D6.4 Motion detection and discrimination accuracy benchmark test in the test-house prototypes

Grant Agreement Number	101099491
Action Acronym	HOLDEN
Action Title	Ethical Design of Holography with Dense wireless Networks (HOLDEN)
Funding Scheme	HORIZON-EIC-2022-PATHFINDEROPEN-01
Version date of the Annex I against which the assessment will be made	13/12/2022
Start date of the project	1/6/2023
Due date of the deliverable	30/11/2025
Actual date of submission	28/11/2025
Responsible	CNR
Contributors	CNR, ADANT, POLIMI
Dissemination level	Public

European Innovation Council

Funded by the European Union

Funded by the European Union. Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or European Innovation Council and SMEs Executive Agency (EISMEA). Neither the European Union nor the granting authority can be held responsible for them. Grant Agreement No: 101099491.

Authors in alphabetical order

Full Name	Organisation	E-mail
R. Bersan	Adant	riccardo.bersan@adant.com
E. Ferro	Adant	emanuele.ferro@adant.com
S. Kianoush	CNR	sanaz.kianoush@cnr.it
U. Milasheuski	CNR	usevaladmilasheuski@cnr.it
F. Fieramosca	POLIMI	federica.fieramosca@polimi.it
M. D'Amico	POLIMI	michele.damico@polimi.it
D. Piazza	Adant	daniele.piazza@adant.com
V. Rampa	CNR	vittorio.rampa@ieiit.cnr.it
S. Savazzi	CNR	stefano.savazzi@cnr.it

Change History

Version	Date	Status	Author (Company)	Description
1.0	21.10.2025	Initial	CNR	First version
2.0	31.10.2025	Revision	ADANT	Second version
3.0	03.11.2025	Revision	ALL	Third version
4.0	17.11.2025	Revision (FINAL)	CNR	Final revision

Executive Summary

The deliverable presents two experimental datasets collected with the purpose of analysing the impact of human body blockage and motions on Electro Magnetic (EM) fields, namely hologram and/or Channel State Information (CSI) frames.

The goal of the first testing activity is assessing the Pattern Reconfigurable Antenna Systems (PARS) technology as a novel opportunity for human- scale passive sensing. Beam management techniques specifically beam steering and beam switching allows to dynamically channelize (or steer) the individual antenna radiation patterns to intended areas. The goal of the setup is to reconstruct an infrastructural snapshot of the environment, namely, to obtain an accurate prediction of the number of human subjects (i.e., the targets) moving in the monitored space as well as to capture the body motion patterns on a given time instant t. The tests have been performed inside the Testhouse environment of ADANT using WiFi 6E devices.

The goal of the second tests is to assess the scalability of the motion detection and people counting scenario in large spaces and using a larger number of devices. The goal of the second setup is again to real-time track of the number of subjects co-present and moving in different areas. These tests have been performed inside a laboratory environment (CNR) and using industrial field devices (IEEE 802.15.4 based with modified firmware).

For each presented scenario we performed an accuracy benchmark test and reported the main results. Datasets and example codes have been published online (Kaggle database)

https://www.kaggle.com/datasets/rvision/people-counting-through-wifi-testhouse-samples

The following contents are covered in this report:

- Short introduction of the HOLDEN project and the collaboration partners
- Measurement dataset and description of the indoor measurement campaign inside the Testhouse environment and the CNR laboratory.
- Accuracy analysis in the considered scenarios targeting infrastructural snapshot application (body motion detection and counting)

Table of Contents

Abbreviations

Abbreviation	Description
3D	three-Dimensional
Aalto	Aalto University
BPA	Back Projection Algorithm
CNR	Consiglio Nazionale delle Ricerche
EC	European Commission
EM	ElectroMagnetic
EU	European Union
FEM	Finite Element Method
GTD	Geometrical Theory of Diffraction
HE	Horizon Europe
HOLDEN	Ethical design of holography in dense wireless networks
MIMO	Multiple Input Multiple Output
MoM	Method of Moments
NN	Neural Network
PEC	Perfect Electric Conductor
RF	Radio Frequency
TUM	Technical University of Munich
TWE	University of Twente
UTD	Uniform geometrical Theory of Diffraction
WP	Work Package
CSI	Channel State Information
CQI	Channel Quality Information
MUSIC	MUltiple SIgnal Classification

1. Introduction

1.1. About HOLDEN

The ubiquitous perception by sensing of objects, subjects, and gestures is a pivotal challenge for future technology: it enables personalized services such as smart living, automated logistics, or interaction through free-space gestures. However, it also challenges ethical and moral boundaries and threatens privacy. HOLDEN proposes a radically new approach to RF-based perception by concisely analysing ethical constraints and privacy risks while re-thinking RF-based sensing. We establish necessary conditions for privacy preserving and ethically compliant sensing and develop new paradigms respecting these constraints.

For the first time ever, HOLDEN constitutes a concentrated effort to explore social aspects of RF-sensing to guide the technological advance, and to derive technology for ethically and privacy compliant perception. The development of ethical and privacy constraints is central to HOLDEN. We use these findings to derive privacy- and ethically compliant concepts for RF-based perception. We will develop a system of distributed multi-antenna devices for simultaneous multitarget recognition and ubiquitous perception with unprecedented accuracy, which constitutes a radical paradigm shift from a technology-centric perspective to a privacy-centric one via a privacy-by-design approach.

HOLDEN achieves this goal along three high risk, complementary, and privacy-centric paths:

Path 1: Continuous-space measurement points: Radio-based 3D vision by holographic image processing of RF wavefronts.

Path 2: Discrete-space measurement points: Advanced 3D beamforming for human-scale recognition and tracking through dense, massive, and connected antenna arrays.

Path 3: Signal processing and learning: High-dimensional tensor processing for the distinction of complex activities and motion from massive-dimensional RF data. The resulting breakthrough approaches and algorithms will be compared against application-level benchmarks via usage scenarios in the fields of logistics, smart living, and free-space.

1.2. Partners

The consortium consists of four academic partners and a high-tech SME partner: (a) Aalto University (AALTO), Finland, (b) Technical University of Munich (TUM), Germany, (c) Consiglio Nazionale Ricerche (CNR), with third party Politecnico di Milano (POLIMI), Italy, (d) University of Twente (TWE), Netherlands, and (e) Adant (Adant), Italy. This consortium features the specialized and complementary expertise required to achieve the project objectives. AALTO will be responsible for the project management (WP1), covered by an experienced and dedicated project manager. Ethical aspects (WP2) will be addressed by TWE (Prof. Ciano Aydin) who is a pioneer in the field. Eventual gender differences in the ethical perception will be considered. TUM pioneered RF holography, which makes TUM (Prof. Thomas Eibert) the ideal leader of WP3. In advanced

distributed signal and information processing, CNR has through Prof. Stefano Savazzi more than 14 years of experience. CNR will lead WP4. Since more than 10 years, AALTO is active in radio sensing and machine learning based activity recognition. This expertise makes AALTO (Prof. Sigg) the ideal leader of WP5. Adant (Daniele Piazza) will contribute to the market analysis, application possibilities, and validation (WP6). Led by AALTO, dissemination with the website as one the media will be addressed by all partners. All academic partners are committed to early publication of results, e.g., via arXiv (open science).

2. Motion detection and people counting through WiFi: Testhouse measurements

Radio Frequency (RF) sensing is an emerging technology paradigm that repurposes existing wireless communication networks, such as WiFi, for imaging and computer vision applications, namely ambient and human sensing. Recent research has demonstrated that RF holography techniques can surpass human vision capabilities in several tasks such as identifying and resolving individuals within dense crowds, interpreting gestures and emotions, and capturing images through walls. The section explores the use of pattern-reconfigurable antenna devices with unmodified WiFi signals for indoor people discrimination, namely counting the number of people co-present in the space. We discuss signal modeling, integration of beam-steering technology, and adaptation for this purpose. A case studies with benchmarmk tests within the ADANT Test-house environment is also presented. Datasets and example codes have been published online.

