# MobiVital: Self-supervised Quality Estimation for UWB-based Contactless Respiration Monitoring


Ziqi Wang*
Qualcomm
wangzq312@g.ucla.edu

Derek Hua
UCLA
derekhua@ucla.edu

Wenjun Jiang
Samsung Research America
wenjun.jiang@samsung.com

Tianwei Xing†
Meta
twxing@ucla.edu

Xun Chen†
Independent Researcher
xunchen@outlook.com

Mani Srivastava‡
UCLA and Amazon
mbs@ucla.edu



## Abstract

Respiration waveforms are increasingly recognized as important biomarkers, offering insights beyond simple respiration rates, such as detecting breathing irregularities for disease diagnosis. Previous works in wireless respiration monitoring have primarily focused on estimating respiration rate, where the breath waveforms are often generated as a by-product. As a result, issues such as waveform deformation and phase inversion have largely been overlooked, reducing the signal's utility for applications requiring breathing waveforms. To address this problem, we present a novel approach, MobiVital, that improves the quality of respiration waveforms obtained from ultra-wideband (UWB) radar data. MobiVital combines a self-supervised autoregressive model for breathing waveform extraction with a biology-informed algorithm to detect and correct waveform inversions. To encourage reproducible research efforts for developing wireless vital signal monitoring systems, we also release a 12-person, 24-hour UWB radar vital signal dataset, with time-synchronized ground truth obtained from wearable sensors. Our results show that the respiration waveforms produced by our system exhibit a 7-34% increase in fidelity to the ground truth compared to the baselines and can benefit downstream tasks such as respiration rate estimation.



## CCS Concepts

• **Human-centered computing** → *Ubiquitous and mobile computing*; • **Computing methodologies** → *Machine learning*.

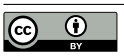
**ACM Reference Format:**

Ziqi Wang, Derek Hua, Wenjun Jiang, Tianwei Xing, Xun Chen, and Mani Srivastava. 2025. MobiVital: Self-supervised Quality Estimation for UWB-based Contactless Respiration Monitoring. In *The 3rd International Workshop on Human-Centered Sensing, Modeling, and Intelligent Systems (HumanSys '25), May 6–9, 2025, Irvine, CA, USA.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3722570.3726878


*The research reported in this paper was conducted at UCLA and Samsung Research America, and is unrelated to this author's current affiliation, Qualcomm.

†The research was conducted when these authors were affiliated with Samsung Research America and is unrelated to this author's current affiliations.

‡The author holds concurrent appointments as an Amazon Scholar and Professor at UCLA, but the work in this paper is not associated with Amazon.





*HumanSys '25, Irvine, CA, USA*



ACM ISBN 979-8-4007-1609-9/2025/05
https://doi.org/10.1145/3722570.3726878


## 1 Introduction

**Motivation.** Respiration is a vital biomarker for assessing health, diagnosing diseases, and monitoring physical exertion. Beyond respiration rate, respiration waveform captures chest movement patterns, offering deeper insights into breath depth, rhythm, and variability [3]. Irregularities in the waveform can signal conditions like sleep apnea, COPD, or Parkinson's disease. Additionally, respiration waveform analysis aids athletic training, diet monitoring, and rehabilitation by ensuring adherence to prescribed breathing exercises [6]. Wearable chestbands provide accurate breath waveform readings by directly measuring chest expansion, but they can cause discomfort and restrict movement. Contactless methods offer a non-invasive alternative leveraging advancements in wireless sensing technologies such as acoustics, Wi-Fi, mmWave, RFID, and IR-UWB radar. In this work, we focus on IR-UWB radar for its low power consumption, high resolution (detecting subtle vibrations), and ranging capability (distinguishing multiple subjects by distance).

**Current Solution.** An IR-UWB respiration sensing system follows three steps: (1) Preprocessing. Complex-valued UWB data are converted to real values using magnitude and phase calculations [13], with background removal filters applied to remove clutters from static objects. (2) Waveform Extraction. UWB radar produces two-dimensional radar matrices, one axis corresponds to object range (distance) and the other corresponds to time. The system needs to detect the distance of the human subject and slice the matrix on the range axis to extract a waveform as the respiration measurement. For target ranging, existing works explore algorithms such as signal power variance [11], autocorrelation [7], and constant false alarm rate (CFAR) detection [13]. (3) Application-specific processing. Filtering [2], mode decomposition algorithms [13], or wavelet analysis [1] are combined with spectrum analysis to extract respiration rates. Literature also reported the usage of end-to-end deep-learning models in this stage [10].

