D3.6 Signal processing tools: segmentation and management of hologram structures

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	31/07/2024
Actual date of submission	31/07/2024
Responsible	CNR
Contributors	CNR, 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
M. D'Amico	POLIMI	michele.damico@polimi.it
F. Fieramosca,	POLIMI	federica.fieramosca@polimi.it
S. Kianoush	CNR	sanaz.kianoush@cnr.it
V. Rampa	CNR	vittorio.rampa@ieiit.cnr.it
U. Milasheuski	CNR, POLIMI	usevaladmilasheuski@cnr.it
A. Prini	CNR	alessio.prini@stiima.cnr.it
S. Savazzi	CNR	stefano.savazzi@cnr.it

Change History

Version	Date	Status	Author (Company)	Description
1.0	12.06.2024	Initial	CNR	First version
2.0	28.06.2024	Revision	POLIMI	Second version
3.0	12.07.2024	Final	CNR	Final version

Executive Summary

The present deliverable studies fundamental signal processing tools used specifically to capture the influence of human body movements and locations on RF data as organized into hologram frames. The hologram frame is a maximally exhaustive representation of the RF radiation field (magnitude and phase) associated to different spatial elements, namely *antennas,* possibly with different *radiation patterns (or beams),* and observed over different *frequency bands, and* consecutive *time symbols*. RF data is thus defined in general as 4-dimensional tensor structure: the goal of this deliverable is to set proper signal processing tools design for efficient management of such structures.

Conventional Bayesian and statistical signal processing tools are designed to extract (or *segment*) the effects of body movements on RF field. Besides, the deliverable also summarizes novel methods and approaches optimized to reconstruct, or synthetically generate, the radiation field. Several model-based methods are proposed as based on Generative Neural Networks (GNN) tools: in particular, the deliverable considers the Variational Autoencoders (VAE) and Generative Adversarial Networks (GAN) techniques. Physics-driven GNNs enable the synthetic generation of RF data and, in general, they can reproduce various radiation field measurements and body motion scenarios based on their trained models.

The goal of GNNs is thus to create a synthetic environment which is used to reproduce hologram frames as well as several Electromagnetic (EM) blockage effects, for example the human body blockage and the alterations induced by body movements on the radio propagation. GNNs also reduce the need for training data as well as personal data collection that could be used maliciously for person (re-)identification. A GNN simulator is developed to integrate several emerging signal processing tools and their adaptations. We also verified the potential of these tools for synthetic generation of different hologram structures, namely RF attenuation data, Received Signal Strength (RSSI), or Channel State Information (CSI) tensors,

An initial measurement campaign has been conducted indoor to validate and test the considered tools. Measurements target indoor localization applications with controllable level of privacy. The deliverable discusses several results and the capability of the proposed tools to reproduce RF data in a practical scenario. The RF datasets and a description of the measurement campaigns will be available in the Deliverable D4.1

The following contents are covered:

- A short introduction of the HOLDEN project and the collaboration partners.
- A theoretical background of the signal processing tools designed for the management of RF tensor and hologram structures/frames.
- Simulations of AI generated data and model-based reconstruction.
- Validation of the generative AI methods for human subject tracking in a realistic scenario.

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 taken into account. 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. Signal processing tools for RF sensing

The RF sensing holography general goal is to extract the EM human body(ies) blockage effects ($\boldsymbol{E}_\theta$) from noisy measurements $\mathbf{S_t}$ of the RF radiation observed at time t. The radiation measurements can be in the form of a hologram structures as detailed in D3.1 or simplified tensor sets for application to dynamic environments. Targeting specifically human subject monitoring, we assume that the target is characterized by an unknown state θ which we want to recover from the body-induced effects on EM field E_θ.

The body effects E_θ can be evaluated in terms of excess attenuations A_θ [1]-[4] as in the example of Figure 1 or magnitude and phase measurements, namely baseband Channel State Information (CSI) C_θ [5]-[7].

In dynamic environments, the body or target state θ generally consists of an ensemble of features, such as body location, size, height, and orientation which depend on the specific sensing application.

In what follows, we provide the necessary background on RF sensing and signal processing tools, focusing specifically on Bayesian methods.

2.1. Bayesian approach and background

The objective of the RF sensing inverse problem is to obtain the posterior distribution $p(E_\theta|S_t)$:

$$p(\boldsymbol{E_\theta}|S_t) = \frac{p(S_t|\boldsymbol{E}_\theta) \cdot p(\boldsymbol{E_\theta})}{p(\boldsymbol{S_t})}$$

of the (unknown) human body blockage effects $\boldsymbol{E}_\theta$ given the measurements $\boldsymbol{S_t}$. Maximum A-Posteriori (MAP) solution to the above problem allows to extract the most likely effects:

$$\left(\widehat{\boldsymbol{E}}_\theta, \hat{\theta}\right) = \arg\max_\theta \{p(\boldsymbol{E}_\theta|\boldsymbol{S_t})\},$$

from which it is possible to recover the subject state $\hat{\theta}$ and any feature of interest, such as target 2D/3D size, position, orientation of the subject or object in the space. Field measurements S_t can be in the form of received power, Received Signal Strength (RSS), or base-band CSI magnitude and phase response.

Observations $\boldsymbol{S_t}$ are perturbed by the body movements according to a prior distribution, $p(\boldsymbol{E_\theta})$, which predicts the effects of the body (i.e., the target) in the state θ as the result of the propagation of the reflected, scattered, and diffracted EM waves.

The above Bayesian approach for solving the radio sensing problem requires the knowledge of the likelihood function $p(\boldsymbol{S_t}|\boldsymbol{E_\theta})$, namely the RF measurement model, and the prior distribution $p(\boldsymbol{E_\theta})$. The likelihood term depends on the data collection process as well as on the impairments introduced by the measurement instrument or by the environment. In localization and human scale sensing applications it is typically modelled as log-normal distributed.

The prior distribution models the initial beliefs on $\boldsymbol{E}_{\boldsymbol{\theta}}$. It is usually hard to represent as it often requires full-wave EM approaches. Model-based reconstruction of the RF data is thus required. Approximated solutions, such as diffraction models [8]-[11] and several variants are in many cases too time-consuming to be of practical use for real-time sensing scenarios. In addition, when it comes to practice, imperfect knowledge of the scenario, small, involuntary, body movements, or changing configurations of the propagation environment, make the prior even more difficult to obtain with an acceptable level of accuracy.

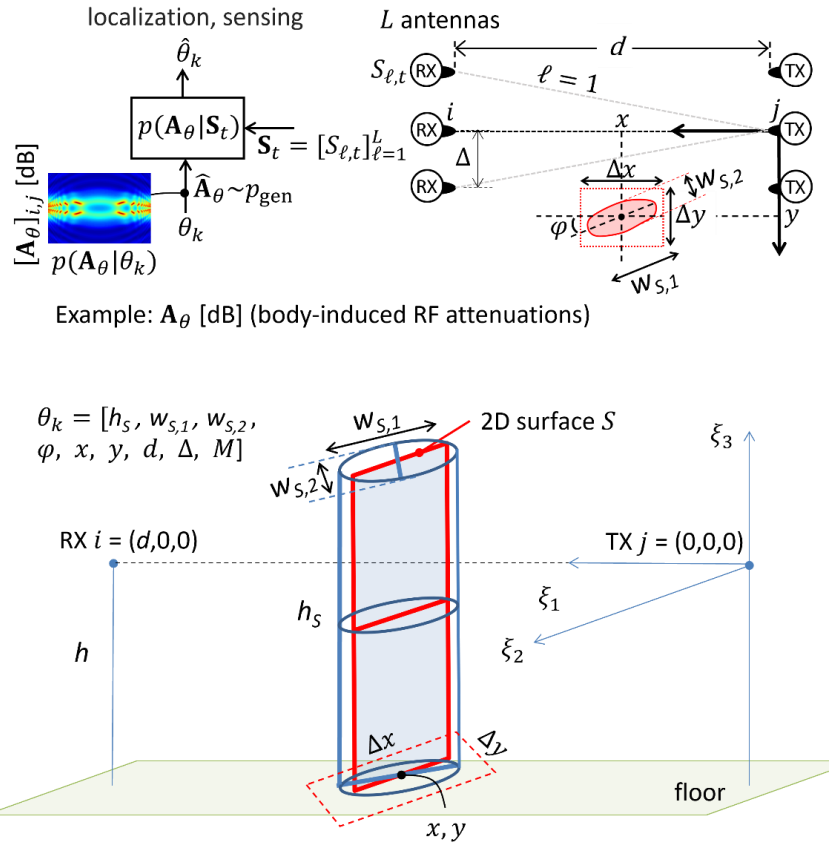


Figure 1: RF sensing general setups: single and multiple antennas, signal processing and model-based reconstruction of data through generative AI

2.2. Bayesian prior modelling of EM effects

In this deliverable, we propose, design and validate several tools targeting the model-based reconstruction of the RF/holography data. The tools are based on EM-informed Generative Neural Networks (GNN) techniques. The goal is to reproduce the EM effects E_θ as sampled from the Bayesian prior probability distribution $p(E_\theta)$. When focusing on dynamic environments, the prior generally quantifies the uncertainties of the body effects arising from imperfect knowledge of the target or body state. It is defined in general as:

$$p(\boldsymbol{E}_{\boldsymbol{\theta}}|\theta_k) = \boldsymbol{E}_{\theta\sim\boldsymbol{p}(\theta|\theta_k)}.$$

In other words, the EM effects are obtained for random instances of body features that follow a probability function. Probability $p(\theta|\theta_k)$ models the uncertainty with respect to the target features θ_k.