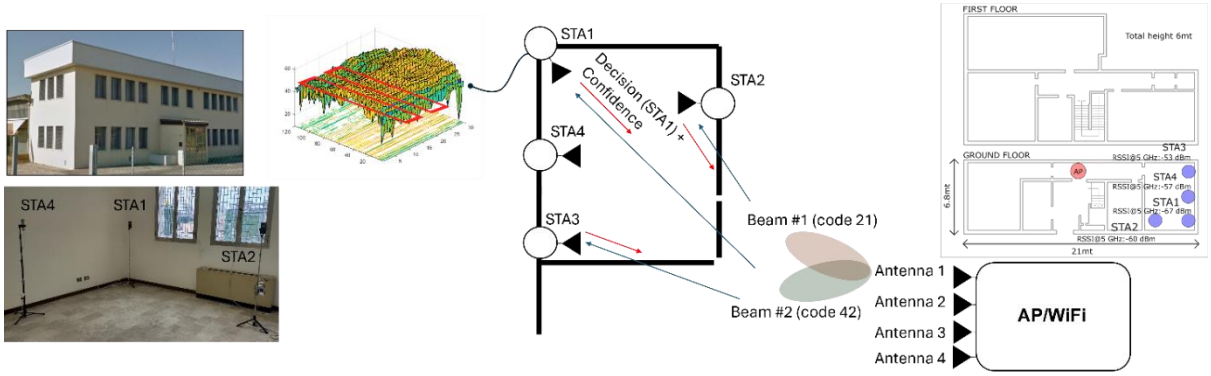

Figure 1: Wi-Fi data collection system for counting use case: ADANT testhouse environment.

2.1. WiFi and Beamsteering technology

Ubiquitous perception through Radio Frequency (RF) signals is a pivotal opportunity for future technology [1]: it enables personalized services such as smart living, remote healthcare, automated logistics or interaction through free-space gestures [2][3]. The ubiquity of Wi-Fi and cellular networks presents a promising platform for the development of innovative sensing tools. Future standards will also introduce dedicated sensing features which, for example, will allow routers to work as frequency modulated continuous wave radios targeting radar applications (see D4.3 and D4.5). Most of the current chip designs support ad-hoc firmware for Channel State Information (CSI) extraction with Multiple-Input Multiple-Output (MIMO) arrangements of the transmitter (TX) and receiver (RX) antennas and Orthogonal Frequency Division Multiplexing (OFDM) subcarriers [4][5]. The CSI describes the phase shift and amplitude attenuation of multiple propagation paths on each subcarrier. The latest IEEE 802.11be standard (Wi-Fi 7) offers a wider subcarrier bandwidth of 160MHz (up to 320MHz), providing at least 120 usable pilot sub- carriers for CSI or Channel Impulse Response (CIR) estimation. Additionally, Wi-Fi signals have been recently exploited to track

daily human movements and behaviors, while as shown in Deliverable D4.3, Wi-Fi signal variations have been shown to differ between different peoplE and can consequently be used for their re-identification [11].

In recent times, Pattern Reconfigurable Antenna Systems (PARS) technologies have been proposed as a novel opportunity for human- scale passive sensing [6]. Beam management techniques specifically beam steering and beam switching allows to dynamically channelize (or steer) the individual antenna radiation patterns to intended areas.

As depicted in Figure 1, each AP antenna system independently provides 2 steerable beams, indicated by beampatterns 21 and 42, respectively. The beam-steering system is deployed on the WiFi Access Point (AP) while the beam management is provided through software control exploiting real-time selection and channelization. Conventional beamforming and smart-antenna systems need precise phase alignment and covariance matrix estimation to determine the optimal beamforming design. However, in PARS, the beam patterns are dynamically altered on each antenna separately and do not require precise phase alignment simplifying product integration which opens unprecedented opportunities for scaling up RF sensing systems as well as for integration of communication and sensing services.

2.2. Initial tests

Considering the advantages of PARS, the HOLDEN project proposes a machine learning based beam-space processing architecture designed to split the learning functions between the deployed station devices (STA) devices and the AP. The STA devices analyze the CSI observed over different antenna beampatterns and utilize an Extreme Gradient Boosting (XGBoost) decision tree model [7] for supervision. Concurrently, a Convolutional Neural Network (CNN) is trained on the AP to generate a final decision, based on the outputs of the local XGBoost trees. The specific machine learning and signal processing architectures/pipelines are described in the following.

Validation tests have been conducted within the Testhouse environment of ADANT, namely a smart home environment of 340 sqm (see Figure 1 on the left).

Overall, the tests aim to achieve two primary objectives:

1) to develop an enhanced residential WiFi AP gateway with beamsteering capabilities (following HOLDEN Innovation 2);
2) to provide motion detection, activity recognition and people counting services via the same residential WiFi test network, supporting various application verticals such as elderly care, occupancy monitoring and building automation [6][9].

Figure 2: Measurement setup and deployment scenarios.

2.3. WiFi data collection system

We focus on the uplink communication of a WiFi communication system consisting of N stations communicating with one WiFi 6E compliant access point (AP) gateway equipped with 4 antennas. We resort to a practical case where the AP gateway is equipped with a smart antenna system and supports beam-steering functions, while the WiFi 4 STAs are equipped with a single antenna.

The AP is a commercially available device currently used by a large Internet Service Provider for its subscribers. It is based on the Broadcom BCM4912, a quad core 64-bit ARM processor, and a combination of three Broadcom BCM6715 WiFi6E radios, which concurrently cover all the 2.4GHz, 5GHz and 6GHz bands with 4x4 MIMO antenna systems and channel bandwidth up to 160MHz. This AP is equipped with smart antenna technologies in the 5GHz and 6GHz bands, by incorporating RF switches between the RF chains and the antenna elements. The design allows for selecting and combining the radiation patterns for different antenna elements placed in the devices, achieving isotropic coverage at each RF chain and effectively covering the holes in radiation patterns that are typically left by the passive antenna design.

As depicted in Figure 2, the WiFi STA devices and the WiFi AP gateway are in Non Line-of-Sight (NLOS) and located in two different rooms. Regarding data collection and processing, the CSI collection duration for training corresponds to 30 minutes for each test, with sampling frequency

of 250 ms, corresponding to 4 CSI samples/sec for each STA (which corresponds to around 30K samples for each test).

Body motions are classified/detected by inspection and real-time analysis of the CSI matrix H response observed by the STA (i) and corresponding to AP antenna beampattern mode (k), at discrete time instants t=1, 2, ... Each time instant corresponds to a WiFi PHY Protocol Data Unit (PPDU) frame. According to this setup, the CSI matrix H observed by the STA device at time t has thus dimension 4 x S with S =120 being the number of usable pilot subcarriers for each frame.

2.4. Tests-counting

The goal of the testing activity is to reconstruct an infrastructural snapshot [9] of the environment, namely, to obtain an accurate prediction of the number (X) of human subjects (i.e., the targets) moving in the monitored space as well as to capture the body motion patterns on a given time instant t. The proposed application case study is meant to design a system that provides augmented sensing functions for real-time tracking of the number of subjects co-present and moving in different areas of the Testhouse environment.

During training and calibration stages, we recorded the CSI corresponding to a single individual moving in the selected area (Figure 2) and gradually increase the number of individuals co-present in the same area up to a maximum of X=8 subjects.

As the number of subjects increases, the body-induced blockage affects the CSI of more subcarriers, resulting in distinct RF fingerprints. During testing, we record people walking, standing, or moving in groups in the same area. The AP serves as edge node and is responsible not only for collecting the CSI but also for its processing. Besides the AP, the proposed architecture allows the STA nodes to deploy and execute simplified models, such as XGBoost, for local computations, as further elaborated in the subsequent discussion.

In the proposed settings, the evaluation of the proposed beam-space processing architectures are implemented on a remote PC: the CSI features are moved/transmitted by using Java Script Object Notation (JSON) serialization and the MQTT publisher/subscriber transport service [10]

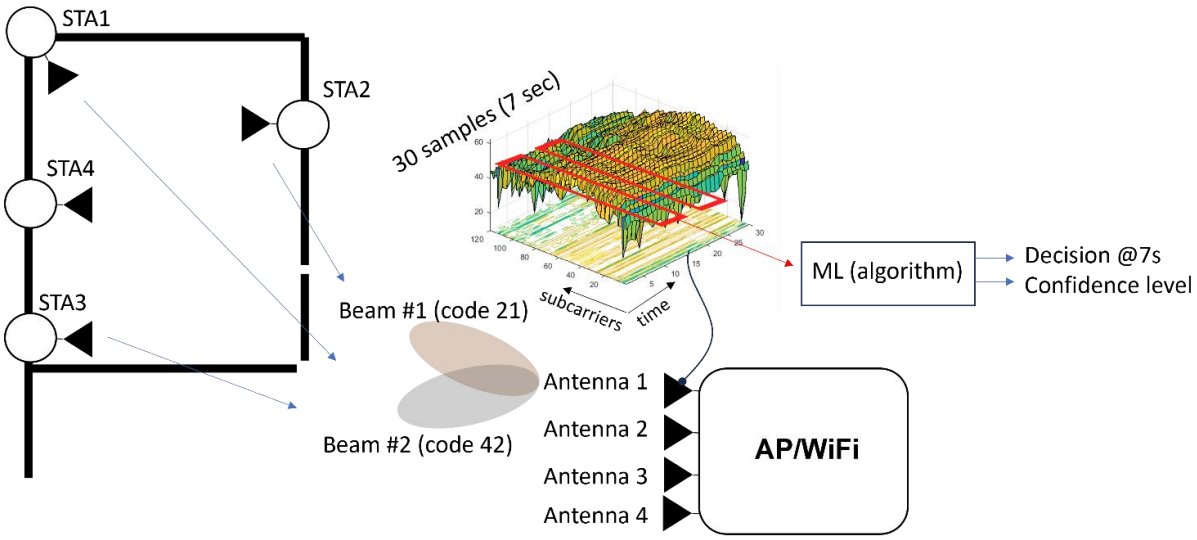

Figure 3: UL-F Uplink CSI data fusion on the AP.