**Problem.** With a focus on estimating the respiration rate, existing works often produce respiration waveform as a by-product. Despite the importance of respiration waveform monitoring, breath waveform morphology issues, i.e., *distortion* and *inversion*, are often overlooked, as they have a limited impact on respiration rate estimation. In Figure 1, we showcase these common signal quality issues in the breath waveform. We attribute these issues to suboptimal waveform extraction: human respiration typically affects a range of more than 50 cm in the UWB data matrix, which produces many possible candidate time series. The time series selected based on the target distance estimation is often suboptimal. To demonstrate

the pervasiveness of this issue, we conduct a simple experiment by selecting the UWB time series with the largest standard deviation. Among the 1874 selected time series, 541 (28.8%) suffer from distortion (correlation with ground truth < 0.5), and 521 (27.8%) are inverted (correlation with ground truth < -0.5). Through dataset collection and analysis, we discovered that a better candidate time series is often available from the same UWB data matrix. If ground truth were available, selecting the best, non-inverted waveform would be straightforward. Leveraging the ground truth, inverted sequences can be fully excluded, and the ratio of low-quality signals being selected drops to 2.4% (24/1874). However, the ground truth respiration waveform is apparently not available during the deployment of such a system. Thus, the challenge lies in *predicting signal quality and detecting inversion without ground truth.*

**Proposed Method.** We introduce MobiVital, which includes: (1) a signal quality predictor to select the best time series and reduce distortion, and (2) a time series inversion detector, both operates without ground truth. **For signal quality estimation**, we propose a self-supervised procedure inspired by a pioneer work [9]. This procedure leverages the generalization limits of machine learning models, which are known to perform well on in-distribution data but poorly on out-of-distribution data. Thus, we train an autoregressive model (using a chuck of the time series to predict its near future) with high-quality respiration data. During deployment, candidate time series are fed into this model: high-quality signals should yield accurate autoregressive predictions, while poor-quality signals generate a high prediction loss. This approach provides a self-supervised quality measure without requiring labels. **For inversion**, we design a lightweight algorithm based on respiration biomechanics [5]. The intuition is that inhalation requires active muscle effort, leading to a chest cavity expansion duration typically below 50% in a breathing cycle. Our pipeline detects inversion by analyzing chest movement patterns from the UWB signal.

We comprehensively evaluate our proposed method, demonstrating that the autoregressive model effectively predicts signal quality. The selected time series shows higher correlation with ground truth compared to baselines and improves respiration rate estimation accuracy. An ablation study further validates the effectiveness of our phase inversion detector.

Developing wireless vital signal monitoring systems often requires complex signal processing pipelines or deep neural networks, necessitating high-quality datasets. However, such datasets are rarely open-sourced. To promote reproducibility, we release a 24-hour IRB-approved dataset[1] from 12 subjects, featuring IR-UWB radar-based vital signal sensing with synchronized wearable sensor ground truth. The code repository of this work is also publicly available[2]. Limited by the length of this paper, we have to omit some theoretical analysis, implementation details, and additional results. We place them in this technical report [8] for interested readers.

The major contributions of this paper are summarized below:

- A 12-person, 24-hour open-source dataset for respiration monitoring using UWB radar.
- A self-supervised autoregressive model to predict radar-measured respiration quality with no knowledge of the ground truth.

[1]Link to the full MobiVital dataset: https://zenodo.org/records/15022885.
[2]Link to the code repository of this paper: https://github.com/nesl/mobivital-public.

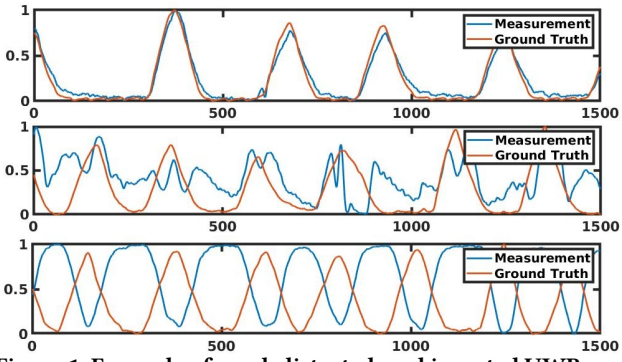



**Figure 1: Example of good, distorted, and inverted UWB measurements.**

- A bio-informed lightweight algorithm to detect if a respiration signal is inverted.

## 2 MobiVital Dataset

We first introduce the MobiVital dataset, detailing the data collection platform, protocol, and key statistics.

