# CONVERGE Challenge: Multimodal Learning for 6G Wireless Communications

ICASSP 2026 SP Grand Challenge

Guidelines for Task 1 – Blockage Prediction

v1.0 (22/12/2025)

Jichao Chen
Communication Systems Department, EURECOM, France
jichao.chen@eurecom.fr

Filipe B. Teixeira
INESC TEC, Faculdade de Engenharia, Universidade do Porto, Portugal
filipe.b.teixeira@inesctec.pt

Francisco M. Ribeiro
INESC TEC, Faculdade de Engenharia, Universidade do Porto, Portugal
francisco.m.ribeiro@inesctec.pt

Ahmed Alkhateeb
Arizona State University, Tempe, USA
alkhateeb@asu.edu

Luis M. Pessoa
INESC TEC, Faculdade de Engenharia, Universidade do Porto, Portugal
luis.m.pessoa@inesctec.pt

Dirk Slock
Communication Systems Department, EURECOM, France
dirk.slock@eurecom.fr

December 22, 2025

**Abstract**

The CONVERGE Challenge: Multimodal Learning for 6G Wireless Communications aims to leverage synchronized visual and wireless data to support communication tasks such as blockage prediction, UE localization, channel prediction, and beam prediction. This report describes the blockage prediction task, the experimental setup used for data collection, the dataset structure and annotations, and the evaluation protocol and participant guidelines.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Task Overview

## 1.1 Description

The Blockage Prediction task aims to forecast future blockage events in a high-frequency (FR2/mmWave) wireless link by leveraging multimodal inputs, including synchronized visual data and wireless measurements. Accurate blockage prediction enables proactive network actions, such as beam switching, handover, or resource reallocation, before link degradation occurs.

## 1.2 Objective

The main objective of this task is to predict a blockage at $t + 142\,\mathrm{ms}$ for each video frame. This time is consistent with the frame rate of 7 fps.

# Chapter 2

# Experimental Setup

## 2.1 Chamber Setup

To collect the data for this task, we have setup the CONVERGE Chamber in an environment similar to the one illustrated in Fig. 2.1.

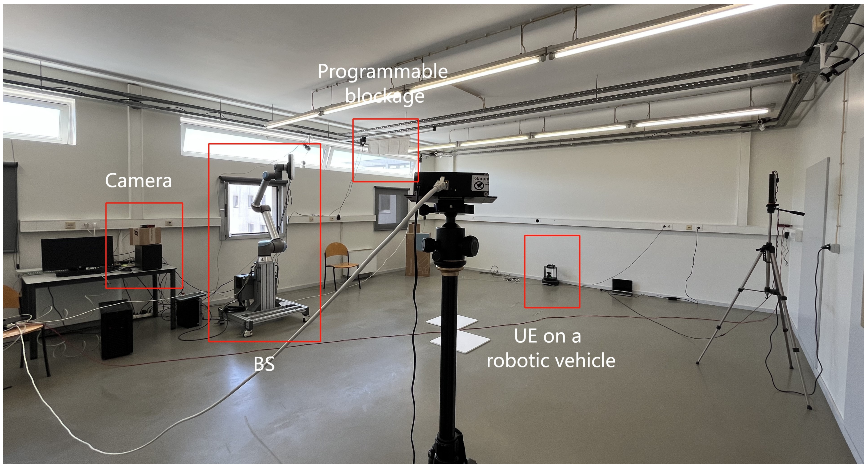


Figure 2.1: CONVERGE chamber with mobile BS, UE, and camera in Porto, Portugal.

## 2.2 Hardware Components

- **Mobile Base Station (BS)**: Mounted on a Universal Robotics UR-10e robotic arm. Implemented with a LiteOn FR2 Radio Unit as the mmWave front-end, integrated with the OpenAirInterface (OAI) software stack to implement the 5G NR gNB. The robotic arm is programmed to do left/right movements to create Line-of-Sight (LoS) / Non-Line-of-Sight (NLoS) conditions.

- **User Equipment (UE)**: RM530F UE from Quectel, mounted on a tripod, fitted with mmWave phased arrays oriented in multiple directions.
- **RF-shield curtain**: Custom-made programmable RF shield curtain capable of dynamically blocking the communication path between the transmitter and receiver. The curtain is capable of adding nearly 30 dB attenuation between the gNB and UE.
- **Nerian 3D Camera**: Precision Time Protocol (PTP) synchronized Nerian Ruby 3D RGB-D camera positioned at the BS to capture the surrounding environment and UE interactions.

## 2.3 Experiments

The data was collected using the CONVERGE software stack and is divided into five experiments.

### 2.3.1 Exp 1. Blockage/non-blockage central beam, with robotic arm

The robotic arm moves left/right to create blockage/non-blockage scenarios between gNB and UE, which are aligned to the central beam. The signal is blocked by an RF-shield curtain providing approximately 30 dB attenuation. There are two blockage/non-blockage runs. Fig. 2.2 shows the mobile gNB setup with the camera mounted on top of the RU and Fig. 2.3 shows a blockage scenario with the robotic arm on the left position, aligned with the UE behind the courtain.


Figure 2.2: Detail of the mobile gNB.

Figure 2.3: Mobile gNB aligned with the UE behind the RF-shield coutain.

### 2.3.2 Exp 2. Blockage/non-blockage non-aligned, with robotic arm

The robotic arm moves left/right to create blockage/non-blockage scenarios between gNB and UE, which are not aligned to the central beam (gNB points slightly to the right). The signal is blocked by an RF-shield curtain providing approximately 30 dB attenuation. There is one run. Fig. 2.4 shows the new alignment of the RU on the blockage position and non-blockage position (Fig. 2.5.