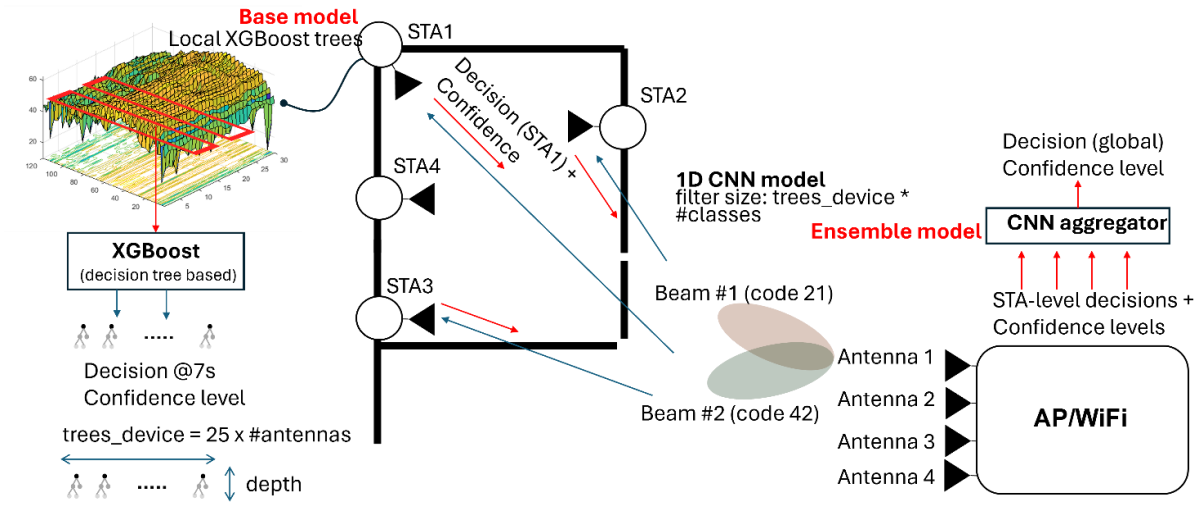

Figure 4: DL-SL: Downlink split learning

2.5. CSI processing pipelines (UL-F and DL-SL)

For all the considered cases, the CSI matrix H is collected over time t for each STA (i), and beam steering profile k (namely two beampatterns k=21 and k=42 are used in these tests). Information about the time-varying Modulation and Coding Scheme (MCS) is also available for each PPDU frame. The approach we followed is to detect and classify: the variations of the estimated CSI amplitude samples △H over a sliding window of duration △T=7 consecutive WiFi PPDU frames - further details are in [5][3].

Two beam-space data processing architectures/pipelines are implemented.

1) **UL-F: Uplink CSI data fusion on the AP**. The first case, shown in Figure 3, involves data collection at the Wi-Fi AP (over uplink) and the subsequent training of a single Machine Learning (ML) model. An XGBoost model is trained on the AP considering all the CSI variations △H as inputs.
2) **DL-SL: Downlink split learning**. The second case, depicted in Figure 4, proposes a split learning architecture where a portion of the ML model is trained locally on each Wi-Fi STA device. The results of this processing are collected by the Wi-Fi AP and then fed to a second aggregation stage. Thus, a second algorithm/model is trained independently. The chosen local model trained on the STAs is the XGBoost, while the model used for aggregating the partial decisions of each STA ise CNN.

	UL-F	DL-SL
Inputs	$\triangle \mathbf{H}_t(i,k), \forall i,k$	$\triangle \mathbf{H}_t(i,k), \forall i,k$
Local Model	n.a.	XGBooost trees per STA: $25 \times M_R$ depth: 5
Aggregator (AP)	XGBooost trees: $25 \times 4 \times M_R$ depth: 5	1D CNN layer filter size: 25×8

Figure 5: UL-F and DL-SL processing: summary of the main local and aggregator (on the AP) model parameters

Notice that the split learning architecture is attractive due to its scalability potential and support to massive RF data processing, since raw CSI are kept on the STA device and never shared [5]. The Table of Figure 5 summarizes the main parameters of the local model (running on the STA) and global/aggregation model (on the AP) used for UL-F and DL-SL architectures.

Scenario	Description
STA i, beam $k = 21$	Counting and subject discrimination are achieved by processing the CSI matrix $\mathbf{H}_t(i,k)$ observed by STA $i = 1, 2, 3, 4$ and beam code $k = 21$
STA i, beam $k = 42$	Counting and subject discrimination are achieved by processing the CSI matrix $\mathbf{H}_t(i,k)$ observed by STA $i = 1, 2, 3, 4$ and beam code $k = 42$
STA i, all beams	Counting and subject discrimination are achieved by processing the CSI matrix $\mathbf{H}_t(i) = [\mathbf{H}_t(i, k = 21), \mathbf{H}_t(i, k = 42)]$ obtained by AP RX antenna switching between beam codes $k = 21$ and $k = 42$ by time division
UL-F	Counting and subject discrimination are achieved by fusing the CSI matrices $\mathbf{H}_t = [\mathbf{H}_t(i = 1), ..., \mathbf{H}_t(i = 4)]$ obtained by all the deployed STA on the AP
DL-SLs	A decision $\widehat{\mathbf{X}}(i,k)$ about the number N of subjects co-present in the room is made locally by each STA i and observed beam k. The results are collected by the AP into a matrix $\widehat{\mathbf{X}} = \left\{\widehat{\mathbf{X}}(i, k = 21), \widehat{\mathbf{X}}(i, k = 42)\right\}_{i=1}^{4}$ and fed to a second processing stage (aggregation stage) and a final decision.

Figure 6: Description of the considered scenarios for benchmark tests.

2.6. Testhouse environment description and scenarios

We now introduce the test-house infrastructure of ADANT and outline the tests for the measurement campaign. We also introduce the test types, set up, and configurations of the CSI processing infrastructure. WiFi CSI data are collected in the Testhouse environment for evaluating algorithms and optimized recognition beams. The goal is to recognize specific movements and locations and improve machine learning algorithms for data classification. The Testhouse environment (see Figure 2) is located in Selvazzano Dentro, Padova (IT) and consists of a two-floor structure with an area of approximately 340 sqm.

In particular, the following features are highlighted: pristine Wi-Fi spectrum free of interference, state-of- the-art equipment for fully automated Over-the-Air (OTA) testing for Wi-Fi and Internet of Things (IOT), compliant with various Wi-Fi standards, including Wi-Fi 6, 6E, and 7. The facility includes eleven rooms and a corridor on each floor that can be used to emulate different types and ranges of installations, ensuring maximum reliability and reproducibility of the results.

The Table of Figure 6 summarizes the CSI processing scenarios considered in the testbed study. The first two setups are designed to evaluate the performance of the motion detection system (people counting) by processing the CSI data observed from individual STAs devices and using a single beamsteering pattern (k=21 or k=42). In the third scenario, counting and motion discrimination are achieved by time-division processing of the CSI matrix. This matrix is obtained by an AP whose RX antennas switch between beams k=21 and k=42. The final CSI processing is implemented by fusing the channel matrices measured for both beamsteering patterns. The last two scenarios refer to UL-F and DL-SL pipelines described previously.