Examples. Some examples are proposed in the following to clarify the approach. First, consider the problem of generating body-induced RF excess attenuation values for a subject located at some (nominal) position $\theta_k = (x, y)$. Involuntary movements, as the result of the complex structure of the human body, strongly affect the RSS and must be adequately considered. Body motions can be represented by random movements in an elementary squared area of size 5-10 cm around the nominal location (x, y). Subject movements can be modelled by setting $p(\theta|\theta_k) = \mathcal{U}_{-\frac{\triangle}{2},\frac{\triangle}{2}}$ with $\triangle = 5 \div 10$ cm [8].

Likewise, let's consider now the problem of subject activity recognition, which requires to real-time track the subject trajectory and orientation w.r.t. the line of sight (LOS) path and its effects on CSI measurements. As reported in [19], diffraction models can be designed to capture the rotation angle φ of the 2D target by varying the size of the absorbing sheet S that represents the body. However, this operation is often expensive in terms of computational time. Rather than simulating each target rotation angle separately, which is not feasible, the proposed generative model can be set to reproduce the CSI for all subject orientations simultaneously. Further examples will be given in the next sections.

3. EM body models

In this section, we discuss relevant diffraction-based EM body models to reproduce the human body blockage effects E_θ, also considering body-induced RF excess attenuations A_θ and magnitude/phase (or CSI) data as special cases. First, we briefly recall the body models proposed in [8] for a single link scenario using scalar diffraction theory considerations. Next, we consider a receiver equipped with an array, i.e. Uniform Linear Array (ULA), of L isotropic receiver antennas. This setup is instrumental for the experimental activities that will be conducted in WP4 (Task T4.4) and WP5. The diffraction model represents the body effects relative to each radio link l, namely $\boldsymbol{E_\theta} = \left[E_{l,\theta}\right]_{l=1}^{L}$, or equivalently $\boldsymbol{A_\theta} = \left[A_{l,\theta}\right]_{l=1}^{L}$ for RSS and $\boldsymbol{C_\theta} = \left[C_{l,\theta}\right]_{l=1}^{L}$ for CSI.

In what follows, we will always assume that the monitored target is in the Fraunhofer's region of both transmitting (TX) and receiving (RX) antennas for all the considered links. Extension to multi-target scenarios can be also inferred according to [26].

3.1. EM models: diffraction

We adopted the body model originally proposed in [11] for a single link scenario and extended for the case of multiple array receivers [26], and more recently for antennas with non-uniform directivity patterns [9].

As depicted in Figure 1, we assume that the length of the radio link (horizontally placed w.r.t. the floor) is given by *d* while *h* is its height from the floor. The effects of floor, walls, ceiling, or other obstacles are not considered. However, with some effort, these obstacles can be included, as shown in [10]. The scalar diffraction theory assumes that the 3D shape of the human body is modelled as a 2D rectangular absorbing sheet with height h_S and traversal size that changes according to a 3D cylinder view, with max. and min. traversal sizes $w_{S,1}, w_{S,2}$, respectively. The target has nominal position coordinates $p = [x, y]$, w.r.t.

the TX position, which is defined by the projection of its barycenter on the horizontal plane that includes the Line-of-Sight (LOS). The 2D target might be also rotated of an angle φ with respect to the LOS direction.

A distribution of Huygens' sources of elementary area dS is located on the absorbing sheet S. The electric field E_θ at the receiver is obtained by subtracting the contribution of the obstructed Huygens' sources from the electric field E_θ of the free-space scenario (with no target in the link area):

$$E_\theta = E_0 - \int_S dE,$$

with time t omitted to simplify the reasoning. The above equation can be rewritten in terms of the field ratio $C_\theta = E_\theta / E_0$, representing the CSI (where λ is the wavelength, and the other quantities are defined in the Figure 1):

$$C_\theta = 1 - j\,\frac{d}{\lambda}\int_S \frac{1}{r_1\, r_2}\exp\left\{-j\frac{2\pi}{\lambda}(r_1 + r_2 - d)\right\} d\xi_2\, d\xi_3.$$

3.2. Full EM simulations of body induced effects on RF field

To allow for a comparative analysis, full-wave electromagnetic simulations have been also performed with Altair FEKO®, a commercial software package that implements, among others, the Method of Moments (MoM), a full-wave solution of Maxwell's integral equations in the frequency domain. Rather than a 2D perfectly absorbing sheet, as assumed in the scalar diffraction models, an anthropomorphic obstacle is here introduced between the antennas. A body-shape model, $h_S = 1.80$ m, $w_{S,1} = 0.52$ m and $w_{S,2} = 0.32$ m, is used while the obstacle is built as shown in Figure 2. The material of the body is retrieved from a simulator available online [28]. Assuming the body mainly constituted by muscles, we set: the relative permittivity $\varepsilon_r = 60$, the dielectric loss tangent tanΔ$= 0.242$, and the mass density $\rho = 1040.0$ kg/m^3. The surface of the humanoid obstacle is discretised using triangles, then the Surface Equivalent Principle (SEP) is applied by introducing equivalent electric and magnetic currents on the surface of the dielectric body.

Most of physical models designed to quantify realistic body motions must also consider imperfect knowledge of the environment, namely small and involuntary body movements which cause the subject to change its orientation or pose w.r.t. to the receivers. All these impairments make body features hard to obtain with an acceptable level of accuracy.

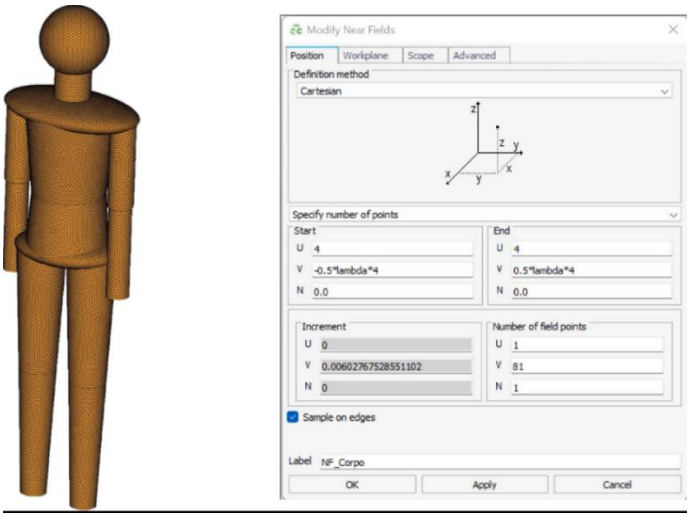


Figure 2: EM body simulation (FEKO interface)

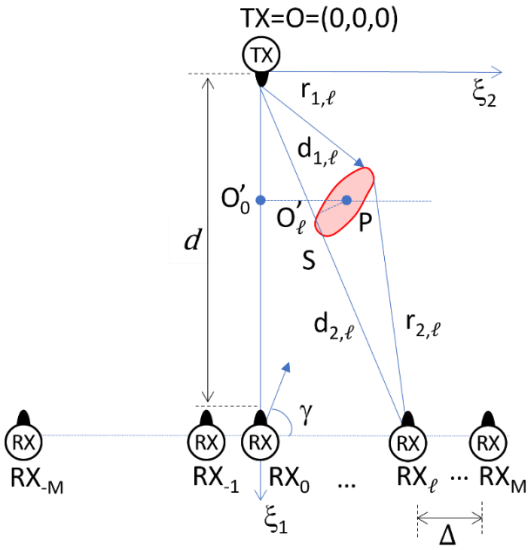


Figure 3: Uniform Linear Array geometry

3.3. Multiple antenna array configurations (MIMO)

We now consider a Uniform Linear Array (ULA) antenna configuration with links ordered as shown in the Figure 3. ULA configurations are typically of interest in holography, tomography and imaging approaches and will be discussed in the following sections. Each l-th antenna of the array is uniformly deployed at mutual distance Δ along a segment orthogonal to the LOS at distance d from the TX and horizontally placed w.r.t. the floor.

