Università di Genova

PhD Program in Science and Technology for Electronic and Telecommunication Engineering

IoT solutions for e-Health applications for care's continuity at home

Caterina Fallani

Thesis submitted for the degree of *Doctor of Philosophy* (38° cycle)

10 March 2026

Prof. Fabio Lavagetto	Tutor
Prof. Igor Bisio	Co-tutor
Prof. Andrea Sciarrone	Co-tutor
Prof. ssa Chiara Garibotto	Co-tutor
Prof. Maurizio Valle	Head of the PhD program

Università di Genova
DITEN DIPARTIMENTO
DI INGEGNERIA NAVALE, ELETTRICA,
ELETTRONICA E DELLE TELECOMUNICAZIONI

IoT solutions for e-Health applications for care's continuity at home

Caterina Fallani

⋈

Committee:

President of the committee
Igor Bisio, Full Professor, University of Genoa, Italy

Reviewers
Mauro De Sanctis, Full Professor, University of Roma, Italy
Giuseppe Araniti, Full Professor, University of Reggio Calabria, Italy

Tutors
Fabio Lavagetto, Full Professor, University of Genoa, Italy
Igor Bisio, Full Professor, University of Genoa, Italy

University of Genoa

Caterina Fallani

IoT solutions for e-Health applications for care's continuity at home

xxi+118 p.

To my family and Davide

Abstract

This research investigates the use of *Channel State Information* (CSI) extracted from Wi-Fi signals as a fine-grained and physically interpretable observable for contactless monitoring of human activity, breathing, and environmental characteristics. The study builds on the premise that the complex amplitude and phase of the wireless channel, traditionally employed for communication diagnostics, also contain rich information about subtle propagation variations induced by materials, objects and human motion.

The thesis is organised into five main chapters. The first chapter introduces the theoretical background and motivation, highlighting the growing interest in device-free sensing and the potential of CSI as a bridge between communication and perception. The second chapter reviews the state of the art, tracing the evolution of CSI-based sensing techniques from amplitude-only analysis to phase-calibrated and deep-learning-driven approaches. The third chapter details the data acquisition methodology, the MATLAB simulation tools, and the signal processing pipeline. The fourth chapter presents the experimental and simulation results across three key domains: (i) material and object identification, (ii) recognition of human movements and activities, and (iii) monitoring of the breath rate. Finally, the fifth chapter provides the general conclusions of the research and outlines the directions for future work.

Under controlled line-of-sight (LoS) conditions and with single subjects, the results demonstrated high stability and temporal coherence. Traditional machine learning models such as Support Vector Machines (SVM) and Random Forests achieved accuracies above 95% in material and object classification. Deep learning architectures, including Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks, surpassed 97% accuracy in movement recognition and breathing detection. These results confirmed that CSI encapsulates a dense, multidimensional representation of the radio channel, capable of capturing both macroscopic motion and sub-centimetric physiological oscillations.

However, experiments also revealed a degradation of performance in non-line-of-sight (NLoS) and multi-subject scenarios, where multipath interference and temporal non-stationarity reduce amplitude and phase coherence. This limitation motivated the exploration of *Reconfigurable Intelligent Surfaces* (RIS) as a means to actively shape and strengthen wireless propagation. A MATLAB-based RIS simulator was implemented to model programmable reflective arrays operating at 28 GHz, where the phase of each element is adjusted to steer or reinforce the signal toward the receiver. Simulations varying both the number and position of RIS elements showed signal-to-noise ratio (SNR) gains of up to 20 dB in obstructed configurations, confirming the RIS's ability to restore phase stability and enhance CSI quality in challenging propagation conditions.

The final chapter discusses the broader implications of this research and identifies future directions. The integration of RIS technology represents an initial step toward adaptive wireless environments capable of dynamically controlling propagation to improve sensing accuracy. Future work will include experimental validation of RIS prototypes in real environments, the adoption of ESP32-based Wi-Fi hardware for full OFDM subcarrier

analysis, and the development of hybrid analytical–data-driven models combining physical channel formulations with machine and deep learning inference. In the long term, the aim is to move toward intelligent, self-optimising wireless systems where communication and perception are seamlessly unified.

Keywords: Channel State Information (CSI), Wi-Fi sensing, machine learning, deep learning, Reconfigurable Intelligent Surfaces (RIS), breath rate monitoring, material and object identification, activity recognition, Line-of-Sight (LoS).

Abstract

La presente ricerca indaga l'utilizzo delle *Channel State Information* (CSI) estratte dai segnali Wi-Fi come osservabili ad alta risoluzione e fisicamente interpretabili per il monitoraggio non invasivo delle attività umane, della respirazione e delle caratteristiche ambientali. L'intero lavoro si fonda sull'idea che l'ampiezza e la fase complesse del canale wireless, tradizionalmente impiegate per la diagnostica delle comunicazioni, racchiudano informazioni preziose sulle variazioni di propagazione dovute alla presenza di oggetti, persone e movimenti di minima entità.

La tesi è articolata in cinque capitoli principali. Il primo capitolo introduce il contesto teorico e le motivazioni della ricerca, il secondo capitolo analizza lo stato dell'arte delle tecniche di rilevamento basate su CSI, il terzo descrive la metodologia di acquisizione, gli strumenti MATLAB sviluppati per la simulazione e l'elaborazione dati, il quarto capitolo presenta i risultati ottenuti attraverso esperimenti e simulazioni nei principali ambiti applicativi come l'identificazione di materiali e oggetti; il riconoscimento di movimenti e attività ed il monitoraggio della frequenza respiratoria, mentre il quinto capitolo è dedicato alle conclusioni generali e ai lavori futuri.

In condizioni di linea di vista (LoS) e con singolo soggetto, i risultati hanno mostrato elevata stabilità e coerenza temporale. I modelli di apprendimento automatico tradizionali, come Support Vector Machine (SVM) e Random Forest, hanno raggiunto accuratezze superiori al 95% nell'identificazione di materiali e oggetti. Le architetture di apprendimento profondo, comprendenti reti convoluzionali (CNN) e ricorrenti (LSTM), hanno superato il 97% nel riconoscimento dei movimenti e nella rilevazione della respirazione. Questi risultati confermano che il CSI costituisce una rappresentazione densa e multidimensionale del canale radio, capace di catturare sia movimenti macroscopici sia oscillazioni sub-centimetriche legate alla fisiologia.

Gli esperimenti hanno tuttavia evidenziato un deterioramento delle prestazioni in condizioni di non linea di vista (NLoS) o in presenza di più soggetti, a causa della perdita di coerenza tra le sottoportanti dovuta alla dinamica multipercorso. Questa constatazione ha motivato l'introduzione delle Superfici Intelligenti Riconfigurabili (RIS) come possibile soluzione per il controllo attivo della propagazione del segnale. È stato utilizzato un simulatore MATLAB per modellare una superficie riflettente programmabile operante a 28 GHz, con elementi dotati di controllo di fase in grado di rinforzare o dirigere il segnale verso il ricevitore. Le simulazioni, effettuate variando il numero di elementi e la posizione della superficie lungo l'asse x, hanno mostrato incrementi del rapporto segnale-rumore (SNR) fino a 20 dB, confermando la capacità del RIS di ripristinare la coerenza di fase e migliorare la qualità del CSI in scenari complessi.

L'ultimo capitolo delinea le prospettive future della ricerca. L'integrazione della tecnologia RIS rappresenta un primo passo verso ambienti wireless capaci di controllare dinamicamente la propagazione del segnale per migliorare l'accuratezza del sensing. Gli sviluppi futuri prevedono la validazione sperimentale di prototipi RIS in ambienti reali, l'adozione di hardware Wi-Fi basato su ESP32 per l'analisi completa delle sottoportanti OFDM e lo sviluppo di modelli ibridi che combinino le formulazioni fisiche del canale

con l'inferenza basata sui dati. A lungo termine, l'obiettivo è quello di realizzare un sistema adattivo e intelligente, in grado di unire comunicazione e percezione in un'unica infrastruttura.

Parole chiave: Informazioni sullo Stato del Canale (CSI), rilevamento Wi-Fi, machine learning, deep learning, Superficie Intelligente Riconfigurabile (RIS), propagazione LoS, monitoraggio della frequenza respiratoria, riconoscimento delle attività, identificazione di oggetti e materiali.

Contents

List of Figures

List of Tables

Notations

c	speed of light in vacuum (3×10^8 m/s)
f_c	carrier frequency (Hz)
f_k	frequency of the k-th OFDM subcarrier (Hz)
λ	wavelength corresponding to f_c ($\lambda = c/f_c$)
t	time (s)
k	subcarrier index ($k = 1, \ldots, N_{\text{sc}}$)
N_{sc}	number of OFDM subcarriers
$N_{\text{tx}}, N_{\text{rx}}$	number of transmit and receive antennas
N	number of RIS reflecting elements
D	distance between transmitter (AP) and receiver (UE) (m)
x_{RIS}	position of RIS along the x-axis (m)
v	velocity vector of the moving subject (m/s)
σ_n^2	noise power
$H_k(t)$	complex channel frequency response (CFR) of the k-th subcarrier at time t
$A_k(t)$	amplitude of the k-th subcarrier, $A_k(t) = \lvert H_k(t) \rvert$
$\phi_k(t)$	phase of the k-th subcarrier, $\phi_k(t) = \arg(H_k(t))$
$\Delta\phi_k(t)$	differential phase between two receive antennas
$\mathbf{H}(t)$	MIMO channel matrix of size $N_{\text{rx}} \times N_{\text{tx}}$
$\rho(A_k, A_{k'})$	correlation coefficient between amplitudes of subcarriers k and k'
$\text{Var}[A_k]$	temporal variance of amplitude on subcarrier k
$\text{PSD}(f)$	power spectral density (dB/Hz)
SNR	signal-to-noise ratio (dB)
$\mathcal{F}\{\cdot\}$	Fourier transform operator
$\mathbb{E}[\cdot]$	statistical expectation operator
$\mathbf{X}$	matrix of extracted CSI features for ML/DL processing
$f(\mathbf{x}; \theta)$	neural network prediction function with parameters θ
$L(\theta)$	loss function during model training
f_b	breathing frequency (Hz)

Acronyms

AP	Access Point
BPSK	Binary Phase Shift Keying
CNN	Convolutional Neural Network
CSI	Channel State Information
CFR	Channel Frequency Response
DL	Deep Learning
DSP	Digital Signal Processing
ESP32	Wi-Fi microcontroller with CSI extraction capability
FFT	Fast Fourier Transform
HR	High Resolution
LoS	Line-of-Sight
LSTM	Long Short-Term Memory network
ML	Machine Learning
MIMO	Multiple Input Multiple Output
NLoS	Non-Line-of-Sight
OA	Overall Accuracy
OFDM	Orthogonal Frequency Division Multiplexing
PSD	Power Spectral Density
RF	Random Forests
RIS	Reconfigurable Intelligent Surface
RX	Receiver
SNR	Signal-to-Noise Ratio
SVM	Support Vector Machine
STIET	Science and Technology for Electronic and Telecommunication Engineering
TX	Transmitter
UE	User Equipment
Wi-Fi	Wireless Fidelity

CHAPTER 1

Introduction

1.1 Basic Principles on e-Health and Channel State Information

Monitoring body movement and vital parameters, such as respiratory rate, is crucial for healthcare management and the prevention of critical diseases, especially in long-term and home care settings. Traditional monitoring techniques usually involve contact-based sensors, such as chest bands or accelerometers. While these solutions are widely used in clinical environments, they have several limitations that prevent their practical use in daily, unsupervised settings. The main drawbacks of contact-based techniques are the physical and psychological discomfort caused by wearing or attaching sensors directly to the body, and the potential alteration of the subject's natural behaviour. Furthermore, such devices require regular maintenance and frequent recharging and are prone to measurement errors caused by incorrect placement or mechanical interference. In home environments, these factors often represent significant obstacles to long-term and continuous use, particularly for elderly people, vulnerable patients and people with disabilities.

To address these limitations, Wi-Fi technology is emerging as an innovative, non-invasive solution for remotely monitoring vital signals and motor activities. Wi-Fi devices, which are already widely used in homes, can be repurposed as passive sensors to detect variations in the wireless communication channel caused by body movement or micro-movements of the chest during respiration. By analysing channel-related information such as channel state information (CSI) or received signal strength indicator (RSSI), it is possible to continuously and accurately extract data related to the user's physiological and behavioural state without requiring physical contact. This approach offers numerous advantages: it is unobtrusive for the user; it does not require structural changes to the environment; it enables the simultaneous monitoring of multiple individuals; and it can be easily integrated into existing infrastructures. Furthermore, using Wi-Fi signals for health-related purposes opens up new possibilities in telemedicine and ambient assisted

living, enabling continuous, discreet health surveillance. This is essential for the early detection of abnormal conditions and management of individual well-being.

1.2 RSSI vs CSI

One of the earliest approaches explored was based on the Received Signal Strength Indicator (RSSI), a metric that is commonly available on commercial devices. RSSI measures the overall power of the received signal and has been widely used in applications such as localisation and presence detection. More recently, it has also been used in preliminary studies aimed at estimating respiratory activity in indoor environments. The early adoption of RSSI was motivated by its ease of access and compatibility with existing Wi-Fi infrastructure, which facilitated experimentation without the need for specialised hardware.

In device-free sensing, the RSSI measurement uses a known channel fading model to estimate the distance between the transmitter and receiver based on signal power. As the propagation distance increases, the signal power decreases, resulting in signal fading. Therefore, signal power can be used to estimate the distance between the transmitter and receiver. The RSSI can be estimated using the Log-Normal Distance Path Loss (LDPL):

$$P(r) = \overline{P_0} + log_{10}(\frac{r}{r_0})^{10n} + X_\sigma \tag{1.1}$$

where $P(r)$ is the received signal power (dB) measured at a distance r; $\overline{P_0}$ is the average received signal power measured at a reference point r_0 and n is the path loss exponent. X_σ is a zero-mean normal random variable that reflects the attenuation in decibels caused by shadowing. The RSSI decreases monotonically with distance. However, due to the random shadowing effect X_σ, this trend only holds on a relatively large scale, which is bounded by the shadowing variance σ. The multipath rich indoor environment causes the RSSI to fluctuate on the order of the signal wavelength. This contributes to large shadowing X_σ. Consequently, it is almost impossible to distinguish locations in the vicinity because the large deviation of multipath shadowing blurs the monotonic trend. This fundamentally limits the accuracy of RSSI-based ranging [10]. A key limitation of RSSI is its inability to represent the effects of multipath propagation. Instead, the wireless propagation channel is represented as a temporal linear filter, known as the Channel Impulse Response (CIR), to provide a complete description of each individual path.

Under the time-invariant assumption, CIR $h(\tau)$ is denoted as:

$$h(\tau) = \sum_{i=1}^{N} \alpha_i e^{-j\sigma_i} \delta(\tau - \tau_i), \tag{1.2}$$

where α_i, σ_i and τ_i are the amplitude, phase and time delay of the i^{th} path, respectively. N is the total number of multipath and $\delta(\tau)$ is the Dirac delta function. Each impulse represents a delayed multipath component that is weighted according to its corresponding amplitude and phase. In the frequency domain, constructive and destructive phase interactions interact with each other, giving rise to frequency-selective fading. This phenomenon is characterised by the Channel Frequency Response (CFR), which encompasses the channel's amplitude-frequency and phase-frequency characteristics.

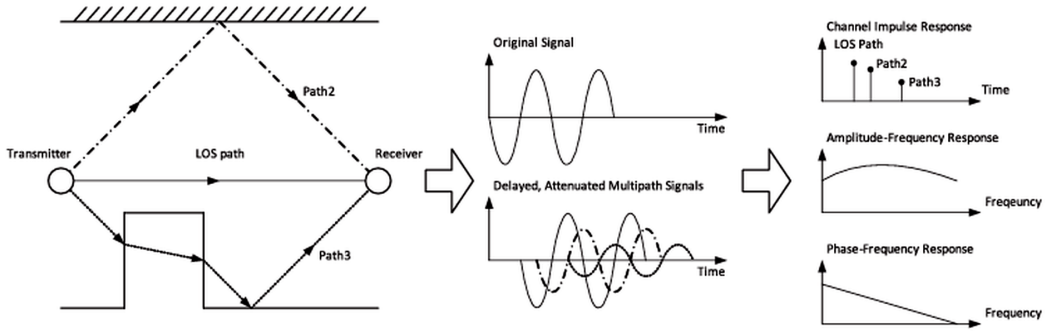

Figure 1.1: *Illustration of multipath propagation in indoor environments. In addition to the direct line-of-sight (LoS) component, reflected paths introduce delayed and attenuated signal replicas, resulting in a composite channel impulse response and frequency-selective amplitude and phase variations captured by CSI.*

Assuming infinite bandwidth, the Channel Impulse Response (CIR) and the Channel Frequency Response (CFR) are equivalent, with the CFR obtained as the Fourier transform of the CIR. Both representations capture the small-scale effects of multipath propagation and are widely used in the characterization and measurement of wireless channels. In the frequency domain f, the corresponding CFR is:

$$H(f) = \sum_{i=1}^{N} \alpha_i e^{-j2\pi f \tau_i}. \tag{1.3}$$

1.2.1 Deriving Channel Response

In the time domain, the received signal $r(t)$ can be expressed as the convolution of the transmitted waveform $s(t)$ with the Channel Impulse Response (CIR) $h(t)$:

$$r(t) = (s * h)(t) = \int_{-\infty}^{+\infty} s(\tau)h(t-\tau)d\tau. \tag{1.4}$$

Applying the Fourier transform, the received signal spectrum $R(t)$ is the multiplication of the transmitted signal spectrum $S(f)$ and the Channel Frequency Response $H(f)$, in the frequency domain.
The Fourier transform and its inverse are defined as follows.

$$\mathcal{F}(\{x(t)\}) = X(f) = \int_{-\infty}^{+\infty} x(t)e^{-j2\pi ft}dt; \tag{1.5}$$

$$\mathcal{F}^{-1}(\{X(f)\}) = x(t) = \int_{-\infty}^{+\infty} X(f)e^{+j2\pi ft}df. \tag{1.6}$$

Applying the Fourier transform to the equation (1.4):

$$R(f) = \int_{-\infty}^{+\infty} r(t)e^{-j2\pi ft}dt = \int_{-\infty}^{+\infty} [\int_{-\infty}^{+\infty} s(\tau)h(t-\tau)d\tau]e^{-j2\pi tf}dt. \tag{1.7}$$

By Fubini's theorem, we can exchange the order of integration:

$$R(f) = \int_{-\infty}^{+\infty} s(\tau)[\int_{-\infty}^{+\infty} h(t-\tau)e^{-j2\pi tf}dt]d\tau. \tag{1.8}$$

Let $u = t - \tau$; thus $t = u + \tau$ and $dt = du$. We have:

$$\int_{-\infty}^{+\infty} h(t-\tau)e^{-j2\pi tf}dt = \int_{-\infty}^{+\infty} h(u)e^{-j2\pi fu}du = e^{-j2\pi f\tau}H(f). \tag{1.9}$$

Substituting, we obtain:

$$R(f) = H(f)\int_{-\infty}^{+\infty} s(\tau)e^{-j2\pi f\tau}d\tau = H(f)S(f). \tag{1.10}$$

Thus, the convolution theorem states:

$$R(f) = S(f) \cdot H(f), \tag{1.11}$$

which shows that convolution in the time domain corresponds to multiplication in frequency domain.
So that

$$H(f) = \frac{R(f)}{S(f)}, \quad h(t) = \mathcal{F}^{-1}\{H(f)\}.$$

In OFDM Wi-Fi, receivers internally estimate the Channel Frequency Response (CFR) on each subcarrier f_k:

$$\mathbf{H}_k = ||\mathbf{H}_k|| e^{j\angle \mathbf{H}_k}, k = 1, \cdots K. \tag{1.12}$$

The vector $\{H_k\}_{k=1}^{K}$ is the Channel State Information (CSI). CSI preserves amplitude and phase at subcarrier resolution, enabling fine-grained characterization of multipath and micro-motions (e.g., respiration), which is not possible with scalar RSSI.
Therefore, we can conclude that the intrinsic limitations of RSSI reduce its effectiveness in scenarios requiring high-precision physiological monitoring. Measurements are strongly influenced by multipath propagation, fading and environmental noise, and they lack detailed information about the structure of the transmitted signal. Its high temporal variability and low sensitivity to micro-movements, such as those associated with breathing or the heartbeat, render it unsuitable for continuous and reliable monitoring in domestic environments. To overcome these limitations, research has increasingly focused on analysing channel state information (CSI), which is accessible via the IEEE 802.11n/ac standards. CSI provides a detailed representation of the wireless communication channel, including amplitude and phase information for each subcarrier in the OFDM modulation scheme. This detailed data enables the accurate detection of signal variations caused by human presence and subtle body movements, even in conditions of non-direct visibility.

Category	RSSI	CSI
Network layer	MAC layer	Physical layer
Time resolution	Packet-level	Multipath cluster scale
Frequency resolution	None	Subcarrier level
Temporal stability	Low	High
Measurement band	RF band	Baseband
Granularity	Coarse-grained	Fine-grained
Universality	Almost all Wi-Fi devices	Some Wi-Fi devices

Table 1.1: *Comparison between Received Signal Strength Indicator (RSSI) and Channel State Information (CSI) as sensing metrics in Wi-Fi systems. RSSI provides a coarse, packet-level measure of signal power at the MAC layer, with no frequency selectivity and limited temporal stability. In contrast, CSI is extracted at the physical layer and captures fine-grained amplitude and phase information at the subcarrier level, reflecting multipath propagation effects in the baseband signal. These characteristics make CSI significantly more informative and suitable for high-resolution wireless sensing applications, such as activity recognition, localisation, and vital-sign monitoring.*

1.3 Channel State Information

Wi-Fi signals can be used to extract Channel State Information (CSI), which provides a detailed and fine-grained description of the wireless propagation channel between a transmitter and a receiver. CSI encapsulates the combined effects of time delay, amplitude attenuation, and phase shift experienced by a transmitted signal as it propagates through the environment. It also reflects key propagation phenomena, such as multipath propagation, shadow fading, and channel-induced distortion.

Modern Wi-Fi standards such as IEEE 802.11n/ac typically employ Multiple-Input Multiple-Output (MIMO) systems, which use multiple transmitting and receiving antennas. MIMO technology enhances diversity gain, array gain, and spectral efficiency (via spatial multiplexing), while simultaneously mitigating co-channel interference.

These systems adopt Orthogonal Frequency Division Multiplexing (OFDM) for modulation. OFDM partitions the available spectrum into multiple orthogonal narrowband subcarriers, each subject to flat fading. This design mitigates the detrimental effects of frequency-selective fading by enabling equalisation of each subcarrier independently. Furthermore, transmitting data in parallel across subcarriers significantly increases throughput while maintaining robustness against multipath delay spread.

1.3.1 Mathematical Model of CSI

In the frequency domain, the narrowband flat-fading MIMO channel is modelled as:

$$\mathbf{Y} = \mathbf{H} \cdot \mathbf{X} + \mathbf{N}, \tag{1.13}$$

where $\mathbf{Y}$ and $\mathbf{X}$ denote the received and transmitted signal vectors, $\mathbf{H}$ represents the Channel's Frequency Response (CFR), and $\mathbf{N}$ is additive white Gaussian noise (AWGN).

The CSI matrix $\mathbf{H}$ has dimensions $N_{tx} \times N_{rx} \times S$, where N_{tx} and N_{rx} are the numbers of transmitting and receiving antennas, and S is the number of subcarriers. Although an IEEE 802.11 OFDM system uses 56 subcarriers over a 20 MHz channel, the Intel 5300 NIC employed in this work reports CSI for only 30 of them.

For each subcarrier k, $\mathbf{H}_k$ is a complex value:

$$\mathbf{H}_k = I_k + jQ_k = ||\mathbf{H}_k||e^{j\angle \mathbf{H}_k}, \tag{1.14}$$

where I_k and Q_k are the in-phase and quadrature components, while $||\mathbf{H}_k||$ and $\angle \mathbf{H}_k$ represent amplitude and phase.

In an indoor NLOS environment, the CFR of subcarrier k can also be expressed as:

$$\mathbf{H}_k = \sum_{i=0}^{K} r_i e^{-j2\pi f_k \tau_i}, \tag{1.15}$$

where K is the number of multipath components, r_i the attenuation, τ_i the propagation delay of the i^{th} path, and f_k the central frequency of subcarrier k.

1.3.2 CSI Structure and Representation

Given a transmitted Wi-Fi signal, CSI can be extracted for each received packet at each subcarrier, yielding a set of channel measurements that describe how the state of each subcarrier changes over time.

Let $\mathbf{H}^k_{M,N}$ denote the CSI matrix for subcarrier k, with M transmitting antennas and N receiving antennas:

$$\mathbf{H}^k_{M,N} = \begin{pmatrix} H_{(1,1)} & \cdots & H_{(1,n)} & \cdots & H_{(1,N)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ H_{(m,1)} & \cdots & H_{(m,n)} & \cdots & H_{(m,N)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ H_{(M,1)} & \cdots & H_{(M,n)} & \cdots & H_{(M,N)} \end{pmatrix} \tag{1.16}$$

Each entry $H_{(m,n)}$ is a complex number representing the amplitude and phase of the channel between transmitting antenna m and receiving antenna n:

$$H_{(m,n)} = \|H_{(m,n)}\| \cdot e^{j\angle H_{(m,n)}}. \tag{1.17}$$

By collecting CSI at the packet level, we obtain $\mathbf{H}^K_{M,N}(t)$ for each time instant t, where K is the set of measured subcarriers. In MIMO-OFDM systems, the multiple signal copies arriving via different paths result in variations of SNR due to multipath fading—traditionally considered a problem for communication, but exploitable for sensing.

1.3.3 Phase Difference Information

Raw CSI phase values from the Intel 5300 NIC are highly random due to asynchrony between transmitter and receiver clocks and oscillators. Two main calibration methods exist: (i) applying a linear transformation to all subcarriers to remove random offsets [11, 12]; (ii) computing the phase difference between two adjacent antennas and subtracting its mean for LOS recognition [13]. While these improve stability or yield zero-mean phase differences, they are not directly optimised for vital sign monitoring.

In the 5 GHz band, the measured phase of subcarrier k can be modelled as [14, 15]:

$$\angle\hat{H}_i = \angle H_i + (\lambda_p + \lambda_s)m_i + \lambda_c + \beta + Z, \tag{1.18}$$

where $\angle H_i$ is the true phase, m_i the subcarrier index, β the initial PLL phase offset, and Z AWGN noise. The terms λ_p, λ_s, and λ_c denote phase errors from:

$$\lambda_p = \frac{2\pi\Delta t}{N}, \tag{1.19}$$

$$\lambda_s = 2\pi\frac{T' - T}{T}\frac{T_s}{T_u}n, \tag{1.20}$$

$$\lambda_c = 2\pi\Delta f T_s n, \tag{1.21}$$

where Δt is the packet boundary detection delay, N the FFT size, T' and T the sampling periods at receiver and transmitter, T_u and T_s the useful and total OFDM symbol durations, n the sampling time offset, and Δf the carrier frequency offset. Since these parameters are unknown and vary between packets, the true phase $\angle H_i$ cannot be directly recovered.

However, the phase difference between two receiving antennas on the same NIC is much more stable, as they share the same clock and downconversion frequency, cancelling common errors:

$$\Delta\hat{H}_k = \Delta H_k + \Delta\beta + \Delta Z, \tag{1.22}$$

where ΔH_k is the true difference, $\Delta\beta$ a constant offset, and ΔZ differential noise.

Theorem 1.3.1. *The measured phase difference on subcarrier k between two receiver antennas is stable, and its mean and variation are given by*

$$\begin{cases} \mathbb{E}(\Delta\angle\hat{H}_k) = \mathbb{E}(\Delta\angle H_k) + \Delta\beta \\ Var(\Delta\angle\hat{H}_k) = Var(\Delta\angle H_k) + 2\sigma^2. \end{cases} \tag{1.23}$$

1.3.4 CSI Acquisition in OFDM

In IEEE 802.11, OFDM divides the spectrum into narrowband subcarriers for parallel transmission. The receiver estimates the CSI values of $\mathbf{H}_{(\mathbf{m},\mathbf{n})}$ given a-priori knowledge of the pre-defined signal $\mathbf{X}$ and the actually received signal $\mathbf{Y}$. Indeed, the Wi-Fi transmitter sends a Long Training Field (LTF) sequence in the packet preamble, which contains predefined symbols for each subcarrier. When LTFs are received, the Wi-Fi receiver estimates the CSI matrix using the received signal and the original LTFs. In addition, a pre-processing phase is required, which includes removing the cyclic prefix, demapping, and OFDM demodulation. Hence, the a priori-known bits sequences sent by the trasmitter are compared with those received on the receiver side. This allows to define $\tilde{\mathbf{H}}_{(\mathbf{m},\mathbf{n})}$ as an

approximation of the CSI matrix $\mathbf{H}_{(\mathbf{m},\mathbf{n})}$. Therefore, the model used for data analysis is:

$$\tilde{\mathbf{Y}} = \tilde{\mathbf{H}}_{(\mathbf{m},\mathbf{n})} \cdot \mathbf{X}. \tag{1.24}$$

1.4 Applications of Channel State Information

The fine-grained nature of channel state information (CSI) has made it a key enabler for a wide range of applications that extend well beyond the optimisation of traditional communication performance. Unlike coarse indicators such as the Received Signal Strength Indicator (RSSI), CSI provides amplitude and phase information for each subcarrier and antenna pair in a Multiple-Input Multiple-Output Orthogonal Frequency Division Multiplexing (MIMO-OFDM) system. This high-resolution view of the wireless channel enables the detection of subtle changes in the propagation environment, facilitating both device-based and device-free sensing applications.

In the field of communications, CSI is already employed for channel estimation, adaptive beamforming, link adaptation and interference mitigation. However, its potential is even more compelling in the context of wireless sensing, where variations in measured CSI are used to infer physical properties, spatial relationships and even human behaviour without the need for wearable sensors or direct contact.

One important application domain is material identification. Different materials interact with radio frequency signals in ways that depend on their electromagnetic properties, such as dielectric constant, conductivity, and loss tangent. These properties affect the signal's attenuation and phase delay, leaving characteristic patterns in the channel frequency response. By analysing these patterns in either the frequency domain or the time-domain channel impulse response, materials can be classified in scenarios ranging from industrial inspection to security screening.

Another prominent area is object localisation. The spatial resolution provided by CSI, especially when combined with MIMO antenna arrays, enables objects or people to be precisely positioned in indoor environments. Reflections from a target modify the multipath structure of the channel, and this information can be used to estimate distances, angles of arrival or trajectories. These capabilities support applications such as asset tracking, indoor navigation and human–robot interaction.

A closely related area is object differentiation, where the objective is to detect the presence of an object and distinguish it from other objects in the environment. Variations in size, shape, material composition and orientation each leave a unique imprint on CSI

measurements. By exploiting these subtle differences, systems can differentiate between multiple objects, even when they are close together. This is a particularly relevant feature in cluttered or dynamic environments.