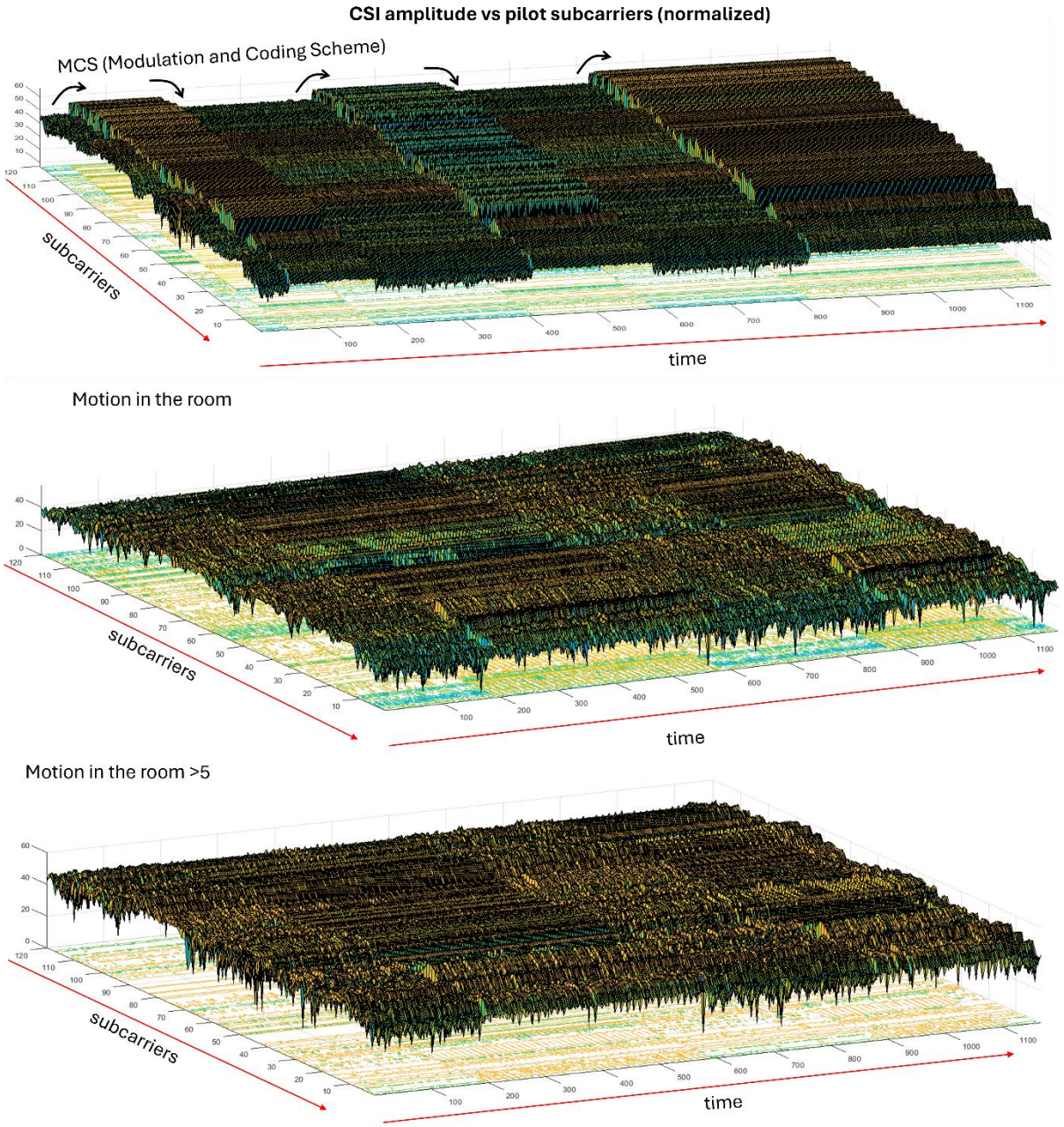

Figure 7: CSI example data: normalized CSI amplitude vs the S = 120 pilot subcarriers for empty environment and different motions example. MCS (modulation and coding format) changes over consecutive PPDU frames.

Algorithm + data	X=1	X=2	X=3	X=4	X=5	X=6	X=7	X=8
STA 1 - beam 21	0.98	0.91	0.89	0.82	0.75	0.78	0.76	0.66
STA 1 - beam 42	0.97	0.95	0.95	0.94	0.93	0.95	0.93	0.84
STA 1 – all beams	**0.98**	**0.93**	**0.97**	**0.95**	**0.94**	**0.94**	**0.94**	**0.86**
STA 2 - beam 21	0.98	0.97	0.89	0.88	0.89	0.83	0.84	0.81
STA 2 - beam 42	0.95	0.95	0.92	0.92	0.96	0.94	0.94	0.90
STA 2 – all beams	**0.98**	**0.97**	**0.98**	**0.96**	**0.96**	**0.95**	**0.93**	**0.94**
STA 3 - beam 21	0.97	0.93	0.86	0.82	0.87	0.80	0.80	0.76
STA 3 - beam 42	0.94	0.89	0.89	0.89	0.92	0.88	0.89	0.87
STA 3 – all beams	**0.97**	**0.91**	**0.92**	**0.93**	**0.92**	**0.90**	**0.92**	**0.89**
STA 4 - beam 21	0.99	0.98	0.86	0.87	0.83	0.84	0.79	0.77
STA 4 - beam 42	0.99	0.92	0.94	0.88	0.89	0.88	0.87	0.83
STA 4 – all beams	**0.99**	**0.98**	**0.96**	**0.93**	**0.90**	**0.91**	**0.91**	**0.90**
UL-F: STA 1 + 4 – all beams – fusion at AP	**0.98**	**0.99**	**0.98**	**0.97**	**0.98**	**0.95**	**0.96**	**0.93**
DL-SL: STA 1 + 4 – all beams - split learning	**0.99**	**0.99**	**0.97**	**0.96**	**0.96**	**0.98**	**0.98**	**0.98**

Figure 8: Counting accuracy measured on all WiFi station devices (STA 1-4), for all beampatterns (beam codes k=21 and k=42) and varying number X of subjects co-present. The performance of UL-F and DL-SL processing pipelines are also highlighted in green. In these examples the CSI is collected from *one* RX antenna at the AP according to typical Integrated Sensing and Communication scenarios.

2.7. Results

Figure 7 shows some examples of the observed normalized CSI amplitudes |H| in log scale across the S=120 subcarriers and for a consecutive number of PPDU frames (t).

Three cases presented are:

1) empty testhouse environment with no subjects moving in the considered room;
2) X=1 subject in motion;
3) X>5 co-present subjects.

The sensitivity of the CSI to the number of targets perturbing the link between STA i=1 and the AP node is noticeable. The abrupt changes observed in the CSI samples within the figure are attributed to the AP node re-selection of the MCS (modulation and coding scheme). While this reselection is beyond the control of the CSI processing tool, its effects are minimized by calculating the variations in estimated CSI amplitude samples across consecutive PPDUs as $\triangle H$.

Algorithm + data	AP/WiFi	AP/WiFi	AP/WiFi	AP/WiFi
STA 1 - beam 21	0.66	0.80	0.86	0.86
STA 1 - beam 42	0.84	0.91	0.93	0.93
STA 1 – all beams	**0.86**	**0.94**	**0.95**	**0.96**
STA 2 - beam 21	0.81	0.86	0.89	0.91
STA 2 - beam 42	0.90	0.94	0.94	0.95
STA 2 – all beams	**0.94**	**0.95**	**0.96**	**0.97**
STA 3 - beam 21	0.76	0.85	0.87	0.87
STA 3 - beam 42	0.87	0.87	0.95	0.93
STA 3 – all beams	**0.89**	**0.93**	**0.95**	**0.94**
STA 4 - beam 21	0.77	0.84	0.86	0.88
STA 4 - beam 42	0.83	0.93	0.95	0.97
STA 4 – all beams	**0.90**	**0.96**	**0.96**	**0.95**
UL-F: STA 1 + 4 – all beams	**0.93**	**0.93**	**0.94**	**0.95**
DL-SL: STA 1 + 4 – all beams - split learning	**0.98**	**0.99**	**0.99**	**0.99**

Figure 9: Counting accuracy measured on all WiFi station devices (STA 1-4), for all beampatterns (beam 21 and 42) and varying RX antennas at the AP node – from 1 to 4. Results are highlighted for X=8 subjects co-present. Performance of UL-F and DL-SL processing pipelines in green.