Ignoring mutual antenna coupling, the CSI observed on each antenna of the array corresponds to the ratio of the electric fields $C_{l,\theta} = E_{l,\theta}/E_0$, thus using the previous equation with:

$$C_{l,\theta} = 1 - j\,\frac{d_l}{\lambda}\int_S \frac{1}{r_{1,l}\,r_{2,l}}\exp\left\{-j\frac{2\pi}{\lambda}\left(r_{1,l} + r_{2,l} - d_l\right)\right\} d\xi_2\,d\xi_3,$$

where $E_{l,0}$ is the EM field received by the same RX node in the reference condition, i.e., the free-space scenario. The term d_l now indicates the distance of the l-th antenna RX of the array from the TX while $d_{1,l}$ and $d_{2,l}$ are the distances of the projection point O'_l (of the barycenter P of the 2D surface S, see Figure 3) from the TX and RX_l nodes. Likewise, $r_{1,l}$ and $r_{2,l}$ are the distances of the generic elementary area dS of the target S from the corresponding TX and RX, respectively. Clearly for M=0, the model reduces to the single-antenna case shown in Sect. 3.1.

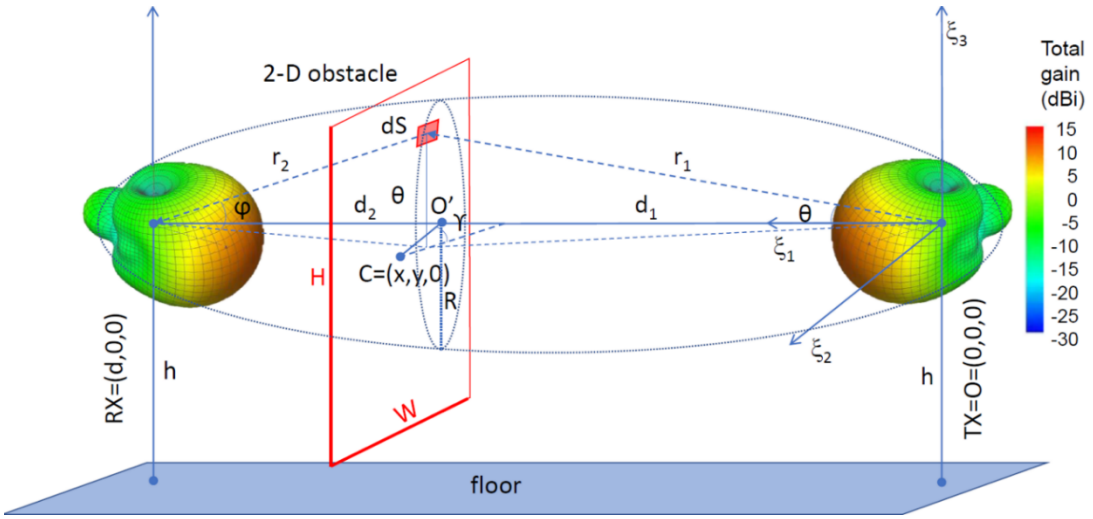


Figure 4: EM diffraction body model including antennas with non-uniform directivity

3.4. Effect of antennas with non-uniform directivity

The previously discussed EM models considered wireless devices equipped with a single or multiple omnidirectional antenna, without exploiting directional properties. Given the interest of HOLDEN project in novel WLAN sensing systems [29]-[34] that leverage antennas with non-uniform and/or re-configurable radiation characteristics, such as beam-switching antennas, it is considered necessary to develop effective EM models and related signal processing that meet these emerging needs. Models will be instrumental for core technology development in WP6.

In the following, based on the results of Section 3.1 and 3.3, we discuss a diffraction model, that includes antenna directivity hypotheses. The evaluation of the impact of antenna radiation patterns is also verified in [9] by exploiting real field measurements. Measurements and dataset used for validation are reported in D4.1.

For a generic non-isotropic antenna, the EM diffraction model previously discussed (Sect. 3.1) must be modified to take into account the antenna radiation pattern $G(\theta,\varphi) = G_0\, f(\theta,\varphi)$, G_0 being the gain and $f(\theta,\varphi)$ the normalized radiation pattern, while θ and φ that are polar coordinates usually referred to the antenna phase center. Considering, as depicted in Figure 4, a directional transmitting and receiving antenna with radiation pattern $G_t(\theta_t,\varphi_t)$ and $G_r(\theta_r,\varphi_r)$, respectively, the field ratio on link l $C_{l,\theta}$ becomes:

$$C_{l,\theta} = 1 - j\,\frac{\lambda d_l}{4\pi}\int_S \frac{1}{r_{1,l}\,r_{2,l}}\sqrt{G_t(\theta_t,\varphi_t)}\,\sqrt{G_r(\theta_r,\varphi_r)}\;\exp\left\{-j\,\frac{2\pi}{\lambda}\left(r_{1,l}+r_{2,l}-d_l\right)\right\} d\xi_2\, d\xi_3.$$

4. Model-based reconstruction of RF data

In this section, we consider the problem of model-based reconstruction of the RF field as support tool for the management and pre-processing of channel state information (CSI) and hologram structures. The methods used in this section are based on a physics-driven generative AI approach which achieve the synthetic generation of RF and hologram recordings to reproduce training data and real RF observations. Synthetic data can replace training data collected using real people. In the following, we discuss a generative AI simulator which integrates several emerging generative neural network techniques and their adaptations. The goal is to verify the potential of these tools for synthetic generation of RF attenuations and magnate/phase baseband RF data.

4.1. Generative Neural Network approach

The generative models considered in the following can reproduce different body-induced EM effects $\boldsymbol{E_\theta}$ as sampled from the ground-truth prior distribution $p(\boldsymbol{E_\theta}|\theta_k)$ described in Sect. 2.2.

In the following, we discuss the adoption of EM-informed Generative Neural Network (GNN) models inspired by VAE and GAN tools. All the GNN models are designed and trained to reproduce the human blockage effects on radio propagation as underpinned by scalar diffraction theory, discussed in Section 2. The models consider different body and link configurations relevant in radio sensing:

1) varying link geometries: distance (*d*), height from the ground (*h*);
2) Multiple-Input Multiple-Output (MIMO) antennas (L);
3) variable body size (size of the human subject, or the obstruction/blockage in case of an object), position, and orientation/posture of the monitored target/body.

The proposed GNNs also comprise of:

1) a generator of body-induced RF signal attenuations, which can be used in several passive motion detection and localization applications target of WP4
2) a generator of EM field samples which can be used for EM full-wave propagation analysis as well as advanced imaging, holography and gesture recognition applications.

The generators are shown to reproduce human body blockage effects under configurations which might be unseen during the training phase, or rather difficult to predict through traditional EM field computing methods. In line with some initial studies [14]-[25], the generators are purely based on sequences of convolutional (and de-convolutional) signal processing layers whose number varies depending on the physical quantity to be reproduced. Therefore, they are well-suited for most of HOLDEN use cases, such as real-time localization applications, as do not need intensive EM wavefield computations. In the following section, VAE and GAN generator methods are analyzed in terms of their accuracy in reproducing EM model diffraction effects, implementation complexity, generation times, and model size.

The implemented generators for all the considered use cases discussed in this deliverable are public available in the following github repository:

https://github.com/labRadioVision/generative_diffraction_model.