CSI is also well suited to behaviour recognition. Human movement modifies the wireless channel through micro-Doppler shifts and changes in multipath propagation. Different activities, such as walking, sitting, gesturing or falling, produce characteristic temporal and spectral patterns in the CSI data. These patterns can be recognised by signal processing pipelines and machine learning models, enabling applications in smart homes, assisted living and health monitoring.

Finally, one of the most promising biomedical applications of CSI is breath monitoring. The small, periodic movements of the chest during respiration produce minute variations in the phase or amplitude of the received signal. By analysing phase differences between antennas, these fluctuations can be isolated from other sources of noise and movement. This enables the contactless estimation of breathing rate and patterns, even when the subject is not within the transmitter's or receiver's line of sight, opening up possibilities for unobtrusive health monitoring in both clinical and home settings.

This dissertation places particular emphasis on five application areas: material identification, object localisation, object differentiation, behaviour recognition and breath monitoring. Although these areas span different domains — from industrial inspection to healthcare — they share a common underlying principle: every object, material or activity leaves a unique electromagnetic signature on the wireless channel. Capturing CSI at high temporal and spectral resolution and applying robust signal processing and learning algorithms enables the extraction of discriminative features that facilitate reliable identification, differentiation and monitoring in real-world environments. This positions CSI as a foundational technology for the next generation of intelligent, pervasive and non-intrusive sensing systems.

CHAPTER 2

Previous work

This chapter presents a detailed review of the research landscape concerning the use of *Channel State Information* (CSI) across the main domains investigated in this doctoral work. In particular, it surveys the methodologies, experimental frameworks, and results reported in the literature for several key application areas: *material recognition*, where variations in CSI are analysed to infer the electromagnetic properties of different materials; *object localisation*, which exploits spatial characteristics of the channel to estimate the position of targets in indoor environments; *object differentiation*, where items located in similar positions are distinguished by their physical or electromagnetic signatures; *movement recognition*, which classifies human or object motion patterns by observing temporal fluctuations in CSI; and *respiration monitoring*, where minute, periodic changes in the wireless channel are used to estimate breathing rate and patterns.

The review discusses the theoretical principles underpinning each application, as well as the signal processing pipelines, machine learning algorithms, and deep learning architectures employed. Particular attention is given to the strengths and limitations of the approaches described in the literature, along with the open challenges that remain to be addressed.

2.1 Material Recognition

Within CSI-based wireless sensing, identifying the material composition of an object has attracted significant attention, due to applications ranging from industrial inspection to security screening. Material identification exploits the fact that electromagnetic waves interact with media in characteristic ways, producing amplitude and phase variations driven by dielectric properties, conductivity, and internal structure. Consequently, recent studies have focused on transforming raw CSI from commodity Wi-Fi devices into stable and discriminative features for reliable classification.

Among the most representative contributions, *WiMi* [6] and *WiMate* [7] address material recognition from complementary perspectives. Both adopt physics-driven modelling and multi-antenna CSI, but differ in the invariances they target, the preprocessing pipeline, and the experimental protocol. *WiMi* builds on the insight that, when a target is placed on the line-of-sight (LoS) path between transmitter and receiver, the material induces differentiable amplitude and phase changes. However, the CSI reported by commercial network interface controllers (NICs) is affected by hardware impairments and cannot be used directly without calibration.

WiMi. Phase stability is obtained by computing the inter-antenna phase difference between two co-located receive antennas (RX) sharing the same clock. This suppresses hardware-induced impairments, including packet boundary delay (PBD), sampling frequency offset (SFO), and carrier frequency offset (CFO), which would otherwise introduce unpredictable phase shifts. Residual multipath is mitigated by selecting subcarriers with low temporal variance, improving the robustness of the phase descriptor.

Amplitude stability is achieved through sequential denoising. First, statistical outliers are removed; then, impulsive noise is attenuated via wavelet-domain filtering. Finally, *WiMi* computes the amplitude ratio between the RX antennas, exploiting the fact that common-mode multipath fluctuations are largely cancelled by the ratio, resulting in a more repeatable amplitude feature.

Based on stabilised phase-difference and amplitude-ratio descriptors, *WiMi* defines a single scalar feature designed to depend primarily on the material attenuation and phase constants, thus reducing sensitivity to target size. The feature is classified using a support vector machine (SVM). Experiments conducted in three indoor environments (library, laboratory, and hall), with ten liquid targets, report an average recognition accuracy of about 96%.

Sensitivity analyses show that performance decreases with transmitter–receiver distance (from 98% at 1 m to 87% at 3 m). Accuracy saturates once more than 20 packets are collected and is largely insensitive to the container material (plastic versus glass) due to baseline subtraction. Conversely, a significant degradation is observed for containers with diameter around 3.2 cm, where diffraction effects violate the underlying penetration assumptions.

WiMate. *WiMate* formalises wave interaction with dissipative media and proposes a descriptor explicitly designed to be invariant to the target position along the LoS path.

Experiments on six solid materials (board, paperboard, nickel, wood chip, iron, and titanium) across five LoS positions (4–16 cm from the transmitter) report an average accuracy of 96.2%. Most positions achieve 100%, while the most challenging case at 16 cm remains close to 90%. Importantly, the feature remains stable across positions and clearly separates metallic from non-metallic materials.

Summary and limitations. Both *WiMi* and *WiMate* rely on differential multi-antenna processing and SVM classifiers operating on low-dimensional, physics-informed features. *WiMi* emphasises robustness across multiple rooms and investigates the impact of container size/material and TX–RX separation; *WiMate* targets position invariance under controlled geometry. Despite the strong performance, both approaches assume LoS propagation and single-material targets, and neither provides an explicit mechanism for adapting to new environments without retraining or reconfiguration.

In-baggage extensions. More recent work extends CSI-based material identification to the more challenging *in-baggage* scenario, where strong multipath, unpredictable object positioning, and environmental changes complicate inference. Representative examples include *Towards In-Baggage Suspicious Object Detection Using Commodity WiFi* [9] and *Environment-Independent In-Baggage Object Identification Using WiFi Signals* [8].

The 2018 study [9] proposes a two-stage pipeline. First, reconstructed complex CSI is used to detect potentially suspicious items (e.g., metals/liquids) versus benign materials (fibres/paper/plastics). Second, differential CSI responses discriminate metals from liquids based on scattering behaviour. The system further estimates liquid volume and metal dimensions by isolating reflected CSI components. Two measurement geometries are considered: a near-transmitter placement to emphasise refraction and reduce geometry influence, and a farther placement to enhance reflection features for size estimation. Preprocessing includes phase correction (unwrapping and linear detrending), impulsive-noise removal, and subcarrier selection via a k-nearest neighbours (KNN) ranking strategy. A six-month campaign (15 metal/liquid items, six bag types) reports $> 95\%$ detection accuracy and $\approx 90\%$ hazardous-category classification, with average errors of ≈ 16 ml (liquid volume) and ≈ 0.5 cm (metal dimensions). Notably, the method operates in NLoS and through obstructions and integrates a risk-assessment component, while still relying on handcrafted features and conventional ML.

The 2021 study [8] targets environment-independent performance with minimal retraining. It introduces two physically motivated features: a *polarisation feature* (ratio of

System	Scenario / Target	Feature basis	Shape handling	Adaptation / generalisation
WiMi (2019)	LoS liquids; size-invariant descriptor	Inter-RX phase difference + amplitude ratio; subcarrier selection; wavelet denoising	Limited (not primary focus)	None; multi-room evaluation (no domain adaptation)
WiMate (2021)	LoS solids; location-invariant descriptor	EM model with fitted constants from calibration data	Not a primary goal	None; validated in a single environment
In-baggage (2018)	Through-baggage metal/liquid detection + risk estimation	Reconstructed complex CSI; inter-RX difference; reflection-based features	Volume/size estimation from reflection	Limited; bag types included in training
Env.-independent in-baggage (2021)	Through-baggage multi-class; object ID within same material class	Polarisation feature + complex CSI difference; CNN feature learning	Shape cues via CSI difference	Adversarial domain adaptation; minimal per-class calibration

Table 2.1: *Comparison of CSI-based material identification systems [6, 7, 8, 9].*

received powers measured by orthogonally polarised antennas), and a *complex CSI difference* between receive antennas capturing shape-related cues. These features are processed by a convolutional neural network (CNN) and combined with adversarial domain adaptation, requiring only one representative object per material class in the new environment. Evaluations on 14 objects, four bag types, and three environments report 97% accuracy in same-environment testing and a 42% improvement under cross-environment conditions with minimal adaptation data. Compared to *WiMi*/*WiMate*, this system explicitly targets domain shifts and extends recognition to specific objects within the same material category.

2.2 Object Localisation

Recent progress in Wi-Fi sensing has shown that CSI can support object detection, movement tracking, and device-free localisation in indoor environments. This section reviews representative studies that address these goals from different methodological perspectives, highlighting their assumptions, feature design choices, and experimental outcomes.

Interleaved subcarrier selection (ISS) for object detection. *A CSI-based Object Detection Scheme using Interleaved Subcarrier Selection* proposes an *interleaved subcarrier selection* (ISS) strategy that retains only subcarriers exhibiting pronounced variance in response to environmental changes (typically associated with object presence or motion). The system uses distributed antennas to increase spatial diversity and trains an SVM on the resulting CSI features. Experiments show that ISS improves detection accuracy with respect to full-spectrum baselines, especially in scenarios where many subcarriers are noisy or weakly informative.

Distributed antennas and compressed CSI. *Performance Evaluation of Object Detection Utilising CSI in WLAN Systems with Distributed Antennas* provides a systematic analysis of design choices related to antenna deployment. Using ray-tracing simulations and the compressed CSI representation defined in IEEE 802.11ac, the authors investigate the impact of antenna placement on detection performance. To mitigate phase wrapping discontinuities, CSI phase is mapped to sine and cosine components before classification. Results indicate that careful antenna placement substantially improves accuracy and that low-resolution CSI (e.g., 4-bit or 6-bit quantisation) can still preserve sufficient information for reliable classification, which is relevant for commodity deployments.

Implicit movement detection from CSI magnitude. *Radio Frequency Signature-based Implicit Object Movement Detection* shifts the focus from static object presence to motion detection. The approach uses only CSI magnitude, avoiding the instability of raw phase and enabling lightweight implementations. Object motion is formulated as a binary hypothesis test based on the Euclidean distance between consecutive CSI frames. A USRP transmitter–receiver pair and a motorised linear stage emulate controlled object motion. The method detects movements of small objects (even below the RF wavelength) with high accuracy across different speeds and target sizes, supporting applications such as ambient-aware security and assisted living.

Device-free localisation with probabilistic modelling. *Passive Indoor Localisation Based on CSI and Naive Bayes Classification* addresses device-free user localisation by building a statistical model of CSI variations collected from distributed antennas. By relying on amplitude-only CSI features, the system avoids phase calibration overhead and achieves sub-metre localisation accuracy in both LoS and NLoS scenarios, demonstrating the potential of low-complexity probabilistic methods for real-time indoor positioning.

Across these studies, three recurring insights emerge: (i) distributed antenna configurations improve spatial diversity; (ii) subcarrier-level feature design is often decisive for robustness; and (iii) there is an inherent trade-off between pipeline complexity and deployment practicality.

System	Objective	Feature type	Classifier / decision rule
Interleaved subcarrier selection (ISS)	Object presence detection	Variance-selected CSI subcarriers	SVM
Distributed antennas (IEEE 802.11ac compressed CSI)	Impact of antenna configuration	Compressed CSI; phase mapped to sin/cos	SVM (RBF kernel)
RF signature movement detection	Implicit movement detection (small objects)	CSI magnitude; frame-to-frame distance	Binary hypothesis test
Passive indoor localisation	Device-free user positioning	CSI amplitude statistics	Naive Bayes classifier

Table 2.2: *Comparison of CSI-based object detection and localisation approaches.*

2.3 Object Differentiation

In addition to object detection and localisation, CSI-based sensing has increasingly focused on *object differentiation* and on improving feature stability in complex indoor environments. Two representative systems, *TagFree* and *Wi-Tar*, illustrate complementary directions: fully passive tag-free differentiation and robust feature design against multipath-induced fluctuations.

TagFree. *TagFree* adopts a completely passive approach, removing the need for RFID tags or active transmitters attached to target objects. The method exploits the natural

interaction between electromagnetic waves and object properties to extract *radiometric signatures* from CSI measurements. Amplitude- and phase-based descriptors are summarised through statistical features (e.g., mean, variance, skewness, and kurtosis) and used to train standard classifiers such as SVMs, Random Forests (RF), and KNN, enabling differentiation even when objects are static. Experimental results in indoor environments report accuracies above 95%, suggesting that passive RF signatures can uniquely capture material- and shape-related cues across different orientations and distances. This tag-free paradigm simplifies deployment and supports privacy-preserving sensing, as it does not require active tagging or continuous user interaction.

Wi-Tar. *Wi-Tar* addresses a key practical limitation of CSI-based sensing, namely feature instability caused by multipath propagation and environmental noise. To improve robustness, it introduces the *CSI Ratio*, defined as an amplitude ratio across selected subcarriers. By design, this ratio mitigates common-mode distortions, yielding features that are more stable than raw amplitude measurements. Wi-Tar further enriches the representation by combining the CSI Ratio with additional descriptors such as signal entropy and signal power, forming a compact and discriminative feature vector. Classification is performed using SVMs trained on adaptive subcarrier-based features. Reported evaluations show improved detection accuracy and reduced false positives with respect to conventional amplitude-only approaches, indicating that stability-oriented feature engineering can enhance performance in dynamic and cluttered indoor scenarios. As a result, Wi-Tar extends the applicability of Wi-Fi sensing toward realistic deployments where channel conditions are inherently time-varying.

Taken together, *TagFree* and *Wi-Tar* highlight two consistent trends in the literature: (i) the shift toward passive, tag-free paradigms for object differentiation, and (ii) the emphasis on noise-resistant CSI features that remain reliable under multipath and environmental variability. These contributions complement earlier LoS-focused material identification systems (e.g., *WiMi* and *WiMate*) by prioritising robustness and deployability rather than focusing exclusively on controlled geometries. Overall, they further confirm the versatility of CSI as a rich physical-layer observable for fine-grained object sensing in real-world indoor environments.

2.4 Behaviour Recognition

WiFi Channel State Information (CSI) has emerged as a mature and versatile sensing modality for device-free human behaviour recognition, offering a privacy-preserving alternative to camera-based and wearable systems. CSI provides per-subcarrier amplitude and phase measurements for OFDM links, enabling the capture of fine-grained multipath variations induced by human motion, gestures, and physiological activities. This integrated analysis revisits seven representative studies, extending the previous review with more detailed methodological descriptions, enriched narrative, and additional visual references.

All examined works are implemented on commodity IEEE 802.11n/ac hardware, typically the Intel 5300 NIC, operating over 20–40 MHz channels with between 30 and 114 subcarriers, and using antenna configurations ranging from 1×1 to 2×3. Sampling rates span from 500 Hz to 1.5 kHz, with carrier frequencies including 5.32 GHz, 5.745 GHz, and 5.808 GHz. In most cases, the transmitter and receiver are placed in fixed positions, and environmental diversity is introduced either through cross-room validation or by moving furniture and objects to vary multipath patterns.

2.4.1 Physics-aware feature extraction approaches.

The first family of methods derives interpretable features grounded in the physics of radio propagation, requiring minimal training data and offering robust generalisation to new environments.

DeFall [1], a fall-detection system, uses short-time Fourier transforms (STFT) to obtain Doppler spectrograms from raw CSI, estimating motion speed and acceleration profiles. These kinematic traces are matched against fall templates using Sparse Learning Normalised Dynamic Time Warping (SLN-DTW) and DTW Barycenter Averaging to account for temporal misalignments. The system maintains high Detection Rate (96% LoS, 94% NLoS) with low False Alarm Rate ($\approx$1.5%) over month-long evaluations.

PhaseBeat [15] extends physics-aware processing to vital sign monitoring. After rigorous phase sanitisation and antenna-pair phase differencing, micro-Doppler components corresponding to chest wall motion are isolated. These are mapped to respiration and heartbeat rates using band-pass filtering and peak detection in the time–frequency domain. The system achieves sub-breath-per-minute (0.23 brpm) and sub-beat-per-minute (0.67 bpm) mean absolute error under static subject conditions.

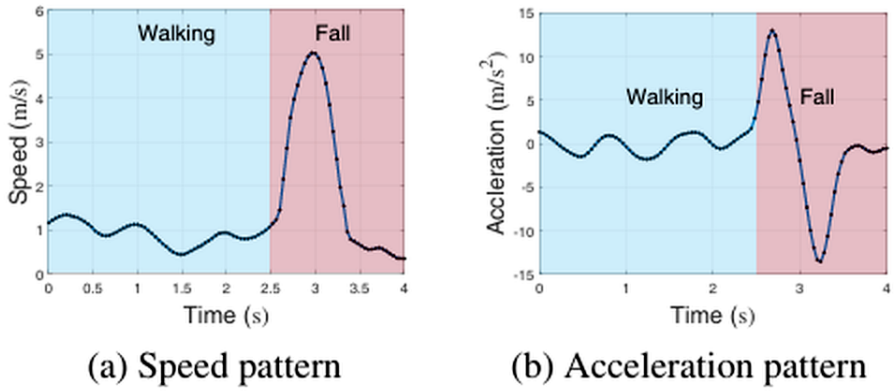

Figure 2.1: *Example of estimated speed and acceleration profiles for a walk–then–fall sequence derived from CSI Doppler features [1].*

While physics-aware designs offer strong interpretability and cross-environment robustness, they are typically task-specific; for example, DeFall [1] cannot directly be adapted to gesture recognition without redefinition of feature templates.

2.4.2 Visual surrogate deep learning approaches.

A second family converts CSI streams into visual representations—often subcarrier–time or subcarrier–frequency matrices—that can be processed by convolutional networks originally developed for image classification.

DTL-CBAM [2] applies a transfer learning strategy, mapping amplitude and phase data into 2D images. These images are processed by a ResNet-18 backbone enhanced with a Convolutional Block Attention Module (CBAM) that adaptively weights feature maps spatially and channel-wise. The network is fine-tuned on the target activity dataset using extensive augmentation to mitigate overfitting. The system reports accuracy $\geq$96.67% across three different indoor environments, demonstrating strong cross-domain capability.

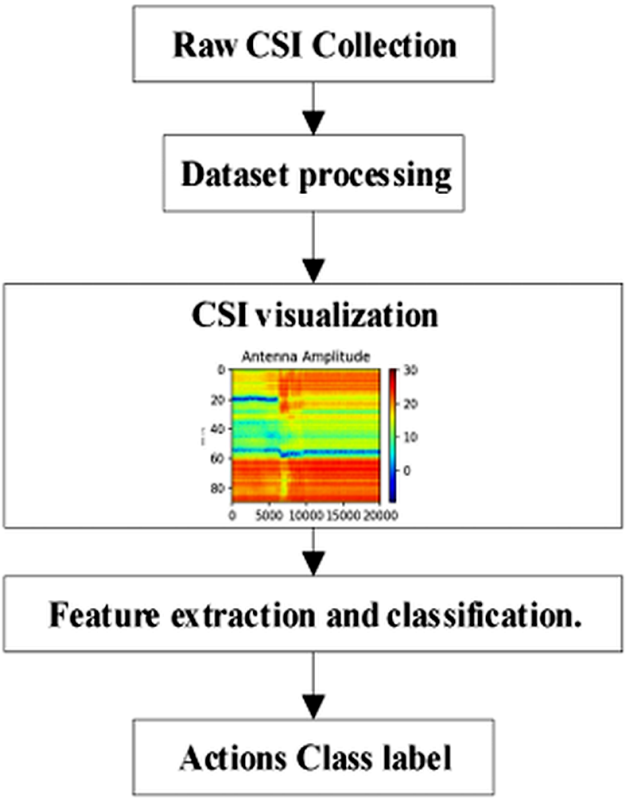

Figure 2.2: *DTL-CBAM processing pipeline. Raw CSI is processed and visualised as CSI amplitude maps, which are then used for feature extraction and classification via attention-enhanced convolutional networks [2].*

True_Detect [3] combines convolutional neural networks for spatial feature extraction with long short-term memory (LSTM) layers for temporal sequence modelling. CSI sequences are segmented into overlapping windows, converted into image-like tensors, and fed into the CNN to extract high-level spatial features. These are concatenated and passed to LSTMs to capture longer temporal dependencies. The architecture achieves ≈97.5% accuracy in cross-environment evaluations, with significant improvement in recognising activities with subtle temporal structure, such as “sit down” versus “stand up”.

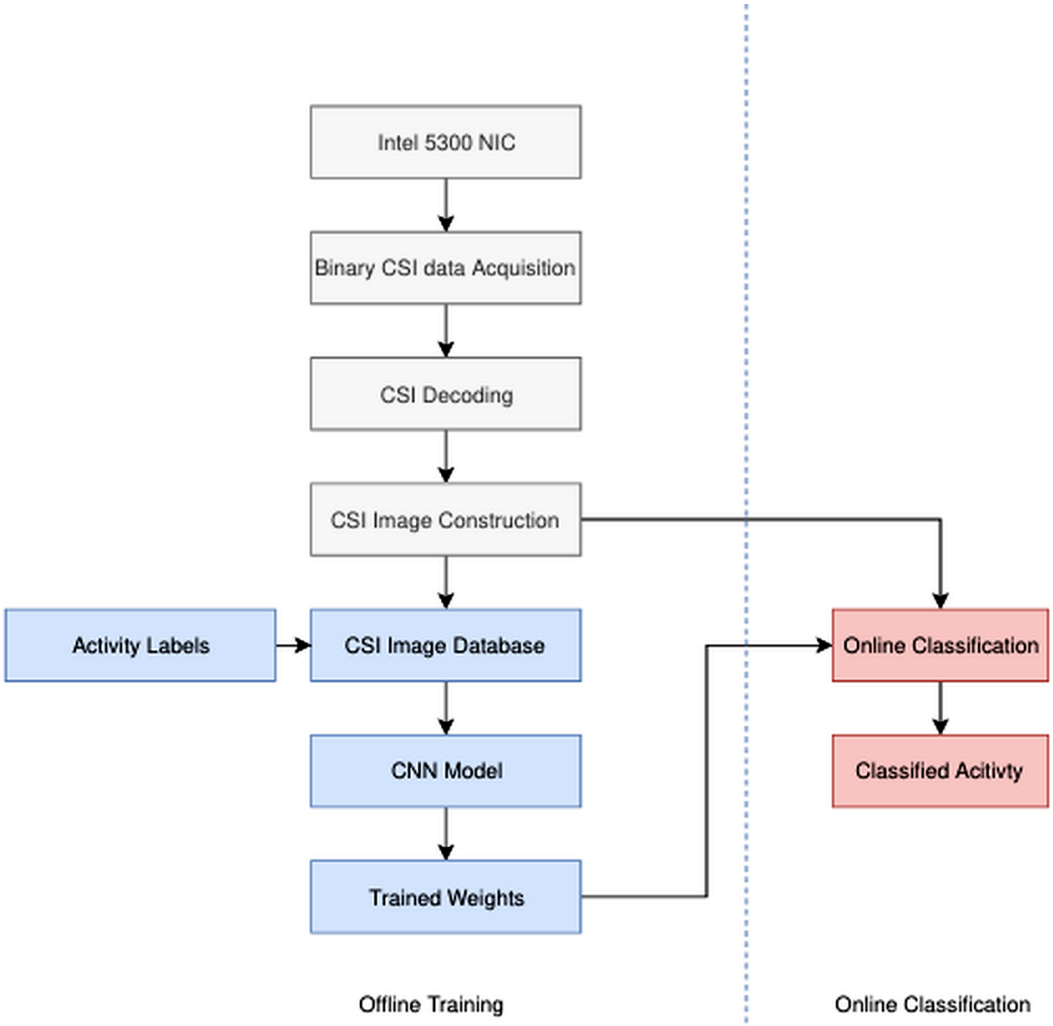

Figure 2.3: *True-Detect processing architecture. CSI data collected from an Intel 5300 NIC are decoded and converted into CSI images. Offline training produces a CNN-based model whose learned weights are then used for online activity classification [3].*

Finally, *HARN-ECNN* [16] focuses on CSI pre-processing. By applying amplitude denoising, phase calibration, and temporal smoothing, the system enhances signal quality before input to an ensemble of CNN classifiers. Each CNN is trained on a different view of the processed CSI (e.g., amplitude-only, phase-only, combined), and their predictions are aggregated. This approach achieves ≈98.12% accuracy on six-class HAR tasks, particularly excelling in activities with low motion intensity due to noise suppression in pre-processing.

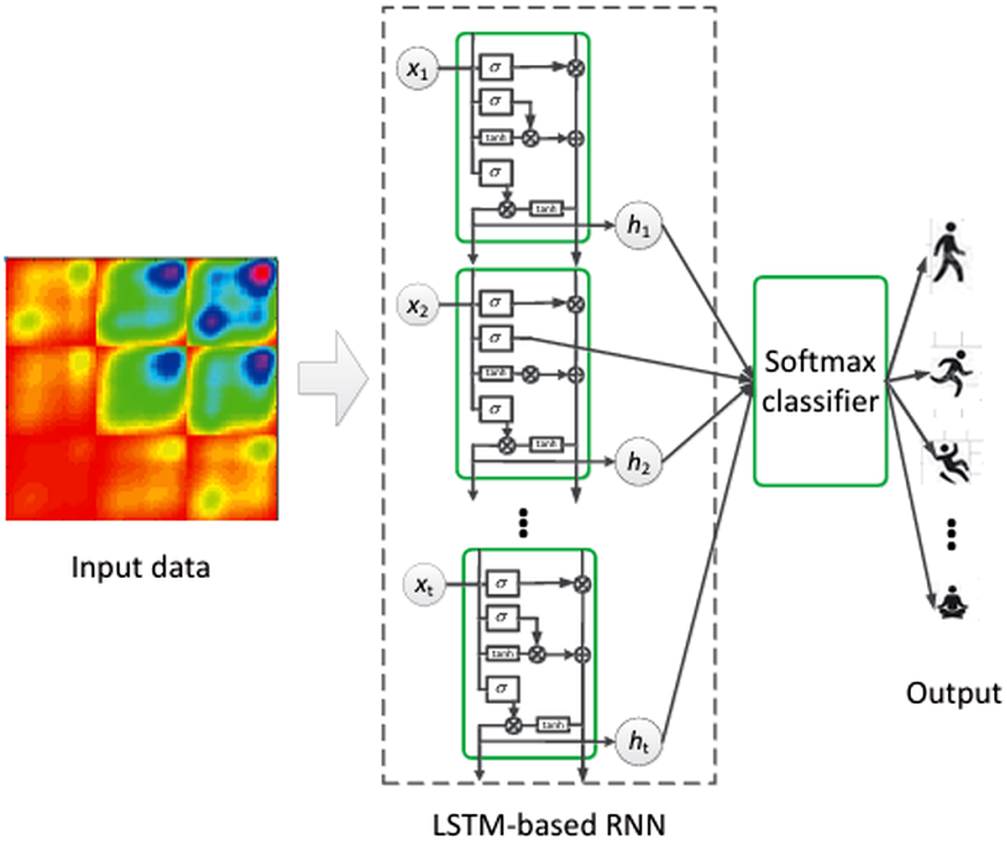

Figure 2.4: *LSTM-based recurrent neural network for CSI-based human activity recognition. CSI images are processed sequentially to extract temporal dependencies, and the resulting hidden states are classified via a softmax layer to infer the performed activity.*

2.4.3 Probabilistic temporal models.

The third family leverages statistical sequence modelling to explicitly represent the temporal evolution of human activities.

GMM–HMM models [4] each activity as a hidden Markov model where states correspond to latent motion stages (e.g., preparation, execution, completion). The emission probabilities are modelled as Gaussian mixtures extracted from phase-difference features, which are more stable across environmental changes than raw amplitude. State transitions encode temporal structure, and Viterbi decoding yields the most probable activity sequence. Average recognition accuracy exceeds 97% in laboratory conditions, with notable robustness to low-intensity activities such as “typing” or “drinking”.

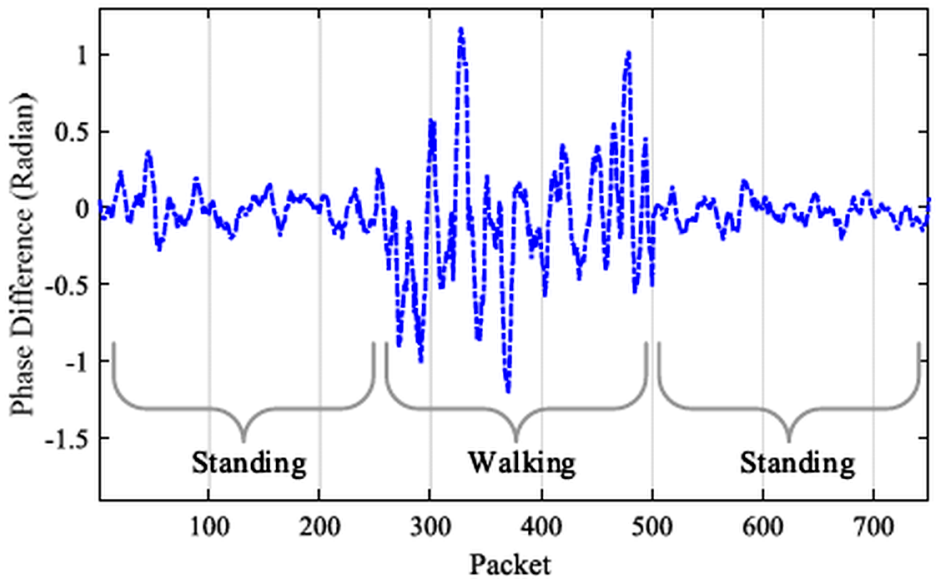

Figure 2.5: *Example of CSI phase-difference time series used for behaviour recognition. Distinct activity segments (standing and walking) induce characteristic temporal patterns that are modelled using GMM–HMM frameworks [4].*

2.4.4 Feature-based machine learning approaches.

An intermediate class between physics-aware and deep learning methods is exemplified by *HBR-CSI* [17], which employs engineered statistical features from CSI amplitude time series in conjunction with classical machine learning classifiers. The system preprocesses raw CSI to remove outliers and mitigate noise, followed by the extraction of time-domain statistics (mean, variance, standard deviation) and frequency-domain descriptors (spectral entropy, dominant frequency components). Feature vectors are then fed into algorithms such as Support Vector Machines (SVM), k-Nearest Neighbours (KNN), and Random Forests (RF), with model selection driven by cross-validation performance. Experiments cover six to eight activity classes, including walking, standing, sitting, lying, and transitional behaviours, recorded in controlled indoor environments using an Intel 5300 NIC at 5.32 GHz with 20 MHz bandwidth. Reported accuracies range from 95% to 97%, with relatively small variance across cross-user and cross-environment splits. The approach offers a good balance between computational cost and classification performance, making it well-suited to resource-constrained deployments. However, its reliance on hand-crafted

features may limit scalability to highly variable or unstructured activity sets without retraining.