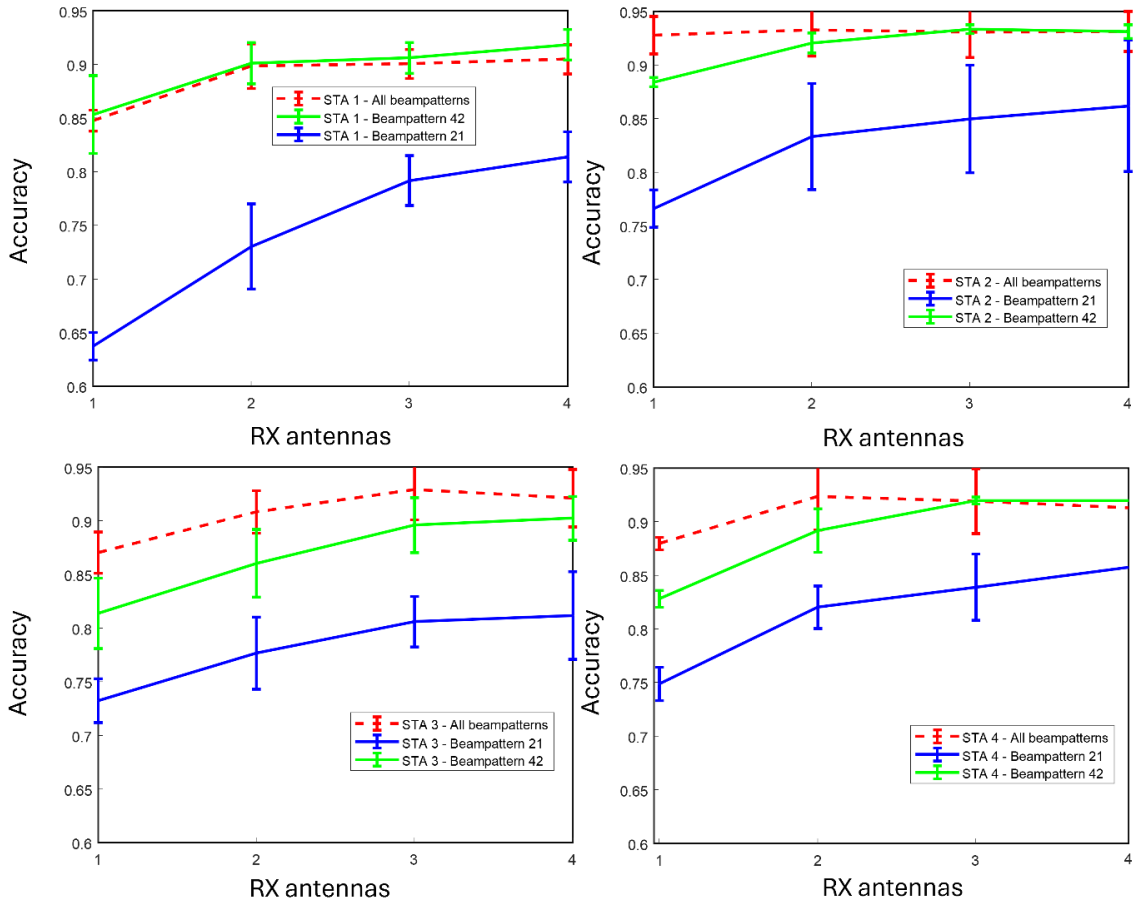

Figure 10: Average accuracy vs the number of RX antennas at the AP for processing taken place on STA 1-4 separately. Green lines correspond to beam 42, blue lines correspond to beam 21. Dashed

lines apply the joint use of both patterns. Error bars are superimposed for each case according to a three-fold cross-validation.

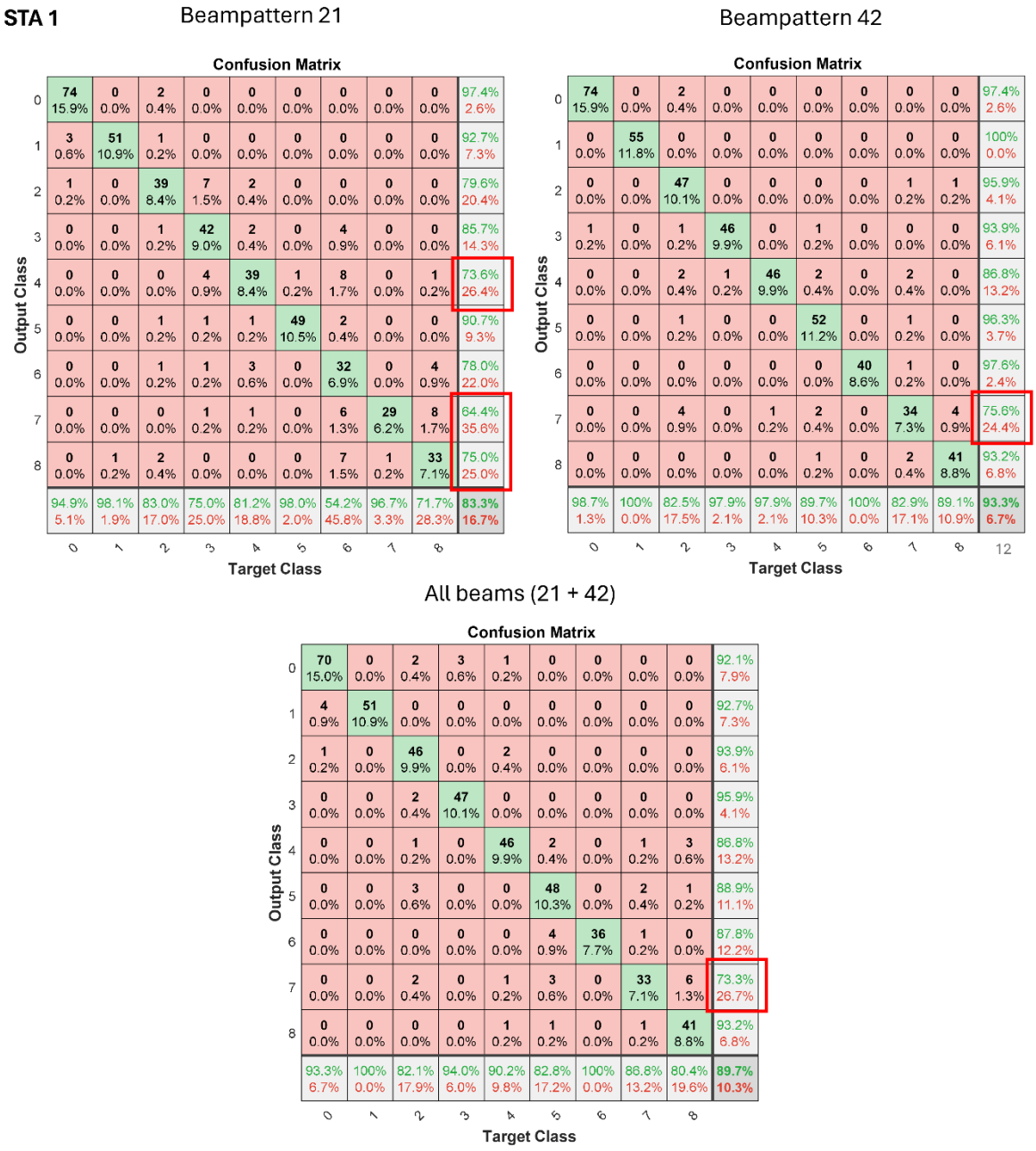

Figure 11: Selected confusion matrices for the motion detection (with up to X=8 subjects): (a) STA i=1, beampattern code k=21; (b) STA i=1, beampattern k=42; (c) STA i=1, beampatterns codes k=21 and k=42; (d) STA i=4, beampattern code k=21; (e) STA i=4, beampattern k=42; (f) STA i=4, beampatterns codes k=21 and k=42.

Figure 12: Selected confusion matrices for the motion detection (with up to X=8 subjects): (a) STA i=4, beampattern code k=21; (b) STA i=4, beampattern k=42; (c) STA i=4, beampatterns codes k=21 and k=42.

The Table of Figure 8 summarizes the counting accuracy obtained by processing the CSI variations △H on individual WiFi STAS (STA i=1,2,3,4) and using one RX antenna at the AP node. This setup is typically adopted in Integrated Sensing and Communication scenarios (ISAC) where some antennas/resources can be reserved for sensing purpose leaving the remaining ones for communication. For each STA, we gather the CSI information on the two beampatterns (code k=21 and k=42) separately and jointly, for computing the counting accuracy, i.e., beampatterns are switching between beam code k=21 and k=42 according to a time division policy. Accuracy is also computed for varying number (X) of subjects co-present, where X=1,...,8 Finally, the counting

performance of UL-F and DL-SL processing pipelines are highlighted in green. From the figure, there is a clear degradation in average performance for beampattern 21 compared to beampattern 42. However, the combined use of both beampatterns is effective in most cases and contributes to increasing the average accuracy by ap- proximately 10%, considering, for example, the counting of X=8 moving targets. The DL-SL architecture appears to perform better than the UL-F approach, which involves CSI data fusion at the AP. As shown in the Table of Figure 9, increasing the number of RX antennas, i.e., from 1 to 4 improves the accuracy as expected while accuracy degradation on beampattern 21 (compare to beam code 42) is still observed.

