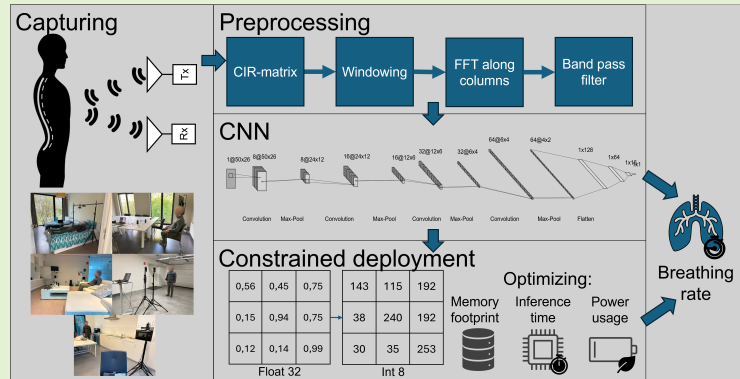


Low-cost Embedded Breathing Rate Determination Using 802.15.4z IR-UWB Hardware for Remote Healthcare

Anton Lambrecht, Stijn Luchie, Jaron Fontaine, Ben Van Herbruggen, Adnan Shahid and Eli De Poorter

Abstract—Respiratory diseases account for a significant portion of global mortality. Affordable and early detection is an effective way of addressing these ailments. To this end, a low-cost commercial off-the-shelf (COTS), IEEE 802.15.4z standard compliant impulse-radio ultra-wideband (IR-UWB) radar system is used to estimate human respiration rates. We propose a convolutional neural network (CNN) specifically adapted to predict breathing rates from ultra-wideband (UWB) channel impulse response (CIR) data, and compare its performance with both other rule-based algorithms and model-based solutions. The study uses a diverse dataset, incorporating various real-life environments to evaluate system robustness. To facilitate future research, this dataset will be released as open source. Results show that the CNN achieves a mean absolute error (MAE) of 1.73 breaths per minute (BPM) in unseen situations, significantly outperforming rule-based methods (3.40 BPM). By incorporating calibration data from other individuals in the unseen situations, the error is further reduced to 0.84 BPM. In addition, this work evaluates the feasibility of running the pipeline on a low-cost embedded device. Applying 8-bit quantization to both the weights and input/output tensors, reduces memory requirements by 67% and inference time by 62% with only a 3% increase in MAE. As a result, we show it is feasible to deploy the algorithm on an nRF52840 system-on-chip (SoC) requiring only 46 KB of memory and operating with an inference time of only 199 ms. Once deployed, an analytical energy model estimates that the system, while continuously monitoring the room, can operate for up to 268.2 days without recharging when powered by a 20 000 mAh battery pack. For breathing monitoring in bed, the sampling rate can be lowered, extending battery life to 313.8 days, making the solution highly efficient for real-world, low-cost deployments.

Index Terms—breathing rate determination, convolutional neural network (CNN), impulse-radio ultra-wideband (IR-UWB), remote healthcare



I. INTRODUCTION

According to Eurostat, respiratory diseases accounted for 6.1% of all deaths in the EU in 2021 [1]. This statistic highlights the significant impact of respiratory issues on public health. Early detection effectively addresses these ailments [2], [3], as timely identification significantly improves recovery prospects and enables more effective symptom management. Proper monitoring of vital signs plays a significant role in facilitating this process, specifically breathing rate (BR) is a key metric for this application [4]. A variety of devices already exist to monitor BR reliably. One example is a

nasal pressure transducer, which measures airflow through the nostrils to track inhalation and exhalation phases. Another is the respiratory effort belt, worn around the chest or abdomen to monitor chest wall movement. While effective and widely used, these devices require direct skin contact, making them unsuitable for patient scenarios like burn victims, where attaching sensors could cause further harm. Additionally, constant monitoring throughout the day using contact-based solutions is inconvenient for individuals. A remote monitoring system could overcome these challenges, enabling respiratory data collection without causing discomfort.

Such systems are particularly valuable for long-term health tracking, including elderly care and sleep quality assessment, the latter of which serves as a vital proxy for physiological stress and alertness [5]. By eliminating the need for wearable hardware, BR can be monitored over extended durations, shifting the focus from occasional snapshots to the analysis of long-term physiological trends.

This study utilizes a non-invasive impulse-radio ultra-

Received 21 February 2026; The research that led to these results has received funding by the DistriMuSe project (HORIZON-KDT-JU-2023-2-RIA) with Grant No 101139769, from the Belgian Defense through contract number 24DEFRA009 and from the FWO PESSO project with Grant No 3G018522.

The authors are affiliated with IDLab, Department of Information Technology, Ghent University-imec, iGent Tower, Technologiepark-Zwijnaarde 126, B-9052 Ghent, Belgium. (email: Anton.Lambrecht@UGent.be)

wideband (IR-UWB) radar system to measure respiratory rates by capturing chest wall reflections. The system does not require on-body sensors and is installed in front of the person. This makes it well-suited for monitoring patients in bed, tracking sedentary users in an office or domestic environment, and observing drivers or passengers during stationary periods.

For non-invasive breathing monitoring, multiple radar technologies are available, including mmWave, frequency-modulated continuous wave (FMCW) radar or ultra-wideband (UWB) radar. In this study, UWB radio technology was chosen for its use of large bandwidth, allowing the transmission of very short radio pulses lasting nanoseconds [6]. These very short UWB pulses and their reflections can be timestamped with high accuracy, allowing the detection of small chest movements. These precise distance measurements of chest movements are crucial for detecting breathing signals. Compared to alternative radar technologies, UWB also requires lower power consumption, is more affordable and generates minimal interference with other radio signals. Moreover, UWB radios can also be used for other purposes, such as communication or localization, making it a suitable technology for multi-functional systems that provide integrated sensing, localization and communication [7], [8]. Analyzing data from the radar hardware involves selecting from various processing techniques. While most studies on BR determination rely on signal processing methods, data-driven based approaches have shown promising results, particularly in challenging conditions [9], [10].

The key contributions of this work are the following:

- A convolutional neural network (CNN)-based processing pipeline is proposed for breathing determination that shows significant accuracy improvement as compared to rule-based, and model-based processing techniques.
- The proposed solution is evaluated with an improved strategy considering multiple environments, subjects, and orientations, providing a more thorough assessment of generalization to unseen conditions.
- The algorithmic complexity of each processing step is analyzed, providing insights into scalability and computational feasibility on resource-constrained devices.
- The effects of model quantization on accuracy, memory usage, execution time, and energy consumption are modeled and evaluated, highlighting practical trade-offs for embedded deployment.
- An analytical energy model is developed, covering the radar hardware, signal-processing pipeline and the model inference stage, enabling accurate predictions of execution time, memory footprint, and energy consumption.
- An open source dataset is provided consisting of UWB breathing data and annotated ground truth measurements captured using a low-cost commercial off-the-shelf (COTS) IR-UWB device.¹

The remainder of the paper is organized as follows. First, Section II provides an overview of the current state of the art. Next, Section III provides a mathematical model of the system. Section IV describes how the dataset was captured. Section V

describes the algorithms developed to extract BR from the channel impulse response (CIR) data. Section VI analyses the obtained results. Afterwards, Section VII describes how the solution can be optimized and deployed on embedded hardware and analyses the reduced footprint, inference time, and power consumption. Finally, Section VIII concludes the paper and describes future work.

II. RELATED WORK

Table I provides an overview of related scientific papers. As seen in the table, most studies employ radar systems designed by Novelda or Pulson devices, which consistently use high bandwidths of 1.4 GHz or higher. A higher bandwidth enhances the precision of the CIRs, aiding in capturing subtle breathing patterns. However, such systems require more powerful downstream processing, reducing the cost-effectiveness of the system. These high-bandwidth devices are also not compliant with the IEEE 802.15.4z standard, lacking standardized bandwidth and center frequency configurations. Additionally, they are significantly more expensive, costing at least 10 times more than commercial standard-compliant transceivers such as Qorvo's DW3000, which is priced under €10. Finally, most of these Novelda and Pulson devices are no longer commercially available.

Furthermore, as discussed in Section I, capturing breathing movements is challenging since individuals are not always stationary or aligned with radar devices. Validation on diverse real-life conditions is essential to claim robustness. Columns 7–13 of Table I summarize these considerations across the reviewed papers, covering tested distances and rotation angles, variation across environments and postures, the presence of multiple subjects in the dataset, multi-person scenarios (multiple subjects measured simultaneously), and subject motion during measurements.