Figure 2.4: New alignment of the RU, blockage position.

Figure 2.5: New alignment of the RU, non-blockage position.

### 2.3.3 Exp 3. Curtain going up with steady robotic arm

The robotic arm is steady while the RF-shield curtain moves up. The signal is blocked by an RF-shield curtain providing approximately 30 dB attenuation. In this experiment the UE is always blocked.

### 2.3.4 Exp 4. Curtain going down with steady robotic arm

The robotic arm and gNB remain steady while the RF-shield curtain moves down to create a non-blockage/blockage scenario. The signal is blocked by an RF-shield curtain providing approximately 30 dB attenuation. There is one non-blockage/blockage run. Fig. 2.6 shows the setup of Exp 4.

Figure 2.6: Courtain going down with steady robotic arm.

### 2.3.5 Exp 5. Curtain going up with steady robotic arm (blockage to non-blockage)

The robotic arm and gNB remain steady while the RF-shield curtain moves up to create a blockage/non-blockage scenario. The signal is blocked by an RF-shield curtain providing approximately 30 dB attenuation. There is one blockage/non-blockage run.

# Chapter 3

# Dataset

## 3.1 Modalities

For each experiment, we collected radio and video data.

### 3.1.1 Radio data

Radio data is collected via (1) the E2 interface and (2) SRS measurements.

**E2 interface** Provides PTP-synced data every 50 ms, including parameters such as:

- `ue_id`: User Equipment identifier (internal UE index)
- `cellid`: Cell identifier
- `in_sync`: Synchronization status of the UE
- `rnti`: Radio Network Temporary Identifier (RNTI)
- `dlBytes`: Downlink bytes (throughput counter)
- `dlMcs`: Downlink Modulation and Coding Scheme (MCS)
- `dlBler`: Downlink Block Error Rate (BLER)
- `ulBytes`: Uplink bytes (throughput counter)
- `ulMcs`: Uplink Modulation and Coding Scheme (MCS)
- `ulBler`: Uplink Block Error Rate (BLER)
- `ri`: Rank Indicator (MIMO rank)
- `pmi`: Precoding Matrix Indicator (PMI)

- `phr`: Power Headroom Report (PHR)
- `pcmax`: Maximum UE transmit power
- `rsrq`: Reference Signal Received Quality (RSRQ)
- `sinr`: Signal-to-Interference-plus-Noise Ratio (SINR)
- `rsrp`: Reference Signal Received Power (RSRP)
- `rssi`: Received Signal Strength Indicator (RSSI)
- `cqi`: Channel Quality Indicator (CQI)
- `pucchSnr`: PUCCH Signal-to-Noise Ratio
- `puschSnr`: PUSCH Signal-to-Noise Ratio
- `dlQm`: Downlink modulation order ($Q_m$)
- `ulQm`: Uplink modulation order ($Q_m$)

**SRS data** Captured every 10 ms for each antenna (0 and 1). The estimated channel is obtained via Least Squares (LS) estimation by dividing received signal by the pilot reference, followed by interpolation to filter noise and estimate subcarriers without an associated pilot. Given 200 MHz bandwidth, 128 PRBs are used. Multiplying PRBs by 12 yields 1536 channel estimates per frame, per antenna, every 10 ms.

### 3.1.2 Video data

Video is captured at 7 fps (one frame every 142 ms), from the Radio Unit (RU) towards the UE. The camera is mounted on top of the RU.

## 3.2 Ground Truth Annotations

Ground truth is provided by manual annotation of blockage states in the `annotations/` folder. Three states are defined:

- `no`: UE fully visible from the gNB camera; no signal degradation expected.
- `partial`: UE partially blocked by the curtain; variable signal conditions expected.
- `full`: UE not visible; strong signal degradation expected.

## 3.3 Dataset Layout

```
dataset/
  calibration/nerian_gnb_1_calib.yaml
  task1/exp1..exp5/
    annotations/*.csv
    radio/E2-*.csv
    radio/SRS-*.json
    video/frames.csv
    video/color/
    video/disparity/
  index.csv
```

The `nerian_gnb_1_calib.yaml` provides a camera calibration kit that allows 3D reconstruction from disparity frames.

## 3.4 Index File

`dataset/index.csv` has one row per scenario with dataset-relative paths.

### Columns

- `task`: `task1, task2 or task3`
- `scenario_id`: `exp1..exp8`
- `scenario_name`: short label
- `video_frames_csv`: path to frame index CSV
- `video_color_dir`: path to color frames directory
- `video_disparity_dir`: path to disparity frames directory
- `radio_e2`: path to E2 CSV
- `radio_srs`: path to SRS JSON
- `annotation`: path to annotation CSV
- `notes`: short scenario description

## 3.5 Timestamps

- Video timestamps are in seconds (float).
- Radio timestamps are in seconds (float) and preserve absolute time. The radio files were normalized to seconds (suffix `_sec`).
- Annotation timestamps are in seconds (float).

## 3.6 Task 1 Annotation Format

The annotation CSV contains per-frame labels:

```
timestamp,state
```

where `state` $\in$ {`no`, `partial`, `full`}. The file name varies per experiment; use `index.csv` to locate it.

## 3.7 Notes on `frames.csv`

`video/frames.csv` uses paths relative to the `video/` directory:

```
color/img_XXXX.png
disparity/img_XXXX.png
```