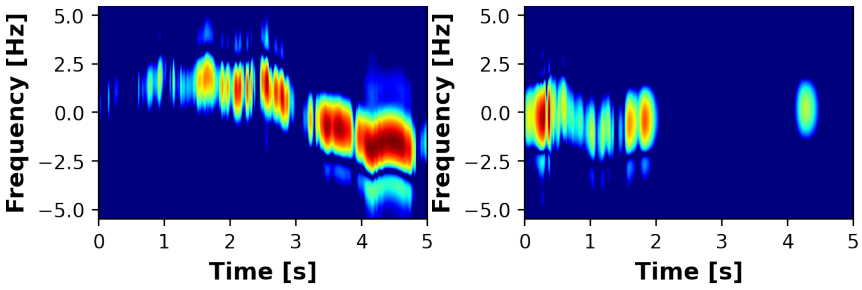
$$\Theta[m,f] = \sum_{n=-\infty}^{+\infty} \sum_{\tau=-\infty}^{+\infty} X\left(k^* + \frac{\tau}{2}, n\right) X\left(k^* - \frac{\tau}{2}, n\right) \times h(\tau) w(n-m) e^{-j2\pi f \tau} \tag{11}$$

Here, $\pm\frac{\tau}{2}$, represents the symmetrical time shifts used for cross-term smoothing in the SPWVD. For the frequency smoothing window $h(\tau)$, we employed a Hann window with a length of 15, whereas for the time smoothing window $w(n-m)$, we used a Kaiser window with a length of 25, and m is the time index. To further reduce artifacts and avoid intensity spikes, the SPWVD outputs were normalized by the number of range bins. This comprehensive approach provides detailed SPWVD-based TF representations, as shown in Fig. 3.5. For dimensionality reduction, we apply Doppler spectrum band-limiting to remove out-of-band noise, retaining only the reduced portion of the Doppler axis between normalised frequencies. Specifically, for the STFT, we maintained the range between [-0.1, 0.1] Hz, whereas for the SPWVD, the range was adjusted to [-0.3, 0.3] Hz. These frequency bands were identified and selected by visually inspecting the maximum significant extent of the Doppler spectra for the specific radar representations.

One of the objectives of this study is to evaluate the applicability of the SPWVD as a spectrogram technique and to examine its characteristics to determine the potential benefits for efficient HARS. Therefore, integrating these four domains, utilising their respective strengths, and addressing their challenges are essential to improve the performance of the R-HARS. This comprehensive approach ensures that the data fed into the CNN architectures are well suited for learning and recognising activity-specific features, thus achieving a balance between processing efficiency and performance accuracy.

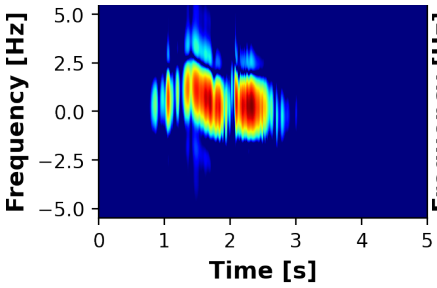
3.3.3 Dataset Settings and Splitting

A critical aspect of this study is the careful division of the dataset into training, validation, and testing sets to avoid data leakage, which is a common issue that can lead to inflated performance metrics. In contrast to the typical 80-20% random split used in other studies [28], where repetitions of the same activity by the same subject may appear in both training and testing sets, we

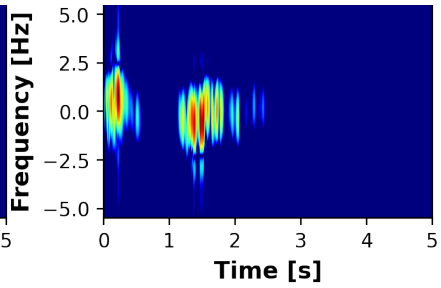


(a) Walking towards and away from radar

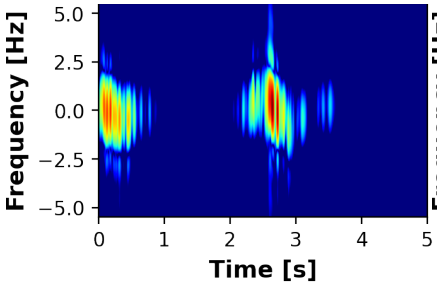
(b) Sitting down



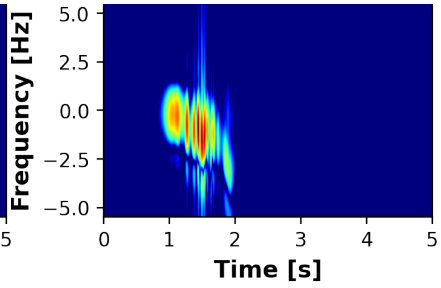
(c) Standing



(d) Picking an object



(e) Drinking water



(f) Fall

Figure 3.5: The SPWVD-based TF maps of a young adult performing different activities.

opted for a subject based split, as illustrated in Table 3.3. This ensured that the subjects used for testing were entirely different from those used in training, thus preventing the model from memorising specific patterns associated with particular individuals rather than learning to recognise the activities. Each subject performed the same activity three times; thus, it is, crucial to split the dataset in such a way that all repetitions from a single subject for a particular activity are confined to a single set. This methodology ensures that the model's performance is evaluated on truly unseen data, thereby providing a more accurate assessment of the model's generalisability.

3.4 Deep Learning for Macro-Activity Recognition

In this section, we present our DL framework for macro-activity recognition using radar data. We explain the models used, the optimisation strategies applied, and how these methods improve the classification accuracy using each radar data representation. The workflow of the DL for R-

Table 3.3: Subject-based dataset splitting.

Total	**Split**	**Subjects**
105 Subjects	Training (75%) *Remaining (25%)	79 Subjects 26 Subjects
***Remaining (25%)**	Validation (75%) Testing (25%)	19 Subjects 7 Subjects
Classes	**Short Name**	**Training Samples**
Walking back and forth	A1	232
Sitting on a chair	A2	234
Standing	A3	234
Picking up an object	A4	232
Drinking	A5	234
Fall	A6	120
Training		1286 Samples
Validation		336 Samples
Testing		126 Samples

HARS is shown in Fig. 3.6, which provides a comprehensive overview of the process. The main goal is to determine the best performing radar map in combination with the most suitable DL model.

To do this, we first train radar maps on a baseline CNN to establish a performance baseline and identify limitations. These limitations are then addressed using TL-based models that leverage pre-trained architectures to improve accuracy and generalisation. With this systematic approach, we aim to optimise the model performance of each radar representation, ensuring robust and accurate classification of human activities in real-world applications.

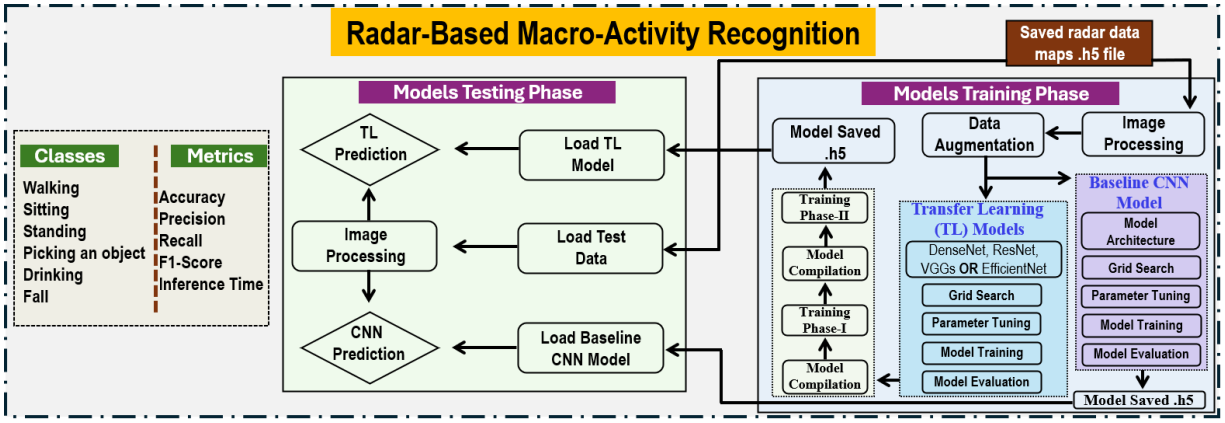


Figure 3.6: Workflow of radar-based human macro-activity recognition using deep and transfer learning models.

3.4.1 Image Processing and Data Augmentation

After applying the preprocessing steps, each technique produced a distinct representation of the radar echo reflections. The radar domain representations must be prepared before being fed into

the CNN models. The first step in the model training pipeline is to standardize all images in the dataset. Each image was resized to a uniform size of 224 × 224 to satisfy the input size requirement of the selected CNN model. The preprocessing stage includes image normalisation and subtraction of the mean RGB value based on the model and training strategies, along with other necessary transformations.

Because the dataset was split manually, as detailed in Table 3.3, there was no further need for splitting or shuffling. However, due to the small size of the training dataset and the presence of class imbalance, we implemented two key techniques to mitigate overfitting: Data Augmentation and Class Weights.

- Data augmentation was applied exclusively to the training set using the data generator function from the Keras library, while the validation and test sets were kept unchanged to ensure fair model evaluation.

- Secondly, class weights from the scikit-learn library were used to address the class imbalance.

This function automatically adjusts the loss function to assign weights to underrepresented classes, ensuring that the model does not disproportionately favor more frequent classes. The augmentation strategy attempts to improve model generalisation by reducing overfitting, and this should result a smaller gap between the training and validation accuracies. The proposed augmented approach was inspired by the work of [216], who demonstrated the effectiveness of a similar technique for image classification tasks. In addition, class weighting was selected for its seamless integration into the model training process, offering a simple yet effective way to handle class imbalance without altering the dataset distribution or increasing the computational overhead.

3.4.2 Convolutional Neural Network

The evaluation begins with a baseline CNN model. The proposed baseline model serves as a starting point for assessing the effectiveness of various radar domain representations and preprocessing techniques. By analysing the performance of this baseline model, we gain initial insights into the ability of each method to enhance data representation for accurate classification. The structure and operation of the baseline CNN model used in this study are illustrated in Fig. 3.7. The model architecture consists of four parts: the input layer, convolutional layer, pooling layer, fully connected (FC) layer and output layer, which are explained in detail below [108]:

- **Convolutional Layer:** This layer applies a series of filters, also called kernels, to the input data. Each filter slides across the input to generate a feature map. By aggregating all generated feature maps, we obtain the output of the convolutional layer. The mathematical

relationship between the input and output in a convolution layer is given by [108]:

$$y_j^l = g^l(v_j^l) = g^l\left(\sum_i (y_i^{l-1} * w_{ij}^l) + c_j^l\right) \tag{13}$$

where y_j^l denotes the output feature map of the j^{th} channel in layer l. The function $g^l(.)$ represents the activation function of the convolution layer, which is ReLU in this study, applied to the net activation v_j^l of the j^{th} channel. The net activation v_j^l is computed as the sum of the convolutions over the input channels y_i^{l-1} using the convolution kernel w_{ij}^l and, the bias term c_j^l.

- **Pooling Layer:** This layer takes the feature maps as input and performs subsampling to reduces their dimensionality. Pooling reduces computational complexity and retains the most significant features. We employ the MaxPooling layer, which selects the maximum value from a defined window over the feature map, effectively reducing the size while preserving the key spatial features. The mathematical relationship between the input and output of the pooling layer is given by [108]:

$$y_j^l = g_{\text{pooling}}^l(v_j^l) \tag{14}$$

Where y_j^l shows the output of the channel j^{th} in pooling layer l, and v_j^l represents the input activation. The function g_{pooling}^l denotes the pooling function, which is MaxPooling in this case, and selects the maximum value from each region of the input feature map.

- **FC Layer:** This layer takes the output from the convolutional and pooling layers and flattens it into a vector. This vector serves as the input to the classification function, which involves a softmax activation function, [108]:

$$v_j^l = \sum_i x_i^l \cdot w_{ij}^l + c_j^l \tag{15}$$

$$y_j^l = g^l(v_j^l) = \max(0, v_j^l) \tag{16}$$

In these equations, v_j^l denots the net activation of the j^{th} neuron in the fully connected layer l, and y_j^l represents the output after applying the activation function. w_{ij}^l is the weight associated with the connection between input i and neuron j, and c_j^l is the bias term. The ReLU activation function, represented by $\max(0, u_j^l)$, ensures non-linearity in the network.

Training the baseline CNN model involved hyperparameter optimisation via grid search and manual tuning to determine the best configuration for each radar map. These parameters included a batch size of 32, the Adam optimiser, the He Uniform kernel initializer, and a patience

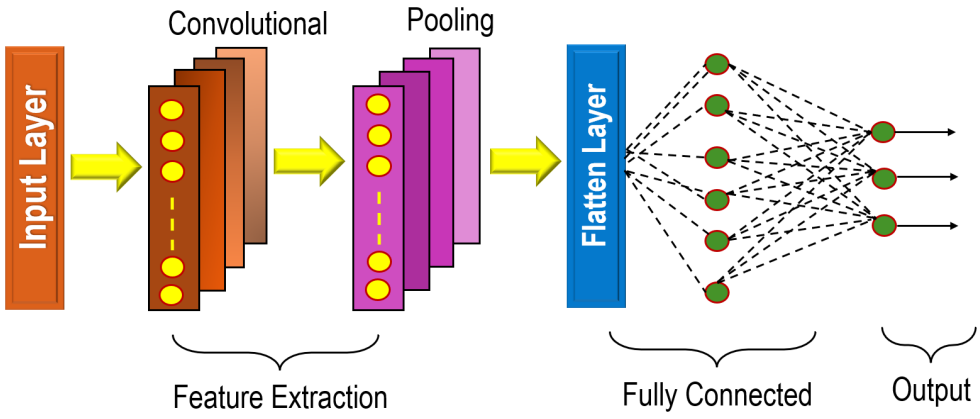


Figure 3.7: Baseline CNN Model Architecture.

learning rate (lr) reduction mechanism for three epochs. To prevent overfitting, we also implemented early stopping with a patience interval of 10 epochs, ensuring that training would stop if the model stopped improving. The total number of epochs for training the baseline CNN was set to 150 across all radar maps.

The model initially had a lr of 1.0×10^{-4} for all radar maps, except for SPWVD, which required a higher lr of 1.0×10^{-3}. In subsequent stages, the lr is decreased as follows:

- TR and STFT: 5.0×10^{-5}
- SPWVD: 1.0×10^{-5}
- RD: kept at 1.0×10^{-4}

To further alleviate overfitting, L2 regularisation with different coefficients is applied:

- TR and STFT: 5.0×10^{-5}
- SPWVD: 1.0×10^{-3}
- RD: 2.0×10^{-3}

The dropout rate for TR and STFT is set to 0.4, for SPWVD is set to 0.5, and for RD is set to 0.2. The FC layers also vary depending on the radar representation, such as: for TR and STFT, two FC layers with 128 and 64 neurons, for RD one FC layer with 128 neurons, and for SPWVD two FC layers with 256 and 128 neurons, respectively, due to its higher complexity.

3.4.3 Transfer Learning Models

Transfer Learning (TL) is a powerful technique in which a model developed for one task is reused as the starting point for a model on a second task [141]. It is particularly beneficial in scenarios with limited labeled data because, it enables faster convergence and better generalisation for new tasks by leveraging knowledge from pre-trained models. In the TL, the pre-trained

models, which has been trained on a large dataset, transfers its learned parameters, especially the weights into the target network. After evaluating the four radar domain representations and preprocessing methods with the baseline CNN, we utilised the same methods to be an input to various pre-trained models, including DenseNet-201, ResNet-34, VGG-16, VGG-19, and EfficientNet-B0, as shown in Fig. 3.6.

DenseNet-201 Model

The densely connected convolutional network (DenseNet) proposed by [217] well known for its outstanding performance on benchmark datasets such as CIFAR-100 and ImageNet [218]. The architecture adopts a densely connected pattern, where each layer is directly connected to all previous layers in a feed-forward manner. Compared with traditional CNNs, this design enhances feature propagation, encourages feature reuse, alleviates the gradient vanishing problem, and reduces the number of parameters. In this study, a variant of DenseNet-201 consisting of 201 trainable layers was selected because it can effectively process radar signals after 2D transformation. DenseNet-201 achieved the best results when trained on STFT-based TF maps, highlighting its potential in HAR. A custom classification head was added to tune the classification model. This head contains:

- Global Average Pooling (GAP) layer for feature aggregation
- Batch normalisation layer for stable training
- Dropout layer with rate of 0.3 to prevent overfitting
- Final dense layer with 6 neurons that shows classes and softmax activation for multi-class classification

ResNet-34 Model

The Residual Network (ResNet) proposed by [145], introduced skip connections to enable deeper CNN architectures by alleviating the vanishing gradient problem. ResNet-34 is a variant with 34 layers and is widely used for image classification and feature extraction due to its high training efficiency and high accuracy [219]. The key innovation of ResNet is the residual building block, which bypasses the convolutional layers using shortcut connections, allowing the model to retain necessary low-level features while learning deep representations. This architecture is particularly beneficial for R-HAR, as subtle motion patterns must be extracted from different radar representations.

In this study, ResNet-34 is used as the backbone model and pre-trained on ImageNet. The classification head consists of a GAP layer for feature extraction, followed by batch normalisation and a dropout layer with a dropout rate of 0.4. The final dense layer is initialised with

He Normal and has six output neurons with softmax activation corresponding to the six activity classes for classification.

VGG-16 Model

Developed by the Visual Geometry Group (VGG) at the University of Oxford [147], VGG-16 is a 16-layer deep CNN architecture known for its simplicity and strong performance in image classification tasks. The main advantage of VGG-16 is that it uses multiple small 3 × 3 convolutional filters instead of larger kernel sizes, which enhances feature extraction while maintaining computational efficiency. The model achieves 92.7% top-5 accuracy on ImageNet, demonstrating its effectiveness in feature learning.

In this study, VGG-16 was used as a pre-trained backbone model, trained on ImageNet, to extract features from radar data. To adapt the architecture for macro-activity recognition, a custom classification head was added, including a GAP layer for feature aggregation, followed by batch normalisation to stabilise training. Dropout layers were added to prevent overfitting, and the final dense layer contained six neurons with softmax activation function for multi-class classification. Apply He normal initialization to all dense layers to improve weight initialization and training stability.

VGG-19 Model

VGG-19 is an extended version of VGG-16 with 19 layers instead of 16. While maintaining the same architectural principles, the additional layers enable deeper feature extraction, which improves recognition performance, especially for complex motion patterns in R-HAR. The model uses small 3×3 convolutional filters to effectively capture spatial features while reducing computational complexity.

In this study, VGG-19 was used as a pre-trained backbone model initialised with weights from ImageNet to extract high-level features from radar data. To adapt the model to HAR, a custom classification head was added, which consists of a GAP layer for feature aggregation followed by batch normalisation to ensure stable training. A dropout layer with a dropout rate of 0.4 was added to mitigate overfitting, and the final dense layer contained six output neurons with softmax activation for activity classification. This architecture leverages the deep feature extraction capabilities of VGG-19, making it well suited for R-HAR tasks.

EfficientNet-Bo Model

EfficientNet is an efficient CNN architecture designed to achieve state-of-the-art accuracy while maintaining computational efficiency [220]. Unlike traditional deep networks, EfficientNet uses compound scaling, adjusting depth, width, and resolution simultaneously to obtain the best balance between performance and efficiency. In addition, it leverages AutoML for automatic hy-

perparameter tuning and efficient convolution operations to reduce computational overhead. For macro-activity recognition using radar data, we selected EfficientNet-B0, which is the baseline model in the EfficientNet family. EfficientNet-B0 is designed using Neural Architecture Search (NAS) and serves as the basis for larger models of increasing complexity. The model uses 3×3 convolutional filters for feature extraction.

We use the pre-trained EfficientNet-B0 model and apply TL to classify R-HAR into six activity categories. To adapt EfficientNet-B0 for HAR, a custom classification head is added, including GAP for feature extraction, followed by batch normalisation to stabilise training. A dropout layer with dropout rate of 0.5 is included for regularisation, and the final dense layer contains six softmax neurons for classification.

3.4.4 Model Training and Optimisation

Each radar representation, is trained sequentially using aformentioned TL based models. The training strategy adopts a two-phase approach to ensure effective feature extraction and fine-tuning for HAR. For each radar map input to the TL model, we optimise the parameters determined by grid search and manual tuning as described in the baseline CNN model. To prevent overfitting and improve generalisation, we apply techniques such as L2 regularisation, dropout layers, and class weights. The dropout rate ranges from 0.3 to 0.6, while the L2 regularisation value is adjusted according to the model and radar representation to balance feature preservation and generalisation. The lr is initialised to 0.0001, and some models reduce it to 0.00002 or 0.00001 during fine-tuning to ensure gradual adaptation. A set of callbacks are used to stabilise training and optimise learning.

- **Model Checkpoint:** Save the best performing model based on validation loss.
- **ReduceLROnPlateau:** Dynamically reduce the lr if validation loss stabilises.
- **EarlyStopping:** Stop training if no improvement is observed within a predefined number of times.
- **Time History:** Record training time for performance evaluation.
- **Learning Rate Warmup:** Gradually increase the lr in the initial epochs to stabilise training.

These callbacks ensure efficient training by dynamically adjusting the lr, saving the best model, and preventing overfitting.

Optimisation Strategy for Each Radar Representation

The optimisation strategy is tailored to the unique features of each radar representation. All models were trained using the Adam optimiser, with the exception of ResNet-34 in phase 1,

which employed SGD. Categorical cross-entropy served as the loss function across all experiments, with further details provided below.

- RT maps contain high feature variability due to their complex and noisy temporal and spatial resolution, requiring larger dense layers (256 neurons), higher dropout (0.6), and stronger L2 regularisation (0.003 for EfficientNet and 0.001 for VGG-16, VGG-19, and ResNet-34). A lr warmup strategy was applied, with EfficientNet warming up for 10 epochs and VGG and ResNet models warming up for 20 epochs before transitioning to fine-tuning. The lr was initially set to 0.0001 and was reduced to 0.00002 in order to fine-tune the weight updates of the control.

- The RD maps exhibits smoother feature transitions, which reduces the complexity of the classification task, allowing the use of smaller classification heads (128 neurons) and moderate L2 regularisation (0.001) to prevent overfitting. Dropout values were set to 0.5 and 0.4 to balance feature preservation and generalisation. A lr warmup of 15 epochs stabilised the training before fine-tuning. The initial lr was set to 0.0001, then reduced to 0.00001 in the second stage to enable gradual adaptation of the learned features from radar data.

- The STFT maps encodes local spectral information and shares features with the RD map, such as smooth feature transitions, allowing similar optimisation methods. The dense layer contains 128 neurons with dropout values of 0.5 and 0.4, respectively, while L2 regularisation was set to 0.001. A lr warmup value period of 15 epochs precedes fine-tuning to ensure smooth adaptation. The lr begins at 0.0001, and decreases to 0.00001 to optimise spectral feature learning.

- The SPWVD maps requires larger dense layers (256 neurons) and moderate dropout (0.5) to effectively capture fine-grained spectro-temporal features, which are more complex and detailed compared to other radar representations. L2 regularisation (0.001) is applied to prevent overfitting. Before fine-tuning, the EfficientNet and VGG models used a lr warmup of 20 epochs to ensure stable weight updates. Unlike other maps, all layers (including batch normalisation layers) are unfrozen during fine-tuning, allowing the models to fully adapt to the radar dataset.

Two-Phase Training Approach for TL Models

Each radar representation was evaluated using five different TL-based models, resulting in 20 model-domain combinations. The architecture on top of these models is shown in Fig. 3.8, and a two-phase training approach is employed for these models, thereby allowing effective feature extraction and fine-tuning of the model.

1. **Phase 1: Feature Extraction:** In Phase 1, the pre-trained model was used as the feature extractor. The base pre-trained model remains frozen, meaning that its weights are not

Figure 3.8: Customized TL-based models architecture, which underwent two-phase training.

updated during this phase. Its role is to extract general features from the radar dataset by using the knowledge acquired by training on the large-scale ImageNet dataset. During this phase, only the new layers, which are added to the base model, are trained, as shown in Fig. 3.8. This allows the model to learn how to classify radar-based data without modifying the pre-trained features.

2. **Phase 2: Fine-Tuning:** In Phase 2, some or all of the base model layers are unfrozen, and its weights are fine-tuned on the radar dataset. This step allows the model to adjust and refine the features it initially learned from ImageNet, thereby making them more relevant for radar data. The fine-tuning process prevents overfitting or excessive bias towards the ImageNet dataset. By gradually adapting the model to the new domain, the model can be better generalised and specialised for radar-based classification, minimising confusion between ImageNet and radar map features. Thus, Phase 1 ensures that the model begins with strong, generalised features, while Phase 2 refines these features to align more closely with the R-HAR dataset, thereby avoiding interference from ImageNet specific characteristics.

Runtime Environment

In this study, we performed data preprocessing on raw radar data and CNN training using Python toolkits on a GPU-accelerated desktop PC. The system was equipped with an 11^{th} Generation Intel® Core™ i7-11700 processor with 8 cores and 16 threads enabled by hyper-threading technology and a base frequency of 2.50 GHz. It was also equipped with 16 GB of RAM and an NVIDIA GeForce RTX 3060 Ti graphics card with 8 GB of memory. To generate radar spectra, we used Python libraries such as Scipy for signal processing, TF for spectrum analysis, and fftpack from scipy for FFT execution. For CNN model training, we used the Keras and TensorFlow frameworks, taking advantage of the multicore and multithreaded CPU capabilities of the workstation and the parallel processing power of the GPU.

3.4.5 Evaluation Metrics

When evaluating the impact of radar domain representation and DL models using TL approaches for real-world deployment settings, both computational efficiency and recognition accuracy must be considered. While additional time components exist, such as radar preprocessing, ADC conversion, and signal processing within radar boards, these remain consistent across all processing methods. Consequently, they are excluded from the comparison since they do not affect the relative performance differences between preprocessing techniques and models.

The time complexity of training a CNN models is theoretically determined by the number of epochs (T), number of parameters (D), and the number of batches (B), as indicated by $O(T.D.B)$ [221]. In our comparative analysis, we measured the training time per epoch to provide an empirical estimate of computational cost. Given that the model was trained for T epochs, the total training time can be approximated by multiplying the average training time per epoch by T. This empirical analysis complements the theoretical complexity by providing a real-world context for the computational effort.

To quantify the prediction or inference time, we compute various statistical measures based on the inference times $t_{\text{infer},i}$ for each sample i in the test set, where N is the total number of samples as follows:

- **Average Inference Time**: The average inference time is calculated by taking the inference time for each individual image and then averaging these values across all images:

$$\bar{t}_{\text{infer}} = \frac{1}{N}\sum_{i=1}^{N} t_{\text{infer, i}} \tag{17}$$

- **Standard Deviation**: The standard deviation of inference time, σ_{infer}, is defined as follows:

$$\sigma_{\text{infer}} = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left(t_{\text{infer, i}} - \bar{t}_{\text{infer}}\right)^2} \tag{18}$$

- **Standard Error**: The standard error of inference time, SE_{infer}, is calculated as follows:

$$\text{SE}_{\text{infer}} = \frac{\sigma_{\text{infer}}}{\sqrt{N}} \tag{19}$$

These metrics provide a comprehensive understanding of the model's time efficiency and consistency, which are critical for real-time applications.

However, fast prediction is insufficient without accurate performance. Thus, recognition accuracy is also a critical metric, especially for applications involving critical tasks such as monitoring elderly people in their homes. High accuracy ensures reliability and safety, reducing the likelihood of false alarms and missed detections. Therefore, for this purpose, we can calculate

accuracy as follows:

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}} = \frac{\sum_{i=1}^{C} \text{TP}_i}{N}. \tag{20}$$

Here C is the number of classes, N is the total number of samples, and TP_i is the number of true positives for class i. To provide a more detailed assessment of the model's performance across different classes, we compute the precision, recall, and F1-score for each class i:

$$\text{Precision}_i = \frac{\text{TP}_i}{\text{TP}_i + \text{FP}_i}, \tag{21}$$

$$\text{Recall}_i = \frac{\text{TP}_i}{\text{TP}_i + \text{FN}_i}, \tag{22}$$

$$\text{F1-Score}_i = 2 \times \frac{\text{Precision}_i \times \text{Recall}_i}{\text{Precision}_i + \text{Recall}_i}, \tag{23}$$

where TP_i shows True Positives for class i, FP_i is the False Positives for class i, and FN_i is False Negatives for class i. These metrics help identify the model's ability to correctly classify instances in each class, thereby minimising FP and FN. Training and validation accuracies, along with confusion matrices, were analysed to assess model performance and generalisation capabilities across different radar preprocessing techniques. These metrics provide comprehensive insights into model effectiveness while identifying potential overfitting or underfitting issues.

3.4.6 Proposed Radar-Based Macro-Activity Recognition Algorithm

The proposed radar-based macro activity recognition system is outlined in Algorithm 1, which covers the entire pipeline from data acquisition through preprocessing, model training, and activity classification.

3.5 Experimental Results and Discussion

This section presents the performance evaluation of baseline CNN and TL models across different radar domain representations. The evaluation focuses on classification metrics and computational efficiency, with particular emphasis on cross-subject generalisation to assess real-world deployment feasibility.

3.5.1 Radar Data Preprocessing Time Analysis

The preprocessing techniques employed in this study showed significant differences in the time required to process radar data events. Among the methods evaluated, SPWVD had the longest preprocessing time (t_{pre}) due to its high computational complexity. This method processes two

Algorithm 1: Proposed Radar-Based Macro-Activity Recognition System

Require: Raw radar data $y[i]$
Ensure: Classified human activities

procedure DATA ACQUISITION AND PREPROCESSING
  Collect raw radar data $\{y[i]\}$ and convert to complex signal $s[i] = y[i] + j \cdot \hat{y}[i]$
  Reshape $s[i] \rightarrow \mathbf{S}$, a 2D matrix of size (N, M)
  Apply Hamming window $w[n]$: $\mathbf{S_w}[i] = \mathbf{S}[i] \cdot w[n]$
  Apply IIR Notch MTI filter: $\mathbf{S_{filtered}}[i] = \text{IIR}(\mathbf{S_w}[i])$
  Perform Range-FFT: $X_n[k] = \text{FFT}(\mathbf{S_{filtered}}[i])$
  Extract desired range bins: $\mathbf{R} = X_n[k]$, for $k \in$ Range bins
  Choose Preprocessing Method:
    Option 1: Generate Range-Time (RT) maps: $\mathbf{RT}(t, r)$
    Option 2: Generate Range-Doppler (RD) maps: $\mathbf{RD}(t, r)$
    Option 3: Apply STFT: $\mathbf{TF}(t, f) = \text{STFT}(\mathbf{R})$
    Option 4: Apply SPWVD: $\mathbf{TF}_{\text{high-res}}(t, f) = \text{SPWVD}(\mathbf{R})$
  Apply Band Limiting (for TF-based methods):
    If Option 3: Apply frequency band limiting to STFT: $\mathbf{TF}_{\text{band}}(t, f) = \mathbf{TF}(t, f)|_{f \in [-0.1, 0.1]\ \text{Hz}}$
    If Option 4: Apply frequency band limiting to SPWVD: $\mathbf{TF}_{\text{band}}(t, f) = \mathbf{TF}_{\text{high-res}}(t, f)|_{f \in [-0.3, 0.3]\ \text{Hz}}$
    If Option 1 or 2: No band limiting applied
  Resize images to 224×224 and normalize: $\mathbf{I} = \text{Norm}(\mathbf{I}_{224 \times 224})$
  Split data: $\mathbf{I}_{\text{train}}, \mathbf{I}_{\text{val}}, \mathbf{I}_{\text{test}} \leftarrow \text{Split}(\mathbf{I})$
end procedure
procedure MODEL SELECTION AND TRAINING
  Choose CNN base model:

  Base pre-trained model $\in \{\text{DenseNet-201}, \text{ResNet-34}, \text{VGG-16}, \text{VGG-19}, \text{EfficientNet-B0}\}$
  Load pre-trained weights: Base Model Weights $\leftarrow$ Pre-trained
  Inputs: $\mathbf{x} \rightarrow$ Base Model Input
  Add Custom Layers:
  $z \leftarrow \text{Dense}(N, \text{activation='softmax'})$ ▷ N: number of activity classes
  Define Model: $\mathscr{M} = \text{Model}(\text{input} = \mathbf{x}, \text{output} = z)$
  Compile Model: $\mathscr{M} \leftarrow \text{Adam}(\alpha = 0.0004), \text{Loss} = \text{Categorical Crossentropy}$
  Train Model: $\mathscr{M}_{\text{trained}} \leftarrow \text{Train}(\mathscr{M}, \mathbf{I}_{\text{train}})$
  Estimate Training Time: $T_{\text{training}} = \text{End Time} - \text{Start Time}$
  Save Model Weights: $W^* \leftarrow W_t$, where $A_{\text{val},t} = \max\{A_{\text{val},1}, \ldots, A_{\text{val},T}\}$
end procedure
procedure ACTIVITY CLASSIFICATION
  Input unseen radar data $\mathbf{y}_{\text{new}} = \mathbf{I}_{\text{test}}$
  Preprocess: $\mathbf{I}_{\text{new}} = \text{Preprocess}(\mathbf{y}_{\text{new}})$
  Predict Activity: $\hat{\mathbf{a}} = \mathscr{M}_{\text{trained}}(\mathbf{I}_{\text{new}})$
  Output Classified Activity: $\mathbf{a} = \arg\max(\hat{\mathbf{a}})$
  Estimate Inference Time: $T_{\text{inference}} = \text{End Time} - \text{Start Time}$
end procedure

independent variables, time and frequency, simultaneously, which increases the computational load. As a result, SPWVD takes an average of 56 seconds to process a 5-second radar event. In comparison, simpler techniques such as RT, RD, and STFT complete preprocessing in approximately 0.5 seconds for the same event duration, as shown in Table 3.4.

Table 3.4: Generation time of radar maps.

Class	Data Length	Preprocessing Techniques	t_{pre} (s)
Walking (A1)	10 s	RT	0.97
		RD	0.88
		STFT	1.07
		SPWVD	113.00
Others (A2-A6)	5 s	RT	0.48
		RD	0.43
		STFT	0.51
		SPWVD	56.00

The large difference in t_{pre} highlights the trade-off between achieving high TF resolution and maintaining computational efficiency. While SPWVD provides superior TF resolution and can capture complex motion dynamics, its computational cost makes it less suitable for real-time applications. In contrast, RT and RD represent computationally efficient but may lose fine-grained motion details, which could affect recognition performance. The t_{pre} reported here were measured using a single-input single-output (SISO) radar sensor operating at a chirp rate of 1000 Hz, where each chirp has a duration of 1 ms. This configuration was chosen because it matches a typical R-HARS, where real-time processing is critical. The ability to efficiently preprocess radar data is particularly important for time-sensitive applications, such as fall detection in elderly care, where delayed event recognition could compromise safety.

3.5.2 Model Performance Across Radar Domains

To evaluate the impact of different preprocessing methods on feature representation, we first established baseline CNN model by training separate networks on each radar domain representation, as described in Section 3.4.2. Building upon these baselines, we then implemented and evaluated the TL models detailed in Section 3.4.3 to determine their effectiveness in improving recognition and classification accuracy across different radar representations. The baseline CNN models were designated as M1, M7, M13, and M19, corresponding to networks trained with RT, RD, STFT, and SPWVD preprocessing techniques, respectively. The TL models, when paired with each radar representation, formed a comprehensive set of 20 model-map pairs, systematically labelled as M2 to M6, M8 to M12, M14 to M18, and M20 to M24. A detailed summary of these model configurations, along with their performance metrics and computational costs, is presented in Tables 3.5 and 3.6.

Table 3.5: Performance metrics comparison of baseline CNN and TL models across four radar representations.