When considering the solutions that use IEEE 802.15.4z-compliant hardware, the following work is most relevant. Farsaei et al. [16] utilize a 500 MHz standard-compliant radar system, but their hardware is proprietary and inaccessible, disqualifying it as COTS. Only 2 real-life considerations are tested using a peak-finding algorithm. In general, validation on BR determination is limited, as only 2 CIR-windows are considered in their validation set. Numan et al. [23] stand out for using the NCJ29D5 radar from NXP. This is a COTS device costing less than €10. The paper only evaluates generalizability to new individuals and over multiple distances. This work focuses more on presence detection and classifying low, normal, and high BRs rather than predicting the exact respiratory rates. BR determination is done with a peak detection-based solution, and its validation is restricted to 9, 10-second CIR windows.

The final column of Table I shows whether a solution uses AI. All but one prior work utilizes a model-based BR estimator. This exception is Zhou et al. [22], which addresses a comparable number of challenges compared to this paper but lacks evaluations in multiple environments or with subjects in different postures. However, it does excel in analyzing non-static scenarios and multi-person setups. Their approach uses

¹<https://gitlab.ilabt.imec.be/datasets/uwb-vital-sign-monitoring>

Low-cost					Real life considerations								
Year	Device	Bandwidth (GHz)	IEEE compliant	COTS	Embedded impl.	Distance	Rotation	Environment	Posture	Multiple persons in dataset	Multi-person	Non static	AI-based
[11]	2017	NVA6201	2.3	-	-	-	-	-	-	✓	-	✓	-
[12]	2019	PulsOn 440	4.3	-	-	✓	-	-	-	✓	-	-	-
[13]	2020	X4	NA	-	-	-	✓	-	-	✓	-	-	-
[14]	2022	X4	NA	-	-	✓	✓	-	-	-	-	-	-
[15]	2022	X4M03	1.4	-	-	✓	-	-	-	✓	-	-	-
[16]	2023	Proprietary	0.5	✓	-	-	-	-	-	✓	✓	-	-
[17]	2020	X4	1.5	-	-	✓	✓	✓	-	-	-	-	-
[9]	2023	NVA6100	NA	-	-	-	-	✓	✓	✓	-	-	-
[10]	2024	X4M03	1.4	-	-	✓	✓	-	-	✓	-	-	-
[18]	2024	X4M05	1.4	-	-	✓	-	✓	-	✓	-	✓	-
[19]	2024	X4M06	NA	-	-	✓	✓	-	✓	-	-	✓	-
[20]	2022	X4m200	1.4	-	-	✓	-	✓	-	✓	✓	-	-
[21]	2017	NVA6201	2.3	-	-	✓	✓	-	-	✓	-	✓	-
[22]	2023	X4M03	1.4	-	-	✓	✓	-	-	✓	✓	✓	✓
[23]	2023	NCJ29D5	0.5	✓	✓	✓	-	-	-	✓	-	-	-
This paper		DW3000	0.5	✓	✓	✓	✓	✓	✓	✓	-	-	✓

TABLE I: Overview of related work, categorized by hardware type and application context. This work is the first to provide a truly low-cost, IEEE 802.15.4z-compliant, off-the-shelf available solution that can run on embedded hardware while coping with a variety of real-life environments and conditions. Indicated by the final column, using an artificial intelligence (AI)-approach in this context has not been explored extensively.

an Attention-Based CNN-long short-term memory (LSTM) model to filter out motion. Afterwards, the filtered data is analyzed with a rule-based BR estimator. In contrast to our work the authors do not analyze the computational costs or provide energy models to evaluate the suitability for embedded implementations. The lack of environmental testing also raises major concerns about generalizability, a key hurdle in AI-driven solutions.

The work described in this paper stands out from all works in Table I by using COTS IEEE 802.15.4z-compliant hardware, validating the system with multiple real-life considerations in mind and also analyzing the processing requirements of the solution. Compared to prior work, this study is the first to focus on analyzing the computational and energy requirement aspects required for embedded hardware compatibility.

III. SYSTEM MODEL

This study employs a radar-based system with a transmitter emitting signals that reflect off objects or people in the room. These reflections will be influenced by respiratory movements of the person in front of the radar and are subsequently captured by the receiver. The varying distance between the radar and the chest of the target is then used to extract the BR [24]. Equation (1) shows mathematically how this distance is modelled.

$$d(t) = d_0 + m(t) = d_0 + m_b \sin(2\pi f_b t) \quad (1)$$

where f_b and m_b are the breathing frequency and amplitude, respectively. The d_0 term represents the distance from the target to the radar. This paper only considers static targets, so in our work this term can be considered a constant. To extract respiratory information, UWB radio pulse signals are transmitted by a transmitter at regular intervals with a time interval in the order of milliseconds. These transmissions are

also called slow-time samples. For each slow-time transmission, multiple reflections are received by a receiver, which are sampled as complex in-phase and quadrature (IQ) values with a resolution in the order of nanoseconds. Because of this higher sampling rate, these are also called fast-time samples. A collection of fast-time samples that correspond to 1 slow-time index forms a CIR. A collection of slow-time indices can then be represented as a matrix with the notation shown in Equation (2).

$$R = r[m, n], 1 \leq m \leq N_s, 1 \leq n \leq N_f \quad (2)$$

where each row corresponds to a CIR, N_s is the number of slow-time samples and N_f is the number of fast-time samples per CIR. Another way to interpret these indices is to consider the slow-time axis as discrete time instants and the fast-time indices as distances from the radar to the reflecting surfaces. Each transmitted signal represents a time step. Within this step, captured fast-time samples correspond to reflections originating from increasingly further distances. Consequently, taking more fast-time samples at a single slow-time index provides reflection data from more distant surfaces. Reflections originate not only from the target but also from static objects in the room. Assuming these objects are static, the received signal can be expressed as Equation (3).

$$r(t, \tau) = \sum_{i=0} A_i p(\tau - \tau_i) + A_p p(\tau - \tau_d(t)) \quad (3)$$

where t and τ are slow- and fast-times indices, respectively. The first term sums the static background reflections (A_i), also called clutter. The second term represents reflections from the target (A_p), where the time-dependent delay $\tau_d(t)$ varies with breathing. Fig. 2a illustrates the amplitude of a single CIR. The first path (FP) corresponds to where the amplitude of the captured IQ values rises above the noise floor. In Fig. 2a, this occurs at index 0.

IV. MEASUREMENT CAMPAIGN AND DATASET DESCRIPTION

For evaluating the different algorithms, data was collected using the Qorvo DW3000, specifically the QM33120WDK1 development kit (DK), acting as a representative low-cost, COTS hardware device. The setup includes separate transmitter and receiver devices, each equipped with a Nordic nRF52840 DK and a DW3000 transceiver shield. The transmitter uses an omnidirectional antenna, while the receiver employs a directional antenna to focus on reflections coming from the desired direction. The system operates at a center frequency of 6.489 GHz with a bandwidth of 500 MHz. The slow-time sampling frequency is 77.5 Hz, and the receiver captures IQ samples at 1 ns intervals. Each fast-time sample represents 30 cm in distance, with each CIR containing 41 fast-time samples and the first peak always aligned at index 3. Samples before the first peak can be considered noise. The maximum distance the system can reach is equal to: $((41 - 4) * 30 \text{ cm}) / 2 = 555 \text{ cm}$. A Plux respiratory effort belt [25] serves as ground truth, measuring the accurate BR through belt stretch.

The dataset was collected in the Homelab in Zwijnaarde [26], a simulated home environment developed by Ghent University and imec for testing Internet of Things (IoT), smart home, and healthcare applications. The lab features various residential rooms, enabling diverse testing conditions. Fig. 1 illustrates the environments used in this project. To avoid saturation at the receiver side, the transmitter and receiver were placed 30 cm apart. In the first environment, targets lay in bed with the UWB hardware positioned 100 cm above their chest. In environments 2 and 3, participants sat on a chair at distances of 120 cm and 180 cm from the radar. Environment 4 involved standing individuals at varying distances (60, 100, 150, 200, and 400 cm) from the radar, with additional measurements at a 45-degree rotation angle for distances 60 cm and 150 cm. In the final environment, participants stood at the same distances without angle variations. For environments 4 and 5, both the transmitter and receiver were placed at a height of 156 cm. In total, 16 different people participated in the data campaign, resulting in a total of 240 different person-setup pairs². In environments 1, 2 and 3, measurements of 2 minutes were taken from each person-setup pair. In environments 4 and 5, only 1 minute CIR windows were captured. Consequently, approximately 288 min of data was obtained.