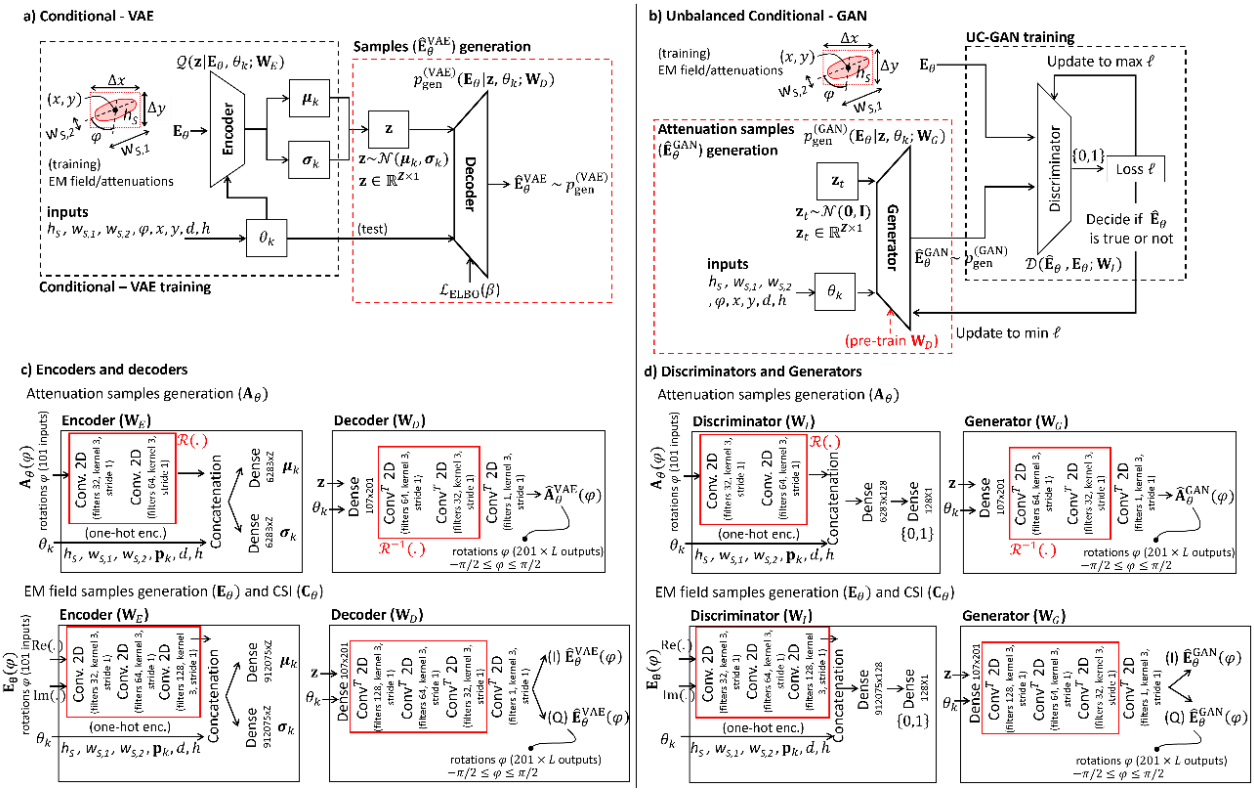


Figure 5: Physics driven Generative Neural Network (GNN) selected tools for reconstruction of RF attenuations and EM field: VAE and GAN

4.2. Variational Autoencoder (VAE) tools

As depicted in Figs. 5a) and 5c), the Conditional AutoEncoder (C-VAE) model uses an encoder (Q), parameterized by the NN parameters $\boldsymbol{W_E}$, which learns the latent space $p_Z(\boldsymbol{z}|\theta_k) \sim \mathcal{N}(\boldsymbol{\mu_k}, \boldsymbol{\sigma_k^2})$ for the inputs θ_k. Latent space is multivariate Gaussian distributed. The encoder is trained using samples $\boldsymbol{E}_\theta$ obtained from an EM body model (in this case the diffraction model discussed in Section 2) and the corresponding body states θ_k.The decoder produces a distribution of samples $\widehat{\boldsymbol{E}_\theta^{\mathrm{VAE}}} \sim p_{\mathrm{gen}}^{\mathrm{VAE}}(\boldsymbol{E}_\theta|\theta_k)$ which is optimized to reproduce the true prior. The goal is to maximize a likelihood metric referred to as Evidence Lower Bound (ELBO) described in [35]. ELBO metric balances the maximization of classical likelihood metric and the minimization of the Kulback-Lieber (KL) divergence which quantifies the difference between the true and the reproduced probability functions. Maximization of the likelihood makes the generated samples more correlated to the latent variables which typically cause the model to be more deterministic. On the other hand, the number of latent variables Z can be tuned to increase the contribution of the KL divergence between the posterior and the prior to the total ELBO and thus increase the randomness of generated samples. In [19], it has been shown that C-VAE generation of samples can be optimized

to improve the randomness of the generation process, also to account for measurements $\boldsymbol{S_t}$ affected by noise and multipath interference.

4.3. Generative Adversarial Network (GAN) tools

GAN training, depicted in Figs. 5 b) and 5 d), is formulated as a min-max problem that can be interpreted as an adversarial game with two players: the *discriminator* $\mathcal{D}\left(\widehat{\boldsymbol{E}_\theta}, \boldsymbol{E}_\theta\right) \in [0,1]$, namely a binary classifier which tries to improve the detection of fake EM field samples and the *generator*, which is trained to fool the discriminator. The goal is to minimize the statistical distance (Jensen-Shannon divergence) between the true and reproduced distributions. This corresponds to maximizing the discriminator loss while minimizing the generator one. More details can be found in [21]. An unbalanced implementation [36] has been recently considered which is specifically designed for reproduction of EM samples.

4.4. Body motion detection and background subtraction with GNN

VAE and GAN models are key enablers for anomaly detection in time series. In the context of HOLDEN project, anomaly detection is used for twofold purposes: 1) to segment RF data time series that might be indicative of body motions; 2) to subtract the background information which is indicative of an empty space (with no targets or subjects moving in the considered area).

The goal of anomaly detection tools is thus to identify the presence of body movements as inducing (body-induced) alterations of the measured RF field. The following body motion segmentation algorithms are identified as potential candidates for future algorithm design and development (WP4 and WP6):

1) **Bayesian anomaly detection**. The goal is to detect a change in the propagation environment as induced by body or target movements in the scene. Considering an indoor environment without targets in the monitored area, the RF data readings $\boldsymbol{S_t}$ (Section 2) are modelled as a stationary process which is typically modelled by a p-th order auto-regressive (AR) model [37]. A change in the environment, for example, single or multiple object/person/people occupying or moving in the corresponding monitored scene, alters the AR parameters that are thus replaced with new terms. Detection of such change can be implemented iteratively by monitoring the cumulative sum (CUSUM) [37].

2) **Array processing based (MUSIC).** These methods apply to receiver antennas organized in large and dense 2D and 3D array layouts, which is a typical setup in the HOLDEN project (see the following sections). RF signals originated from the transmitter might reflect off multiple objects when approaching the receiver array AP. Considering that the presence of the subject moving in the environment modifies the Angle of Arrival (AoA) paths, the MUSIC (MUltiple SIgnal Classification) algorithm is typically adopted [41] to disentangle the multipath components and estimate the AoA of the dominant propagation paths. Monitoring of the changes of such AoA paths is effective to obtain segmentation of body motions (see implementation and data examples in [42]).

3) **AI-based tools.** VAE-LSTM [38] and Anomaly Transformers [39] are identified as further candidates. The considered tools adopt an unsupervised approach that does not need prior knowledge about body-induced effects on the data. VAE-LSTM tools for example apply Long Short-Term Memory network-based Variational Auto-Encoder to extract the sequential features of the input multivariate time series RF measurements. Similarly to what done in the previous sections, the VAE model is trained to reproduce the RF data observations and convert them into a low-dimensional embedding (latent space) where possible anomalies induced by body movements are inspected. Both VAE-LSTM and Anomaly Transformer allow to detect and segment anomalies induced by body motions and occurring over both short and long periods.

5. Full-wave EM simulation analysis

In line with typical setups considered in the HOLDEN project, we report and describe here a simulation dataset which supports the analysis of the human body blockage effects on the EM field observed by receiver antennas organized in large and dense 2D and 3D array layouts. The dataset is also intended to support the analysis and the validation of array processing techniques for passive indoor localization which will be the core of WP4 activities. Full dataset will be published in D4.1.

The simulation scenario described below involves an Hertzian dipole radiating at the nominal frequency f = 2.4868 GHz and located at h = 0.99 m from the floor. The radiated field is measured over an array of points organized in a 3D layout as depicted in Figure 6. The 3D layout comprises of multiple 2D arrays that are uniformly deployed on co-located surfaces positioned at varying distances from the source. Each surface is composed of a grid of 90 × 180 elements, also spaced $\Delta = \lambda/10$ apart. The first measurement surface, i.e. the one closest to the source, is located 4 m away (distance d = 4 m) from the source itself (and its center is aligned with the source), and it is followed by other 49 additional surfaces spaced at intervals of $\Delta = \lambda$. The body is modelled as an anthropomorphic target described previously in Sect. 3.2.

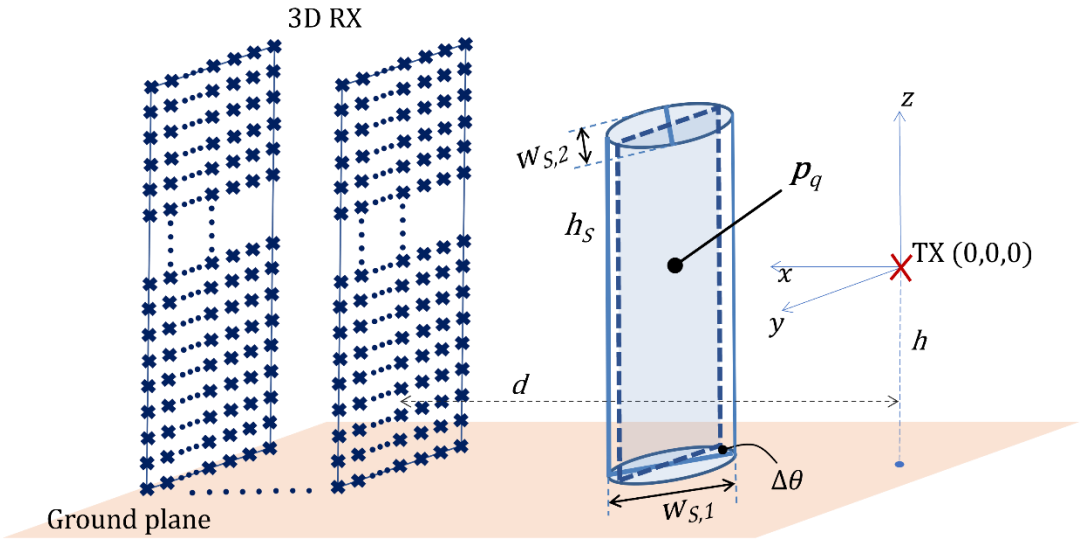


Figure 6: Full wave EM simulation of human body blockage observed by a dense 2D array of receivers.