In our experiments, we configure the SLMX4 UWB radar to send probing pulse at 50 Hz, covering a range of 0.3–6.3 meters with a spatial resolution of 5 cm, dividing the space into 120 bins. The resulting radar data is $D_{uwb} \in \mathbb{C}^{120\times 50t}$, where $t$ is time in seconds. The radar connects to a Raspberry Pi via SPI, which also controls an inertial measurement unit (IMU) via I2C. The IMU samples at 100Hz, generating $D_{imu} \in \mathbb{R}^{6\times 100t}$, where three axes capture linear acceleration and three record rotational movement. The Raspberry Pi, radar, and IMU are mounted on a cheeseplate (Figure 3), which can be tripod-mounted or handheld.

For ground truth respiration waveforms, we use the NeuLog NUL-236 respiration belt sensor, widely adopted in related works. The belt detects chest movement via air pressure changes converted into electronic signals. Additionally, our dataset includes heart rate and blood pulse waveforms from the NeuLog NUL-208, a plethysmograph-based finger clip sensor. Both NeuLog sensors

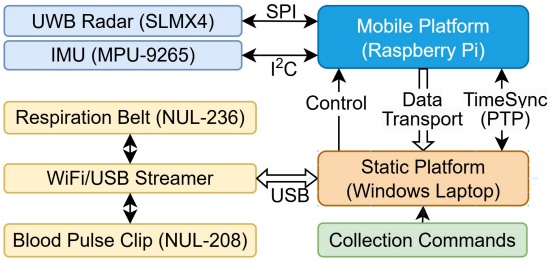



**Figure 2: MobiVital data collection platform.**

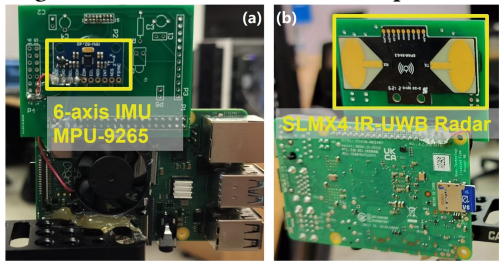



**Figure 3: Mobile platform hardware of MobiVital: (a) backside. (b) front-side.**

| Volunteer | A | B | C | D | E | F | G | H | I | J | K | L |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Duration/mins (Tripod) | 112.5 | 79.5 | 107.0 | 104.0 | 64.0 | 55.0 | 68.0 | 71.5 | 76.0 | 64.0 | 76.0 | 59.5 |
| Duration/mins (Handheld) | 94.5 | 27.0 | 107.0 | 98.5 | 32.0 | 28.0 | 32.0 | 27.0 | 32.0 | 36.0 | 32.0 | - |

**Table 1: Dataset Statistics.**


**Figure 4: Experiment scenarios: (a) tripod, (b) handheld.**

operate at 50Hz, producing $D_{heart} \in \mathbb{R}^{1\times 50t}$ and $D_{breath} \in \mathbb{R}^{1\times 50t}$. These sensors connect to a NeuLog USB streaming unit, controlled via an HTTP API on a Windows laptop.

Time synchronization between the Windows laptop and Raspberry Pi is achieved using Precision Time Protocol (PTP), with the Raspberry Pi as the server (Figure 2). During data collection, control scripts on the laptop start all sensors simultaneously. Each session lasts approximately five minutes, divided into 30-second sub-sessions due to buffer constraints. All data is timestamped, transferred to the laptop, resampled to 50Hz using a low-pass anti-aliasing filter followed by decimation, and aligned based on timestamp proximity, using the UWB radar timestamps as the reference.

Now, we provide details about our data collection protocol. Our study is IRB-approved (UCLA IRB#23-000754). We recruited 12 participants for our study. During data collection, participants were instructed to sit on a chair 1.5 meters away from the sensor and breathe normally. They were advised to relax, avoid controlling their breath, and refrain from body movements for about 20 minutes while data was collected. Participants could watch videos, listen to music, or meditate. A researcher assisted in fitting the wearable sensors: a heart rate and pulse sensor clipped to the left-hand pinky finger, and a respiration monitor belt secured around the lower ribs and diaphragm, inflated to a comfortable fit.

The experiment platform was placed on a tripod with the UWB radar facing the subject's chest, as shown in Figure 4(a). Recent research has highlighted challenges posed by relative motion between the sensor and the subject [12]. To address this, we also include a handheld case where a researcher holds the sensing platform, blending involuntary hand motion into the data (see Figure 4(b)). While not the primary focus of this paper, we will release this part of the data alongside the tripod data. Dataset statistics are shown in Table 1. In total, we collected 937 minutes of tripod data from 12 subjects and 546 minutes of handheld data from 11 subjects, ensuring diversity across users and different respiration patterns on various days.