To ensure variability in BRs, subjects followed a randomized breathing pattern guided by an app, producing rates between 6 and 20 breaths per minute (BPM). This range was selected to align with established physiological benchmarks. Porter et al. define the nominal resting respiratory rate for adults as 12–20 BPM [27]. By extending the experimental range down to 6 BPM, the study incorporates the detection of bradypnea (abnormally slow breathing), a clinical indicator correlated with specific stress responses and sleep-disordered breathing. High-frequency respiratory rates exceeding 20 BPM (tachypnea) were excluded from this campaign, as such rates

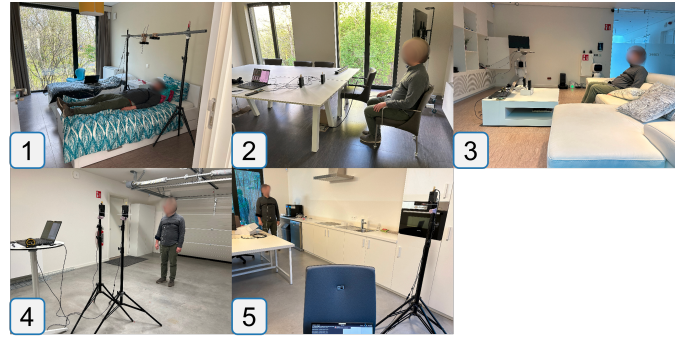


Fig. 1: Samples were collected across various environments with 16 participants. Environment 1 had subjects lying down. Environments 2 and 3 had them sitting. In the final 2 environments, subjects stood at multiple distances and angles from the radar. Defining a setup as a distinct combination of environment, distance, angle, and posture resulted in 15 different setups.

typically occur during physical exertion or acute clinical distress. Given that this work focuses on a non-intrusive baseline for resting-state and long-term trend monitoring, these higher frequencies were deemed outside the current scope.

V. METHODOLOGY

This section outlines the proposed processing pipeline, beginning with preprocessing steps and followed by different prediction algorithms. The prediction approaches are categorized into 2 types: rule-based peak-finding methods that serve as reference and the proposed model-based machine learning method.

A. Preprocessing

The preprocessing pipeline is shown in Figs. 2a to 2e. This pipeline processes the samples captured as IQ samples. For each slow-time sample, 41 fast-time samples are taken, forming a single CIR. Once collected, the amplitude of these samples is calculated. The resulting amplitude versus fast-time is shown in Fig. 2a. While phase information is available, it is discarded because phase stability across unsynchronized, low-cost UWB hardware is often unreliable. Furthermore, the high sensitivity of phase data makes it more vulnerable to environmental noise and multipath effects in complex settings. These CIRs are then combined into a matrix, of which a heatmap representation is shown in Fig. 2b. Each row in the matrix corresponds to a CIR (the “slow time”), and each column captures variations across time at a certain distance (the “fast time”). The maximum value of the slow-time index of the matrix is 120 seconds, representing all CIR samples collected from a single person in a specific environment and setup. To calculate the breathing rate over a shorter duration than the full experiment, a 30-second windowing operation is applied to the CIR matrix with a stride of 15 seconds. This window size was empirically determined to be an optimal balance between the dataset size required for model training and the accuracy of the Fast Fourier Transform (FFT) that

²Informed consent was obtained from all participants prior to the start of the measurement campaign.

will be applied later in the pipeline. Using an overlapping windowing approach results in more training samples, which is advantageous for training a neural network (NN). The 15-second stride is used only during training to augment the dataset. At deployment, the update rate is configurable by adjusting the stride between successive windows. Careful splitting between train and validation samples in our validation strategy will ensure no data leakage is introduced, which is further detailed in Section VI. An example of a window is shown in Fig. 2c. In this representation, it is clear that a periodic signal is present around range bin 8. In the next step, an FFT is applied to each column, generating a frequency spectrum for each distance bin. The obtained representation is shown on Fig. 2d, which is referred to as a range-frequency map. Peaks at the target breathing frequency are expected at the range index where a person is present. To focus on breathing rate-specific frequencies, signals below 0.09 Hz or above 0.5 Hz should be filtered out. This band-pass operation also suppresses low-frequency components arising from static objects and slow-varying factors such as temperature or background changes. When zooming in to this frequency range, peaks are indeed present in the example. Since the first peak index is present at index 3, the first 4 range bins can also be removed. Both the frequency filtering and range bin removal can be implemented with an image crop operation. The result of this operation is shown in Fig. 2e, showing which movement frequencies are present at each distance. The final step involves applying Min-Max normalization to scale all values between 0 and 1. This ensures consistent data scales across all CIR windows, which will benefit NN approaches. The min and max values are derived solely from the training data to prevent data leakage. The normalized CIR window then serves as input to the breathing detection algorithms, which in our case will be either rule-based techniques or neural networks.

B. Peak finding techniques

As noted in Section II, the most relevant low-cost or IEEE 802.15.4z-compliant studies are the papers by Numan *et al.* [23] and Farsaei *et al.* [16], both of which use rule-based approaches. To enable comparison, this first category of prediction methods includes 3 different rule-based frequency peak-finding techniques. The first one closely resembles the approach by Farsaei *et al.*, from which the others are 2 variations. All algorithms use the range-frequency maps as input, which can be mathematically represented as: $M = m_{ij}, 0 < i < n; 0 < j < p$. Here, m_{ij} is an element in the matrix, where i represents the row index (frequency bin) and j the column index (range bin). n is the total number of frequency bins and p is the number of fast-time indices. f_i will be defined as the frequency associated with row i .

(i) The first approach identifies the breathing frequency by locating the row index i_{max} of the maximum value in the matrix. Mathematically, this can be described as Equation (4). Once i_{max} is found, the corresponding frequency bin is: $f_{pred} = f_{i_{max}}$. From now on, this method is called the **highest peak algorithm**.

$$i_{max} = \arg \max_i (\max_j \{m_{ij}\}) \quad (4)$$

(ii) In the second method, the distance-frequency map is first accumulated in 1 row with the operation described in Equation (5). After this, the highest frequency index can be determined with: $i_{max} = \arg \max_i (a_i)$. This solution is the **accumulated highest peak algorithm**. The rationale is that scattering causes signals influenced by breathing movements to propagate to more distant surfaces. When these reflections are captured, the frequency component of the breathing signal is also present at these distances. By accumulating, it is hypothesized that the breathing frequency will become more pronounced.

$$a_i = \sum_{j=1}^p m_{ij}, j = 1, 2, \dots, p \quad (5)$$

(iii) The last algorithm employs the same accumulated row resulting from Equation (5). This time, the predicted frequency is determined by the weighted average operation described in Equation (6). Consequently, this solution is called the **accumulated weighted average algorithm**. Weighted averaging is also tested since this could alleviate the influence of large noise components.

$$f_{pred} = \sum_{i=1}^n \frac{a_i}{\sum_{k=1}^n a_k} f_i \quad (6)$$

C. Convolutional neural network based machine learning

Among existing studies, only Zhou *et al.* [22] proposes an attention-based CNN-LSTM solution evaluated on a similar amount of real-world considerations as our dataset. A more simple CNN was chosen in this paper due to its significantly lower processing requirements, aligning better with the low-cost aspect of this paper. The model architecture and optimized parameters are detailed in Table II. Various configurations of parameter sizes, kernel shapes, and layer counts were tested, with the best-performing design summarized in the table. A grid search was used to optimize normalization and regularization techniques, including batch normalization, dropout, and kernel regularization. The best-performing settings are also provided in Table II. To ensure the robustness of the results, each model was trained 10 times over. The model that scored the median validation loss was selected. All subsequent reported results follow this training strategy.

Although not previously applied to IR-UWB-based BR estimation, several resource-efficient model-based approaches exist, such as Random Forests and lightweight CNNs like MobileNet. To provide constrained-AI baselines, both methods were applied to our dataset. The Random Forest was implemented on flattened input features, while MobileNet used pretrained convolutional layers (ImageNet) with adapted fully connected layers for regression. To provide an additional CNN baseline without an explicit emphasis on deployability, the ResNet architecture was also evaluated with pretrained convolutional layers (ImageNet) and modified dense layers. Our proposed CNN (141 KB) is significantly more compact