Figure 10 analyses in detail the increase in performance that can be achieved from the combined use of the two available beampatterns for each AP antenna. The figure shows the accuracy obtained by combining all beams, compared to that achievable using separate beams (green and blue lines). Finally, in Figure 11 and Figure 12 we analysed the corresponding confusion matrices chosen here as metrics to effectively assess how the system can discriminate between different targets as well as to quantify fairness (parity see D4.3), namely whether the system maintains an acceptable level of accuracy for different number of people (X) co-present in the space. The confusion matrices are analyzed in detail in Figure 11 for STA i=1 and in Figure 12 for STA i = 4, respectively. The benefits derived from the joint use of both beampatterns are highlighted (in green) and are particularly noticeable when the number of targets to be discriminated against is high (X > 4).

2.8. Datasets and examples

The dataset adopted for the benchmark tests as well as example codes written in Python have been published online

https://www.kaggle.com/datasets/rvision/people-counting-through-wifi-testhouse-samples

More specifically in Figure 15 we describe the WiFi CSI data. The dataset contains a list of files in the following format

feature_all_selected_drop$C$_antenna$Y$_samples$Z$_classes$X$.mat

where C is the number of samples discarded, Y is number of antennas to consider the AP side (1 to 4), Z is number of consecutive WiFi PPDU frames used for processing, X is number of classes (counting people from 0 to 8)

Each .mat file contains the following data:

1) *antenna_selected*: the specific antenna at the AP side (from 1 to 4)
2) *classes_nA_ant$B$:* true classes for each sample (true number of people) - Dimension correponds to the number of samples (PPDU frames). A is end device address (from 1 to 4 end devices). B is antenna beampattern configuration at the AP side (two options: 21 and 42)
3) *features_nA_ant$B$:* CSI features (standard deviations of the CSI ΔH over 120 consecutive subcarriers) - Dimension 120 x the total number of samples
4) *subcarriers_nA_ant$B$*: WiFi subcarriers indicator - Dimension 1 x the number of subcarriers = 120

Finally in https://www.kaggle.com/code/rvision/ul-and-dl-processing-split-learning-example (screenshot in Figure 16 we reported the example software used to implement UL-F and DL-SL signal processing architectures. The software allows to:

1) acquire the CSI data from the .mat files;
2) on device training an XGBoost model on each STA device;
3) training an XGBoost model on each STA device while considering all the beams jointly
4) fuse raw CSI data on the WiFi AP, and training an XGBoost model i on the AP (UL-F implementation);
5) sharing the outputs of the locally trained XGBoost trees with the AP where the XGBoost tree are the base models of the ensemble structure created on the AP. The trained XGBoost trees (or base models) are loaded from a shared folder;
6) learning and running a Convolutional Neural Network (CNN) based aggregator (ensemble model) on the AP using the outputs of the local XGBoost models (this is used to implement DL-SL architecture and signal processing pipeline).

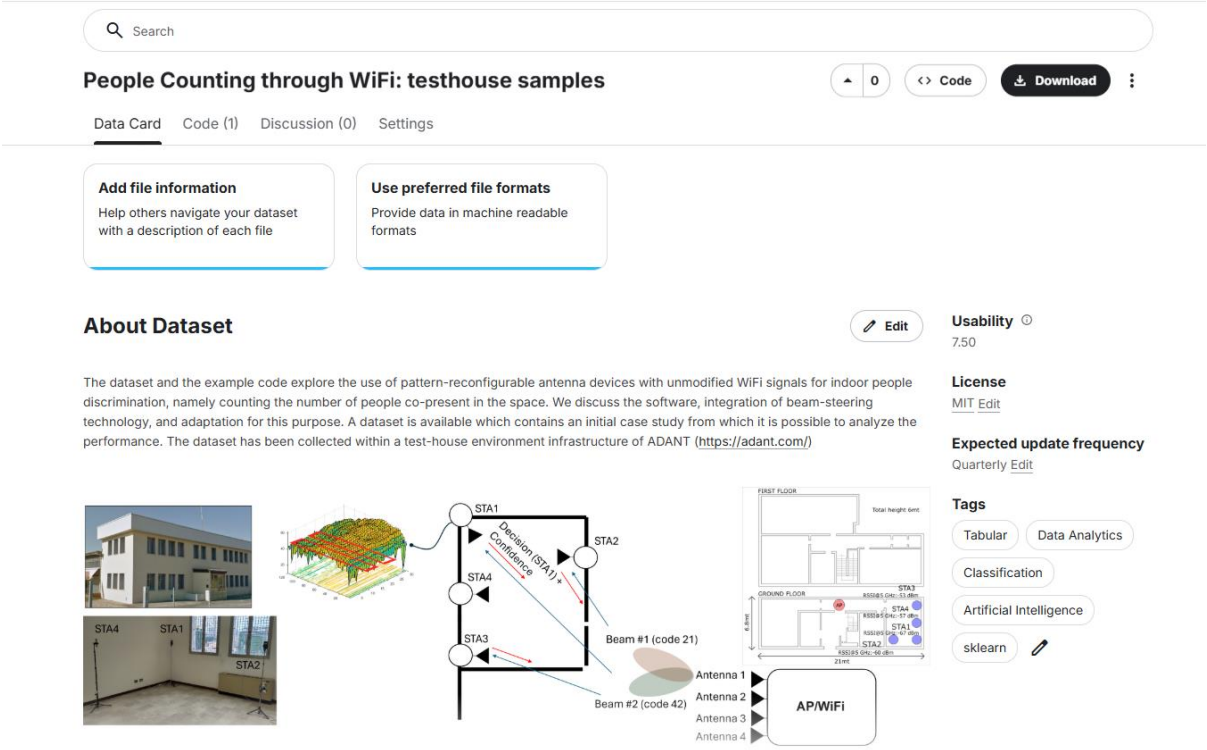

Figure 13: Published dataset online (Kaggle database)