Model-Map Pairs	Maps	Models	Accuracy (%)			Precision	Recall	F1-score	Model Size (MB)	Memory Footprint (MB)
			Train	Valid	Test					
M1	RT	Baseline CNN	55.81	64.58	64.29	0.6553	0.6429	0.6347	78.19	26.03
M2		DenseNet-201	78.92	81.25	89.68	0.8993	0.8968	0.8968	213.61	70.93
M3		ResNet-34	88.02	88.69	89.68	0.9016	0.8968	0.8968	245.16	81.59
M4		VGG-16	92.61	91.07	92.06	0.9241	0.9206	0.9186	169.32	56.40
M5		VGG-19	93.31	92.26	91.27	0.9126	0.9127	0.9090	230.87	76.91
M6		**EfficientNet-B0**	**86.47**	**85.42**	**93.65**	**0.9383**	**0.9365**	**0.9366**	**50.75**	**16.73**
M7	RD	Basline CNN	62.60	65.48	72.22	0.7338	0.7222	0.7094	78.08	25.99
M8		DenseNet-201	85.06	85.11	84.92	0.8662	0.8492	0.8477	223.39	74.19
M9		ResNet-34	87.09	85.71	84.92	0.8457	0.8492	0.8440	246.11	81.91
M10		VGG-16	86.63	88.69	93.65	0.9372	0.9365	0.9360	169.32	56.40
M11		**VGG-19**	**90.90**	**90.48**	**93.65**	**0.9432**	**0.9365**	**0.9371**	**230.10**	**76.65**
M12		EfficientNet-B0	79.05	85.12	92.06	0.9189	0.9206	0.9195	48.86	16.10
M13	STFT	Baseline CNN	58.01	78.87	79.37	0.7887	0.7937	0.7867	78.19	26.03
M14		**DenseNet-201**	**98.75**	**97.02**	**96.03**	**0.9627**	**0.9603**	**0.9602**	**217.01**	**72.06**
M15		ResNet-34	91.06	95.54	93.65	0.9357	0.9365	0.9359	246.49	82.04
M16		VGG-16	96.35	95.54	95.24	0.9529	0.9524	0.9519	169.32	56.40
M17		VGG-19	97.67	95.83	95.24	0.9530	0.9524	0.9521	230.10	76.65
M18		EfficientNet-B0	94.79	93.15	94.44	0.945	0.9444	0.9442	48.86	16.10
M19	SPWVD	Baseline CNN	59.72	76.49	80.95	0.8119	0.8095	0.8011	151.98	50.62
M20		DenseNet-201	87.55	88.69	90.48	0.9080	0.9048	0.9038	224.92	74.70
M21		**ResNet-34**	**90.75**	**92.56**	**92.06**	**0.9184**	**0.9206**	**0.9188**	**245.16**	**81.59**
M22		VGG-16	91.91	93.75	91.27	0.9129	0.9127	0.9125	170.08	56.65
M23		VGG-19	90.20	92.26	90.48	0.9060	0.9048	0.9047	230.87	76.91
M24		EfficientNet-B0	81.57	89.58	88.10	0.8778	0.8810	0.8782	50.75	16.73

RT Domain

As shown in Table 3.5, applying TL models to RT representations significantly improves classification accuracy compared to the baseline CNN. Among them, EfficientNet-B0 (M6) performs well, achieving a peak recognition accuracy of 93.65%, a significant improvement of 29.36% over the baseline CNN model (M1), which achieves an accuracy of 64.29%. Fig. 3.9 shows this significant improvement, highlighting the ability of TL models to extract and effectively utilise features from RT representations. In addition, VGG-16 and VGG-19 perform well, achieving

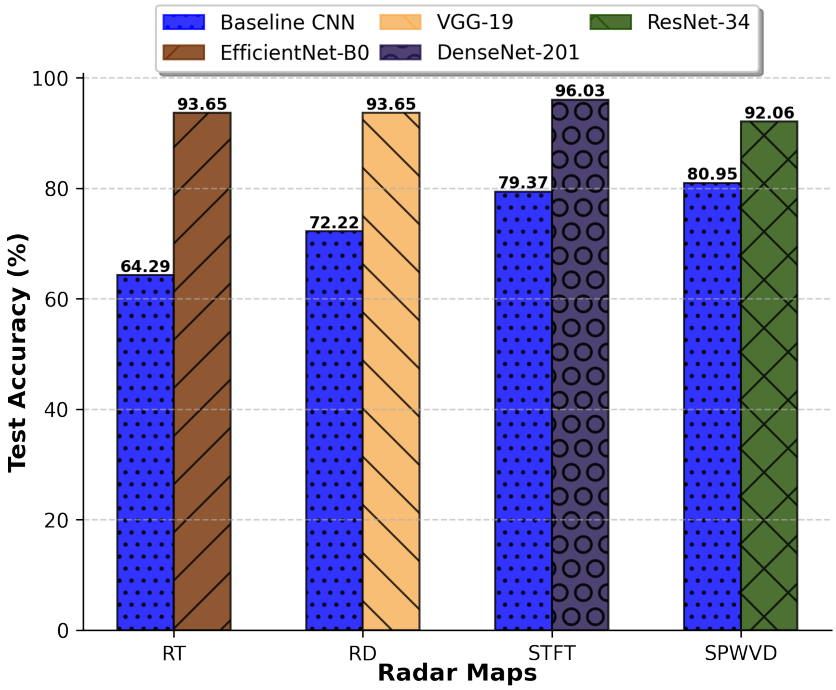


Figure 3.9: Comparison between baseline CNN and best performing TL models for each radar preprocessing techniques.

accuracies of 92.06% and 91.27%, respectively. However, denser and more complex models such as DensNet-201 and ResNet-34 are less effective on RT maps, suggesting that they may not be the most suitable choice for R-HARS that rely on RT preprocessing.

RD Domain

The RD maps also show significant gain when used as input to the TL model. The TL model improves the classification performance by 21.43% compared to the baseline CNN (M7) that achieves 72.22% accuracy, as shown in Fig. 3.9. The recognition accuracy for VGG-16 and VGG-19 reaches up to 93.65%, as illustrated in Table 3.5 . These results highlight the ability of the TL model to effectively exploit the detailed range and Doppler information embedded in the RD representation. Overall, the excellent performance of the TL model on the RD map suggests that this representation is a promising candidate for real-time, computationally efficient R-HARS.

STFT Domain

Combining the TL model with the STFT preprocessing method achieves consistently high classification accuracy across all tested architectures. DenseNet-201 (M14) emerges as the best performing model with a peak accuracy of 96.03%, demonstrating its ability to extract relevant TF features necessary for macro-activity recognition. This performance is significantly better than

the baseline CNN (M13), which achieves an accuracy of 79.37%, as shown in Fig. 3.9. These findings further confirm the effectiveness of the STFT as a robust TF representation, especially when paired with a well-optimised CNN architecture.

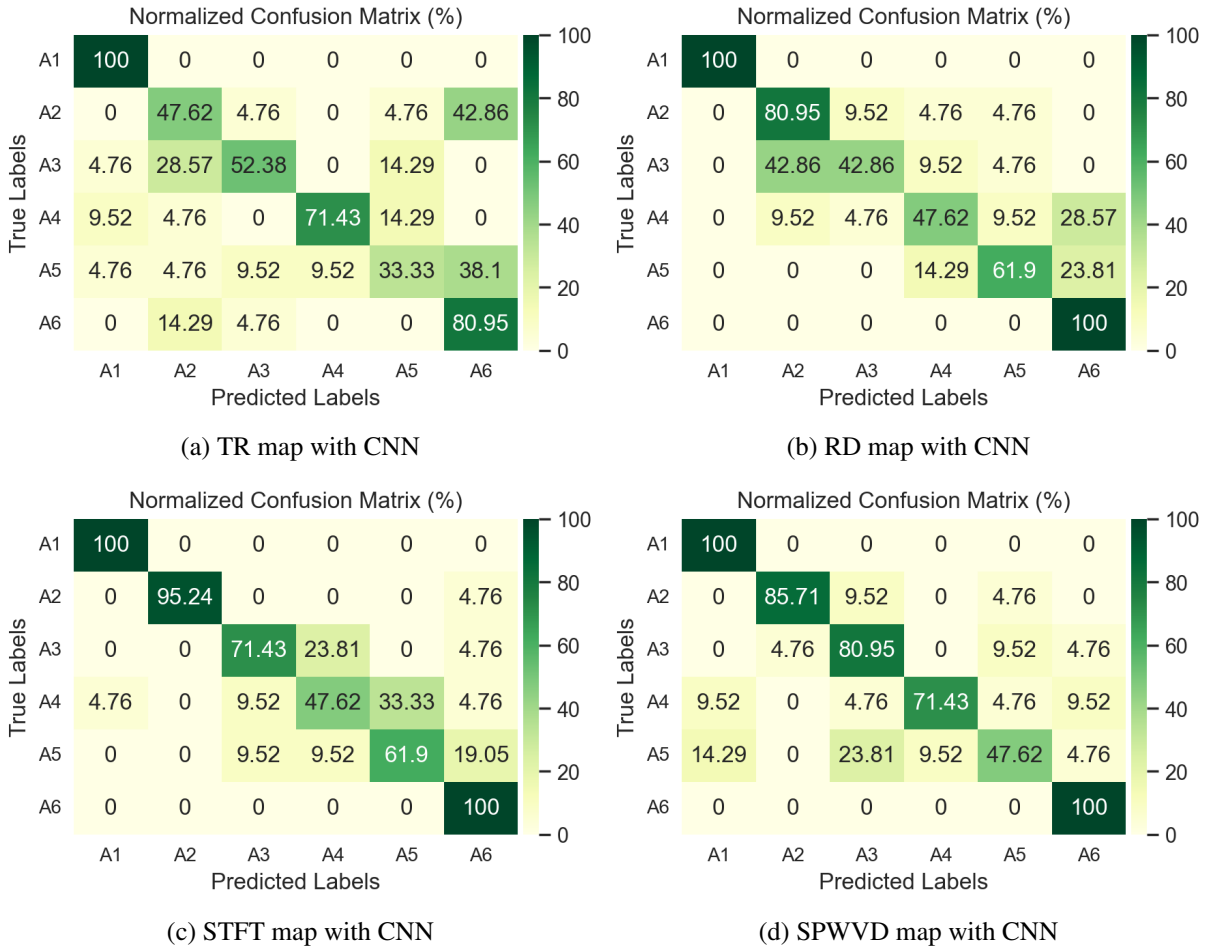
SPWVD Domain

Despite providing rich TF information, SPWVD-based representations do not always outperform other preprocessing methods (e.g., STFT and RD) when combined with TL models. As shown in the Fig. 3.9, the highest recognition accuracy in this category is ResNet-34 (M20), which achieves 92.06%. While this is an 11.11% improvement over the baseline CNN (M19) with an accuracy of 80.95%, it still does not reach the performance of STFT or even RT when fused with EfficientNet-B0.

A major contributing factor to this limitation is that SPWVD is more computationally expensive and takes longer to preprocess, as described in Table 3.4. Although SPWVD can capture detailed Doppler features, the added complexity does not necessarily translate into higher classification accuracy. As described in Table 3.5, model map pairs such as M21 and M23 achieve competitive accuracies of approximately 90.48% and 91.27%, respectively. However, their counterparts using STFT or RD generally outperform them, further confirmation the conclusion that while SPWVD is a rich representation, its practical benefits may be offset by its computational requirements.

3.5.3 Fall Event Detection and False Alarm Reduction

Accurately identifying fall events and minimising false alarms are critical requirements in R-HARS. Most radar domain preprocessing methods enable the baseline CNN to achieve high or perfect fall detection accuracy, as shown in Fig. 3.10. Notably, RT representation yielded 80.95% fall detection, whereas RD, STFT, and SPWVD achieved 100% fall detection. However, the RT-based CNN model exhibited a high false positive rate, with 42.86% of A2 (sitting) instances incorrectly classified as falls, indicating a significant overlap in signal features between these activities. TL models has led to substantial improvements, as seen in Fig. 3.11, the EfficientNet-B0 model trained on RT representation completely eliminated misclassification of A2 as falls. Overall, the TL models consistently outperformed the baseline CNN by enhancing inter-class separation and reducing misclassification rates, particularly in critical fall detection tasks. Among the high-resolution methods, SPWVD (Fig. 3.11d) demonstrated strong performance with high fall detection accuracy and reduced confusion with adjacent classes. Despite a slight residual misclassification of A5 (drinking) as A6 (fall), it remained significantly more robust than the baseline. Nevertheless, the computational overhead of SPWVD preprocessing remains a challenge, particularly for real-time deployment. These findings suggest that although RT is computationally efficient, its discriminative power significantly improves when



(a) TR map with CNN

(b) RD map with CNN

(c) STFT map with CNN

(d) SPWVD map with CNN

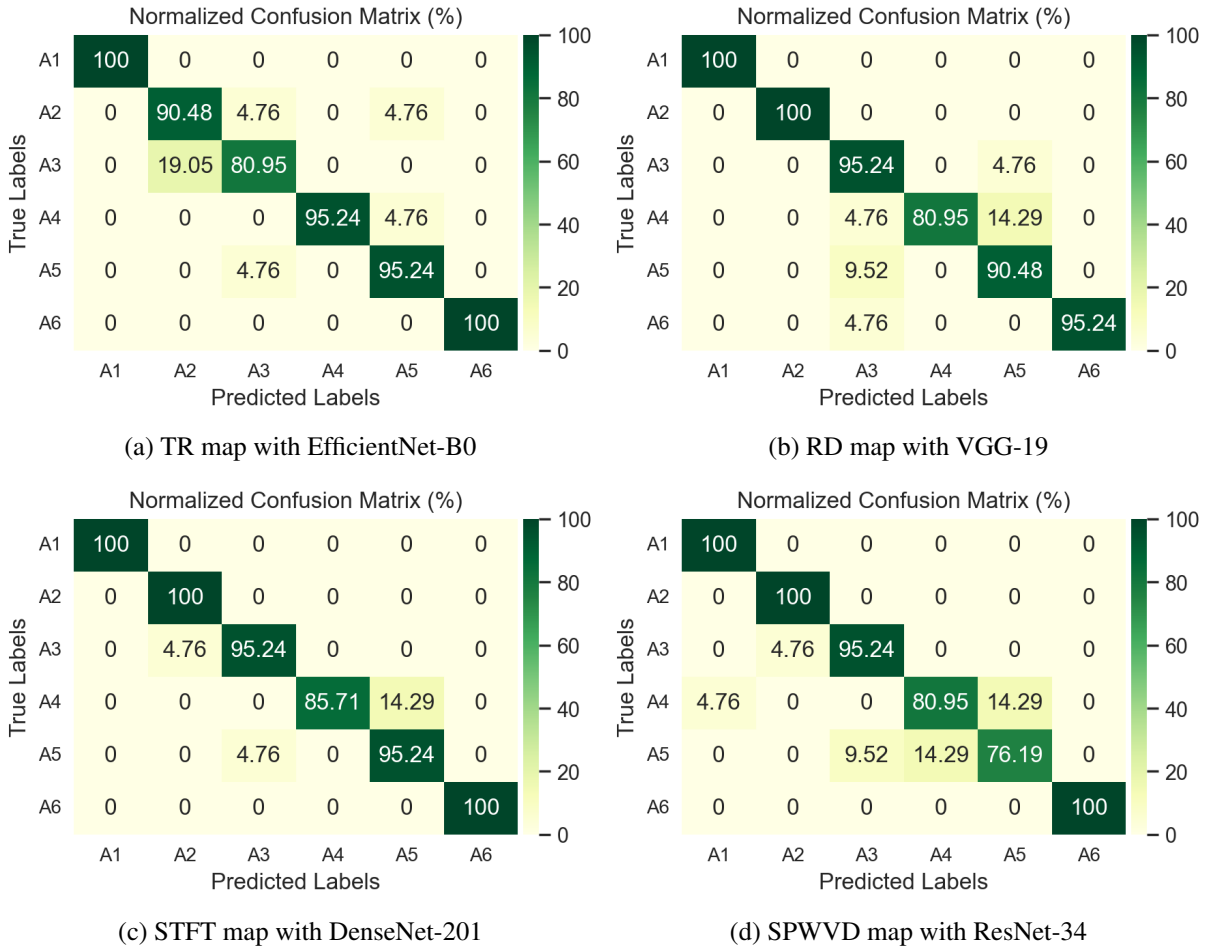
Figure 3.10: Confusion Matrices for the baseline CNN model using radar domains as input.

paired with advanced TL models. Similarly, STFT and RD representations combined with TL successfully minimised false positives, further reinforcing the benefits of integrating with radar data.

3.5.4 Analysis of Computational Efficiency

Computational efficiency is a key factor in real-world settings, especially in terms of training and inference time. Training time determines how quickly a model learns from a radar data, while inference time measures how quickly predictions are made. To evaluate training efficiency, we examined the time required per epoch for each model to process a specific radar domain representation, as illustrated in Table 3.6. Among the evaluated models, the baseline CNN (M7) trained with the RD representation exhibited the shortest training time, taking only 6.71 seconds to complete one epoch. While the training duration varied between the TL models due to architectural differences, the uniform input image size of 224 × 224 pixels ensured relatively consistent training times for all radar representations.

As shown in Table 3.6, the average inference time varied little, ranging from 31.14 ms (SP-



(a) TR map with EfficientNet-B0

(b) RD map with VGG-19

(c) STFT map with DenseNet-201

(d) SPWVD map with ResNet-34

Figure 3.11: Confusion Matrices for the selected best TL models using radar domains as input.

WVD) to 32.86 ms (STFT). SPWVD has the smallest fluctuation in inference time, as evidenced by its smaller standard deviation (STF) and standard error (SE). Despite the higher preprocessing overhead, SPWVD's inference latency remains comparable to other techniques. This consistency can be attributed to the standardised input size (224×224 pixels), which helps mitigate the computational complexity differences between different preprocessing methods, thereby ensuring consistency in prediction regardless of the model selected.

VGG-16 based models (M4, M10, M16, and M22) exhibit relatively low inference times while maintaining strong recognition accuracy, making them particularly suitable for latency-sensitive applications. In contrast, DenseNet-201 based models (M2, M8, M14, and M20) require longer inference times due to their deeper architectures and increased computational requirements. Although DenseNet-201 achieves the highest classification accuracy, its extended processing time may pose a challenge in scenarios where fast decision making is critical. While inference latencies below 100 ms may be negligible for event windows of 5 seconds, these latencies may become significant when deploying models on low-power edge devices with limited computational resources. Models based on EfficientNet-B0 (M6, M12, M18, and M24) achieve

Table 3.6: Comparison of radar maps with different TL based networks and their computational costs.

Pairs	Training Time/epoch (s)	Inference Time/sample (ms)	STF (ms)	SE (ms)
M1	8.55	31.99	5.78	0.51
M2	22.18	71.17	7.35	0.65
M3	7.81	46.89	11.10	0.98
M4	9.85	38.19	2.63	0.23
M5	11.76	39.93	5.62	0.50
M6	10.84	55.31	9.79	0.82
M7	6.71	31.65	2.31	0.20
M8	21.31	70.42	8.12	0.72
M9	7.57	42.02	7.21	0.64
M10	9.77	37.82	3.20	0.28
M11	11.77	39.09	5.55	0.49
M12	10.72	51.47	5.41	0.48
M13	8.74	32.86	2.69	0.23
M14	22.36	70.68	7.18	0.63
M15	7.86	45.38	9.74	0.86
M16	10.13	37.14	2.85	0.25
M17	11.78	39.69	5.76	0.51
M18	10.85	52.82	9.61	0.85
M19	7.03	31.14	1.86	0.16
M20	22.36	70.68	7.48	0.66
M21	8.00	43.04	7.35	0.65
M22	10.02	37.29	2.75	0.24
M23	11.89	40.54	4.86	0.43
M24	10.53	52.08	9.72	0.86

a balance between accuracy and computational efficiency. These models provide competitive classification performance while maintaining moderate inference times, making them a practical choice for applications where accuracy and real-time performance are equally important, as shown in Table 3.6.

3.5.5 Implications for Real-world Deployment

For real-world deployments, particularly for low-power edge devices used in applications such as elderly monitoring and smart security systems, achieving a balance between recognition accuracy and computational efficiency is essential. Systems must make prompt decisions to prevent adverse outcomes. To quantitatively assess this trade-off, we define a simple balance metric which can be computed as: $S_{\text{bal}} = \frac{\text{Accuracy (\%)}}{t_{\text{pre}}+t_{\text{infer}}}$. This metric is used to evaluate the efficiency of each model-domain pair (MDP) by jointly considering recognition performance and computational latency.

The top-performing MDPs, along with their computed S_{bal} values and corresponding deployment suitability, are listed in Table3.7. This metric reflects the accuracy per unit inference latency, where a higher score indicates a more efficient balance between performance and re-

Table 3.7: Top model-domain pairs with highest balanced scores and their suitability for deployment.

MDP	S_{bal}	Best Use Case
M6	91.34	Real-time applications where moderate processing time is acceptable.
M11	**101.89**	**Lowest total processing time; ideal for latency-critical scenarios.**
M14	84.19	Maximum accuracy where moderate latency is acceptable.
M21	0.81	Offline analysis; not suitable for real-time deployment.

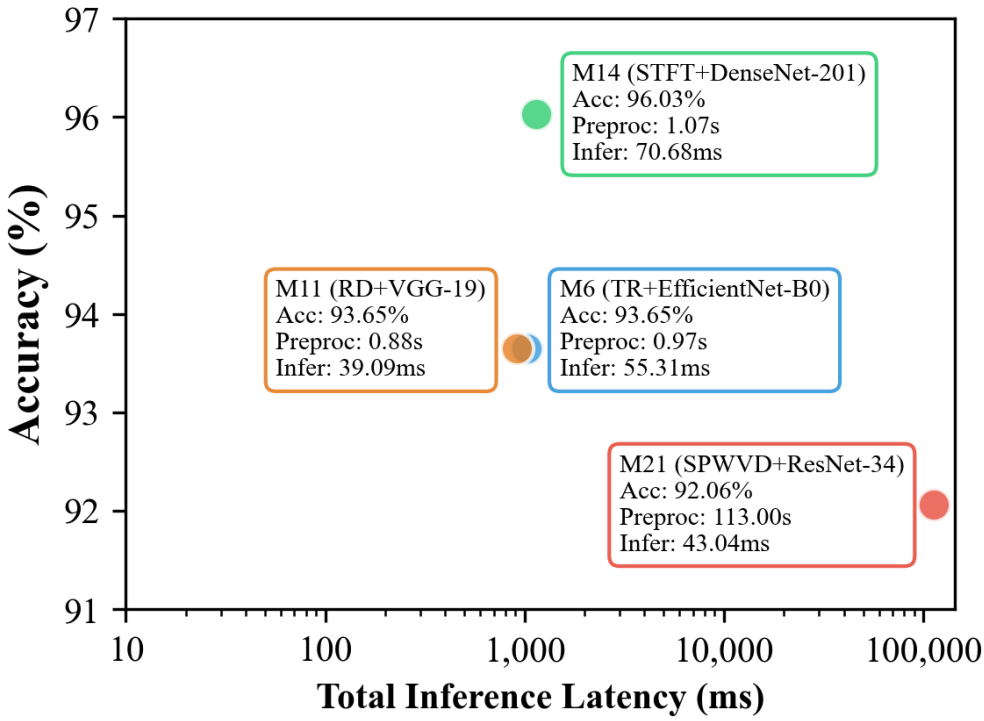


Figure 3.12: Accuracy vs. total inference latency for top MDPs. Here Acc = Accuracy, Preproc = t_{pre}, Infer = t_{infer}.

sponsiveness. The detailed performance and inference latency characteristics are shown in Fig. 3.12. Different MDPs are optimal for different application contexts. M11 (RD + VGG-19) achieved the highest score of 101.89, making it ideal for low-latency, real-time systems. M6 (TR + EfficientNet-B0) offers a strong balance between speed and accuracy, whereas M14 (STFT + DenseNet-201) delivers maximum recognition accuracy with a modest latency overhead. In contrast, M21 (SPWVD + ResNet-34) suffers from an extremely long S_{bal}, rendering it unsuitable for responsive deployments. This evaluation framework allows system designers to move beyond accuracy only comparisons and incorporate inference latency aware decision making. Developers can use S_{bal} to objectively assess whether a given MDP offers an efficient trade-off for their deployment scenario.

3.5.6 Cross-Frequency Validation of the Proposed Framework

To assess the generalisability of our framework beyond a single dataset and radar configuration, we selected the two most suitable radar representations, RD and STFT maps, paired with their best-performing models, VGG-19 and DenseNet-201, respectively. We utilised two publicly available datasets collected using a 24 GHz Ancortek radar and a 77 GHz mmWave FMCW radar, which are detailed in [213] and listed as item 5 in our dataset explanation (see Section 3.3.1). These datasets contain 11 daily activities recorded from six subjects, with each subject performing 10 repetitions of each activity. This initially provided 660 samples; however, after removing corrupted raw files, 600 samples were used for the analysis.

Owing to the limited number of subjects and samples, we implemented a Leave-One-Subject-Out Cross-Validation (LOSO-CV) strategy, which aligns well with our subject-based evaluation approach used for the primary dataset. For each of the six folds, we trained on five subjects and tested on the excluded subject, maintaining our commitment to evaluating performance on completely unseen individuals. The raw radar data were processed into RD and STFT representations using identical preprocessing pipelines as applied to our primary 5.8 GHz dataset. We then trained the VGG-19 and DenseNet-201 models using the same hyperparameters and training procedures established in our primary experiments.

Table 3.8 presents the cross-frequency generalisation performance of the optimal MDPs. The reported accuracies represent the average performance across all six test subjects with their STD. The results demonstrate that our preprocessing techniques and TL models maintain strong performance across different radar frequencies, with STFT as domain and DenseNet-201 as model, achieving particularly impressive recognition accuracy of 96.42%±2.12 using 24 GHz radar raw data. This cross-frequency validation confirms the robustness of our approach and its potential applicability in diverse R-HARS, regardless of the operating frequency.

Table 3.8: Generalisation performance of optimal radar domain-model pairs across different frequencies using LOSO-CV.

Domain	Model	Freq.	Acc. (%)	Precision	Recall	F1-score
RD	VGG-19	24 GHz	86.74±2.20	0.8772	0.8674	0.8655
		77 GHz	89.70±2.21	0.9074	0.8970	0.8974
STFT	DenseNet-201	24 GHz	96.42±2.12	0.9658	0.9642	0.9641
		77 GHz	91.71±1.31	0.9228	0.9171	0.9164

Overall, the quantified performance across both radar frequencies was strong. In particular, the STFT maps with DenseNet-201 model achieved an excellent accuracy of 96.42% at 24 GHz and maintained a high accuracy of 91.71% at 77 GHz. These values are considered very good for R-HAR tasks, particularly given the challenging cross-frequency validation setup. The RD maps with VGG-19 model also demonstrated reliable performance, achieving 86.74% at 24

GHz and improving slightly to 89.70% at 77 GHz. The decrease in accuracy for the STFT maps with DenseNet-201 pair at 77 GHz can be attributed to multiple factors. Higher-frequency radar signals are more sensitive to environmental clutter, noise, and multipath interference, which can degrade signal quality. They also suffer from reduced penetration and effective range compared to lower frequencies like 24 GHz. In addition, the 77 GHz dataset includes more variability and slightly lower-quality samples, which may further challenge model generalisation. These aspects collectively contribute to the modest but consistent performance reduction observed at 77 GHz.

3.6 Summary

This chapter presents a comprehensive evaluation of DL models for non-intrusive macro-activity recognition using various radar-domain representations. The analysis focused on balancing the classification performance with computational efficiency to address challenges **C1** and **C2** outlined in Section 1.3. To mitigate the challenge of limited data diversity in radar-based activity recognition, this study transformed raw signals from a single radar sensor into multiple domain representations, including range-time (RT), range-Doppler (RD), and time-frequency (TF) representations using STFT and SPWVD. This multi-domain preprocessing strategy enriches the input feature space, enabling the identification of the most informative representations for real-world deployment scenarios. The study also demonstrated effective model domain adaptation through TL, wherein pre-trained architectures originally developed for natural image classification were successfully adapted to radar-based inputs. This approach has proven valuable in scenarios in which labelled radar data are scarce. Analysis of the preprocessing time revealed that high-resolution techniques, such as SPWVD, capture rich Doppler information but impose a significant computational burden. In contrast, RT and RD maps offer a favourable trade-off, achieving faster preprocessing with competitive classification accuracy.

Among the evaluated models, TL approaches, particularly EfficientNet-B0, VGG-16, VGG-19, and DenseNet-201, consistently outperformed the baseline CNN across all radar domains. EfficientNet-B0 has emerged as the most balanced architecture in terms of accuracy and inference latency, making it a strong candidate for real-time applications. Although deeper networks, such as DenseNet-201, achieved higher accuracy, their increased inference time reduced their suitability for real-time deployment. Models such as VGG-16 and EfficientNet-B0 were found to provide a better balance between computational cost and recognition performance, which is critical for practical use. Multi-frequency validation using datasets collected at 24 GHz and 77 GHz confirmed that the proposed preprocessing pipeline and TL models generalise effectively across different radar frequencies. This reinforces the adaptability of the framework to varied sensing conditions and its potential for integration into multiple radar platforms. Fall detection analysis reveals that TL models effectively reduce false alarms while improving system

reliability, establishing a foundation for more accurate and reliable HARS in safety-critical applications, such as AAL. Class imbalance remains a common challenge in R-HAR, especially for critical events such as falls. In this study, the limited fall data reflects the ethical and practical constraints of collecting such samples from elderly participants. Although data augmentation and class weighting were used to reduce their impact, future work should aim to expand the datasets for better representation of minority classes.

The comprehensive analysis presented in this study, while demonstrating the effectiveness of various radar processing techniques and TL approaches, has been conducted primarily on GPU-based systems with high computational demands and large models with millions of parameters. These resource-intensive requirements pose significant challenges for the real-world deployment of resource-constrained edge devices, where memory limitations, power constraints, and processing capabilities are severely restricted. Having established the optimal radar signal processing techniques and identified the most promising model architectures for macro-activity recognition, the next critical step involves adapting these findings for edge-based deployment scenarios that maintain recognition accuracy while meeting the stringent computational constraints of edge-computing environments.

Chapter 4

Radar-based Vital Sign Monitoring

This chapter investigates the use of radar technology for vital sign monitoring by focusing on micro-activity monitoring in healthcare applications, thereby addressing challenge **C1**. As outlined in Section 1.3, one of the core contributions of this thesis is enabling non-intrusive multi-scale HAR using radar sensor. While radar has already proven effective for macro-activity recognition, such as walking, sitting, and fall detection, it also offers a unique capability to detect micro-activities that reflect the vital signs of the human body. These micro-activities include subtle physiological movements such as chest wall displacements caused by respiratory cycles and cardiac activity, which correspond to respiratory rate (RR) and heart rate (HR), respectively. These vital signs serve as key indicators of health status and are integral to continuous monitoring in both clinical and non-clinical settings. The ability of radar to unobtrusively capture both physical and physiological signals positions it as a versatile solution for multi-scale HAR applications. This chapter examines two radar systems specifically designed for micro-activity estimation and details the robust signal processing strategies employed to accurately extract and estimate HR and RR under real-world conditions.

4.1 Introduction

Accurate monitoring of human micro-activities such as HR and RR, is essential for safeguarding health, particularly in vulnerable populations. Patients with chronic conditions such as sleep apnoea require continuous physiological monitoring, while individuals trapped during natural disasters may need to be detected through barriers such as walls or debris. Similarly, elderly patients and infants in intensive care units (ICUs) benefit from contactless systems because conventional wearable or wired sensors can cause discomfort, skin irritation, or restricted mobility [222]. To address these needs, radar-based biomedical sensing has emerged as a powerful solution, capable of remotely capturing vital signs by detecting subtle physiological movements, such as chest wall displacements. Recent progress in this area includes the use of a 24 GHz radar sensor for tracking breathing patterns in patients with sleep apnoea [223]. These advancements

lay the groundwork for broader healthcare applications, particularly in elderly care and critical care environments, where continuous and comfortable monitoring is vital for patient safety.

The underlying principle of radar-based vital sign monitoring relies on the transmission and analysis of electromagnetic signals that interact with the human body. When a radar transmitter emits periodic signals, these waves reflect off targets within the radar cross-section (RCS) and return to the receiver with modified characteristics that encode information about target movement and displacement [80]. In biomedical applications, even tiny movements, such as chest expansion during breathing or surface vibrations from the heartbeat, can cause detectable changes in the reflected signals. By strategically positioning radar sensors within clinical environments, healthcare providers can establish ambient-intelligence platforms that automatically extract health-related data streams without requiring patient cooperation or awareness [25]. This approach marks a significant shift from traditional monitoring methods, allowing for the creation of smart healthcare environments in which diagnostic information is gathered seamlessly and integrated into comprehensive patient care systems.

Over the past few decades, three main types of radar systems have been employed for vital sign monitoring: pulse radar (also known as UWB radar) [224], CW Doppler radar [225], and mmWave FMCW radar [25,226]. In one study [224], a 3.3 GHz UWB radar was used to extract the HR and RR by applying a time-varying filter to reduce noise, although the detection accuracy was not evaluated. While UWB radar can effectively capture physiological signals, it faces challenges related to system integration and low-power operation due to its high bandwidth requirements. In contrast, CW and FMCW radars offer better integration, lower power consumption, and improved performance in proximity detection. Among these, the FMCW radar stands out for its high range and Doppler resolution, as well as its stronger resistance to noise compared to pulse-based systems. A key advantage of the FMCW radar is its ability to distinguish reflections from different distances, making it suitable for monitoring the vital signs of multiple individuals simultaneously which is a major strength highlighted in recent research [178].

In this study, we employed two radar sensors, UWB radar and mmWave FMCW radar, for vital sign estimation. These radars were used for data collection and analysis, leveraging their unique characteristics to improve the accuracy and stability of contactless vital sign monitoring. The performance of both radar systems is affected by factors such as the distance between the radar and the target, as well as the position and orientation of the target relative to the radar. Several studies have explored the use of mmWave radars for HR monitoring [25,178]. However, most of these approaches require the subject to remain seated or lie at a fixed distance, as even minor body movements can exceed the amplitude of chest displacements, making accurate HR detection more challenging. In contrast, this study proposes an mmWave radar system designed to collect data from individuals in a standing position, with the radar placed at varying distances and orientations to evaluate performance under more flexible and realistic conditions. When designing radar-based healthcare application for vital sign detection, it is important to validate

the extracted biomedical signals against a clinically established gold standard. Therefore, in this study, signals from both radar systems were recorded simultaneously using a gold standard medical device. This approach enabled quantitative correlation analysis of the physiological signals extracted by the radar and the reference measurements. This validation step ensured reliability, accuracy, and clinical applicability in real-world healthcare scenarios.

4.2 Contributions

Radar-based vital sign monitoring relies on the reflection of electromagnetic signals from the human body to measure HR and RR. The phase shift of the reflected signal is proportional to the distance the signal has travelled [25]. By analysing this phase change, the target distance can be determined and vital sign information can be extracted. Therefore, an important factor to consider in radar-based vital sign monitoring is the operating distance of the radar and the signal processing techniques used to detect chest wall motion and target presence. These detected signals are then processed to extract and estimate human micro-level activities. In summary, the main contributions of this work are highlighted as follows:

- This study investigated the effect of distance on the accuracy of radar-based vital sign estimation using UWB radar signals to reliably estimate RR, and HR. An empirical study was conducted using a single subject to evaluate the effectiveness of the system, with data collected at three different distances, and validated using medical-grade equipment. A simple traditional signal processing method based on the FFT was used to determine the HR and RR. Additionally, radar-based features extracted from the raw data were used to identify and classify three breathing patterns: hypopnoea (or breath-holding), normal breathing, and shallow or elevated breathing. This work demonstrates the feasibility of using UWB radar for vital sign monitoring and lays the foundation for further development.

- A major advance in this study is the transition from UWB radar to more sophisticated FMCW radar operating in the mmWave spectrum. Due to its superior raw data conversion capabilities, this radar technology provides enhanced capabilities including target detection, distance measurement, and high data rates. These attributes make mmWave radar particularly suitable for remote monitoring applications such as elderly care and patient monitoring. An advanced signal processing framework was developed to extract and estimate vital signs from radar data.

- To address a critical research gap, this study extensively evaluated mmWave radar at different height positions (including one meter, chest height, and two meters) to determine the optimal radar position for accurate estimation of vital signs. In addition, a unique experimental setup was designed in which the target remained in a standing posture while the radar position was changed. This provide valuable insights into the ideal radar placement

for optimal performance. This study was conducted using an advanced technology setup incorporating a DCA1000 data acquisition module. This module enables real-time data transmission and seamlessly interfaces with emerging 5G and B5G network standards, making it a promising solution for future wireless medical applications.

4.3 Vital Sign Monitoring Using Ultra-wideBand Radar

UWB radars offer several advantages over microwave Doppler radars, including lower human exposure risk, and higher signal-to-noise ratio (SNR). These advantages make UWB radars ideal for contactless vital sign estimation, especially in healthcare and biomedical applications [224]. Recently, there has been a growing interest in contactless physiological monitoring using UWB radars [172, 173]. Vital signs are estimated by detecting the time-of-flight changes of narrow pulses emitted by the radar, which are reflected from the human chest and return to the receiver. In addition to vital sign monitoring, UWB radars have also been used for human localisation and gait analysis [227], demonstrating their versatility in HAR applications. Despite these advantages, radar-based monitoring techniques are susceptible to low SNR and clutter effects, as reported in recent studies [228]. To address this challenge, we used a commercial off-the-shelf (COTS) UWB radar device, a Walabot, which has recently emerged as an affordable solution with potential RR monitoring capabilities [229].

4.3.1 System Model and Preliminaries

This section presents an overview of the Walabot radar system used for contactless vital-sign monitoring. It outlines the operating principles, technical specifications, and medical-grade reference sensor employed for validation. The system's performance was assessed by comparing its output with that of the reference sensor to evaluate its effectiveness as a non-invasive alternative to traditional wearable devices. Fig. 4.1 illustrates the complete workflow of the proposed system, which includes radar signal acquisition, raw data processing, and vital sign estimation.

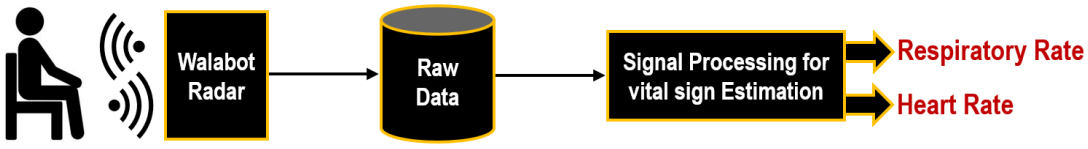


Figure 4.1: The Walabot radar captures raw data, which is processed to estimate vital signs.

Overview of UWB Radar

The Walabot developer version [230] used in this study for vital sign monitoring, is a 3D imaging sensor that employs UWB technology, operating in the frequency range of 5–8 GHz, with a

transmit power of 5dBm [231]. Its core is the Vayyar YYR2401 A3 system-on-chip (SoC) integrated circuit (IC), supported by an array of 18 linearly polarised broadband antennas, as shown in Fig. 4.2, [231]. The radar is powered by simply connecting a USB wire to the host device, which is also used to send data between the radar and the host device using the cypress controller, as shown in Fig. 4.3A. The given frequency range is good enough to detect direct distance within 10 meters based on the gradient of radar pulse. The transmit chirp of Walabot radar was characterised by [230]:

$$x(t) = A\sin\left(2\pi\left(\frac{f_1 - f_0}{2T}(t-t_0)^2 + f_0(t-t_0)\right)\right) \tag{4.1}$$

Where A represents the amplitude of the linear frequency, f_0 represents the starting frequency of the linear frequency modulation, f_1 represents the ending frequency, t_0 represents the starting time and T corresponds to the linear frequency modulation period.