Across the seven systems, representation choice strongly influences model design and resulting performance. Physics-aware approaches (DeFall, PhaseBeat [1], [15]) require careful signal calibration but offer the clearest interpretability and are less dependent on large labelled datasets. They excel in tasks where the target activity has a strong, distinctive kinematic or micro-Doppler signature.

Visual surrogate deep learning methods (DTL-CBAM, True-Detect, HARN-ECNN [2, 3, 16]) benefit from the maturity of CNN architectures, enabling scalable multi-class recognition. However, their dependence on the fidelity of CSI-to-image transformation means that noise, sampling artefacts, or misaligned subcarriers can significantly degrade accuracy. Pre-processing, as in HARN-ECNN, plays a crucial role in mitigating these risks.

Probabilistic models (GMM–HMM [4]) bridge interpretability and statistical learning, capturing temporal dynamics explicitly while allowing some resilience to environmental variation through feature engineering. Their performance is competitive with deep learning in small-class problems but may underperform in highly diverse activity sets unless combined with more expressive features.

When comparing quantitative results, HARN-ECNN [16] achieves the highest multi-class accuracy (≈98.12%), closely followed by True-Detect [3] and DTL-CBAM [2](≥96.67%–97.5%). DeFall achieves similar performance in binary fall detection but with the advantage of temporal stability over months without retraining. PhaseBeat [15] demonstrates the fine granularity achievable with CSI, with sub-unit errors in respiration and heart rate, albeit in static scenarios. GMM–HMM [4] remains competitive (>97%) with modest computational cost, making it suitable for low-power deployments.

Critical observations.

Generalisation vs. specificity: Physics-aware methods generalise well to new rooms but are often limited to a single well-defined task. Deep learning methods can handle a broader activity set but require domain adaptation.

Data requirements: Template matching and probabilistic models can operate effectively with fewer labelled samples, whereas deep learning requires larger datasets or transfer learning.

Computational complexity: GMM–HMM and template-based methods are lightweight and suitable for embedded devices. CNN-based pipelines are heavier, though model compression could mitigate this.

Robustness to noise: Pre-processing steps (phase calibration, amplitude denoising) are essential in all methods; their absence can lead to significant performance drops.

Method	Strengths	Weaknesses
DeFall	High DR/FAR stability over months; interpretable Doppler features; minimal retraining	Task-specific; not scalable to multi-class HAR without redesign
PhaseBeat	Fine-grained physiological sensing; sub-unit MAE; interpretable phase-based features	Requires static subjects; sensitive to phase calibration errors
DTL-CBAM	High multi-class accuracy; attention mechanism highlights discriminative CSI regions	Requires large labelled data or transfer learning; sensitive to CSI-to-image mapping
True_Detect	Combines spatial and temporal modelling; strong cross-environment performance	Computationally heavy; potential overfitting without augmentation
HARN-ECNN	Signal enhancement boosts low-motion activity recognition; ensemble improves robustness	Added pre-processing complexity; ensemble increases inference time
GMM–HMM	Lightweight; interpretable temporal structure; good with limited data	Lower performance in complex multi-class tasks; feature engineering required
HBR-CSI	Balanced accuracy across environments; simple implementation	Limited scalability; less robust to subtle activity differences

Table 2.3: *Strengths and weaknesses of the reviewed methods.*

In summary, the reviewed works represent complementary points in the design space for CSI-based behaviour recognition. Physics-aware approaches prioritise interpretability and robustness with minimal data, deep learning visual surrogates exploit the representational power of CNNs for rich activity sets, and probabilistic models offer a lightweight compromise. Selection of the optimal strategy depends on task specificity, available data, and deployment constraints, with pre-processing quality emerging as a decisive factor in real-world performance.

2.5 Respiration and Heartbeat Monitoring

The exploitation of commodity Wi-Fi infrastructure for contactless respiration and heartbeat monitoring has attracted significant attention due to its non-invasive nature, low cost, and ubiquity. Over the past decade, the field has shifted from coarse, power-level sensing to sophisticated channel-state-based methods capable of recovering sub-centimeter periodic motions. This evolution can be grouped into four methodological families:

1. **RSS-based methods**, which track variations in total received signal strength.

2. **CSI amplitude-based methods**, leveraging per-subcarrier magnitude changes.
3. **CSI phase-based methods**, exploiting relative or calibrated absolute phase.
4. **Hybrid and learning-based methods**, combining amplitude and phase features with data-driven inference.

Research on Wi-Fi-based vital sign sensing has progressed from small-scale demonstrations of breathing detection using raw signal strength to sophisticated platforms capable of reconstructing detailed respiratory waveforms, monitoring heart rate, and tracking multiple individuals simultaneously. The systems reviewed here illustrate how iterative innovation has moved from signal-processing heuristics to physics-based models, advanced calibration strategies, and deep learning pipelines, each step broadening the scope of what is possible with off-the-shelf hardware.

The first generation of work relied on Received Signal Strength (RSS), a coarse measure of signal power reported by every Wi-Fi device. Among these systems, UbiBreathe [18] stands out as a clear demonstration of the latent sensing power in basic RSS metrics. By continuously monitoring RSS over time and applying a sliding-window FFT to isolate oscillations in the 0.1–0.5 Hz band, the system reconstructs a respiration waveform while suppressing low-energy components and long-term drifts. Its design incorporates α-trimmed mean filtering to stabilize noisy readings and a convergence rule to decide when an RR estimate is reliable, while apnea detection is achieved through wavelet denoising and dynamic thresholds. The key innovation of UbiBreathe [18] is its fusion of measurements from multiple APs: majority voting across RSS streams significantly improves reliability and cuts false alarms, achieving apnea detection rates of over 96%. Extensive experiments demonstrate that this pipeline works well in varied environments, including offices and homes, with reliable operation at distances of 8 m through walls and up to 11 m in LoS conditions. The entire system runs on commodity smartphones and laptops, requiring no custom drivers, firmware changes, or CSI extraction. While its accuracy is impressive given the constraints, this approach has inherent limitations: the resolution of RSS is coarse, environmental changes can cause substantial noise, and the technique heavily depends on geometry and line-of-sight alignment. UbiBreathe [18] nonetheless represents a milestone: it proved that ubiquitous hardware and careful signal processing alone could deliver medically relevant respiratory monitoring, making it attractive for large-scale, low-cost deployments.

As this research matured, attention turned to Channel State Information (CSI), which exposes the full frequency response of each Wi-Fi OFDM subcarrier, providing per-subcarrier amplitude and phase data. CSI offers far more granular insight into multipath propagation, revealing subtle variations caused by the cyclic expansion and contraction of the chest during breathing. Early work exploiting CSI amplitude marked a leap forward in both accuracy and robustness. Wi-Sleep [19] was one of the pioneering systems to showcase this potential, focusing specifically on sleep monitoring. It captured breathing oscillations and posture changes overnight with low packet rates, using commodity Intel 5300 NICs with multiple antennas. Wi-Sleep's pipeline employed Hampel-based outlier removal, interpolation to address packet jitter, and wavelet filtering to preserve periodic components while suppressing noise, combined with subcarrier selection to isolate those with clear respiratory patterns. This research demonstrated that Wi-Fi signals, traditionally thought of only as a medium for data transmission, could rival specialized sensing devices such as pressure mats or camera-based systems, without the privacy concerns associated with video or the discomfort of wearable sensors.

Building on this insight, Liu et al. proposed Contactless Respiration Monitoring via Off-the-Shelf WiFi Devices [20], which expanded CSI amplitude analysis to include posture classification and apnea detection. This study introduced variance-based subcarrier selection and careful Channel Frequency Response (CFR) aggregation, achieving 85–94% accuracy in classifying posture states and 80–100% accuracy in detecting posture transitions. By systematically analyzing performance across sleeping positions, Liu's work provided one of the earliest comprehensive studies demonstrating that CSI amplitude alone is sufficient not just for breathing detection but also for clinically meaningful features like apnea recognition. These results highlight how amplitude-only CSI processing pipelines—simple to implement and requiring no calibration of phase—provide reliable respiration tracking, even at relatively low sampling rates.

While amplitude-based CSI is powerful, it remains sensitive to environmental factors. The next advance exploited the phase difference between antennas on the same device, an approach that cancels common-mode distortions from hardware oscillators and yields a far more stable breathing signal. The PhaseBeat system exemplifies this methodology [15]. Its pipeline begins with environment detection and calibration to account for static multipath, followed by subcarrier selection to emphasize stable measurements and wavelet analysis to resolve both respiration and heartbeat components. PhaseBeat [15] demonstrated that Wi-Fi sensing could deliver sub-bpm respiratory accuracy and reliably detect heartbeats with errors around 1 bpm, even through walls or in cluttered corridors. This achievement

is significant: heartbeat detection requires sensitivity to oscillations of less than one millimeter, which PhaseBeat accomplished with commodity Wi-Fi hardware.

TensorBeat [21] extended this methodology from single-user to multi-user monitoring. Rather than treating CSI as a set of time series, TensorBeat [21] reshapes CSI phase-difference measurements into a three-dimensional tensor indexed by time, antenna, and subcarrier. Applying Canonical Polyadic (CP) decomposition, the system separates breathing signals belonging to different individuals, then uses a stable roommate matching algorithm to pair tensor components with individuals. Respiration frequencies are extracted via simple peak detection on the separated signals. TensorBeat [21] demonstrated robust tracking of up to five users simultaneously, even in through-wall setups, marking one of the first scalable, multi-user Wi-Fi vital sign sensing platforms. This approach showed that commodity Wi-Fi could act as a "breathing radar," leveraging indoor multipath instead of treating it as a hindrance.

In parallel, BreathTrack [22] advanced the state of the art by addressing a key limitation of CSI: the instability of absolute phase on commodity hardware. Wi-Fi NICs introduce random phase errors due to oscillator drift, packet detection delay, and carrier and sampling frequency offsets, making raw phase nearly unusable. BreathTrack [22] solved this through a combination of hardware calibration—using coaxial cables and splitters to create a reference path from transmitter to receiver—and software correction of CFO, SFO, and PDD based on phase differences between a reference antenna and the sensing antennas. After calibration, BreathTrack [22] applied sparse recovery over a joint Angle-of-Arrival (AoA) and Time-of-Flight (ToF) space to identify and track the dominant reflection path from the subject's chest. Once isolated, the phase of this path provided a clean respiration waveform. BreathTrack [22] achieved median RR accuracy above 99% and produced faithful breathing waveforms suitable for tidal volume estimation and other advanced metrics. This work represents the upper bound of what commodity Wi-Fi can achieve with careful calibration and sophisticated signal processing, bridging the gap between academic experiments and clinical-grade measurement.

The development of machine learning pipelines, exemplified by BreatheSmart [23], further accelerated this field. Instead of manually designing subcarrier selection rules and handcrafted signal-processing pipelines, BreatheSmart [23] fused amplitude and phase features in a deep LSTM model trained to classify breathing patterns and estimate RR. However, its most impactful contribution lies in evaluation methodology: BreatheSmart [23] systematically studied the effects of sampling rates, attenuation levels, and environmental conditions, providing publicly available datasets and standardized metrics for

reproducibility. It reported 98.7% RR accuracy and 99.5% classification accuracy for breathing patterns, but more importantly, it established experimental practices that the community has increasingly adopted to ensure fair comparisons between systems.

The research community has also recognized that geometry profoundly influences Wi-Fi sensing. The study "Human Respiration Detection with Commodity Wi-Fi Devices: Do User Location and Body Orientation Matter?" [24] introduced a Fresnel-zone model for respiration detectability, showing that even small changes in transmitter and receiver placement can shift users between regions of high and low sensitivity. Experimental validation confirmed that respiration signal strength varies predictably with location and orientation, reframing multipath effects as a design parameter rather than a nuisance. Today, Fresnel-zone modeling informs placement strategies in multi-user scenarios and underpins systems like TinySense [25] and TensorBeat [21].

TinySense [25] offers a compelling example of how geometry-aware processing can simplify multi-user sensing. Instead of performing heavy tensor decomposition, it transforms CSI into the delay domain via IFFT, applies a Time-of-Arrival threshold to discard reflections from distant objects, and selects subcarriers with moderate multipath sensitivity. This lightweight design isolates signals from two nearby individuals, improving accuracy from about 70% in baseline methods to nearly 88%. With its modest hardware requirements—a single commodity router and Intel 5300 NIC—TinySense [25] illustrates a practical trade-off: with intelligent filtering, multi-user performance can be achieved without computationally expensive algorithms.

The trajectory of this field is now clear. Systems such as UbiBreathe [18] showed that breathing detection is possible with nothing more than clever RSS processing, while BreathTrack [22] demonstrated waveform-level fidelity through meticulous calibration and sparse recovery. In between, studies like Wi-Sleep, Contactless Respiration Monitoring and PhaseBeat [19, 20, 15] refined CSI amplitude and phase-difference processing, extracting clinically meaningful signals with commodity devices. TensorBeat [21] pushed these methods to multi-user scalability, while TinySense introduced lightweight alternatives for smaller deployments. Fresnel modeling transformed empirical "trial-and-error" placement strategies into principled design, and BreatheSmart [23] closed the loop by offering high-accuracy learning-based sensing with standardized evaluation protocols. Together, these studies mark the evolution of Wi-Fi vital sign monitoring into a mature, science-driven discipline.

Today's systems can be broadly characterized by their sensing methodology and trade-offs. RSS-based techniques are the simplest and most deployable, suitable for basic

home apnea screening but limited in precision. CSI amplitude pipelines are easy to implement and effective for applications like overnight monitoring and posture tracking. Phase-difference processing adds stability, enabling heartbeat monitoring and multi-user scalability. Absolute-phase calibration with sparse AoA/ToF recovery offers laboratory-grade fidelity and waveform reconstruction, albeit with heavier setup requirements. Machine learning now provides accuracy comparable to advanced pipelines while abstracting physical modeling, though interpretability remains a challenge. Across all approaches, geometry plays a decisive role: multipath, orientation, and Fresnel-zone effects can either obscure or amplify breathing signals, emphasizing that Wi-Fi sensing is a joint problem of signal processing and wireless propagation physics.

The progression of Wi-Fi sensing research reflects a deliberate convergence between signal processing, modeling, and machine learning. Rather than replacing one another, these paradigms complement each other: models like the Fresnel-zone analysis provide intuition for sensor placement and hardware design; calibration-based methods like BreathTrack [22] set benchmarks for physical fidelity; and learning-based systems like BreatheSmart [23] demonstrate scalability and robustness across environments. Together, they position Wi-Fi as a promising technology for unobtrusive, ubiquitous health monitoring, paving the way for clinical adoption and integration into smart homes, hospitals, and assisted living facilities.

System	Sensing Principle	Setup	Processing Highlights	Key Results
UbiBreathe [18]	RSS	Commodity APs, smartphones; multi-AP fusion	Sliding FFT; α-trimmed averaging; wavelet apnea detection	RR error <1 bpm; apnea >96%; 8–11 m range
Wi-Sleep [19]	CSI amplitude	Intel 5300 NIC; low packet rate; bedroom	Hampel outlier removal; wavelet filtering	Contactless sleep monitoring; posture/rollover detection
Contactless Resp. Monitoring [20]	CSI amplitude	Off-the-shelf Wi-Fi; CFR aggregation	Variance-based subcarrier selection; posture classification	85-94% posture accuracy; 80–100% posture change detection
PhaseBeat [15]	CSI phase difference	Multi-antenna RX; indoor & through-wall	Calibration; wavelet analysis; subcarrier selection	RR error ~0.25 bpm; HR error ~1 bpm
TensorBeat [21]	Phase diff. + tensor	Multi-antenna RX; CP decomposition	Tensor factorization; stable matching	Multi-user monitoring (up to 5)
BreathTrack [22]	Calibrated phase + AoA/ToF	Coax calibration; reference antenna; sparse solver	Hardware/software calibration; sparse dominant path	>99% RR accuracy; waveform recovery
TinySense [25]	Delay-domain ToA filtering	Router + Intel 5300; 500 pps	IFFT; ToA gating; subcarrier pruning	Two-user monitoring; ~88% accuracy
BreatheSmart [23]	ML on amp.+phase	COTS Wi-Fi; varied attenuation	LSTM; standardized datasets	98.7% RR accuracy; reproducible benchmark
Fresnel Study [24]	Propagation modeling	Intel 5300 5.24 GHz; spatial grid	Fresnel zone modeling; sensitivity maps	Placement/orientation sensitivity mapping

Table 2.4: *Comparison of representative Wi-Fi-based respiration and heartbeat monitoring systems.*

CHAPTER 3

Proposed contribution

3.1 Introduction

This chapter introduces the core research work of the thesis, which examines the use of *Channel State Information* (CSI) as a sensing resource for identifying objects, recognizing human motion patterns, and monitoring vital signs such as respiration. Modern Wi-Fi networks, now ubiquitous in homes, workplaces, and public environments, rely on Orthogonal Frequency Division Multiplexing (OFDM) and Multi-Input Multi-Output (MIMO) techniques. These approaches divide the radio channel into multiple narrowband subcarriers and use several antennas to improve data throughput and link robustness. As a by-product, they provide access to highly detailed measurements of the propagation channel: CSI captures the complex amplitude and phase response of each subcarrier–antenna pair, describing how signals travel, reflect, and scatter within the surrounding space.

Traditional indicators such as the Received Signal Strength Indicator (RSSI) condense all propagation effects into a single value, whereas CSI preserves the multidimensional structure of the channel. This fine-grained view makes CSI particularly sensitive to small variations in the physical environment. Over the past decade, this property has encouraged extensive research into reusing standard Wi-Fi deployments as cost-effective, non-intrusive sensing systems. By exposing multipath dynamics, CSI effectively transforms indoor spaces into sensing volumes, where objects, human motion, and even physiological processes leave detectable signatures on the wireless channel. Within this framework, the present chapter focuses on three closely related areas of application:

Object recognition and environmental sensing: CSI measurements can reveal the presence and even material properties of static objects by modeling their influence on multipath propagation.

Movement and activity recognition: Human motion introduces distinct Doppler shifts and time-varying multipath patterns. By analyzing temporal CSI dynamics,

it is possible to detect actions such as walking, running, falling, or gesturing, and even distinguish between multiple people.

Breathing and vital sign monitoring: The exceptional sensitivity of CSI to small-scale changes enables non-contact measurement of physiological activity. Respiratory motion, on the order of millimeters, are detectable through subcarrier-level phase and amplitude analysis. Advanced signal processing and machine learning pipelines can reconstruct breathing waveforms, estimate respiration rate, and identify apnea events in real time.

In this chapter, we see how CSI-based detection systems have been implemented in realistic environments. This work demonstrates that Wi-Fi infrastructure can serve as a detection backbone that preserves privacy and is cost-effective, transforming wireless networks into a dual-purpose platform for both communication and e-Health applications.

3.2 Methodology

This section will demonstrate how the system can collect and analyse CSI data, and the problems encountered when setting up the system. MATLAB was used to analyse the collected CSI data and extract features from it, resulting in a large amount of data from

The content of this chapter has been published in the following papers:

[1] I. Bisio, C. Fallani, C. Garibotto, A. Grattarola, F. Lavagetto, A. Sciarrone, S. Zappatore, M. Zerbino, "Performance Evaluation of Machine/Deep Learning-Based Object Recognition Techniques Leveraging Channel State Information Using a Real Testbed," in IEEE Access, vol. 12, pp. 98680-98692, 2024, doi: 10.1109/ACCESS.2024.3428612.

[2] I. Bisio, C. Fallani, C. Garibotto, H. Haleem, F. Lavagetto, M. Hamedani, A. Schenone, A. Sciarrone, M. Zerbino, "AI-Enabled Internet of Medical Things: Architectural Framework and Case Studies," in IEEE Internet of Things Magazine, vol. 8, no. 2, pp. 121-128, March 2025, doi: 10.1109/IOTM.001.2400150.

[3] I. Bisio, C. Fallani, C. Garibotto, F. Lavagetto, A. Sciarrone and M. Zerbino, "Analysis of CSI-based Human Activity Recognition for Contactless Patients Monitoring," GLOBECOM 2024 - 2024 IEEE Global Communications Conference, Cape Town, South Africa, 2024, pp. 438-443, doi:10.1109/GLOBECOM52923.2024.10901592.

[4] I. Bisio, C. Fallani, C. Garibotto, A. Grattarola, F. Lavagetto, A. Sciarrone, M. Zerbino, "AI-Driven Estimation of Breathing Frequency Through CSI Analysis," ICC 2025 - IEEE International Conference on Communications, Montreal, QC, Canada, 2025, pp. 934-939, doi: 10.1109/ICC52391.2025.11161454.

which a system was built that can differentiate objects, recognise human behaviours and monitor breathing.

3.2.1 The employed CSI tool

The software platform used to capture the CSI data is the Linux 802.11n CSI tool. This tool captures the signal from an Intel Wi-Fi 5300 NIC equipped with three antenna, using a custom-modified firmware and open-source Linux wireless drivers, as shown in the figure 3.1.

Figure 3.1: *The Intel Wireless Wi-Fi Link 5300. This 802.11n device has three transmit/receive antennas, operates on both 2.4GHz and 5GHz frequency bands, and supports up to three spatial streams for a maximum bitrate of 450Mbps [5].*

The CSI Tool extracts per-subcarrier channel measurements based on the IEEE 802.11n-2009 standard, where the Long Training Fields (LTFs) of each packet preamble are used to estimate the channel response.
In the default 20 MHz mode, the Intel 5300 NIC measures CSI on 30 subcarriers, evenly spaced across the 56 subcarriers of the channel, whereas in 40 MHz mode, measurements span 30 subcarriers over 114 carriers. Each CSI sample is represented as a complex number $\mathbf{H}_k = ||\mathbf{H}_k||e^{j\angle\mathbf{H}_k}$,, where $||\mathbf{H}_k||$ and $\angle\mathbf{H}_k$ denote the amplitude and phase of the k-th subcarrier. The CSI values are estimated by comparing predefined symbols in

the packet preamble with the received versions, providing a channel estimate for each transmit–receive antenna pair.
For each received packet, the CSI Tool outputs a 3D tensor $\tilde{\mathbf{H}}$ of dimensions $N_{\text{tx}} \times N_{\text{rx}} \times 30$, where N_{tx} and N_{rx} are the numbers of transmitting and receiving antennas, respectively. Each Channel State Information (CSI) sample is represented as a complex value, with its real and imaginary components encoded using 8-bit signed integers. Along with the CSI data, supplementary metadata, including packet timestamps and the received signal strength indicator (RSSI) for each antenna, are also stored. The acquisition tool further enables fine-grained configuration of capture parameters, such as:

Wi-Fi channel (frequency);

Channel bandwidth (20 MHz or 40 MHz);

Transmission rate and antenna configuration.

Captured data is logged in binary format and can be exported to MATLAB or Octave for offline processing.

In all experiments described in this chapter, we used the CSI Tool in its default configuration, mode on channel 64 (5.32 GHz center frequency). CSI was collected at packet rates up to 500 Hz to ensure high temporal resolution, which is crucial for movement recognition and vital sign estimation. Both amplitude and phase were extracted from the raw CSI; however, unless otherwise specified, only amplitude information was used for baseline classification tasks.

Throughout my PhD research, I have explored the potential of using Channel State Information (CSI) from off-the-shelf Wi-Fi devices for environmental sensing. The research has focused on three main areas: object recognition; the classification of various human activities, such as walking, running, sitting, sleeping, falling and standing up; and the contactless monitoring of breathing rate. The results so far have been encouraging, demonstrating that commodity Wi-Fi networks can be used for more than just communication, and can also be employed for accurate sensing. This has significant potential applications in health monitoring, security, and smart home systems. However, the experiments have also revealed some clear limitations, particularly when users are out of line of sight or when the environment is cluttered. These conditions create strong multipath effects that degrade performance and highlight the need for better channel control. For this reason, I have started investigating the use of Reconfigurable Intelligent Surfaces (RIS), a technology that shapes the propagation environment to improve link quality. Although

this is still an early part of the study, it suggests a promising direction for future work. Integrating RIS with CSI-based sensing could make these systems more robust and reliable, even in challenging indoor environments, representing a natural evolution of this research.

3.2.2 CSI modes operation

The Intel 5300 CSI Tool, which extends the functionality of the Intel WiFi Link 5300 NIC, supports four operational modes that define how packets are transmitted, captured, and analyzed. Each mode is tailored to different experimental or deployment needs:

Client Mode: In this configuration, the Intel 5300 NIC functions as a standard Wi-Fi client connected to an access point (AP). CSI is extracted from packets transmitted by the AP to the client, making this mode ideal for experiments where the sensing device operates passively within an existing wireless infrastructure. The client mode replicates a realistic network environment and is commonly used in deployments where no modification of the AP is possible.

Access Point (AP) Mode: In AP mode, the Intel 5300 NIC itself acts as the network's master access point, managing connections with one or more clients (stations). This mode is useful for controlled sensing scenarios because it centralizes network management, allowing precise control of transmission parameters (e.g., packet rate, bandwidth). Here, CSI is extracted from packets transmitted by connected clients, and the AP determines when and how packets are sent, enabling reproducible measurements.

IBSS (Ad-Hoc) Mode: The Independent Basic Service Set (IBSS), or ad-hoc mode, creates a peer-to-peer network without requiring a dedicated AP. Each station manages its own communication schedule, and packets are transmitted directly between devices. This configuration is particularly useful for laboratory experiments or field deployments where a central AP is unavailable or undesirable, enabling low-latency and flexible communication between devices.

Monitor Mode: In monitor mode, the NIC captures all Wi-Fi frames in the vicinity, including those not addressed to the device itself. This provides a "sniffer" capability for CSI acquisition, making it valuable for passive sensing and network diagnostics. Monitor mode also allows packet injection, though transmitting high-throughput (HT) frames often requires driver modifications. Unlike other modes, CSI packets in

monitor mode use fixed source and destination MAC addresses for the transmitter and receiver, simplifying offline analysis. However, hardware support for full monitor mode functionality varies by device.

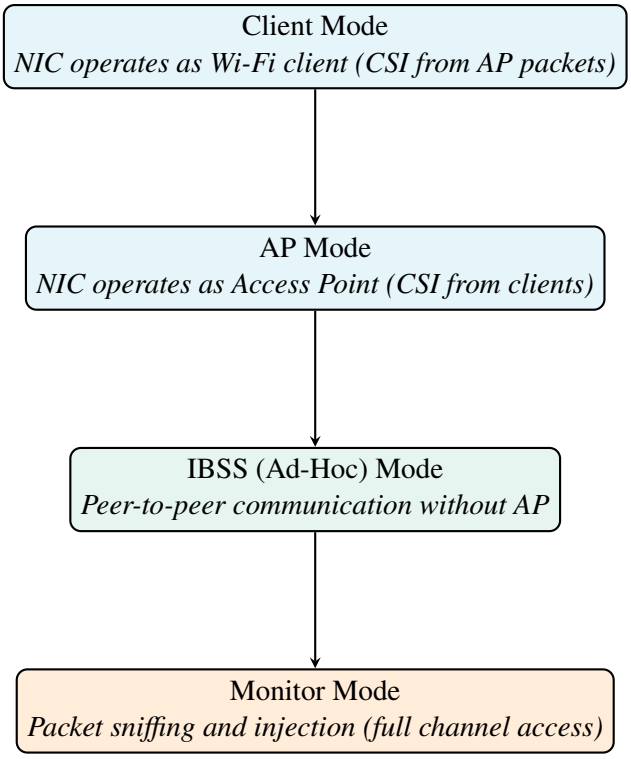

Figure 3.2: *Operational modes supported by the Intel 5300 CSI Tool. Client, AP, and IBSS modes provide different communication topologies for CSI acquisition within standard Wi-Fi networks, while Monitor mode enables full packet capture and fine-grained control of transmission parameters, making it particularly suitable for controlled sensing experiments.*

3.2.3 Monitor Mode Configuration

In this work, the Monitor mode is adopted as the standard communication mode between devices, as it allows precise control over various transmission settings such as packet rate, number of transmitted packets, number of active antennas, HT mode, and guard interval length (short or long). The setup procedure follows the step-by-step guide described in [26], starting from two computers with the CSI Tool installed.

Once one computer was assigned as the transmitter and the other as the receiver, the Lorcorn library was installed on both systems. Lorcorn is a widely used library that enables packet injection and parameter control for custom wireless transmissions. The default

transmitter and receiver scripts provided with the CSI Tool were then adapted to suit the needs of the experiments. Finally, two Bash scripts—one for each machine—were created to simplify and automate the process of launching the CSI Tool, applying the desired parameters, and starting data collection. This configuration ensured that experiments could be reproduced reliably and data acquisition could be carried out efficiently.

Once the PC's are turned on, first thing to do is to run the Bash script at the receiver side. At the transmitting side a bash script is run. Choosing an appropriate value for the monitor mode transmission rate is a critical step in ensuring that Channel State Information (CSI) can be correctly extracted. This is because only packets transmitted at High Throughput (HT) rates, as defined by the IEEE 802.11n standard, carry the necessary preamble fields from which the CSI can be computed. The transmission rate setting itself is encoded in a hexadecimal value whose bits specify various transmission parameters, including modulation scheme, coding rate, channel width, and whether short or long guard intervals are used. If a rate that does not correspond to an 802.11n HT setting is chosen, the receiver will not be able to decode the required training symbols and, as a result, will be unable to estimate the CSI for that packet.

For this thesis, the transmission rate used in monitor mode is `0x4101`, following the configuration recommended in [26]. This choice ensures compatibility with the CSI Tool and allows reproducible CSI measurements across experiments. The meaning of this hexadecimal value, which encodes the modulation and transmission parameters, is detailed in [26] and has been included here for reference.

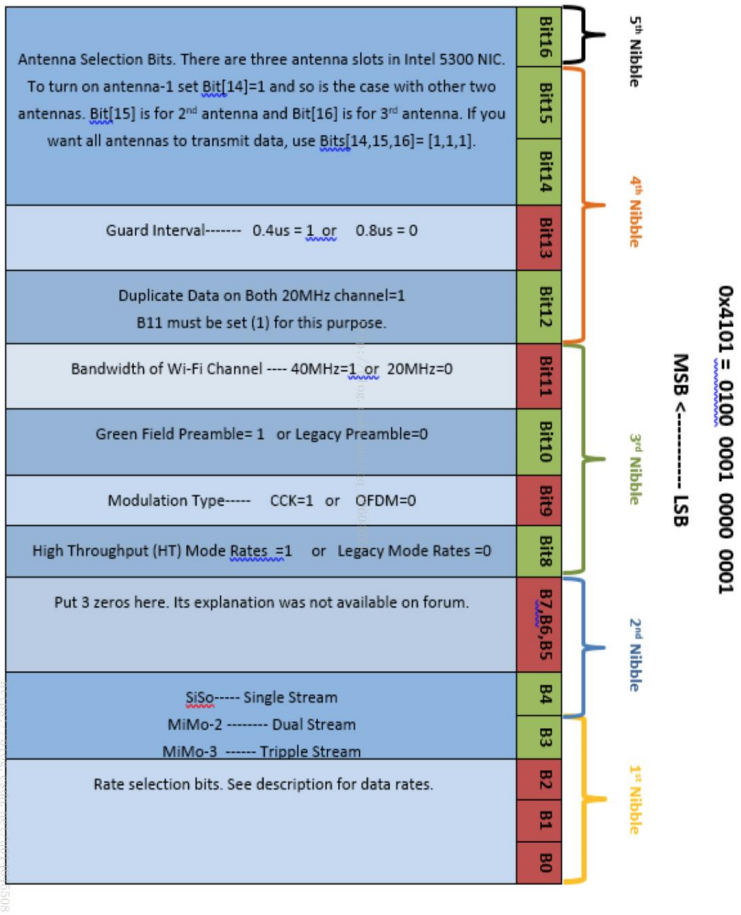

Figure 3.3: *Bit-level structure of the transmission rate mask used in monitor mode for the Intel 5300 CSI Tool. The hexadecimal value* `0x4101` *configures IEEE 802.11n High Throughput (HT) transmission with selected bandwidth, guard interval, antenna configuration, and modulation parameters required for CSI extraction.*

In 3.3 it can be seen that:

Bits 0-2: rate selection bits. Three bits are used to choose among 8 possible data rates.