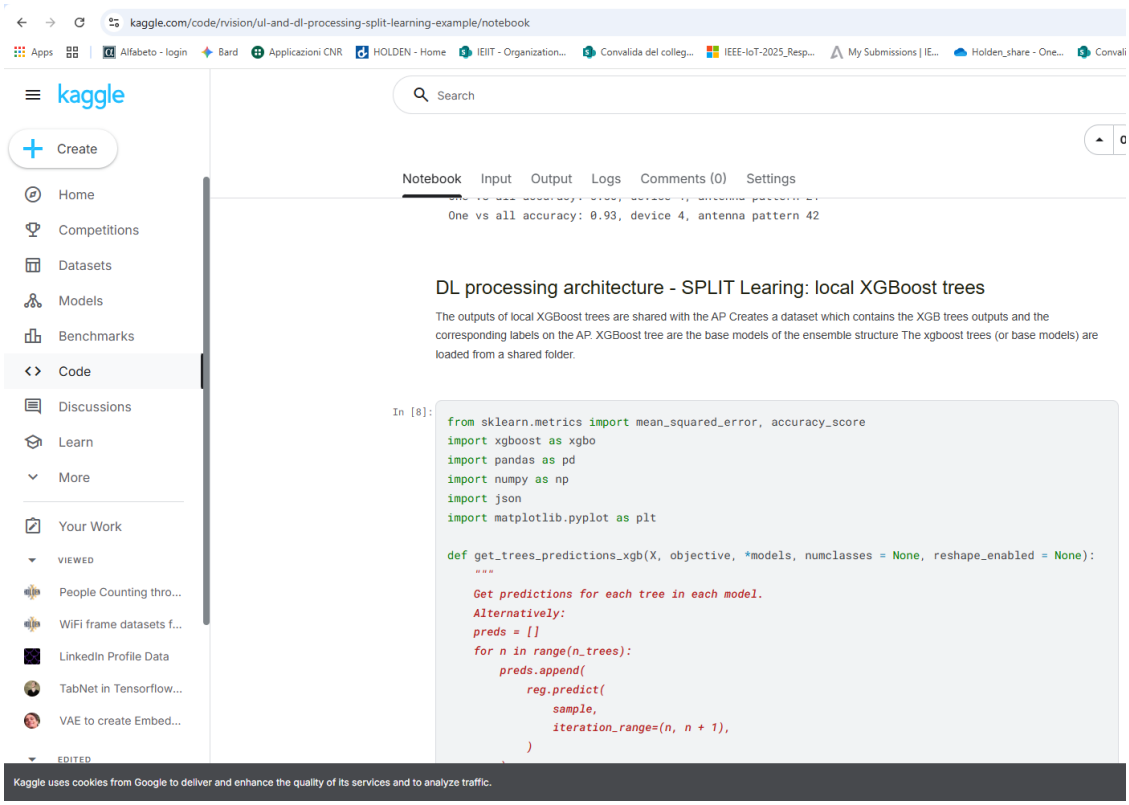

Figure 14: Published examples of UL-F and DL-SL signal processing pipelines

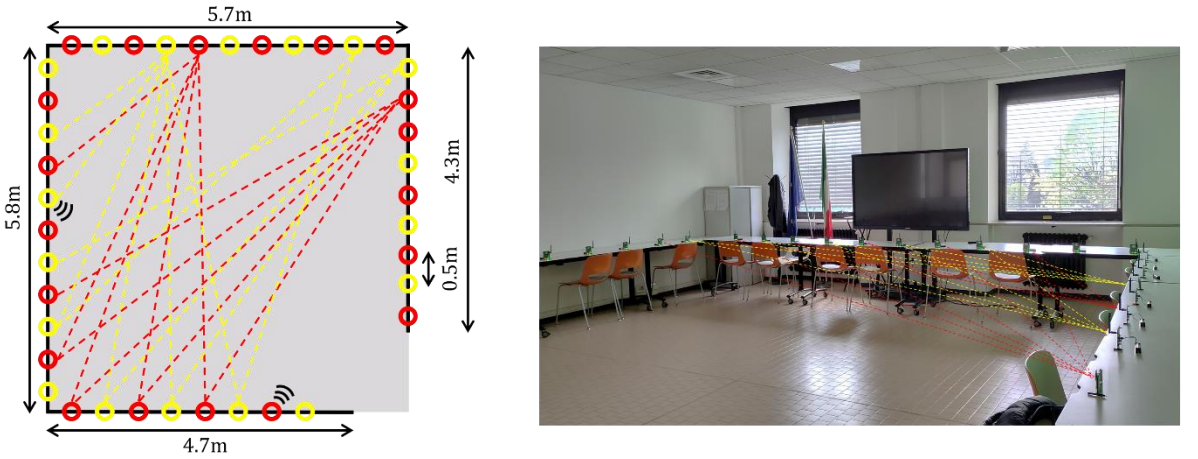

Figure 15: Infrastructure snapshot scenario: deployment and tests in an indoor environment (presented in D4.3).

3. People counting in dense IoT networks

In this section we complete the case study described in Deliverable D4.3. The study is designed to verify the performance of the motion detection tools over larger spaces (compared to the previous tests) and for increasing number of antenna elements. We also verified the proposed signal processing tools using Received Signal Strength (RSS) information rather than the full CSI as inputs. In addition, differently from the previous cases the RSS is collected at 2.4GHz. This setup is typical in IoT application use cases. We conducted an experimental campaign using two IEEE 802.15.4-based battery-powered RF networks for IoT applications operating at 2.4GHz. Both networks were deployed over a large 49 sqm indoor area (Figure 15), in accordance with typical open space scenarios. The two networks, represented as red and yellow nodes in (Figure 15), collect Received Signal Strength Indicator (RSSI) measurements simultaneously on different channels at 2.41 GHz and 2.43 GHz (IEEE channels 12 and 16), allowing approximation as a single network deployment due to negligible attenuation differences for very close frequencies [20] (i.e., 20 MHz in our case). Further details about the specific setup can be found in [16][17].

In what follows, we describe the experimental setup, the deployed data collection and signal processing system, the accuracy observed when counting the number of people present. We refer to Deliverable D4.3 for further details on the graph neural network tool (Deep Graph Convolutional Neural Network – DGCNN) and the statistical parity metric which can be used to assess the ability of the AI model to produce fair results.

3.1. Scenario and graph neural network model description

The network operates using a synchronous, time-slotted approach. The RF nodes are programmed to transmit Physical Protocol Data Unit frames, namely beacons, at pre-assigned time slots with duration 0.5ms and guard time of 0.15ms. During each time slot, the network nodes measure the RSS of the received beacons. Collected samples are then transmitted to a central Access Point (AP), which acts as the hub for data processing and analysis. A snapshot, namely one RSS measurement for each link of the network is collected by the AP every 60ms, which corresponds to one IEEE 802.15.4 superframe. Notice that the transmitter (TX) and receiver (RX) roles can be switched in each snapshot according to the assigned network topology. The protocol is implemented on a testbed using nodes equipped with the low-power SoC JN5189 transceiver from NXP [18]. Differently from the previous system, the testbed is optimized to work with a very large number of devices as it supports a maximum of 60 devices simultaneously collected to cover a building.

Although the Media Access Control (MAC) sublayer of the IEEE 802.15.4 standard has been customized for RF sensing, it still maintains full compatibility with the base standard, and several popular machine-type communication protocols such as RFC 4944 (6LoWPAN), Time Slotted Channel Hopping (TSCH) solutions, i.e., 6TiSCH [19]

The system operating frequency can be selected at run time, before deployment using standard IEEE 802.15.4 channels. Multiple networks might therefore operate in parallel over different, i.e., adjacent channels, without suffering from interference. Two networks (red and yellow nodes in (Figure 15) were deployed in the monitored area. Both networks operate on two distinct frequencies, namely 2.41GHz (channel 12 of the IEEE 802.15.4 standard) and 2.43GHz (channel 16). The combined network integrates a total of 38 antennas (nodes/edges) with 0.5m spacing and 342 links. It must be noted that the combined network can be approximated as a single network working at 2.42GHz since the attenuation effects due to the targets for a frequency difference of 10MHz are negligible.

As reported in D4.3, the RF signals and EM field samples are organized into graph structures with nodes and edges corresponding to the antennas of the radio devices and the active communication links, respectively. This choice is motivated by node-invariant properties (i.e., permutation invariance) of RF sensing as applied to dense networks. The problem of multi-target detection is formulated here as a classical graph classification problem. The presence of targets modifies the RF field and gives rise to different populations of graphs and node features which encode information about the targets as well as their physical characteristics (i.e., locations, size, and motions). In D4,.3 a flexible classification tool based on the DGCNN is proposed and discussed. The model is now applied to real measurements.

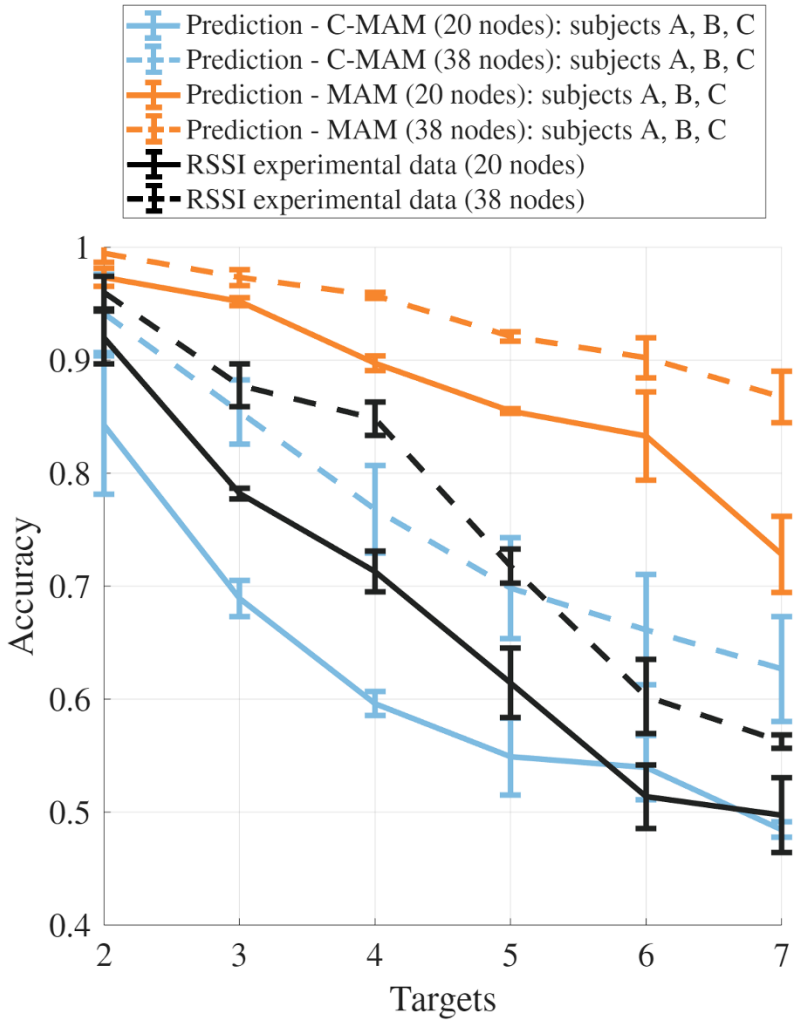

Figure 16: Average counting accuracy of the DGCNN model (D4.3) for varying number of targets (from X=2 to X=7) and RF nodes (20 - solid lines - and 38 - dashed lines) using attenuation samples generated by C-MAM (blue) and MAM (orange) models. C-MAM and MAM models are described in [16][17]. Predicted accuracy is compared with RSS experimental results (black curves). Error bars for C-MAM and MAM reflect the variability due to different subjects (Subject A, B, and C defined in D4.3).