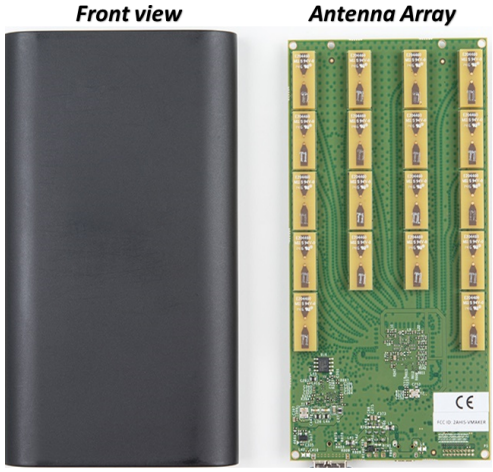


Figure 4.2: Walabot radar sensor with front view showing the black outer shell (left) and an internal antenna array board comprising 18 linearly polarised antennas (right).

Although the Walabot radar is widely used for applications such as detecting wires and pipes behind walls, its design also offers distinct advantages for vital sign monitoring [232]. Its user-friendly design and respiratory monitoring capabilities made it a practical choice for this study. Walabot's compact, integrated architecture combines the transmitter and receiver into a single unit, simplifying the system setup and minimising the alignment challenges. This makes it particularly well suited for research on user positioning and system configurations in real-world scenarios. Walabot used a parameter called a "Profile" to identify which antenna pairs to use for a specific application. Four profile options are available: short-range imaging, trackers, sensors, and sensor narrow. The sensor narrow profile provides lower resolution and higher capture rates, making it particularly suitable for applications involving motion detection, including RR monitoring [231].

Radar Arena Setup & Calibration: Arena is a term used to describe the area scanned by a device. To ensure a precise operation, the orientation of the arena must follow specific guidelines, as shown in Fig. 4.3. Using Cartesian coordinates parameters X, Y, and Z, Arena are defined for short-range imaging profiles, which typically correspond to the size of the device [233]. X and Y parameters were defined from -10 to 10 cm, whereas Z was defined from 1 to 20 cm. For all other profile categories, the Arena is specified using spherical coordinates parameters: R, *theta* (θ), and *phi* (ϕ) [231].

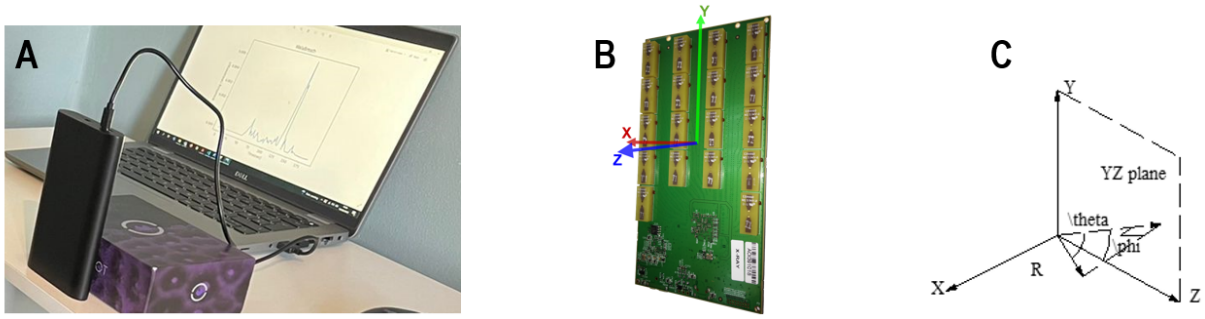


Figure 4.3: Walabot Radar and coordinate systems: A) Connection between the Walabot and host device, B) Cartesian coordinates from the antenna array, C) Cartesian and spherical coordinates.

In this context, R specifies the Z range (1 to 1000 cm), while θ and ϕ determine the horizontal (Y-axis) and vertical (X-axis) angles ranging from 0 to 90 degrees, respectively. When developing a new application, the user can choose to configure the arena by using either Cartesian or Spherical coordinates. The application automatically converts spherical coordinates to Cartesian coordinates when required, using the following equations [233]:

$$X = R \cdot sin\theta \tag{4.2}$$

$$Y = R \cdot cos\theta \cdot sin\phi \tag{4.3}$$

$$Z = R \cdot cos\theta \cdot cos\phi \tag{4.4}$$

Regardless of the coordinate system used, each parameter was associated with a resolution parameter that was configurable in the range of 0.1 to 10. The resolution parameter determines the validity of related parameters. The ranges of all the parameters depend on the specific features used [231].

The calibration process must be performed before starting the application, as calibration ensures that the device ignores walls and reflections. It should be performed in a quiet environment with no moving objects to improve accuracy. Proper calibration of the Walabot radar prior to sensing improves detection reliability [233]. Table 4.1 lists the key parameters configured for the Walabot radar during data acquisition, employed in this study. These parameters define the radar's initialisation, sensing range, and filtering techniques to ensure accurate detection within

the specified range.

Table 4.1: Configured parameters of the Walabot radar during calibration.

Parameter	Configured Value
API to Read Signals	GetImageEnergy()
Arena Parameters	Wlbt.SetArenaTheta(-0.1, 0.1, 1) Wlbt.SetArenaPhi(-0.1, 0.1, 1) Wlbt.SetArenaR(20, 80, 0.2)
Applied Filter	Wlbt.SetDynamicImageFilter (Wlbt.FILTER_TYPE_DERIVATIVE)

Radar Specifications & Parameters

Table 4.2 lists the specifications and parameters of the Walabot radar during the vital sign experimental setup.

Table 4.2: Radar specifications for collecting vital sign data

Parameters	Values
Center Frequency	6.5 GHz
Bandwidth	3 GHz
PRF	53.5 MHz
Range	0 - 10 meters
Range Resolution	Up to 2mm
Angular Resolution	Up to 2mm
Frames Rate	1-10 frames per second
Sampling rate (fs)	4 Hz

Medical-Grade Reference Sensor

In our experiment, we used a respiration belt from GDX-RB, Vernier, USA [234], as shown in Fig. 4.4, as the reference sensor to measure RR that worn around the upper abdomen, just below the chest. The belt is highly sensitive and accurate, with a frequency response range of 10 Hz, making it suitable for capturing low-frequency vibrations caused by breathing. Upon receiving the data, a Python script was used to conduct a frequency spectral analysis using the FFT technique. This would determine the dominant frequency component of the force signal, which corresponds to the subject's RR. This allows us to establish a reliable reference signal for validating the RR obtained from the radar, by comparing the RR measurements obtained from both sensors, we can assess the performance of the radar.

Figure 4.4: Gold-standard medical grade respiration belt that worn around the upper abdomen.

Experimental Setup

The Walabot radar was initialised using Python programming and the API's to establish a connection between the radar and the PC via a micro USB cable, as shown in Fig. 4.5A. Once the calibration process was completed, the sensing area (Arena) was configured, and the system was ready for data collection. During data acquisition, each captured signal was stored in a comma-separated value (CSV) file for easy post-processing and analysis. To extract vital sign information, breathing energy was generated using the GetImageEnergy() function from the Walabot API. In addition, various API functions were used to configure the system and receive signals, as summarised in Table 4.3.

Table 4.3: Different API's of the Walabot during data collection of vital sign estimation

Walabot API's	**Description**
Walabot Init()	To initialise the Walabot for start
Walabot Connect()	Start communication with Walabot
Walabot Set Parameters()	Set the Arena size
Walabot Calibrate()	Calibrate the Walabot
Walabot Start()	Start the system for capturing
Walabot Trigger()	Scan according to profile and record the signal
Walabot stop()	Stop capturing data
Walabot disconnect()	Disconnect from the running phase

Data Collection

The experiments were conducted in an open room with an area of 16×13 square feet, as shown in Fig. 4.5B, which shows a top view of the experimental setup. The data collection process was divided into two phases to ensure a comprehensive validation of the UWB radar-based micro-activity monitoring.

In the first phase, data was collected at three different distances between the radar and the participant: 55, 60, and 65 cm. The experimental setup is shown in Fig. 4.5A, where the participant is sitting 60 cm away from the radar sensor, while wearing a medical grade respiration belt

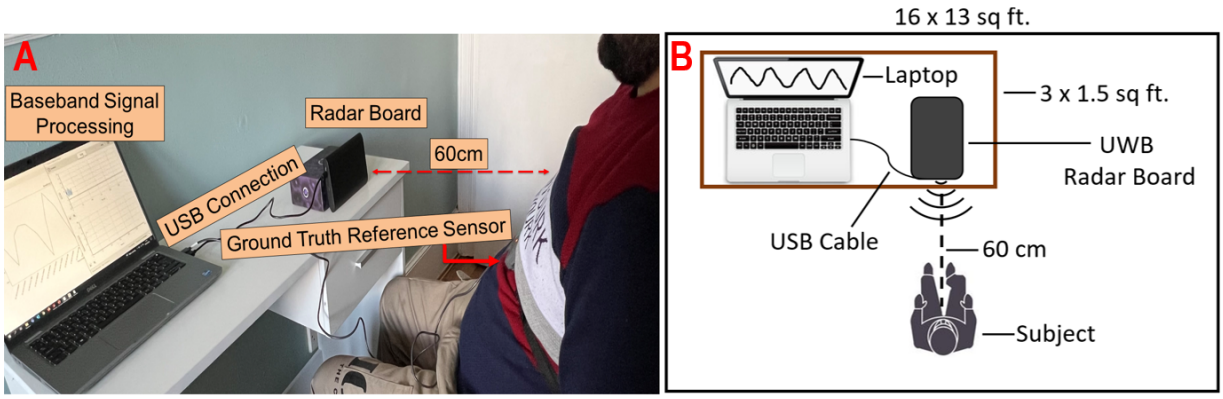


Figure 4.5: Experimental Setup: A) A person seated 60 cm in front of the radar during data collection, B) Top view of the experimental setup.

as a reference sensor. Each experiment lasted one minute, during which the radar and the reference sensor simultaneously recorded the breathing signal to ensure accuracy and consistency. The sampling frequency of the Walabot radar was configured to 4 Hz to capture subtle chest movements associated with breathing.

In the second phase, data were collected while the participant maintained a fixed distance of 60 cm from the radar. This phase assessed the system's ability to detect different breathing patterns. Participants performed the following breathing patterns:

- Slow breathing
- Normal breathing
- Elevated breathing

Each breathing pattern was maintained and collected for one minute, followed by a one minute break before transitioning to the next pattern. Throughout this phase, the participant continued to wear a medical grade respiration belt, ensuring reliable ground truth validation. This protocol was designed to evaluate the radar accuracy across varying respiratory conditions and its potential for real-world deployment. Throughout the data collection process, a real-time breathing pattern graph was displayed on the laptop screen to provide visual feedback on the respiratory activity, as depicted in Fig. 4.6. The breathing signal was recorded simultaneously during post-processing. FFT based signal processing techniques were applied to extract and estimate RR and HR.

4.3.2 Data Preprocessing for Vital Sign Estimation

The entire pipeline for preprocessing the radar data to estimate HR, RR, and breathing patterns is illustrated in Fig. 4.7. This diagram demonstrates how raw radar data undergoes systematic signal processing to extract vital signs with high accuracy. The process begins with the acquisition of raw radar signals, which capture complex, multidimensional reflections from the human

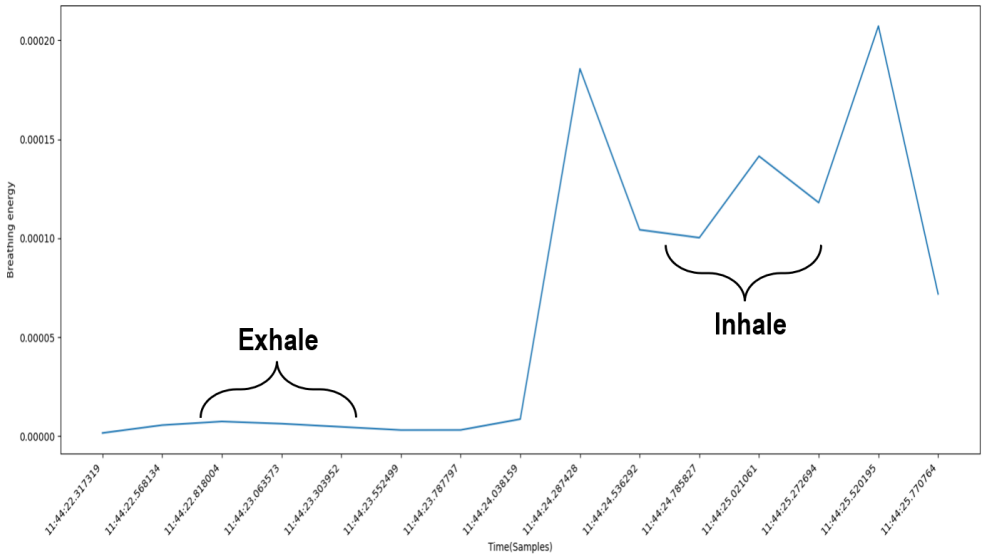


Figure 4.6: The capture of a Real-time graph that shows breathing application.

chest due to micro-movements caused by breathing and heartbeat. However, these signals are not directly usable for frequency-based analyses because of their complex structures. To prepare the data for processing, a flattening operation was applied to convert the radar reflections into one-dimensional time series. This transformation ensures that all relevant physiological information is structured into a format suitable for subsequent filtering and spectral analysis, as shown in Fig. 4.8.

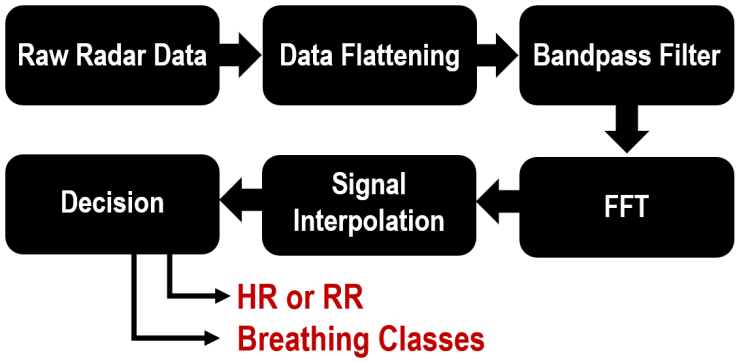


Figure 4.7: Signal processing pipeline for vital sign estimation from raw radar data.

For filtering, we employed a second-order Butterworth bandpass filter with different cutoff frequencies for specific estimations. For RR estimation, the cutoff frequency was set to a range of 0.16 – 0.33 Hz to capture respiration oscillations. For the HR estimation, a higher frequency range 0.8 – 2.0 Hz was used to isolate the heartbeat signal. The application of these bandpass filters ensures that the extracted signal components correspond only to respiratory and cardiac activity, removing disturbances from random body movements or background noise. To verify the effectiveness of the bandpass filter, its frequency response was computed using freqz() function in Python, ensuring that the selected cutoff frequencies correctly isolated respiratory or

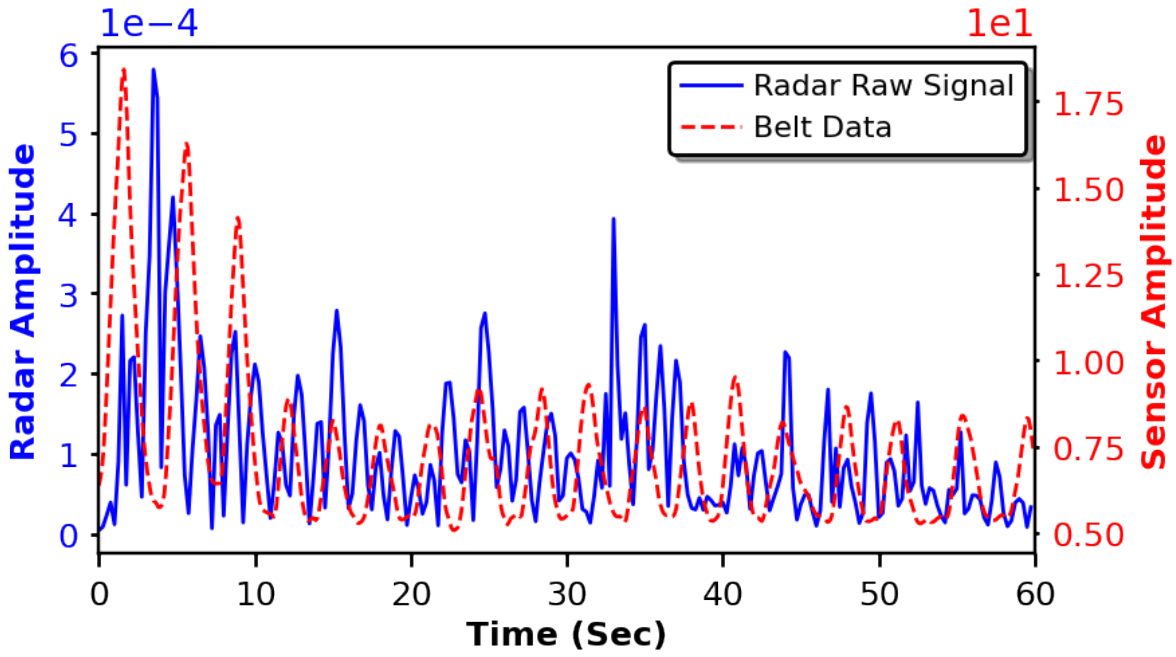


Figure 4.8: Time series Raw data of radar and belt after data flattening.

cardiac oscillations while suppressing unwanted noise components.

Once the signal was filtered, an FFT was performed to convert the time domain data into the frequency domain. This transformation allows the identification of the dominant periodic components that correspond to respiration and heartbeat cycles. To improve the spectral resolution and enhance the peak detection stability, a sixty second window was used before applying FFT. The FFT spectrum reveals the strongest peaks at the dominant frequencies. The largest peak in the 0.16 – 0.33 Hz range corresponds to the RR. The largest peak in the 0.8 – 2.0 Hz range corresponds to the HR. The fundamental frequency of these peaks was then converted into beats per minute (Bpm) for HR and breaths per minute (Brpm) for RR using the following equation [235]:

$$Bpm = DominantFrequency \times 60. \tag{4.5}$$

To further refine the accuracy of peak detection, cubic signal interpolation was applied after FFT. The interpolation technique was used to smooth spectral fluctuations, reduce noise effects, and improve frequency resolution for better peak estimation. Additionally, post-processing filters were used to remove sudden variations or outliers, ensuring the stability of the final estimated HR and RR values.

The same processing pipeline was used to classify different breathing patterns, and the input dataset was changed accordingly. By adjusting the cutoff frequencies, the dominant frequency for each breathing pattern was extracted and mapped to predefined categories based on Bpm ranges. The classification criteria are summarised in Table 4.4, which differentiates between low, normal, and elevated breathing based on their respective Brpm and frequency ranges [236]. The

final decision making process was used, where the extracted HR and RR values were classified into their respective physiological categories.

Table 4.4: Breathing classification based on frequency and Brpm ranges with key physiological indicators.

Breathing Class	**Frequency Range (Hz)**	**Brpm**	**Key Characteristics**
Slow breathing	< 0.2	< 12	May indicate sleep apnea
Normal Breathing	0.2 - 0.33	12 - 20	Stable, effortless
Elevated Breathing	> 0.33	> 20	Linked to anxiety

4.3.3 Experimental Results and Discussion

In this section, we present and evaluate the experimental results obtained during the preprocessing of UWB radar data. The accuracy of the radar derived measurements was verified using a medical grade reference sensor. Specifically, we focus on the estimation of HR and RR, followed by an analysis of the breathing pattern classification. For the RR validation, we used a respiration belt as the reference sensor; thus, comparisons were limited to the RR derived from the radar data.

Micro-Activity Estimation

The primary objective of the study was to estimate radar-based micro-activities that correspond to RR and HR, and then compare the outcome with a ground truth reference sensor (a respiration belt) and evaluate the accuracy of the system. The filtered signals from both the radar and respiration belt are displayed in Fig. 4.9, demonstrating a periodic breathing pattern over a sixty-second interval. The synchronised waveforms indicate that the radar effectively captured the chest motion associated with respiration.

To estimate RR, FFT was applied to the filtered data, converting it into the frequency domain. This transformation enabled the identification of the dominant breathing frequency, which correspond to the RR in Hz. The frequency spectra obtained from both the radar and respiration belt are presented in Fig. 4.10. The dominant frequency component identified from the FFT spectra for RR estimation is 0.28 Hz which, when converted to Brpm using Eq. 4.5, yields a value of 16.8 Brpm, falling within the normal breathing range. For the reference belt sensor, the corresponding value is 0.30 Hz, which translates to 18 Brpm. The absolute error in the RR estimation between the two sensors was 1.2 Brpm. This minor discrepancy can be attributed to slight motion artifacts and radar sensitivity to small chest displacement. Nevertheless, radar-based estimation remains within the normal physiological breathing range and demonstrates close alignment with the reference sensor.

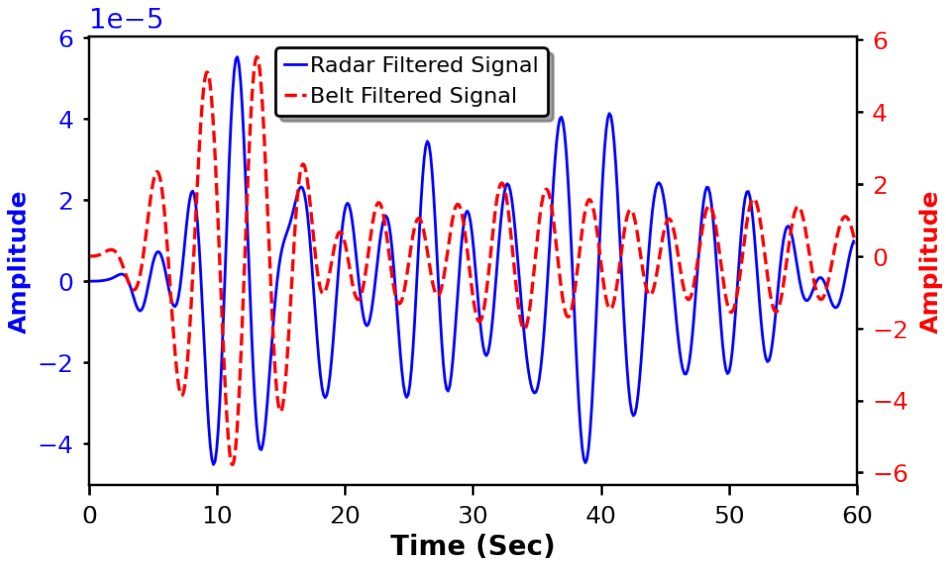


Figure 4.9: Output of the bandpass filter applied to both radar and reference sensor signals.

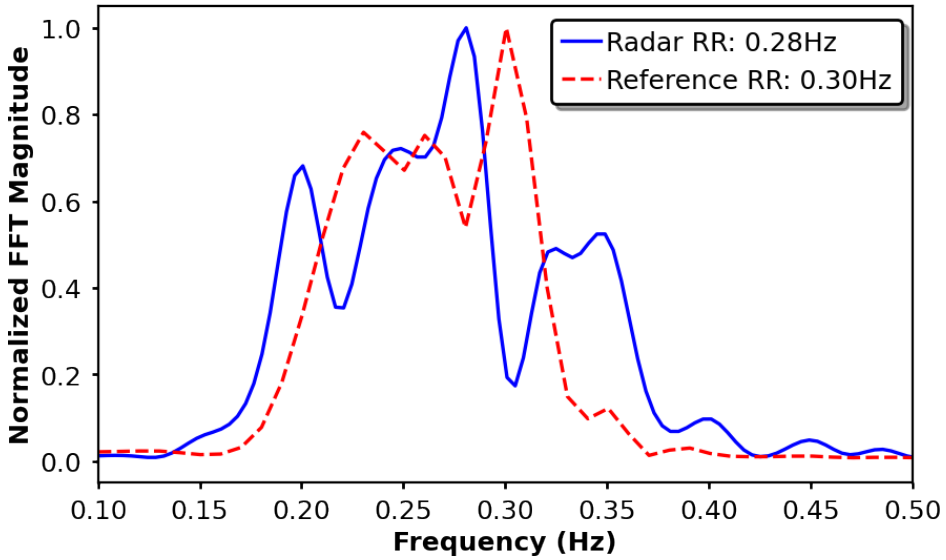


Figure 4.10: Comparison of RR estimates obtained from radar signals and a reference respiration belt using FFT-based analysis.

The human HR typically ranges from 60 to 240 Bpm, which corresponds to 1 – 4 Hz in the frequency domain. However, to isolate the heartbeat signal, a second-order Butterworth bandpass filter with a cut-off frequency of 1 – 2 Hz was applied to the radar signal. This filtering step removes lower frequency components below 1 Hz and higher frequency components above 2 Hz. After filtering, an FFT was performed to obtain the frequency spectrum, which allowed us to extract the dominant HR frequency component. The resulting HR spectrum is shown in Fig. 4.11, where a distinct peak in the range of 1 – 2 Hz corresponds to the estimated HR. Among the identified frequency peaks, the one with the highest magnitude was selected for HR estimation.

The extracted frequency was found to be 1.25 Hz, which, when converted to Bpm using the same Eq. 4.5, yields an estimated HR of 75 Bpm, which falls within the normal resting HR range for adults, that collected during normal breathing. Unlike RR, HR estimation was not validated with a reference sensor in this experiment because the respiration belt used for RR measurement does not provide HR values. However, the detected HR falls within a physiologically reasonable range, indicating that the radar-based approach is effective for HR monitoring.

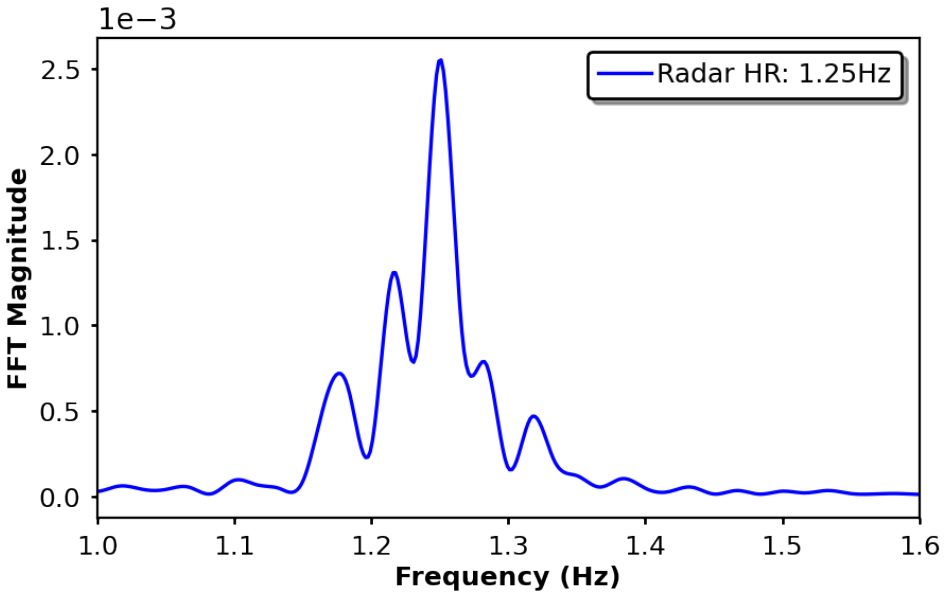


Figure 4.11: Radar-based HR estimation following FFT and interpolation processing.

Effect of Distance on RR Estimation: Another focus of this study was to evaluate the effect of distance on the radar-based RR estimation. To investigate this, experiments were conducted at three different distances between the radar and the participant: 55 cm, 60 cm, and 65 cm. The results indicate that as the distance varies, the accuracy of the RR estimation fluctuates, showing deviations from the ground-truth respiration belt measurements. Table 4.5 summarises the experimental results, presenting the RR and HR frequencies, the estimated Brpm, and Bpm from both the radar and the reference sensor, along with the absolute error at each distance.

Table 4.5: Comparison of radar-based micro-activity estimates with reference measurements at varying distances.

Distance	**Radar HR**		**Radar RR**		**Ref. RR**	**Range**	**Error**
	Freq (Hz)	**Bpm**	**Freq (Hz)**	**Brpm**	**Brpm**	**Brpm**	**Brpm**
55 cm	1.26	75.6	0.20	12.0	15.6		3.6
60 cm	1.25	75.0	0.28	16.8	18.0		1.2
65 cm	1.07	64.2	0.19	11.4	13.8		2.4

The results showed that the most accurate RR estimation occurred at 60 cm, where the absolute error was only 1.2 Brpm. This suggests that, at this optimal distance, the radar receives strong, unobstructed reflections from chest motion, leading to a higher SNR and improved detection accuracy. Additionally, the HR estimation at 60 cm was 75.0 Bpm, which falls within the expected physiological range. At 55 cm and 65 cm, the absolute errors in RR estimation increased to 3.6 Brpm and 2.4 Brpm, respectively. At shorter distances, the radar near-field effects and stronger multipath reflections likely introduced noise, leading to a higher deviation from the reference sensor measurements.

Similarly, at longer distances, weaker radar reflections and increased attenuation reduce the signal strength, making it more challenging to extract an accurate respiratory frequency. Additionally, the HR estimation at 65 cm dropped to 64.2 Bpm, which is lower than the HR values estimated at 55 cm (75.6 Bpm) and 60 cm (75.0 Bpm). This reduction in HR accuracy at longer distances may be attributed to weaker radar signal reflections, leading to a less reliable heartbeat detection. These results suggest that, while radar-based vital sign estimation is feasible across different distances, performance is optimal at 60 cm, likely because of the balance between strong reflections, minimal signal attenuation, and reduced noise interference.

Breathing Pattern Classification

In this study, RR estimation was used to classify three distinct breathing patterns: low, normal, and elevated. The classification was based on frequency ranges and Brpm, as summarised in Table 4.4.

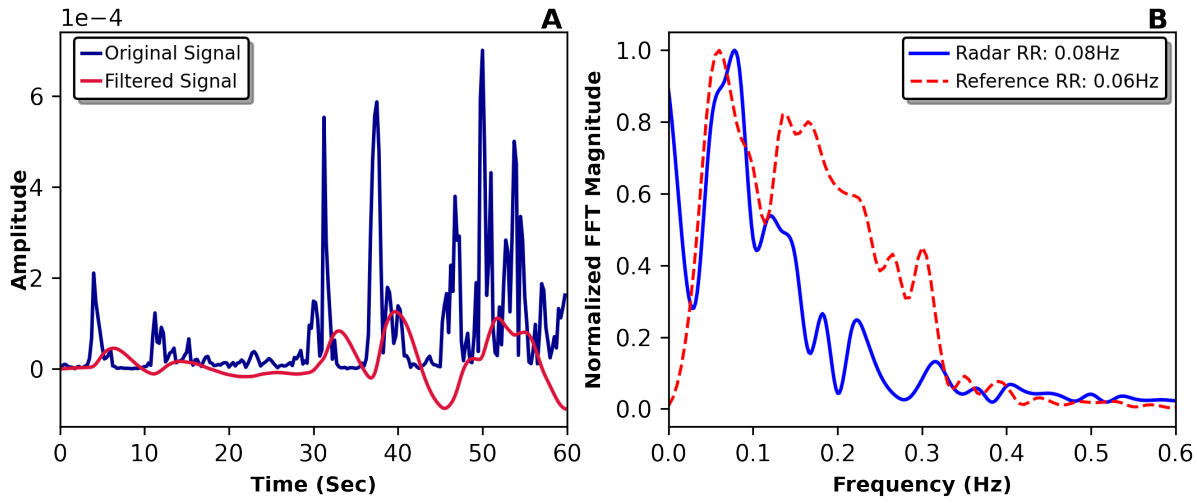


Figure 4.12: Slow breathing: (A) Original and filtered radar signal; (B) estimated spectrum of radar and reference sensor.

Slow breathing Estimation: Hypopnea, or slow breathing, is characterised by RR value below 12 Brpm, corresponding to a frequency of 0.16 Hz. Fig. 4.12 illustrates the time domain

representation of the raw and filtered radar signals. In subplot A, the raw signal appears noisy, while the filtered signal clearly highlights the periodic breathing cycles. Subplot B compares the radar estimated RR of 0.08 Hz (4.8 Brpm) to the reference RR of 0.06 Hz (3.6 Brpm), confirming slow breathing class. In addition, the estimated HR during slow breathing is shown in Fig. 4.13, indicates a dominant frequency of 1.12 Hz, corresponding to 67.2 Bpm. This slightly reduced HR is consistent with the expected physiological response to slow breathing. These results validate the ability of radar-based monitoring to detect hypopnea patterns and their physiological impact.

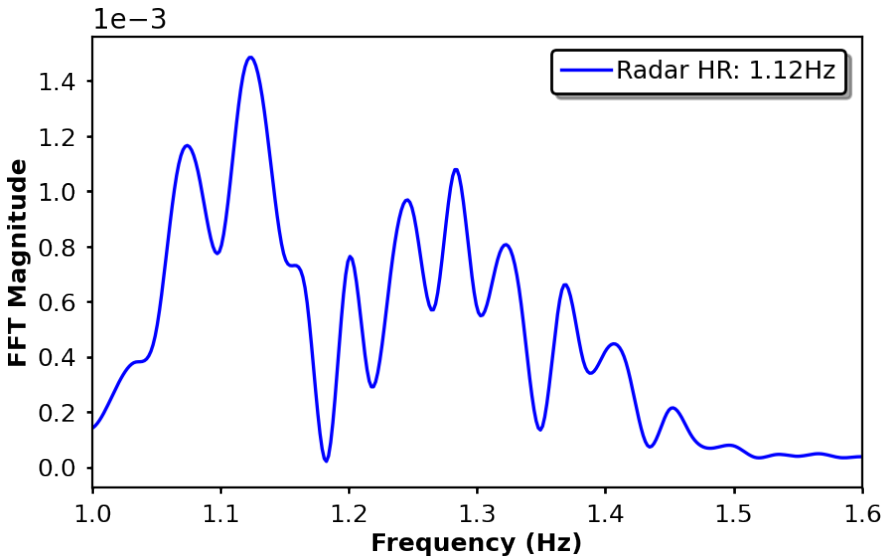


Figure 4.13: Estimated HR during Slow breathing class.

Normal Breathing Estimation: Normal breathing of an adult is typically lying in the range of 12-20 Brpm, corresponding to a frequency range of 0.16-0.33 Hz. Fig. 4.14 illustrates the time domain representation of the raw and filtered radar signals. In subplot A, the raw signal appears noisy, while the filtered signal clearly highlights a stable breathing pattern. Subplot B compares the radar estimated RR of 0.20 Hz (12 Brpm) to the reference RR of 0.23 Hz (13.8 Brpm), confirming normal breathing. In addition, the estimated HR during normal breathing, shown in Fig. 4.15, indicates a dominant frequency of 1.34 Hz, corresponding to 80.4 Bpm. This HR value is within the expected range of resting HR and reflects stable cardiovascular activity. These results validate the radar's ability to accurately detect normal breathing patterns and their associated physiological responses.

Elevated Breathing Estimation: Hyperpnea or elevated breathing is defined as a RR exceeding 20 Brpm, corresponding to a frequency above 0.33 Hz. Fig. 4.16 shows the time domain representation of the raw and filtered radar signals. In subplot A, the respiratory cycles occur

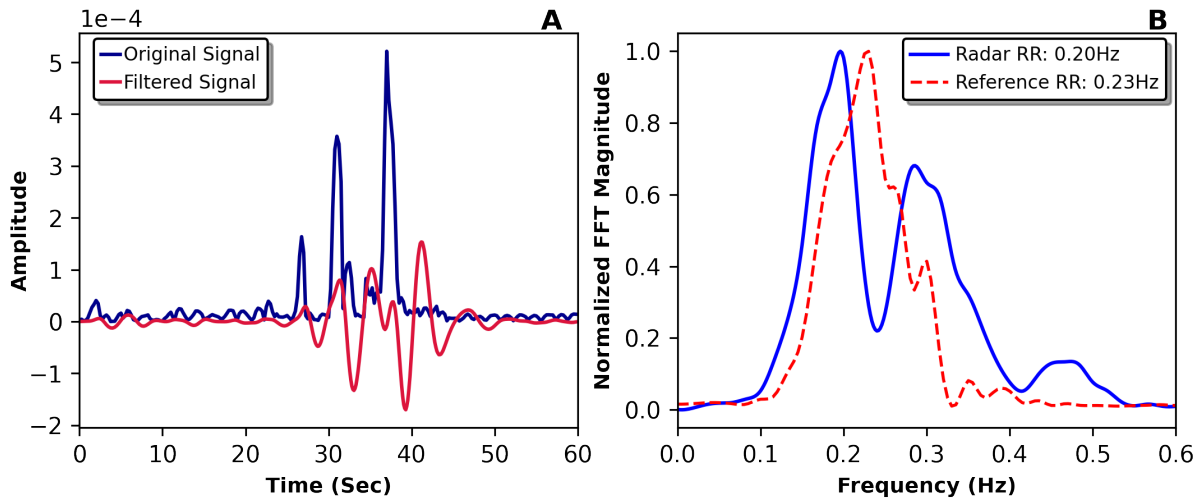


Figure 4.14: Normal Breathing: (A) Original and filtered radar signal; (B) Estimated spectrum of radar and reference sensor.