Bits 3-4: stream type bits. Two bits are used to choose among three different stream types: SISO (one stream), MIMO-2 (two streams) and MIMO-3 (three streams).

Bits 5-7: all zeros bits.

Bit 8: Legacy Mode Rates=0 or High Throughput (HT) Mode Rates=1. According to this bit the values of data rates in the first three bits change.

Bit 9: modulation type bit. Complementary code keying = 1 or Othogonal frequency modulation = 0.

Bit 10: Preamble type bit. Green Field Preamble = 1 or Legacy Preamble = 0. With Legacy Preamble the frame transmission is understandable by all 802.11 stations including legacy 802.11a/b/g stations. On the other hand, the Green-Field mode frame format does not contain any legacy frame preamble and stations will not be capable of deciphering them.

Bit 11: Bandwidth of Wi-Fi channel bit: 40 MHz = 1 or 20 MHz = 0.

Bit 12: Duplicate data on both 20 MHz channel bit. Bit 11 must be = 1 if bit 12 = 1.

Bit 13: Guard interval bit: short guard interval (0.4 microseconds) = 1 or normal guard interval (0.8 microseconds) = 0.

Bits 14-16: Transmitting antenna selection bits: when antenna A is chosen to transmit the rate mask should be `0x04000`, when antenna B is chosen the rate mask should be `0x08000`, when antenna C is chosen the rate mask should be `0x10000`.

At this stage, the receiver side receives packets and collects CSI data.

3.2.4 Data Processing

To process and analyze the CSI data, MATLAB R2024b is installed on a workstation running Ubuntu 14.04 LTS, ensuring compatibility with the Intel 5300 CSI Tool. Once the CSI log files are collected from experiments, they are loaded into MATLAB for parsing and visualization. To correctly interpret the binary log files produced by the CSI Tool, the directory *linux-80211n-csitool-supplementary* downloaded during the tool's installation, is added to the MATLAB path together with all its subfolders. This supplementary package contains MATLAB scripts specifically designed to parse the raw CSI data, extract amplitude and phase information, and organize it into structured arrays suitable for analysis. After updating the MATLAB path, the CSI data could be imported directly into the workspace, where it is visualized and preprocessed prior to applying feature extraction and machine learning techniques. The code is written as follows: First, it parses the Intel 5300 CSI .dat capture file and discards invalid entries. Using the hardware timestamps,

it then estimates the effective sampling frequency, which is later used for time-axis and spectral analyses. The complex per-subcarrier CSI of each packet is stacked into a 4-D tensor $C \in \mathbb{C}^{T\times N_{\text{tx}}\times N_{\text{rx}}\times 30}$. Then it computes the amplitude in decibels $A = 20\log_{10}\big(|C|+\varepsilon\big)$ (with a small ε to avoid $\log(0)$) and, for every link and subcarrier, it evaluates the time mean μ and standard deviation σ. To prioritize stable carriers, it ranks subcarriers with score $= \mu - \alpha\sigma$, where α controls the stability penalty. It is possible to visualize mean and score via link×subcarrier heatmaps (rainbow colormap) and to provide per-link plots of mean and mean±std across the 30 subcarriers. To mitigate non-Gaussian noise and session-length effects it also compares alternative robust scores: a one-sided standard-error bound $\mu - z\,\sigma/\sqrt{N}$, the 5th percentile, and median $- k\,\text{MAD}$ (with $k \approx 1.4826$). Finally, it reports the Top-K subcarriers per link and rank links by their average score.

3.2.4.1 Step-by-step logic

1. **File selection and tool setup.** The script prompts the user to select a CSI `.dat` file and adds the Intel 5300 CSI Tool MATLAB helpers to the path so that `read_bf_file` and `get_scaled_csi` are available.

2. **Read & validate CSI packets.** `read_bf_file` parses the `.dat`. Empty cells are discarded and hardware timestamps (in μs) are collected for each valid packet. If no valid packets remain, the script aborts.

3. **Sampling frequency estimate.** Inter-packet spacings `dts` (in seconds) are computed from timestamps and their mean is used to estimate the sampling rate f_s. This is crucial for spectral analysis and for constructing time axes.

4. **CSI tensor assembly.** For each packet, `get_scaled_csi` returns a complex CSI matrix of size $[N_{\text{tx}} \times N_{\text{rx}} \times 30]$. Stacking all packets yields a 4-D tensor

$$C \in \mathbb{C}^{T\times N_{\text{tx}}\times N_{\text{rx}}\times 30},$$

   where T is the number of valid packets (time dimension) and 30 is the number of reported subcarriers for the Intel 5300.

5. **Convert to amplitude in dB.** Amplitude is computed as

$$A = 20\log_{10}(|C| + \varepsilon),$$

with a small ε to avoid $\log(0)$. The result is a time series in dB for every (TX, RX, subcarrier).

6. **Time statistics per link and subcarrier.** For each [TX, RX, subcarrier], the script computes the time mean and standard deviation. Links are then flattened so that $[N_{\text{tx}} \times N_{\text{rx}}]$ becomes $L = N_{\text{tx}}N_{\text{rx}}$ rows, enabling link×subcarrier heatmaps: `mean_link_sc` $\in \mathbb{R}^{L\times 30}$ and `std_link_sc` $\in \mathbb{R}^{L\times 30}$.

7. **Stability-aware ranking.** The main score is

$$\texttt{score\_link\_sc} = \texttt{mean\_link\_sc} - \alpha_{\text{Std}} \cdot \texttt{std\_link\_sc},$$

prioritizing subcarriers that are strong (high mean in dB) and stable (low std). The weight α_{Std} controls the variability penalty.

8. **Alternative robust scores.** To handle non-Gaussian noise and different record lengths, three robust scores are computed in addition to the primary one:

SE bound (one-sided LCB): $\mu - z\,\sigma/\sqrt{N}$ (length-aware; depends on sample count N and chosen z).

5th percentile: robust lower-tail indicator (shape-agnostic).

Median $-k$·MAD: robust to outliers; with $k \approx 1.4826$ (MAD $\approx \sigma$ under Gaussianity).

9. **Rankings and console output.** For each score, the script prints the Top-K subcarriers per link and the best links overall (mean score across subcarriers).

3.2.4.2 Variables and parameters visible in the MATLAB Workspace

`fn`, `fp`, `fullPath` Selected filename, folder, and full path to the `.dat` file.

`addpath(...)` Enables `read_bf_file` and `get_scaled_csi`.

`alphaStd` Weight for the std penalty in mean$-\alpha\cdot$std. Larger values emphasize stability; smaller values prioritize raw amplitude.

`topK_sc` Number of subcarriers to list per link in the Top-K ranking.

`plot_time_heatmap` Whether to draw the time × subcarrier heatmap for the best link.

raw Cell array of decoded packets from `read_bf_file`.

valid Filtered list of non-empty packets.

ts Hardware timestamps `timestamp_low` (in μs) for each valid packet.

dts Differences of consecutive timestamps converted to seconds.

fs Estimated sampling frequency (Hz), $f_s = 1/\texttt{mean(dts)}$.

csi1 CSI of the first valid packet; used to read tensor dimensions.

$N_{\mathbf{tx}}, N_{\mathbf{rx}}, N_{\mathbf{sc}}$ Number of TX antennas, RX antennas, and subcarriers.

T Number of valid packets (time length).

L Number of links, $L = N_{\text{tx}} \cdot N_{\text{rx}}$.

$C \in \mathbb{C}^{T \times N_{\mathbf{tx}} \times N_{\mathbf{rx}} \times 30}$ Complex CSI (voltage) tensor over time.

$A \in \mathbb{R}^{T \times N_{\mathbf{tx}} \times N_{\mathbf{rx}} \times 30}$ Amplitude in dB ($20 \log_{10} |C|$).

meanAmp $[N_{\text{tx}} \times N_{\text{rx}} \times 30]$ Time mean of amplitude per link and subcarrier.

stdAmp $[N_{\text{tx}} \times N_{\text{rx}} \times 30]$ Time standard deviation of A.

mean_link_sc $(L \times 30)$ `meanAmp` reshaped to link × subcarrier.

std_link_sc $(L \times 30)$ `stdAmp` reshaped likewise.

labels_link $(L \times 1)$ String labels for each link row.

score_link_sc $(L \times 30)$ Stability-aware score in dB units: `mean_link_sc` $- \alpha_{\text{Std}} \cdot$ `std_link_sc`.

order_link Indices of links sorted by average score (descending).

best_link Index of the highest-scoring link.

tx_sel, rx_sel TX and RX indices corresponding to `best_link`.

Amp_link $(T \times 30)$ Time × subcarrier amplitude (dB) for the best link.

tvec $(T \times 1)$ Time vector in seconds for plotting $(0, \ldots, (T-1)/f_s)$.

z_quantile = 1.64 One-sided ~ 95% quantile used in the SE bound.

k_mad = 1.4826 Scaling so that MAD $\approx \sigma$ under Gaussianity.

SE_link_sc Standard error $\sigma/\sqrt{N}$.

score_SE_link_sc $\mu - z\,\mathrm{SE}$.

perc5_link_sc 5th percentile over time (dB) for each link, subcarrier.

score_P5_link_sc Same as `perc5_link_sc`.

median_link_sc Time median (dB).

mad_link_sc MAD (dB) scaled by `k_mad`.

score_MAD_link_sc median $- k \cdot$ MAD (higher $\Rightarrow$ strong and robust).

Sall Cell array holding the four score matrices for plotting.

metrics Names of the four metrics shown in the console report.

SCORES Cell array of the four score matrices aligned with `metrics`.

It is important to note the following: in IEEE 802.11n transmissions over a 20 MHz channel, 52 subcarriers are used to convey data, 4 act as pilots, and the remaining 8 are null or guard tones, including the DC carrier at the center frequency. However, the *Intel 5300* network interface card, when used with the publicly available *Linux 802.11n CSI Tool*, does not expose the full set of 52+4 subcarriers. Instead, the firmware and driver are hard-coded to report only 30 evenly spaced subcarriers, selected deterministically to reduce the amount of CSI data while maintaining a representative sampling of the channel frequency response. Importantly, these 30 subcarriers are fixed and not randomly chosen at run time; their indices remain constant across different captures and correspond to a known subset of the full OFDM grid (as documented in the CSI tool source code).

Because the CSI measurements and all subsequent signal processing depend on the quality and stability of these reported subcarriers, we first performed baseline measurements in an empty environment using the MIMO configuration (3×3, three transmit and three receive antennas). This preliminary step allowed us to evaluate which of the 30 reported subcarriers provide the highest and most stable signal amplitude, minimizing temporal fluctuations and noise. The analysis was conducted by computing, for each

TX→RX link and subcarrier, the average amplitude μ and its variability σ over time. A stability-aware score was then defined as:

$$\text{score} = \mu - \alpha\,\sigma,$$

where α is a tunable parameter controlling the penalty applied to unstable subcarriers (large σ). Alternative robust metrics (e.g., lower confidence bounds $\mu - z \cdot \sigma/\sqrt{N}$, the 5^{th} percentile, and median$-k \cdot$ MAD) were also considered to mitigate the impact of outliers and non-Gaussian noise.

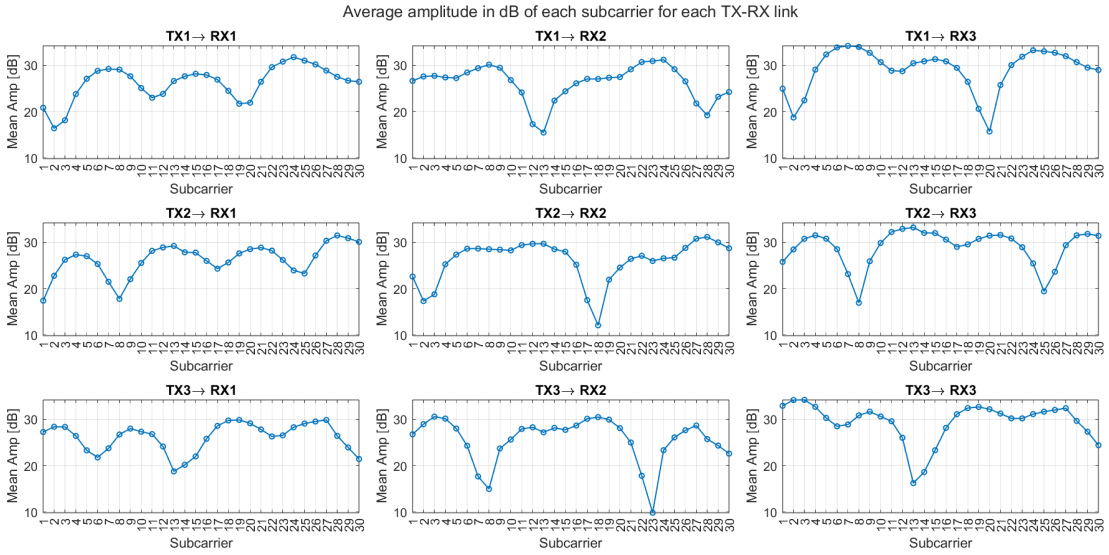

Figure 3.4: *Average CSI amplitude (in dB) for each OFDM subcarrier and each transmit–receive (TX→RX) antenna pair in an empty indoor environment under line-of-sight (LoS) conditions. The curves highlight the frequency-dependent channel response and the variability across MIMO links, providing a baseline for identifying stable and informative subcarriers used in subsequent sensing experiments.*

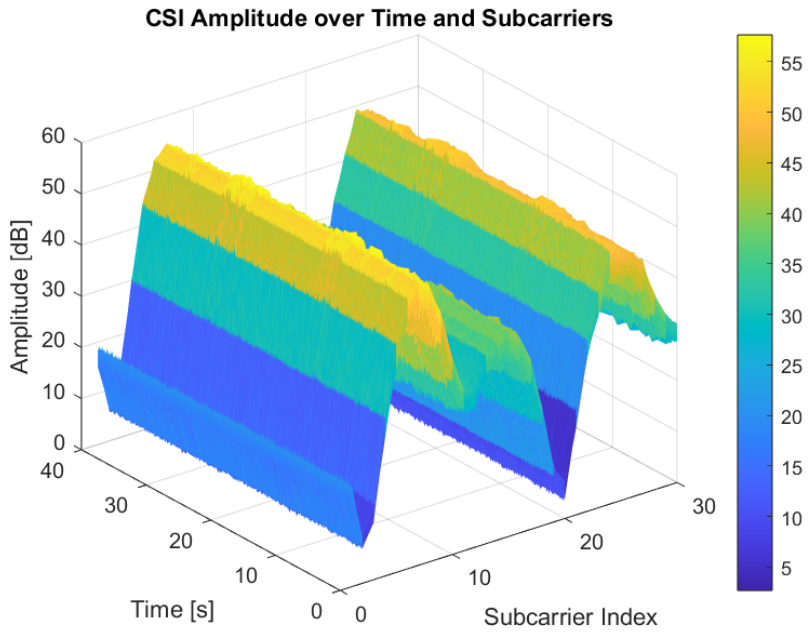

Figure 3.5: *Average CSI amplitude (in dB) for each OFDM subcarrier and each transmit–receive (TX→RX) antenna pair in an empty indoor environment under line-of-sight (LoS) conditions. The curves highlight the frequency-dependent channel response and the variability across MIMO links, providing a baseline for identifying stable and informative subcarriers used in subsequent sensing experiments.*

Figures 3.5 shows the baseline Channel State Information (CSI) measurements collected in an empty environment using a MIMO configuration with no objects or people present in the Line-of-Sight (LoS) path. These measurements are essential for understanding the natural stability and hardware-induced variations of the channel before performing breathing, object or activity recognition experiments.

The absence of movement produces a channel that is highly stable and coherent over time: the traces remain flat, without the periodic variations typically caused by respiration or body motion. This temporal stability provides a reliable baseline for selecting the most robust subcarriers to be used in subsequent sensing experiments.

These baseline measurements are crucial for:

1. Identifying the most stable and reliable subcarriers to be used in further experiments.

2. Quantifying the natural background noise, which should be compensated during preprocessing to isolate the low-frequency components related to breathing or human activity.

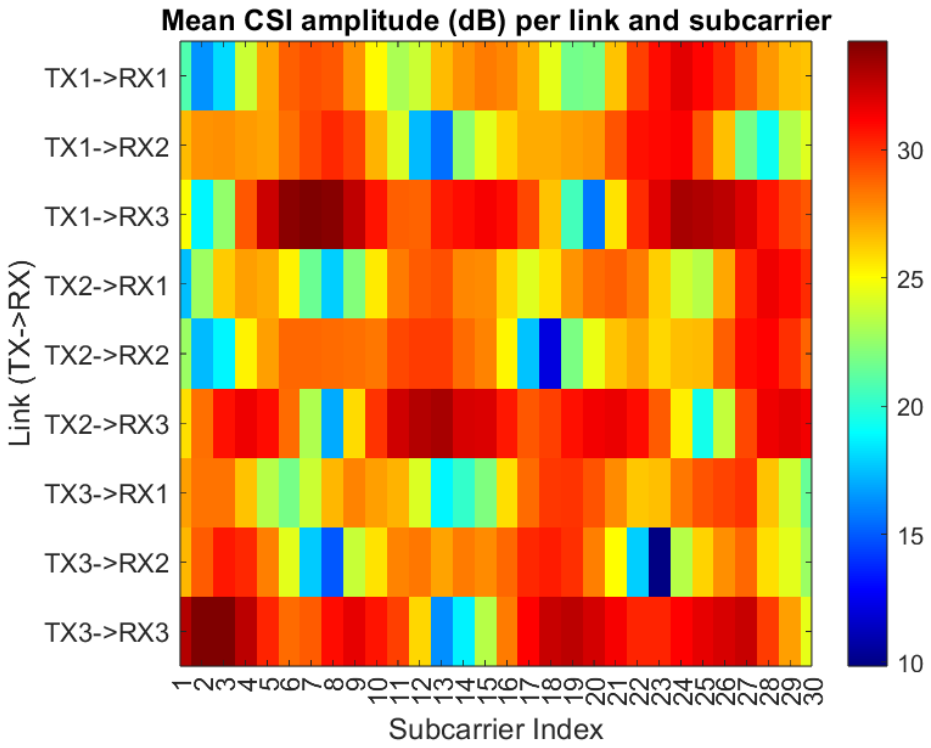

Figure 3.6: *Heatmap of the mean CSI amplitude (in dB) computed for each transmit–receive (TX→RX) link across the 30 available subcarriers.*

Figure 3.6 shows a heatmap of the received signal amplitude for each MIMO transmission link between a transmitting and a receiving antenna. The columns represent the subcarrier indices reported by the Intel 5300 NIC. The colour scale shows the average received signal amplitude in decibels: warmer tones (red) show higher power levels, while cooler tones (blue) show weaker channel responses. This visualisation clearly shows the spatial–frequency structure of the channel and helps to select the most stable and energetically significant subcarriers across all links. Consequently, all subsequent analyses and recognition tasks presented in this dissertation use the identified subcarriers, thereby ensuring that the input CSI data originates from the most stable and representative portion of the channel frequency response.

CHAPTER 4

Experimental Results

4.1 Setup

All measurements were carried out in an indoor laboratory environment using two PCs equipped with Intel WiFi Link 5300 network interface cards (NICs) running the *Intel 5300 CSI Tool*. Each device was fitted with three external antennas, enabling MIMO (3×3) operation. The transmitter and receiver were configured to send and capture IEEE 802.11n High Throughput (HT) packets on channel 64 (5.32 GHz) to ensure the availability of fine-grained Channel State Information (CSI). Two different physical arrangements were adopted depending on the sensing task:

Material and Object Recognition: The transmitter and receiver PCs were placed 2 m apart in line-of-sight (LoS) conditions. The target object was positioned midway between the devices so that the electromagnetic waves directly interacted with the material under test. This compact configuration enabled precise analysis of how different materials and object categories affect the Wi-Fi channel, while keeping multipath effects controlled.

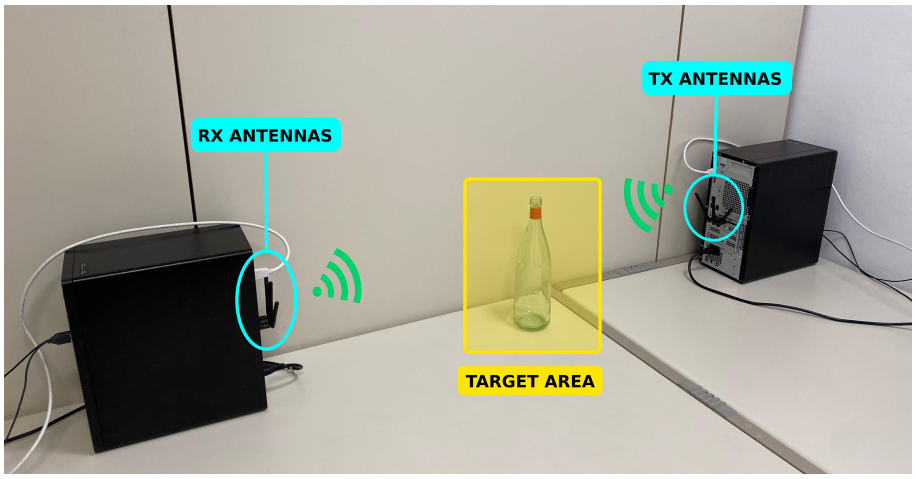

Figure 4.1: small *Experimental setup, considering a glass bottle as sample target.*

Movement and Breathing Recognition: The configuration shown in Figure 4.2 was used. The two PCs were positioned at opposite ends of the room, approximately 5.80 m apart along the main LoS path, with a room width of 3.20 m. This longer propagation path was chosen to better capture the impact of human movements and the subtle chest displacements caused by breathing. The extended LoS link increases the channel sensitivity to small motion variations.

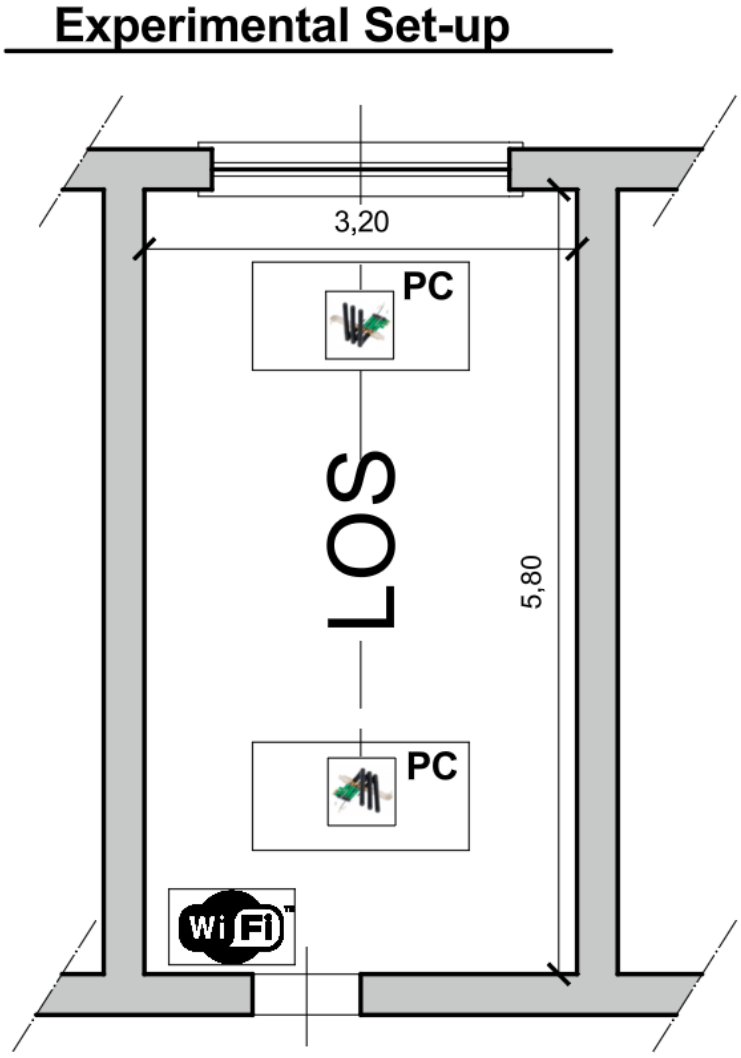

Figure 4.2: *Experimental set-up used for CSI data collection. The two PCs are equipped with Intel 5300 NICs and configured according to the task under analysis.*

This dual-configuration approach ensured that the experimental campaign could accurately evaluate Wi-Fi CSI sensing under both object-focused and human-centric scenarios, adapting the spatial layout to the specific detection task.

4.2 The CSI-based object recognition approach

The process begins with raw CSI collection: the transmitter sends packets and the receiver extracts the complex frequency response for each subcarrier. This rich channel information is highly sensitive to the presence, material, and motion of nearby objects, forming the basis of all subsequent sensing tasks. Raw CSI readings are often noisy and unstable due to hardware imperfections, interference, and multipath propagation. To mitigate these effects, we employ an ad-hoc pre-processing pipeline, as shown in Figure 4.3 4.4 :

Outlier removal: abrupt spurious spikes caused by packet loss or hardware glitches are detected and removed by appling an Hampel filter.

Noise suppression: a median filter is applied to reduce random fluctuations while preserving correlated channel dynamics.

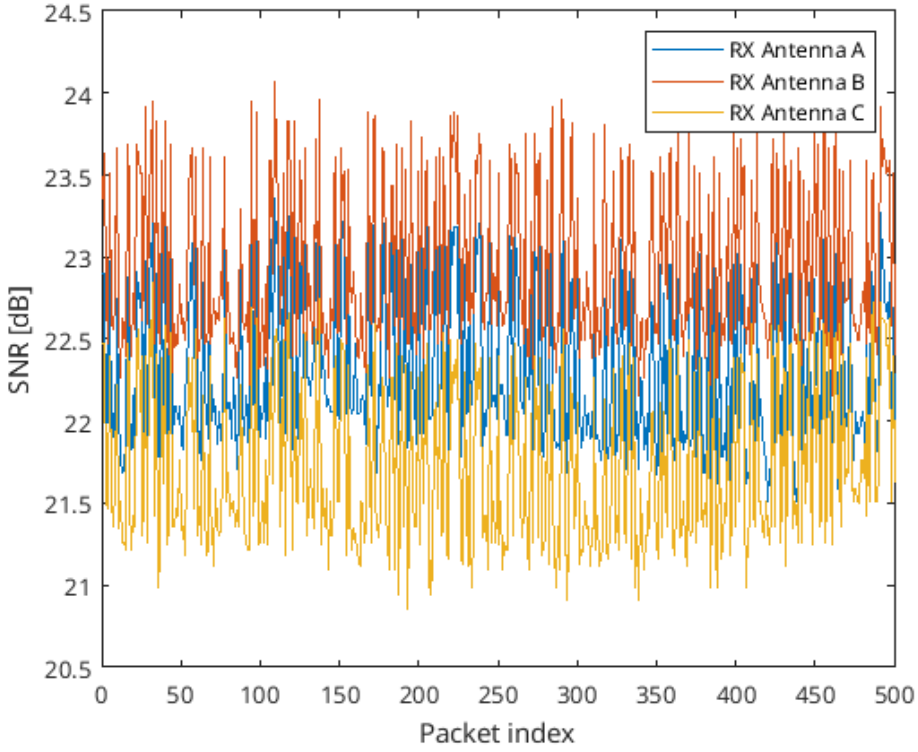

Figure 4.3: *SNR vs. Packet Index before outlier and noise removal.*

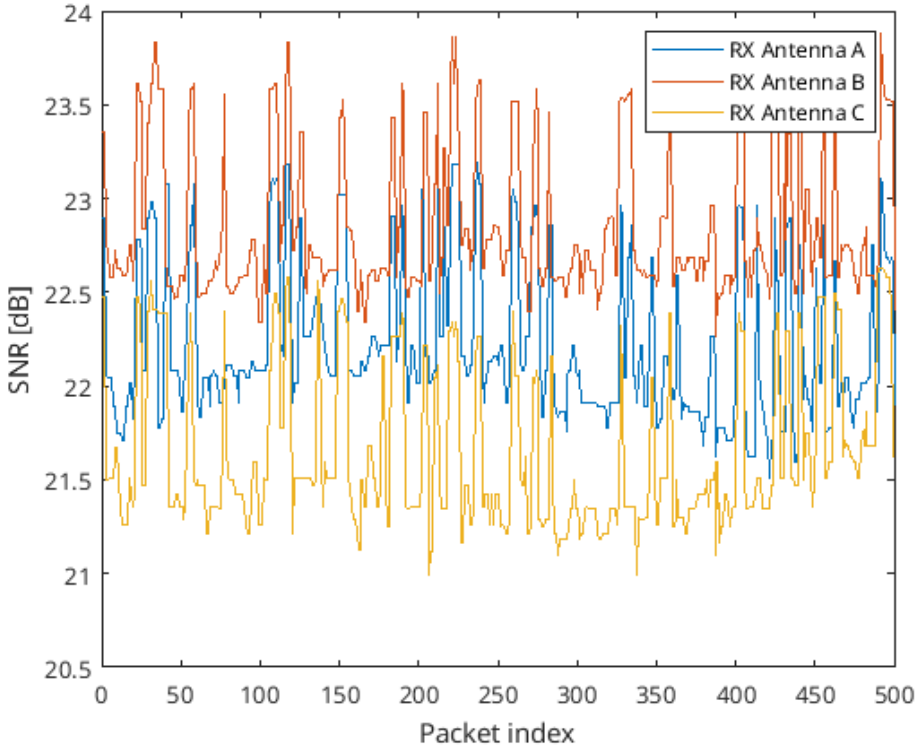

Figure 4.4: *SNR vs. Packet Index after outlier and noise removal.*

This cleaning step is essential to stabilize the CSI before feature extraction.

Feature extraction For the machine-learning–based approach, the cleaned CSI amplitude is transformed into a compact set of statistical descriptors summarizing its temporal behavior. The features include:

Interquartile Range (IQR):

$$\text{IQR} = Q_3 - Q_1 \tag{4.1}$$

is the difference between the third (75th percentile) and first (25th percentile) quartiles. These are often abbreviated to Q3 and Q1,respectively. The IQR is used to represent the middle (50%) spread of the data.