# 3 System Design

## 3.1 Overview

The system design for MobiVital is illustrated in Figure 5. The input to the system is the entire complex UWB matrix (120 x 1500). As shown in literature, both the magnitude and phase of a UWB signal contain respiration waveform information. Therefore, MobiVital

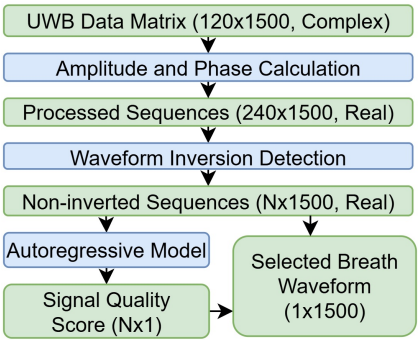



**Figure 5: Overview of the MobiVital system design. $N$ is the number of sequences classified as noninverted.**

computes the magnitude and phase for each distance bin and juxtaposes them as parallel channels, resulting in 240 possible candidates. The phase is calculated based on the angle of the complex I/Q value, followed by an unwrap operation to handle gaps around $2\pi$. These 240 candidates are then filtered using our Inversion Detector, Algorithm 1, which removes any sequences classified as inverted. The remaining candidates are scored by the autoregressive predictor. Finally, MobiVital selects the sequence with the highest MobiVital score and returns both the selected sequence and its score.

## 3.2 Autoregressive Predictor

MobiVital leverages an autoregressive deep learning model to automatically select the best sequence to represent the user's respiration, among all the candidate time series. The intuition of this method is to leverage the tendency of deep neural networks to overfit the distribution of their training datasets. This tendency is often considered to be detrimental, causing neural networks to work well on data with a similar distribution to the training dataset, and work poorly if the distribution of data is different. However, we see opportunities in this generalization limit. If we can train a neural network that learns the dynamics or the distribution of high-quality respiration data, the neural network's performance becomes an indicator: a good performance suggests "in distribution" (high signal quality).

The autoregressive model *training* pipeline is detailed in part a of Figure 6(a). The word "autoregressive" means the task of the model is to reconstruct part of the sequence given some history values. Since we want the model to learn the dynamics of respiration waveforms, we construct a training dataset using (1) ground truth respiration waveform from the on-body sensors and (2) high-quality UWB-measured respiration waveform. High-quality sequences are defined as sequences that have a correlation coefficient with the ground truth respiration waveform higher than a threshold $r_0 = 0.9$.

The sequences in the training dataset are then processed using a sliding window of 4.5 seconds. Inside the window, the first 4 seconds (200 samples) become the "history", and the remaining 0.5 seconds (25 samples) become the "future". Then the sliding window moves forward in increments of 0.5 seconds to generate more

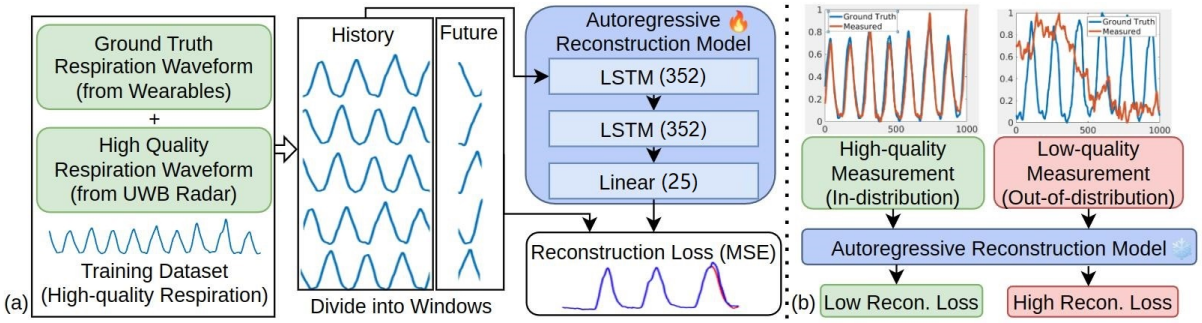



**Figure 6: Autoregressinve predictor. Part A shows the autoregressive model training pipeline. Part B shows the model's behavior during deployment when exposed to high and low-quality input sequence examples.**

**Algorithm 1:** Breath Waveform Inversion Detector

**Input:** Candidate time series $y \in \mathbb{R}^T, r_{th} = 1.0$
**Output:** Indicator of Waveform Inversion $inv \in [0, 1]$