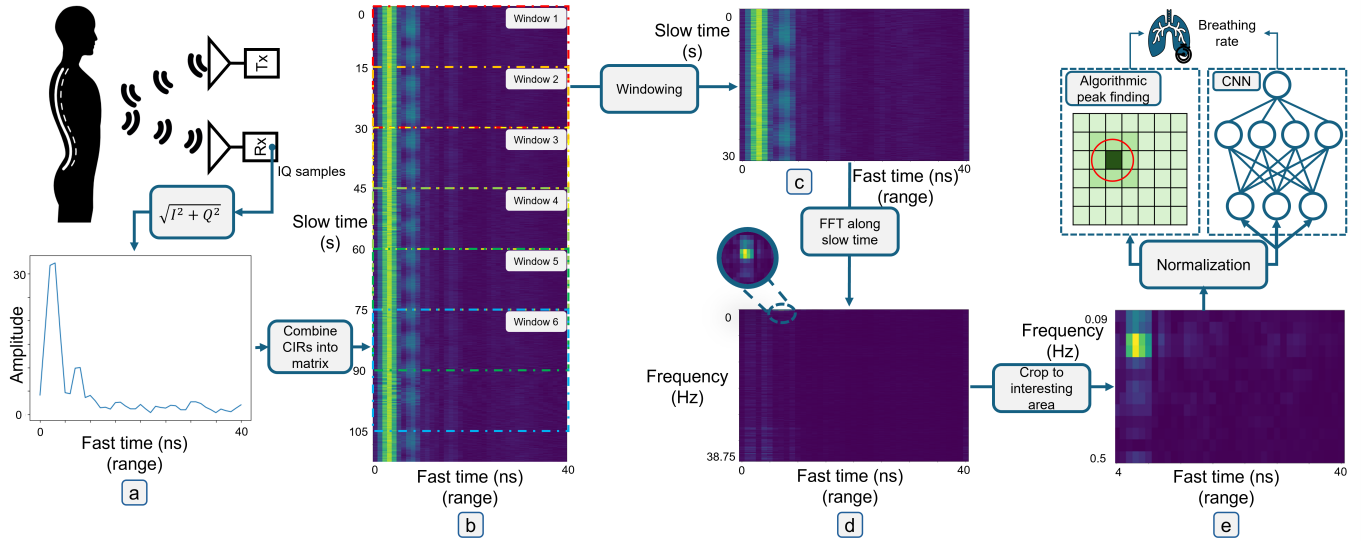


Fig. 2: Visualization of processing pipeline. IQ samples coming from the receiver will serve as input for the pipeline. Each sub-representation of the data is described in Section V and is indicated with a letter at the bottom of the subfigure.

Layer	Parameters
CIR_window (input)	size = (36, 13, 1)
conv2d	#filters = 8, kernel size = (10, 10), activation = relu, padding = same, kernel regularizer = L2
max_pooling2d	kernel size = (2, 2), strides = (2, 2)
conv2d_1	#filters = 16, kernel size = (8, 8), activation = relu, padding = same, kernel regularizer = L2
max_pooling2d_1	kernel size = (2, 2), strides = (2, 2)
conv2d_2	#filters = 32, kernel size = (4, 4), activation = relu, padding = same, kernel regularizer = L2
max_pooling2d_2	kernel size = (2, 2), strides = (2, 1)
conv2d_3	#filters = 64, kernel size = (2, 2), activation = relu, padding = same, kernel regularizer = L2
max_pooling2d_3	kernel size = (2, 2), strides = (2, 1)
dense	dim = 64, activation = relu, kernel regularizer = L2
dropout	dropout = 0.2
dense_1	dim = 16, activation = relu, kernel regularizer = L2
dropout_1	dropout = 0.2
dense_2 (output)	dim = 1

TABLE II: Architecture of the proposed CNN for breathing rate prediction.

than the MobileNet (23.7 MB) and ResNet (89.6 MB) implementations. Efficiency is achieved through a tailored 36×13 input and a shallow four-layer structure. Aggressive pooling and minimal dense layers eliminate memory bottlenecks while effectively capturing frequency components across range bins. This compact design reduces overfitting risk and ensures compatibility with resource-limited hardware. More complex architectures (e.g., LSTM or transformers) were not considered, as they are incompatible with the embedded, resource-limited focus of this work.

VI. RESULTS AND ANALYSIS

This section evaluates the proposed CNN and compares it to the representative rule-based solutions. This is done by applying the algorithms to the dataset described in Section IV, comprising samples from 16 participants across 15 setups. Each of these participant-setup combinations forms a unique pair. For evaluating the generalizability of the solutions, specific pairs are excluded from the training data and used as validation samples. As a pair is always fully included or excluded from the training set, no data leakage will occur.

Three validation strategies are used to evaluate the generalizability of the model architecture. The leave-one-person-out cross-validation experiment assesses how well the proposed CNN handles unseen individuals. The leave-one-situation-out cross-validation experiment evaluates its ability to adapt to new setups. Finally, the leave-one-person-situation-pair-out cross-validation strategy examines how the system performs with new combinations of individuals and setups. These strategies can also be interpreted as follows. If the system, trained on a certain set of person-situation pairs, is deployed to track an entirely new person, the leave-one-person-out strategy will estimate its performance. For deployment in a previously unseen setup, the leave-one-situation-out analysis indicates the expected performance. In the leave-one-person-situation-pair-out strategy, the scenario involves introducing a new setup, where measurements for some participants already in the dataset are used as calibration data. The prediction is that the calibration data will enhance the model's performance in the new setup.

A. Generalization to new persons

The first analysis evaluates the leave-one-person-out cross-validation strategy. In this approach, 1 participant is excluded from the training set in each fold and used for validation. This process is repeated for all 16 participants, producing 1

model per participant. The results are summarized in Table III, showing the mean L1 errors and their standard deviation. For the highest peak method, mean L1 errors range from 3.10 BPM to 5.58 BPM. The accumulated distance-frequency map approach reduces mean L1 errors slightly, ranging from 2.75 BPM to 5.32 BPM. The weighted average algorithm shows L1 errors from 3.29 BPM to 4.46 BPM. Among the rule-based methods, selecting the highest peak from the accumulated frequency bins proves most effective.

The table also includes results from other AI-based approaches previously mentioned in Section V. The results in the rightmost column of Table III show the performance of our proposed model. Overall, model-based solutions clearly outperform rule-based methods. Fig. 3 highlights conditions where the model-based solutions surpass traditional approaches. In ‘simple’ conditions with little noise (Fig. 3a), the breathing frequency peak is easily identified by most algorithms. In more challenging scenarios, however, noise components may dominate the breathing signal (Fig. 3b). In these cases, the model-based solutions perform better by not only analyzing the highest peak, but also taking the higher harmonic components into account. Among all investigated solutions, our proposed CNN achieves the best performance, with mean L1 errors between 0.97 BPM and 2.51 BPM, demonstrating a clear improvement over both rule-based methods and current AI solutions. This clearly demonstrates that the architecture proposed in this work is well-optimized for effective performance on IR-UWB data.

When compared to prior work, Farsaei et al. [16] and Numan et al. [23] report mean L1 errors of <1 BPM and 1.22 ± 0.63 BPM, respectively. In comparison, the rule-based approaches evaluated here show worse L1 scores. Our CNN comes closer to this error but still results in a higher L1 error. This difference can be attributed to the inclusion of significantly more challenging scenarios and greater variability in BRs in our study. A fairer comparison is with the results from Zhou et al. [22], who report a mean L1 error of 2.3 ± 1.9 BPM across various challenging scenarios. Our proposed CNN outperforms their results, with all validation folds in Table III achieving lower mean L1 scores.

Collecting datasets in healthcare settings is resource-intensive and requires sufficient participant diversity to ensure statistical validity and generalizability. To assess how performance scales with sample size, we employed an extended leave-k-people-out cross-validation with multiple permutations of training-testing splits. Results show that L1 error consistently decreases as the number of training participants increases, though with diminishing returns: expanding from 1 to 5 participants reduced error by 0.262 BPM, from 5 to 10 by 0.261 BPM, and from 10 to 16 by 0.154 BPM. This indicates that while additional data may enhance performance, the incremental gains are expected to diminish quickly as more individuals are added.

B. Generalization to new situations

The leave-one-setup-out cross-validation strategy evaluates generalizability to unseen setups. In each fold, 1 setup is

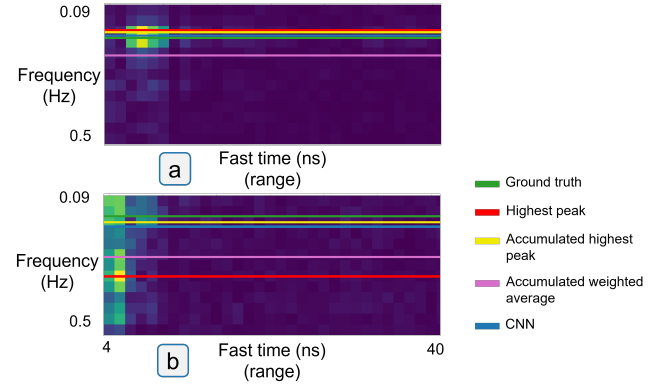


Fig. 3: The breathing rate estimates of the rule-based algorithms and the proposed CNN are visualized on a range-frequency map. In a simple scenario with a visible peak, shown in (a), the predictions from all methods closely align. However, in noisier scenarios where the peak is less distinct, as shown in (b), the CNN predictions are significantly closer to the ground truth.

entirely excluded for validation. Fig. 4 presents the results, where red box plots show the CNN performance, and the green, yellow, and purple plots represent the highest peak, accumulated highest peak, and accumulated weighted average algorithms. The box plots display L1 errors across validation samples, with the x-axis indicating the setup parameters in the validation set. The results confirm that the CNN generally outperforms rule-based methods. Only in setup 1 do the highest peak and accumulated highest peak algorithms slightly outperform the CNN. However, this difference is negligible.