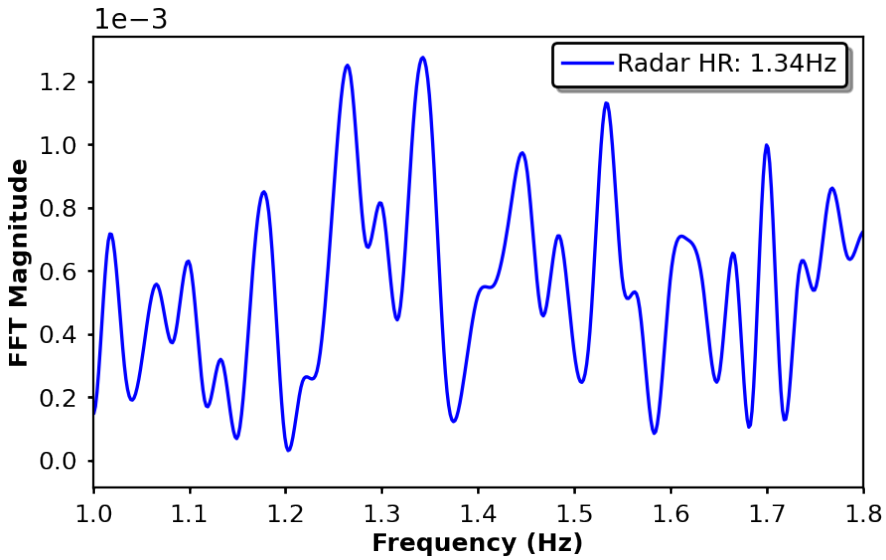


Figure 4.15: Estimated HR during normal breathing class.

more frequently than normal breathing, and the filtered signal can clearly capture the breathing pattern. Subplot B compares the radar estimated RR of 0.52 Hz (31.2 Brpm) to the reference RR of 0.57 Hz (34.2 Brpm), confirming the classification of elevated breathing. In addition, the estimated HR during hyperpnea, shown in Fig. 4.17, shows a dominant frequency of 1.80 Hz, corresponding to 108 Bpm. This increase in HR reflects the expected physiological response to an increase in RR, which may be due to higher metabolic demands and activation of the sympathetic nervous system. These results demonstrate the ability of radar to accurately detect elevated breathing patterns and their associated physiological effects.

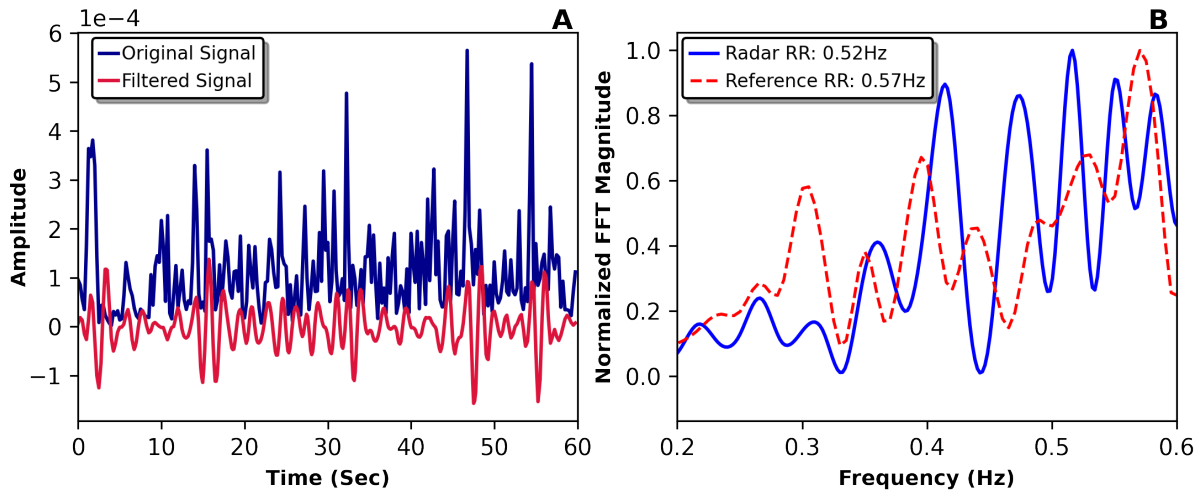


Figure 4.16: Elevated Breathing: (A) Original and filtered radar signal; (B) Estimated spectra of radar and reference sensor.

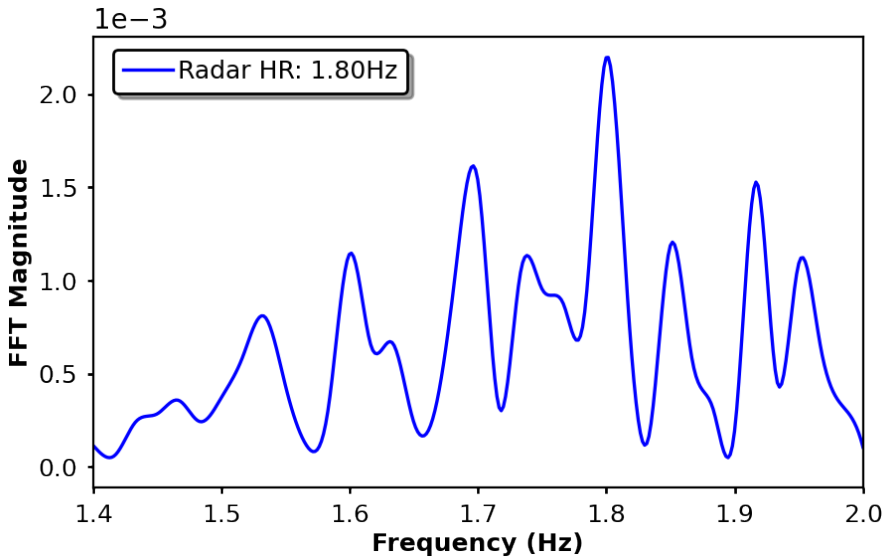


Figure 4.17: Estimated HR during elevated breathing class.

4.4 Vital Sign Monitoring Using Millimetre Wave Radar

Building on the previous discussion, this section explores a study that investigates the use of mmWave radar for non-invasive monitoring of micro-activities or vital sign specifically HR and RR. In the previous sections, we examined the use of UWB radar for vital sign estimation. While promising, this study was constrained by a limited experimental setup, as data were collected from only a single subject seated at varying distances, reducing the generalisability of the findings. To address these limitations and enhance the robustness and accuracy of micro-activity monitoring, we adopted a more advanced sensing approach using mmWave radar. Operating in the mmWave band, this technology offers higher resolution, improved sensitivity to minute chest

movements, and better suitability for diverse subject testing and real-world applications [25]. Among the various radar types, mmWave FMCW radars offer a combination of capabilities not available in single alternative systems. A key advantage is that they operate in the mmWave frequency range, which experiences high attenuation. This feature ensures high isolation between multiple co-located radars, even if they are only a few meters apart [178].

In addition, mmWave radars are highly sensitive to small displacements in the mm range, making them particularly suitable for detecting chest wall motion associated with breathing and heartbeat. Another important feature of these radars is their ability to distinguish between reflections at different ranges, allowing for multi-subject vital sign detection. This capability, highlighted in [237], offers a significant advantage over CW radars, which suffer from multipath fading due to their inability to distinguish between reflections at different ranges. FMCW radars mitigate this problem by reducing the likelihood of multiple reflections interfering with the main signal, especially in environments with minimal scattering.

Additionally, FMCW radar signals are inherently more robust against thermal noise than UWB radars. Similar to frequency modulation (FM) signals, vital sign information in the FMCW radar is encoded in the received phase, making it less susceptible to noise interference. This enhanced noise immunity contributes to the reliability of radar-based vital sign monitoring, even under non-ideal conditions. Despite these advantages, an important aspect to consider is the radar's ability to detect vital signs when it is placed at elevated positions relative to the subject. This is particularly relevant in scenarios where the radar must remain fixed, regardless of variations in the subject height or position. Previous studies explored various orientations and use cases. For instance, [238] analysed radar performance across multiple orientations, including front, back, left, and right, at varying distances and concluded that position and orientation significantly impact accuracy.

Another study [239] examined non-contact RR detection while a subject was on a treadmill, demonstrating the feasibility of radar-based vital sign monitoring during physical activity, which holds promise for sports and fitness applications. Similarly, [240] investigated the use of radar in automotive environments, where a sensor mounted on the top of a windshield successfully detected a driver's vital signs, reinforcing the potential of non-contact monitoring in real-world settings. However, these studies did not evaluate radar performance at different heights, leaving a gap in the understanding of its effectiveness when placed at elevated positions.

Our research aims to bridge this gap by systematically evaluating the performance of the FMCW radar at different heights. To conduct this study, we employed an advanced setup featuring the DCA1000 data acquisition module, which facilitates real-time data transmission and aligns with emerging 5G/B5G network standards. Additionally, we incorporated two distinct signal processing methods, FFT and peak counting, to analyse the acquired data. FFT is well suited for frequency domain analysis, enabling precise identification of vital sign patterns, whereas the peak count method provides a more straightforward approach for detecting respira-

tory cycles. By comparing these techniques in terms of computational complexity and detection accuracy, our study aims to optimise signal-processing strategies for mmWave radar-based vital sign monitoring.

4.4.1 System Overview

The system model, as depicted in Fig. 4.18, was developed to evaluate the accuracy and reliability of the FMCW radar in monitoring vital signs at three distinct heights: 1 m, 2 m, and the standard chest height, based on the subject's height measurements. The setup consists of an FMCW radar sensor paired with a DCA1000 data acquisition module, which transmits IF signals to a PC for preprocessing and vital sign estimation. Additionally, reference sensors, including an in-ear sensor and respiration belt, were used for benchmarking, enabling error quantification. The following subsections provide a detailed breakdown of each system component and its role in the framework.

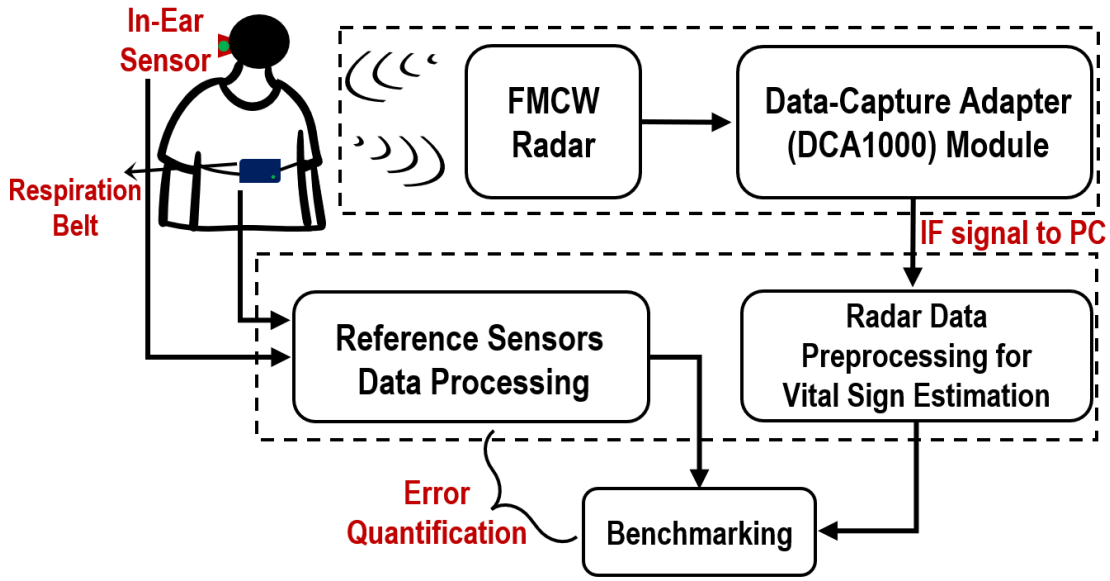


Figure 4.18: Block Diagram of our Proposed System Model.

mmWave Radar Sensor

In this study, we employed the IWR6843AOP radar sensor from Texas Instruments (TI) [241], which features an antenna-on-package (AoP) design that eliminates the need for external antennas or additional connectors. This self contained architecture minimises signal degradation from external interference, thereby ensuring high fidelity measurements of human respiration. As shown in Fig. 4.19A illustrates the radar sensor with its compact antenna configuration, whereas Fig. 4.19B presents the sensor connected to the DCA1000 data acquisition module used for transmitting IF signal to the connected PC with high data rate.

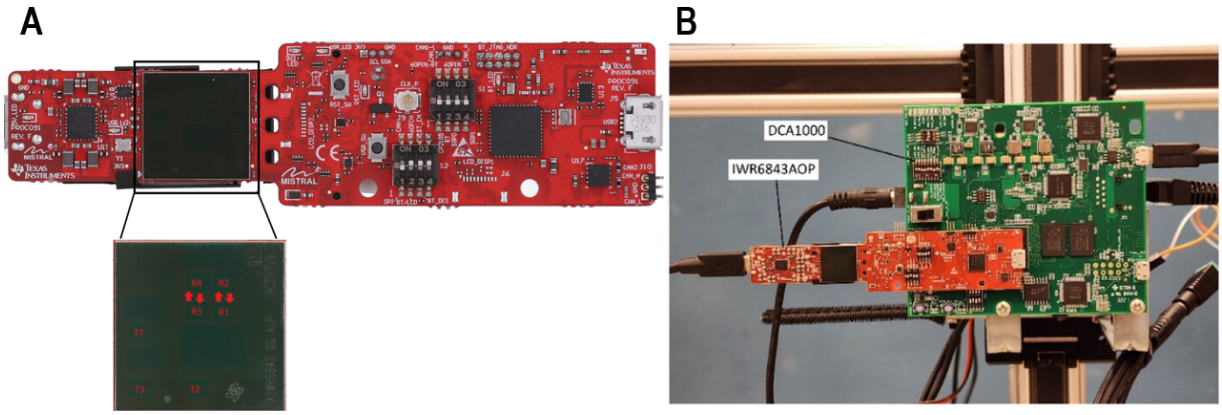


Figure 4.19: IWR6843 radar sensor board with DCA1000. (A) illustrates the radar sensor with its antenna configuration, while (B) shows the sensor connected to the DCA1000 data capture card.

The radar sensor is equipped with three transmitting antennas (3Tx) and four receiving antennas (4Tx). However, in this study, we utilised 1Tx and all 4Rx antennas to optimise the chest displacement detection during breathing. The radar operated at a starting frequency of 60.25 GHz with a bandwidth of 3.75 GHz, providing high resolution phase and amplitude measurements. To further detail the system's functionality, Table 4.6 summarises the key technical specifications of the FMCW radar sensor and the settings optimised for vital sign extraction and estimation.

Table 4.6: FMCW mmWave radar settings for vital sign estimation.

Parameters	**Values**
Center Frequency	60.25 GHz
Bandwidth	3.75 GHz
No. of ADC samples	256
No. of Frames	2400
No. of Chirps per frame	2
Frame Periodicity	50 msec
Chirp Duration	64 μsec
Maximum range	10 meters
Range Resolution	0.05 meter
Radar Memory	1.5 MB

mmWave Radar Basics: A detailed overview of FMCW radar sensor is provided in Chapter 2 (see Section 2.2), here, we summarise only the essential concepts relevant to vital sign estimation. The FMCW radar system transmits a chirp signal towards the subject's body and receives

the reflected signal, which can be mathematically expressed as [25]:

$$s(t) = e^{j(2\pi f_c t + \pi \frac{B}{T} t^2)}, \tag{4.6}$$

$$r(t) = e^{j(2\pi f_c (t-t_d) + \pi \frac{B}{T} (t-t_d)^2)}, \tag{4.7}$$

where f_c is the carrier frequency, B is the bandwidth of the chirp, T is the chirp duration and B/T is the frequency modulation slope of the chirp. The received signal is a delayed t_d and attenuated version of the transmitted signal, that shows $t_d = \frac{2R}{c}$, R shows the range of the target, and c shows the speed of light.

To extract the range, phase, and vital sign information, the received signal was mixed with a copy of the transmitted signal and filtered using a mixer. This mixing process generates a beat signal b(t) of a target at range R as:

$$b(t) = e^{j(4\pi \frac{BR}{cT} t + \frac{4\pi}{\lambda} R)} \tag{4.8}$$

Here $f_b = 4\pi \frac{BR}{cT}$ shows the beat frequency and $\phi_b = \frac{4\pi}{\lambda} R$ represents the phase shift, which encodes the displacement information of the subject's chest due to respiration and heartbeat.

For vital sign monitoring of a single subject, we analyse the phase variation of the beat signal $b(t)$, which is sinusoidal and consists of both frequency f_b and phase ϕ_b:

$$b(t) = e^{j(f_b t + \phi_b)} \tag{4.9}$$

To measure small scale vibrations such as chest displacement during respiration, we tracked the phase changes of the FMCW radar signal over time. If an object moves by a small displacement ΔR, the corresponding phase change between consecutive measurements is:

$$\Delta\phi_b = \frac{4\pi}{\lambda} \Delta R \tag{4.10}$$

For example, at a wavelength $\lambda = 4mm$, a displacement as small as ΔR=1 mm results in a phase change of $\Delta\phi_b = \pi$. To extract this phase information, we take the FFT of the beat signal and compute the phase at the target range bin. The displacement signal can then be derived from the phase variations over time:

$$x(m, nT_s) = \frac{\lambda}{4\pi} \phi_b(m, nT_s) \tag{4.11}$$

Where m is the range bin, n is the chirp index, and T_s is the time interval between consecutive chirps. It is assumed that the vibrations $x(t)$ are small, ensuring that the target remains in the same range bin throughout the measurement period. The resulting beat signal, which contains I and Q components, was collected via the DCA1000 module for further processing.

DCA1000 Data Capture Module

DCA1000EVM is an evaluation module (EVM) designed by TI [242] for high speed data acquisition from FMCW radar sensors. It serves as an interface between the radar sensor and the host PC, enabling the real-time capture and storage of raw ADC data for post-processing. The DCA1000 connects to the radar module through a 60 pin high-density samtec connector, ensuring a low latency data transfer. The captured data were streamed via a one Gbps ethernet interface to the host PC. The module is equipped with an field-programmable gate array (FPGA) to handle data processing, and a USB interface for additional configurations. The DCA1000 supports two configuration modes.

- **Hardware switch configuration mode:** Configured via on-board switches to define data logging and transfer settings.
- **Software CLI configuration mode:** Configured via JSON files and executed using a command-line interface (CLI) on Windows or Linux systems.

To ensure seamless operation, the DCA1000 provides raw and multi-mode logging for diverse data acquisition requirements and support for different LVDS modes (4-lane / 2-lane) to optimise data capture. The flexible stop conditions include byte count, frame count, duration-based, or infinite streaming. Error detection and LED indicators for status monitoring (e.g. data transfer, buffer overflow, and LVDS path errors). The captured data are stored in a binary format (.bin), which can later be processed to extract phase, range, and vital sign information as detailed in subsequent sections.

Experimental and Data Collection Setup

This study investigated the performance of an FMCW radar system for detecting RR and HR at multiple heights. The experiment was conducted in a controlled laboratory environment at the University of Glasgow, UK, and involved 10 test subjects who remained stationary while the radar recorded their vital signs. The radar system was placed at three different heights:

1. 1 meter above the ground
2. Chest height of the subject (baseline, varied per individual)
3. 2 meters above the ground

To ensure accuracy, the radar measurements were compared with of those two reference sensors. A respiration belt was worn around the upper abdomen of the subject for the RR estimation, and an in-ear sensor was inserted into each subject's ear for HR estimation.

Each experimental session lasted 120 second per radar position for each subject, ensuring that all sensors were properly synchronised and operational before the data collection began.

The radar was positioned 0.8 meters away from the subject in all cases for consistency. When placed at the chest height, it was mounted to maintain a zero tilt angle, ensuring that its main beam was perpendicular to the chest of the subject. However, at a height of 1 m, the radar was tilted upward towards the chest, forming an average elevation angle of 25.3°, as illustrated in Fig. 4.20A. At a height of 2 m, the radar was tilted downward at an average depression angle of 41.5°, as shown in Fig. 4.20B. These variations allowed for a comparative analysis of the radar performance across different heights and angles.

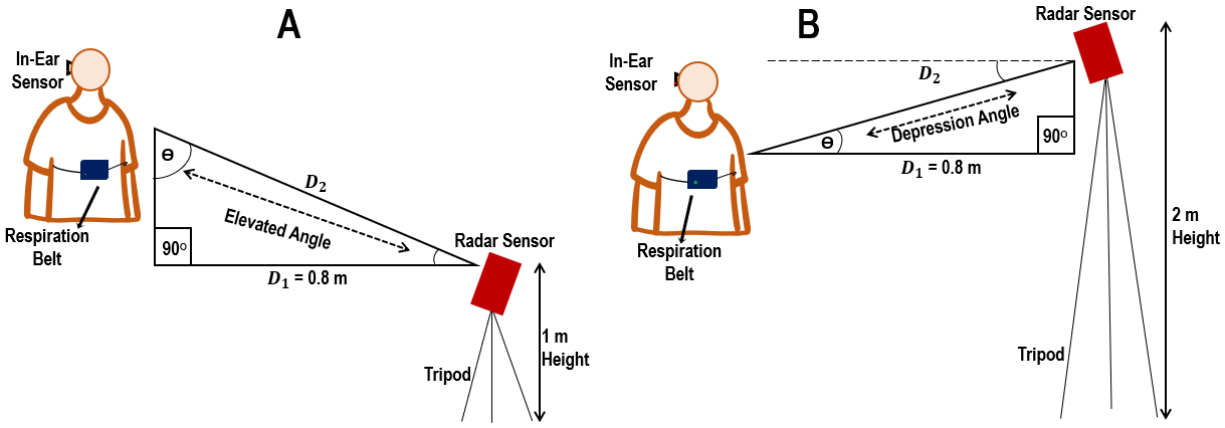


Figure 4.20: Radar positioning at different heights and corresponding tilt angles. (A) Radar elevated at 1 m with an elevated angle, while (B) shows radar elevated at 2 m with a depression angle.

During data collection, test subjects were instructed to stand still and breathe normally, and their RR and HR were recorded. All sensors were synchronised and tested each time before the start of data collection from a new subject to ensure proper consistency and accuracy of the sensors. Radar data were collected from three predetermined heights and stored for post-processing. Two Dell laptops were used for the data acquisition and storage.

- Laptop 1 Collected respiration belt and in-ear sensor data via USB, controlled through a Python script.
- Laptop 2 Managed radar data acquisition, storing raw binary (.bin) files, which were later parsed and processed using MATLAB and Python for vital sign estimation.

The experimental environment and data collection setup are shown in Fig. 4.21, showing the radar at different heights and the subject positioning during the experiment.

4.4.2 Radar Data Preprocessing for Vital Sign Estimation

Extracting vital signs from radar data is a multi-stage process involving systematic signal processing, as shown in Fig. 4.22. The process consists of three key steps: Human detection, Phase extraction, and Signal separation for HR and RR estimation. Each stage plays a vital role in ensuring accurate and reliable detection of human heartbeat and breathing signals.

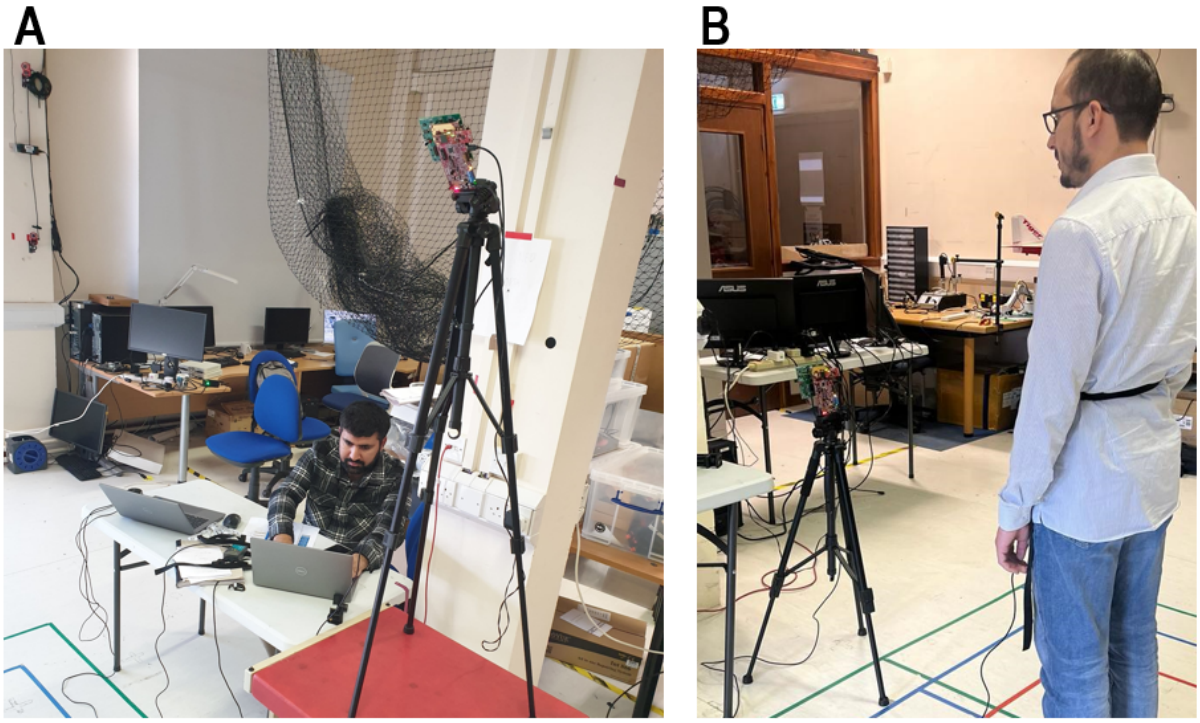


Figure 4.21: Experimental setup during data collection. (A) shows the data collection environment with radar at 2 meters height, while (B) illustrates a subject positioned for measurements when the radar was at 1 meter height.

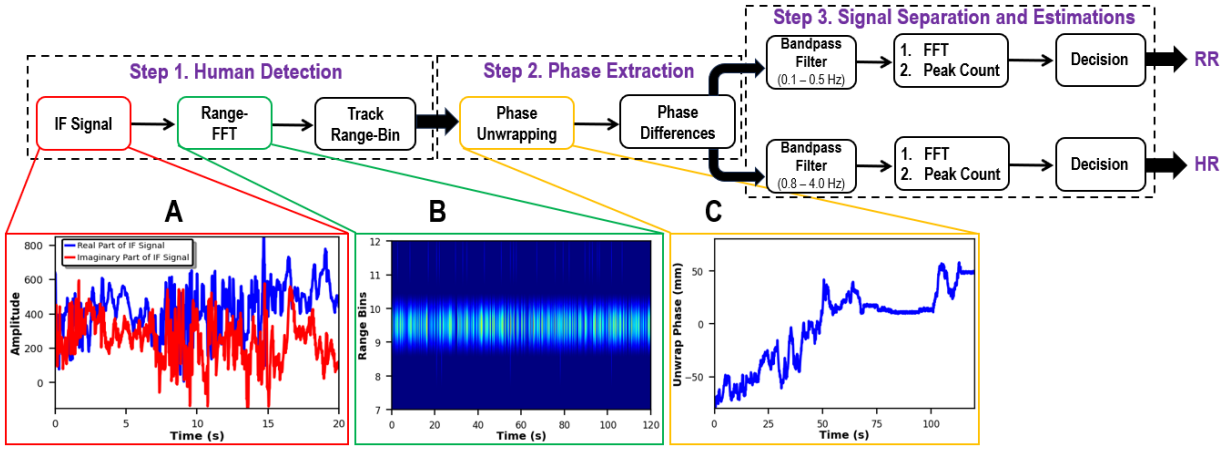


Figure 4.22: Block diagram illustrating radar data Preprocessing steps for vital sign estimation.

Human Detection: The first step in radar data processing is to detect and localise the subject, as only the range bins corresponding to the subject's position contain relevant vital sign information. This is achieved using Range-FFT, which transforms the received time-domain signal into a range profile, separating the reflected signals based on their distance from the radar. In an FMCW system, each column in the range profile represents the FFT of one chirp and each row corresponds to the beat frequency at a specific target distance, known as a range bin [95]. Over multiple chirps, the range profile formed a slow time axis. For stationary subjects, the subject's chest location was identified within a specific range bin. For example, in Fig. 4.22B, the

subject is primarily located in bins 9-11, with the strongest signal at bin 9.5. The optimal range bin can be selected by selecting the bin with the highest signal magnitude or maximum energy in the detected region. Once the subject's range bin is determined, vital sign signal processing algorithms can proceed.

Phase Extraction: After identifying the subject's range bin, phase information was extracted to capture the chest displacement caused by respiration and heartbeat. The phase values derived from the complex range profile data were measured over time, assuming that the subject remained within the same range bin. If the subject moves, the algorithm re-locks onto the new range bin after a short delay.

The extracted phase values are initially wrapped within the range $[-\pi, \pi]$, causing discontinuities when the phase exceeds this range. Phase unwrapping corrects these jumps by adding or subtracting 2π whenever the phase difference between consecutive values exceeds $\pm\pi$, ensuring a smooth, continuous signal. To further refine the signal, the phase difference was computed by subtracting successive unwrapped phase values. This step enhances the heartbeat signal while reducing low-frequency drifts, which can impact the accuracy. The phase variations correspond directly to the chest displacement, enabling the estimation of respiration and heartbeat induced movement.

As shown in Fig. 4.22C, the unwrapped phase can be converted to displacement distance over the entire 120 second recording period, where large displacements represent breathing and smaller superimposed displacements correspond to heartbeats. At the 70 second mark, the subjects were instructed to hold their breath, allowing for an isolated observation of the heartbeat waveform. To approximate the maximum line-of-sight (LOS) motion of the target, we used the following relation [178]:

$$d_{\max} = \frac{\pi R_{\text{pp}} f_x}{f_{\text{PRF}}}. \tag{4.12}$$

Where R_{pp} is the peak-to-peak sinusoidal motion, f_x is the motion frequency, and f_{PRF} is the pulse repetition frequency of the radar.

For an extreme case of breathing, assuming R_{pp}= 2cm and f_x=0.5Hz, this constrains $f_{\text{PRF}} \leq 13$ Hz for the unwrapping process to remain valid. Because the unwrapped phase is directly proportional to the displacement distance, the chest displacement can be expressed as [178]:

$$R[n] = \frac{c\Phi[n]}{4\pi f_c} \tag{4.13}$$

where f_c is the center frequency of the EM wave, c is the speed of light in the propagation medium, and $\Phi[n]$ represents the unwrapped phase at time step n. These chest displacement variations, originating from the breaths and heartbeats, can be observed in Fig. 4.22C, prior to any filtering or signal processing. Although some noise may still be present, the distinct

breathing and heartbeat waveforms remain partially identifiable.

Signal Separation and Estimations: After the phase extraction, the next step involves separating the respiratory and heartbeat signals and estimating their respective frequencies. Because both signals were superimposed in the extracted phase data, they were isolated based on their distinct spectral characteristics using bandpass filtering and spectral estimation techniques. The extracted phase values first underwent bandpass filtering to eliminate noise and isolate the frequency ranges associated with RR and HR.

A fourth order Butterworth filter is applied with cutoff frequencies set between 0.1 – 0.5 Hz for RR (6 – 30 Brpm) and 0.8 – 4 Hz for HR (48 – 240 Bpm). To address motion artifacts or large body movements, segments with high-amplitude distortions were either scaled down or discarded if their energy exceeded a predefined threshold, ensuring that only clean and reliable data were used for further analysis. Once the signals are filtered, spectral estimation techniques such as FFT and peak detection are applied to identify the dominant frequency components corresponding to RR and HR. FFT computes the power spectral density (PSD) of the filtered signals and identifies the strongest frequency component associated with either RR or HR. The RR, for instance, can be calculated using the equation [235]:

$$RR = \frac{\Delta f}{2 f_m} \times 60 \tag{4.14}$$

Where Δf represents the detected frequency component and f_m is the modulation frequency.

Additionally, time-domain peak detection analyses the periodic variation in the chest displacement signal by identifying peak-to-peak intervals corresponding to breathing or heartbeat cycles, thereby providing another estimate of RR and HR. To enhance accuracy, a confidence metric is computed by incorporating factors such as the SNR, peak prominence, and consistency across multiple spectral techniques. The final HR and RR values are determined based on the confidence metric from these spectral estimation techniques, ensuring the robustness and accuracy of the vital sign measurements.

Quantification Across Reference Sensors

To evaluate the accuracy of the FMCW radar sensor in monitoring RR and HR, we compared its performance with that of reference sensors. The same respiration belt previously used in UWB-based vital sign estimation was used for RR measurements in this study. For HR estimation, an in-ear sensor from cosinuss Germany [243], was employed that wirelessly transmitted real-time data to a dedicated server equipped with a built-in Wi-Fi module. The in-ear sensor system recorded HR data in real time and stored preprocessed values along with the corresponding timestamps. The HR data were retrieved from the server in the CSV format at a sampling rate of 1 Hz and transferred to a local PC for further analysis. To ensure consistency in the

performance evaluation, an equal number of RR and HR values were collected from both the radar and reference sensors at each radar height.

The resulting discrepancies in the measurements were analysed to benchmark the radar's accuracy in detecting the vital signs. To systematically quantify the errors between the radar-based and reference sensor measurements, we calculated two statistical error metrics:

1. Mean Absolute Error (MAE): Measures the average deviation between radar and reference sensor readings.

2. Root Mean Square Error (RMSE): Captures the variance and overall accuracy of radar-based estimations.

These benchmarking metrics provide insights into the precision and reliability of the radar system in capturing vital signs at different height positions.

4.4.3 Results and Discussion

This section presents the key findings of our study on extracting vital signs from raw mmWave radar data using two primary signal processing techniques: FFT and peak count analysis. We evaluated the performance of these methods across three radar heights (1 m, chest height, and 2 m) to assess their accuracy in detecting micro-activities and determine the reliability of mmWave-based FMCW radar for non-invasive vital sign monitoring. The experimental setup utilised two Dell laptops, each equipped with an 11th generation Intel core i7 processor (3.00 GHz) and 16 GB of RAM, for data storage and preprocessing of both radar and reference sensor data. Our analysis aimed to identify which signal processing method delivers the most consistent and accurate measurements across different height configurations. The following subsections provide a detailed comparative analysis of FFT and peak count techniques, examining their respective strengths, limitations, and suitability for multi-height radar-based micro-activity monitoring.

Performance Comparison of HR and RR

Fig. 4.23 presents 120 seconds of breathing and heartbeat waveforms, extracted from raw radar data after applying a bandpass filter, as described in the previous section. To validate the radar-derived estimates of RR and HR, a thorough comparison with reference sensor readings was performed. This comparison allows us to assess the accuracy and reliability of radar-based measurements under different processing techniques. In Fig. 4.24 , we present the estimated RR and HR obtained using the FFT and peak counting methods, as well as the reference sensor benchmark. While the analysis was performed for all ten subjects at three different radar altitudes, the results shown here correspond to one subject at 2 m altitude as a representative example of data consistency across conditions. The comparison shows that the estimates obtained using the

peak counting method are in closer agreement with the reference sensor measurements for HR and RR. This suggests that the peak counting method has a higher accuracy in detecting vital signs than the FFT-based method. But still FFT based outcomes fall under the normal RR and HR estimations.

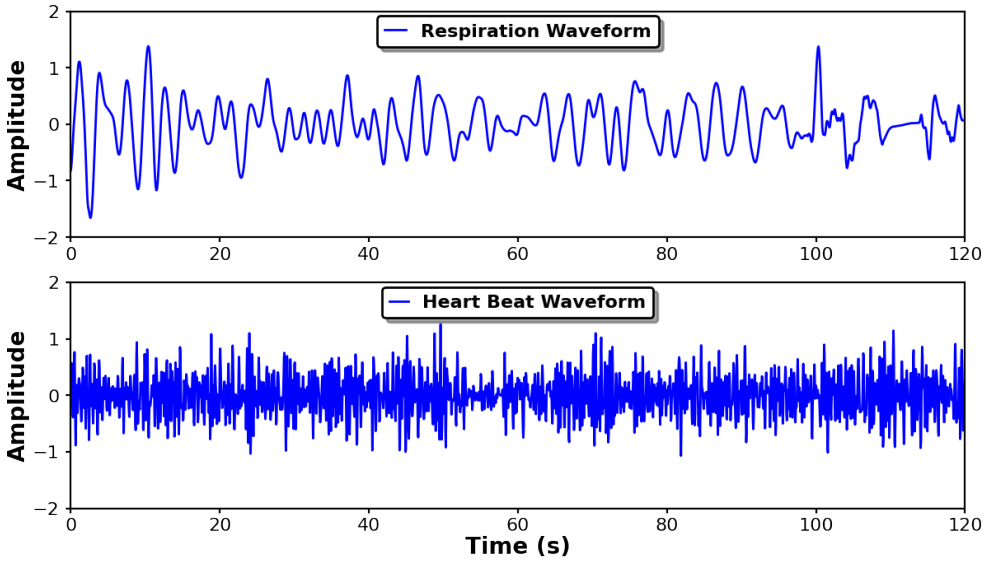


Figure 4.23: Filtered HR (bottom) and RR (top) waveforms over 120 seconds.

Statistical Analysis of Error Metrics

To assess the accuracy of RR and HR estimations using the FFT and Peak Count methods, we analysed the MAE and RMSE values at three radar heights: 1 m, chest level, and 2 m. Fig. 4.25, illustrate the error comparisons, showing that placing the radar at 2 m, which is closer to the standing height of the subjects (160 - 185 cm), results in lower errors compared to 1 m. This suggests that aligning the radar height with the subject's height, such as using 1 m for seated or lying positions, could further optimise the accuracy.