FEKO® simulations were conducted with the body positioned at various distances and orientations relative to the source and the measurement points, allowing the collection of both real and imaginary samples of the EM field at the receivers' location.

The goal is to set a simulator environment for model-based reconstruction of the amplitude and phase of the CSI samples as well as the creation of a dataset (T4.3) which supports the analysis of the human body blockage effects on the EM field observed by receiver antennas organized in large and dense array layouts.

A simple tool for body imaging is also explored in this section using purely statistical considerations. The tool is based on the analysis of mean and standard deviation of the simulated RF attenuation samples. The tool is also useful to assess the impact of involuntary movements of the body on EM samples.

The images of Figure 7 report the average and standard deviation map of the EM field amplitudes collected by receivers on the 2D surface array. In this example, the body is in position (x = 0 m, y = 0.25 m). Figure 7(a) further highlights the overlapping between the average attenuation map and the body rotated by each Δθ to help relate the final pattern with some of the movements considered. Analyzing both maps, the body footprint is clearly distinguishable as well as parts of the torso and head. The regions with the highest attenuation are those consistent over translations and rotations of the body.

When the target moves, only a subset of receivers is affected by such movements. Using the relative position of these receivers in the 2D array manifold, it is possible to trace back the offset of the body blockage from the LOS path. Also, by observing the corresponding standard deviation maps, it is possible, for example, to distinguish specific human body parts most susceptible to movements and variability, i.e., the arm movements and shape.

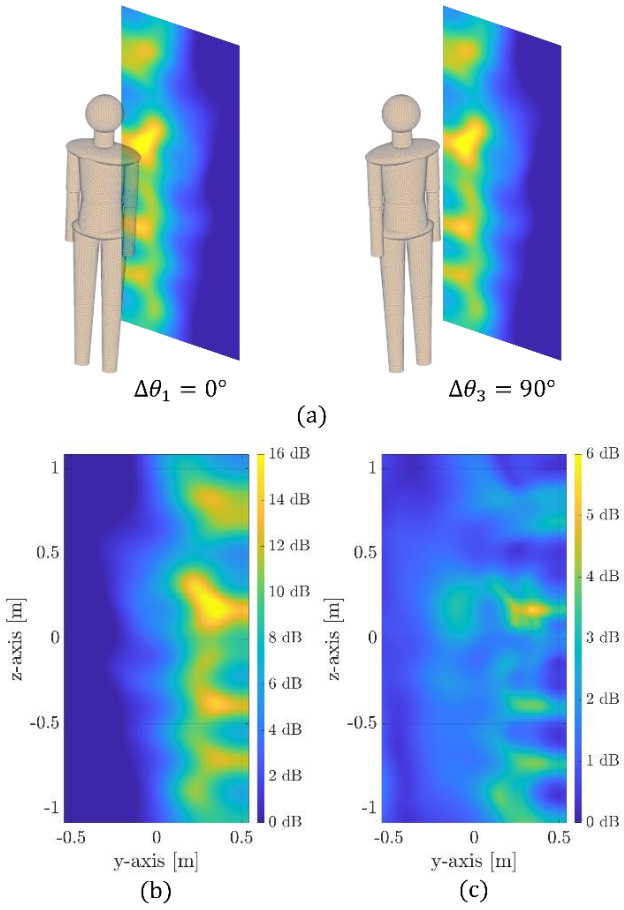


Figure 7: Imaging maps of attenuations observed on a 2D surface grid of 90x180 elements. (a) Positions and orientations of the simulated body highlighted; (b) average attenuation map and (c) corresponding standard deviations.

6. Validation of generative tools for detection of body motions

The generation of samples using the GNN tools proposed in Sect. 3 is assessed here at carrier frequencies 2.4 GHz and 5GHz. First, we analyze the reproduction accuracy against EM body diffraction. Next, we compare the results with an EM simulator (FEKO® software) and the body model detailed in Section 4. As depicted in the Figure 8, the RX node is equipped with a Uniform Linear Array (ULA) of 9 omni-directional antennas and spaced at $\lambda/2$ while the TX node is equipped with a single antenna. All the links of the array are horizontally placed at height h=0.99 m from the ground. Distance between the TX and the central RX of the array is d=4 m.

The C-VAE tool of Sect. 4.2 is trained here using examples of EM diffraction from: 1) 75 marked locations inside the first Fresnel's ellipsoid of the central link of the array; 2) varying blockage dimension in the interval 1.65 - 2.00 m. Notice that the chosen blockage dimensions are reasonable when considering human bodies. In the following, we verify the ability of the proposed C-VAE model to generate CSI samples to reproduce the array excess attenuation response of the conventional linear beamforming processing. Further details are provided in [26]-[36].

6.1. Reproduction of body induced diffraction effects

In this section, we verify the ability of the C-VAE model to reproduce the results of EM field simulation obtained by diffraction and MoM (Method of Moment) methods (i.e., using Altair FEKO tools). The TX is an hertzian dipole transmitting at 2.4868 GHz, the RX is composed by 9 isotropic antenna elements. For FEKO® simulations, the same anthropomorphic obstacle is used as described in Sect. 4.

Figure 8 shows the body-induced array response, in terms of the excess attenuation as a function of the direction of arrival (DoA) and for different values of the y displacement of the target (w.r.t. the central LOS) and x=2 m. The array signal processing extracts the response for varying DoA and is based on Fast Fourier Transform (FFT) with 257 points. The figure shows two responses (red and green lines) for corresponding nominal displacements y=0.25 m and y=-0.25 m, respectively. Dashed lines show the responses obtained using EM body diffraction, while solid lines represent the reproduced response using C-VAE. Notice that, in this example, each array response represents a random movement of the body in the surrounding of the nominal displacement. The maximum response of the array, namely the dominant DoA, is perturbed by the presence of the subject and such alteration is generally well reproduced by the C-VAE model.