Rule-based results in “simple” scenarios (setups 1 and 2) are comparable to prior work by Farsaei et al. [16] and Numan et al. [23]. However, in more challenging conditions not addressed by these prior studies, their performance degrades sharply, while the proposed CNN maintains reasonable accuracy. Validation folds for environments 1, 2, and 3 involved entirely unseen surroundings. The model’s strong performance in these cases demonstrates its robustness to environmental variation, a common challenge for many AI-based methods. Comparable evidence of generalizability is also absent in Zhou et al. [22].

Finally, we consider a situation whereby a situation has been observed, but a new person is monitored for which the CNN was not trained. This “leave-one-person-setup-pair-out” cross-validation strategy assesses performance when data from a specific person-setup pair is excluded. The results are represented by the blue box plots in Fig. 4. These box plots show that when training includes data from other individuals in the same setup, the CNN significantly improves its performance in fully unseen setups. With this strategy, only setup 10 has a Q3 above 2 BPM, indicating robust performance in nearly all scenarios.

C. Influence of situation parameters

Based on the results in Fig. 4, several conclusions can be made concerning the influence of setup parameters on CNN

Person	Rule based						AI based		Constrained AI based					
	Highest peak		Accumulated highest peak		Accumulated weighted average		ResNet50		Random forest		Mobilenet		CNN (Our solution)	
	mean (BPM)	std (BPM)	mean (BPM)	std (BPM)	mean (BPM)	std (BPM)	mean (BPM)	std (BPM)	mean (BPM)	std (BPM)	mean (BPM)	std (BPM)	mean (BPM)	std (BPM)
1	3.94	4.20	3.13	2.79	3.89	2.18	2.04	1.51	2.05	1.30	2.03	1.49	1.65	1.24
2	4.20	4.02	4.05	3.85	4.34	2.51	2.59	1.6	2.40	1.43	2.55	1.72	2.08	1.52
3	5.38	4.34	5.32	4.89	3.54	2.50	3.09	2.2	2.82	2.31	2.77	2.1	2.51	2.33
4	4.31	4.14	2.87	3.32	3.85	2.27	2.47	1.28	2.13	1.41	2.47	1.29	1.51	1.30
5	4.57	4.25	3.76	3.97	3.57	2.39	2.7	1.56	2.34	1.36	2.59	1.56	1.59	1.27
6	4.15	3.48	3.83	3.54	3.64	2.63	2.74	1.21	2.57	1.32	2.69	1.63	2.06	1.5
7	3.61	3.15	2.97	3.18	3.57	1.88	1.97	1.19	1.68	1.19	1.96	1.24	0.97	0.83
8	4.94	4.19	3.58	3.70	3.80	2.66	2.88	1.4	2.46	1.48	2.87	1.37	2.17	1.57
9	5.58	4.67	4.10	3.67	3.99	2.56	2.57	1.4	2.47	1.52	2.58	1.39	2.05	1.80
10	4.19	4.53	2.75	3.06	4.26	2.32	2.23	1.38	1.88	1.22	2.18	1.39	1.35	1.18
11	3.10	3.40	3.19	3.65	4.25	2.34	2.42	1.6	2.16	1.30	2.34	1.75	1.64	1.48
12	3.67	3.14	3.33	2.99	4.46	2.19	2.3	1.64	2.12	1.26	2.28	1.3	1.83	1.34
13	4.19	4.49	2.98	2.89	4.10	2.27	2.02	1.32	1.94	1.29	2.01	1.51	1.89	1.29
14	3.29	3.26	2.80	3.10	3.98	2.21	2.27	1.44	2.20	1.52	2.28	1.42	1.54	1.38
15	3.52	3.78	2.86	3.41	3.29	2.42	2.56	1.45	1.99	1.30	2.52	1.52	1.34	1.11
16	3.61	3.74	3.01	3.92	4.22	2.07	2.32	1.96	2.24	1.20	2.29	1.68	1.19	0.97
Average	4.14	3.92	3.41	3.50	3.92	2.34	2.45	1.51	2.22	1.40	2.40	1.52	1.71	1.38

TABLE III: Mean and standard deviation of the L1 error obtained when predicting the breathing rate of unseen persons. The CNN performs better than rule-based and other AI-based techniques, with mean accuracies between 0.97 and 2.51 BPM and lower standard deviations.

performance. In this analysis, the results from the leave-one-person-setup-pair-out cross-validation are considered. First, the posture of the target person significantly impacts performance. Analyzing individuals lying down is the easiest due to minimal random body movement (RBM). Sitting down introduces additional challenges due to increased reflection angles and slightly higher RBM. Standing is the most difficult posture, consistently yielding poorer results at the same distances. The effect of distance is best represented in the scenarios where the individuals were standing. The accuracy generally decreases as the distance between the target and radar increases from 100 to 400 cm due to the decreased strength of the reflected signal. However, the CNN still provides an acceptable mean accuracy of 1.29 BPM at a 4-meter distance. Interestingly, there is a significant performance drop at a distance of 60 cm. This results from the fixed radar height of 156 cm (as noted in Section IV) for environments 4 and 5. This height exceeds typical chest level, causing signals to arrive at steeper angles for shorter individuals, especially at close range. Additionally, at 60 cm, target reflections may overlap with first peak signals from the transmitter. For use cases where shorter ranges are required, the transmitter and receiver can be set up closer together. This would alleviate first peak overlap for shorter ranges but would require the transmission power to be lower to limit the saturation of the receiver. Finally, regarding the body angle, the CNN provides an acceptable mean error of 0.86 BPM even when the target person is positioned at a 45-degree angle towards the radar. This demonstrates that the CNN can handle such scenarios effectively.

D. Window Length Sensitivity Analysis

The window duration T determines the spectral resolution of the FFT preprocessing step ($\Delta f = 1/T$). While longer windows provide finer granularity, they increase temporal

smoothing and increase computational complexity. Evaluating T between 20 and 40 s showed that MAE generally decreases as spectral resolution improves (from 1.84 BPM at 20 s to 1.62 BPM at 30 s and 1.57 BPM at 40 s). Despite the modest accuracy gain at larger window sizes, larger range-frequency maps increase input dimensionality, memory usage (from 141 KB at 30 s to 174.3 KB at 40 s), and inference time, while also extending settling time. Consequently, window length constitutes a design trade-off between estimation accuracy and responsiveness. A 30-second window was selected as a balanced operating point, but the optimal choice remains application-dependent.

VII. DEPLOYING ON EMBEDDED HARDWARE

Finally, this section thoroughly evaluates the feasibility of deploying the proposed solution on constrained hardware. First, the radar's slow-time sampling frequency is analyzed. The second subsection will give an assessment of the preprocessing pipeline's computational complexity. In the following sections, the resource usage of the CNN-based solution is examined. The evaluation is conducted by deploying the model on the nRF52840 system-on-chip (SoC) to verify whether the model can run directly on the receiver device. In the final section, the power usage of the complete solution is described.

A. Required sampling rate

This section addresses the question of how often the UWB transmitter should transmit a packet. The maximum sampling rate of our system is empirically determined to be 77.5 Hz³. A higher sampling frequency requires sending more packets,

³The maximum sampling rate is determined by the speed at which the CIR can be read out over the serial bus. Currently, 41 fast-time IQ samples are collected per CIR. Increasing the number of collected fast-time IQ samples will increase the maximum radar detection range but will decrease the maximum sampling rate at which CIR values can be read out.