Additionally, Table 4.7 presents the mean error metrics for both FFT and Peak Count techniques, confirming that the Peak Count consistently outperforms FFT with lower error values across all height configurations. The Chest_ref column provides a baseline reference for RR and HR errors at chest height, whereas the RR_all and HR_all columns aggregate the errors across all height positions, reinforcing the adaptability of the radar system for non-contact vital sign monitoring. Further statistical analysis explored the relationship between subject height and error variations, as shown in Table 4.8. The correlation coefficients indicated a weak negative correlation between subject height and errors at 1 m radar height, whereas a weak positive correlation was observed at chest height and 2 m. The correlation coefficients (r) for MAE and RMSE ranged from -0.13 to 0.33, suggesting that subject height had a minimal impact on radar measurement accuracy. The mean MAE values range from 4.1 to 5.9, and mean RMSE values range from 5.2 to 6.7, with the STD and variance showing relatively stable consistency across

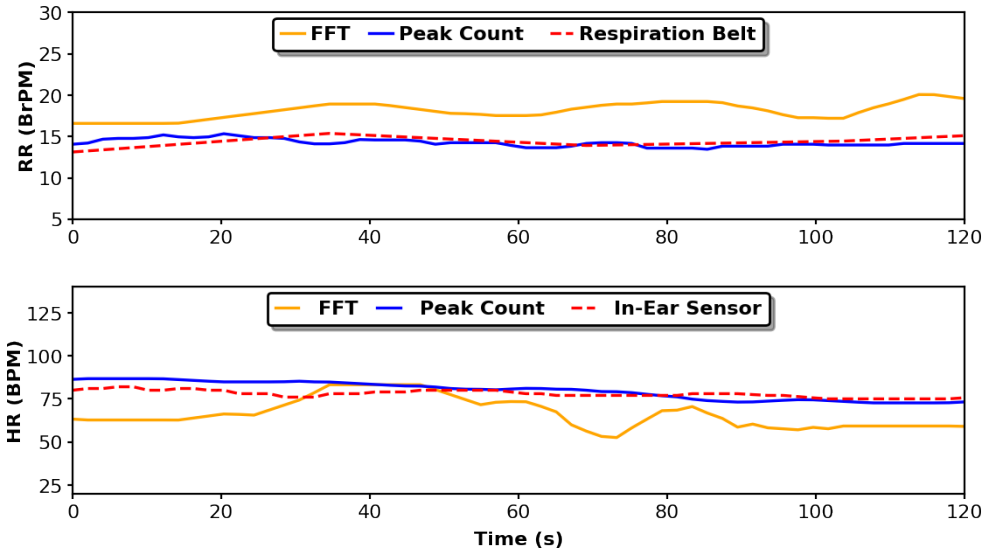


Figure 4.24: Estimated HR (bottom) and RR (top) over 120 seconds: Radar FFT and Peak count methods vs. reference sensors.

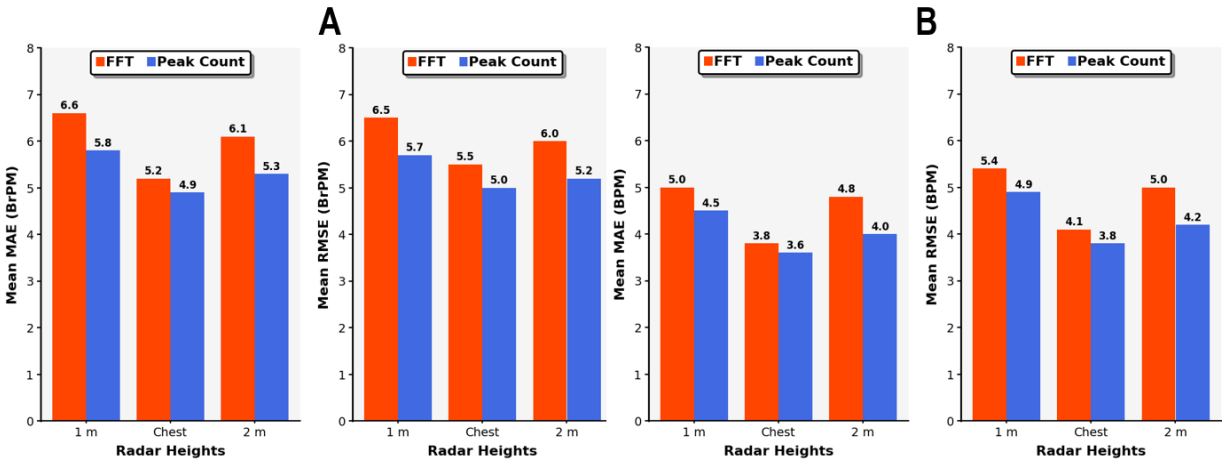


Figure 4.25: Mean Error Metrics. A) Comparison of RR mean error metrics across radar heights. B) Comparison of HR mean error metrics across radar heights.

different radar heights. These findings validate the robustness of the radar system and signal processing techniques and demonstrate their capability to reliably estimate vital signs across various subject positions and conditions.

Overall, the Peak Count method outperformed FFT-based spectral estimation, consistently yielding lower MAE and RMSE values. Additionally, radar height placement significantly influences error rates, with higher radar elevations (closer to the subject's height) reducing the errors. The weak correlation between subject height and error metrics suggests that the accuracy of the radar system is largely independent of subject height, further confirming its adaptability and reliability in non-invasive HAR applications.

Table 4.7: Mean error metrics for RR and HR estimations using FFT and Peak Count methods at 1 m, 2 m, and all heights.

Errors	Chest_ref	RR	HR	RR_all	HR_all
FFT					
MAE	4.5	6.3	4.9	5.4	4.7
RMSE	4.8	6.2	5.2	5.5	5
Peak Count					
MAE	4.2	5.5	4.2	4.9	4.2
RMSE	4.4	5.4	4.5	4.8	4.4

Table 4.8: Correlation and statistical analysis between radar height and error metrics.

Metrics	1 m	Chest	2 m
Var MAE	2.3	5.2	5.1
Var RMSE	2.5	6.5	4.8
STD MAE	1.51	2.2	2.25
STD RMSE	1.58	2.5	2.20
r MAE	-0.13	0.25	0.33
r RMSE	-0.06	0.19	0.27
Mean MAE	5.7	4.09	5.9
Mean RMSE	6.7	5.24	6.7

4.5 Summary

This chapter presents a comprehensive study on two different radar technologies, UWB and mmWave FMCW radar, for non-intrusive micro-activity monitoring, that tackle the challenge **C1**, as discussed in **Section 1.3**. In the first phase, we evaluated HR and RR estimation using UWB radar by collecting data from a single subject in a sitting position under various breathing patterns. This initial study allowed us to analyse radar performance, refine signal processing techniques, and establish a baseline algorithm for subsequent advanced signal processing. The UWB radar evaluation provided critical insights into radar-based physiological monitoring, helping to prepare a signal processing pipeline for more complex scenarios. In the second phase, we extended our study to the FMCW radar for multi-subject, multi-position vital sign monitoring. This phase involved 10 subjects with radar data collected at three different heights (1 m, chest level, and 2 m) to examine the impact of radar placement on RR and HR estimation.

We developed a comprehensive signal processing pipeline, including range bin selection, phase extraction, phase unwrapping, and frequency based signal separation. The Peak Count method consistently outperformed the FFT, demonstrating lower MAE and RMSE values and greater robustness in extracting vital signs. Additionally, while radar placement influenced the

error rates, statistical analysis revealed a weak correlation between subject height and estimation accuracy, confirming the versatility and adaptability of the radar system across different scenarios. This chapter highlights the progression from UWB to FMCW radar, demonstrating the capabilities, limitations, and necessary optimisations for effective radar-based vital sign monitoring. Building on this foundation, this study addresses the challenge of using radar for non-intrusive, multi-scale HAR by accurately estimating micro-activities, alongside macro-activity recognition. The findings confirm that the radar is a reliable and accurate tool for non-contact physiological monitoring, reinforcing its potential as a preferred solution for AI-driven, non-invasive HAR applications in healthcare and remote monitoring.

Chapter 5

Edge-Optimised Privacy Preserved Macro-Activity Recognition

This Chapter marks another contribution of this thesis: the transition from radar data processing to the real-time deployment of a robust R-HARS for healthcare applications. This addresses challenges **C2** and **C3**, as detailed in Section 1.3. Our prior evaluation of four radar preprocessing techniques using TL identified two optimal candidates, namely RD and STFT-based TF domains. These were selected because of their balance between computational efficiency and recognition performance, making them suitable for resource-constrained environments. These domains are integrated with DL models and optimised for low-latency inference on edge devices. Energy efficiency is a key deployment challenge, requiring comprehensive model optimisation strategies that minimise computational overhead while maintaining recognition accuracy. Additionally, determining the optimal pairing between radar domain representations and TL models ensures strong performance under constrained conditions. Although radar technology is inherently privacy-preserving, we further enhance data security by incorporating a local differential privacy (LDP) mechanism to enable secure on-device inference. By bridging the gap between research and implementation, this study demonstrates the feasibility of deploying optimised and privacy-aware radar-based macro-activity recognition systems on edge platforms. This advancement paves the way for practical applications of AAL and beyond.

5.1 Introduction

The integration of artificial intelligence (AI), the Internet of Things (IoT), and Edge computing has transformed real-time decision-making in various fields, especially in healthcare. AI-enabled edge computing systems address major global challenges including resource limitations, high energy consumption, and security vulnerabilities [244]. However, traditional cloud-based solutions have powerful computing capabilities, they also have significant drawbacks, including high communication latency, increased energy consumption, and potential privacy risks [181].

These limitations make cloud-based AI unsuitable for real-time, resource-constrained IoT applications that require fast processing and secure data transmission [245].

Edge intelligence (EI) has emerged as an effective solution to overcome these challenges. EI integrates AI and edge computing to perform real-time data processing directly on edge devices without relying on remote cloud servers [246]. This approach can reduce latency, save bandwidth, and enhance privacy by keeping sensitive data locally. Compared with traditional cloud-based AI that requires large amounts of data to be transmitted to remote servers for processing, edge AI performs inference locally, making it suitable for real-time applications such as healthcare monitoring and emergency response systems. Despite its advantages, deploying DL models on edge devices still faces significant challenges due to limited computing resources and power constraints. Inference on edge devices requires optimised AI models that can run efficiently within hardware limitations, ensuring a balance between accuracy, speed, and energy consumption.

Although EI enables real-time decision-making, training DL models directly on edge devices is impractical due to their limited computational and memory capacities. To address this, we adopt a TL approach in which models are pre-trained and optimised on high-performance GPU-based systems before being deployed to edge devices for inference. This significantly reduces the computational burden on resource-constrained platforms while maintaining inference accuracy and efficiency. Following optimisation, the models are deployed on single-board computers (SBCs), enabling low-latency inference with minimal energy and memory consumption. A prominent application of this approach in healthcare is HAR, particularly fall detection. By combining FMCW radar sensing with edge AI, the system delivers accurate motion detection while preserving user privacy. The deployment of DNNs on SBCs ensures that critical events such as falls can be detected in real time, allowing immediate alerts and timely medical intervention [247].

However, the complexity of DL models and complex radar datasets pose challenges related to computational efficiency and deployment feasibility. TL and model compression techniques provide effective solutions by reducing the computational requirements during inference. Instead of manually optimising models, the proposed approach leverages pre-trained networks and fine-tunes them for specific healthcare tasks, reducing the need for extensive training and enabling scalable, real-time AI solutions. By processing radar data directly at the edge, the system minimises latency, computational overhead, and security risks, making it a viable solution for applications such as AAL and real-time patient monitoring at the hospitals. The proposed approach demonstrates the potential of edge AI to transform healthcare by enabling cost-effective, intelligent, and real-time decision making in resource-constrained environments.

5.2 Contributions

The key contributions of this Chapter is summarised as follows:

- Developed an optimised signal processing pipeline that generates radar-based RD and TF representations on edge devices, reducing both computational and communication costs while improving real-time inference efficiency.
- TL-based approaches were employed with comprehensive GPU-based energy profiling during training to minimise computational costs, enhance model generalisation, and reduce the carbon footprint for large-scale DL deployment. Framework generalisability was validated using publicly available 24GHz and 77GHz radar datasets on GPU systems for accuracy and performance assessment. The real-time edge deployment energy consumption was precisely measured using DC current monitoring with INA219 sensors integrated with edge devices during inference and radar signal processing.
- Applied post-training quantisation (PTQ) techniques to achieve significant model size reduction and computational cost optimisation while maintaining recognition accuracy, enabling efficient inference on low-power edge devices, such as Raspberry Pi (RPi) and Jetson Nano (JNano), with minimal communication overhead.
- Local differential privacy (LDP) directly at the edge for inference outputs, ensuring privacy-compliant predictions without server-side data exposure or model retraining requirements. This approach maintains real-time HAR performance while eliminating sensitive data transmission costs and regulatory compliance risks.
- Delivered a comprehensive decision framework that optimally combined radar domain representation, TL architecture, and edge hardware selection, enabling systematic trade-off between performance, energy efficiency, communication costs, and privacy preservation for robust healthcare and AAL deployments.

5.3 Proposed Method

In this section, we present the proposed methodology for a R-HARS designed for real-time inferences on edge devices. As illustrated in Fig. 5.1, the conceptual collaborative learning framework follows a structured pipeline, beginning with data acquisition from an FMCW radar, followed by data preprocessing at the edge, DNNs modeling, and model compression techniques to optimise edge deployment. A key component of this framework is the integration of LDP alongside radar data, ensuring privacy preservation by securing both data transmission and model inference at the edge. Unlike traditional DP methods, LDP is applied near the edge before inference, ensuring that predictions remain privacy-preserving while maintaining the model

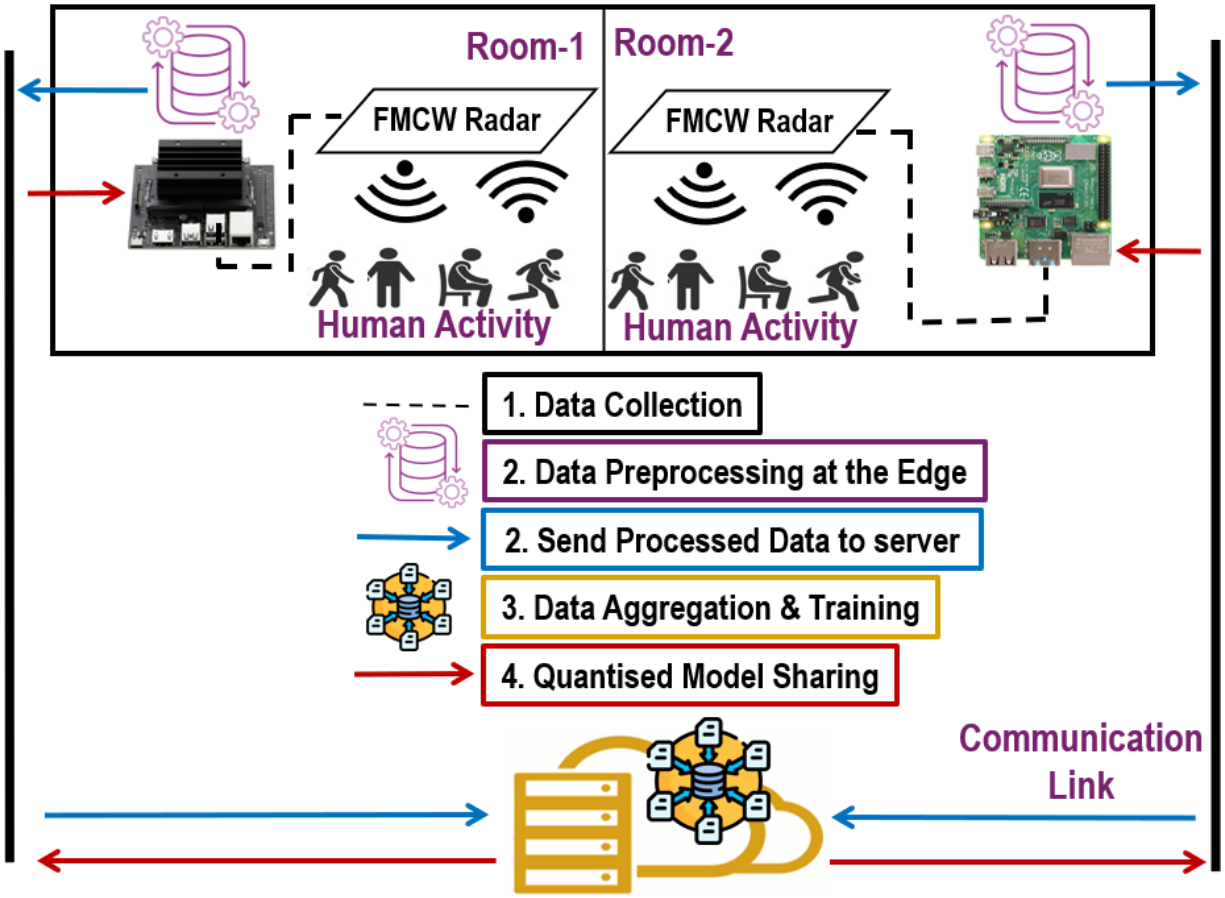


Figure 5.1: Conceptual collaborative learning framework for HAR using FMCW radar connected to edge devices.

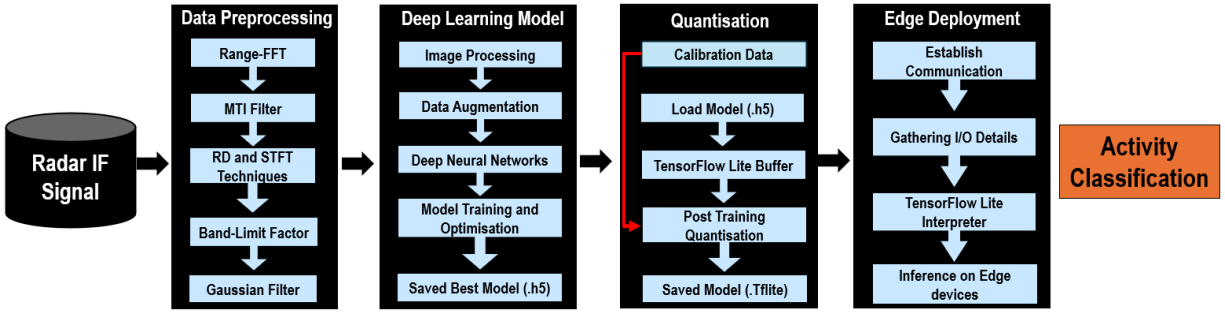
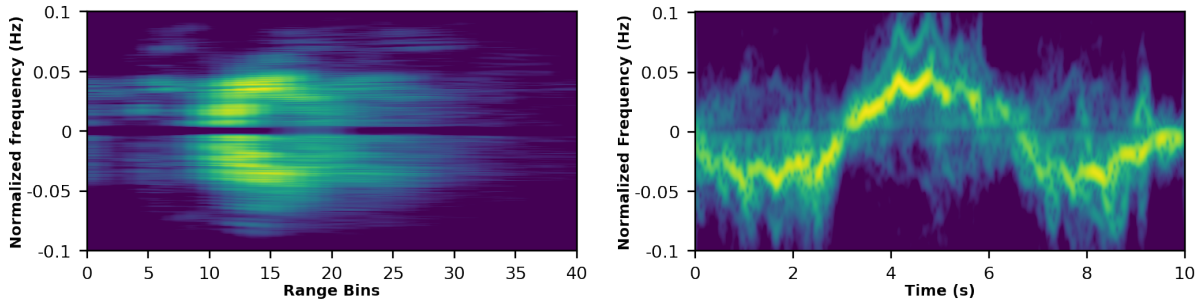


Figure 5.2: Proposed methodology from data acquisition to edge deployment for radar-based macro-activity recognition.

utility. The following subsections provide a detailed explanation of each step in the proposed system, as shown in Fig. 5.2.

5.3.1 Data Preprocessing

As previously described in Chapter 3, the data acquisition process and dataset splitting strategy for FMCW radar remain the same in this Chapter, and the same dataset is used for analysis. Our previous investigation in Chapter 3 showed that two specific data preprocessing techniques



(a) RD map

(b) TF-based STFT map

Figure 5.3: Example of the walking class, acquired using RD and STFT techniques, with Doppler band-limiting and gaussian filtering applied for smoothing.

performed well among the available options for R-HARS. Based on this analysis, we adopt the RD technique and the TF-based STFT approach as the main preprocessing methods in this Chapter, as shown in Fig. 5.2.

To maintain consistency across radar domain representations, the same preprocessing workflow was applied to both the RD and STFT outputs. This included two key steps: Gaussian filtering [248], and Doppler band-limiting. A Gaussian filter with a standard deviation of $\sigma = 0.5$ was applied to suppress high-frequency noise and smooth the edges of the radar maps. This step reduces abrupt intensity variations, leading to cleaner spatial and temporal features, which in turn facilitates more stable convergence during model training. In addition, a Doppler band-limiting filter was applied within the frequency range of -0.1 Hz to 0.1 Hz. This range was selected to preserve the most relevant low-frequency motion components typically associated with macro-activities, such as walking, while discarding irrelevant background noise and micro-movements. Band-limiting helps reduce input dimensionality and focuses the model on discriminative motion patterns, thereby improving feature extraction efficiency and reducing overfitting. As shown in Fig. 5.3A, the RD map highlights Doppler intensity concentrated within range bins 1–40, clearly delineating motion signatures associated with walking. Similarly, Fig. 5.3B presents the STFT representation for a 10-second activity window, where frequency-localised motion components are retained after filtering. This refined preprocessing pipeline enhances the quality of the training data by preserving the high-impact motion features while eliminating redundant or noisy information, ultimately supporting more efficient and robust model learning.

5.3.2 Deep Neural Networks

In this study, we adopted TL based strategy, as detailed in the previous Chapter 3, but with different model selections. The purpose of using various models is to evaluate the edge deployment feasibility of baseline and compressed models. Specifically, we tried four different models to analyse their performance on edge devices, deployment constraints, and inference latency.

The pre-trained DNNs used in this study include Inception-v3, Xception, DenseNet-169, and ResNet-18.

The same class imbalance handling technique discussed in Chapter 3, was applied in this study. Similarly, data augmentation strategies were implemented across all models to enhance the generalisation and adaptability to radar-based data [216]. Augmentation techniques include rotation, shift, scaling, and horizontal flipping to simulate real-world radar variations and to prevent overfitting. For models such as DenseNet-169, Xception, and ResNet-18, STFT based spectrograms were rotated by 10°, offset by 20% in width and height, scaled by 10-20%, and flipped horizontally to ensure robustness to variations. RD based maps were rotated by $5-10^{\circ}$, offset by 10-20% in width and height, scaled by 10%, and flipped horizontally where applicable, further improving the ability to learn features from motion-induced Doppler shifts.

In this study, we employed a set of streamlined yet effective DNNs, each enhanced with a custom classification head specifically designed for fine-tuning radar-based macro-activity recognition tasks. While the backbone networks share a common feature extraction architecture, the custom layers appended to each model are tailored to improve the learning efficiency and generalisation. The classification head begins with a GAP layer, which reduces the spatial dimensions of feature maps while retaining essential semantic information. This is followed by batch normalisation to stabilise the activation distributions and accelerate training convergence. To mitigate overfitting, we applied a dropout layer with a high dropout rate (50-60%), which randomly deactivates neurones during training. Subsequently, a fully connected dense layer is introduced with a ReLU activation function and L2 regularisation, encouraging sparsity and penalising complex weights. An additional round of batch normalisation and dropout further strengthens model robustness and generalisation. Finally, a softmax layer produces the output probabilities across six predefined macro-activity classes, completing a structured and regularised pipeline optimised for real-time R-HARS.

Model Configuration and Hyperparamters Tuning

To ensure compatibility with all selected pre-trained models, radar-based representations were transformed to 224×224 pixels and aligned with the input size requirements of ImageNet pre-trained models. This standardisation allowed the models to effectively leverage TL, utilising ImageNet pre-trained weights for feature extraction before fine-tuning the radar-based macro-activity dataset. The training process was conducted using RD-based and STFT-based radar maps employing four DNNs.

Each model was initialised with pre-trained ImageNet weights and fine-tuned for radar-based classification. Training was conducted using the Adam optimiser, with a learning rate (lr) of 0.0001 for DenseNet-169, and Xception, whereas Inception-v3 and ResNet-18 used a slightly higher lr of 0.0002 for improved convergence. The batch size was set to 32 for all models except DenseNet-169, which used a batch size of 16 owing to the memory constraints. Categorical

cross-entropy loss was applied to optimise the classification performance. The other parameter settings are presented in Table 5.1.

Table 5.1: Parameter settings during DL models training.

Models	Inception-v3	Xception	DenseNet-169	ResNet-18
RD Maps				
Batch Size	32	32	16	32
Learning rate	0.0002	0.0001	0.0001	0.0002
No. of Neurons	128	256	512,256	256,256
L2 regulariser	0.004	0.004	0.0005	0.004
Drop out rate	0.5	0.5-0.6	0.5	0.4
STFT Maps				
Batch Size	32	32	16	32
Learning rate	0.0002	0.0001	0.0001	0.0002
No. of Neurons	128	256	256,256	256,256
L2 regulariser	0.004	0.004	0.004	0.004
Drop out rate	0.5	0.5-0.6	0.4-0.5	0.4

To improve the training efficiency, callbacks were implemented to enhance the model convergence and prevent overfitting, which are details as follows:

- **ModelCheckpoint:** The best model was selected based on validation accuracy.
- **ReduceLROnPlateau:** Adjusted lr dynamically when validation loss plateaued (factor=0.2, patience=7, min lr=1e-7).
- **EarlyStopping:** Training was stopped if validation loss did not improve (patience=10, restore best weights=True).
- **CSVLogger:** Logged training progress for further analysis.

Each model was trained for 100 epochs to ensure optimal performance through adaptive lr and batch size adjustments. The fully connected layers were fine-tuned while retaining the pre-trained convolutional layers, allowing the models to adapt to radar-specific feature learning effectively. This comparative approach provides insights into the performance of various CNN architectures for radar spectrogram classification, optimising feature extraction, and classification for both STFT and RD-based domains.

Energy Tracking during Training

The Emission Tracker from CodeCarbon [249], is a Python API designed to measure and track the energy consumption and carbon footprint of DL model during training on GPU based system. It monitors power usage across key hardware components, including GPU, CPU, and RAM, and

calculates electricity consumption in kilowatt-hour (kwh) and carbon dioxide (CO_2) emissions, expressed as kilograms of (CO_2)-equivalents [kgCO_2eq]. During training, it leverages the pynvml library to track Nvidia GPU power usage, while CPU power is estimated using either Intel power gadget for windows and mac or Intel RAPL files for Linux. If direct CPU tracking is unavailable, the tracker resorts to fallback mode based on predefined power consumption values. Additionally, the RAM energy consumption is approximated based on memory allocation. The CO_2 emissions are determined by multiplying the energy consumed by the carbon intensity of the electricity grid, which is inferred from global databases or country specific data when available [249]. Code carbon runs power measurements at configurable time intervals by default with every 15 second and logs the results into a CSV file for further analysis. This tool provides a lightweight and efficient approach for quantifying and optimising the environmental impacts of computational workloads.

5.3.3 Post-Training Quantisation

In this study, we employed model compression technique such as PTQ in this case [194], to optimise DL models for edge deployment. As shown in Fig. 5.4, PTQ involves calculating scaling factors to convert a full precision or floating-point 32-bit (FP32) model to a lower precision format after training. A representative dataset, called the calibration data, is used to capture the activation distribution of each activation tensor. These distributions are then used to calculate scale values for each tensor, and the weight distribution is used to determine the weight scale [193]. Using PTQ, we developed three versions of the compressed model: a non-quantised version that is directly converted to TensorFlow-Lite (TF-Lite), a FP16 floating point (FP16) quantised version, and INT8 integer (INT8) quantised version. While quantisation may result in a slight loss in accuracy due to the lower precision of weights and activations, the goal is to minimise this loss while significantly improving model efficiency. These modifications adjust the bit width of operations to optimise the model for efficient edge deployment.

To prepare the models for quantisation, radar-based representations were preprocessed by loading and normalising the images using a standard preprocessing pipeline inspired by ImageNet techniques. The pixel values were first scaled between 0 and 1, followed by mean subtraction and STD normalisation to standardise the input features. Class labels were extracted from the image filenames and converted into a categorical format using one-hot encoding to match the classification task requirements. Once preprocessing was completed, different quantisation strategies were applied using TF-Lite to convert the trained model, reduce the memory footprint, and improve inference efficiency. It is important to note that this preprocessing approach is a general implementation and may vary across different models. Each model requires specific preprocessing steps tailored to its input requirements. For FP32, the model was directly converted and saved in TF-Lite format without modifications. In FP16 quantisation, an optimisation setting was enabled to support FP16, leading to lower memory consumption according to

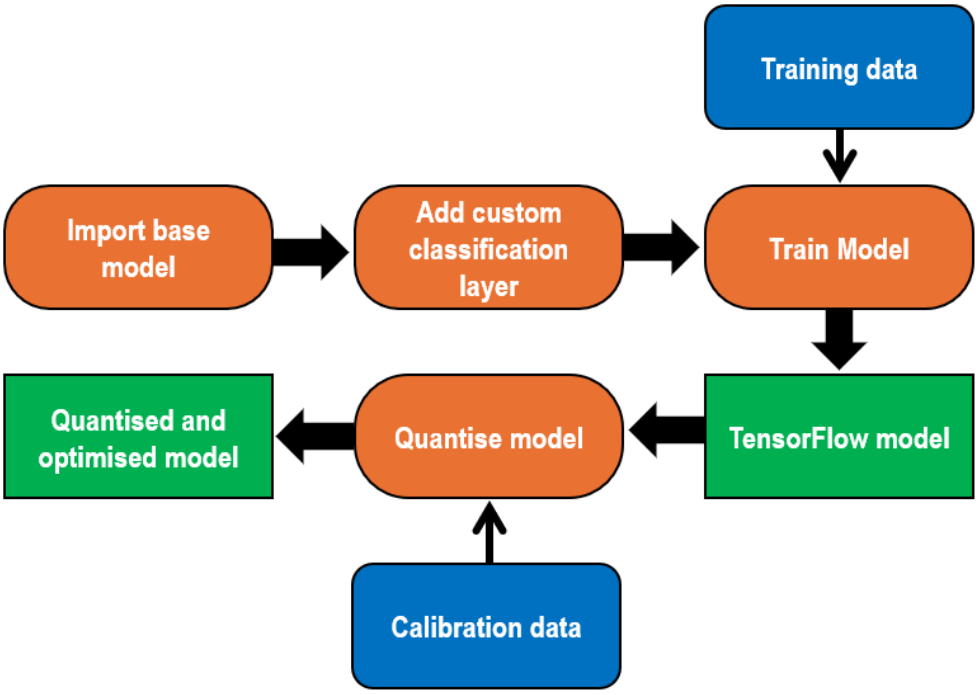


Figure 5.4: Post-training quantisation workflow with transfer learning approach.

the following equation:

$$x_{\text{fp16}} = \text{convert}(x_{\text{fp32}}), \tag{5.1}$$

where the conversion function reduces the precision of the fp numbers by modifying the exponent and mantissa bit sizes according to the IEEE 754 standard [184]. For INT8 quantisation, the converter was configured to enforce an INT8 representation for weights and activations using a representative dataset to ensure calibration. The INT8 quantisation process is as follows [250]:

$$x_{int8} = \text{round}\left(\frac{x_{\text{fp32}}}{s}\right) + z, \tag{5.2}$$

Here x_{fp32} is the original model, s is the scaling factor determined during calibration, and z is the zero point offset for adjusting the range. To ensure that the quantised values remain within valid integer limits, the following clamping function is applied:

$$x_{int8} = \text{clamp}(x_{int8}, -128, 127) \tag{5.3}$$

For general integer quantisation, the final representation is computed as:

$$x_{int} = \text{clamp}(\lfloor x/s \rceil - z, p_{min}, p_{max}) \tag{5.4}$$

Here clamping ensures that the values do not exceed the allowed limits, and P_{min} and P_{max} define the valid integer range.

5.3.4 Local Differential Privacy

In scenarios where raw data must remain private even before leaving the user's device, local differential privacy (LDP) provides a powerful solution. By adding controlled randomness to the outputs, LDP ensures that individual data points remain indistinguishable from one another before inference, preserving privacy while maintaining the usability of the predictions. Unlike traditional DP methods [251], which are often applied in centralised settings to protect data during model training, the LDP operates at the edge device before inference [252]. This ensures that the output predictions remain privacy-preserving, mitigating the risk of sensitive information leakage without affecting the model training phase [253]. As illustrated in Fig. 5.1, before generating predictions, LDP is applied to the output, ensuring that even if an adversary gains access to the inference results, they cannot infer sensitive information about the original data. Formally, LDP guarantees that a randomised mechanism $M : X \rightarrow R$, which maps an input domain X to an output range R, satisfies epsilon ε-LDP if, for any two neighboring input data points $x, x' \in X$, and for all measurable sets $S \subseteq R$, the following inequality holds [254]:

$$Pr[M(x) \in S] \leq e^{\varepsilon} \cdot Pr[M(x') \in S], \tag{5.5}$$

where ε represents the privacy budget, which governs the level of oneness between perturbed outputs, and $P_r[\cdot]$ is the probability of $(\cdot)$. A smaller ε value offers stronger privacy guarantees, as the perturbed outputs become highly similar, making it difficult for an adversary to differentiate between them. Conversely, a higher ε value weakens privacy protection by making the outputs more distinguishable. In this study, LDP is applied before inference at the edge, where the perturbation probability of retaining or modifying a class label is determined by β, given by [254]:

$$\beta = \frac{e^{\varepsilon/K} - 1}{e^{\varepsilon/K} - 1 + C} \tag{5.6}$$

where K represents the number of perturbation steps, and C is the total number of class labels. This equation dynamically adjusts the perturbation probability based on the privacy budget and classification complexity, ensuring a balance between privacy protection and model utility.

Noise Injection Mechanism for LDP

To enforce the LDP at the edge, a randomised response mechanism was employed, which introduced controlled noise into the inference process. This noise injection ensures that, even if an adversary gains access to the predictions, they cannot confidently determine whether they correspond to the actual class or a perturbed label. The mechanism operates by selectively perturbing the predicted labels according to a Bernoulli-distributed random variable, ensuring controlled randomness while preserving the usability [254]. For each predicted class label, the

noise-injection mechanism is as follows:

$$\hat{y} = \begin{cases} y, & \text{with probability } \beta \\ y', & \text{with probability } (1-\beta), \end{cases} \tag{5.7}$$

where y represents the original predicted label, and y' is a randomly selected label from the set of possible classes. The probability of retaining the true label is determined by β, while the probability of perturbing the label is $1-\beta$. This ensures that each output remains privacy-preserving without completely degrading the inference accuracy. By applying LDP at the edge before inference, the system ensures strong privacy guarantees while maintaining model performance. Because perturbation is applied only to inference outputs, it does not interfere with the training process on the server, allowing the model to learn from clean, unaltered data.

In addition, applying privacy protection at the edge reduces the communication overhead, as no differentially private transformations are required during data transmission. The localised application of LDP ensures that privacy is enforced at the exact point of decision-making, preventing the potential exposure of sensitive user data while still enabling real-time inference. This approach is particularly beneficial for edge-based HAR, where privacy concerns arise owing to continuous monitoring and personal data collection. By integrating LDP at the inference stage, the system achieves a robust trade-off between privacy and utility, ensuring that privacy-preserving decisions can be made directly at the edge, while leveraging DL models trained on a central server.

5.3.5 Deployment Setup

In this study, we employed three different hardware platforms along with various software frameworks to facilitate DL model training, optimisation, and deployment, as shown in Table 5.2. The DNN training process was conducted on a GPU-based system, leveraging TensorFlow and Keras, which are widely used in both academia and industry owing to their flexibility and computational efficiency [255]. These frameworks provide a good trade-off between platform compatibility and performance, making them ideal for DL applications. All data preprocessing, model training, optimisation, quantisation, and LDP processing were performed on the GPU-based system to accelerate computation, significantly reducing the processing time compared with CPU-based systems. The GPU hardware used in this study is detailed in Chapter 3, and further specifications are summarised in Table 5.2. For DL framework compatibility, we used: TensorFlow 2.10 and Keras 2.10 for training and model development. Scikit-learn for data preprocessing and feature engineering. TF-Lite for model quantisation and edge deployment, and Python 3.11 as the programming environment.

Table 5.2: DL Model Training and deployment Platforms.

Platforms	**Nvidia GeForce RTX 3090**	**RPi 4 Model B**	**Nvidia Jetson Nano**
Processor	1.5 GHz	64-bit Quadcore	1.43 GHz Quadcore
Framework (OS)	Cuda 11.4	Debian 64-bit Os	Ubunto 20.04
Memory	8 GB	4 GB	2 GB
Usage	Training	Edge Inference	
Type	Installed in CPU	Itself a CPU	GPU + CPU
DL Framework	Full Tensorflow	Tensorflow and TF-Lite	

Edge Hardware Configuration

To evaluate the performance of edge devices for real-time inference, this study employed two distinct SBCs: the RPi 4 and the NVIDIA JNano. The device characteristics are listed in Table 5.2. Each SBC was tested for its ability to execute DL models efficiently, with a focus on inference latency, energy consumption, and deployment feasibility in edge computing environments. The RPi is used in this study, as a cost-effective and reliable edge computing platform for real-time HAR using FMCW radar data [256]. In a practical deployment scenario, the FMCW radar continuously captures raw IF signals, which are subsequently processed directly on the RPi. A custom Python script running on the device executes the optimised TF-Lite model, enabling low-latency prediction of human macro-activities in real time, while operating under constrained computational and memory resources.