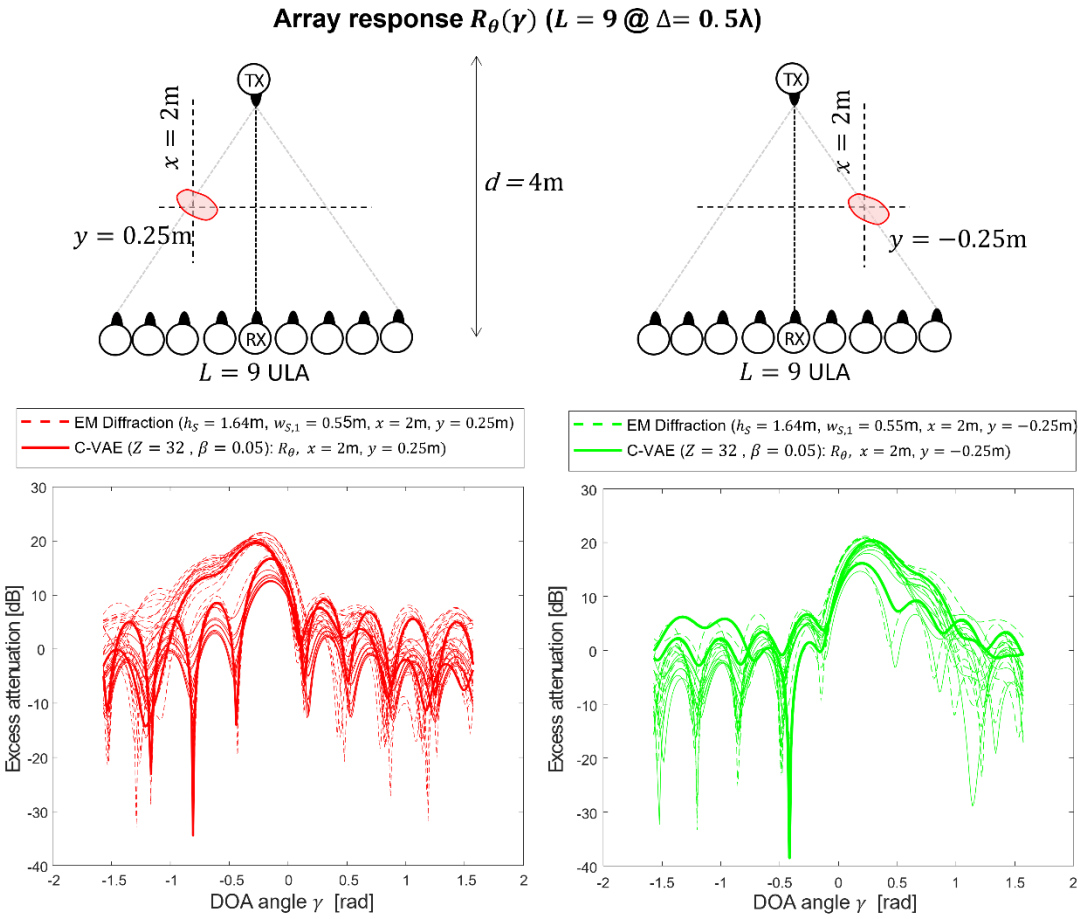


Figure 8: Array layout setups (top) and corresponding responses (bottom). Target is a human body with dimensions 1.65 m size 0.55 m / 0.25 m. Two subject positions are considered, namely x=2 m, y=0.25 m (left) and x=2 m y=-0.25m (right). Responses are obtain obtained using VAE generated field samples (green/red solid lines) and compared with the array response obtained with the EM body diffraction model (dashed lines) of Sect. 3.1.

Generation times of C-VAE are compared in the Table of Figure 9 with those obtained by classical diffraction models [11] as well as popular EM body simulator tools (Altair FEKO® software). VAE-based generation of the samples is about one order of magnitude faster than EM diffraction model computation, which also depends on the chosen numerical integration configurations (i.e., tiled integration method, and absolute error tolerance), target size, and antenna configuration. The generative model can produce up to up to 100-150 samples per second, which is reasonable considering typical body movement speeds (max. 1 m/s). The tool is therefore suitable for application in dynamic scenarios, which is the goal of WP4.

Model	**Configuration** (L links)	**Generation time*** [sec / I+Q sample / link]
C-VAE (E_θ generation)	$Z = 32$ $Z = 16$	5.3×10^{-2} 6.1×10^{-2}
EM Diffraction (omnidirectional antennas)	$\epsilon = 10^{-6}$ (num. integ.) $\epsilon = 10^{-3}$	0.78 3.1
Full wave (i.e., EM FEKO® simulations)	$L = 81$	> 240

Figure 9: CSI model-based generation time analysis: C-VAE vs diffraction and FEKO® simulations.

In Figure 10, we report the body-induced excess attenuation values measured by each antenna element for the corresponding subject positions (black, red and green colors). Responses consider the effect of small body movements (5 cm) around the nominal subject position. FEKO® simulated results are shown in solid lines, while, in dashed lines, we reported the corresponding C-VAE generated attenuation samples reproducing EM body diffraction effects. In general, the GNN mostly underestimates the FEKO® results, due to different hypotheses about body composition and shape, namely about 4 dB for positions on the LOS and 2 dB for positions outside (y = 0.25 m, y = −0.25 m). A possible remedy, to be investigated in the future, is to train the GNN network using examples from both diffraction model and FEKO® simulation samples, with the goal of improving the GNN generalization capabilities.

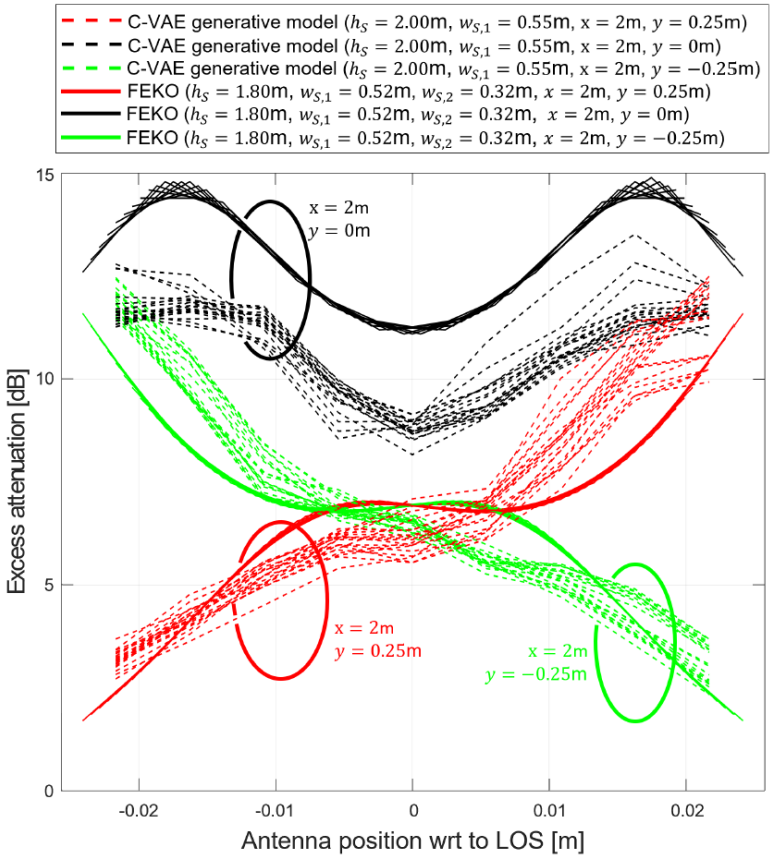


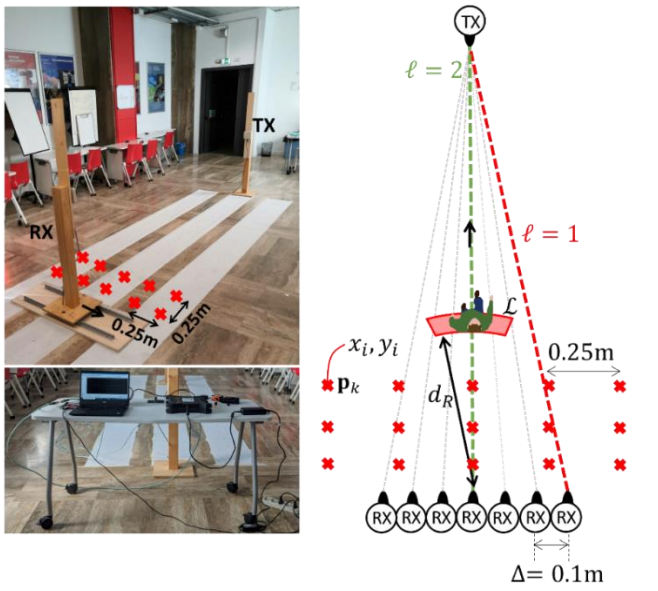
Figure 10: Excess attenuation values obtained by the FEKO® software (solid) and compared with C-VAE generation (dashed) for varying antennas of the array (9 antennas) and target dimensions. Three nominal subject positions are considered namely x=2 m, y=0.25 m (red), x=2 m, y=0 m (black) and x=2 m, y=-0.25 m (green). For all cases, attenuation responses account for small body movements around the nominal positions.

6.2. Reproduction of measured RF data

A specific experimental use case is considered in this section. The goal is to demonstrate the effectiveness of the proposed GNN tools to reproduce the EM data and to support a real-time localization process. The EM-informed generative tool has been thus validated with measurements taken in a hall of size 6.1 m x 14.4 m as shown in Figure 11. Both TX and RX are equipped with directional antennas with parameters summarized in the table embedded in the same figure. A mechanical handling mechanism is used to move the RX antenna at specified positions where RF measurements on multiple links are collected. The target is located in K=75 marked positions which belong to a regular 2D grid as shown in the same figure.

A tracking-generator-enabled spectrum analyzer is used to measure the RSS in the 2.4-2.5 GHz band and over 81 frequencies with 1.25 MHz spacing. For each frequency and target position under test, 500 consecutive time samples are acquired in 1 minute (with 120 ms sampling time). The human target (one of the authors who volunteered) has height 1.80 m and traversal max. and min.

sizes approximated as 0.55 m and 0.25 m, respectively. Further details on the dataset structure will be published in D4.1.



Start/stop frequency: 2.4GHz – 2.5GHz	**Power**: 0 dBm
Frequency spacing: 1.25MHz	**Antenna HPBW**: H (60 deg.), V (76 deg.)
Resolution: 100 kHz	**Antenna gain**: 9 dBi

Figure 11: Measurement setup, TX/RX antennas, linear guide system for RX antenna positioning. Explored target positions around the Fresnel area.