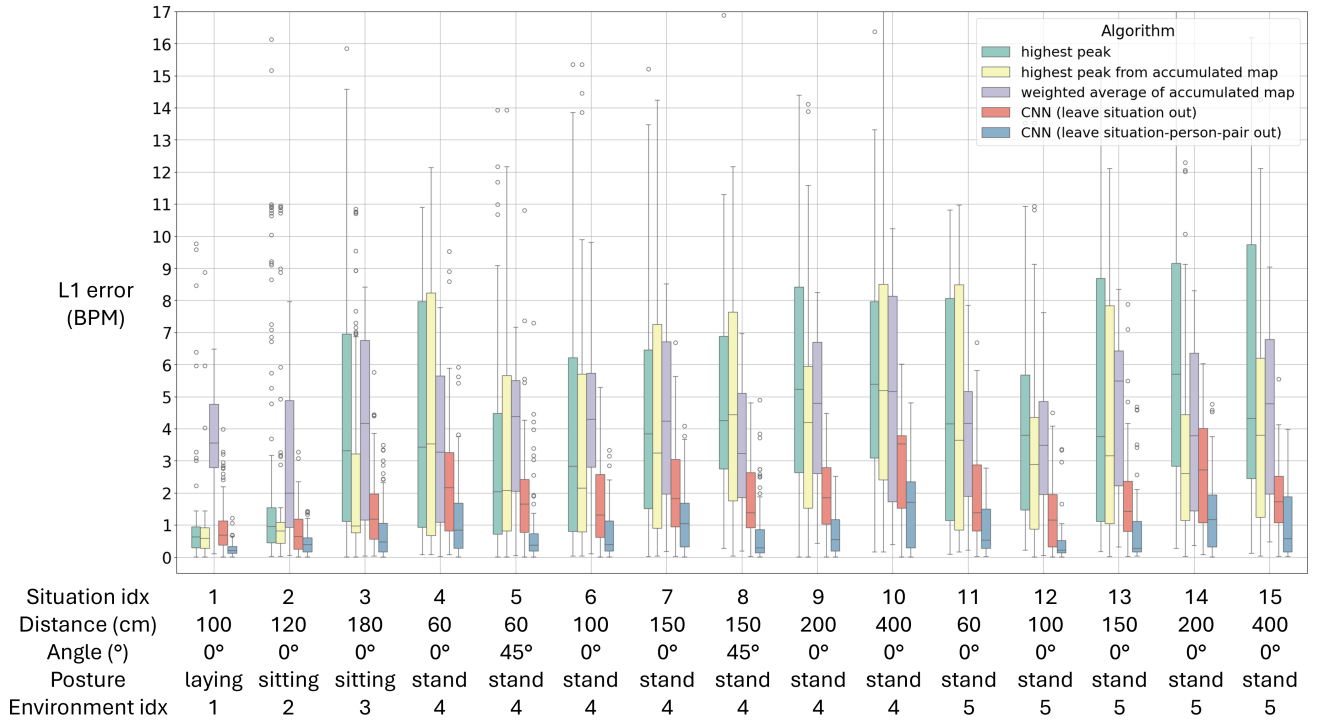


Fig. 4: Breathing rate prediction errors in different situations (environment, posture, distance and body angle). The green, yellow and purple plots are the results when applying the rule-based methods described in Section V-B. The red plots correspond to the L1 errors of the CNN for fully unseen situations. The blue box plots correspond to the L1 errors of the CNN for predicting the breathing rate for a new unseen person when other persons have already been observed in the specific situation (e.g. some form of pretraining on other individuals was possible). The CNN achieves a mean absolute error (MAE) of 1.73 BPM in unseen situations, significantly outperforming rule-based methods (3.40 BPM). When pretraining is possible, the MAE of the CNN is further reduced to 0.84 BPM.

which is not desirable for battery-powered embedded solutions. The red curve in Fig. 5 illustrates the leave-one-setup-out cross-validation results on subsampled datasets. Limiting the sampling rate to 20 Hz has minimal impact on L1 error, but reducing it further significantly increases error. Given that human BRs peak at 0.5 Hz, Nyquist’s theorem requires a minimal sampling rate of 1.0 Hz. The blue curve, representing the errors obtained when only validating on situation 1, shows that lower sampling rates have less impact in less complex conditions, with performance declining only at 4 Hz.

B. Algorithmic complexity of preprocessing pipeline

This section theoretically analyzes which parameters most profoundly impact the computation complexity of the preprocessing pipeline. First, the pipeline applies an FFT to each 30-second range bin. This FFT operation has a complexity of $O(n \log n)$, with n the array length [28]. Since an FFT is performed for each fast-time index, indicated by m , the total complexity becomes $O(mn \log n)$. Next, a cropping operation reduces the window size from (n, m) to (k, r) , where k is the number of frequency bins and r is the number of range bins after cropping. This slicing operation has a complexity of $O(kr)$. Finally, a normalization operation is applied by using an element-wise subtraction and division operation. Both these

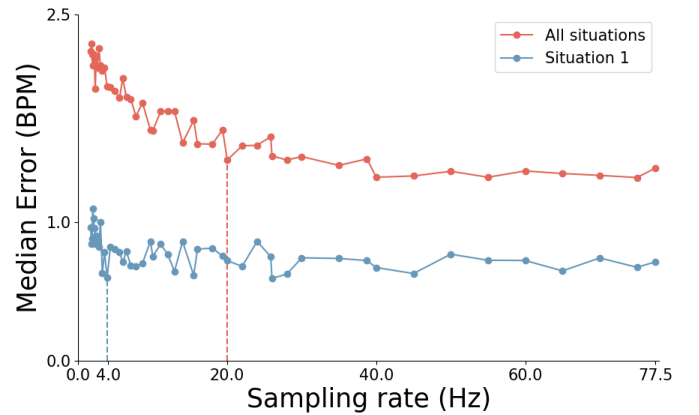


Fig. 5: Median L1 error for different sampling rates. Lowering the sampling rate negatively affects L1 error. For easy situations (situation 1), the sampling rate can be lowered to 4 Hz. When also considering more difficult situations, sampling rate should remain above 20 Hz.

operations have a complexity of $O(kr)$. All previous steps can be summarized as $O(mn \log n + 3kr)$. Since m and r are radio configuration-dependent constants, this simplifies to $O(n \log n + k)$. The dominant term, being $O(n \log n)$, indi-

cates an acceptable computational load suitable for constrained hardware.

C. Memory usage of the CNN model

One consideration with the CNN-based solution is its higher memory requirements compared to rule-based approaches. As the model size grows, the memory demand increases, limiting deployment on embedded devices. The proposed CNN architecture occupies 141 KB. This is already an acceptable model size as the nRF52840 SoC, with 256 KB static random access memory (SRAM) and 1 MB flash memory available, can run this model. Further memory reduction can be achieved through quantization. This technique lowers the precision of the model parameters, converting the standard 32-bit floating-point values to lower-bit formats. In this case, the model parameters will be converted to 8-bit unsigned integers. In addition to quantizing the model parameters, the input and output tensors are also quantized, resulting in the entire network operating in 8-bit integer format. After this operation, the model size is reduced by 67 % or to 46 KB, at the cost of only a 3.15 % increase in MAE due to the mismatch between the float32 to int8 transformation operation, making it suitable for severely memory-constrained devices.

D. Execution time of the system

Next, we experimentally measure the execution time of the different approaches. Using the rule-based techniques results in an inference time of 0.139 ms. The inference time of the CNN is higher at 523 ms. However, since 30-second windows are processed, the CNN's inference time is still significantly shorter than the considered time window. Reducing the inference time is still beneficial for embedded devices because this leaves more processing time for other tasks and allows the hardware to enter energy-saving states sooner. Quantization, as described earlier, not only reduces memory usage but also lowers inference time. When deploying the quantized model on the nRF52840 SoC, the inference time is reduced to 199 ms, which is a 62% reduction. The nRF52840 SoC includes a floating-point unit (FPU), which helps handle floating-point operations. On microcontrollers without this support, the benefits of the quantized model would be even more pronounced.

E. Energy consumption

1) Experimental measurement of the energy consumption:

This section analyzes the energy consumption of the proposed solution. Practical measurements were obtained by deploying the complete processing pipeline on the nRF52840 DK platform and monitoring current draw with a Keysight N6705B DC Power Analyzer at a supply voltage of 5 V (USB-powered). Reported currents represent averages, while power values are derived from precise area-under-the-curve calculations. These measurements were combined into a full-system energy model (Equation (7)), which accounts for platform (P_{plat}), transmitter (P_{TX}), and receiver (P_{RX}) consumption. For this analysis, preprocessing and model inference are

	V (V)	P (W)	T (s)
Practical			
Platform	5	0.00923	continuous
Preprocessing	5	0.01016	0.110
Model	5	0.02287	0.523
Quantized model	5	0.02146	0.199
$TX_{20.Hz}$	5	0.04474	continuous
$RX_{20.Hz}$	5	0.06373	continuous
$TX_{4.Hz}$	5	0.04357	continuous
$RX_{4.Hz}$	5	0.05468	continuous
Theoretical			
$TX_{20.Hz}$	3	0.000838	continuous
$RX_{20.Hz}$	3	0.001252	continuous
$TX_{4.Hz}$	3	0.000168	continuous
$RX_{4.Hz}$	3	0.000251	continuous

TABLE IV: Voltage and power consumption of components in the processing pipeline. Preprocessing and model stages consume power intermittently, while other modules draw continuous power. Practical results are based on measured energy usage from deployed components, whereas theoretical results represent lower bounds after applying the optimizations described in Section VII-E.

assumed to occur once every 30 s, so they are included as $E_{preproc}$ and E_{model} , respectively. Equation (7) is structured so that readers can recompute the energy budget for any other update rate by replacing the 30-second interval in the $\lfloor t/30 \rfloor$ term with the desired stride. As transmitter and receiver are separate devices, P_{plat} is counted twice, with all processing executed on the receiver side.