Variance:

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2 \tag{4.2}$$

quantifies signal fluctuation power. In the formula above, μ represents the mean of the data points, x_i is the value of an individual data point, and N is the total number of data points. It is important to note that because the formula works with the square of the deviations, variance is always going to be a positive number, or zero. If the variance is zero, then all entries likely have the same value. Similarly, a large variance infers that the numbers in the set are far from the mean and each other.

Median: the median formula differs depending on whether the number of observations N is odd or even. It represents the middle value in an ordered dataset:

$$\text{median}(x) = \begin{cases} x_{(\frac{N+1}{2})}, & \text{if } N \text{ is odd,} \\ \dfrac{x_{(\frac{N}{2})} + x_{(\frac{N}{2}+1)}}{2}, & \text{if } N \text{ is even.} \end{cases} \tag{4.3}$$

Range: the range of a dataset is the difference between the largest and smallest values in that dataset.

$$R = \max(x) - \min(x) \tag{4.4}$$

where $\max(x)$ and $\min(x)$ are the higher and lower observed value, respectively.

Skewness: degree of distortion from the symmetrical bell curve or the normal distribution. It measures the lack of symmetry in data distribution. It differentiates extreme values in one versus the other tail. A symmetrical distribution will have a skewness of 0.

$$\beta_1 = \frac{1}{N} \sum_{i=1}^{N} \frac{(x_i - \mu)^3}{\sigma^3} \tag{4.5}$$

where N is the number of data in the data set, x_i is the value of the i^{th} data sample, μ and σ are the mean value and the standard deviation, respectively.

Kurtosis: measures whether the data are peaked or flat relative to a normal distribution. Data sets with high kurtosis tend to have a distinct peak near the mean, decline rather rapidly and have heavy tails. Data set with low kurtosis tend to have a flat top near the mean rather than a sharp peak

$$\beta_2 = \frac{1}{N} \sum_{i=1}^{N} \frac{(x_i - \mu)^4}{\sigma^4}. \tag{4.6}$$

These metrics capture spread, asymmetry, peakedness, and overall variability, effectively reducing the high-dimensional CSI streams to interpretable feature vectors.

Classification We consider different machine learning methods to be applied to the extracted feature set, belonging to hyperplane-based and decision tree solutions: *Random Forest* (RF), *Support Vector Machine* (SVM), and the *J48* decision tree (an implementation of the C4.5 algorithm). Each model follows a distinct learning paradigm, leading to different strengths, weaknesses, and levels of interpretability. A clear understanding of their differences is crucial for selecting the most suitable classifier in Wi-Fi CSI sensing applications, where datasets may be limited, noisy, and high dimensional.

Random Forest is an ensemble learning algorithm that constructs a large number of decision trees and aggregates their predictions to produce a final output. Each tree is trained on a randomly drawn *bootstrap sample* of the dataset, and at every split only a randomly selected subset of features is considered. This process introduces diversity among trees and mitigates overfitting, which is a common issue with single decision trees. Predictions are obtained by *majority voting* among all trees. Mathematically, if $\{T_b(x)\}_{b=1}^{B}$ are the predictions from B independently trained trees, the Random Forest classifier outputs:

$$\hat{y} = \text{mode}\{T_b(x)\}_{b=1}^{B}. \tag{4.7}$$

RF offers several advantages for CSI-based sensing:

Robustness to noise and variability: bagging and random feature selection help to stabilize performance in the presence of multipath fading and hardware imperfections.

Scalability to high-dimensional feature sets: RF can handle the large number of subcarriers and statistical descriptors without explicit feature selection.

Feature importance analysis: the algorithm provides a measure of how much each feature (e.g., variance or skewness of a specific subcarrier) contributes to classification.

Its main drawbacks are the increased computational cost when a large number of trees is used and reduced interpretability compared to a single decision tree.

Support Vector Machines are margin-based classifiers that aim to find the optimal separating hyperplane between different classes. For linearly separable data, SVM

maximizes the margin between the closest points of opposite classes, known as *support vectors*. Through the *kernel trick*, SVM can implicitly map data into high-dimensional feature spaces, allowing the classifier to model complex, non-linear decision boundaries. For a binary classification problem, the SVM optimization problem is:

$$\min_{\mathbf{w},b,\xi_i} \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{N}\xi_i \quad \text{s.t.} \quad y_i(\mathbf{w}^\top\phi(x_i)+b) \geq 1-\xi_i,\ \xi_i \geq 0, \tag{4.8}$$

where $\phi(\cdot)$ is the feature mapping defined by the chosen kernel function and C controls the trade-off between maximizing the margin and minimizing classification errors. SVM is well suited for CSI-based sensing because:

Strong generalization: the margin maximization principle provides good performance on unseen data.

Capability to handle non-linear separability: with appropriate kernels (e.g., radial basis function), SVM can model complex interactions between subcarriers and features.

Effective with high-dimensional data: it remains reliable even when the number of features exceeds the number of samples.

However, SVMs require careful parameter tuning (kernel type, regularization C, and kernel hyperparameters) and can be sensitive to outliers or class imbalance. Moreover, their decision boundaries are less interpretable than those of tree-based methods.

J48 (C4.5 Decision Tree) is the Java implementation of the well-known C4.5 decision tree algorithm. It builds a single, interpretable tree by recursively partitioning the dataset using the feature that provides the highest *information gain ratio* (IGR), a metric that balances class separation power and attribute bias. Once the tree is fully grown, a *pruning* step removes branches that do not improve generalization, reducing overfitting. The information gain ratio for a feature A is defined as:

$$\text{IGR}(A) = \frac{\text{Gain}(A)}{\text{SplitInfo}(A)}, \tag{4.9}$$

where

$$\text{Gain}(A) = H(\mathcal{S}) - \sum_{v \in \text{values}(A)} \frac{|\mathcal{S}_v|}{|\mathcal{S}|} H(\mathcal{S}_v), \tag{4.10}$$

and $H(\cdot)$ is the Shannon entropy. J48 is particularly attractive for CSI-based recognition because:

High interpretability: the resulting decision tree can be analyzed to understand which subcarriers or features (e.g., variance or IQR) drive classification decisions.

Low computational cost: training and inference are lightweight and suitable for real-time systems.

Minimal parameter tuning: it works effectively with default settings.

Nevertheless, single decision trees are prone to overfitting when the dataset is noisy or highly variable, and their accuracy is generally lower than that of ensemble methods or well-tuned SVMs. The three models represent a trade-off between accuracy, robustness, and interpretability:

Random Forest tends to achieve the highest classification accuracy for CSI-based material and object recognition thanks to its ensemble nature and resistance to noise, while still allowing feature importance analysis.

SVM is particularly effective in high-dimensional, non-linear problems but requires parameter tuning and lacks interpretability.

J48 offers full interpretability and fast execution, making it valuable for understanding the role of specific features and for lightweight applications, albeit at the cost of lower robustness and accuracy compared to RF and SVM.

Table 4.1 summarizes their key characteristics.

Model	Type	Strengths	Limitations
Random Forest	Ensemble of decision trees (bagging + random feature selection)	High accuracy, robust to noise, works well with many features, provides feature importance	Computationally heavier than single trees, less interpretable
Support Vector Machine (SVM)	Margin-based classifier with kernel trick	Strong generalization, handles non-linear patterns, effective in high-dimensional spaces	Requires parameter tuning, sensitive to outliers and class imbalance, less interpretable
J48 (C4.5)	Single decision tree	Interpretable, fast training and inference, minimal tuning	Lower accuracy in complex tasks, prone to overfitting on noisy data

Table 4.1: *Comparison of the classification models used for CSI-based recognition.*

In summary, Random Forest is generally the most robust and accurate for the proposed Wi-Fi CSI sensing framework, SVM offers strong performance in non-linear and high-dimensional problems, while J48 provides a simple and interpretable model for understanding feature relevance and rapidly prototyping recognition tasks.
The extracted features are fed into several supervised classifiers:

Random Forest (RF): ensemble of decision trees robust to noise and feature interactions.

Support Vector Machine (SVM): effective in high-dimensional feature spaces.

C4.5 (J48): interpretable decision tree classifier.

These models learn to discriminate between materials, categories, or specific objects based on the extracted statistical features. In parallel, a deep learning–based approach is explored, where the pre-processed CSI sequences are directly provided to a Long Short-Term Memory (LSTM) neural network. This approach avoids manual feature engineering, automatically learning temporal patterns and discriminative representations. It is particularly suitable for complex, time-dependent tasks such as movement detection and respiratory monitoring. Depending on the chosen classification path and model, the system addresses three recognition levels:

Material Recognition (MR): determining the material composition (e.g., glass vs. plastic).

Category Recognition (CR): classifying an object into a general category (e.g., bottle, box).

Object Recognition (OR): identifying specific instances of objects, even within the same category.

This architecture deliberately separates signal pre-processing from classification, enabling two complementary strategies. The feature-based machine learning branch offers interpretability and robustness when the dataset is limited or domain knowledge is available. The deep learning branch (LSTM) provides an end-to-end solution capable of automatically learning complex temporal relationships, but it requires more data and computational resources. By comparing these two paths under the same experimental framework, we can fairly evaluate the trade-offs between classic feature-driven approaches and modern deep architectures for Wi-Fi CSI sensing tasks, ranging from object/material recognition to human movement and breathing detection.

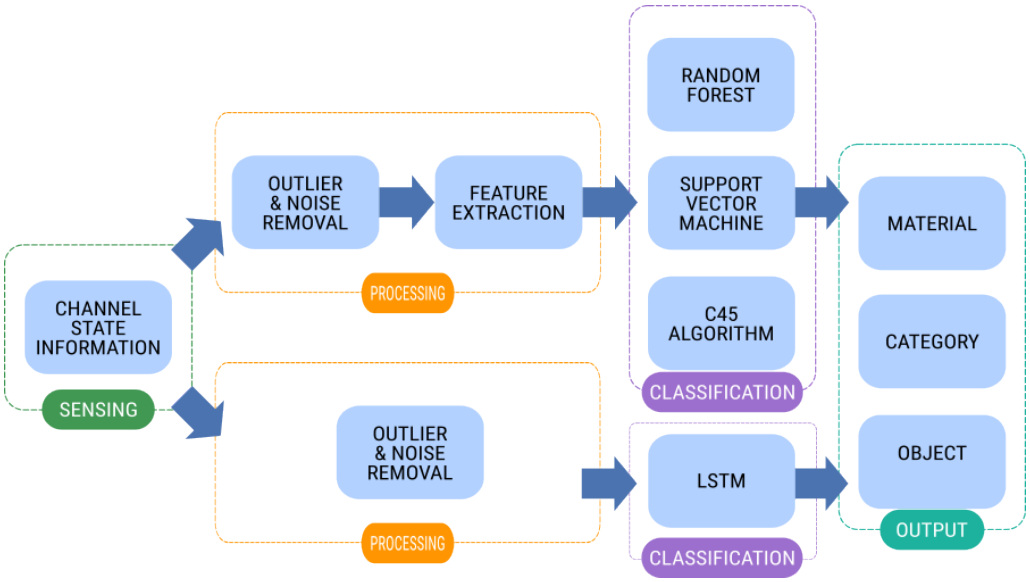

Figure 4.5: *Object recognition architecture, tested with different classifiers for the selected recognition tasks.*

4.2.1 Experimental validation

4.2.1.1 Material recognition

The first part of the experimental results focused on the task of material recognition (MR), that is identifying the physical composition of an object based on its impact on the wireless

channel. Seven distinct materials, chosen to represent a wide range of electromagnetic properties, were included in the tests:

Aluminum – highly conductive and reflective.

Glass – dielectric but with relatively low absorption at 5 GHz.

Metals – strong reflectors with high conductivity.

Nylon – light, low-loss polymer with weak reflection.

Paper – porous and weakly reflective.

Plastic – low permittivity, weak reflection but stable surface.

Wood – lossy, irregular material with moderate absorption.

Each material was placed within the line-of-sight (LoS) path between the Wi-Fi transmitter and receiver, and CSI was captured over long sequences to ensure statistical stability. From each measurement, pre-processing was applied to remove outliers and suppress noise, followed by the feature extraction pipeline described previously. The resulting feature vectors were then used to train and evaluate four classification approaches: *Random Forest* (RF), *Support Vector Machine* (SVM), *J48* (C4.5 decision tree), and a deep learning model based on a *Long Short-Term Memory* (LSTM) network. A 20-fold cross-validation procedure was employed to guarantee reliable performance estimation.

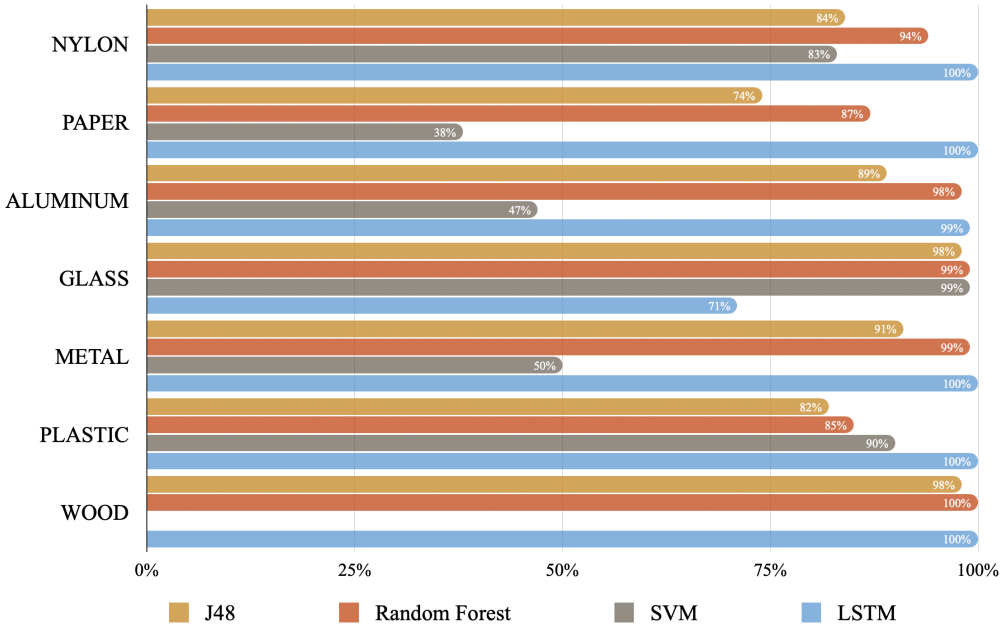

Figure 4.6: *Bar plot showing the average accuracy for material recognition.*

Figure 4.6 reports the average classification accuracy obtained by the four models on the complete dataset. The results show that material differentiation is highly feasible using Wi-Fi CSI features: all classifiers achieve competitive performance, but with significant differences in robustness and precision.

The **Random Forest** classifier reached an average accuracy of **92.7%**, outperforming both SVM and J48. Its ensemble structure mitigates overfitting and exploits the diverse statistical descriptors extracted from CSI.

The **Support Vector Machine (SVM)** achieved an average accuracy of **85.7%**. While it effectively handles the high-dimensional feature space, its sensitivity to parameter tuning (kernel, regularization) may limit robustness in the presence of channel variability.

The **J48 decision tree (C4.5)** performed the worst among the three classical algorithms, with an average accuracy of **61.1%**. Its single-tree structure struggles when the dataset contains noisy or overlapping feature distributions.

The **LSTM deep learning model** reached the highest overall accuracy with **94.8%**, slightly surpassing Random Forest. Its temporal modeling capability allowed it to

learn discriminative patterns in the raw CSI sequences without relying on hand-crafted features.

These results demonstrate that both *Random Forest* and *LSTM deep learning* are particularly effective for CSI-based material recognition. The feature-based RF is competitive with LSTM despite requiring significantly fewer training samples and less computational complexity. A more detailed analysis reveals that classifier performance varies across materials, reflecting their different electromagnetic signatures:

Plastic: RF achieved **100% correct classification** on plastic objects, while SVM and J48 showed weaker performance. Plastic's relatively stable and low-loss response appears to match well with the statistical feature set exploited by RF.

Glass: both RF, SVM, and J48 recognized glass with very high accuracy (**98–99%**), significantly outperforming the LSTM network, which dropped by approximately **25%**. This suggests that, for highly stable and non-dispersive materials, explicit statistical descriptors (median, variance, IQR) remain more discriminative than deep learned temporal patterns.

Metals and aluminum: all classifiers performed well, but RF and LSTM slightly outperformed SVM and J48, likely because strong reflections from metals produce clear, low-variance CSI signatures.

Nylon and paper: these more absorptive, weakly reflective materials were harder to classify. Here, LSTM showed a slight advantage due to its ability to exploit subtle temporal variations, while RF still maintained good performance by combining features such as variance and kurtosis.

Wood: this material showed high classification accuracy for both RF (100%), LSTM (100%) and J48 (98%), whereas SVM and underperformed. The heterogeneous internal structure of wood and its intermediate electromagnetic absorption cause moderate signal variability, which is well captured by the feature ensemble (variance, skewness and kurtosis) of RF and J48, and further exploited by LSTM through temporal modelling. SVM's weaker results indicate its difficulty in handling the non-linear and multi-modal nature of the CSI distribution for this material.

Material	**SVM**	**C4.5 (J48)**	**Random Forest**	**LSTM**
Aluminum	55	83	91	**96**
Glass	98	99	**100**	75
Metal	60	86	**93**	95
Nylon	48	79	87	**97**
Paper	52	82	88	**96**
Plastic	99	98	**100**	99
Wood	65	82	90	**98**
Average	61.1	85.7	92.7	**94.8**

Table 4.2: *Per-material classification accuracy [%] for the four evaluated algorithms.*

These findings provide valuable insights into the design of Wi-Fi-based sensing systems. Firstly, statistical features oriented towards stability, such as variance, interquartile range (IQR) and skewness, have proven highly informative when dealing with materials that exhibit consistent electromagnetic behaviour, such as plastic and glass. At the same time, deep learning models, particularly Long Short-Term Memory (LSTM) networks, demonstrate their effectiveness in the presence of subtle, time-dependent micro-variations, as observed in more challenging materials such as nylon or paper. Their ability to autonomously model complex temporal dependencies enables them to capture intricate dynamics that handcrafted statistical features may fail to represent. Finally, ensemble learning methods, particularly the Random Forest classifier, offer a practical advantage in that they achieve performance levels close to those of deep learning models while being significantly easier to train, tune and interpret. Taken together, these observations confirm that recognising materials using Wi-Fi CSI is highly feasible, even with inexpensive off-the-shelf hardware. Random Forest offers an excellent balance between accuracy, robustness and interpretability, whereas LSTM can further improve performance at the expense of increased computational complexity and reduced transparency. These considerations informed the methodological choices for subsequent recognition tasks, in which both Random Forest and LSTM were retained as primary classification strategies, selected according to the sensing scenario's complexity and temporal characteristics. Overall, the experimental campaign confirms that material recognition using Wi-Fi CSI is highly feasible even with off-the-shelf hardware. Random Forest offers a strong balance of accuracy, robustness, and interpretability, while deep LSTM models can further improve performance at the cost of higher training complexity and reduced transparency. These insights informed the choice of classifiers for subsequent recognition tasks, where both RF

and LSTM were retained as strong candidates depending on the complexity of the sensing scenario.

4.2.1.2 Category recognition

A second set of experiments was conducted to assess the ability of the proposed Wi-Fi CSI–based sensing framework to recognize different object categories. In this evaluation, we considered seven representative classes of common everyday items: *backpack*, *book*, *bottle*, *box*, *can*, *laptop*, and *wood piece*. The aim of this campaign was to verify whether the classification pipeline, previously validated for material differentiation, could also generalize to the recognition of object types that combine both material composition and structural geometry. The results, reported in Figure 4.7, highlight marked differences among the tested classifiers.

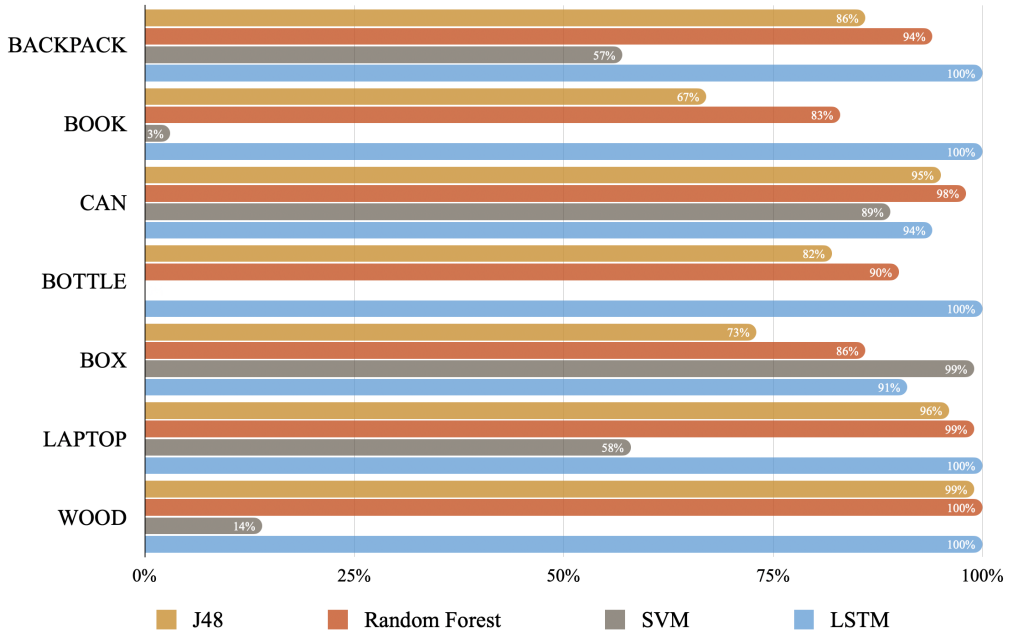

Figure 4.7: *Bar plot showing the average accuracy for category recognition.*

Overall, the **Random Forest (RF)** algorithm demonstrates the most reliable performance among the classical machine learning models, achieving an average recognition accuracy of **90.4%** across all categories. The decision tree–based **C4.5 (J48)** classifier follows with a moderately lower average accuracy of **81.3%**, while the **Support Vector Machine (SVM)** exhibits a significantly weaker performance, with an overall average of only **47.5%**. On the other hand, the deep learning approach based on **Long Short-Term**

Memory (LSTM) networks achieves the highest accuracy overall, reaching **97.1%**, thus confirming its strong ability to model subtle temporal dynamics in the CSI signal.

Classifier	**SVM**	**C4.5 (J48)**	**Random Forest**	**LSTM**
Average Accuracy [%]	47.5	81.3	90.4	97.1

Table 4.3: *Average category recognition accuracy for the considered classifiers.*

A closer examination of the per-category results reveals some particularly striking trends. For instance, the **Book** category proves challenging for traditional machine learning methods: SVM achieves only **3%** accuracy, and C4.5 reaches **67%**, whereas RF performs better at **83%**, and LSTM achieves a perfect **100%**. Similarly, the **Bottle** category is weakly classified by SVM (**8%**) but is recognized with much higher accuracy by RF (**90%**), C4.5 (**82%**), and almost perfectly by LSTM (**99%**). Other categories such as **Can**, **Box**, and **Laptop** also show a similar trend, where SVM struggles considerably (ranging between 57% and 89%) while both RF and C4.5 achieve competitive results, and LSTM consistently outperforms all other approaches, often achieving or approaching **100%** accuracy. Finally, for the **Wood piece** category, SVM drops to only **14%**, whereas C4.5 achieves **99%**, RF reaches **100%**, and LSTM again achieves **100%**.
These results confirm several key observations: first, the poor performance of SVM across most categories suggests that the CSI features employed here are not linearly separable in the original feature space and that kernel-based separation struggles to model the complex propagation patterns associated with different object shapes and compositions. In contrast, **Random Forest** proves highly competitive, benefiting from its ensemble structure and ability to handle heterogeneous, non-linear relationships between features. Its performance remains close to that of the deep LSTM model while retaining easier interpretability and lower computational complexity. **LSTM**, however, remains the most accurate solution overall, confirming that its capacity to learn temporal dependencies and model subtle micro-variations in the CSI signal is advantageous when object categories induce fine-grained and time-varying multipath effects. From a practical standpoint, these findings suggest that RF remains an excellent compromise between accuracy, interpretability, and computational cost, particularly when the dataset size or computational resources are limited. However, in scenarios where maximum classification performance is required and sufficient training data are available, the LSTM-based approach is clearly preferable, as it consistently outperforms traditional machine learning models in recognizing complex object categories.

4.2.1.3 Object recognition

The final and most challenging classification task concerns specific object recognition. Unlike material or category recognition, which aim to group items by shared physical properties or broad semantic classes, this task requires the system to distinguish between *individual objects*, even when they share the same material and general shape. The experimental campaign considered eleven representative everyday objects: *backpack*, *book*, *rigid (hardcover) book*, *cardboard box*, *aluminum can*, *1 L glass bottle*, *glass box*, *0.5 L plastic bottle*, *plastic box*, *laptop*, and a *wood piece*. This scenario is particularly demanding because it pushes the classifiers to exploit subtle differences in the Channel State Information (CSI) patterns caused by object geometry, thickness, surface finishing, and minute variations in electromagnetic absorption and reflection.

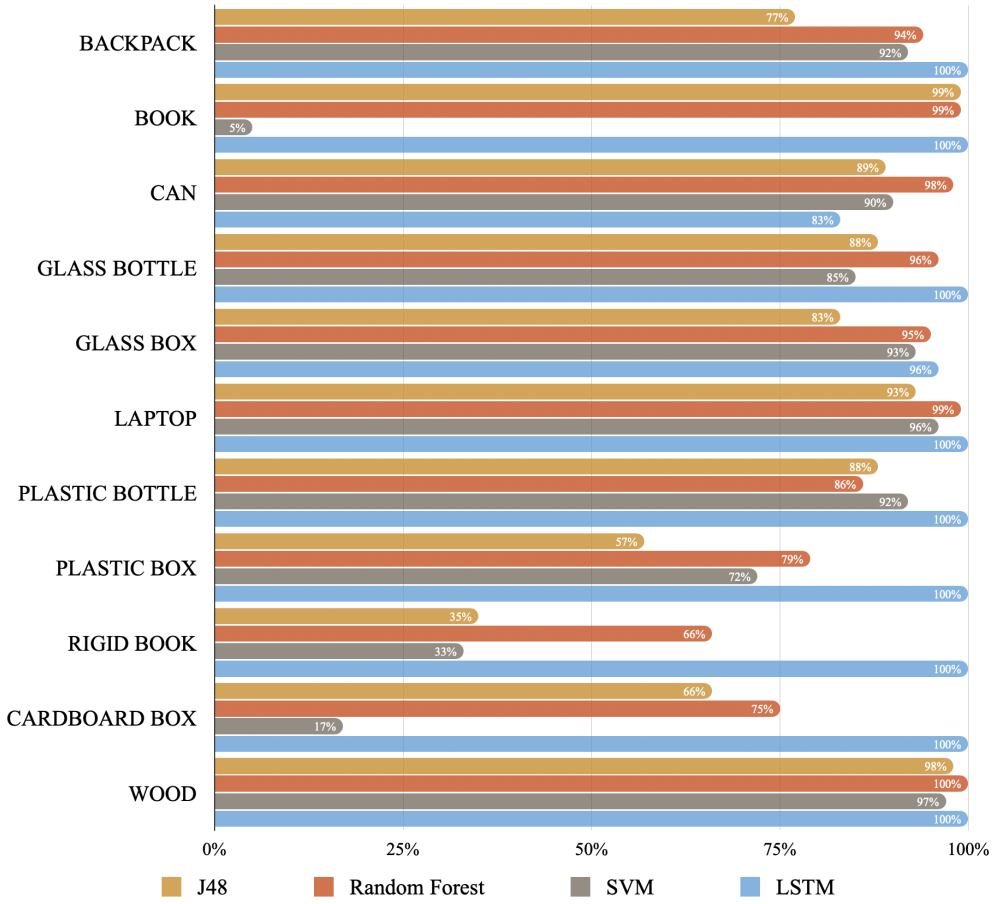

Figure 4.8: *Bar plot showing the average accuracy for object recognition.*

Figure 4.8 reports the per-object accuracy achieved by the considered classification algorithms. The results reveal a consistent performance hierarchy among the evaluated models. The **Long Short-Term Memory (LSTM)** network once again outperforms all other solutions, achieving an impressive **98.1%** average accuracy across all eleven objects. Among the traditional machine learning methods, the **Random Forest (RF)** classifier remains the strongest, with an average accuracy of **89.7%**. The **C4.5 (J48)** decision tree achieves moderately lower accuracy, averaging **79.5%**, while the **Support Vector Machine (SVM)** shows the weakest performance with an average of **70.2%**.

A closer inspection of the per-class results highlights the relative strengths and weaknesses of each model. For objects with distinctive electromagnetic signatures, such as the *backpack*, *wood piece*, or *aluminum can*, RF and LSTM both reach near-perfect accuracy. However, when discriminating between very similar objects — for example, a *book* versus a *rigid (hardcover) book*, or a *glass bottle* versus a *glass box* — the advantage of LSTM becomes clear. Its ability to model subtle temporal dependencies and fine-grained multipath variations enables it to consistently outperform the handcrafted feature-based models. In contrast, SVM struggles significantly in these ambiguous scenarios, with recognition rates as low as 5–35% for some book-related classes. J48 performs better than SVM but shows limited ability to separate classes when small differences in shape or cover rigidity are the only distinguishing factors. RF strikes a good balance, achieving robust performance across most classes while maintaining interpretability and ease of training.

Classifier	**SVM**	**C4.5 (J48)**	**Random Forest**	**LSTM**
Average Accuracy [%]	70.2	79.5	89.7	98.1

Table 4.4: *Average accuracy in specific object recognition for the evaluated classifiers.*

When considering the broader experimental campaign across all recognition levels, several trends emerge. Random Forest consistently demonstrates strong performance across the three investigated tasks — **92.7%** in material recognition, **90.4%** in category differentiation, and **89.7%** in specific object recognition — proving to be a reliable and interpretable choice when using handcrafted features. C4.5 provides moderately good results but is less robust to complex, non-linear separations. SVM, while occasionally effective in simpler classification scenarios, consistently underperforms in the highly variable wireless sensing context. Finally, the LSTM model achieves the highest accuracy across all tasks (**94.8%** for material, **97.1%** for category, and **98.1%** for object recognition), confirming that deep sequential modeling of the CSI is particularly well-suited for

fine-grained recognition problems where subtle temporal dynamics and multipath signatures must be exploited.

Overall, these results demonstrate that Wi-Fi CSI can be effectively leveraged for specific object identification even using low-cost, off-the-shelf MIMO hardware. While Random Forest remains an excellent trade-off between accuracy, interpretability, and training complexity, deep learning methods such as LSTM provide a substantial performance gain when the goal is maximum precision in challenging, fine-grained recognition tasks.

4.3 Activity Recognition

To evaluate the ability of the proposed Wi-Fi CSI sensing framework to detect and classify human movements, a dedicated experimental campaign was carried out in a controlled indoor environment. The devices were positioned on desks at opposite ends of an otherwise empty room, facing each other in line-of-sight (LoS) configuration, so that a human subject could move within the propagation path. The employed methodology is illustrated in Figure 4.9.