$y_s \leftarrow SavitzkyGolayFilter(y, order = 5, frame\_len = 100)$ ;
/* Smooth the original waveform */
$P_{pos} \leftarrow FindPeaks(y_s, prominence = 0.1)$
$W_{pos} \leftarrow PeakWidth(y_s, P_{pos}, rel\_height = 0.5)$
$w_{pos} \leftarrow Average(W_{pos})$
$P_{inv} \leftarrow FindPeaks(-y_s, prominence = 0.1)$
$W_{inv} \leftarrow PeakWidth(-y_s, P_{inv}, rel\_height = 0.5)$
$w_{inv} \leftarrow Average(W_{inv})$
**if** $w_{pos}/w_{inv} < r_{th}$ **then** $inv \leftarrow 0$ **else** $inv \leftarrow 1$

histories and futures. Our autoregressive reconstruction model is a lightweight 2-layer LSTM model followed by a linear layer, as LSTM models are regarded as effective and efficient solutions for sequential problems. The model takes the history as the input and predicts the future sequence of 0.5s. The Mean Squared Error (MSE) between the predicted segment and the future segment is used as the loss function. This MSE loss is only used during training. The details of the model structures are shown in Figure 6(a). The model hyperparameters and history/future lengths are optimized using Bayesian hyperparameter turning (see [8] for additional details).

Figure 6(b) shows the autoregressive model's behavior during deployment time (with the model's weights frozen). The model accurately reconstructs the time series if the candidate measurement is high-quality, i.e., the signal "looks like" the respiration signals the model was trained on. However, if the input sequence is low-quality and the model was not trained on similar data, the model struggles to reconstruct the sequence, resulting in a high reconstruction loss.

## 3.3 Inversion Detector

The UWB candidate time series may be inverted with respect to the direction of chest movement. While this issue is often overlooked in prior works focused on breathing rate estimation, it is crucial for our goal of producing high-quality waveforms. Therefore, we propose a biology-informed algorithm to detect inverted UWB sequences, utilizing the morphology of respiration.

The algorithm is inspired by the biological nature of human respiration. During inhalation, chest and abdominal muscles contract to draw air into the lungs. Upon exhalation, these muscles relax, allowing the lungs to deflate [5]. Since holding air in the chest requires active muscle effort, normal respiration waveforms typically show spikes during active inhalation-exhalation, with longer gaps between these spikes. In other words, if we define the deflated chest position as "0" and the inflated position as "1", the "duty cycle" $\frac{t_1}{t_1+t_0}$ should be less than 0.5.

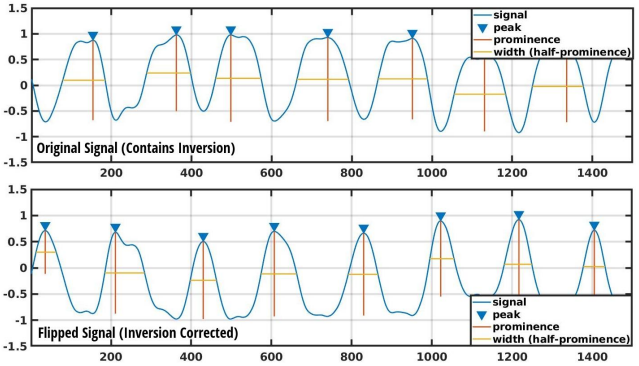



**Figure 7: Example of the breath waveform inversion detector. The upper original signal has an average peak width of 119.46 and the mirrored version below has 70.17. The signal is considered to be inverted.**

Based on this observation, we propose the algorithm detailed in Algorithm 1. The procedure starts by applying a Savitzky-Golay finite impulse response (FIR) smoothing filter with a polynomial order of 5 and a frame length of 100 to process the candidate sequence, removing jitters and microspikes. After smoothing, we identify the peaks in the signal and calculate the average peak width $W_{pos}$. The peak width is determined at height $h = h_{peak} - 0.5P$, where $P$ is the prominence of the peak. The same procedure is then applied to the flipped version of the signal, yielding $W_{inv}$. The signal is classified as inverted if $W_{pos}/W_{inv} \leq r_{th}$, and as non-inverted otherwise.