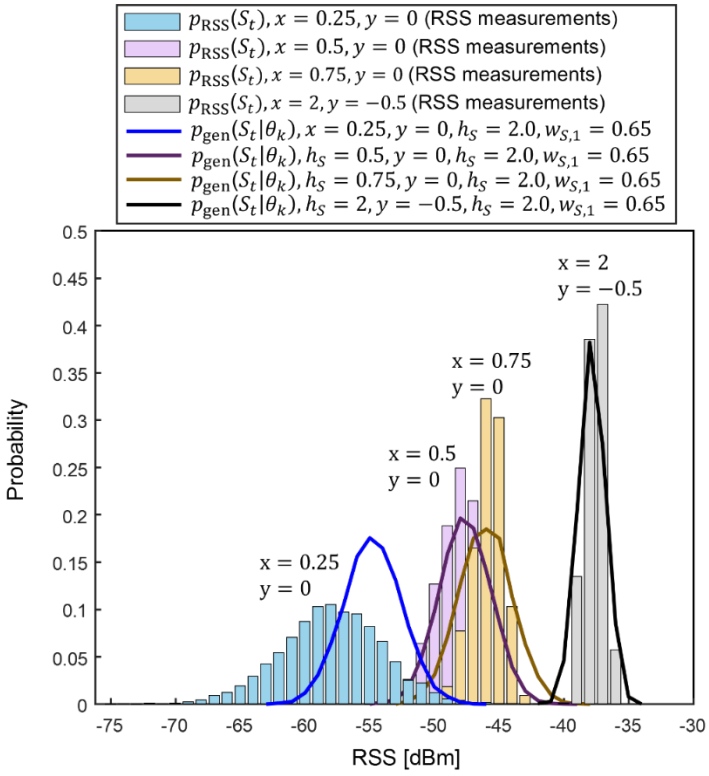


Figure 12: Generated RSS samples (solid lines) compared with RSS measurements (histograms) for link $\ell = 2$ in Figure 11 and considering 4 nominal positions of the human blockage, namely x=0.25m, y=0 (blue color), x=0.5m, y=0 (violet color), x=0.75m, y=0 (yellow color).

Figure 12 compares the generated RSS samples (by following the procedure illustrated in [40]) with RSS measurements at 2.4 GHz for the scenario in Figure 11 and link $\ell = 2$. The subject is standing while performing small movements around 4 nominal positions, namely x=0.25 m, y=0 (blue color), x=0.5 m, y=0 (violet color), x=0.75 m, y=0 (yellow color). Position outside the Fresnel area is x=2 m, y=-0.5 (black color). The generative model can be effectively used to predict the true RSS values for both the considered links with average error of 2.7 dB and max standard deviation error of 2.9 dB. On the other hand, as also previously observed, the C-VAE generation tool seems to under-estimate the observed RSS values for target positions close to the transmitter or receiver, i.e., x=3 m, since it is trained using diffraction models.

Distance	Estimated prior	C-VAE prior	Uniform prior
0.5m	0.95/0.98	0.92/0.85	0.67/0.48
1m	0.92/0.91	0.76/0.66	0.35/0.20
1.5m	0.92/0.90	0.87/0.67	0.41/0.55
2m	0.89/0.88	0.70/0.64	0.47/0.30
2.5m	0.78/0.85	0.47/0.69	0.19/0.18
3m	0.86/0.89	0.73/0.77	0.18/0.19
3.5m	0.76/0.74	0.62/0.83	0.19/0.26

Figure 13: Precision/recall metrics for distance estimation. Estimated prior from calibration data, C-VAE prior and uniform prior cases compared.

6.3. Passive localization example

In this section, we discuss an example of passive localization. The goal is to detect the distance d_R of the target from the multi-antenna RX device (Figure 11) in real-time, as in typical radar applications.

The proposed use case is critical in industrial scenarios where human workers operate in areas featuring increasing levels of safety or privacy. Enforcing safety/privacy constraints requires the real-time monitoring and tracking of human subjects. The estimated position of the target relative to RX is obtained as shown in Sect. 2.

In the following, we analyze the precision $p_1(d_R)$ and recall$p_2(d_R)$ metrics: the recall metric measures how often the algorithm correctly identifies the target distance from all the true positive counts, while the precision indicates how often the algorithm is correct when predicting the target

distance. In particular, the Table of

Distance	Estimated prior	C-VAE prior	Uniform prior
0.5m	0.95/0.98	0.92/0.85	0.67/0.48
1m	0.92/0.91	0.76/0.66	0.35/0.20
1.5m	0.92/0.90	0.87/0.67	0.41/0.55
2m	0.89/0.88	0.70/0.64	0.47/0.30
2.5m	0.78/0.85	0.47/0.69	0.19/0.18
3m	0.86/0.89	0.73/0.77	0.18/0.19
3.5m	0.76/0.74	0.62/0.83	0.19/0.26

Figure 13 analyzes precision and recall for varying distance d_R from the RX and compares three cases:

1) **Estimated prior**: assumes the prior probability $p(\boldsymbol{E_\theta})$ being estimated from large calibration measurements (ideal case, human subjects involved during training data collection);
2) **C-VAE prior**: adopts $p_{gen}^{\mathrm{VAE}}(\boldsymbol{E_\theta})$ to generate synthetic samples using the C-VAE generator tool (no human subjects involved during training);
3) **Uniform prior**: the prior is replaced with a uniform probability function.

Note that scenario (1) gives the best performance, as expected; on the other hand, it requires time-consuming data collection and a calibration stage which might be not feasible in practice. Case (3) corresponds to the worst-case scenario since no prior information on body-induced attenuations are available. Finally, case (2) does not need any calibration (human subjects are thus not involved in calibration activities) as it uses the C-VAE tool to real-time generate synthetic samples from the prior. From the results in the table, the performance of C-VAE prior scenario approaches the estimated prior case, with an average drop of about 10%. This initial analysis will be expanded through the HOLDEN project and will serve as the basis for some of the case studies in WP6.

7. Conclusion

This deliverable discussed fundamental signal processing tool for management of hologram structures in the form of RF tensor databases. We introduced classical Bayesian tools for RF sensing focusing specifically on passive radio localization. Next, we proposed the use of EM-informed Generative Neural Network (GNN) tools deployed for model-based reconstruction and prediction of body-induced EM effects, such as diffraction. The deliverable reported the generation time analysis and the explored applications are in dynamic tracking of body motions, specifically focusing on passive radio sensing and localization. A Variational Auto-Encoder (VAE) tool, namely the Conditional VAE (C-VAE), is designed to generate samples of the targeted EM model through latent variable encoding/decoding neural network operations. Adaptations of Generative Adversarial Networks (GAN) are also considered for comparative analysis. The proposed methods implement model-based reconstruction of the RF samples and are set to reproduce both RF signal attenuations, i.e., Received Signal Strength (RSS), as well as base-band Channel State Information (CSI) for full EM analysis of human body blockage.

GNNs produce observations of CSI sampled from the Bayesian prior probability which supports Bayesian estimation problems. They are an essential component in state-of-the-art segmentation and motion detection tools targeting anomaly detection analysis over multidimensional time-series data. The possibility of generating large CSI tensor structures representing the full RF radiation field is useful in holographic methods and microwave imaging techniques based on Synthetic Aperture Radar (SAR). Finally, the proposed tools have been proved to be effective in reproducing body motions in user selectable locations. The property is instrumental to privacy selective sensing policies. Generated samples might also serve as synthetic training data for supervised or semi-supervised machine learning tasks. They thus reduce the need of personal data collection, that could be used maliciously for person (re-)identification.

Beside the advantages, the proposed generators require hyper-parameter optimization to achieve a good compromise between average reproduction accuracy and generalization capability. The former measures how close the prediction is to the training observations, while the latter quantifies the randomness of generated samples which is useful to predict effects not seen during training. When compared with real measurements, the generated tools appear to underestimate some human blockage effects, i.e. for small targets placed close to the receivers. This opens the room for possible improvements. The considered generative systems are currently trained to reproduce EM blocking effects obtained by scalar diffraction methods. Training the system with different EM blockage models, such as full-wave solutions or Method of Moments (MoM), and/or using more accurate body models, might increase the generalization capabilities.

We expect the proposed models to become indispensable tools in different scenarios all relevant in HOLDEN project, and particularly WP3, WP4 and WP6.

3. References

[1] M. Youssef, et al., "Challenges: Device-free passive localization for wireless environments," 13th Annu. ACM Int. Conf. Mobile Comput. Netw. (MobiCom'07), pp. 222–229, Sep. 2007.

[2] J. Wilson, et al., "Radio tomographic imaging with wireless networks," IEEE Trans. on Mobile Comp., vol. 9, no. 5, pp. 621–632, May 2010.

[3] M. Seifeldin, et al., "Nuzzer: A large-scale device-free passive localiza- tion system for wireless environments," IEEE Trans. Mobile Comput., vol. 12, no. 7, pp. 1321–1334, Jul. 2013.

[4] B. Rodrigues, et al., "BluePIL: A Bluetooth-based PassIve localization method," 2021 IFIP/IEEE International Symposium on Integrated Network Management (IM'21), pp. 28–36, May 2021.