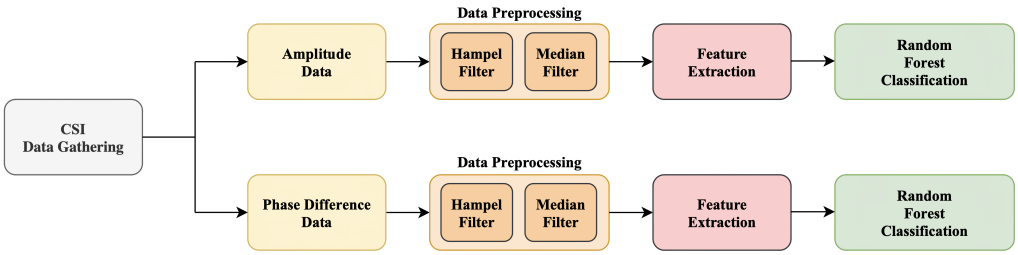

Figure 4.9: *Block diagram of the proposed CSI-based sensing methodology. After CSI data gathering, the processing pipeline is split into two parallel branches operating on amplitude and inter-antenna phase difference information, respectively. Both branches apply robust preprocessing based on Hampel and median filtering to suppress outliers and noise, followed by feature extraction and Random Forest classification. This dual-stream design enables a comparative analysis of amplitude- and phase-based representations under identical experimental conditions.*

As illustrated in Fig. 4.9, the sensing pipeline is intentionally structured in two parallel branches to isolate and compare the contributions of amplitude- and phase-based CSI features, while preserving identical preprocessing and classification stages. During the experiments, a single person was positioned between the two PCs and performed a set of predefined activities (e.g., standing still, sitting down, walking, and running). Each activity was executed continuously for a duration of 60 seconds, while Channel State

Information was continuously collected. Six different volunteers participated in the data collection to ensure variability in body size and movement style. The Intel 5300 CSI tool was configured to operate under different antenna configurations: *SISO* (single-input single-output), *SIMO* (single-input multiple-output), *MISO* (multiple-input single-output) and full *MIMO* (multiple-input multiple output), using channel 64, centered at 5.320 GHz. This setup enabled the comparative analysis of how different link configurations influence sensing performance. Each received Wi-Fi packet provides the estimated frequency-domain channel response, stored in the complex CSI matrix $\tilde{\mathbf{H}}$, which contains amplitude and phase values for each transmit–receive antenna pair and subcarrier. As CSI streams are typically affected by hardware jitter, environmental interference, and impulse noise, a two-stage filtering pipeline was applied before feature extraction. First, a *Hampel filter* was employed to remove abrupt outliers, followed by a *median filter* to suppress random fluctuations while preserving the underlying channel dynamics. These steps were applied to both the amplitude and the phase information to ensure stability and improve the discriminative quality of the signals. After cleaning, the CSI time series were divided into non-overlapping segments, and for each segment a set of well-established statistical descriptors was computed. These features include: median, variance, skewness, kurtosis, range, interquartile range (IQR), signal power, and standard deviation. Such metrics are widely used in signal analysis and have proven effective in characterizing both stationary and dynamic channel behavior. Importantly, these operations were carried out on both the *amplitude* and the *phase difference* between receiving antennas. Phase difference was calculated as suggested in [21, 4] by subtracting the instantaneous phase of two receiving antennas for each subcarrier i and packet p:

$$\angle \tilde{H}_{i,p}^{(j,k)} = \angle \tilde{H}_{i,p}^{(j)} - \angle \tilde{H}_{i,p}^{(k)}, \quad i \in [1, 30],\ j, k \in [1, 3],\ j \neq k. \tag{4.11}$$

This representation cancels common-mode phase noise and emphasizes relative changes due to movement, making it particularly robust against hardware-induced phase drift.
The following figures illustrate example time-series segments (10 seconds each) comparing amplitude and phase difference signals under various activities: standing vs. walking (Fig. 4.11), standing vs. sitting (Fig. 4.10), and walking vs. running (Fig. 4.12).

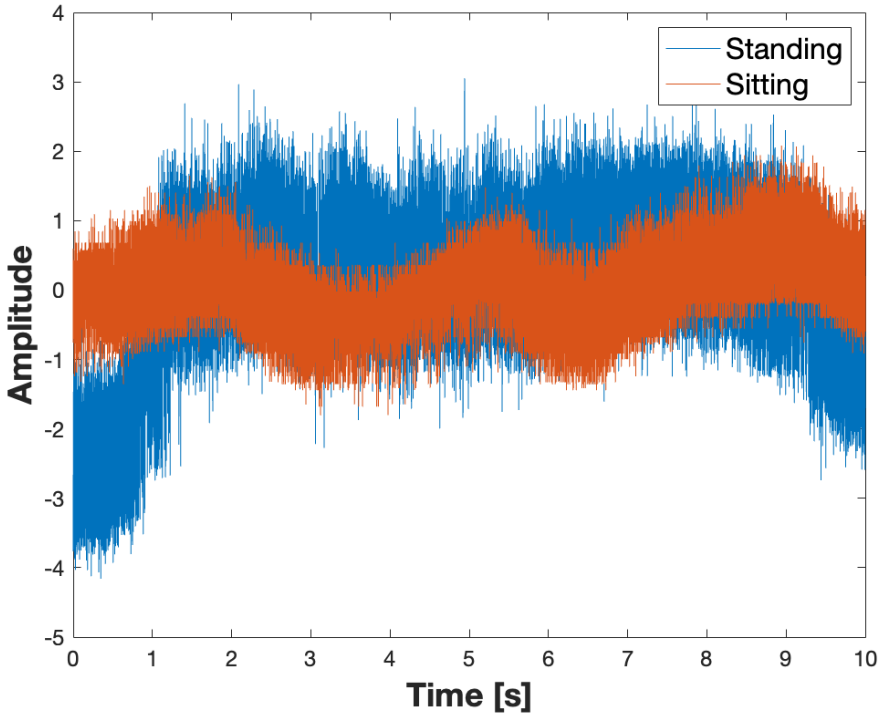

Figure 4.10: *Comparison between CSI amplitude variation in the standing vs sitting case.*

When comparing standing and sitting, the amplitude traces exhibit relatively low overall variability, confirming that both activities induce only minor perturbations in the wireless channel. The standing condition (blue curve) shows slightly higher fluctuation levels than sitting, with moderate amplitude swings and small random spikes. In contrast, the sitting condition produces a more stable and compact amplitude profile, characterized by a narrower dynamic range and reduced random noise. This behavior indicates that static postures create a relatively constant multipath environment, with the sitting position further attenuating micro-movements of the body.

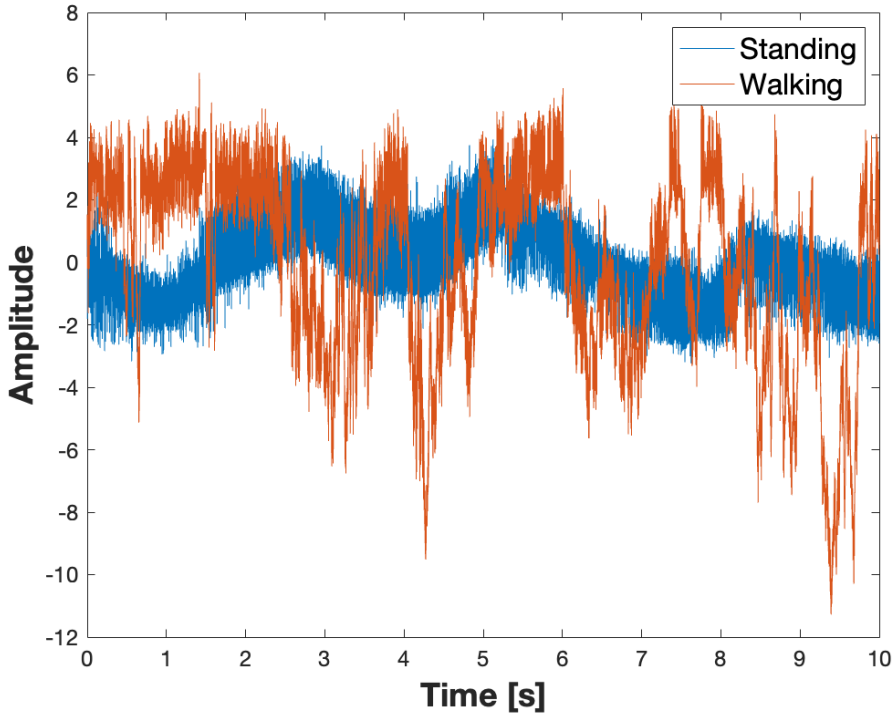

Figure 4.11: *Comparison between CSI amplitude variation in the standing vs walking case.*

A more pronounced contrast emerges when standing and walking are compared. While the standing trace remains relatively compact and stable, the walking trace (orange curve) exhibits strong amplitude fluctuations, including deep fades and sharp peaks. The movement of the limbs introduces periodic and irregular changes to the multipath structure while walking, producing rapid constructive and destructive interference in the wireless channel. Consequently, the amplitude signal during walking displays a significantly larger dynamic range and volatility than in the static condition.

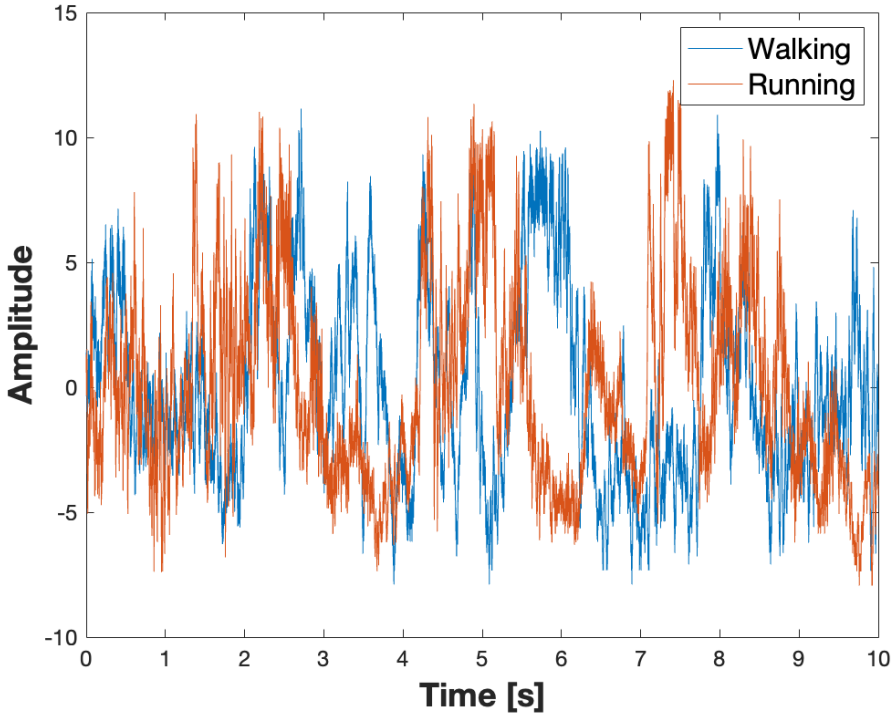

Figure 4.12: *Comparison between CSI amplitude variation in the walking vs running case.*

The difference between walking and running is even more evident. Running generates highly irregular and bursty amplitude variations, with extreme peaks and deep fades far exceeding those observed during walking. While walking induces quasi-periodic fluctuations associated with the stride cycle, running results in more chaotic and abrupt amplitude transitions due to faster body motion, stronger Doppler effects, and frequent multipath disruptions. The amplitude signal appears considerably less stable and more power-dispersed under high-mobility conditions.

These observations confirm that amplitude-based CSI analysis can effectively distinguish between static and dynamic activities. Static actions such as standing or sitting yield compact, low-variance signals, whereas walking and running progressively increase amplitude volatility and dynamic range due to stronger and more frequent body-induced multipath perturbations. Nevertheless, amplitude alone may be insufficient to discriminate activities with similar motion dynamics or to ensure robustness under noisy conditions. This limitation motivates the complementary exploitation of phase-difference analysis and feature engineering in the proposed sensing framework.

From a qualitative perspective, phase difference provides a clearer and more stable separation between different activities. Static conditions, such as standing, produce nearly constant phase-difference traces, whereas even subtle movements, such as sitting down, generate distinctive, high-contrast patterns. By contrast, amplitude signals tend to fluctuate even when the subject is stationary, and this variability can obscure the differences between slow and moderate movements. These experiments confirm that phase difference is a highly discriminative feature for recognising human movement, outperforming raw amplitude in terms of both stability and contrast between activity classes. While amplitude information remains valuable, especially in environments with strong reflections, it is more susceptible to small-scale fading and noise. Combining robust pre-processing (Hampel + median filtering) with feature modelling on amplitude and phase-difference signals provides reliable input for subsequent machine or deep learning classification models.

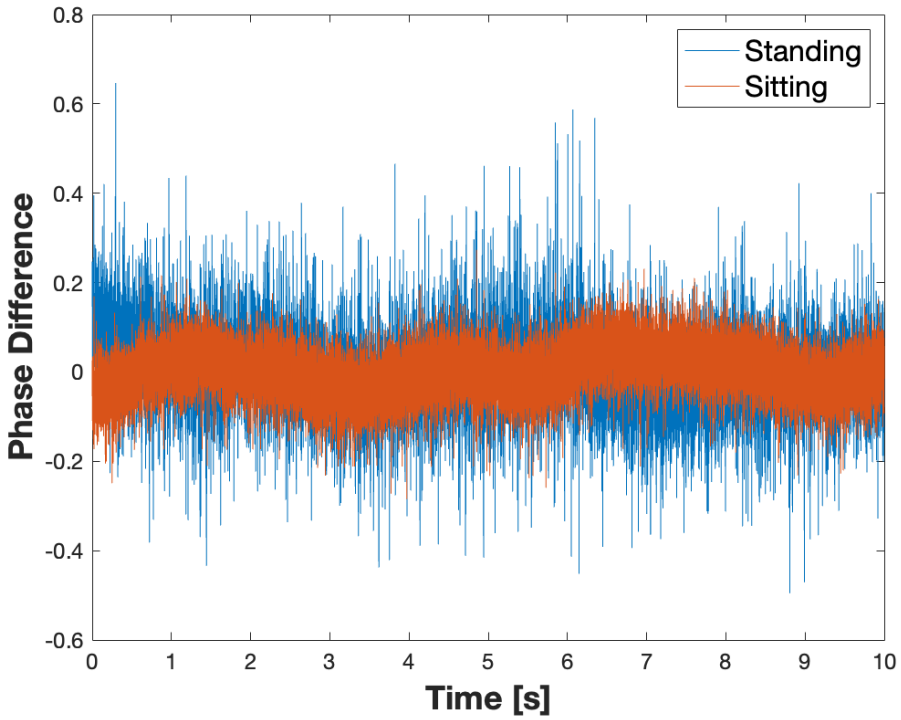

Figure 4.13: *Comparison between CSI phase difference variation in the standing vs sitting case.*

In the standing vs sitting scenario (Figure 4.13), the phase difference remains relatively stable in both cases, with only minor fluctuations around zero. However, the sitting condition exhibits a slightly more compact and centred distribution, whereas the standing trace shows minor low-frequency drifts and occasional spikes. This suggests that, while

the phase difference remains fairly constant in static conditions, subtle changes in body geometry (e.g. sitting posture) can still introduce measurable phase shifts.

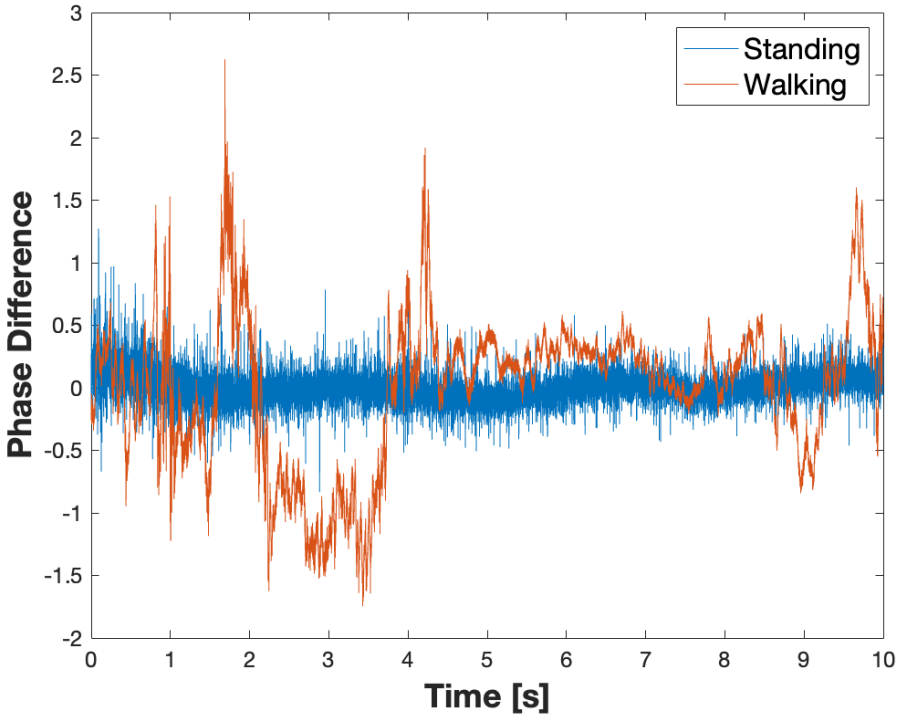

Figure 4.14: *Comparison between CSI phase difference variation in the standing vs walking case.*

In the standing vs. walking case (Figure 4.14), a clear divergence between the two signals is observed. While the standing phase remains narrow and stable, the walking trace exhibits significant fluctuations, with peaks exceeding ±1.5 rad. These rapid variations are caused by the dynamic multipath effect introduced by the movement of the legs and arms, which alter the propagation geometry and thus induce non-stationary phase perturbations.

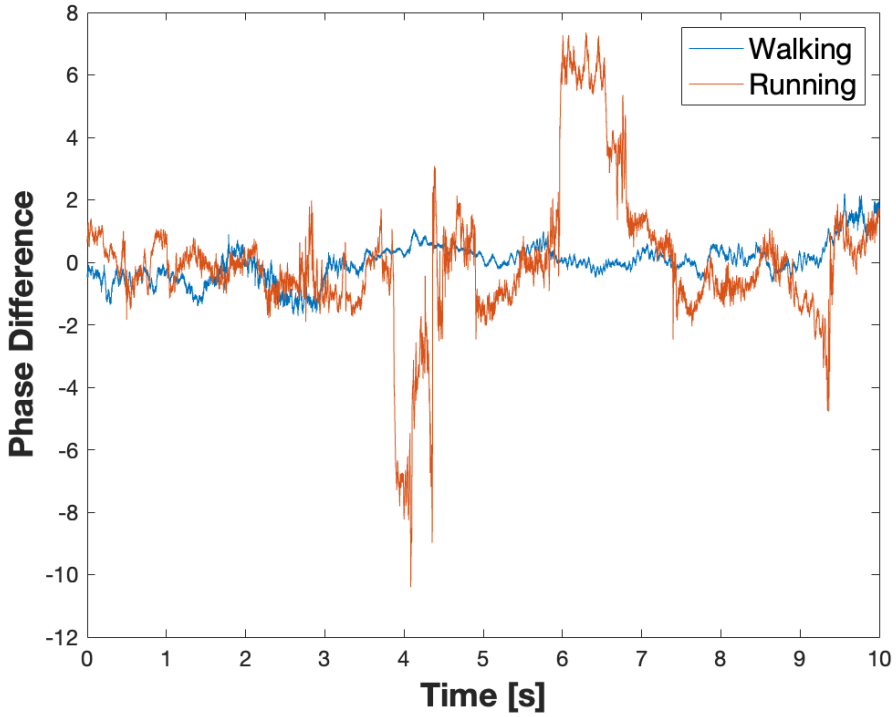

Figure 4.15: *Comparison between CSI phase difference variation in the walking vs running case.*

Finally, the walking vs. running comparison (Figure 4.15) reveals the strongest temporal dynamics. The running trace shows abrupt and large swings of up to ±10 rad, far exceeding the smoother, lower-amplitude oscillations of walking. This dramatic increase reflects the more intense and frequent body movements during running, which cause rapid path length variations and strong Doppler-induced phase shifts.

Overall, these results emphasise that phase difference is a highly sensitive indicator of human motion. While amplitude is affected by noise, even in static scenarios, phase difference remains stable when no movement occurs, reacting strongly and distinctly to different motion intensities. This makes phase difference a particularly valuable feature for discriminating between static and dynamic human activities and for distinguishing between different motion patterns, such as walking and running.

4.3.1 Experimental validation

Tables 4.11 and 4.12 report the mean classification accuracy achieved under different radio link configurations when using either *amplitude* or *phase difference* information extracted from the CSI. Table 4.11 shows results averaged for each individual participant, while

Table 4.12 summarizes the overall accuracy when the entire dataset, including data from all subjects, is considered.

Link Mode	Amplitude Information	Phase Information
SISO	86.38%	-
SIMO	87.39%	90.90%
MISO	71.07%	-
MIMO	73.74%	99.49%

Table 4.5: *Mean accuracy using different radio link modes (individual)*

Link Mode	Amplitude Information	Phase Information
SISO	84.10%	-
SIMO	81.94%	84.63%
MISO	68.66%	-
MIMO	63.44%	99.92%

Table 4.6: *Mean accuracy using different radio link modes (entire dataset)*

The first notable observation concerns the comparison between amplitude-based and phase-based features. When amplitude alone is used, SISO and SIMO links already achieve reasonably high accuracy on individual subjects, with average success rates of 86.38% and 87.39%, respectively (Table 4.11). MISO shows weaker performance (71.07%), while MIMO surprisingly underperforms in amplitude-only mode (73.74%). This reduction in performance is due to the higher dimensionality and increased noise sensitivity of raw amplitude measurements when multiple antenna pairs are used without a phase-stabilising mechanism. Conversely, adopting phase difference as the main descriptor drastically improves MIMO-based classification. Using phase information, MIMO achieves mean accuracy of **99.49%** on single-subject tests and **99.92%** when aggregating data from multiple individuals (see Tables 4.11 and 4.12). This remarkable performance confirms the stability and discriminative power of phase difference, which is less affected by variations in hardware gain variations and more sensitive to small path-length changes induced by human motion. Further comparisons between Tables 4.11 and 4.12 highlight the generalisation capability of the proposed approach. As expected, all link modes perform slightly better when training and testing are performed using data from the same individual.

However, the decrease in accuracy when transitioning to multi-subject testing is relatively modest for SIMO and negligible for MIMO when using phase information. Specifically, MIMO with phase difference decreases only from 99.49% to 99.92%, which is effectively stable, demonstrating that the phase-based features generalise well across different body shapes and movement styles. In contrast, amplitude-based features are more subject-dependent: SIMO drops from 87.39% to 81.94%, and SISO from 86.38% to 84.10%, likely due to the variations in received signal strength caused by different body sizes and movement patterns.

In addition, the following tables present the confusion matrices for selected representative scenarios, allowing a more granular understanding of patterns.

Actual \ Predicted	Lie	Fall	Walk	Run	Sit	Stand
Lying	**93%**	0%	0%	1%	6%	0%
Falling	0%	**96%**	3%	0%	1%	0%
Walking	0%	1%	**99%**	0%	0%	0%
Running	1%	0%	0%	**99%**	0%	0%
Sitting	8%	1%	0%	0%	**87%**	4%
Standing	1%	1%	0%	0%	4%	**94%**
Overall Accuracy = 94.89%						

Table 4.7: *Confusion matrix for activity recognition — Individual SIMO case (Amplitude information)*

Actual \ Predicted	Lie	Fall	Walk	Run	Sit	Stand
Lying	**80%**	2%	0%	0%	13%	5%
Falling	1%	**80%**	4%	8%	3%	4%
Walking	0%	3%	**83%**	13%	0%	1%
Running	0%	5%	7%	**87%**	0%	1%
Sitting	10%	0%	0%	0%	**81%**	9%
Standing	6%	2%	0%	0%	12%	**80%**
Overall Accuracy = 81.94%						

Table 4.8: *Confusion matrix for activity recognition — Multiple individuals SIMO case (Amplitude information)*

Actual \ Predicted	Lie	Fall	Walk	Run	Sit	Stand
Lying	**100%**	0%	0%	0%	0%	0%
Falling	0%	**99%**	0%	~0%	0%	~0%
Walking	0%	0%	**99%**	~0%	0%	0%
Running	0%	~0%	~0%	**99%**	0%	~0%
Sitting	0%	0%	0%	0%	**99%**	~0%
Standing	~0%	0%	0%	~0%	~0%	**99%**
Overall Accuracy = 99.44%						

Table 4.9: *Confusion matrix for activity recognition — Individual MIMO case (Phase information)*

Actual \ Predicted	Lie	Fall	Walk	Run	Sit	Stand
Lying	**100%**	0%	0%	0%	0%	0%
Falling	0%	**99%**	0%	0%	~0%	~0%
Walking	0%	0%	**99%**	~0%	0%	0%
Running	0%	0%	~0%	**99%**	0%	~0%
Sitting	0%	0%	0%	0%	**100%**	0%
Standing	0%	0%	0%	0%	0%	**100%**
Overall Accuracy = 99.92%						

Table 4.10: *Confusion matrix for activity recognition — Multiple individuals MIMO case (Phase information)*

The confusion matrices in Tables 4.7–4.10 offer additional insight into the classifier's behavior. When using SIMO with amplitude features (Tables 4.7 and 4.8), the system already achieves solid per-class accuracy on individual subjects (overall 94.89%), but misclassifications appear for visually similar static postures such as *sitting* vs. *standing* (up to 8% confusion) and occasional confusion between *lying* and *sitting*. When multiple subjects are included (Table 4.8), the accuracy drops to 81.94%, with more frequent confusion between static classes and occasional mislabeling of dynamic actions such as *walking* and *running*.

In contrast, the MIMO setup with phase difference (Tables 4.9 and 4.10) achieves near-perfect classification: all activities—including subtle differences between *standing*, *sitting*, and *lying*, are recognized with 99% or higher accuracy in both single-subject and multi-subject cases. Dynamic activities (*walking*, *running*, *falling*) show almost no confusion, underscoring the robustness of phase-based MIMO sensing. In summary, three important findings emerge:

> Phase difference is the most reliable feature for human activity recognition. Its use nearly eliminates misclassification even in cross-subject scenarios and fully exploits the spatial diversity of MIMO.

MIMO configurations are highly advantageous when paired with phase-based features. While amplitude-only MIMO suffers from increased complexity and noise, MIMO with phase difference delivers state-of-the-art accuracy.

Classical machine learning can rival deep learning. A well-tuned Random Forest using engineered statistical features reaches performance levels comparable to or better than complex deep models, while requiring significantly less computational effort and data.

These results confirm that robust Wi-Fi CSI–based human activity recognition can be achieved with relatively simple models, provided that proper feature engineering and MIMO phase information are exploited.

4.4 Breath rate monitoring

Patient monitoring represents a cornerstone of modern healthcare, ensuring the continuous and accurate assessment of vital signs to support early diagnosis and timely clinical intervention. Among the key physiological indicators, the *respiratory rate* plays a central role: abnormal breathing patterns can signal a wide spectrum of health conditions, including respiratory disorders, cardiovascular dysfunction, infections, and post-operative complications. Continuous respiratory monitoring is particularly valuable in the management of chronic diseases such as asthma, Chronic Obstructive Pulmonary Disease (COPD), and heart failure, where tracking breathing dynamics allows clinicians to evaluate disease progression and treatment effectiveness. In perioperative care, respiration monitoring is equally critical to ensure adequate oxygenation and to detect early signs of respiratory distress. Beyond the clinical setting, the measurement of breathing patterns also provides actionable insights for athletes and for patients undergoing physical rehabilitation.

A normal respiratory cycle is composed of two fundamental phases: *inspiration* (inhalation) and *expiration* (exhalation). During inspiration, the diaphragm and intercostal muscles contract, enlarging the thoracic cavity and reducing intrathoracic pressure, which draws air into the lungs. Expiration is usually passive in healthy individuals, driven by the relaxation of the diaphragm and the elastic recoil of lung tissue. Each complete sequence of inspiration followed by expiration constitutes one *breath*. In healthy adults at rest, the respiratory rate typically ranges from *12 to 20 breaths per minute (bpm)*. Well-trained athletes may exhibit lower resting rates (8–12 bpm), while children present higher

frequencies (20–30 bpm depending on age). Regular, sinusoidal breathing with a stable rate in these ranges is generally regarded as an indicator of normal respiratory function.

Deviations from these values can indicate underlying pathology:

Tachypnea (rapid breathing, > 20 bpm) may arise from fever, infection, heart failure, pulmonary disease, or anxiety.

Bradypnea (slow breathing, < 12 bpm) can occur in neurological disorders, respiratory depression, or under the influence of sedative drugs.

Abnormal and periodic breathing patterns—such as Cheyne–Stokes respiration or apneic episodes—are frequently associated with severe cardiac or neurological disease and sleep-related disorders. Accurate, continuous estimation of respiratory rate and early detection of such anomalies are therefore essential to prevent complications and guide clinical decision-making.

Recent advances in the Internet of Medical Things (IoMT) have expanded the possibilities for respiratory monitoring. *Wearable* systems, including chest straps, adhesive patches, and textile-integrated sensors, measure thoracic expansion directly but can be uncomfortable or intrusive for long-term use. *Contactless* solutions, such as camera-based motion tracking, infrared thermal imaging, and radar-based sensors, offer non-invasive alternatives but may face challenges related to privacy, cost, and environmental sensitivity.

In this context, we propose a fully contactless respiration monitoring approach based on Wi-Fi Channel State Information (CSI). The method leverages the fine-grained CSI measurements available on off-the-shelf hardware to detect the minute, periodic changes in the wireless channel induced by chest wall motion during breathing. Importantly, the same experimental setup adopted for human activity recognition is employed: two tower PCs, each equipped with Intel 5300 NICs and three antennas, are positioned at opposite sides of an otherwise empty room in a Line-of-Sight (LoS) configuration. This arrangement creates a stable propagation environment and maximizes sensitivity to small periodic movements caused by respiration.

The underlying principle is that the expansion and contraction of the thorax and abdomen modulate the radio channel, introducing low-frequency oscillations in both the *amplitude* and the *phase difference* of the CSI. To exploit these signals, the same preprocessing pipeline used for activity recognition is applied: outliers are removed using Hampel filtering, random fluctuations are suppressed via median filtering, and statistical features are extracted from the cleaned amplitude and phase sequences. This processing step provides stable, interpretable descriptors of the respiratory signal.

To estimate breathing frequency and detect potential anomalies, both traditional machine learning algorithms (Random Forest, Support Vector Machines, and J48 decision trees) and deep learning models (Long Short-Term Memory networks) are evaluated. All analyses are performed using the subcarrier identified in the preliminary calibration phase as the most reliable and informative, ensuring that the input CSI reflects the most stable portion of the channel frequency response for respiratory monitoring.