In contrast, NVIDIA JNano is a compact system-on-module (SoM) designed for executing multiple DL networks concurrently. It provides a better balance between computational power and energy efficiency, making it well suited for embedded AI applications. Operating under a low power requirement of less than 5 W, JNano excels in tasks such as image classification, object detection, semantic segmentation, and HAR [257]. To further enhance the DL inference efficiency on the JNano, this study utilised TF-Lite, an inference optimisation toolkit along with PTQ which can run quantised and optimised models on JNano efficiently.

By employing both the RPi 4 and JNano, this study provides a comparative analysis of their performance in real-time HAR tasks. The RPi 4, which is energy-efficient and cost-effective, is limited by its computational capabilities and is best suited for simpler tasks. On the other hand, JNano, with its superior processing power, demonstrates better performance in handling complex DL models, albeit at a slightly higher cost and energy consumption. This comparison highlights the trade-offs between computational capability, energy efficiency, and cost, and offers valuable insights for selecting an appropriate edge device based on specific application requirements.

The INA219 sensor was employed for real-time current, voltage, and power monitoring

in an edge hardware setup [258]. It is a high-precision bidirectional power monitoring circuit capable of measuring up to 26V and 3.2A with a resolution of 0.8mA. Operating within a voltage range of 3.0V to 5.5V, the sensor uses the I²C communication protocol, enabling seamless data acquisition by embedded systems. The INA219 breakout board integrates a 0.1 ohm shunt resistor, labelled as R100 on sensor, which measures voltage drop across the load, allowing current computation using Ohm's Law: $I = VR$, where I represents the current, V is the measured voltage drop, and R is the shunt resistance. With the default 0.1 ohm resistor, the maximum measurable current was 3.2A. For higher current applications, the shunt resistor can be replaced with a lower-value resistor while considering power dissipation:$P = I^2R$ where P represents the heat dissipation in watts.

RPi and JNano both operate at 5V/3A and share the same connection scheme. However, for simplicity, Fig. 5.5 illustrates only the RPi setup. A USB Type-C breakout board was used to supply 5V/3A power, which was also monitored using INA219. The $V_{\text{in}}+$ terminal of INA219 connects to the positive terminal of the power source, while $V_{\text{in}}-$ connects to the RPi, which power up the devices, and enabling the measurement of current consumption. To interface the INA219 sensor with the RPi or JNano, the following connections were established:

1. The VCC of INA219 is connected to 3.3V (pin 1) on the RPi.
2. The GND of INA219 is connected to the GND (Pin 39) to ensure a common reference voltage.
3. The SDA (I²C data line) of INA219 is connected to GPIO2 (pin 3).
4. The I²C clock line (SCL) of INA219 is connected to GPIO3 (Pin 5).

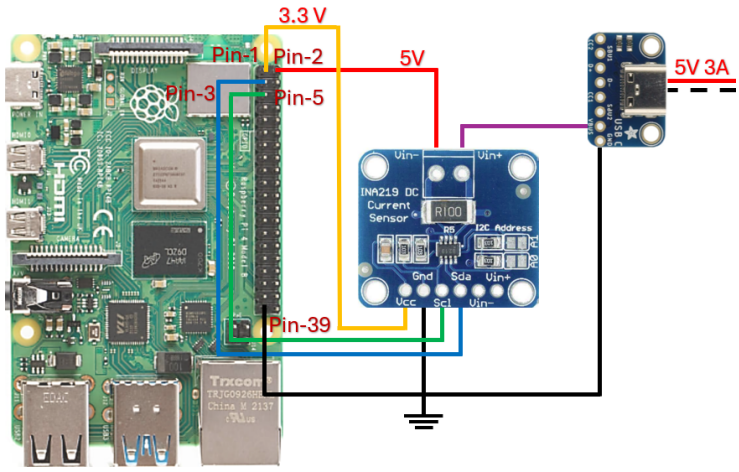


Figure 5.5: Connection setup between the RPi, INA219 sensor, and Type-C breakout board for device powering and real-time current and power measurement.

5.3.6 Performance and Computational Metrics

A rigorous evaluation of the predictive model is essential to demonstrate the effectiveness of the edge-optimised R-HARS. To achieve this, various metrics were analysed to comprehensively assess the efficiency and effectiveness of the system.

Predictive Model Evaluation

The HAR task is treated as a multiclass classification problem based on the collected radar data. To assess the model's performance, we evaluated both training and validation accuracy, which measures how well the model learns from the training data and generalises to unseen validation data. Training accuracy represents the proportion of correctly classified instances in the training dataset, whereas validation accuracy helps assess the model's generalisation ability and detect potential overfitting.

While accuracy is a commonly used metric during training and testing of a model, it can be misleading in the presence of class imbalance owing to the accuracy paradox. To ensure a balanced evaluation, we also consider additional performance metrics, including: Precision= $\frac{TP}{TP+FP}$, Recall= $\frac{TP}{TP+FN}$ and F1-score= $2X\frac{PrecisionXRecall}{Precision+Recall}$, where TP (True Positives), FP (False Positives), and FN (False Negatives) quantify the model's classification performance. Precision measures the number of predicted positive instances that are actually correct, recall evaluates how well the model captures actual positive cases, and the F1-score provides a harmonic mean of precision and recall for a balanced assessment. These metrics offer a deeper insight into the predictive power of the model, particularly for handling variations in human activities captured by radar signals.

Energy Efficiency and Latency Analysis

To evaluate the computational efficiency and environmental impact of the training phase, we analysed GPU energy consumption, power usage, and CO_2 emissions. The total energy consumed by the GPU during training was estimated using integrated power monitoring tools, such as NVIDIA's pynvml via nvidia-smi with help of codecarbon. The energy consumption was computed as follows: $E_{\text{GPU}} = P_{\text{GPU}} \times T$, where P_{GPU} represents the power consumption of GPU during training in watts, and T is the training duration. The CO_2 emissions generated by GPU energy consumption were estimated using: $CO_{2\text{GPU}} = E_{\text{GPU}} \times C$, C is the carbon intensity of electricity $(kgCO_2/kWh)$, which depends on the energy source of the local grid. If specific grid data is unavailable, a default world average of 0.475 $(kgCO_2/kWh)$ is used [249].

To account for the complete energy usage of the system, the CPU and RAM power consumption are included in the total energy estimation. CPU power is measured using an Intel Power Gadget or inferred from Intel RAPL files, whereas RAM power consumption is approximated as: $P_{\text{RAM}} = 3W\text{Per}8GB$ [249]. The total system energy consumption during training is then

calculated as:

$$E_{\text{total}} = E_{\text{GPU}} + E_{\text{CPU}} + E_{\text{RAM}}, \tag{5.8}$$

where each term represents the energy consumed by the GPU, CPU, and RAM during the training period. For edge deployment, the energy consumption and power usage were evaluated to measure the efficiency of inference under real-time conditions. Because edge devices have limited computational resources, an optimised and quantised model was deployed to reduce both power consumption and latency. The INA219 sensor was used to measure power consumption during inference. The total energy consumed during inference is then calculated as: $E_{\text{edge}} = P_{\text{edge}} \times t_{\text{inf}}$, where P_{edge} is total power in mW measure during inference using INA219 sensor, and t_{inf} is the inference time per sample in second.

Finally, the latency of the system was evaluated based on the time required for data capture by the radar (t_{cap}), data preprocessing (d_{px}), and inference (t_{inf}) on the edge device. The total end-to-end latency D is expressed as:

$$D = t_{cap} + d_{px} + t_{inf}. \tag{5.9}$$

This latency metric helps to analyse the trade-off between on-device inference and energy efficiency.

5.4 Results and Discussion

In this section, we present a comprehensive analysis of our evaluation metrics, performance outcomes, and computational measurements obtained during the DL model training and edge deployment. Our research aims to assess the effectiveness of radar-based macro-activity recognition for real-time applications, particularly in AAL and Assisted Smart Living (ASL). We evaluated these systems in multiple dimensions that include accuracy, energy efficiency, and adaptability when processing radar data. Given the increasing emphasis on privacy-preserving technologies, radar-based sensing offers a privacy-aware alternative to the conventional vision-based systems. However, to further enhance privacy, we incorporated LDP, ensuring both model and data privacy during inference. Finally, we provide a thorough energy efficiency comparison and latency analysis, highlighting the practical advantages of performing on-device inference on edge devices.

5.4.1 Computational Efficiency and Model Performance

A comprehensive analysis of the outcomes measured and estimated using GPU based server during the offline training and testing phases. Our study evaluated four DL models in conjunction with two radar representations to determine both the efficiency and performance characteristics

of R-HARS intended for real-world settings. When developing edge AI systems, two complementary aspects require careful assessment: classification accuracy and computational efficiency (measured using metrics such as inference time, power consumption, and energy-precision ratio). The classification accuracy depends critically on proper model training, avoiding both overfitting and underfitting, to ensure that the models generalise effectively to unseen data when deployed on edge devices. During the training phase, we measured the computational efficiency of each DL model to identify the optimal combination of model architecture and radar representation for future deployment scenarios.

Energy-aware Training Efficiency

During training, we assessed the computational efficiency of the system by measuring GPU power consumption, energy usage, and CO_2 emissions, as summarised in Table 5.3. The results indicate that Xception exhibits the highest GPU power usage, consuming 155.33 W with RD and 157.88 W with STFT, leading to increased energy consumption and environmental impact. In contrast, ResNet-18 consumed 100.74 W with RD and 94.21 W with STFT, indicating that STFT-based training requires less power. Although Xception and DenseNet-169 require more power and energy for training, their higher computational demand may be justified by their classification performance, which is evaluated in the next section.

Table 5.3: Energy consumption and environmental impact of DL models during training with different radar representations.

Data Type	Models	GPU Power (W)	Energy (kWh)	Energy/ Epoch (kWh)	CO_2 (g.CO_2eq/s)
RD	Inception-v3	133.55	0.018	4.50×10^{-4}	0.0116
	Xception	155.33	0.043	8.52×10^{-4}	0.0133
	DenseNet-169	137.73	0.046	8.54×10^{-4}	0.0119
	ResNet-18	100.74	0.007	2.94×10^{-4}	0.0094
STFT	Inception-v3	139.24	0.019	4.40×10^{-4}	0.0120
	Xception	157.88	0.070	8.57×10^{-4}	0.0133
	DenseNet-169	119.35	0.028	7.91×10^{-4}	0.0108
	ResNet-18	94.21	0.006	2.68×10^{-4}	0.0091

DL Model Performance

The performance of the models was evaluated using unseen test data. As discussed earlier (see Chapter 3), the dataset was split such that the test set consisted of 126 samples from seven subjects, who were completely excluded from the training and validation sets. This setup ensured that the generalisation capability of the models could be assessed. All four models were tested using unseen data. For the RD maps, the highest recognition accuracy of 95.24% was achieved

by the Xception model. In comparison, for the STFT maps, the Inception-v3 model performed the best at 98.41%. A summary of the accuracy and other performance metrics for all models is provided in Table 5.4.

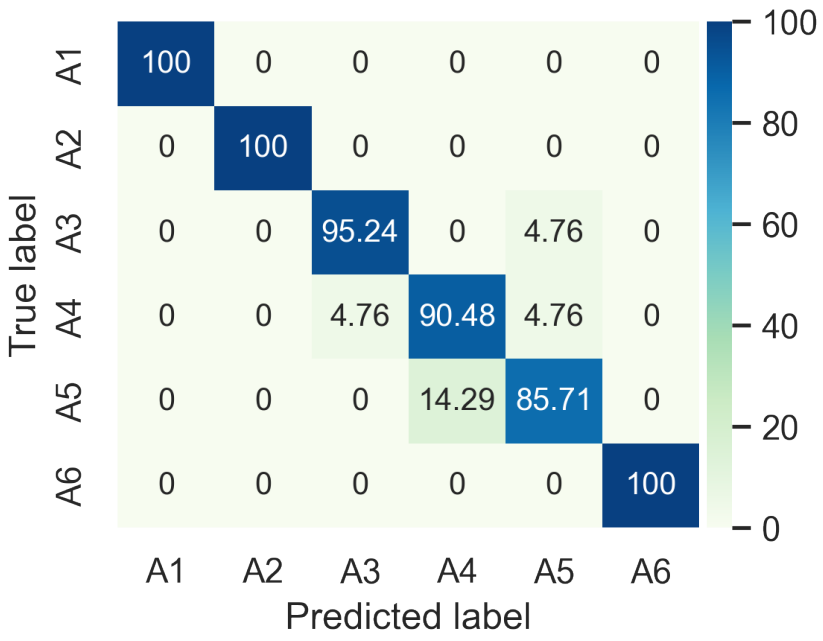
Table 5.4: Performance metrics of all four models with their radar representations.

Data Type	Models	Accuracy (%)	Precision	Recall	F1-score
RD	Inception-v3	92.86	0.9295	0.9286	0.9274
	Xception	95.24	0.9527	0.9524	0.9524
	DenseNet-169	93.65	0.9381	0.9365	0.9351
	ResNet-18	94.44	0.9441	0.9444	0.9434
STFT	Inception-v3	98.41	0.9845	0.9841	0.9841
	Xception	94.44	0.9508	0.9444	0.9440
	DenseNet-169	96.83	0.9688	0.9683	0.9678
	ResNet-18	96.03	0.9598	0.9603	0.9597

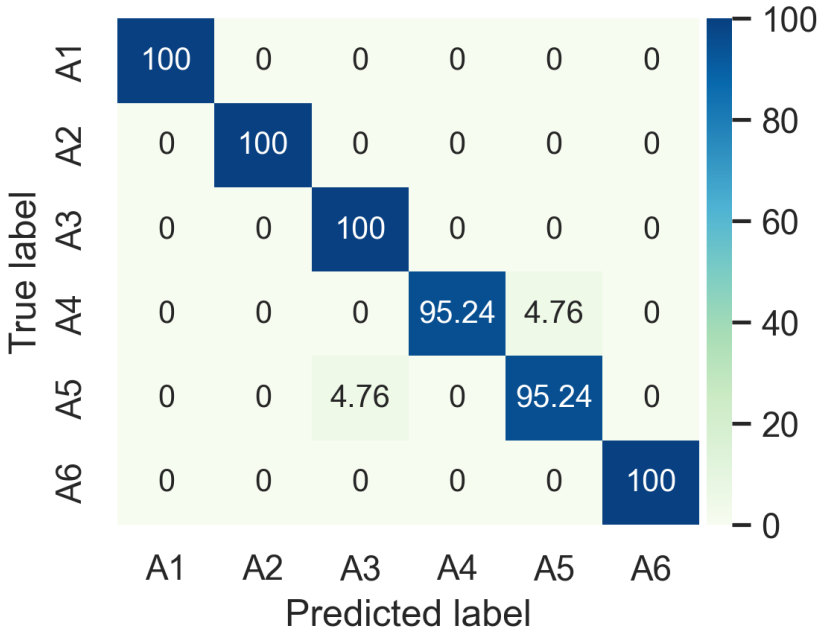
The confusion matrices for the best-performing models such as Xception for the RD maps and Inception-v3 for the STFT maps, are shown in Fig. 5.6. Both models achieved 100% recognition accuracy for fall detection, which is crucial because misclassification of falls can have serious consequences. In addition, no class was misclassified as a fall, reinforcing the reliability of these models for critical activity recognition. However, some misclassifications were observed, like picking of an object (A4) and drinking water (A5) activities were occasionally confused, particularly in the Xception model using the RD maps. This can be attributed to the similar hand movements involved in both actions, which led to misinterpretations of the radar signal.

Generalisation on other Radar Datasets

To validate the effectiveness of our framework across different radar frequencies, we applied the same cross-frequency generalisation strategy described in Chapter 3. The dataset and preprocessing pipeline remained consistent; however, the models used in this evaluation were Xception for RD and Inception-v3 for STFT, as they achieved the highest recognition accuracy on the 5.8 GHz radar dataset. Table 5.5 presents the multi-frequency generalisation performances of these optimal domain-model pairs. The reported accuracies reflect the average performance of the six test subjects, along with their standard deviations. The results demonstrate that our preprocessing methods and TL models maintain strong performance across varying radar frequencies, with STFT and Inception-v3 achieving an impressive accuracy of 94.39% ± 2.40 using 24 GHz radar data. This multi-frequency evaluation confirmed the robustness of our approach and its potential applicability in diverse radar-based macro-activity recognition system, regardless of the operating frequency. It also highlights the feasibility of using different radar frequencies in edge-based HAR scenarios.



(a) Xception model with RD map



(b) Inception-v3 model with STFT map

Figure 5.6: Confusion matrices of models that shows highest accuracy for RD and STFT maps as input.

Impact of Quantisation on Model Performance

In this study, we applied three different PTQ methods: no quantisation, FP16, and INT8 quantisation. For RD maps, all PTQ methods significantly reduced the model sizes, making them more suitable for deployment on resource-constrained edge devices, as shown in Table 5.6. The

Table 5.5: Generalisation performance of optimal radar domain-model pairs across different frequencies using LOSO-CV.

Domain	Model	Freq.	Acc. (%)	Precision	Recall	F1-score
RD	Xception	24 GHz	88.22±2.63	0.8913	0.8823	0.8821
		77 GHz	84.17±2.06	0.8570	0.8418	0.8404
STFT	Inception-v3	24 GHz	94.39±2.40	0.9530	0.9493	0.9496
		77 GHz	92.75±1.34	0.9319	0.9275	0.9262

sizes remained consistent across the RD and STFT inputs, with INT8 achieving the highest compression up to ∼12x reduction factor in size. However, INT8 quantisation caused slight accuracy drops for the RD maps, most notably for DenseNet-169, which decreased from 93.65% to 89.68%. However, the performance remains within an acceptable range for real-time edge deployment. In contrast, for STFT maps, all models maintained their accuracy across the different PTQ methods, confirming that the quantisation effects depend on the input representation. Regardless of the input type, all the models benefited from reduced sizes, further enhancing their feasibility for edge deployment.

Table 5.6: Effect of PTQ on model size and accuracy for different input representations.

Model	Model Size (MB)				Accuracy (%)
	FP32	No-quant	FP16	INT8	FP32 (RD) / INT8 (RD)
Inception-v3	253.36	84.10	42.11	21.52	92.86 / 93.65
Xception	245.04	81.33	40.71	21.62	95.24 / 93.65
DenseNet-169	157.02	51.50	25.89	13.38	93.65 / 89.68
ResNet-18	130.58	43.42	21.73	11.00	94.44 / 93.65

5.4.2 Edge System Assessment

The computational cost of deploying DL models on edge devices has been evaluated based on the inference time, power consumption, model size, and Energy-Precision Ratio (EPR). These metrics provide critical insights for selecting the optimal combination of radar data representation and DL models for edge-based system deployment. The power consumption was measured using an INA219 sensor connected to both RPi and JNano to ensure accurate measurement of the current and power during model inference. The EPR metric, which provides a complete evaluation of the accuracy-energy efficiency trade-offs, was computed as: $EPR = Error \times EPI$, where Error represents the classification error rate (1 - accuracy) and EPI (Energy Per item) is calculated by multiplying inference time with incremental power consumption.

To establish meaningful comparisons, baseline power measurements were first recorded during the idle state by averaging the readings over 10 iterations. For inference, all metrics were

Table 5.7: FP16 quantised model computational cost on Edge Platforms.

Model	Inference Time ± SD (ms)		Energy ± SD (mWh)		EPR (mWh)	
	RPi	JNano	RPi	JNano	RPi	JNano
RD maps						
Inception-v3	659.58 ± 0.82	503.56 ± 2.23	0.23 ± 0.03	0.34 ± 0.008	0.016	0.024
Xception	1138.10 ± 1.71	966.98 ± 13.90	0.46 ± 0.04	0.63 ± 0.061	0.022	0.030
DenseNet-169	821.89 ± 2.04	668.67 ± 4.55	0.27 ± 0.04	0.44 ± 0.076	0.017	0.028
ResNet-18	452.98 ± 20.64	358.45 ± 2.66	0.17 ± 0.02	0.24 ± 0.014	0.009	0.013
STFT maps						
Inception-v3	672.38 ± 2.27	516.51 ± 1.80	0.23 ± 0.025	0.33 ± 0.008	0.004	0.005
Xception	1139.80 ± 1.38	977.10 ± 8.79	0.46 ± 0.037	0.63 ± 0.035	0.026	0.035
DenseNet-169	846.98 ± 46.73	668.55 ± 3.21	0.29 ± 0.074	0.43 ± 0.015	0.009	0.014
ResNet-18	450.85 ± 9.01	358.39 ± 3.12	0.16 ± 0.023	0.24 ± 0.013	0.006	0.010

computed over 40 iterations to ensure the statistical reliability. Table 5.7, presents the computational cost of the FP16 quantised models. For the RD maps, ResNet-18 achieved the best balance of accuracy and efficiency, with the lowest EPR values of 0.009 mWh for RPi and 0.013 mWh for JNano, while maintaining the second fastest inference times of 453 ms on RPi and 358 ms on JNano. Despite Xception's strong classification accuracy, it exhibits the highest inference times and power consumption, leading to EPR values approximately 2.5 times higher than those of ResNet-18. For STFT maps, Inception-v3 emerged as the most efficient model, achieving exceptionally low EPR values of 0.004 mWh for RPi and 0.005 mWh for JNano owing to its high classification accuracy of 98.41%. Despite its moderate energy consumption, the significantly low error rate compensates for power requirements, making it the optimal choice for the STFT domain. DenseNet-169 and ResNet-18 also demonstrated competitive performance, with EPR values notably lower than Xception, further reinforcing their suitability for edge deployment.

The performance of the INT8 quantised models, summarised in Table 5.8, demonstrates substantial improvements in inference time and energy efficiency across all architectures. Compared to FP16, INT8 quantisation reduces the inference time by approximately 50% while simultaneously decreasing the energy consumption by 30-50%. For the RD maps, ResNet-18 remained the most efficient model, achieving the lowest EPR values of 0.007 and 0.008 mWh for RPi and JNano, respectively. The slight accuracy drop observed with INT8 quantisation from 94.44% to 93.65% was compensated by significant reductions in inference time and power consumption. For STFT maps, INT8 quantisation delivers even more impressive gains, particularly for Inception-v3, which achieves an EPR of only 0.002 mWh on RPi and 0.003 mWh on JNano. This marks a 50% reduction in EPR compared with its FP16 counterpart, with no loss in classification accuracy. Similarly, ResNet-18 with STFT maps demonstrated exceptional energy

Table 5.8: INT8 quantised model performance metrics with Energy-Precision Ratio on Edge Platforms.

Model	Inference Time ± SD (ms)		Energy ± SD (mWh)		EPR (mWh)	
	RPi	JNano	RPi	JNano	RPi	JNano
RD map						
Inception-v3	323.19 ± 0.41	301.78 ± 1.43	0.12 ± 0.015	0.20 ± 0.005	0.009	0.014
Xception	453.42 ± 0.87	492.35 ± 8.32	0.20 ± 0.02	0.32 ± 0.031	0.010	0.015
DenseNet-169	387.45 ± 1.12	367.91 ± 2.87	0.14 ± 0.02	0.28 ± 0.045	0.009	0.018
ResNet-18	221.35 ± 9.86	201.43 ± 1.58	0.13 ± 0.01	0.14 ± 0.008	0.007	0.008
STFT map						
Inception-v3	328.64 ± 1.13	308.65 ± 1.07	0.12 ± 0.013	0.20 ± 0.005	0.002	0.003
Xception	454.77 ± 0.69	498.48 ± 5.13	0.20 ± 0.019	0.33 ± 0.024	0.011	0.018
DenseNet-169	403.12 ± 23.19	369.87 ± 1.93	0.18 ± 0.037	0.41 ± 0.009	0.006	0.013
ResNet-18	220.63 ± 4.57	202.51 ± 1.68	0.08 ± 0.012	0.14 ± 0.007	0.003	0.006

efficiency, achieving an EPR of 0.003 mWh for RPi and 0.006 mWh for JNano.

The best models based on EPR values are listed in Table 5.9, reflecting the optimal trade-off between the classification accuracy and energy efficiency. For RD representations, ResNet-18 consistently emerged as the best model, maintaining high accuracies of 94.44% for FP16 and a reduced 0.8% for INT8. This demonstrates the lowest EPR values across both edge platforms. For STFT representations, Inception-v3 proved to be the most efficient, particularly with INT8 quantisation, where it achieved the lowest EPR observed in the entire evaluation while maintaining an exceptional accuracy of 98.41%.

Table 5.9: Best performing domain-model pairs based on EPR with accuracy and inference time values.

Data Type	Bit Precision	Platform	Best Model	Accuracy (%)	Inference Time (ms)	EPR (mWh)
RD	FP16	RPi	ResNet-18	94.44	453.0	0.009
		JNano	ResNet-18	94.44	358.5	0.013
	INT8	RPi	ResNet-18	93.65	221.4	0.007
		JNano	ResNet-18	93.65	201.4	0.008
STFT	FP16	RPi	Inception-v3	98.41	672.4	0.004
		JNano	Inception-v3	98.41	516.5	0.005
	INT8	RPi	Inception-v3	98.41	328.6	0.002
		JNano	Inception-v3	98.41	308.7	0.003

An in-depth analysis of ResNet-18 with the RD domain, as illustrated in Fig. 5.7, highlights the advantages of quantisation. The model size was significantly reduced from 130.58 MB (32-

bit) to 11.00 MB (INT8), while the accuracy remained remarkably stable, dropping only slightly from 94.44% to 93.65%. Additionally, inference time is reduced by ∼50%, from 453 ms to 221 ms on RPi and from 358 ms to 201 ms on JNano, while energy consumption decreases by over 20% from 0.17 mWh to 0.13 mWh on RPi. These improvements make ResNet-18 (INT8 model) an ideal candidate for resource-constrained edge deployment, offering a strong balance between efficiency and accuracy.

However, a key limitation of our study is that Inception-v3 (32-bit) cannot be evaluated on both edge devices because of its large model size, which causes both systems to become unresponsive during inference. Consequently, we only present the results for ResNet-18 (32-bit), as its size remained within acceptable limits for edge deployment. This highlights a critical constraint: larger DL models may not always be feasible for direct deployment on low-power edge devices without quantisation. The STFT with INT8 Inception-v3 on RPi demonstrated the lowest EPR values with the highest accuracy of 98.41%, making it the best choice for real-time macro-activity recognition using FMCW radar. This result further emphasises the importance of radar data representation selection because STFT maps paired with efficient quantisation strategies can deliver high-performance HAR models suitable for real-world applications.

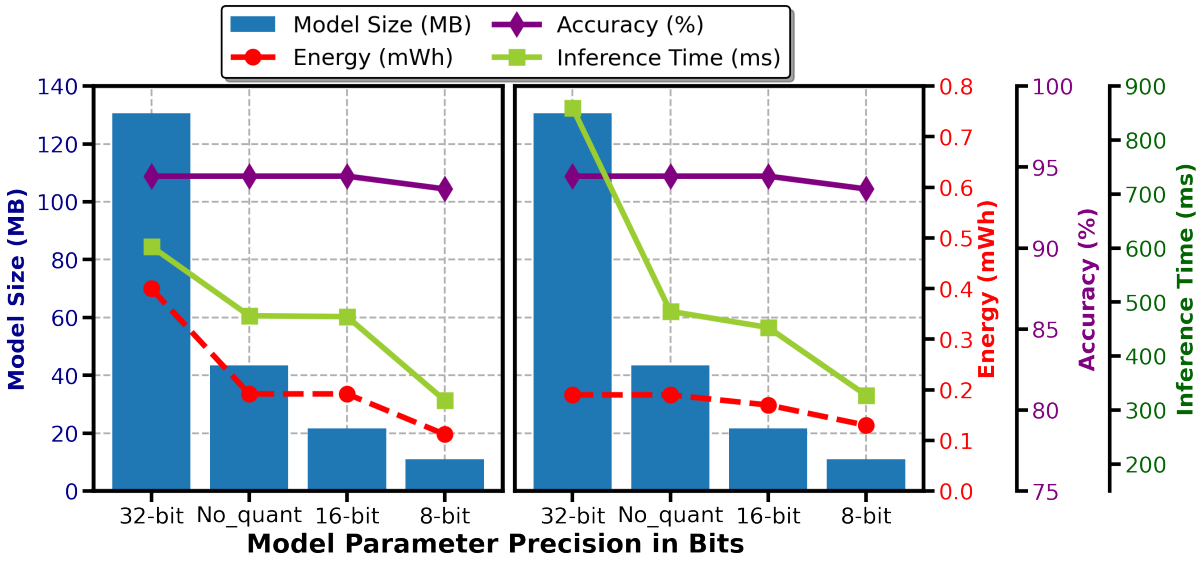


Figure 5.7: The comparison between the reduction in communication overhead, computed as energy estimates for different quantisation using RD as input with ResNet-18 on JNano (left) and RPi (right).

Performance of Radar Representations on Edge Devices

The evaluation of radar domain processing on edge platforms highlights the key considerations for battery-operated systems. As shown in Table 5.10, RD processing consistently outperformed STFT in terms of efficiency, with ∼20% faster processing times and 16% lower energy consumption on RPi, regardless of activity duration. This efficiency advantage is crucial for real-

time applications, as RD processing for a 10 seconds activity completes in just 2.13 seconds, consuming only 0.61 mWh, enabling thousands of inferences on a single battery charge. On the JNano, the energy advantage of RD processing is even more pronounced, consuming 40-50% less energy than STFT, making it the most power-efficient option for edge-based HAR deployment. Although STFT remains viable, its higher energy demand may be a limiting factor for battery-powered devices.

Table 5.10: Computational metrics for radar domain processing on edge platforms over 20 iterations.

Edge Device	Data Type	Activity Length	Processing Time (s)	Energy (mWh)
RPi	RD	10 s	2.13 ± 0.046	0.61 ± 0.05
		5 s	1.10 ± 0.034	0.29 ± 0.026
	STFT	10 s	2.60 ± 0.035	0.73 ± 0.094
		5 s	1.40 ± 0.045	0.37 ± 0.051
JNano	RD	10 s	2.36 ± 0.027	0.97 ± 0.26
		5 s	1.20 ± 0.014	0.35 ± 0.03
	STFT	10 s	2.65 ± 0.024	1.60 ± 0.073
		5 s	1.38 ± 0.013	0.80 ± 0.035

End-to-End Edge System Evaluation

The final evaluation of end-to-end system performance integrates both radar domain processing and model inference, providing a complete assessment of energy efficiency and latency across edge devices. As summarised in Table 5.11, the optimal configuration for battery-operated radar-based HAR is a RPi running INT8 quantised ResNet-18 with RD processing, offering the lowest total energy consumption of 0.42 mWh for 5-second activities, and 0.74 mWh for 10-second activities, and fastest response times of 1.32 second and 2.35 second, respectively. This configuration ensures real-time feasibility, as processing and inference are completed within the activity duration. While JNano provides faster inference times, its higher energy consumption, particularly for STFT data up to 1.80 mWh for 10s activities, makes it less ideal for battery-powered applications. Notably, STFT with Inception-v3 on RPi delivers competitive performance of 0.49 mWh for 5-second activities, and 0.85 mWh for 10-second activities, but still consumes $\sim$16% more energy than RD processing with ResNet-18.

5.4.3 Impact of Local Differential Privacy on Model Performance

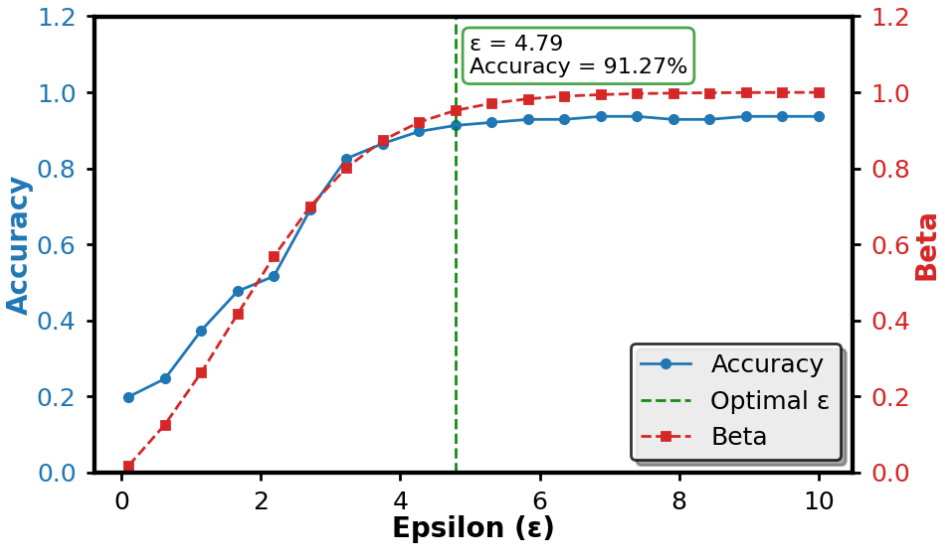
To assess the privacy preservation trade-off in our proposed macro-activity recognition edge-based system, we applied LDP to the top-performing INT8 quantised models of ResNet-18 for

Table 5.11: An end-to-end edge system analysis in terms of computational cost for R-HAR.

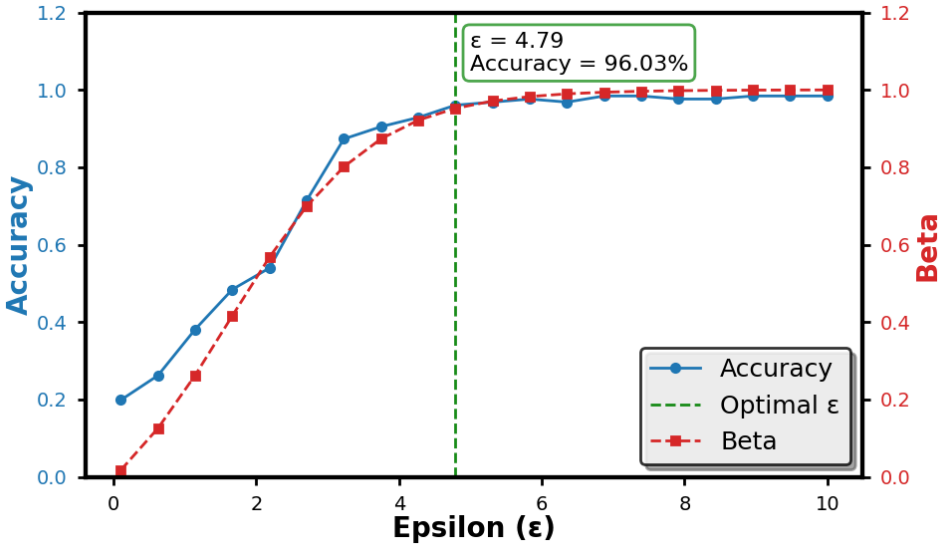
Edge Device	Data Type	Activity Length	Model	Precision	Total Energy (mWh)	Total Latency (s)
RPi	RD	10s	ResNet-18	FP16	0.78	2.58
			ResNet-18	INT8	0.74	2.35
		5s	ResNet-18	FP16	0.46	1.55
			ResNet-18	INT8	0.42	1.32
	STFT	10s	Inception-v3	FP16	0.96	3.27
			Inception-v3	INT8	0.85	2.93
		5s	Inception-v3	FP16	0.60	2.07
			Inception-v3	INT8	0.49	1.73
JNano	RD	10s	ResNet-18	FP16	1.21	2.72
			ResNet-18	INT8	1.11	2.56
		5s	ResNet-18	FP16	0.59	1.56
			ResNet-18	INT8	0.49	1.40
	STFT	10s	Inception-v3	FP16	1.93	3.17
			Inception-v3	INT8	1.80	2.96
		5s	Inception-v3	FP16	1.13	1.90
			Inception-v3	INT8	1.00	1.69

RD and Inception-v3 for STFT, which demonstrated high classification accuracy and energy efficiency on edge devices. Fig. 5.8, presents two subplots: (A) RD with ResNet-18 and (B) STFT with Inception-v3, illustrating the relationship between the privacy budget (ε), perturbation probability (β), and model accuracy, highlighting the impact of privacy preservation on classification performance. The results show a clear trade-off between privacy protection and model accuracy. At low ε values between 0.1–1.0, where privacy guarantees are strongest, accuracy drops significantly from 19.84–26.19% for RD, and 20–26% for STFT owing to the high level of randomisation introduced by LDP. As ε increases from 2.0 to 3.0, both β and accuracy improve, with a notable inflection point at $\varepsilon \approx 3.0$, where the accuracy rises to 82.54% for RD and 87.30% for STFT. Beyond $\varepsilon = 4.0$, the accuracy for both models surpasses 89%, demonstrating that privacy protection can be maintained without severely degrading the performance.

For RD with ResNet-18, the accuracy stabilises at 93.65% for $\varepsilon \geq 6.87$, closely aligned with its non-private performance of 93.65%. Similarly, for the STFT with Inception-v3, the accuracy plateaus at 98.41% for $\varepsilon \geq 7.0$, matching the non-privacy baseline model. This suggests that choosing ε between 4.0 and 5.0 provides an optimal balance between privacy and utility, ensuring meaningful privacy protection while preserving the classification performance. For edge-optimised DL model deployments, setting $\varepsilon = 4.79$ achieves 91.27% accuracy for RD and 96.03% for STFT, demonstrating that our energy-efficient R-HARS successfully integrates



(a) ResNet-18 with RD map



(b) Inception-v3 with STFT map

Figure 5.8: LDP curves illustrating the relationship between privacy budget (ε), perturbation probability (β), and classification accuracy for INT8 quantised models.

privacy protection without significantly compromising performance, as shown in Fig. 5.8

5.5 Summary

This Chapter presents an edge-optimised, energy efficient, and privacy-preserved radar-based macro-activity recognition system that addresses the key challenges **C2** and **C3**, by deploying DL models on resource-constrained edge devices using two different radar domain representation. The training efficiency analysis revealed that ResNet-18 with RD maps consumed the least

energy at 0.007 kWh, while the same model with STFT required 0.006 kWh, both operating at 100.74 and 94.21 watts of GPU power, respectively. Among the evaluated models, Xception with RD maps achieved the highest RD accuracy of 95.24%, whereas Inception-v3 with STFT achieved the highest recognition accuracy of 98.41%. However, model compression via INT8 quantisation resulted in an approximately 90% reduction in model size, a 50% reduction in inference time compared to FP16, and 30–50% lower energy consumption, making INT8 quantisation the most efficient approach for edge deployment. For real-time HAR on edge devices, ResNet-18 INT8 with RD maps has emerged as the most practical solution in which the inference speed is critical. It achieved fast inference times of 221 ms on RPi and 201 ms on JNano, maintaining good accuracy of 93.65%, low energy consumption of 0.13 mWh on RPi and 0.14 mWh on JNano, and a minimal memory footprint of 11 MB.