[5] S. Savazzi, et al., "On the use of stray wireless signals for sensing: A look beyond 5G for the next generation of industry," *Computer*, vol. 52, no. 7, pp. 25-36, July 2019, doi: 10.1109/MC.2019.2913626.

[6] O. Kaltiokallio, et al., "A Novel Bayesian Filter for RSS-Based Device-Free Localization and Tracking," IEEE Transactions on Mobile Computing, vol. 20, no. 3, pp. 780–795, Mar. 2021

[7] S. Palipana, et al., "Recent advances in RF-based passive device-free localisation for indoor applications," Ad Hoc Netw., vol. 64, pp. 80– 98, 2017

[8] V. Rampa, et al., "EM models for passive body occupancy inference," *IEEE Antennas and Wireless Propagation Letters*, vol. 16, pp. 2517-2520, 2017, doi: 10.1109/LAWP.2017.2728369.

[9] F. Fieramosca, et al., "On the impact of the antenna radiation patterns in passive radio sensing," *IEEE Antennas and Wireless Propagation Letters*, vol. 23, no. 2, pp. 503-507, Feb. 2024, doi: 10.1109/LAWP.2023.3327955.

[10] F. Fieramosca, et al., "Modelling of the Floor Effects in Device-Free Radio Localization Applications," 17th EU Conf. on Ant. and Prop. (EuCap), Florence, Italy, pp. 1–5, Mar. 2023, doi: 10.23919/EuCAP57121.2023.10133182.

[11] V. Rampa, et al., "Electromagnetic Models for Passive Detection and Localization of Multiple Bodies," *IEEE Transactions on Antennas and Propagation*, vol. 70, no. 2, pp. 1462-1475, Feb. 2022, doi: 10.1109/TAP.2021.3111405.

[12] B. R. Hamilton, et al., "Propagation modeling for radio frequency tomography in wireless networks," IEEE J. Sel. Topics Signal Process., vol. 8, no. 1, pp. 55–65, Feb. 2014.

[13] S. Kianoush, et al., "Pre-deployment performance assessment of device-free radio localization systems," *2016 IEEE International Conference on Communications Workshops (ICC)*, Kuala Lumpur, Malaysia, 2016, pp. 1-6, doi: 10.1109/ICCW.2016.7503754.

[14] Z. Zhao, et al., "Generative Models for Inverse Imaging Problems: From mathematical foundations to physics-driven applications," IEEE Signal Processing Magazine, vol. 40, no. 1, pp. 148–163, Jan. 2023.

[15] T. Shan, et al., "Physics-informed supervised residual learning for electromagnetic modeling," IEEE Transactions on Antennas and Propagation, vol. 71, no. 4, pp. 3393–3407, Apr. 2023.

[16] L. Lu, et al., "Physics-informed neural networks with hard constraints for inverse design," SIAM Journal on Scientific Computing, vol. 43, no. 6, B1105-B1132, 2021.

[17] P. Di Barba, "Future Trends in Optimal Design in Electromagnetics," IEEE Transactions on Magnetics, vol. 58, no. 9, pp.1–4, Sep. 2022.

[18] M. Baldan, et al., "Physics-Informed Neural Networks for Inverse Electromagnetic Problems," IEEE Transactions on Magnetics, vol. 59, no. 5, pp.1–5, May 2023

[19] S. Savazzi et al., "A physics-informed generative model for passive radio-frequency sensing," *2023 IEEE-APS Topical Conference on Antennas and Propagation in Wireless Communications (APWC)*, Venice, Italy, 2023, pp. 056-061, doi: 10.1109/APWC57320.2023.10297498, [Online] Available: https://arxiv.org/abs/2310.04173.

[20] S. Bond-Taylor, et al., "Deep generative modelling: A comparative review of VAEs, GANs, normalizing flows, energy-based and autoregressive models," IEEE Trans. Pattern Anal. Mach. Intell., vol. 44, no. 11, pp. 7327–7347, Nov. 2021.

[21]I. Goodfellow, et al., "Generative adversarial nets," 27th Int. Conf. Neural Inf. Process. Syst., pp. 2672–2680, 2014.

[22] B. Bartlett, "A "generative" model for computing electromagnetic field solutions," 2018. [Online]. Available: https://cs229.stanford.edu/proj2018/report/233.pdf

[23] T. Shan, et al., "Study on a Fast Solver for Poisson's Equation Based on Deep Learning Technique," IEEE Transactions on Antennas and Propagation, vol. 68, no. 9, pp. 6725–6733, Sep. 2020.

[24] L. Lu, et al., "DeepXDE: A deep learning library for solving differential equations," SIAM Rev., vol. 63, no. 1, pp. 208–228, 2021.

[25] A. Khan, et al., "Physics Informed Neural Networks for Electromagnetic Analysis," IEEE Transactions on Magnetics, vol. 58, no. 9, pp. 1–4, art. no. 7500404, Sep. 2022

[26] V. Rampa et al., "An EM body model for Device-Free Localization with multiple antenna receivers: A first study," *2023 IEEE-APS Topical Conference on Antennas and Propagation in Wireless Communications (APWC)*, Venice, Italy, 2023, pp. 164-169, doi: 10.1109/APWC57320.2023.10297446, [Online] Available: https://arxiv.org/abs/2405.09558.

[27] "RSA500 Series Real Time Spectrum Analyzers," [online], accessed: Feb 6, 2024. [Online] Available: https://www.tek.com/en/products/spectrum-analyzers/rsa500

[28]Body Tissue Dielectric Parameters: https://www.fcc.gov/general/body-tissue-dielectric-parameters, FCC, US. Accessed: 11 Oct. 2023.

[29] D. Halperin, et al., "Predictable 802.11 packet delivery from wireless channel measurements," ACM SIGCOMM Computer Communication Review, vol. 41, no. 4, p. 159–170, 2011.

[30] Y. Xie, et al., "Precise power delay profiling with commodity Wi-Fi,"IEEE Transactions on Mobile Computing, vol. 18, no. 6, pp. 1342–1355, 2018.

[31] S. Chauhan, et al., "IEEE 802.11be: A Review on Wi-Fi 7 Use Cases,"Proc. of the 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO'21), pp. 1–7, 2021.

[32] L. Zhang, et al., "DeFi: Robust Training-Free Device-Free Wireless Localization With WiFi,"IEEE Transactions on Vehicular Technology, vol. 67, no. 9, pp. 8822–8831, Sept. 2018.

[33] S. Shukri, et al., "Enhancing the radio link quality of device-free localization system using directional antennas,"Proc. of the 7th International Conference on Communications and Broadband Networking, pp. 1–5, Apr. 2019.

[34] D. Garcia, et al., "POLAR: Passive object localization with IEEE 802.11ad using phased antenna arrays,"Proc. of the IEEE Conference on Computer Communications, pp. 1838–1847, Jul. 2020

[35] D. P. Kingma, et al., "Auto-encoding variational Bayes," Proc. of Int. Conf. Learn. Represent (ICLR), 2013. [Online] Available: https://arxiv.org/abs/1312.6114

[36] F. Fieramosca, V. Rampa, M. D'Amico, S. Savazzi, "Physics-Informed Generative Neural Networks for RF Propagation Prediction with Application to Indoor Body Perception," 18th European Conference on Antennas and Propagation (EuCAP 2024), pp. 1-5, Glasgow, Scotland, Mar. 17-22, 2024, DOI: 10.23919/EuCAP60739.2024.10501077, [Online] Available: https://arxiv.org/abs/2405.02131.

[37] Savazzi, S. Kianoush, V. Rampa and U. Spagnolini, "Cellular Data Analytics for Detection and Discrimination of Body Movements," *IEEE Access*, vol. 6, pp. 51484-51499, 2018, doi: 10.1109/ACCESS.2018.286970.

[38] S. Lin, R. Clark, R. Birke, S. Schönborn, N. Trigoni and S. Roberts, "Anomaly Detection for Time Series Using VAE-LSTM Hybrid Model," ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 2020

[39] Jiehui Xu and Haixu Wu and Jianmin Wang and Mingsheng Long "Anomaly Transformer: Time Series Anomaly Detection with Association Discrepancy," International Conference on Learning Representations, 2022

[40] Stefano Savazzi, et al. "Electromagnetic-informed generative models for passive RF sensing and perception of body motions," *IEEE Open Journal of Antennas and Propagation*, doi: 10.1109/OJAP.2024.3407199.

[41] M. Santoboni, R. Bersan, S. Savazzi, A. Zecchin, V. Rampa and D. Piazza, "Wireless LAN sensing with smart antennas," *2022 16th European Conference on Antennas and Propagation (EuCAP)*, Madrid, Spain, 2022, pp. 1-5, doi: 10.23919/EuCAP53622.2022.9769414, , [Online] Available: https://arxiv.org/abs/2205.00973.

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

8. List of figures