4.4.1 Experimental validation

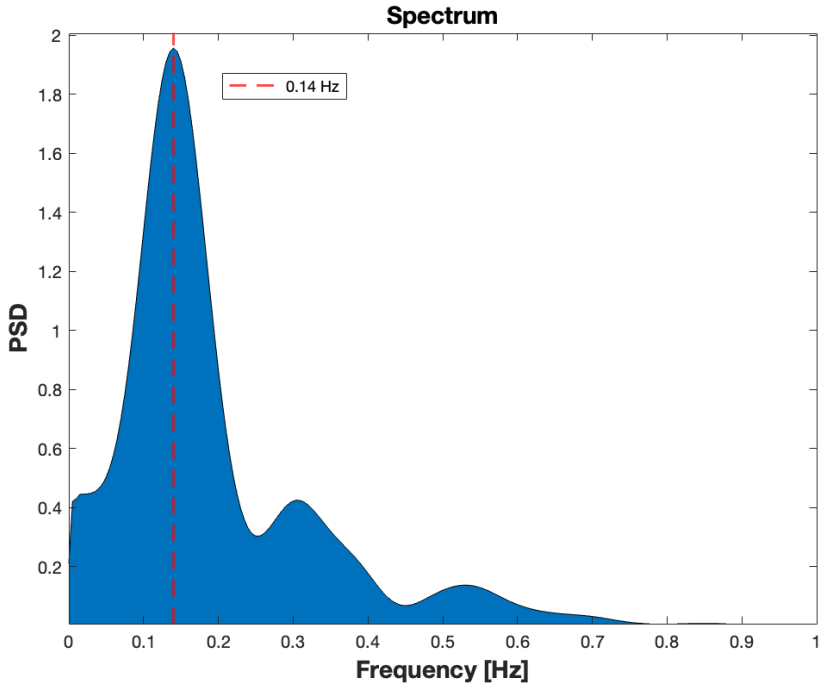

Figure 4.16: *Power spectral density (PSD) of a breathing signal recorded at 0.14 Hz (7 s per cycle). The dashed red line highlights the frequency corresponding to the expected breathing rate.*

Figure 4.16 illustrates an example of a breathing signal collected while a subject was standing still between the transmitter and receiver, breathing at a rate of 0.14 Hz (one breath every 7 s). The power spectral density (bottom panel) reveals a distinct peak precisely at the expected frequency. A similar trend is observed in recordings with faster cycles, such as 0.20 Hz (5 s per cycle), confirming that frequency-domain analysis effectively isolates the periodic component associated with respiration, even when the raw signal appears irregular.

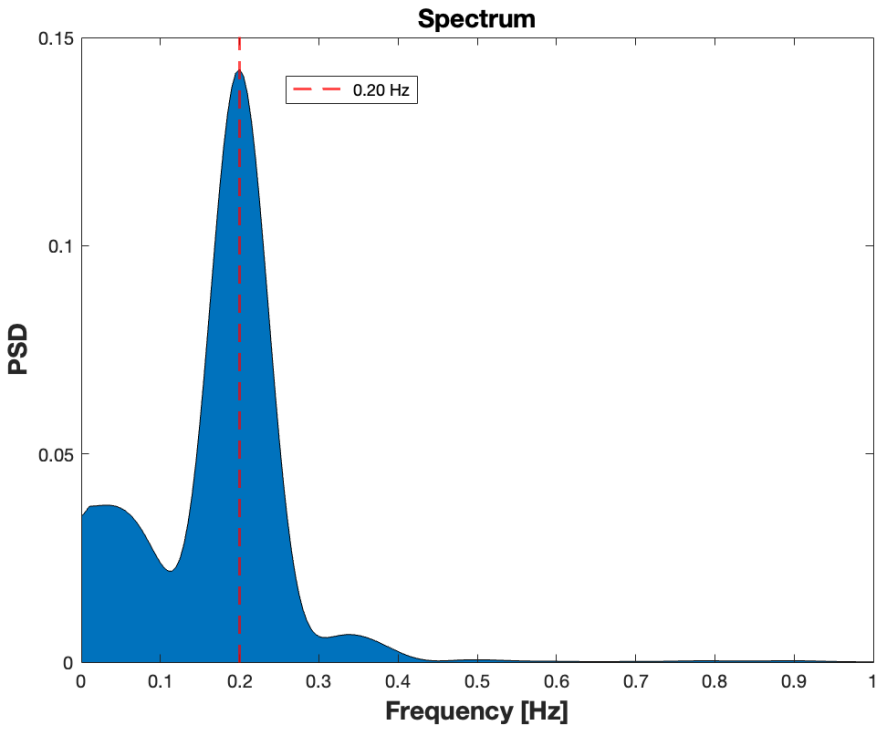

Figure 4.17: *Power spectral density (PSD) of a breathing signal recorded at 0.20 Hz (5 s per cycle). The dashed red line highlights the frequency corresponding to the expected breathing rate.*

To further exploit this information for automatic classification, we evaluated multiple AI algorithms for recognizing breathing rates. Statistical features—including variance, median, interquartile range (IQR), skewness, kurtosis, and range—were extracted from both the amplitude and the phase difference of the CSI, providing compact yet descriptive representations of the underlying respiratory dynamics. The classification performance of Random Forest (RF), Random Tree (RT), C4.5 (J48), and a Long Short-Term Memory (LSTM) network was then compared.

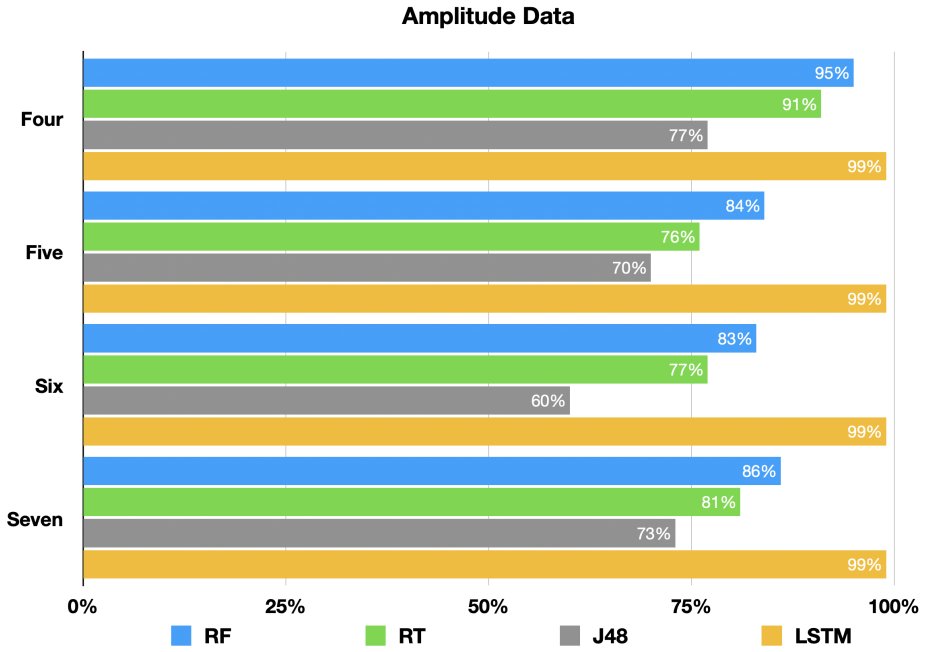

Figure 4.18: *Classification accuracy for different breathing rates using CSI amplitude features.*

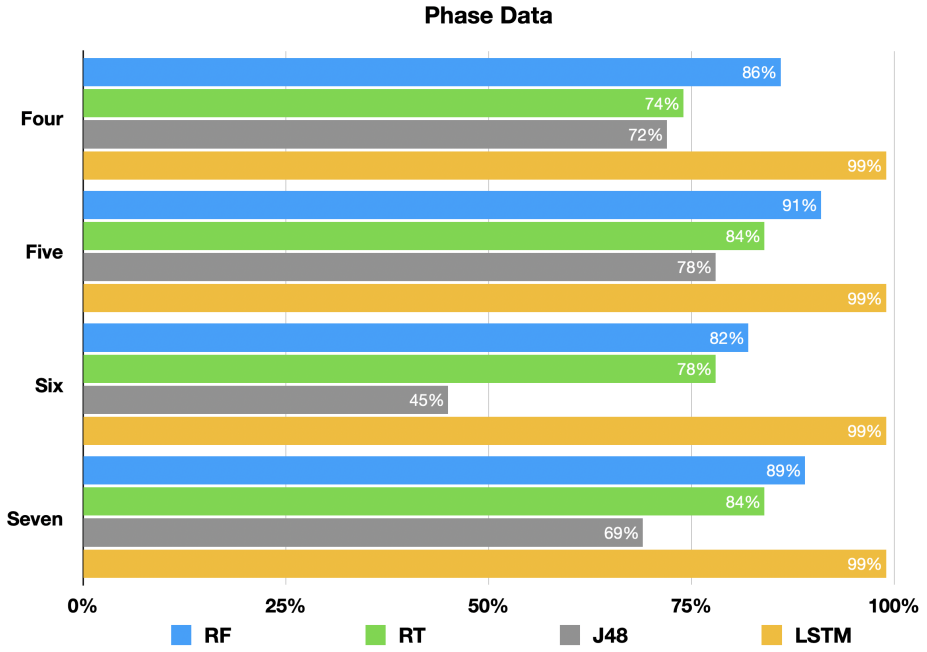

Figure 4.19: *Classification accuracy for different breathing rates using CSI phase difference features.*

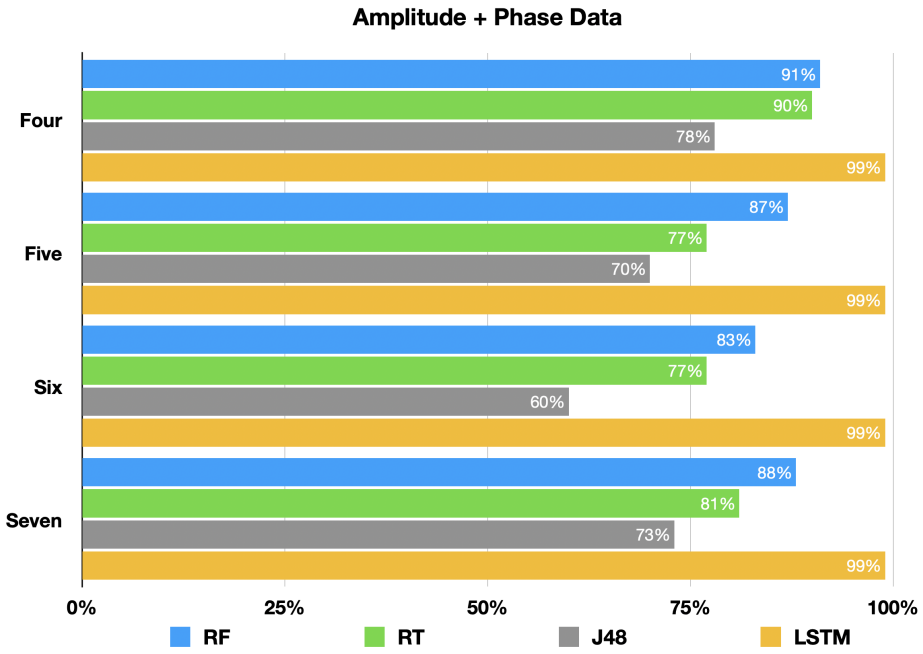

Figure 4.20: *Classification accuracy when combining amplitude and phase features.*

Figures 4.18–4.20 compare the performance of the considered classifiers across different breathing rates (4, 5, 6, and 7 s per cycle). Several key observations can be made:

LSTM consistently achieves the highest accuracy, exceeding 99% in all cases and proving highly effective at modeling the quasi-periodic temporal dynamics of respiration.

Random Forest provides the best results among classical ML models, reaching accuracies between 83% and 95%, with minimal sensitivity to the type of input (amplitude, phase, or their combination).

Random Tree and J48 underperform compared to RF, with J48 showing a marked drop in the more challenging cases (down to 45% with phase-only features for a 6 s cycle), confirming its limited capacity to generalize in the presence of signal variability.

Phase features slightly outperform amplitude when used alone, indicating that phase difference is more robust to static noise and better captures the fine-grained cyclic changes induced by respiration.

Combining amplitude and phase does not significantly outperform using phase alone, suggesting that phase information carries most of the discriminative power for this task.

Tables 4.11 and 4.13 report the confusion matrices obtained for the Random Forest classifier under the three feature configurations: amplitude only, phase only, and combined amplitude + phase. The overall performance summary across all algorithms is reported in Table 4.14.

Actual \ Predicted	Four	Five	Six	Seven
Four	**95%**	2%	3%	0%
Five	2%	**84%**	4%	10%
Six	0%	15%	**83%**	2%
Seven	0%	13%	1%	**86%**
Overall Accuracy = 85.46%				

Table 4.11: *Confusion matrix for breathing rate recognition (RF, Amplitude data)*

Actual \ Predicted	Four	Five	Six	Seven
Four	**86%**	5%	4%	5%
Five	1%	**91%**	2%	6%
Six	4%	7%	**82%**	7%
Seven	2%	8%	1%	**89%**
Overall Accuracy = 87.77%				

Table 4.12: *Confusion matrix for breathing rate recognition (RF, Phase data)*

Actual \ Predicted	Four	Five	Six	Seven
Four	**91%**	3%	3%	3%
Five	1%	**87%**	4%	8%
Six	3%	10%	**83%**	4%
Seven	2%	8%	2%	**88%**
Overall Accuracy = 86.69%				

Table 4.13: *Confusion matrix for breathing rate recognition (RF, Amplitude + Phase data)*

	RF	RT	J48	LSTM
Amplitude	85%	79%	70%	99%
Phase	88%	82%	69%	99%
Amp + Phase	87%	80%	70%	99%
Global Mean Accuracy = 83.92% (ML methods)				

Table 4.14: *Overall average classification accuracy for breathing rate recognition*

Several insights emerge from these results. First, Random Forest shows strong and balanced performance across all breathing rates. Using amplitude-only features (Table 4.11), RF achieves an overall accuracy of 85.46%, but misclassifications increase for slower rates: the 6 s and 7 s cycles are more frequently confused with adjacent classes (e.g., 6 s often predicted as 5 s). This reflects the challenge of distinguishing very low-frequency signals in the presence of CSI noise.

When using phase difference features (Table 4.12), the performance improves to 87.77%. Phase data prove more stable and better suited to capture the small, periodic displacements of the chest wall. Notably, the confusion between adjacent slow rates (6 s vs. 7 s) is reduced, confirming that phase provides a clearer temporal signature than amplitude alone.

Combining amplitude and phase (Table 4.13) yields 86.69% accuracy, slightly below the phase-only case. This suggests that the additional amplitude information does not significantly enrich the discriminative power and may introduce some variability. Thus, phase difference emerges as the most informative single feature source for breathing analysis.

The global comparison (Table 4.14) provides a broader perspective across all classifiers. Random Forest (RF) consistently outperforms other traditional ML models, reaching 85–88% accuracy depending on the input. Random Tree (RT) performs slightly worse, likely due to its single-tree structure, which lacks the ensemble stability of RF. C4.5 (J48) shows the weakest results (69–70%), struggling with the subtle inter-class differences typical of low-frequency respiratory signals. In contrast, the deep learning model (LSTM) achieves almost perfect recognition (99% in all cases), leveraging its ability to model temporal dependencies and smooth out noise across long sequences.

These findings confirm that phase-based CSI features provide the most reliable representation of respiratory dynamics, outperforming amplitude in terms of robustness and class separability. Random Forest emerges as the most practical and interpretable machine learning option, providing high accuracy without complex tuning or risk of overfitting. However, when near-perfect classification is required — for example, in clinical-grade monitoring — deep temporal models such as LSTM are clearly superior, achieving 99% accuracy across all tested breathing rates.

The confusion patterns also reveal that errors primarily occur between adjacent breathing rates (e.g., 5 s vs. 6 s, 6 s vs. 7 s), where the spectral peaks are close and small fluctuations or noise can shift the estimated dominant frequency. This suggests that improving

frequency resolution and stabilizing the pre-processing stage could further enhance machine learning–based classification.

Overall, the results demonstrate that even with conventional machine learning, CSI can reliably distinguish between normal respiratory rates, but deep learning provides the robustness and precision necessary for real-time, contactless medical applications.

4.5 Limitation of CSI-based sensing. The introduction of RIS for NLoS scenarios

Throughout this phase of the research, Channel State Information (CSI) has proved to be an exceptionally sensitive probe of how wireless signals interact with their surroundings. By analysing both the amplitude and the phase of the channel response across OFDM subcarriers and antenna pairs, we have been able to capture minute variations produced by materials, human movement and even physiological rhythms. The experimental campaign progressed in stages, from identifying static objects, to recognizing dynamic gestures, and finally to detecting low-frequency periodic behavior such as breathing.

Under line-of-sight (LoS) and carefully controlled single-subject conditions, the CSI data were remarkably stable. The amplitude and phase remained consistent from packet to packet, allowing straightforward interpretation and reliable learning through both traditional machine-learning algorithms and deep-learning networks trained on time–frequency features. Quantitatively, the standard deviation of amplitude between adjacent subcarriers stayed below half a decibel, and correlation coefficients regularly exceeded 0.9, confirming that the channel maintained strong spatial coherence. Within this stable regime, breathing signals appeared as narrow spectral lines around 0.2–0.4 Hz in the power spectral density of the differential phase, and detection accuracy in single-subject trials exceeded 95%.

The situation changed noticeably as the experimental geometry became more complex. When the line of sight was obstructed, when multiple people were present, or when the subject moved within the room, the channel lost much of its temporal and spatial regularity. Amplitude fluctuations grew by an order of magnitude, and phase unwrapping was often interrupted by multipath discontinuities. In multi-person sessions, overlapping reflections blurred the spectral signature of respiration and occasionally misled the motion-classification models. Algorithms that had performed almost flawlessly in LoS dropped by roughly 20–30% in accuracy, and deep networks required extra data augmentation or retraining to recover part of their robustness.

These observations underline how strongly CSI-based sensing depends on the physical stability of the propagation environment. Once that stability is compromised, the channel statistics no longer conform to the assumptions implicit in most data-driven models. For this reason, the next step in the study turned towards methods capable of exerting some degree of control over the channel itself. A Reconfigurable Intelligent Surface (RIS) was therefore introduced as a way to enhance it under non-line-of-sight (NLoS) conditions. By adaptively reflecting and phase-shaping the incoming waves, an RIS can strengthen otherwise weak links and restore part of the coherence lost in multipath propagation. The following chapter explores this approach in greater detail, describing the MATLAB-based simulator used to model RIS-assisted scenarios and the first results obtained in preparation for hardware implementation.

4.6 Reconfigurable Intelligent Surface and MATLAB-Based Simulator

The introduction of Reconfigurable Intelligent Surfaces (RIS) marked a natural evolution of the research activity described in the previous chapters. The experimental results based on CSI alone had demonstrated excellent performance in line-of-sight (LoS) conditions, but also exposed a clear limitation in non-line-of-sight (NLoS) scenarios, where the received signal became weak and unstable due to multipath fading and shadowing. To mitigate this degradation, I started studying RIS as a controllable propagation interface designed to reconfigure the wireless environment [27, 28, 29, 30]. This allows the reflection and phase of the incident electromagnetic field to be adjusted in real time, enabling the received wave to be reinforced constructively at the receiver.

An RIS can be modelled as an N-element planar array, where each element introduces a programmable phase shift θ_n on the impinging signal. Considering a transmitter (access point, AP) and a receiver (user equipment, UE), the total baseband channel for subcarrier k becomes

$$H_k^{(\mathrm{tot})} = H_k^{(\mathrm{dir})} + \mathbf{h}_{\mathrm{rx},k}^{\top} \boldsymbol{\Phi}_k \mathbf{h}_{\mathrm{tx},k}, \tag{4.12}$$

where $H_k^{(\mathrm{dir})}$ denotes the direct AP–UE path, $\mathbf{h}_{\mathrm{tx},k}$ and $\mathbf{h}_{\mathrm{rx},k}$ are the AP–RIS and RIS–UE channel vectors respectively, and $\boldsymbol{\Phi}_k = \mathrm{diag}(e^{j\theta_{1,k}}, \ldots, e^{j\theta_{N,k}})$ defines the RIS configuration matrix [31]. By properly selecting $\theta_{n,k}$, the cascaded reflected signal can be aligned in

phase with the direct path, yielding a power gain that ideally scales with N^2 under far-field conditions.

From a sensing perspective, such coherent reinforcement has two important consequences: it improves the received SNR in NLoS conditions, and it stabilises the amplitude and phase statistics of CSI, thereby restoring the conditions necessary for effective breathing and motion recognition [32]. To test this principle experimentally, a dedicated MATLAB simulator was used, following and extending the reference structure presented in the MathWorks documentation.

4.6.1 MATLAB Simulation Environment

The simulator builds upon the `Phased Array System Toolbox`, which provides the necessary channel modelling primitives for free-space propagation and array geometries. Three `phased.FreeSpace` links are instantiated to model the direct path (LoS), the AP–RIS hop, and the RIS–UE hop [33].

The transmitted waveform consists of a simple BPSK baseband sequence, stored in `xt`, which provides a controlled probe of the propagation environment under deterministic conditions. The parameter `Ncparam` defines the set of RIS configurations to be tested, corresponding to different numbers of reflecting elements N. This allows the simulator to examine how the received signal strength evolves as the RIS aperture increases. Each instance of the `phased.FreeSpace` object represents a linear free-space channel, modelled according to the Friis transmission equation and accounting for path loss, phase delay, and the frequency-dependent attenuation of the signal.

The transmitter (AP), the RIS, and the receiver (UE) are positioned in three-dimensional Cartesian coordinates, forming a simple triangular layout on the horizontal plane. All propagation channels share the same sample rate and carrier frequency, ensuring that their phases remain consistent and that the composite signal at the receiver is coherently combined across paths.

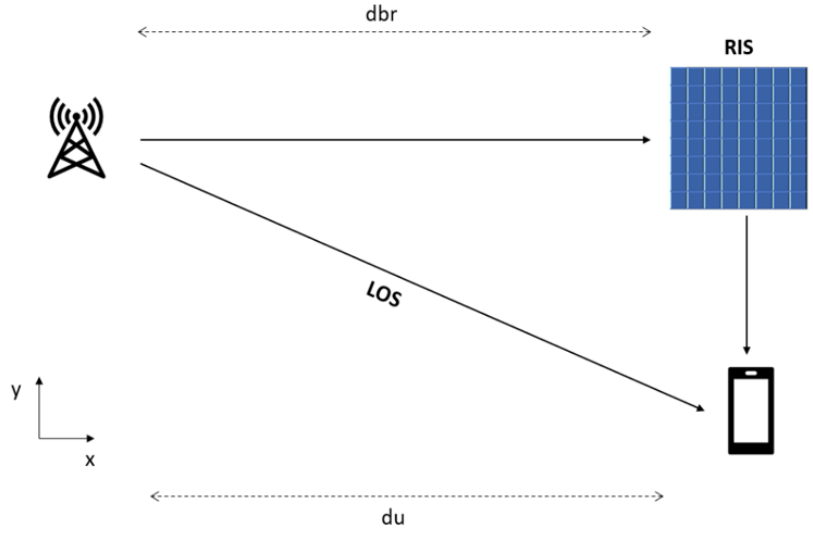

Figure 4.21: *System geometry used in the RIS-assisted simulation. The access point (AP) transmits both a direct line-of-sight (LoS) signal and a reflected signal via the reconfigurable intelligent surface (RIS). The parameters d_{br} and d_u denote the AP–RIS and AP–user distances, respectively.*

The main control logic is handled by the function `helperRISSimulation`, which performs the entire simulation sequence. For each selected number of RIS elements, the function computes three distinct propagation scenarios:

1. **LoS-only:** a direct transmission from the AP to the UE, without any RIS contribution;
2. **RIS-only:** transmission via the RIS, with the direct link intentionally blocked;
3. **LoS+RIS:** simultaneous presence of both the direct and the reflected components.

In each case, the transmitted waveform is propagated through the respective channel configuration, and the received signal is calculated. When the RIS is active, each reflecting element is assigned a phase shift that compensates for the total propagation delay along the AP–RIS–UE path, steering the reflected wavefront toward the receiver. The optimal phase of the n-th element can be expressed as

$$\theta_n^{\star} = -\frac{2\pi f_c}{c}\left(d_n^{(\text{AP–RIS})} + d_n^{(\text{RIS–UE})}\right) + \arg(H^{(\text{dir})}), \tag{4.13}$$

where $d_n^{(\text{AP–RIS})}$ and $d_n^{(\text{RIS–UE})}$ denote the distances between the element and the AP or UE, respectively. This choice of phase ensures that all reflections arrive in phase at the receiver, thereby maximising constructive interference and improving the received power.

The simulation outputs three SNR curves—one for each propagation condition (LoS-only, RIS-only, and LoS+RIS). As anticipated, the SNR grows with the number of RIS elements, approximately following a quadratic dependence on N under ideal conditions. For small apertures, the effect of the RIS is marginal, but beyond roughly fifty to sixty elements the reflected signal becomes sufficiently coherent to yield a marked improvement, with the RIS-only configuration approaching the LoS+RIS performance.

In the extended version of the simulator developed for this work, additional control routines were implemented to vary not only the number of elements but also the spatial position of the RIS along the x-axis and the incidence and reflection angles between the AP, RIS, and UE. For each value of x_{RIS}, the program recomputes all path distances and updates the phase configuration of the RIS elements. The resulting SNR distribution,

$$\text{SNR}(x_{\text{RIS}}, \theta_{\text{inc}}, \theta_{\text{ref}}) = 10 \log_{10}\left(\frac{\mathbb{E}[|y(t)|^2]}{\sigma_n^2}\right), \tag{4.14}$$

was then evaluated across a grid of spatial and angular parameters to assess the RIS's ability to enhance the received signal in different non-line-of-sight (NLoS) configurations.

The first simulation campaign was conceived as an exploratory exercise, aimed at understanding how the MATLAB *Reconfigurable Intelligent Surface* (RIS) simulator operates and what kind of performance can be expected based on established theoretical models. In this initial phase, only a single simulation approach was tested, focusing on a simplified geometry that allowed for a clear interpretation of the results and for validation of the simulation workflow before extending the analysis to more complex scenarios.

The results confirmed the theoretically expected behaviour of RIS-assisted systems. The signal-to-noise ratio (SNR) curves exhibited a distinct and consistent growth trend as the number of RIS elements increased. For small array sizes, the contribution of the reflected component was minimal; however, once the surface comprised approximately fifty or more elements, the coherent superposition of the reflected wavefronts produced significant gains, approaching those of the direct line-of-sight (LoS) link. Under ideal phase alignment conditions, the simulated RIS achieved SNR improvements of up to 15–20 dB, in close agreement with analytical predictions and previously published experimental results.

A more detailed parametric analysis was performed by varying the spatial position of the RIS along the x-axis, as well as the incidence and reflection angles between the transmitter, the surface, and the receiver.

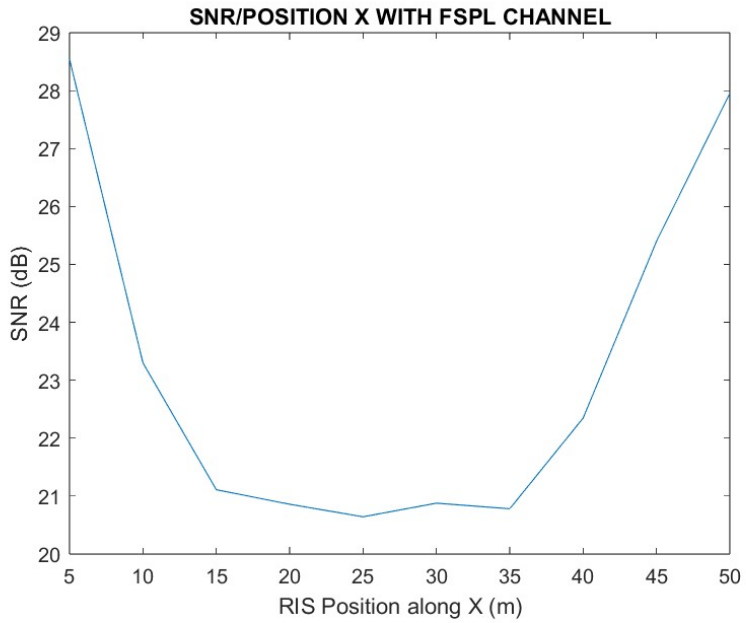

Figure 4.22: *Variation of the signal-to-noise ratio (SNR) as the RIS position changes along the x-axis under an FSPL channel model. The curve shows the existence of an optimal placement region where constructive interference between direct and reflected components maximises SNR, while misalignment leads to degradation.*

The resulting SNR maps revealed a well-defined region of constructive interference, surrounded by areas of reduced gain corresponding to phase misalignment. These spatial patterns illustrate the geometric basis of coherent field synthesis: as the surface is displaced from its optimal location, the differences in propagation path lengths between the RIS elements increase. Identifying such optimal configurations provides not only a physical understanding of RIS-assisted propagation but also valuable guidelines for practical placement in indoor sensing scenarios.

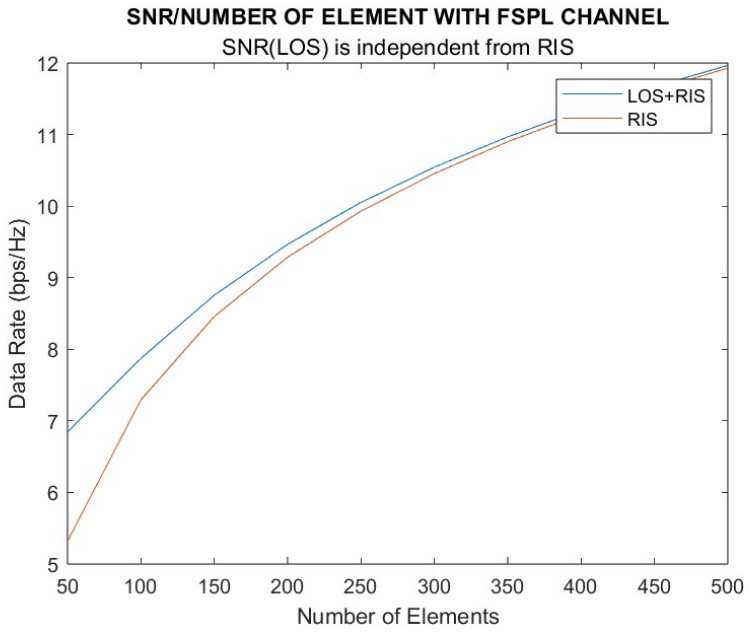

Figure 4.23: *Data rate performance as a function of the number of RIS elements under a free-space path loss (FSPL) channel. The comparison between the LoS+RIS and RIS-only configurations highlights how increasing the number of reflecting elements enhances the overall channel capacity. The line-of-sight (LoS) component remains constant and independent of the RIS.*

As shown in Fig. 4.23, the achievable data rate increases steadily with the number of RIS elements, confirming the theoretical expectation that a larger reflecting surface enhances the effective channel gain through coherent combination of the reflected wavefronts. At small array sizes, the contribution of the RIS remains limited, since the individual reflections are too weak to provide significant constructive interference at the receiver. As the surface grows, however, the cumulative effect of the elements strengthens the reflected link, leading to a near-logarithmic increase in data rate that eventually saturates when the reflected and direct components reach comparable power levels. The comparison between the LoS+RIS and RIS-only configurations highlights that, while the direct line-of-sight (LoS) path dominates at small RIS sizes, the reflected path becomes increasingly relevant as the number of controllable elements rises, approaching the performance of the direct channel. This behaviour demonstrates the potential of large-scale reconfigurable surfaces to significantly improve spectral efficiency in propagation-limited environments.