Conversely, Inception-v3 INT8 with STFT provided the best accuracy-oriented solution, delivering exceptional accuracy of 98.41% with a reasonable inference time of 329 ms on RPi, and 309 ms on JNano, low EPR values of 0.002 mWh on RPi and 0.003 mWh on JNano, and an acceptable trade-off between energy consumption and accuracy. The end-to-end system evaluation demonstrated that the RD-based ResNet-18 model could process over 35,000 activity classifications on a single 3,000 mAh (15,000 mWh) battery charge, making it the most practical solution for long-term real-time HAR deployment. Additionally, the LDP evaluation revealed that choosing ε between 4.0 and 5.0 preserves the accuracy above 91% while maintaining privacy guarantees. This ensures that our energy-efficient R-HARS integrates privacy protection without significantly compromising classification performance. The successful implementation of macro-activity recognition confirms the effectiveness of radar signal preprocessing and model deployment strategies at the edge. However, monitoring subtle micro-movements such as respiration and heartbeat introduces new challenges in radar signal sensitivity and resolution, which are addressed in the next chapter.

Chapter 6

Conclusion and Future Work

This final chapter concludes the thesis by reflecting on the core contributions and their impact on addressing the three central challenges of non-invasive AI-driven HAR, as outlined in Section 1.3. These challenges include optimising radar signal preprocessing and domain representations for macro-activity recognition, enabling energy-efficient and privacy-preserving edge deployment, and achieving accurate, contactless monitoring of micro-level physiological signals. This thesis advances radar-based sensing by leveraging the radar's inherent ability to capture motion across multiple scales through novel signal processing pipelines, DL frameworks, and edge-optimised system architectures. This chapter synthesises key findings, discusses their broader implications for healthcare and intelligent environments, and outlines promising directions for future research.

6.1 Summary of Contributions

In Chapter 3, a R-HAR framework is developed to address the challenges **C1** and **C2**, with a specific focus on macro-activity recognition. This study explored the preprocessing of radar signals into four domain representations: RT, RD, STFT, and SPWVD, to enhance activity classification. Both a baseline CNN and several DL models with a focus on TL approach, were employed using pre-trained weights to improve feature extraction. The baseline CNN served as a practical starting point, demonstrating that even with simpler architectures, radar-based signal processing can yield meaningful results. However, for real-world deployment, more accurate and computationally efficient solutions are essential to satisfy the latency and energy constraints. A subject-wise data splitting strategy ensured the evaluation of unseen individuals, directly tackling generalisation under non-IID conditions (C2). The TL models significantly outperformed the baseline, with DenseNet-201 achieving 96.03% accuracy using STFT. Although SPWVD yielded rich Spectro-temporal features, its 56-second preprocessing time per 5-second event limited its suitability for real-time applications, unlike RT, RD, and STFT, which required approximately 0.5 seconds. The fall detection performance further validated the reliability of

the TL approach, with notable reductions in false positives compared to the baseline. Computationally efficient model-domain pairs such as EfficientNet-B0 with RT and VGG-19 with RD achieved 93.65% accuracy, offering a strong trade-off between performance and inference cost. A decision framework was proposed to guide the selection of radar domains and models according to the application requirements. To assess cross-frequency generalisability, two optimal pairs such as RD with VGG-19 and STFT with DenseNet-201, were validated on datasets collected at 24 and 77 GHz. Using a LOSO-CV strategy, they achieved 89.70% and 96.42% recognition accuracy, respectively, confirming their robustness across different radar systems and deployment environments.

In Chapter 4 addresses challenge **C1** by extending the proposed HARS to micro-activity monitoring, focusing on non-invasive estimation of vital signs such as HR and RR. This study demonstrated that radar technology, in addition to recognising macro-activities, can reliably capture subtle physiological signals without physical contact. Two radar platforms were explored: UWB radar for the initial validation and mmWave FMCW radar for the final deployment. Dedicated signal-processing pipelines were developed for each to estimate vital signs under different operational conditions. Initial experiments using UWB radar evaluated the effect of varying subject distances (55, 60, and 65 cm) on RR estimation accuracy. Results showed that 60 cm yielded the most accurate readings, with an absolute error of just 1.2 Brpm compared to a medical-grade reference sensor. At 55 cm and 65 cm, the errors increased to 3.6 and 2.4 Brpm, respectively. In addition to continuous tracking, the system was able to distinguish between three distinct breathing patterns: slow breathing (approximately 4.8 Brpm), normal breathing (12 Brpm), and elevated breathing (31.2 Brpm). These classifications were validated using a respiration belt, confirming the effectiveness of the system in identifying varied respiratory conditions.

Building on these findings, mmWave radar experiments were conducted with ten participants, and the system performance was evaluated at different installation heights, including 1 m, chest height, and 2 m. Two signal-processing approaches were compared: FFT and peak counting. The peak counting method consistently outperformed FFT, delivering lower estimation errors across both HR and RR. It achieved an average MAE of 4.2 Bpm for HR and 4.9 Brpm for RR, while FFT recorded higher errors. The analysis also showed that placing the radar at approximately 2 m, aligning with the subject's standing height, produced the most reliable results. Subject height had a minimal effect on accuracy, indicating robust performance across diverse users. These results confirm the potential use of mmWave radar for real-time unobtrusive vital sign monitoring. The system offers a practical and non-invasive alternative to wearable sensors, making it suitable for applications in healthcare monitoring, especially in elderly care or scenarios requiring continuous observation. Together with the macro-activity recognition covered in earlier chapters, this study demonstrates that radar technology can serve as a unified, multi-scale sensing platform for non-intrusive HAR and health monitoring.

Finally, Chapter 5 presents a collaborative edge-AI framework for R-HAR that operates

independently without cloud infrastructure dependencies. This study specifically address challenges **C2** and **C3** by focusing on macro-activity recognition in real-world environments. Building upon the radar domain insights from Chapter 3, this study evaluated RD and STFT inputs across different DL models. The experimental results demonstrated strong performance: RD input paired with the Xception model achieved 95.24% recognition accuracy, whereas STFT input combined with Inception-v3 reached 98.41% accuracy. To assess the generalisability of the proposed framework, the study evaluated the optimal model-domain pairs on two additional publicly available datasets collected using FMCW radar at 24 GHz and 77 GHz frequencies. For the 24 GHz dataset, the RD-Xception combination achieved 88.22% accuracy, while the performance on the 77 GHz dataset reached 84.17%. The STFT with Inception-v3 pairing demonstrated superior generalisation, achieving 94.39% accuracy on the 24 GHz dataset and 92.75% on the 77 GHz dataset. These evaluations employed the LOSO-CV strategy to address the inherent challenges of non-IID data and heterogeneous characteristics typical of radar datasets.

To support efficient deployment on edge devices, model compression techniques were applied using PTQ, which compresses the full 32-bit models to 16-bit and 8-bit integer formats. The 16-bit quantised models achieved a substantial reduction in model size while maintaining an accuracy comparable to the full 32-bit precision. The 8-bit quantised models further reduced the memory footprint and improved the inference latency by approximately 50%, with only minimal loss in classification accuracy. On-device evaluations showed that ResNet-18 with RD achieved 93.65% accuracy with an EPR value of 0.007 mWh, while Inception-v3 with STFT reached 98.41% accuracy and the lowest EPR of 0.002 mWh. These models were successfully deployed on Raspberry Pi and Jetson Nano, confirming the framework's capability to operate in real-time within resource-constrained, battery-powered environments. The proposed framework was further validated through end-to-end system integration, including radar signal acquisition, domain transformation, compressed inference, and output generation. When deployed on a Raspberry Pi, the system completed a full inference cycle for a 5-second activity window in just 1.32 seconds with a total energy cost of 0.42 mWh, confirming real-time responsiveness. These findings establish the practicality of the system for continuous HAR in smart healthcare and AAL environments. Given the sensitive nature of human activity data, privacy preservation is a critical aspect of system design. The study incorporated LDP at the inference stage, ensuring that data remained protected on-device before any transmission. A privacy-utility trade-off analysis revealed that even with a moderate noise level ($\varepsilon = 4.79$), the system retained over 91% classification accuracy for RD and 96% for STFT. This balance between accuracy and privacy reinforces the system's readiness for edge deployment in healthcare and surveillance contexts. Moreover, the inherently non-visual and non-intrusive nature of radar sensing enhances user privacy compared to camera-based systems, making radar an ideal modality for privacy-aware AI-driven HAR applications.

6.2 Limitations and Future Research Direction

This thesis addressed the critical challenges for non-invasive AI-driven HAR, offering innovative solutions for the multi-scale monitoring of radar in capturing both macro-movements for activity recognition and micro-movements for vital sign estimation. This research advanced radar signal processing techniques and optimised edge resource efficiency while implementing cross-subject validation to assess model generalisability across individuals with diverse physiological and behavioural characteristics. The practical implementations throughout this work served as proof of concept for combining edge technologies, including the analysis of platforms capable of storing data locally on edge devices without requiring constant internet connectivity. Despite these significant contributions, several areas remain unexplored, creating opportunities for future research to build upon current findings. The following section outline the limitations of the present study and potential directions for enhancing radar-based sensing frameworks, exploring emerging privacy-preserving techniques, and addressing new challenges. These directions aim to refine the scalability, robustness, and applicability of radar-based sensing systems in diverse real-world scenarios.

R-HAR studies primarily utilise publicly available datasets collected in controlled indoor environments, which may not fully capture the complexity and variability of real-world scenarios. Additionally, the relatively small number of participants and limited data samples can significantly affect the generalisation capability of the developed ML algorithms, whether classical or DL methods are employed. This limitation restricts the robustness and scalability of the proposed system when deployed in diverse environments. Future work should focus on creating more comprehensive datasets collected with mmWave band frequency radars in outdoor environments using robust data collection protocols. Expanding the dataset to include a wider range of participants with diverse physical characteristics, activity patterns, and environmental conditions would enhance the generalisation capabilities of the models. Moreover, investigating the performance of the model across different SNRs and aspect angles would provide valuable insights into its robustness in real-world applications where noise levels fluctuate and target orientations vary.

Another critical limitation is related to the radar signal processing and representation approach adopted in this study. The current methodology transforms raw radar data into 2D images for DL model training, which may not fully capture the rich spatio-temporal information present in radar signals. This dimensional reduction could potentially limit the extraction of the complex features necessary for distinguishing subtle differences between similar activities or vital sign patterns. Future research should investigate 3D and 4D radar representations that incorporate additional dimensions, such as time, Doppler, and angular information. Furthermore, exploring fusion techniques for different radar representations could provide complementary information and enhance the overall performance of DL models. Advanced architectures, such as multi-modal transformer models, can be incorporated to effectively capture complex relation-

ships across diverse radar representations. These models, with their self-attention mechanisms, have shown remarkable success in capturing dependencies in sequential data, and can significantly improve feature extraction and classification performance in R-HAR frameworks.

Another limitation of the current framework is the absence of explainability mechanisms that are essential for interpreting and understanding the decisions made by models. In many real-world applications, such as healthcare, energy management, and wireless communications, providing insights into why a model makes a specific prediction or decision is crucial for building trust and facilitating adoption. Explainability is particularly challenging in AI-driven radar systems because of the complex nature of radar signal processing and the "black box" characteristics of DL models. Additionally, the study did not address the model's response to unknown classes, which is an essential aspect of ML models deployed in dynamic environments. The ability to accurately identify or appropriately handle signals that do not belong to any of the training categories is crucial for practical applications. Future work should focus on the development of lightweight and privacy-preserving explainability techniques tailored for radar-based sensing applications. These techniques include attention visualisation, feature importance analysis, and model-agnostic interpretation methods, which provide insights into the decision-making process without compromising privacy or computational efficiency. Moreover, investigating open-set recognition approaches and anomaly detection mechanisms would enhance the ability of the models to handle unknown or unseen activities in real-world deployments.

Vital sign estimation studies have several significant limitations. First, the work concentrated primarily on scenarios in which subjects were nearly stationary, which does not reflect real-world conditions in which individuals make random body movements that can adversely affect radar echo signals. Current algorithms struggle to extract physiological signals when subjects are in motion, as non-biological motions dominate radar reflections. Additionally, these studies were not implemented in real-time environments with diverse participant populations, limiting their immediate applicability in clinical settings. The research also did not adequately explore the distinction between human vital signs and periodic motions generated by background sources, such as fans or furniture, which can influence RR estimation. To achieve practical utility for vital sign monitoring, future research should focus on comprehensive testing in real-world healthcare settings with diverse participant cohorts, developing algorithms capable of extracting vital signs during movement, implementing multi-scale radar systems to improve the accuracy of detecting multiple individuals, and evaluating performance in practical circumstances, such as elderly care facilities, where the technology could provide the most benefit.

Another critical limitation is the deployment and edge computing aspects. The current approach uses edge devices only for inference, whereas training is conducted on server-side GPUs. This strategy increases bandwidth usage and limits system efficiency for data transfer and model deployment, thereby limiting the efficiency of the system in resource-constrained environments. In addition, transferring models to edge devices makes them susceptible to attacks,

potentially compromising privacy and security. Future research should explore the implementation of federated learning environments where local model training is performed on edge devices and only model updates are shared with the central server. This approach significantly reduces the communication overhead, enhances privacy, and enables real-time learning. The adoption of lightweight and energy-efficient models, such as spiking neural networks, could further optimise edge computing performance, particularly when deployed on neuromorphic hardware designed for low-power, event-driven processing.

Future research should focus on expanding R-HARS by integrating advanced ML models, developing multi-modal sensor fusion frameworks, and improving real-time processing capabilities. One promising direction is the incorporation of multi-modal transformer models, which have shown exceptional success in capturing complex relationships across diverse data modalities such as text, images, and time-series signals. Applying transformers in HAR can enhance feature extraction and predictive performance, particularly in decentralised learning environments. However, the high computational cost of transformers poses challenges for edge deployment, necessitating research into optimisation techniques, such as model pruning, quantisation, and efficient attention mechanisms. Another avenue for exploration is the enhancement of radar-based security applications, including intruder detection, fall detection, and human identification by integrating biometric recognition with activity analysis. Additionally, future studies should investigate novel antenna designs and sub-terahertz frequencies to extend the sensing range of RF systems and improve their penetration through obstacles. Exploring bio-inspired computing models, such as neuromorphic processing, can further optimise the power efficiency and real-time adaptability. For privacy and security, emerging techniques such as homomorphic encryption and quantum key distribution offer promising alternatives to LDP, potentially providing stronger privacy guarantees without significant performance degradation. By addressing these limitations and pursuing these future directions, radar-based sensing technology can reach its full potential as a transformative tool for human-centric applications in healthcare, smart environments, and security systems.

Acknowledgement for the Use of AI Tools

I acknowledge the use of two language assistance tools, **ChatGPT** by OpenAI and **Grammarly**, in the proofreading of this thesis. These tools were employed solely to enhance the clarity, coherence, and academic tone of the written text. Their contributions were limited to improving language and flow, while the intellectual content and findings of this work remain entirely my own.

Bibliography

[1] "Home - office for national statistics." [Online]. Available: https://www.ons.gov.uk/

[2] L. Bosque-Mercader and L. Siciliani, "The association between bed occupancy rates and hospital quality in the english national health service," *The European Journal of Health Economics*, vol. 24, no. 2, pp. 209–236, 2023.

[3] "Bed blocking in the nhs and why it's a problem." [Online]. Available: https://www.sparktsl.com/blog/bed-blocking-in-the-nhs-and-why-its-a-problem

[4] "Statistics » bed availability and occupancy – kh03." [Online]. Available: https://www.england.nhs.uk/statistics/statistical-work-areas/bed-availability-and-occupancy/bed-availability-and-occupancy-kh03/

[5] H. Storf, T. Kleinberger, M. Becker, M. Schmitt, F. Bomarius, and S. Prueckner, "An event-driven approach to activity recognition in ambient assisted living," in *Ambient Intelligence: European Conference, AmI 2009, Salzburg, Austria, November 18-21, 2009. Proceedings*. Springer, 2009, pp. 123–132.

[6] "Government to invest £30 million in innovative technology for nhs - gov.uk." [Online]. Available: https://www.gov.uk/government/news/government-to-invest-30-million-in-innovative-technology-for-nhs

[7] D. Mohan, D. Z. Al-Hamid, P. H. J. Chong, K. L. K. Sudheera, J. Gutierrez, H. C. Chan, and H. Li, "Artificial intelligence and iot in elderly fall prevention: A review," *IEEE Sensors Journal*, vol. 24, no. 4, pp. 4181–4198, 2024.

[8] A. Dey, S. Rajan, G. Xiao, and J. Lu, "Radar-based human activity recognition using multi-domain multi-level fused patch-based learning," *IEEE Transactions on Instrumentation and Measurement*, 2024.

[9] "Tunstall | innovative telehealth & telecare services." [Online]. Available: https://www.tunstall.co.uk/

[10] "Canary care." [Online]. Available: https://www.canarycare.co.uk/

[11] "Howz | technology enabled remote care." [Online]. Available: https://www.howz.com/

[12] S. Czaja, S. Beach, N. Charness, and R. Schulz, "Older adults and the adoption of healthcare technology: Opportunities and challenges," *Technologies for active aging*, pp. 27–46, 2013.

[13] Y. Yu, B. Jain, G. Anand, M. Heidarian, A. Lowe, and A. Kalra, "Technologies for non-invasive physiological sensing: Status, challenges, and future horizons," *Biosensors and Bioelectronics: X*, vol. 16, p. 100420, 2024.

[14] M. Al-Faris, J. Chiverton, D. Ndzi, and A. I. Ahmed, "A review on computer vision-based methods for human action recognition," *Journal of imaging*, vol. 6, no. 6, p. 46, 2020.

[15] I. Ullmann, R. G. Guendel, N. C. Kruse, F. Fioranelli, and A. Yarovoy, "A survey on radar-based continuous human activity recognition," *IEEE Journal of Microwaves*, vol. 3, no. 3, pp. 938–950, 2023.

[16] C. Yang, X. Wang, and S. Mao, "Tarf: Technology-agnostic rf sensing for human activity recognition," *IEEE journal of biomedical and health informatics*, vol. 27, no. 2, pp. 636–647, 2022.

[17] S. A. Shah and F. Fioranelli, "Rf sensing technologies for assisted daily living in healthcare: A comprehensive review," *IEEE Aerospace and Electronic Systems Magazine*, vol. 34, no. 11, pp. 26–44, 2019.

[18] B. Suthar and B. Gadhia, "Human activity recognition using deep learning: a survey," in *Data Science and Intelligent Applications: Proceedings of ICDSIA 2020*. Springer, 2021, pp. 217–223.

[19] Y. Jia, Y. Guo, G. Wang, R. Song, G. Cui, and X. Zhong, "Multi-frequency and multi-domain human activity recognition based on sfcw radar using deep learning," *Neurocomputing*, vol. 444, pp. 274–287, 2021.

[20] A. Dey, S. Rajan, G. Xiao, and J. Lu, "Measurement methodology: Radar-based human activity recognition: Is it ready for aging in place?" *IEEE Instrumentation & Measurement Magazine*, vol. 26, no. 7, pp. 12–19, 2023.

[21] S. R. Shakya, C. Zhang, and Z. Zhou, "Comparative study of machine learning and deep learning architecture for human activity recognition using accelerometer data," *Int. J. Mach. Learn. Comput*, vol. 8, no. 6, pp. 577–582, 2018.

[22] X. Li, Y. He, and X. Jing, "A survey of deep learning-based human activity recognition in radar," *Remote sensing*, vol. 11, no. 9, p. 1068, 2019.

[23] J. Zhang, J. Tao, and Z. Shi, "Doppler-radar based hand gesture recognition system using convolutional neural networks," in *Communications, Signal Processing, and Systems: Proceedings of the 2017 International Conference on Communications, Signal Processing, and Systems*. Springer, 2019, pp. 1096–1113.

[24] A. Shastri, N. Valecha, E. Bashirov, H. Tataria, M. Lentmaier, F. Tufvesson, M. Rossi, and P. Casari, "A review of millimeter wave device-based localization and device-free sensing technologies and applications," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 3, pp. 1708–1749, 2022.

[25] M. Alizadeh, G. Shaker, J. C. M. De Almeida, P. P. Morita, and S. Safavi-Naeini, "Remote monitoring of human vital signs using mm-wave fmcw radar," *IEEE Access*, vol. 7, pp. 54 958–54 968, 2019.

[26] T. Sukianto, M. Wagner, S. Seifi, C. Carbonelli, and M. Huemer, "An uncertainty aware semi-supervised federated learning framework for radar-based hand gesture recognition," in *2024 21st European Radar Conference (EuRAD)*. IEEE, 2024, pp. 168–171.

[27] R. El Hail, P. Mehrjouseresht, D. M.-P. Schreurs, and P. Karsmakers, "Radar-based human activity recognition: A study on cross-environment robustness," *Electronics*, vol. 14, no. 5, p. 875, 2025.

[28] Z. Li, J. Le Kernec, Q. Abbasi, F. Fioranelli, S. Yang, and O. Romain, "Radar-based human activity recognition with adaptive thresholding towards resource constrained platforms," *Scientific Reports*, vol. 13, no. 1, p. 3473, 2023.

[29] S. Huan, L. Wu, M. Zhang, Z. Wang, and C. Yang, "Radar human activity recognition with an attention-based deep learning network," *Sensors*, vol. 23, no. 6, p. 3185, 2023.

[30] S. Z. Gürbüz, B. Erol, B. Çağlıyan, and B. Tekeli, "Operational assessment and adaptive selection of micro-doppler features," *IET Radar, Sonar & Navigation*, vol. 9, no. 9, pp. 1196–1204, 2015.

[31] Y. Chen, J. Li, E. Blasch, and Q. Qu, "Future outdoor safety monitoring: Integrating human activity recognition with the internet of physical–virtual things," *Applied Sciences*, vol. 15, no. 7, p. 3434, 2025.

[32] N. Rashid, B. U. Demirel, and M. A. Al Faruque, "Ahar: Adaptive cnn for energy-efficient human activity recognition in low-power edge devices," *IEEE Internet of Things Journal*, vol. 9, no. 15, pp. 13 041–13 051, 2022.

[33] J. Chen and X. Ran, "Deep learning with edge computing: A review," *Proceedings of the IEEE*, vol. 107, no. 8, pp. 1655–1674, 2019.

[34] F. Savvidou, S. A. Tegos, P. D. Diamantoulakis, and G. K. Karagiannidis, "Passive radar sensing for human activity recognition: A survey," *IEEE Open Journal of Engineering in Medicine and Biology*, 2024.

[35] M. Czerkawski, C. Ilioudis, C. Clemente, C. Michie, I. Andonovic, and C. Tachtatzis, "Interference motion removal for doppler radar vital sign detection using variational encoder-decoder neural network," in *2021 IEEE Radar Conference (RadarConf21)*. IEEE, 2021, pp. 1–6.

[36] X. Ye, K. Sakurai, N.-K. C. Nair, and K. I.-K. Wang, "Machine learning techniques for sensor-based human activity recognition with data heterogeneity—a review," *Sensors (Basel, Switzerland)*, vol. 24, no. 24, p. 7975, 2024.

[37] P. Rashidi and A. Mihailidis, "A survey on ambient-assisted living tools for older adults," *IEEE journal of biomedical and health informatics*, vol. 17, no. 3, pp. 579–590, 2012.

[38] M. Muaaz, S. Waqar, and M. Pätzold, "Radar-based passive step counter and its comparison with a wrist-worn physical activity tracker," in *International Conference on Intelligent Technologies and Applications*. Springer, 2021, pp. 259–272.

[39] C. ALBARRAN MORILLO, D. Jawla, J. D. Kelleher, M. Demichela *et al.*, "Safety-critical systems in the automotive sector: pros and cons in the current state-of-the-art of human performance assessment," in *Proceedings of the 32nd European Safety and Reliability Conference (ESREL 2022)*. RESEARCH PUBLISHING, SINGAPORE, 2022, pp. 3226–3233.

[40] H. Zhou, C. Tawk, and G. Alici, "A multipurpose human–machine interface via 3d-printed pressure-based force myography," *IEEE Transactions on Industrial Informatics*, 2024.

[41] S. Nižetić, P. Šolić, D. L.-d.-I. Gonzalez-De, L. Patrono *et al.*, "Internet of things (iot): Opportunities, issues and challenges towards a smart and sustainable future," *Journal of cleaner production*, vol. 274, p. 122877, 2020.

[42] A. Rahaman, M. M. Islam, M. R. Islam, M. S. Sadi, S. Nooruddin *et al.*, "Developing iot based smart health monitoring systems: A review." *Rev. d'Intelligence Artif.*, vol. 33, no. 6, pp. 435–440, 2019.

[43] N. Gupta, S. K. Gupta, R. K. Pathak, V. Jain, P. Rashidi, and J. S. Suri, "Human activity recognition in artificial intelligence framework: a narrative review," *Artificial intelligence review*, vol. 55, no. 6, pp. 4755–4808, 2022.

[44] G. Bhola and D. K. Vishwakarma, "A review of vision-based indoor har: state-of-the-art, challenges, and future prospects," *Multimedia Tools and Applications*, vol. 83, no. 1, pp. 1965–2005, 2024.

[45] M. Mohtadifar, M. Cheffena, and A. Pourafzal, "Acoustic-and radio-frequency-based human activity recognition," *Sensors*, vol. 22, no. 9, p. 3125, 2022.

[46] I. M. Pires, N. M. Garcia, E. Zdravevski, and P. Lameski, "Daily motionless activities: A dataset with accelerometer, magnetometer, gyroscope, environment, and gps data," *Scientific Data*, vol. 9, no. 1, p. 105, 2022.

[47] G. Lin, W. Jiang, S. Xu, X. Zhou, X. Guo, Y. Zhu, and X. He, "Human activity recognition using smartphones with wifi signals," *IEEE Transactions on Human-Machine Systems*, vol. 53, no. 1, pp. 142–153, 2022.

[48] G. V. Doviak and T. Sledevič, "Towards human activity recognition in smart environments through thermal imaging," in *2024 IEEE Open Conference of Electrical, Electronic and Information Sciences (eStream)*. IEEE, 2024, pp. 1–4.

[49] Y. Karayaneva, S. Sharifzadeh, Y. Jing, and B. Tan, "Human activity recognition for ai-enabled healthcare using low-resolution infrared sensor data," *Sensors*, vol. 23, no. 1, p. 478, 2023.

[50] M. Koch, T. Pfitzinger, F. Schlenke, F. Kohlmorgen, R. Groll, and H. Wöhrle, "Recognition of human activities based on ambient audio and vibration data," *IEEE Access*, 2024.

[51] O. D. Lara and M. A. Labrador, "A survey on human activity recognition using wearable sensors," *IEEE communications surveys & tutorials*, vol. 15, no. 3, pp. 1192–1209, 2012.

[52] F. S. Abuhoureyah, Y. C. Wong, and A. S. B. M. Isira, "Wifi-based human activity recognition through wall using deep learning," *Engineering Applications of Artificial Intelligence*, vol. 127, p. 107171, 2024.

[53] S. Waqar, M. Muaaz, and M. Pätzold, "Human activity signatures captured under different directions using siso and mimo radar systems," *Applied Sciences*, vol. 12, no. 4, p. 1825, 2022.

[54] J. Liu, H. Liu, Y. Chen, Y. Wang, and C. Wang, "Wireless sensing for human activity: A survey," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 3, pp. 1629–1645, 2019.

[55] S. Qiu, H. Zhao, N. Jiang, Z. Wang, L. Liu, Y. An, H. Zhao, X. Miao, R. Liu, and G. Fortino, "Multi-sensor information fusion based on machine learning for real applications in human activity recognition: State-of-the-art and research challenges," *Information Fusion*, vol. 80, pp. 241–265, 2022.

[56] E. Garcia-Ceja, M. Riegler, T. Nordgreen, P. Jakobsen, K. J. Oedegaard, and J. Tørresen, "Mental health monitoring with multimodal sensing and machine learning: A survey," *Pervasive and Mobile Computing*, vol. 51, pp. 1–26, 2018.

[57] P. Susarla, A. Mukherjee, M. L. Cañellas, C. Á. Casado, X. Wu, O. Silvén, D. B. Jayagopi, M. B. López *et al.*, "Non-contact multimodal indoor human monitoring systems: A survey," *Information Fusion*, vol. 110, p. 102457, 2024.

[58] A. Ray, M. H. Kolekar, R. Balasubramanian, and A. Hafiane, "Transfer learning enhanced vision-based human activity recognition: a decade-long analysis," *International Journal of Information Management Data Insights*, vol. 3, no. 1, p. 100142, 2023.

[59] M. Batool, S. S. Alotaibi, M. H. Alatiyyah, K. Alnowaiser, H. Aljuaid, A. Jalal, and J. Park, "Depth sensors-based action recognition using a modified k-ary entropy classifier," *IEEE Access*, vol. 11, pp. 58 578–58 595, 2023.

[60] Y. Zhou, H. Huang, S. Yuan, H. Zou, L. Xie, and J. Yang, "Metafi++: Wifi-enabled transformer-based human pose estimation for metaverse avatar simulation," *IEEE Internet of Things Journal*, vol. 10, no. 16, pp. 14 128–14 136, 2023.

[61] B. Fu, N. Damer, F. Kirchbuchner, and A. Kuijper, "Sensing technology for human activity recognition: A comprehensive survey," *Ieee Access*, vol. 8, pp. 83 791–83 820, 2020.

[62] M.-T. Vi, D.-N. Tran, V. T. Thuong, N. N. Linh, and D.-T. Tran, "Efficient real-time devices based on accelerometer using machine learning for har on low-performance microcontrollers," *Computers, Materials and Continua*, vol. 81, no. 1, pp. 1729–1756, 2024.

[63] P. Yang, C. Yang, V. Lanfranchi, and F. Ciravegna, "Activity graph based convolutional neural network for human activity recognition using acceleration and gyroscope data," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 10, pp. 6619–6630, 2022.

[64] Q. Xu, X. Wei, R. Bai, S. Li, and Z. Meng, "Integration of deep adaptation transfer learning and online sequential extreme learning machine for cross-person and cross-position activity recognition," *Expert Systems with Applications*, vol. 212, p. 118807, 2023.

[65] S. Waqar, M. Muaaz, and M. Pätzold, "Direction-independent human activity recognition using a distributed mimo radar system and deep learning," *IEEE Sensors Journal*, 2023.

[66] Z. Wu, Z. Cao, X. Yu, J. Zhu, C. Song, and Z. Xu, "A novel real-time multi-person activity recognition system using millimeter-wave siso radar data," in *International Conference on Radar Systems (RADAR 2022)*, vol. 2022. IET, 2022, pp. 696–701.

[67] N. Avazov, R. Hicheri, and M. Pätzold, "A trajectory-driven siso mm-wave channel model for a human activity recognition," in *2021 17th International Conference on Wireless and*

Mobile Computing, Networking and Communications (WiMob). IEEE, 2021, pp. 133–138.

[68] A. Rahdar, M. Chahoushi, and S. A. Ghorashi, "Efficiently improving the wi-fi-based human activity recognition, using auditory features, autoencoders, and fine-tuning," *Computers in Biology and Medicine*, vol. 172, p. 108232, 2024.

[69] L. M. Dang, K. Min, H. Wang, M. J. Piran, C. H. Lee, and H. Moon, "Sensor-based and vision-based human activity recognition: A comprehensive survey," *Pattern Recognition*, vol. 108, p. 107561, 2020.

[70] R. Raj and A. Kos, "Different techniques for human activity recognition," in *2022 29th International Conference on Mixed Design of Integrated Circuits and System (MIXDES)*. IEEE, 2022, pp. 171–176.

[71] J. Phipps, B. Passage, K. Sel, J. Martinez, M. Saadat, T. Koker, N. Damaso, S. Davis, J. Palmer, K. Claypool *et al.*, "Early adverse physiological event detection using commercial wearables: challenges and opportunities," *NPJ Digital Medicine*, vol. 7, no. 1, p. 136, 2024.

[72] A. D. Singh, S. S. Sandha, L. Garcia, and M. Srivastava, "Radhar: Human activity recognition from point clouds generated through a millimeter-wave radar," in *Proceedings of the 3rd ACM Workshop on Millimeter-wave Networks and Sensing Systems*, 2019, pp. 51–56.

[73] S. Mekruksavanich and A. Jitpattanakul, "Smartwatch-based human activity recognition using hybrid lstm network," in *2020 IEEE SENSORS*. IEEE, 2020, pp. 1–4.

[74] Y. Zhu, J. Yu, F. Hu, Z. Li, and Z. Ling, "Human activity recognition via smart-belt in wireless body area networks," *International Journal of Distributed Sensor Networks*, vol. 15, no. 5, p. 1550147719849357, 2019.

[75] "Mercury Thermometers | US EPA." [Online]. Available: https://www.epa.gov/mercury/mercury-thermometers

[76] L. S. Mokatren, R. Ansari, A. E. Cetin, A. D. Leow, O. A. Ajilore, H. Klumpp, and F. T. Y. Vural, "Eeg classification by factoring in sensor spatial configuration," *IEEE Access*, vol. 9, pp. 19 053–19 065, 2021.

[77] C. K. H. Ne, J. Muzaffar, A. Amlani, and M. Bance, "Hearables, in-ear sensing devices for bio-signal acquisition: a narrative review," *Expert Review of Medical Devices*, vol. 18, no. sup1, pp. 95–128, 2021.

[78] M. G. Moghaddam, A. A. N. Shirehjini, and S. Shirmohammadi, "A wifi-based method for recognizing fine-grained multiple-subject human activities," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–13, 2023.

[79] R. Ou, Y. Chen, and Y. Deng, "Wiwalk: Gait-based dual-user identification using wifi device," *IEEE Internet of Things Journal*, vol. 10, no. 6, pp. 5321–5334, 2022.

[80] S. Ahmed and S. H. Cho, "Machine learning for healthcare radars: Recent progresses in human vital sign measurement and activity recognition," *IEEE Communications Surveys & Tutorials*, 2023.

[81] Y. J. Choo, G. W. Lee, J. S. Moon, and M. C. Chang, "Application of non-contact sensors for health monitoring in hospitals: a narrative review," *Frontiers in Medicine*, vol. 11, p. 1421901, 2024.

[82] S. Dong, L. Wen, Y. Ye, Z. Zhang, Y. Wang, Z. Liu, Q. Cao, Y. Xu, C. Li, and C. Gu, "A review on recent advancements of biomedical radar for clinical applications," *IEEE Open Journal of Engineering in Medicine and Biology*, 2024.

[83] B. R. Mahafza, *Introduction to radar analysis*. Chapman and Hall/CRC, 2017.

[84] H. Rahman, *Fundamental principles of radar*. CRC Press, 2019.

[85] B. Zhou, Y. Lin, J. Le Kernec, S. Yang, F. Fioranelli, O. Romain, and Z. Zhao, "Simulation framework for activity recognition and benchmarking in different radar geometries," *IET Radar, Sonar & Navigation*, vol. 15, no. 4, pp. 390–401, 2021.

[86] Y. Richter, J. Gerasimov, N. Balal, and Y. Pinhasi, "Tracking of evasive objects using bistatic doppler radar operating in the millimeter wave regime," *Remote Sensing*, vol. 14, no. 4, p. 867, 2022.

[87] M. Jankiraman, *FMCW radar design*. Artech House, 2018.

[88] F. Gini, "Grand challenges in radar signal processing," *Frontiers in Signal Processing*, vol. 1, p. 664232, 2021.

[89] J. D. Taylor, "Ultra-wideband radar overview," in *Introduction to ultra-wideband radar systems*. CRC Press, 2020, pp. 1–10.

[90] H. Li, A. Shrestha, H. Heidari, J. Le Kernec, and F. Fioranelli, "A multisensory approach for remote health monitoring of older people," *IEEE Journal of Electromagnetics, RF and Microwaves in Medicine and Biology*, vol. 2, no. 2, pp. 102–108, 2018.

[91] V. C. Chen, *The micro-Doppler effect in radar*. Artech house, 2019.

[92] T. Instruments, "The fundamentals of millimeter wave radar sensors," 2020.

[93] Y. Shao, Y. Dai, L. Yuan, and W. Chen, "Deep learning methods for personnel recognition based on micro-doppler features," in *Proceedings of the 9th International Conference on Signal Processing Systems*, 2017, pp. 94–98.

[94] M. Zhang, B. Li, H. Liu, and C. Zhao, "Federated learning for radar gesture recognition based on spike timing-dependent plasticity," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 60, no. 2, pp. 2379–2393, 2024.

[95] J. Wang, "Cfar-based interference mitigation for fmcw automotive radar systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 8, pp. 12 229–12 238, 2021.

[96] Q. Zheng, L. Yang, Y. Xie, J. Li, T. Hu, J. Zhu, C. Song, and Z. Xu, "A target detection scheme with decreased complexity and enhanced performance for range-doppler fmcw radar," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–13, 2020.

[97] Y. Lin, H. Li, and D. Faccio, "Human multi-activities classification using mmwave radar: Feature fusion in time-domain and pcanet," *Sensors*, vol. 24, no. 16, p. 5450, 2024.