Per-component measurements are summarized in Table IV. The nRF52840 DK consumes 1.84 mA at 5 V. Preprocessing is executed once every 30 s, drawing 2.03 mA for 0.110 s. CNN inference requires 4.57 mA for 0.523 s, which is reduced to 4.29 mA and 0.199 s after quantization, yielding substantial savings. At a sampling rate of 20 Hz, which is shown in Section VII-A to provide robust operation under diverse conditions, the total energy consumption per 30 s is $E_{30sec} = 3.813 J$, corresponding to an average power of 0.127 W. To put this in intuitive terms, a typical AAA cell stores roughly 5400 J, so the system consumes one AAA's worth of energy every 11.8 h. When the sampling rate is reduced to 4 Hz for simpler scenarios, transmitter and receiver consumption decreases, lowering E_{30sec} to 3.507 J (average power 0.117 W). Preprocessing at this rate showed no measurable difference. Under this configuration, the system consumes one AAA's worth of energy every 12.8 h.

$$E_{system} = \underbrace{(P_{plat} + P_{TX}) \cdot t}_{\text{transmitter}} + \underbrace{(P_{plat} + P_{RX}) \cdot t + (E_{preproc} + E_{model}) \cdot \lfloor \frac{t}{30} \rfloor}_{\text{receiver}} \quad (7)$$

2) Theoretical energy consumption model: In the previous calculations, the degrees of freedom are limited by the development boards used. For practical deployments, energy consumption can be significantly reduced.

(i) A first optimization is to deploy the transmitter and receiver on 1 platform. By doing this, only 1 P_{plat} term

remains in the energy calculation. To this end, different antenna designs should minimize self-interference between the transmitting and receiving antenna [29].

(ii) Secondly, note that the transmitter and receiver make up 84 % of the total power consumption. Part of this can be attributed to the fact that a DK is used that is not fully optimized for energy consumption. Using the information stated in the datasheet of the DW3000 transceivers [30], a theoretical minimal current profile can be calculated for both the transmitter and receiver. The current consumption of the transceiver will vary in time depending on the state in which the transceiver is. Equation (8) shows the calculation of the theoretical minimal energy consumption when the transceiver is in transmitter mode. Equation (9) calculates the used energy for the receiver. In these formulas, E_{PLL} represents the energy consumed by the phase locked loop (PLL) phase, E_{TX_SHR} and E_{RX_SHR} denote the energy used while transmitting and receiving the synchronization header (SHR), and $E_{TX_PHR_PSDU}$ and $E_{RX_PHR_PSDU}$ correspond to the energy consumed during the transmission and reception of the physical layer header (PHR) and physical layer service data unit (PSDU). Additionally, E_{PR_HUNT} is the energy used during the preamble hunt phase, which involves the receiver searching for and detecting the preamble signal to synchronize with the incoming data transmission. E_{RC} is the energy consumed in the receiver chain during the IDLE_RC phase, and E_{SLEEP} represents the energy used when the transceiver is in a low-power sleep state.

$$E_{TX_min} = E_{PLL} + E_{TX_SHR} + E_{TX_PHR_PSDU} + E_{SLEEP} \quad (8)$$

$$E_{RX_min} = E_{PLL} + E_{PR_HUNT} + E_{RX_SHR} + E_{RX_PHR_PSDU} + E_{PLL} + E_{RC} + E_{SLEEP} \quad (9)$$

(iii) Thirdly, another major improvement can be made for the receivers' E_{PR_HUNT} term. The DK receiver automatically reverts to the preamble hunt phase after receiving a packet. In the context of this solution, the timings between the transmitter and receiver are precisely defined. Also, as stated in optimization 1, both transmitter and receiver are deployed on a single platform, making it possible to share the same clock. This enables us to minimize this preamble hunt phase, leading to a negligible phase hunt duration.

(iv) A final optimization concerns IDLE current usage of the transceivers. The development kit hardware lacks deep sleep capability and instead defaults to the PLL state, which consumes significantly more power. By implementing this functionality in the deployment version of the system, power consumption will be significantly reduced.

Theoretical durations and current usages are presented in Table V. Note that the transceiver operates at 3V, so all energy calculations were performed with this voltage in mind. The transmitter starts in the PLL state, using 18 mA for 0.020 ms. It then transmits the SHR, which is 256 symbols long, taking 0.270 ms and consuming 48 mA. Next, the PHR and PSDU

State	State name	T (ms)	I (mA)	E (mJ)
Transmitter				
1	PLL lock	0.020	18	0.00108
2	TX SHR	0.270	48	0.03888
3	TX PHR/PSDU	0.016	40	0.0019
4 (4Hz)	SLEEP	249.694	0.00026	0.00019
4 (20Hz)	SLEEP	49.694	0.00026	0.000039
Receiver				
1	PLL lock	0.020	18	0.00108
2	PR_HUNT	0.000	70	0
3	RX SHR	0.24	78	0.05616
4	RX PHR/PSDU	0.016	70	0.00336
5	PLL lock	0.02	18	0.00108
6	IDLE RC	0.036	8	0.000864
7 (4 Hz)	SLEEP	249.668	0.00026	0.00019
7 (20 Hz)	SLEEP	49.668	0.00026	0.000039

TABLE V: Theoretical duration and current usage of transmitter and receiver in the different states of the energy model.

are sent. For CIR collection, empty packets can be sent, but the PSDU will still contain the medium access control (MAC) header and reed solomon (RS) bits, taking 0.016 ms and consuming 40 mA. The transmitter then transitions into a deep sleep state for the rest of the 250 or 50-millisecond period, based on the configured sampling rate. In this case, 4 Hz and 20 Hz are shown. This state is very efficient, with a consumption of only 0.00026 mA. The receiver also starts in the PLL state and then skips the HUNT state. After this, the receiving device sequentially acquires the SHR, PHR, and PSDU, ending with another PLL, IDLE RC, and deep sleep state. When considering the system with a sampling rate of 20 Hz, 1 propagation sums to 0.0419 mJ for the transmitter and 0.0626 mJ for the receiver. Multiplying this by 20 gives the power consumption per second as stated in Table IV. These new values give a new power consumption per 30 seconds: $E_{30_sec_min} = 0.345 J$. Consequently, on average 0.0115 W is used, meaning the optimized system consumes one AAA's worth of energy (5400 J) every 130.5 h. A more realistic deployment option is a common battery pack of 20 000 mAh. This configuration could enable the system to work for 268.2 days without recharging. In less complex situations, where a sampling rate of 4 Hz is sufficient, P_{TX_min} and P_{RX_min} decrease even further. Consequently, $E_{30_sec_min}$ becomes 0.295 J, enabling continuous monitoring for approximately 313.8 days. This allows for BR monitoring in locations without power outlets, such as movable beds, furniture, and desks.

VIII. CONCLUSION

This research developed and validated a non-intrusive system for accurate BR monitoring with potential applications in long-term remote patient monitoring, and presence detection in smart homes. A CNN model, based on an architecture specifically adapted for IR-UWB CIR data and trained on a new open-source dataset, outperformed existing methods. Robustness was confirmed across distances up to 4 m (average L1 error 2.41 BPM, reduced to 1.29 BPM with calibration), body angles (0.86 BPM at 45°), unseen environments, and a variety of postures.

Complexity analysis verified the preprocessing pipeline is lightweight, and the CNN was effectively quantized with minimal accuracy loss, enabling deployment on a broader range of platforms. A novel analytical energy model was developed, showing that at 20 Hz the system consumes one AAA's worth of energy every 11.8 h, extending to every 130.5 h with theoretical optimizations. With a common 20 000 mAh battery pack, the optimized system operates for 268.2 days at 20 Hz, extending to 313.8 days at 4 Hz. Results also show that sampling at 4 Hz is sufficient in less complex scenarios, whereas a higher rate of 20 Hz is recommended for analyzing more complex cases.