## 3.4 MobiVital Score

As the last step, the MobiVital score is a metric produced by the autoregressive model used to determine the quality of each possible candidate sequence. Alg. 2 describes the calculation algorithm. The signal $y$ is firstly chopped into a history set $T^{his}$ and a future set $T^{fut}$. For all the entries in $T^{his}$, we use the autoregressive model to predict a future $t_{pred}$. Then, we calculate the correlation coefficient $r_{c,i}$ between $t_{pred}$ and $T_i^{fut}$. Averaging over all the correlation coefficient $r_{c,i}$'s gives us the MobiVital score $m_s \in [-1, 1]$. In this

**Algorithm 2:** MobiVital Score Calculation

**Input:** Candidate time series $y \in \mathbb{R}^T$
**Output:** Mobivital score $m_s \in [-1, 1]$

$T^{his}, T^{fut} \leftarrow SlidingWindow(y)$
$m_s \leftarrow 0, len \leftarrow ||T^{his}||$
LOOP FOR $i$ FROM 1 TO $len$
  $t_{pred} \leftarrow AutoregssiveModel(T_i^{his})$
  $m_s \leftarrow m_s + CorrCoeff(T_i^{fut}, t_{pred})/len$

algorithm, no additional ground truth signal is used and the signal waveform itself serves as the ground truth, that is why we call the pipeline "self-supervised".

## 4 Evaluations

**Dataset Creation.** Our dataset includes 12 subjects, each contributing more than hour of tripod-mounted sensor data. To assess our model's generalization ability, we reserve data from users G, H, I, and J for final evaluations, while the remaining 8 users' data are used for training and hyperparameter optimization.

**Metric.** The correlation coefficient $r$ is used as the metric to quantify the similarity between a UWB sequence and the ground truth. It is robust to constant offsets and scaling but sensitive to time shifts, noise, and distortion, making it ideal for evaluating signal quality. Two time series need to have not only aligned peak locations but also similar morphologies (shape and trend) to receive a high $r$ score. The bounded range $r \in [-1, 1]$ also provides a standardized measure superior to L1- or L2-norms.

**Baselines.** We evaluate three widely used baseline methods: Variance [1, 2, 11], Signal-to-Noise Ratio (SNR) Estimation [4], and Constant False Alarm Rate (CFAR) [13]. All methods follow the same initial pre-processing: a loop-back filter removes static clutters [13], followed by a detrending algorithm to eliminate polynomial trends. The methods then diverge in their approaches, as detailed below.

- Variance: Calculates the variance of each distance bin's signal and selects the bin with the highest peak after applying peak detection. Operates on amplitude only.
- CFAR: Performs FFT on each bin's time series to generate a range-FFT map, then applies CFAR to identify the most significant peak. Operates on amplitude only.
- SNR: Computes the ratio of respiration-band energy (0.2-0.7 Hz) to total energy after FFT, selecting the bin with the highest SNR. Uses both amplitude and phase.

For fairness, our inversion detected algorithm 1 is applied to all the baseline methods as well. For CFAR and Variance, if a time series is determined to be inverted, the time series will be flipped, as these two baselines require peak detection on the entire landscape. In MobiVital and SNR, a time series loses its "candidacy" if determined inverted, since these two baselines can individually score each sequence. Finally, an oracle gives the theoretical upper bound performance by selecting the sequence with the highest correlation to the ground truth (unavailable in real-life applications).

### 4.1 Quantitative Results

**MobiVital score as a surrogate of signal quality.** For the score to be a good surrogate of the signal quality, it must ideally be bounded and have a positive linear relationship with the ground truth correlation. The MobiVital score satisfies both of these criteria. Fig. 8 shows the MobiVital score on all the time series from the eight subjects used for developing MobiVital. The score has an almost 1 to 1 relationship with the ground truth correlation. The score is also strictly bounded from -1 to 1 inclusive. Thus, the MobiVital score can be used to measure the quality of the signal, without any knowledge of the ground truth respiration waveform.

**Comparison with the baselines.** In Table 2, we compare the performance of MobiVital with the three baselines. The table shows that *MobiVital has a clear advantage over the baselines* in waveform

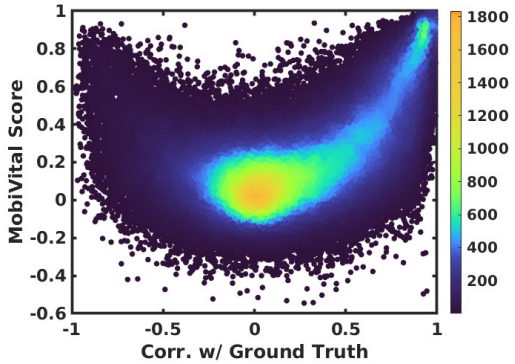



**Figure 8: MobiVital score on all candidate time series. This figure shows that MobiVital score can be a good surrogate of the time series quality.**

quality and is the closest to the oracle. MobiVital improves upon the closest baseline, SNR, by 7%, achieving a score of 0.819. Without our inversion detection algorithm, the gap increases to 34%. The advantage is even greater over Variance and CFAR, likely because these methods prioritize the strongest human body reflections, which do not always correspond to the best respiration signal.