[98] S. Z. Gurbuz, M. M. Rahman, Z. Bassiri, and D. Martelli, "Overview of radar-based gait parameter estimation techniques for fall risk assessment," *IEEE Open Journal of Engineering in Medicine and Biology*, 2024.

[99] Y. Luo, Y.-J. Chen, Y.-Z. Zhu, W.-Y. Li, and Q. Zhang, "Doppler effect and micro-doppler effect of vortex-electromagnetic-wave-based radar," *IET Radar, Sonar & Navigation*, vol. 14, no. 1, pp. 2–9, 2020.

[100] D. Tahmoush, "Review of micro-doppler signatures," *IET Radar, Sonar & Navigation*, vol. 9, no. 9, pp. 1140–1146, 2015.

[101] L. Tang, Y. Jia, Y. Qian, S. Yi, and P. Yuan, "Human activity recognition based on mixed cnn with radar multi-spectrogram," *IEEE Sensors Journal*, vol. 21, no. 22, pp. 25 950–25 962, 2021.

[102] J. Y. Chen and B. Z. Li, "The short-time wigner–ville distribution," *Signal Processing*, vol. 219, p. 109402, 2024.

[103] K. Chen, D. Zhang, L. Yao, B. Guo, Z. Yu, and Y. Liu, "Deep learning for sensor-based human activity recognition: Overview, challenges, and opportunities," *ACM Computing Surveys (CSUR)*, vol. 54, no. 4, pp. 1–40, 2021.

[104] Z. Li, Y. Liu, B. Liu, J. Le Kernec, and S. Yang, "A holistic human activity recognition optimisation using ai techniques," *IET Radar, Sonar & Navigation*, vol. 18, no. 2, pp. 256–265, 2024.

[105] E. Alpaydin, *Introduction to machine learning*. MIT press, 2020.

[106] S. B. u. d. Tahir, A. B. Dogar, R. Fatima, A. Yasin, M. Shafiq, J. A. Khan, M. Assam, A. Mohamed, and E.-A. Attia, "Stochastic recognition of human physical activities via augmented feature descriptors and random forest model," *Sensors*, vol. 22, no. 17, p. 6632, 2022.

[107] A. Y. Shdefat, N. Mostafa, Z. Al-Arnaout, Y. Kotb, and S. Alabed, "Optimizing har systems: Comparative analysis of enhanced svm and k-nn classifiers," *International Journal of Computational Intelligence Systems*, vol. 17, no. 1, p. 150, 2024.

[108] L. Alzubaidi, J. Zhang, A. J. Humaidi, A. Al-Dujaili, Y. Duan, O. Al-Shamma, J. Santamaría, M. A. Fadhel, M. Al-Amidie, and L. Farhan, "Review of deep learning: concepts, cnn architectures, challenges, applications, future directions," *Journal of big Data*, vol. 8, pp. 1–74, 2021.

[109] A. Shrestha and A. Mahmood, "Review of deep learning algorithms and architectures," *IEEE access*, vol. 7, pp. 53 040–53 065, 2019.

[110] E. Buber and D. Banu, "Performance analysis and cpu vs gpu comparison for deep learning," in *2018 6th International Conference on Control Engineering & Information Technology (CEIT)*. IEEE, 2018, pp. 1–6.

[111] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[112] W. Yang, X. Zhang, Y. Tian, W. Wang, J.-H. Xue, and Q. Liao, "Deep learning for single image super-resolution: A brief review," *IEEE Transactions on Multimedia*, vol. 21, no. 12, pp. 3106–3121, 2019.

[113] J. Tang, S. Li, and P. Liu, "A review of lane detection methods based on deep learning," *Pattern Recognition*, vol. 111, p. 107623, 2021.

[114] M. Shafiq and Z. Gu, "Deep residual learning for image recognition: A survey," *Applied Sciences*, vol. 12, no. 18, p. 8972, 2022.

[115] A. Bashar *et al.*, "Survey on evolving deep learning neural network architectures," *Journal of Artificial Intelligence*, vol. 1, no. 02, pp. 73–82, 2019.

[116] O. Sharma, "Deep challenges associated with deep learning," in *2019 international conference on machine learning, big data, cloud and parallel computing (COMITCon)*. IEEE, 2019, pp. 72–75.

[117] W. Ding, X. Guo, and G. Wang, "Radar-based human activity recognition using hybrid neural network model with multidomain fusion," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, no. 5, pp. 2889–2898, 2021.

[118] I. Khan, A. Guerrieri, E. Serra, and G. Spezzano, "A hybrid deep learning model for uwb radar-based human activity recognition," *Internet of Things*, vol. 29, p. 101458, 2025.

[119] M. Agarwal, L. Saba, S. K. Gupta, A. Carriero, Z. Falaschi, A. Paschè, P. Danna, A. El-Baz, S. Naidu, and J. S. Suri, "A novel block imaging technique using nine artificial intelligence models for covid-19 disease classification, characterization and severity measurement in lung computed tomography scans on an italian cohort," *Journal of Medical Systems*, vol. 45, no. 3, p. 28, 2021.

[120] X. Li, Y. He, F. Fioranelli, and X. Jing, "Semisupervised human activity recognition with radar micro-doppler signatures," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–12, 2022.

[121] M. A. Khatun, M. A. Yousuf, S. Ahmed, M. Z. Uddin, S. A. Alyami, S. Al-Ashhab, H. F. Akhdar, A. Khan, A. Azad, and M. A. Moni, "Deep cnn-lstm with self-attention model for human activity recognition using wearable sensor," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 10, pp. 1–16, 2022.

[122] G. Klarenbeek, R. Harmanny, and L. Cifola, "Multi-target human gait classification using lstm recurrent neural networks applied to micro-doppler," in *2017 European Radar Conference (EURAD)*. IEEE, 2017, pp. 167–170.

[123] S. Yao, Y. Zhao, H. Shao, D. Liu, S. Liu, Y. Hao, A. Piao, S. Hu, S. Lu, and T. F. Abdelzaher, "Sadeepsense: Self-attention deep learning framework for heterogeneous on-device sensors in internet of things applications," in *IEEE INFOCOm 2019-IEEE conference on computer communications*. IEEE, 2019, pp. 1243–1251.

[124] S. P. Singh, M. K. Sharma, A. Lay-Ekuakille, D. Gangwar, and S. Gupta, "Deep convlstm with self-attention for human activity decoding using wearable sensors," *IEEE Sensors Journal*, vol. 21, no. 6, pp. 8575–8582, 2020.

[125] M. S. Seyfioğlu and S. Z. Gürbüz, "Deep neural network initialization methods for micro-doppler classification with low training sample support," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 12, pp. 2462–2466, 2017.

[126] Y. Kim and T. Moon, "Human detection and activity classification based on micro-doppler signatures using deep convolutional neural networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 1, no. 13, pp. 8–12, 2016.

[127] J. Zhu, H. Chen, and W. Ye, "A hybrid cnn–lstm network for the classification of human activities based on micro-doppler radar," *Ieee Access*, vol. 8, pp. 24 713–24 720, 2020.

[128] M. Wang, Y. D. Zhang, and G. Cui, "Human motion recognition exploiting radar with stacked recurrent neural network," *Digital Signal Processing*, vol. 87, pp. 125–131, 2019.

[129] H. Du, Y. He, and T. Jin, "Transfer learning for human activities classification using micro-doppler spectrograms," in *2018 IEEE International Conference on Computational Electromagnetics (ICCEM)*. IEEE, 2018, pp. 1–3.

[130] Y. Lang, Q. Wang, Y. Yang, C. Hou, D. Huang, and W. Xiang, "Unsupervised domain adaptation for micro-doppler human motion classification via feature fusion," *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 3, pp. 392–396, 2018.

[131] Y. Kim, J. Park, and T. Moon, "Classification of micro-doppler signatures of human aquatic activity through simulation and measurement using transferred learning," in *Radar Sensor Technology XXI*, vol. 10188. SPIE, 2017, pp. 324–329.

[132] Y. Yang, C. Hou, Y. Lang, D. Guan, D. Huang, and J. Xu, "Open-set human activity recognition based on micro-doppler signatures," *Pattern Recognition*, vol. 85, pp. 60–69, 2019.

[133] P. Cao, W. Xia, M. Ye, J. Zhang, and J. Zhou, "Radar-id: human identification based on radar micro-doppler signatures using deep convolutional neural networks," *IET Radar, Sonar & Navigation*, vol. 12, no. 7, pp. 729–734, 2018.

[134] H. Zhou, Y. Zhao, Y. Liu, S. Lu, X. An, and Q. Liu, "Multi-sensor data fusion and cnn-lstm model for human activity recognition system," *Sensors*, vol. 23, no. 10, p. 4750, 2023.

[135] Y. Shao, S. Guo, L. Sun, and W. Chen, "Human motion classification based on range information with deep convolutional neural network," in *2017 4th International Conference on Information Science and Control Engineering (ICISCE)*. IEEE, 2017, pp. 1519–1523.

[136] Z. Zhang, Z. Tian, and M. Zhou, "Latern: Dynamic continuous hand gesture recognition using fmcw radar sensor," *IEEE Sensors Journal*, vol. 18, no. 8, pp. 3278–3289, 2018.

[137] H. Karimi, T. Derr, and J. Tang, "Characterizing the decision boundary of deep neural networks," *arXiv preprint arXiv:1912.11460*, 2019.

[138] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, and Q. He, "A comprehensive survey on transfer learning," *Proceedings of the IEEE*, vol. 109, no. 1, pp. 43–76, 2020.

[139] K. Yang, T. Yang, Y. Yao, and S. dong Fan, "A transfer learning-based convolutional neural network and its novel application in ship spare-parts classification," *Ocean and Coastal Management*, vol. 215, p. 105971, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0964569121004543

[140] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 248–255.

[141] J. Park, R. J. Javier, T. Moon, and Y. Kim, "Micro-doppler based classification of human aquatic activities via transfer learning of convolutional neural networks," *Sensors*, vol. 16, no. 12, p. 1990, 2016.

[142] H.-C. Shin, H. R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, and R. M. Summers, "Deep convolutional neural networks for computer-aided detection: Cnn architectures, dataset characteristics and transfer learning," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1285–1298, 2016.

[143] M. S. Seyfioglu, B. Erol, S. Z. Gurbuz, and M. G. Amin, "Dnn transfer learning from diversified micro-doppler for motion classification," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 55, no. 5, pp. 2164–2180, 2019.

[144] M. S. Seyfioğlu, A. M. Özbayoğlu, and S. Z. Gürbüz, "Deep convolutional autoencoder for radar-based classification of similar aided and unaided human activities," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 54, no. 4, pp. 1709–1723, 2018.

[145] Z. Wu, C. Shen, and A. Van Den Hengel, "Wider or deeper: Revisiting the resnet model for visual recognition," *Pattern recognition*, vol. 90, pp. 119–133, 2019.

[146] S. Ahmed and M. Bons, "Edge computed nilm: a phone-based implementation using mobilenet compressed by tensorflow lite," in *Proceedings of the 5th international workshop on non-intrusive load monitoring*, 2020, pp. 44–48.

[147] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[148] H. Rahmani and A. Mian, "Learning a non-linear knowledge transfer model for cross-view action recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 2458–2466.

[149] R. Shu, H. H. Bui, H. Narui, and S. Ermon, "A dirt-t approach to unsupervised domain adaptation," *arXiv preprint arXiv:1802.08735*, 2018.

[150] A. Ray and M. H. Kolekar, "Transfer learning and its extensive appositeness in human activity recognition: A survey," *Expert Systems with Applications*, vol. 240, p. 122538, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417423030403

[151] A. R. Sanabria and J. Ye, "Unsupervised domain adaptation for activity recognition across heterogeneous datasets," *Pervasive and Mobile Computing*, vol. 64, p. 101147, 2020.

[152] E. Al Hadhrami, M. Al Mufti, B. Taha, and N. Werghi, "Transfer learning with convolutional neural networks for moving target classification with micro-doppler radar spectrograms," in *2018 International Conference on Artificial Intelligence and Big Data (ICAIBD)*. IEEE, 2018, pp. 148–154.

[153] Y. Yang, Y. Zhang, H. Ji, B. Li, and C. Song, "Radar-based human activity recognition under the limited measurement data support using domain translation," *IEEE Signal Processing Letters*, vol. 29, pp. 1993–1997, 2022.

[154] M. Chakraborty, H. C. Kumawat, S. V. Dhavale *et al.*, "Application of dnn for radar micro-doppler signature-based human suspicious activity recognition," *Pattern Recognition Letters*, vol. 162, pp. 1–6, 2022.

[155] S. Mavaddati, "Voice-based age, gender, and language recognition based on resnet deep model and transfer learning in spectro-temporal domain," *Neurocomputing*, vol. 580, p. 127429, 2024.

[156] E. M. El Houby, "Covid-19 detection from chest x-ray images using transfer learning," *Scientific Reports*, vol. 14, no. 1, p. 11639, 2024.

[157] S. G. Dhekane and T. Ploetz, "Transfer learning in human activity recognition: A survey," *arXiv preprint arXiv:2401.10185*, 2024.

[158] "Vital Signs (Body Temperature, Pulse Rate, Respiration Rate, Blood Pressure) | Johns Hopkins Medicine." [Online]. Available: https://www.hopkinsmedicine.org/health/conditions-and-diseases/vital-signs-body-temperature-pulse-rate-respiration-rate-blood-pressure

[159] S. Asgarzadeh, A. Ebadi, A. S. Shahrbabaki, S. Safari, S. H. Aghili, M. F. Ranjbar, and S. Sadeghi, "National early warning score in predicting adverse outcomes for patients admitted to emergency department; a prognostic accuracy study," *Archives of Academic Emergency Medicine*, vol. 12, no. 1, p. e1, 2023.

[160] S. Lee, Y.-D. Park, Y.-J. Suh, and S. Jeon, "Design and implementation of monitoring system for breathing and heart rate pattern using wifi signals," in *2018 15th IEEE Annual Consumer Communications & Networking Conference (CCNC)*. IEEE, 2018, pp. 1–7.

[161] F. C. Commission, "Revision of part 15 of the commission's rules regarding ultra-wideband transmission systems," Federal Communications Commission, Washington, D.C., First Report and Order FCC 02-48, 2002, eT Docket 98-153. Adopted: February 14, 2002. Released: April 22, 2002. [Online]. Available: https://transition.fcc.gov/Bureaus/Engineering_Technology/Orders/2002/fcc02048.pdf

[162] X. Liang, J. Deng, H. Zhang, and T. A. Gulliver, "Ultra-wideband impulse radar through-wall detection of vital signs," *Scientific reports*, vol. 8, no. 1, p. 13367, 2018.

[163] A. Vorobyov, E. Daskalaki, E. Le Roux, J. Farserotu, and P. Dallemagne, "Contactless vital signs sensing: a survey, preliminary results and challenges," in *2020 XXXIIIrd General Assembly and Scientific Symposium of the International Union of Radio Science*. IEEE, 2020, pp. 1–4.

[164] M. S. Raheel, F. Tubbal, R. Raad, P. Ogunbona, J. Coyte, C. Patterson, D. Perlman, S. Iranmanesh, N. Odeh, and J. Foroughi, "Contactless vital sign monitoring systems: a comprehensive survey of remote health sensing for heart rate and respiration in internet of things and sleep applications," *Sensors and Diagnostics*, vol. 3, no. 7, pp. 1085–1118, 2024.

[165] M. Kebe, R. Gadhafi, B. Mohammad, M. Sanduleanu, H. Saleh, and M. Al-Qutayri, "Human vital signs detection methods and potential using radars: A review," *Sensors*, vol. 20, no. 5, p. 1454, 2020.

[166] T. Hall, D. Y. Lie, T. Q. Nguyen, J. C. Mayeda, P. E. Lie, J. Lopez, and R. E. Banister, "Non-contact sensor for long-term continuous vital signs monitoring: A review on intelligent phased-array doppler sensor design," *Sensors*, vol. 17, no. 11, p. 2632, 2017.

[167] M. Bahache, J. P. Lemayian, W. Wang, and J. M. Hamamreh, "An inclusive survey of contactless wireless sensing: A technology used for remotely monitoring vital signs has the potential to combating covid-19," *RS Open Journal on Innovative Communication Technologies*, vol. 1, no. 2, 2020.

[168] T. Lauteslager, M. Maslik, F. Siddiqui, S. Marfani, G. D. Leschziner, and A. J. Williams, "Validation of a new contactless and continuous respiratory rate monitoring device based on ultra-wideband radar technology," *Sensors*, vol. 21, no. 12, p. 4027, 2021.

[169] X. Yang, Y. Yu, H. Qian, X. Zhang, and L. Zhang, "Body orientation and vital sign measurement with ir-uwb radar network," in *2020 42nd Annual International Conference*

of the IEEE Engineering in Medicine & Biology Society (EMBC). IEEE, 2020, pp. 485–488.

[170] S. Wu, T. Sakamoto, K. Oishi, T. Sato, K. Inoue, T. Fukuda, K. Mizutani, and H. Sakai, "Person-specific heart rate estimation with ultra-wideband radar using convolutional neural networks," *IEEE Access*, vol. 7, pp. 168 484–168 494, 2019.

[171] M. Le, "Heart rate extraction based on eigenvalues using uwb impulse radar remote sensing," *Sensors and Actuators A: Physical*, vol. 303, p. 111689, 2020.

[172] Z. Duan and J. Liang, "Non-contact detection of vital signs using a uwb radar sensor," *IEEE Access*, vol. 7, pp. 36 888–36 895, 2019.

[173] X. Shang, J. Liu, and J. Li, "Multiple object localization and vital sign monitoring using ir-uwb mimo radar," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 56, no. 6, pp. 4437–4450, 2020.

[174] J.-Y. Park, Y. Lee, R. Heo, H.-K. Park, S.-H. Cho, S. H. Cho, and Y.-H. Lim, "Preclinical evaluation of noncontact vital signs monitoring using real-time ir-uwb radar and factors affecting its accuracy," *Scientific Reports*, vol. 11, no. 1, p. 23602, 2021.

[175] J.-M. Muñoz-Ferreras, J. Wang, Z. Peng, C. Li, and R. Gómez-García, "Fmcw-radar-based vital-sign monitoring of multiple patients," in *2019 IEEE MTT-S International Microwave Biomedical Conference (IMBioC)*, vol. 1, 2019, pp. 1–3.

[176] E. Turppa, J. M. Kortelainen, O. Antropov, and T. Kiuru, "Vital sign monitoring using fmcw radar in various sleeping scenarios," *Sensors*, vol. 20, no. 22, p. 6505, 2020.

[177] S. Yao, J. Cong, D. Li, and Z. Deng, "Non-contact vital sign monitoring with fmcw radar via maximum likelihood estimation," *IEEE Internet of Things Journal*, 2024.

[178] Y. Wang, W. Wang, M. Zhou, A. Ren, and Z. Tian, "Remote monitoring of human vital signs based on 77-ghz mm-wave fmcw radar," *Sensors*, vol. 20, no. 10, p. 2999, 2020.

[179] J. Liu, Y. Li, C. Li, C. Gu, and J.-F. Mao, "Accurate measurement of human vital signs with linear fmcw radars under proximity stationary clutters," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 15, no. 6, pp. 1393–1404, 2021.

[180] M. He, J. Ma, C. Wan, Z. Cao, and Z. Cui, "Vital sign detection system based on multi-vital box fitting approximation," in *IGARSS 2023-2023 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2023, pp. 7467–7470.

[181] A. Wazwaz, K. Amin, N. Semary, and T. Ghanem, "Dynamic and distributed intelligence over smart devices, internet of things edges, and cloud computing for human activity

recognition using wearable sensors," *Journal of Sensor and Actuator Networks*, vol. 13, no. 1, p. 5, 2024.

[182] P. P. Ray, "A review on tinyml: State-of-the-art and prospects," *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no. 4, pp. 1595–1623, 2022.

[183] J. Zhu, X. Lou, and W. Ye, "Lightweight deep learning model in mobile-edge computing for radar-based human activity recognition," *IEEE Internet of Things Journal*, vol. 8, no. 15, pp. 12 350–12 359, 2021.

[184] C. Gianoglio, A. Mohanna, A. Rizik, L. Moroney, and M. Valle, "On edge human action recognition using radar-based sensing and deep learning," *IEEE Transactions on Industrial Informatics*, vol. 20, no. 3, pp. 4160–4172, 2023.

[185] S. S. Yadav, S. Anand, M. Adithya, D. S. Nikitha, and C. S. Thakur, "tinyradar: Lstm-based real-time multi-target human activity recognition for edge computing," in *2024 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, 2024, pp. 1–5.

[186] D. Kankipati, M. Munasala, D. S. Nikitha, S. S. Yadav, S. Rao, and C. S. Thakur, "tinyradar for gesture recognition: A low-power system for edge computing," in *2023 IEEE Asia Pacific Conference on Circuits and Systems (APCCAS)*. IEEE, 2023, pp. 75–79.

[187] L. Chen, Y. Chen, J. Xi, and X. Le, "Knowledge from the original network: restore a better pruned network with knowledge distillation," *Complex & Intelligent Systems*, pp. 1–10, 2021.

[188] A. Gordon, E. Eban, O. Nachum, B. Chen, H. Wu, T.-J. Yang, and E. Choi, "Morphnet: Fast & simple resource-constrained structure learning of deep networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 1586–1595.

[189] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, vol. 25, 2012.

[190] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein *et al.*, "Imagenet large scale visual recognition challenge," *International journal of computer vision*, vol. 115, pp. 211–252, 2015.

[191] T. Brown, B. Mann, N. Ryder, M. Subbiah, J. D. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell *et al.*, "Language models are few-shot learners," *Advances in neural information processing systems*, vol. 33, pp. 1877–1901, 2020.

[192] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1–9.

[193] B. Rokh, A. Azarpeyvand, and A. Khanteymoori, "A comprehensive survey on model quantization for deep neural networks in image classification," *ACM Transactions on Intelligent Systems and Technology*, vol. 14, no. 6, pp. 1–50, 2023.

[194] H. Wu, P. Judd, X. Zhang, M. Isaev, and P. Micikevicius, "Integer quantization for deep learning inference: Principles and empirical evaluation," *arXiv preprint arXiv:2004.09602*, 2020.

[195] M. S. Seyfioğlu, A. M. Özbayoğlu, and S. Z. Gürbüz, "Deep convolutional autoencoder for radar-based classification of similar aided and unaided human activities," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 54, no. 4, pp. 1709–1723, 2018.

[196] L. Zhu, D. Liu, X. Li, J. Lu, L. Wei, and X. Cheng, "An efficient hardware architecture for epileptic seizure detection using eeg signals based on 1d-cnn," in *2021 IEEE 14th International Conference on ASIC (ASICON)*. IEEE, 2021, pp. 1–4.

[197] T. Barakbayeva and F. M. Demirci, "Fully automatic cnn design with inception and resnet blocks," *Neural Computing and Applications*, vol. 35, no. 2, pp. 1569–1580, 2023.

[198] M. Yang, T. Guo, T. Zhu, I. Tjuawinata, J. Zhao, and K.-Y. Lam, "Local differential privacy and its applications: A comprehensive survey," *Computer Standards & Interfaces*, vol. 89, p. 103827, 2024.

[199] K. Zhang, X. Song, C. Zhang, and S. Yu, "Challenges and future directions of secure federated learning: a survey," *Frontiers of computer science*, vol. 16, pp. 1–8, 2022.

[200] C. Ding, S. Guo, G. Cui, and X. Yang, "A parameter estimation and deep learning hybrid extraction network for multi-directional human activity recognition based on mmwave radar," *IEEE Internet of Things Journal*, 2024.

[201] Y. L. Coelho, F. d. A. S. dos Santos, A. Frizera-Neto, and T. F. Bastos-Filho, "A lightweight framework for human activity recognition on wearable devices," *IEEE Sensors Journal*, vol. 21, no. 21, pp. 24 471–24 481, 2021.

[202] H. Abedi, A. Ansariyan, P. P. Morita, A. Wong, J. Boger, and G. Shaker, "Ai-powered noncontact in-home gait monitoring and activity recognition system based on mm-wave fmcw radar and cloud computing," *IEEE Internet of Things Journal*, vol. 10, no. 11, pp. 9465–9481, 2023.

[203] X. Yang, W. Gao, X. Qu, P. Yin, H. Meng, and A. E. Fathy, "A lightweight multi-scale neural network for indoor human activity recognition based on macro and micro-doppler features," *IEEE Internet of Things Journal*, 2023.

[204] W.-L. Hsu, J.-X. Liu, C.-C. Yang, and J.-S. Leu, "A fall detection system based on fmcw radar range-doppler image and bi-lstm deep learning," *IEEE Sensors Journal*, 2023.

[205] V. Tuytte, A. Heidari, L. Werthen-Brabants, T. Dhaene, and I. Couckuyt, "Optimized data transmission for radar-based edge-cloud human activity recognition via quantization," in *2024 21st European Radar Conference (EuRAD)*. IEEE, 2024, pp. 145–148.

[206] J. Tian, Y. Zou, F. Liu, C. Yuan, Y. Zhong, J. Lai, and D. Li, "Human-computer interactions using low-cost millimeter wave radar sensors," in *2023 IEEE 20th International Conference on Mobile Ad Hoc and Smart Systems (MASS)*. IEEE, 2023, pp. 584–591.

[207] S. Z. Gurbuz and M. G. Amin, "Radar-based human-motion recognition with deep learning: Promising applications for indoor monitoring," *IEEE Signal Processing Magazine*, vol. 36, no. 4, pp. 16–28, 2019.

[208] T. Han, W. Kang, and G. Choi, "Ir-uwb sensor based fall detection method using cnn algorithm," *Sensors*, vol. 20, no. 20, p. 5948, 2020.

[209] D. Gusland, J. M. Christiansen, B. Torvik, F. Fioranelli, S. Z. Gurbuz, and M. Ritchie, "Open radar initiative: Large scale dataset for benchmarking of micro-doppler recognition algorithms," in *2021 IEEE Radar Conference (RadarConf21)*, 2021, pp. 1–6.

[210] F. Fioranelli, J. Le Kernec, and S. A. Shah, "Radar for health care: Recognizing human activities and monitoring vital signs," *IEEE Potentials*, vol. 38, no. 4, pp. 16–23, 2019.

[211] M. Ritchie, R. Capraru, and F. Fioranelli, "Dop-net: a micro-doppler radar data challenge," *Electronics Letters*, vol. 56, no. 11, pp. 568–570, 2020.

[212] G. Bhavanasi, L. Werthen-Brabants, T. Dhaene, and I. Couckuyt, "Patient activity recognition using radar sensors and machine learning," *Neural Computing and Applications*, vol. 34, no. 18, pp. 16 033–16 048, 2022.

[213] S. Z. Gurbuz, M. M. Rahman, E. Kurtoglu, T. Macks, and F. Fioranelli, "Cross-frequency training with adversarial learning for radar micro-doppler signature classification (rising researcher)," in *Radar Sensor Technology XXIV*, vol. 11408. SPIE, 2020, pp. 58–68.

[214] A. Safa, F. Corradi, L. Keuninckx, I. Ocket, A. Bourdoux, F. Catthoor, and G. G. Gielen, "Improving the accuracy of spiking neural networks for radar gesture recognition through preprocessing," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 34, no. 6, pp. 2869–2881, 2021.

[215] L. Tang, Y. Jia, Y. Qian, S. Yi, and P. Yuan, "Human activity recognition based on mixed cnn with radar multi-spectrogram," *IEEE Sensors Journal*, vol. 21, no. 22, pp. 25 950–25 962, 2021.

[216] J. Liang, R. He, and T. Tan, "A comprehensive survey on test-time adaptation under distribution shifts," *International Journal of Computer Vision*, pp. 1–34, 2024.

[217] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.

[218] X. Yu, N. Zeng, S. Liu, and Y.-D. Zhang, "Utilization of densenet201 for diagnosis of breast abnormality," *Machine Vision and Applications*, vol. 30, pp. 1135–1144, 2019.

[219] L. Gao, X. Zhang, T. Yang, B. Wang, and J. Li, "The application of resnet-34 model integrating transfer learning in the recognition and classification of overseas chinese frescoes," *Electronics*, vol. 12, no. 17, p. 3677, 2023.

[220] M. Jannat, R. Karim, N. Z. Islam, A. N. Chy, and A. K. Muhammad Masum, "Human activity recognition using ensemble of cnn-based transfer learning models," in *2023 IEEE International Conference on Computing (ICOCO)*, 2023, pp. 112–117.

[221] S. Takano, "Chapter 2-traditional microarchitectures," *Thinking machines, S. Takano, Ed. Academic Press*, pp. 19–47, 2021.

[222] S. Yoo, S. Ahmed, S. Kang, D. Hwang, J. Lee, J. Son, and S. H. Cho, "Radar recorded child vital sign public dataset and deep learning-based age group classification framework for vehicular application," *Sensors*, vol. 21, no. 7, p. 2412, 2021.

[223] N. Du, K. Liu, L. Ge, and J. Zhang, "Apnearadar: A 24ghz radar-based contactless sleep apnea detection system," in *2017 2nd International Conference on Frontiers of Sensors Technologies (ICFST)*. IEEE, 2017, pp. 372–376.

[224] S. A. T. Hosseini and H. Amindavar, "Uwb radar signal processing in measurement of heartbeat features," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2017, pp. 1004–1007.

[225] C. Li, V. M. Lubecke, O. Boric-Lubecke, and J. Lin, "A review on recent advances in doppler radar sensors for noncontact healthcare monitoring," *IEEE Transactions on microwave theory and techniques*, vol. 61, no. 5, pp. 2046–2060, 2013.

[226] D. Wang, S. Yoo, and S. H. Cho, "Experimental comparison of ir-uwb radar and fmcw radar for vital signs," *Sensors*, vol. 20, no. 22, p. 6695, 2020.

[227] C. Hadjipanayi, M. Yin, A. Bannon, A. Rapeaux, M. Banger, S. Haar, T. S. Lande, A. H. McGregor, and T. G. Constandinou, "Remote gait analysis using ultra-wideband radar technology based on joint range-doppler-time representation," *IEEE Transactions on Biomedical Engineering*, 2024.

[228] Z. Liang, M. Xiong, Y. Jin, J. Chen, D. Zhao, D. Yang, B. Liang, and J. Mo, "Non-contact human vital signs extraction algorithms using ir-uwb radar: A review," *Electronics*, vol. 12, no. 6, p. 1301, 2023.

[229] W. Han, S. Dai, and M. R. Yuce, "Real-time contactless respiration monitoring from a radar sensor using image processing method," *IEEE Sensors Journal*, vol. 22, no. 19, pp. 19 020–19 029, 2022.

[230] R. N. Khushaba and A. J. Hill, "Radar-based materials classification using deep wavelet scattering transform: A comparison of centimeter vs. millimeter wave units," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 2016–2022, 2022.

[231] "Walabot 3D Sensor | Walabot API: Imaging Features." [Online]. Available: https://api.walabot.com/_features.html

[232] F. Ayaz, M. S. Khan, S. Hussain, W. Ahmad, F. Kawsar, M. A. Imran, and A. Zoha, "Contact-free vital sign estimation using ultra-wide band radar," in *2022 29th IEEE International Conference on Electronics, Circuits and Systems (ICECS)*, 2022, pp. 1–4.

[233] A. Chebrolu, "Fallwatch: A novel approach for through-wall fall detection in real-time for the elderly using artificial intelligence," in *2021 Third International Conference on Transdisciplinary AI (TransAI)*. IEEE, 2021, pp. 57–63.

[234] "Go direct® respiration belt - vernier." [Online]. Available: https://www.vernier.com/product/go-direct-respiration-belt/?srsltid=AfmBOopfpYWBuoY7kLdXK_GoAU9JmxeVZGp8dwyd0z-zO7fJfE9QBGJM

[235] N. J. Napoli, V. R. Rodrigues, and P. W. Davenport, "Characterizing and modeling breathing dynamics: Flow rate, rhythm, period, and frequency," *Frontiers in Physiology*, vol. 12, p. 772295, 2022.

[236] M. A. Russo, D. M. Santarelli, and D. O'Rourke, "The physiological effects of slow breathing in the healthy human," *Breathe*, vol. 13, no. 4, pp. 298–309, 2017.

[237] S. Wang, A. Pohl, T. Jaeschke, M. Czaplik, M. Köny, S. Leonhardt, and N. Pohl, "A novel ultra-wideband 80 ghz fmcw radar system for contactless monitoring of vital signs," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2015, pp. 4978–4981.

[238] S. Iyer, L. Zhao, M. P. Mohan, J. Jimeno, M. Y. Siyal, A. Alphones, and M. F. Karim, "mm-wave radar-based vital signs monitoring and arrhythmia detection using machine learning," *Sensors*, vol. 22, no. 9, p. 3106, 2022.

[239] T. Zheng, Z. Chen, S. Zhang, C. Cai, and J. Luo, "More-fi: Motion-robust and fine-grained respiration monitoring via deep-learning uwb radar," in *Proceedings of the 19th ACM conference on embedded networked sensor systems*, 2021, pp. 111–124.

[240] T. Zheng, Z. Chen, C. Cai, J. Luo, and X. Zhang, "V2ifi: In-vehicle vital sign monitoring via compact rf sensing," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 4, no. 2, pp. 1–27, 2020.

[241] "Iwr6843 data sheet, product information and support | ti.com." [Online]. Available: https://www.ti.com/product/IWR6843

[242] "Dca1000evm evaluation board | ti.com." [Online]. Available: https://www.ti.com/tool/DCA1000EVM

[243] "In-ear sensors - cosinuss°." [Online]. Available: https://www.cosinuss.com/en/products/in-ear-sensors/

[244] Y. Himeur, A. N. Sayed, A. Alsalemi, F. Bensaali, and A. Amira, "Edge ai for internet of energy: Challenges and perspectives," *Internet of Things*, vol. 25, p. 101035, 2024.

[245] C. Banbury, V. J. Reddi, P. Torelli, J. Holleman, N. Jeffries, C. Kiraly, P. Montino, D. Kanter, S. Ahmed, D. Pau *et al.*, "Mlperf tiny benchmark," *arXiv preprint arXiv:2106.07597*, 2021.

[246] X. Wang, Y. Han, C. Wang, Q. Zhao, X. Chen, and M. Chen, "In-edge ai: Intelligentizing mobile edge computing, caching and communication by federated learning," *Ieee Network*, vol. 33, no. 5, pp. 156–165, 2019.

[247] S. Deng, H. Zhao, W. Fang, J. Yin, S. Dustdar, and A. Y. Zomaya, "Edge intelligence: The confluence of edge computing and artificial intelligence," *IEEE Internet of Things Journal*, vol. 7, no. 8, pp. 7457–7469, 2020.

[248] I. J. Tsang, F. Corradi, M. Sifalakis, W. Van Leekwijck, and S. Latré, "Radar-based hand gesture recognition using spiking neural networks," *Electronics*, vol. 10, no. 12, p. 1405, 2021.

[249] B. Courty, V. Schmidt, S. Luccioni, Goyal-Kamal, MarionCoutarel, B. Feld, J. Lecourt, LiamConnell, A. Saboni, Inimaz, supatomic, M. Léval, L. Blanche, A. Cruveiller, ouminasara, F. Zhao, A. Joshi, A. Bogroff, H. de Lavoreille, N. Laskaris, E. Abati, D. Blank, Z. Wang, A. Catovic, M. Alencon, M. Stęchły, C. Bauer, L. O. N. de Araújo,

JPW, and MinervaBooks, "mlco2/codecarbon: v2.4.1," May 2024. [Online]. Available: https://doi.org/10.5281/zenodo.11171501

[250] J. Lang, Z. Guo, and S. Huang, "A comprehensive study on quantization techniques for large language models," in *2024 4th International Conference on Artificial Intelligence, Robotics, and Communication (ICAIRC)*. IEEE, 2024, pp. 224–231.

[251] M. A. Husnoo, A. Anwar, N. Hosseinzadeh, S. N. Islam, A. N. Mahmood, and R. Doss, "A secure federated learning framework for residential short-term load forecasting," *IEEE Transactions on Smart Grid*, vol. 15, no. 2, pp. 2044–2055, 2023.

[252] Q. Miao, W. Jing, and H. Song, "Differential privacy–based location privacy enhancing in edge computing," *Concurrency and Computation: Practice and Experience*, vol. 31, no. 8, p. e4735, 2019.

[253] C. Xu, J. Ren, D. Zhang, and Y. Zhang, "Distilling at the edge: A local differential privacy obfuscation framework for iot data analytics," *IEEE Communications Magazine*, vol. 56, no. 8, pp. 20–25, 2018.

[254] T. Qi, F. Wu, C. Wu, L. He, Y. Huang, and X. Xie, "Differentially private knowledge transfer for federated learning," *Nature Communications*, vol. 14, no. 1, p. 3785, 2023.

[255] E. Cai, D.-C. Juan, D. Stamoulis, and D. Marculescu, "Neuralpower: Predict and deploy energy-efficient convolutional neural networks," in *Asian Conference on Machine Learning*. PMLR, 2017, pp. 622–637.

[256] E. Tavanti, A. Rizik, A. Fedeli, D. D. Caviglia, and A. Randazzo, "A short-range fmcw radar-based approach for multi-target human-vehicle detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–16, 2021.

[257] X. Feng, Y. Weng, W. Li, P. Chen, and H. Zheng, "Damun: A domain adaptive human activity recognition network based on multimodal feature fusion," *IEEE Sensors Journal*, vol. 23, no. 18, pp. 22 019–22 030, 2023.

[258] M. El Hadi, A. Ouariach, R. Essaadaoui, A. El Moussaouy, and O. Mommadi, "Rc time constant measurement using an ina219 sensor: creating an alternative, flexible, low-cost configuration that provides benefits for students and schools," *Physics Education*, vol. 56, no. 4, p. 045015, 2021.