In addition to the quantitative SNR trends obtained in the present simulations, the results suggest potential avenues for exploiting technology as an active component in future wireless systems. Although the RIS was assumed to be purely passive in the simulated configuration, its capacity to enhance the SNR through spatial reconfiguration

indicates its potential role in actively stabilising wireless channels. In non-line-of-sight (NLoS) conditions, where multipath propagation leads to rapid temporal fluctuations in channel state information (CSI), an actively controlled RIS could act as a reconfigurable reflector to reintroduce a dominant, deterministic path component into the received signal. This would increase the average received power and regulate the statistical behaviour of amplitude and phase, thereby improving channel coherence and predictability. Thus, the RIS evolves from a passive power-boosting device into an intelligent, environment-aware interface that can dynamically shape and condition the propagation medium — a feature that could be particularly valuable for robust sensing and communication in highly dynamic or obstructed NLoS scenarios.

These preliminary results can be regarded as proof of concept for integrating RIS technology into CSI-based sensing and communication frameworks. MATLAB simulations show that coherent, phase-aligned reflections can mitigate multipath fading, improve signal coherence and enhance the detectability of low-frequency components associated with human motion or respiration. Although the present analysis is restricted to a simplified simulation-based configuration, it establishes a quantitative and conceptual baseline for subsequent experimental work. The observed SNR behaviour provides valuable insights for designing future tests involving programmable RIS prototypes and commercial Wi-Fi transceivers, facilitating a systematic comparison between theoretical predictions and empirical performance. Ultimately, these insights will facilitate the integration of RIS-assisted propagation models into CSI-driven machine learning architectures, paving the way for adaptive, environment-aware wireless sensing systems that can operate effectively even under challenging non-line-of-sight (NLoS) conditions.

CHAPTER 5

General conclusion

This research has undertaken a comprehensive exploration of *Channel State Information* (CSI) as a fine-grained and physically interpretable descriptor for contactless monitoring of human presence, movement, and physiological activity. The study began with the observation that modern Wi-Fi networks, beyond their communication purpose, continuously probe the surrounding environment through electromagnetic waves whose amplitude and phase encode the interaction with objects and living bodies. By analysing these subtle modulations in CSI, it becomes possible to infer dynamic properties of the scene—such as position, motion, or respiration—without requiring the person to wear or carry any sensor, demonstrating that commodity Wi-Fi infrastructure can be repurposed as a non-invasive sensing platform.

The thesis is organised into four main chapters, each reflecting a progressive deepening of the topic. The *first chapter* introduces the general context of wireless sensing within the framework of Internet of Things (IoT) applications for healthcare and smart environments, outlining the motivations, objectives, and overall structure of the research. The *second chapter* presents an extensive review of the state of the art in CSI-based sensing, highlighting the main contributions and limitations of existing systems for object recognition, motion analysis, and vital sign monitoring. The *third chapter* is devoted to the methodological framework, detailing the data acquisition setup, the processing pipeline for amplitude and phase extraction, and the use of machine learning and deep learning architectures for feature classification. Finally, the *fourth chapter* reports the experimental results, obtained through a combination of controlled laboratory tests and MATLAB simulations, including the first implementation of a Reconfigurable Intelligent Surface (RIS) simulator for signal enhancement in non-line-of-sight (NLoS) conditions, explicitly addressing one of the main limitations identified in current CSI-based sensing systems.

The investigation was conceived as a gradual progression from theory to experiment. The first stage was devoted to consolidating a rigorous theoretical and methodological foundation. Following an extensive review of the scientific literature on wireless sensing,

a series of MATLAB simulations and controlled measurements were designed to observe how CSI responds to static and dynamic perturbations of the propagation channel. These preliminary tests served not only to validate the acquisition pipeline but also to highlight the sensitivity of CSI to even minute spatial or temporal variations in the environment, confirming its suitability for fine-grained sensing tasks. Building on these results, the research was organised around three main thematic axes: (i) the identification of materials and objects through their electromagnetic signatures, (ii) the recognition of human movements and gestures, and (iii) the monitoring of respiratory activity as a proxy for vital signs.

Each of these domains required a specific adaptation of the acquisition and processing framework. The first experiments, centred on material and object identification, exploited the spatial diversity of the channel to distinguish between different dielectric and conductive surfaces in line-of-sight (LoS) conditions. By analysing amplitude attenuation and phase delay across multiple subcarriers, it was possible to classify objects according to their composition and geometry, achieving high reproducibility across trials and classification accuracies close to 98%. The second group of studies addressed body motion recognition. Here, distributed antennas and fine-grained time–frequency representations were used to capture transient fluctuations in CSI caused by gestures and displacements. The time-evolution of subcarrier phase and amplitude revealed repeatable patterns linked to specific movements, demonstrating the feasibility of CSI-based behaviour recognition with recognition accuracies exceeding 95% in controlled scenarios.

The final and most delicate phase of experimentation focused on respiration monitoring, which demanded both higher sensitivity and temporal stability. By isolating the low-frequency components of the amplitude and phase's variations, corresponding to the periodic expansion and contraction of the thoracic cavity, the system was able to extract breathing rates in real time with mean estimation errors below 0.05 Hz for stationary subjects. This step marked a significant transition from macroscopic motion analysis to the detection of micro-movements in the sub-centimetric range. The results confirmed that, under controlled conditions, the power spectral density of the CSI phase exhibits narrow peaks around the natural breathing frequency, validating the method's potential for unobtrusive physiological sensing.

Across all these experimental stages, the research integrated signal processing, feature engineering, and machine learning (ML) methods to assess the limits and possibilities of commercial Wi-Fi hardware as a sensing platform. The findings consistently demonstrated that CSI constitutes a dense and multidimensional representation of the radio channel,

capable of encoding fine spatial and temporal variations with remarkable precision. In particular, when appropriately filtered and normalised, CSI captures both large-scale variations associated with body displacement and small-scale oscillations induced by respiratory motion, enabling a unified sensing framework across heterogeneous tasks. The ability to extract such diverse levels of information from a single physical quantity underscores the potential of CSI as a versatile and cost-effective tool for contactless monitoring, paving the way toward non-invasive, continuous observation of human activity in everyday environments.

5.0.1 Synthesis of the Main Results

The results obtained throughout this research confirm that *Channel State Information* (CSI) can be effectively employed as a high-resolution observable for contactless sensing in indoor environments. Across the different phases of the work—object recognition, movement analysis, and physiological monitoring—CSI has demonstrated a remarkable ability to encode the micro-variations of the electromagnetic field caused by materials and human activity, with experimentally validated performance levels comparable to those reported in the state of the art. Each experimental stage contributed to consolidating a unified view of how amplitude and phase statistics evolve under varying propagation conditions, and how these can be exploited by both machine learning (ML) and deep learning (DL) models to infer meaningful information about the environment and its occupants.

Object and Material Identification

The first group of experiments addressed the identification of objects and materials using differential CSI features. The dataset included aluminium, glass, metals, nylon, paper, plastic, wood, placed in controlled line-of-sight (LoS) configurations between a Wi-Fi access point (AP) and a receiver. The pre-processing pipeline included amplitude normalisation, wavelet denoising, and subcarrier selection to suppress multipath artefacts. ML classifiers such as Support Vector Machines (SVM) and Random Forests were trained on the extracted amplitude–phase ratios and inter-antenna phase differences, allowing a systematic evaluation of the discriminative power of CSI-derived features.

The results showed that materials with different electromagnetic constants produce distinct CSI signatures: conductive materials (such as metals) caused strong attenuation

and rapid phase transitions, while dielectrics introduced frequency-dependent phase delays. The SVM model achieved an average classification accuracy of *98%* across ten material categories, maintaining above 90% even under slight changes of container shape or measurement distance, confirming both robustness and repeatability of the proposed approach. The confusion matrices revealed that the remaining misclassifications occurred mainly between materials of similar permittivity (e.g., glass and water). These results are consistent with those reported in *WiMi* and *WiMate* systems, validating the reproducibility of material identification using commodity Wi-Fi NICs and positioning the obtained performance at the level of state-of-the-art solutions.

Activity recognition

The second experimental phase focused on the recognition of dynamic movements. CSI was collected using distributed antennas positioned around a confined indoor area, with subjects performing a predefined set of gestures (walking, sitting, running, lying, falling, standing). The CSI streams were transformed into spectrograms through short-time Fourier analysis and fed into different model families. Classical ML classifiers (SVM, J48 and Random Forest) were used for baseline comparisons, while recurrent (LSTM) network is adopted to capture temporal dependencies and spectral evolution, enabling a direct comparison between feature-based and sequence-based learning paradigms.

The results confirmed that CSI amplitude and phase dynamics are highly correlated with the type and direction of motion. In single-person LoS conditions, the LSTM achieved an average recognition accuracy of *96%*, surpassing the best traditional model (SVM, 89%) and demonstrating the advantage of temporal modelling for dynamic activities. LSTM networks, trained on time-domain phase sequences, proved particularly robust to temporal misalignment and were able to distinguish between repetitive and non-repetitive gestures. The spectral energy distribution of CSI magnitude also revealed characteristic Doppler signatures: walking produced distinct low-frequency components (0.5–2 Hz), whereas arm gestures caused higher-frequency modulations up to 4 Hz, providing a physical interpretation of the learned features.

When tested under more complex NLoS or multi-person scenarios, accuracy decreased by approximately 20–25%, primarily due to phase decorrelation and reduced coherence among subcarriers. Nevertheless, combining amplitude-based and differential-phase features partially mitigated the degradation, maintaining classification accuracy at acceptable

levels even under adverse propagation conditions, confirming that multi-antenna diversity and spatial filtering remain crucial for generalisation.

Breath rate monitoring

The final and most challenging phase involved extracting respiratory signals from CSI data. In this setting, the subjects were positioned in front of the transmitting antennas at distances ranging from 1.0 to 2.5 metres. Both static and moving conditions were tested: first with the subjects seated or standing and breathing normally, and then while walking. The unwrapped and detrended phase signals were analysed using power spectral density estimation (the Welch method), with the breathing frequency typically falling within the 0.2–0.4 Hz band, corresponding to normal adult respiration rates.

For stationary subjects, the spectral peak corresponding to respiration was clearly visible in the differential phase between the receiving antennas. Shallow neural models trained on temporal–spectral features achieved detection accuracies above 94%, with mean absolute errors in estimated breathing rate below 0.05 Hz, indicating clinically relevant precision for non-contact monitoring. When CNNs were employed that had been trained on spectrogram inputs, performance reached 97% accuracy, demonstrating strong robustness to small postural variations. In contrast, when the subject began to move, the spectral peaks broadened, reflecting the increased non-stationarity of the propagation channel and highlighting the intrinsic limits of CSI-based respiration sensing under dynamic conditions. These observations confirmed the physical limitations of CSI-based respiration monitoring when the reflective surfaces of the body move unpredictably or the line of sight (LoS) component is lost.

General Trends

These experimental results collectively demonstrate that CSI provides a comprehensive and informative representation of the wireless channel; however, its effective utilisation hinges critically on propagation stability. In line-of-sight (LoS) scenarios with a single subject, the amplitude and phase of each subcarrier evolve smoothly over time, yielding high inter-tone correlation (above 0.9) and low amplitude variance (below 0.5 dB), conditions under which learning-based models achieve their highest accuracy. Under these conditions, both machine learning (ML) and deep learning (DL) models learn stable mappings from

CSI statistics to semantic classes, achieving accuracies comparable to those obtained in specialised radar systems.

However, when moving to non-line-of-sight (NLoS) configurations or multi-person environments, the temporal coherence of CSI collapses. Amplitude variance increases by an order of magnitude and inter-subcarrier correlation drops sharply, causing signal distortions that even deep architectures struggle to disentangle, leading to systematic performance degradation across all sensing tasks. This degradation is not solely due to model overfitting, but also to the intrinsic instability of the channel, which is constantly changing as objects and people move within it.

Despite these limitations, a combination of differential features, multi-antenna diversity and temporal filtering was capable of maintaining an acceptable level of accuracy, even under mild NLoS or dynamic conditions, demonstrating partial resilience of CSI-based sensing when supported by appropriate signal processing. These results support the idea that, while CSI is highly sensitive to channel geometry, it can still provide reliable information when aided by appropriate preprocessing and model regularisation.

Overall, this body of work concludes that CSI-based sensing is a technically viable and scientifically sound approach to passive monitoring, provided its dependence on propagation geometry is recognised and mitigated through both algorithmic and physical-layer solutions. Reconfigurable Intelligent Surfaces (RIS) technology naturally follows from this observation, as it enables the controlled manipulation of reflection paths. This promises to restore the coherence and signal-to-noise ratio (SNR) necessary for robust CSI-based inference, even in non-line-of-sight or highly dynamic environments.

Integration of Reconfigurable Intelligent Surfaces

The introduction of *Reconfigurable Intelligent Surfaces* (RIS) represented a turning point in this thesis. By enabling programmable control over the phase of reflected electromagnetic waves, a RIS can reconfigure the propagation paths and enhance the received signal power in areas that would otherwise suffer from deep fading, effectively acting as a controllable extension of the wireless channel.

To verify the feasibility of this approach, a dedicated MATLAB simulator was used using the `Phased Array System Toolbox`. The simulator models the access point (AP), RIS, and receiver (UE) as distinct propagation nodes and computes the total received field by combining the direct and reflected components. Each RIS element is assigned a tunable phase shift, and by aligning these phases with the direct path, constructive

interference can be achieved at the receiver, resulting in SNR gains of up to 15–20 dB in simulated NLoS conditions.

Through systematic sweeps of the RIS position, number of elements, and angular orientation, the simulations revealed measurable improvements in the signal-to-noise ratio (SNR). From the perspective of CSI-based sensing, these results are significant. The RIS does not merely boost received power but also stabilises the temporal coherence of the channel, reducing random fluctuations and restoring the structured variability that ML and DL models rely on, thereby extending the operational range of CSI-based sensing beyond favourable LoS scenarios. By providing a controllable propagation layer, the RIS effectively extends the applicability of CSI-based sensing to previously inaccessible NLoS environments.

5.0.2 Future Work and Perspectives

The results presented in this thesis lay the foundations for several promising research directions that extend beyond the current stage of CSI-based sensing. Although the experiments and simulations demonstrated the feasibility of recognising materials, movements, and respiratory patterns through Wi-Fi channel analysis, they also revealed fundamental challenges linked to channel instability, multipath propagation, and limited frequency resolution, which currently constrain scalability and generalisation across heterogeneous environments. The next phase of work should therefore focus on consolidating these findings into more generalisable, adaptable, and physically grounded systems.

A first line of development concerns the integration of Reconfigurable Intelligent Surfaces (RIS) into experimental setups. The MATLAB simulator provided in this work has shown that even under idealised conditions, a passive and programmable surface can significantly increase the received signal power and stabilise the amplitude and phase statistics of the CSI, thereby improving both signal-to-noise ratio and temporal coherence. However, this proof-of-concept must be verified on real hardware. The practical implementation will involve the design of a controllable RIS prototype equipped with programmable phase shifters, connected to the Wi-Fi transmission chain through low-cost microcontrollers. Such a system will make it possible to actively manipulate the propagation geometry, aligning reflected paths to reinforce the desired signal components and mitigate destructive interference under realistic deployment constraints.

In particular, one objective will be to test how RIS-assisted propagation affects breathing detection and movement classification in non-line-of-sight (NLoS) settings. If the

simulated gains of 15–20 dB can be replicated experimentally, RIS technology could restore the signal-to-noise ratio (SNR) necessary for stable physiological monitoring even in the presence of obstacles or multiple subjects, enabling sensing scenarios that are currently infeasible with conventional Wi-Fi deployments. Beyond its primary function as a reflective amplifier, the RIS can be envisaged as an active, reconfigurable sensing component capable of directing and shaping the wireless field selectively. By dynamically adjusting its reflection pattern, it can monitor specific spatial regions or individuals, effectively transforming the propagation medium itself into a programmable sensor and introducing a new degree of freedom in wireless sensing design. This approach would be a significant step towards creating truly adaptive and perceptive wireless environments.

A second development is the adoption of ESP32-class devices and full-subcarrier acquisition. While the Intel 5300 NIC, which is used in most CSI research, only exposes 30 subcarriers, the new generation of ESP32 chips provides access to all 64 (or more) OFDM tones. This improves the frequency-domain resolution and robustness of the analysis. Exploiting the complete subcarrier set will enable more accurate estimation of phase slopes and amplitude decay, which are critical for identifying materials or tracking subtle motions, particularly in noisy or multipath-rich environments. Furthermore, using open-source firmware on these platforms enables the integration of real-time RIS control loops and local machine learning inference directly on embedded hardware, paving the way for closed-loop, self-adaptive sensing systems. This paves the way for compact, self-contained sensing modules.

A third perspective involves strengthening the analytical and mathematical modelling of CSI variations. Current work has relied primarily on empirical observation and machine learning to detect correlations between CSI dynamics and physical phenomena. Future research could incorporate simplified electromagnetic and statistical models to provide analytical baselines for comparison and to formalise the link between physical motion and observed CSI perturbations. For example, small-scale motion could be modelled as a phase modulation process driven by periodic displacements, and breathing could be represented as a quasi-harmonic perturbation of path length. Analytical derivations of the expected power spectral density or signal variance could be used to validate or calibrate ML and DL estimators, quantifying their bias and variance relative to theoretical limits and improving model interpretability. This hybrid approach, which links physics-based reasoning with data-driven learning, would enhance both interpretability and transferability across environments.

Another avenue for expansion is the introduction of multi-modal sensing frameworks, where CSI data are fused with complementary information from inertial, acoustic, or visual sensors. By correlating wireless signals with other physical modalities, it will be possible to isolate motion-induced perturbations more precisely and compensate for environmental noise that cannot be resolved through radio-frequency sensing alone. In the context of healthcare, this fusion could support continuous, unobtrusive monitoring of multiple physiological parameters, such as heart rate, posture, and respiratory rhythm, within a single integrated platform, increasing reliability and clinical relevance.

Finally, this thesis also opens up a methodological reflection on the role of physical modelling in data-driven sensing. While deep learning has proven extremely effective in discovering patterns from complex data, the stability and interpretability of these models still depend on the underlying physics of the signal. A balanced approach, where analytical models guide network design and training, can yield systems that are not only accurate but also explainable and transferable across devices, environments, and use cases. In this sense, the future of CSI-based sensing may lie in the convergence of three paradigms: physics-based modelling, machine learning, and reconfigurable propagation.

In conclusion, this doctoral research has demonstrated that Wi-Fi Channel State Information can serve as a versatile and precise tool for contactless sensing, capable of capturing environmental dynamics with high sensitivity. At the same time, it has revealed the dependence of this technology on propagation conditions and channel geometry, highlighting the need for active control through RIS and for deeper theoretical understanding to ensure robustness and scalability. The perspectives outlined here, experimental validation of RIS prototypes, full-subcarrier data acquisition, analytical modelling, and integration with intelligent wireless systems, constitute a coherent roadmap for the continuation of this work. The ultimate goal is to turn ordinary wireless infrastructures into distributed sensing networks that can perceive, interpret, and interact with the physical world seamlessly and intelligently, bridging the gap between communication and perception.

Acknowledgements

Completing this doctoral journey has been a deeply formative experience, both professionally and personally. Firstly, I would like to express my sincere gratitude to my supervisor, Professor Fabio Lavagetto, for his constant guidance, constructive criticism and intellectual generosity. His ability to balance scientific rigour with curiosity has been a constant source of inspiration throughout my research.

I am also very grateful to my co-supervisors, Prof. Igor Bisio, Prof. Andrea Sciarrone, and Prof. Chiara Garibotto, for their patience and insight, which helped me refine many of the ideas that shaped this work. Special thanks also go to my colleagues Sheida, Matteo, Halar, Mohammad, Junaid, and everyone in the DSP lab for their collaboration and for the moments of research and friendship we shared. Working alongside such talented and motivated individuals has made these years both challenging and rewarding.

On a more personal note, I would like to thank my family for their unconditional support, encouragement and understanding, particularly during the most challenging periods of my PhD. Their presence and confidence have been my greatest source of strength.

I would like to express my deepest gratitude to my special people: Luca, Martina and Davide. Without them, I would not have been able to carry on. Thank you for always being there for me, for brightening up my days and planning them with me, for watching over me and for believing in me. Thank you, my students for your affection.

I would like to dedicate this thesis to all those who believe in the power of education and research as a means of improving lives and broadening horizons.

Publication record

1. International Journal Papers

[1] I. Bisio, C. Fallani, C. Garibotto, A. Grattarola, F. Lavagetto, A. Sciarrone, S. Zappatore, M. Zerbino, "Performance Evaluation of Machine/Deep Learning-Based Object Recognition Techniques Leveraging Channel State Information Using a Real Testbed," in IEEE Access, vol. 12, pp. 98680-98692, 2024, doi: 10.1109/ACCESS.2024.3428612.
[2] I. Bisio, C. Fallani, C. Garibotto, H. Haleem, F. Lavagetto, M. Hamedani, A. Schenone, A. Sciarrone, M. Zerbino, "AI-Enabled Internet of Medical Things: Architectural Framework and Case Studies," in IEEE Internet of Things Magazine, vol. 8, no. 2, pp. 121-128, March 2025, doi: 10.1109/IOTM.001.2400150.

2. International Conference Papers

[3] I. Bisio, C. Fallani, C. Garibotto, F. Lavagetto, A. Sciarrone and M. Zerbino, "Analysis of CSI-based Human Activity Recognition for Contactless Patients Monitoring," GLOBECOM 2024 - 2024 IEEE Global Communications Conference, Cape Town, South Africa, 2024, pp. 438-443, doi: 10.1109/GLOBECOM52923.2024.10901592.
[4] I. Bisio, C. Fallani, C. Garibotto, A. Grattarola, F. Lavagetto, A. Sciarrone, M. Zerbino, "AI-Driven Estimation of Breathing Frequency Through CSI Analysis," ICC 2025 - IEEE International Conference on Communications, Montreal, QC, Canada, 2025, pp. 934-939, doi: 10.1109/ICC52391.2025.11161454.

Participation in Research Projects

Research project 1, HIPPOCRATES project, Research Projects of National Interest (PRIN2022) call, Project No. 2022YSRWEZ.

Description: The breath monitoring component of this research was developed within the framework of the RAISE – Robotics and AI for Socioeconomic Empowerment project, funded by the European Union–NextGenerationEU. The project aims to integrate robotics, artificial intelligence, and advanced sensing technologies to promote innovation and social well-being through intelligent, data-driven systems. This work was also partially supported by the Italian Ministry of University and Research (MUR) under the HIPPOCRATES project.

Activity: the research contributed to this effort by developing algorithms and experimental methodologies for respiratory signal extraction.

Teaching and other activities

Reviewing activity for the following international journals: IEEE Internet of Things Journal.

Bibliography

[1] Y. Hu, F. Zhang, C. Wu, B. Wang, and K. J. Ray Liu, "A wifi-based passive fall detection system," in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 1723–1727.

[2] M. Ding, Q. Wang, and C. Wang, "Deep transfer learning for actions recognition with wifi signals," in *2022 IEEE 10th International Conference on Information, Communication and Networks (ICICN)*, 2022, pp. 406–411.

[3] M. Sulaiman, S. A. Hassan, and H. Jung, "True detect: Deep learning-based device-free activity recognition using wifi," in *2020 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, 2020, pp. 1–5.

[4] X. Cheng, B. Huang, and J. Zong, "Device-free human activity recognition based on gmm-hmm using channel state information," *IEEE Access*, vol. 9, pp. 76 592–76 601, 2021.

[5] D. Halperin, "Simplifying the configuration of 802.11 wireless networks with effective snr," 01 2013.

[6] F. Chao, J. Xiong, L. Chang, J. Wang, X. Chen, D. Fang, and Z. Tang, "Wimi: Target material identification with commodity wi-fi devices," 07 2019, pp. 700–710.

[7] Y. Gu, Y. Zhu, J. Li, and Y. Ji, "Wimate: Location-independent material identification based on commercial wifi devices," in *2021 IEEE Global Communications Conference (GLOBECOM)*, 2021, pp. 01–06.

[8] C. Shi, T. Zhao, Y. Xie, T. Zhang, Y. Wang, X. Guo, and Y. Chen, "Environment-independent in-baggage object identification using wifi signals," 08 2021.

[9] C. Wang, J. Liu, Y. Chen, H. Liu, and Y. Wang, "Towards in-baggage suspicious object detection using commodity wifi," 05 2018, pp. 1–9.

[10] Z. Yang, Z. Zhou, and Y. Liu, "From rssi to csi: Indoor localization via channel response," *ACM Comput. Surv.*, vol. 46, no. 2, Dec. 2013. [Online]. Available: https://doi.org/10.1145/2543581.2543592

[11] X. Wang, L. Gao, and S. Mao, "Csi phase fingerprinting for indoor localization with a deep learning approach," *IEEE Internet of Things Journal*, vol. 3, pp. 1–1, 12 2016.

[12] G. Wang, Y. Zou, Z. Zhou, K. Wu, and L. M. Ni, "We can hear you with wi-fi!" ser. MobiCom '14. New York, NY, USA: Association for Computing Machinery, 2014, p. 593–604. [Online]. Available: https://doi.org/10.1145/2639108.2639112

[13] C. Wu, Z. Yang, Z. Zhou, K. Qian, Y. Liu, and M. Liu, "Phaseu: Real-time los identification with wifi," in *Proceedings of the IEEE conference on computer communications (INFOCOM)*, 2015.

[14] Y. Xie, Z. Li, and M. Li, "Precise power delay profiling with commodity wifi," in *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*, ser. MobiCom '15. New York, NY, USA: Association for Computing Machinery, 2015, p. 53–64. [Online]. Available: https://doi.org/10.1145/2789168.2790124

[15] X. Wang, C. Yang, and S. Mao, "Phasebeat: Exploiting csi phase data for vital sign monitoring with commodity wifi devices," in *2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS)*, 2017, pp. 1230–1239.

[16] Z. Shi, J. A. Zhang, R. Xu, and G. Fang, "Human activity recognition using deep learning networks with enhanced channel state information," in *2018 IEEE Globecom Workshops (GC Wkshps)*, 2018, pp. 1–6.

[17] Z. Tang, A. Zhu, Z. Wang, K. Jiang, Y. Li, and F. Hu, "Human behavior recognition based on wifi channel state information," in *2020 Chinese Automation Congress (CAC)*, 2020, pp. 1157–1162.

[18] H. Abdelnasser, K. A. Harras, and M. Youssef, "Ubibreathe: A ubiquitous non-invasive wifi-based breathing estimator," in *Proceedings of the 16th ACM International Symposium on Mobile Ad Hoc Networking and Computing*, ser. MobiHoc '15. New York, NY, USA: Association for Computing Machinery, 2015, p. 277–286. [Online]. Available: https://doi.org/10.1145/2746285.2755969

[19] X. Liu, J. Cao, S. Tang, and J. Wen, "Wi-sleep: Contactless sleep monitoring via wifi signals," in *2014 IEEE Real-Time Systems Symposium*, 2014, pp. 346–355.

[20] X. Liu, J. Cao, S. Tang, J. Wen, and P. Guo, "Contactless respiration monitoring via off-the-shelf wifi devices," *IEEE Transactions on Mobile Computing*, vol. 15, no. 10, pp. 2466–2479, 2016.

[21] X. Wang, C. Yang, and S. Mao, "Tensorbeat: Tensor decomposition for monitoring multiperson breathing beats with commodity wifi," *ACM Trans. Intell. Syst. Technol.*, vol. 9, no. 1, Sep. 2017. [Online]. Available: https://doi.org/10.1145/3078855

[22] D. Zhang, Y. Hu, Y. Chen, and B. Zeng, "Breathtrack: Tracking indoor human breath status via commodity wifi," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 3899–3911, 2019.

[23] S. Mosleh, J. B. Coder, C. G. Scully, K. Forsyth, and M. O. A. Kalaa, "Monitoring respiratory motion with wi-fi csi: Characterizing performance and the breathesmart algorithm," *IEEE Access*, vol. 10, pp. 131 932–131 951, 2022.

[24] H. Wang, D. Zhang, J. Ma, Y. Wang, Y. Wang, D. Wu, T. Gu, and B. Xie, "Human respiration detection with commodity wifi devices: do user location and body orientation matter?" in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, ser. UbiComp '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 25–36. [Online]. Available: https://doi.org/10.1145/2971648.2971744

[25] P. Wang, B. Guo, T. Xin, Z. Wang, and Z. Yu, "Tinysense: Multi-user respiration detection using wi-fi csi signals," in *2017 IEEE 19th International Conference on e-Health Networking, Applications and Services (Healthcom)*, 2017, pp. 1–6.

[26] Csitool monitor mode. [Online]. Available: https://www.programmersought.com/article/797254612/

[27] S. Basharat, S. A. Hassan, H. Pervaiz, A. Mahmood, Z. Ding, and M. Gidlund, "Reconfigurable intelligent surfaces: Potentials, applications, and challenges for 6g wireless networks," *IEEE Wireless Communications*, vol. 28, no. 6, pp. 184–191, 2021.

[28] R. Singh, A. Kaushik, W. Shin, G. C. Alexandropoulos, M. Toka, and M. Di Renzo, "Indexed multiple access with reconfigurable intelligent surfaces: The reflection tuning potential," *IEEE Communications Magazine*, vol. 62, no. 4, pp. 120–126, 2024.

[29] V. Degli-Esposti and H. Bertoni, "Evaluation of the role of diffuse scattering in urban microcellular propagation," in *Gateway to 21st Century Communications Village. VTC 1999-Fall. IEEE VTS 50th Vehicular Technology Conference (Cat. No.99CH36324)*, vol. 3, 1999, pp. 1392–1396 vol.3.

[30] C. Liaskos, S. Nie, A. Tsioliaridou, A. Pitsillides, S. Ioannidis, and I. Akyildiz, "A new wireless communication paradigm through software-controlled metasurfaces," *IEEE Communications Magazine*, vol. 56, no. 9, pp. 162–169, 2018.

[31] S. Hu, F. Rusek, and O. Edfors, "Beyond massive mimo: The potential of data transmission with large intelligent surfaces," *IEEE Transactions on Signal Processing*, vol. 66, no. 10, pp. 2746–2758, 2018.

[32] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Communications Magazine*, vol. 58, no. 1, pp. 106–112, 2020.

[33] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface-aided wireless communications: A tutorial," *IEEE Transactions on Communications*, vol. 69, no. 5, pp. 3313–3351, 2021.

IoT solutions for e-Health applications for care's continuity at home

Caterina Fallani