Future research can address multi-person detection, non-static subjects, and the implementation of motion gating techniques to identify and filter out unreliable data during high-energy body movements. Exploring foundation models pre-trained on large-scale radar datasets could also enhance the system's ability to distinguish subtle respiratory patterns from complex body movements in non-static scenarios. Additionally, expanding the experimental scope to include higher respiratory rates and clinical populations would be essential for expanding the use case to acute respiratory distress scenarios. Finally, explainable AI represents a highly promising direction, particularly in the context of healthcare applications.

REFERENCES

- [1] Eurostat. (2024) Respiratory diseases statistics. [Online]. Available: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Respiratory_diseases_statistics
- [2] J. Bousquet, C. Tanasescu, T. Camuzat, J. Anto, F. Blasi, A. Neou, S. Palkonen, N. Papadopoulos, J. Antunes, B. Samolinski, P. Yiallouris, and T. Zuberbier, "Impact of early diagnosis and control of chronic respiratory diseases on active and healthy ageing: A debate at the European Union Parliament," *Allergy*, vol. 68, no. 5, pp. 555–561, May 2013.
- [3] K. Kostikas, D. Price, F. S. Gutzwiller, B. Jones, E. Loeffroth, A. Clemens, R. Fogel, R. Jones, and H. Cao, "Clinical Impact and Healthcare Resource Utilization Associated with Early versus Late COPD Diagnosis in Patients from UK CPRD Database," *International Journal of Chronic Obstructive Pulmonary Disease*, vol. 15, pp. 1729–1738, 2020.
- [4] S. A. Kayser, R. Williamson, G. Siefert, D. Roberts, and A. Murray, "Respiratory rate monitoring and early detection of deterioration practices," *British Journal of Nursing*, vol. 32, no. 13, pp. 620–627, Jul. 2023.
- [5] S. Choshen-Hillel, A. Ishqer, F. Mahameed, J. Reiter, D. Gozal, A. Gileles-Hillel, and I. Berger, "Acute and chronic sleep deprivation in residents: Cognition and stress biomarkers," *Medical Education*, vol. 55, no. 2, pp. 174–184, Feb. 2021. [Online]. Available: <https://asmeipublications.onlinelibrary.wiley.com/doi/10.1111/medu.14296>
- [6] M. Cheraghinia, A. Shahid, S. Luchie, G.-J. Gordebeke, O. Caytan, J. Fontaine, B. V. Herbruggen, S. Lemey, and E. D. Poorter, "A Comprehensive Overview on UWB Radar: Applications, Standards, Signal Processing Techniques, Datasets, Radio Chips, Trends and Future Research Directions," *IEEE Communications Surveys & Tutorials*, vol. 27, no. 4, pp. 2283–2324, Aug. 2025. [Online]. Available: <https://ieeexplore.ieee.org/document/10738385/>
- [7] B. Van Herbruggen, S. Luchie, J. Jockué, R. Wilsens, S. Lemey, and E. De Poorter, "Wip paper: Uwb-based integrated sensing and communication (isac) for robotic applications," in *EWSN2024, the International Conference on Embedded Wireless Systems and Networks*, 2024.
- [8] D. Coppens, A. Shahid, and E. De Poorter, "UWB TDoA Error Correction Using Transformers: Patching and Positional Encoding Strategies," *IEEE Transactions on Wireless Communications*, vol. 25, pp. 7000–7013, 2026. [Online]. Available: <https://ieeexplore.ieee.org/document/11237034/>
- [9] Q. Li, J. Liu, R. Gravina, W. Zang, Y. Li, and G. Fortino, "A uwb-radar-based adaptive method for in-home monitoring of elderly," *IEEE Internet of Things Journal*, vol. 11, no. 4, pp. 6241–6252, 2024.
- [10] W. Mu, J. Zhang, X. Jiang, K. Wang, L. Li, and L. Zhang, "Ap-hr: Amplitude and phase joint optimization-based heartbeat and respiration separation algorithm using ir-uwb radar," in *2024 IEEE International Workshop on Radio Frequency and Antenna Technologies (iWRAF&AT)*, 2024, pp. 265–270.
- [11] S. K. Leem, F. Khan, and S. H. Cho, "Vital sign monitoring and mobile phone usage detection using ir-uwb radar for intended use in car crash prevention," *Sensors*, vol. 17, no. 6, 2017.
- [12] Z. Duan and J. Liang, "Non-contact detection of vital signs using a uwb radar sensor," *IEEE Access*, vol. 7, pp. 36 888–36 895, 2019.
- [13] Y. Zhang, X. Li, R. Qi, Z. Qi, and H. Zhu, "Harmonic multiple loop detection (hmlD) algorithm for not-contact vital sign monitoring based on ultra-wideband (uwb) radar," *IEEE Access*, vol. 8, pp. 38 786–38 793, 2020.
- [14] M. Husaini, L. M. Kamarudin, A. Zakaria, I. K. Kamarudin, and M. A. Ibrahim, "Non-contact vital sign monitoring during sleep through uwb radar," in *2022 10th International Conference on Information and Communication Technology (ICICT)*, 2022, pp. 157–162.
- [15] A. Lopes, D. F. N. Osório, H. Silva, and H. Gamboa, "Equivalent pipeline processing for ir-uwb and fmcw radar comparison in vital signs monitoring applications," *IEEE Sensors Journal*, vol. 22, no. 12, pp. 12 028–12 035, 2022.
- [16] A. Farsaei, B. Meyer, A. Sheikh, M. El Soussi, P. Zhang, G. K. Ramachandra, J. Govers, and M. Hijdra, "An iee 802.15.4z-compliant ir-uwb radar system for in-cabin monitoring," in *2023 IEEE 34th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2023, pp. 1–5.
- [17] D. Wang, S. Yoo, and S. H. Cho, "Experimental comparison of ir-uwb radar and fmcw radar for vital signs," *Sensors*, vol. 20, no. 22, 2020.
- [18] P. Wang, X. Ma, R. Zheng, L. Chen, X. Zhang, D. Zeghlache, and D. Zhang, "Slprof: Improving the temporal coverage and robustness of rf-based vital sign monitoring during sleep," *IEEE Transactions on Mobile Computing*, vol. 23, no. 7, pp. 7848–7864, 2024.
- [19] J. Kim and S. Kim, "Proposed signal processing method for continuous respiration monitoring using uwb radar," in *2024 IEEE International Conference on Consumer Electronics (ICCE)*, 2024, pp. 1–4.
- [20] X. Dang, J. Zhang, and Z. Hao, "A non-contact detection method for multi-person vital signs based on ir-uwb radar," *Sensors*, vol. 22, no. 16, 2022.
- [21] F. Khan and S. H. Cho, "A detailed algorithm for vital sign monitoring of a stationary/non-stationary human through ir-uwb radar," *Sensors*, vol. 17, no. 2, 2017.
- [22] J. Zhou, "Mvsm: Motional vital signs monitoring with ir-uwb radar and attention-based cnn-lstm network," in *2023 8th International Conference on Intelligent Computing and Signal Processing (ICSP)*, 2023, pp. 679–683.
- [23] P. E. Numan, H. Park, J. Lee, and S. Kim, "Machine learning-based joint vital signs and occupancy detection with ir-uwb sensor," *IEEE Sensors Journal*, vol. 23, no. 7, pp. 7475–7482, 2023.
- [24] S. Venkatesh, C. Anderson, N. Rivera, and R. Buehrer, "Implementation and analysis of respiration-rate estimation using impulse-based uwb," in *MILCOM 2005 - 2005 IEEE Military Communications Conference*, 2005, pp. 3314–3320 Vol. 5.
- [25] Plux, "Respiration (PZT) sensor datasheet," 2020. [Online]. Available: <https://support.pluxbiosignals.com/wp-content/uploads/2021/11/Respiration.PZT.Datasheet.pdf>
- [26] iDlab. (2024) iDlab homelab. [Online]. Available: <https://homelab.ilabt.imec.be/>
- [27] R. Porter and D. D. Graham, "Abnormal Respirations," in *StatPearls*. Treasure Island (FL): StatPearls Publishing, 2025. [Online]. Available: <http://www.ncbi.nlm.nih.gov/books/NBK470309/>
- [28] E. Rajaby and S. M. Sayedi, "A structured review of sparse fast fourier transform algorithms," *Digital Signal Processing*, vol. 123, p. 103403, 2022. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1051200422000203>
- [29] G.-J. Gordebeke, S. Lemey, O. Caytan, M. Boes, J. Jockué, S. Van De Velde, C. Marshall, E. De Poorter, and H. Rogier, "Time-Domain-Optimized Antenna Array for High-Precision IR-UWB Localization in Harsh Urban Shipping Environments," *IEEE Sensors Journal*, vol. 24, no. 5, pp. 5561–5577, Mar. 2024.
- [30] Qorvo, "DW3000 Datasheet," 2020. [Online]. Available: <https://www.qorvo.com/products/d/da008142>