3.2. Multi-target discrimination: accuracy results

The proposed scenario assumes static RF links as in typical IoT smart-home environments. The DGCNN model is trained on multiple snapshots that capture the continuous, time-varying changes in the RF signals as the result of human movements. In what follows, we compare the crowd sensing performance under two scenarios:

1) the DGCNN model trained with measured RSS attenuations obtained from up to X=7 real subjects;

2) the DGCNN model trained using attenuation samples generated by two EM models discussed in D3.6 and WP4, labelled as Multibody Additive Model (MAM) and a composite version (C-MAM) described in detail in [16][17]. The proposed generative models are used

to reproduce the environment in Figure 15: Infrastructure snapshot scenario: deployment and tests in an indoor environment (presented in D4.3). as well as the subject physical dimensions. The procedure is illustrated in [20]. The objective is to evaluate the predictive potential of the EM models discussed in previous deliverables in comparison to real-world measurements and the specific setup.

Looking to the measurements and the comparison with EM models, the following conclusions can be made:

1) The trend of the accuracy obtained through the RSS measurements is well aligned with the prediction of the C-MAM model while MAM model results can be considered as an upper bound to real performance. The proposed C-MAM effectively captures the mutual coupling of targets as co-present in the same Fresnel zone.
2) For a limited number of deployed nodes/antennas (20), C-MAM predictions tend to slightly underestimate accuracy (by approximately 10%) for target counts below N=4, likely due to the absence of multipath effects, which are relevant in real environments and might result in performance improvements.
3) When the number of nodes increases to 38, the predictions from C-MAM align more closely with the experimental data for intermediate values of the number of targets (2-5), but an overestimation of around 5% can be observed when the number of targets is larger than 5.
4) The uncertainty related to the subjects' physical dimensions increases with the number of co-present targets, a trend that is consistently reflected in both model predictions and real measurements.
5) Unlike the previous test with WiFi devices (Section 2), using RSS for motion detection and counting of co-present targets reduces te performance by approximately 10% for the same number of subjects to be discriminated. This performance reduction can be compensated for by increasing the number of RF nodes (links).

4. Conclusions

This deliverable presented experimental validation and benchmark tests for human motion detection and people counting based on RF sensing technologies developed within the HOLDEN project. Two complementary experimental setups were implemented and analyzed: a WiFi-based Testhouse environment leveraging Pattern Reconfigurable Antenna Systems (PARS), and a dense IoT network scenario based on IEEE 802.15.4 devices for large-scale coverage.

The Testhouse measurements demonstrated that beam-steering antennas can effectively enhance spatial discrimination and coverage uniformity, improving people counting accuracy through adaptive beam management. The experiments compared uplink fusion (UL-F) and distributed split learning (DL-SL) schemes, showing that DL-SL achieves superior scalability and privacy preservation, as raw Channel State Information (CSI) remains local to devices. Joint use of multiple beampatterns increased the accuracy by approximately 10% in multi-target scenarios, confirming the advantage of beam diversity.

The IoT network experiments extended the analysis to dense sensor deployments and graph-based neural processing. Results confirmed that electromagnetic-informed models (C-MAM) accurately predict real-world performance trends, validating their use as generative tools for RF data synthesis and system pre-training. Performance degradation observed with RSS-based sensing compared to CSI ($\approx$10%) was compensated by increasing the number of active nodes and leveraging graph convolutional learning.

Overall, the results confirm the feasibility of privacy-preserving, infrastructure-embedded RF perception for motion and occupancy awareness in smart environments. The developed datasets and example codes published on open platforms (Kaggle) ensure reproducibility and support further community research.

2. References

[1] Moustafa Youssef, Matthew Mah, and Ashok Agrawala. Challenges: device-free passive localization for wireless environments. In Proceedings of the 13th Annual ACMInternational Conference on Mobile Computing and Networking, MobiCom '07,

[2] Dan Wu, Daqing Zhang, Chenren Xu, Hao Wang, and Xiang Li. Device-free wifi human sensing: From pattern-based to model-based approaches. IEEE Communications Magazine, 55(10):91–97, 2017.

[3] S. Savazzi, et al., " Holography with Dense Wireless Networks: A Case for Ethical Design," IEEE Comm. Magazine to appear 2025. Available online: https://www.techrxiv.org/doi/full/10.36227/techrxiv.175393563.33121883

[4] Rui Du, Haocheng Hua, Hailiang Xie, Xianxin Song, Zhonghao Lyu, Mengshi Hu, Narengerile, Yan Xin, Stephen McCann, Michael Montemurro, Tony Xiao Han, and Jie Xu. An overview on ieee 802.11bf: Wlan sensing. IEEE Communications Surveys I& Tutorials, 27(1):184–217, 2025.

[5] R. Bersan et al, WiFi-Based People Counting Using Beam-Steerable Antennas: A Test-bed Study. ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp Com- panion '25), October 12–16, 2025, Espoo, Finland.

[6] Stefano Savazzi, Vittorio Rampa, Sanaz Kianoush, and Daniele Piazza. Pattern reconfigurable antennas for passive motion detection: Wifi test-bed and first studies. In 2019 IEEE 30th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), pages 1–6, 2019.

[7] Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16, page 785–794, New York, NY, USA, 2016. Association for Computing Machinery.

[8] Rathin Chandra Shit, Suraj Sharma, Deepak Puthal, Philip James, Biswajeet Pradhan, Aad van Moorsel, Albert Y. Zomaya, and Rajiv Ranjan. Ubiquitous localization (ubiloc): A survey and taxonomy on device free localization for smart world. IEEE Communications Surveys I& Tutorials, 21(4):3532–3564, 2019.

[9] Sage Cammers-Goodwin, Nolen Gertz, and Ciano Aydin. Moralizing radio frequency (rf) photography using techno-moral scenarios. In 2024 IEEE 29th International Conference on Emerging Technologies and Factory Automation (ETFA), pages 1–5, 2024.

[10] ISO/IEC. Standard iso/iec 20922:2016 information technology - message queuing telemetry transport (mqtt) v3.1.1, 2016.

[11] D. Avola, et al. "Person Re-Identification Through Wi-Fi Extracted Radio Biometric Signatures," IEEE Transactions on Information Forensics and Security, vol. 17, pp. 1145-1158, 2022

[12] Stefano Savazzi. (2024). WiFi frame datasets for body motion discrimination [Data set]. Kaggle, [Online] Available: https://doi.org/10.34740/KAGGLE/DS/4802891.

[13] C. Guo, G. Pleiss, Y. Sun, and K. Q. Weinberger, "On calibration of modern neural networks," 2017. [Online]. Available: https://arxiv.org/abs/1706.04599

[14] Stefano Savazzi, and Riccardo Bersan. (2025). People Counting through WiFi: testhouse samples [Data set]. Kaggle. https://www.kaggle.com/datasets/rvision/people-counting-through-wifi-testhouse-samples Example code: https://www.kaggle.com/code/rvision/ul-and-dl-processing-split-learning-example

[15] V. Rampa, et al., "Electromagnetic Models for Passive Detection and Localization of Multiple Bodies," IEEE Trans. on Antennas and Propagation, vol. 70, no. 2, pp. 1462–1745, 2022.

[16] F. Fieramosca et al. Passive RF Sensing for People Counting with Dense IoT Networks ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp Com- panion '25), October 12–16, 2025, Espoo, Finland.

[17] F. Fieramosca, RF sensing with dense IoT network graphs: An EM-informed analysis. IEEE Internet of Things Journal (under revision) https://www.techrxiv.org/users/30892/articles/1349298-rf-sensing-with-dense-iot-network-graphs-an-em-informed-analysis

[18] IEEE 802.15.4 low power wireless MCU Rev. 1.3, product data sheet JN5189 [Online] Available: https://www.nxp.com/docs/en/nxp/ data-sheets/JN5189.pdf, May 2021.

[19] Elsas, R., Van Leemput, D., Hoebeke, J. et al. "Multi-Modal Industrial IoT Networks: Recent Advances and Future Challenges," Wireless Pers. Commun. no. 140, pp. 1–24, 2025.

[20] S. Savazzi, F. Fieramosca, S. Kianoush, M. D'Amico and V. Rampa, "Electromagnetic-Informed Generative Models for Passive RF Sensing and Perception of Body Motions," in *IEEE Open Journal of Antennas and Propagation*, vol. 5, no. 4, pp. 958-973, Aug. 2024, doi: 10.1109/OJAP.2024.3407199

6. List of figures