| Method | w/ Inv Det. | w/o Inv Det. |
|---|---|---|
| MobiVital | **0.819** | **0.816** |
| SNR | 0.745 | 0.475 |
| CFAR | 0.516 | 0.218 |
| Variance | 0.514 | 0.225 |
| Oracle | - | 0.943 |

**Table 2: Average score of methods**

**Ablation study on the inversion detector.** Table 2 and Figure 9 show the improvement the inversion algorithm brings. The inversion algorithm doubles the performance for CFAR, Variance, and SNR. MobiVital sees a smaller 0.03 gain since the autoregressive model already rejects some inverted signals (the model is trained on a dataset without any inverted signals). The inversion detection algorithm is very lightweight with minimal computation overhead.

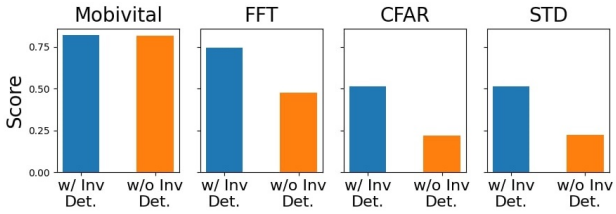



**Figure 9: Comparison of methods with and without the inversion detection algorithm.**

**Downstream task: respiration rate estimation.** We evaluate MobiVital 's impact on downstream tasks, using respiration rate (RR) estimation as an example. Ground truth RR is derived using a Savitzky–Golay filter and peak detection. Table 3 shows that MobiVital significantly reduces RR error compared to baselines, confirming that higher-quality signals improve downstream analysis accuracy and reliability.

## 5 Limitations and Future Work

**Improving Dataset Diversity**. The dataset used to develop MobiVital contains 12 users, and the data from the tripod-mounted

| Seq. Selection Method | MobiVital (Proposed) | SNR | CFAR | Variance |
|---|---|---|---|---|
| RR Error (bpm) | 0.68 | 1.02 | 1.73 | 1.78 |

**Table 3: Comparison of the respiration rate estimation error with different sequence selection methods.**

platform is about 12 hours. Despite being a relatively large user study, the current dataset is limited to an ideal lab setting with static-sitting subjects. Also, as a dataset collected in a university, the experiment subjects can have demographic biases such as age and health conditions. Expanding experiments to more realistic conditions, including real-world field studies and large-scale experiments involving diverse populations and medical treatments, would significantly enhance the practical effectiveness of the proposed dataset.

**Building More Applications and A Real-time System.** We have demonstrated that MobiVital improves the time series quality of UWB-respiration signals, leading to more accurate respiration rate calculation. It would be valuable to further investigate whether applications that require fine-grained respiration waveform analysis, such as inhale-exhale ratio estimation and tidal volume measurement, can benefit from MobiVital, which is a primary motivation of this work. Another important aspect is the feasibility of building a real-time system. MobiVital's training and evaluation are conducted offline on a PC with pre-collected datasets. However, we have designed MobiVital with efficiency in mind: the autoregressive model is simplified to a two-layer LSTM, and the inversion detection algorithm is computationally lightweight. Integrating MobiVital into existing mobile applications, particularly "breath coaching" apps on smartphones and smartwatches, could enable real-time respiration sensing and feedback. These apps currently rely on verbal or visual cues to guide users in respiration training for activities like yoga and meditation. A system that remotely monitors respiration and provides real-time guidance would significantly enhance user experience and training effectiveness.

**Mitigating Sensor-subject Relative Motion.** Motion interference in vital sign monitoring systems is becoming an increasingly interesting topic. Researchers have studied mitigating the movement of the sensing platform [12] and the body movements of the subjects [14]. The MobiVital system mostly considers static subjects during sleeping, sedentary activities, or meditations. However, A part of our dataset is collected with a handheld sensor platform, where the relative motion between the sensor and the subject contaminates the data and may even mask the vital signals. We hope future research efforts can lead us to high-quality respiration waveforms when the user is exercising, or using robot/cellphone-mounted sensors. Finally, conducting a longitudinal evaluation that captures body variations throughout the day or after different physical activities would offer deeper insights into the robustness and real-world applicability of MobiVital.

## 6 Conclusions

In this work, we highlight previous overlooked signal quality issues in UWB-based respiration monitoring, such as deformation and inversion. We introduce MobiVital, a system that uses a self-supervised autoregressive model and a bio-informed algorithm to generate high-quality respiration waveforms. Our MobiVital score effectively serves as a signal quality surrogate. The system design strategically leverages the limited generalization of neural networks, turning a common limitation into an engineering advantage. Future work can enhance MobiVital by improving dataset diversity, optimizing for real-time mobile applications in breath coaching, and addressing the challenges caused by sensor-subject relative motion to expand applicability to dynamic scenarios like exercise or wearable sensing.

## Acknowledgments

The research reported in this paper was sponsored in part by the Army Research Laboratory (ARL) under Cooperative Agreement W911NF1720196, and the NIH mHealth Center for Discovery, Optimization and Translation of Temporally-Precise Interventions (mDOT) under award 1P41EB028242. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the funding agencies.

## References

[1] Muhammad Husaini, Latifah Munirah Kamarudin, Ammar Zakaria, Intan Kartika Kamarudin, Muhammad Amin Ibrahim, Hiromitsu Nishizaki, Masahiro Toyoura, and Xiaoyang Mao. 2022. Non-contact breathing monitoring using sleep breathing detection algorithm (SBDA) based on UWB radar sensors. *Sensors* 22, 14 (2022), 5249.

[2] Jong Deok Kim, Won Hyuk Lee, Yonggu Lee, Hyun Ju Lee, Teahyen Cha, Seung Hyun Kim, Ki-Min Song, Young-Hyo Lim, Seok Hyun Cho, Sung Ho Cho, et al. 2019. Non-contact respiration monitoring using impulse radio ultrawideband radar in neonates. *Royal Society open science* 6, 6 (2019), 190149.

[3] Janosch Kunczik, Kerstin Hubbermann, Lucas Mösch, Andreas Follmann, Michael Czaplik, and Carina Barbosa Pereira. 2022. Breathing pattern monitoring by using remote sensors. *Sensors* 22 (2022), 8854.

[4] Siyun Liu, Qingjie Qi, Huifeng Cheng, Lifeng Sun, Youxin Zhao, and Jiamei Chai. 2022. A vital signs fast detection and extraction method of UWB impulse radar based on SVD. *Sensors* 22, 3 (2022), 1177.

[5] National Heart Lung and Blood Institute. 2022. How the Lungs Work - How Your Body Controls Breathing | NHLBI, NIH. https://www.nhlbi.nih.gov/health/lungs/body-controls-breathing.

[6] Andrea Nicolò, Carlo Massaroni, Emiliano Schena, and Massimo Sacchetti. 2020. The importance of respiratory rate monitoring: From healthcare to sport and exercise. *Sensors* 20, 21 (2020), 6396.

[7] Hongming Shen, Chen Xu, Yongjie Yang, Ling Sun, Zhitian Cai, Lin Bai, Edward Clancy, and Xinming Huang. 2018. Respiration and heartbeat rates measurement based on autocorrelation using IR-UWB radar. *IEEE transactions on circuits and systems II: express briefs* 65, 10 (2018), 1470–1474.

[8] Ziqi Wang, Derek Hua, Wenjun Jiang, Tianwei Xing, Xun Chen, and Mani Srivastava. 2025. MobiVital: Self-supervised Time-series Quality Estimation for Contactless Respiration Monitoring Using UWB Radar. *arXiv preprint arXiv:2503.11064* (2025).

[9] Zongxing Xie, Ava Nederlander, Isac Park, and Fan Ye. 2023. Short: RF-Q: Unsupervised Signal Quality Assessment for Robust RF-based Respiration Monitoring. In *Proceedings of the 8th ACM/IEEE International Conference on Connected Health: Applications, Systems and Engineering Technologies*. 158–162.

[10] Zongxing Xie, Hanrui Wang, Song Han, Elinor Schoenfeld, and Fan Ye. 2022. DeepVS: A deep learning approach for RF-based vital signs sensing. In *Proceedings of the 13th ACM international conference on bioinformatics, computational biology and health informatics*. 1–5.

[11] Yanni Yang, Jiannong Cao, Xiulong Liu, and Xuefeng Liu. 2019. Multi-breath: Separate respiration monitoring for multiple persons with UWB radar. In *2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC)*, Vol. 1. IEEE, 840–849.

[12] Fusang Zhang, Jie Xiong, Zhaoxin Chang, Junqi Ma, and Daqing Zhang. 2022. Mobi2Sense: empowering wireless sensing with mobility. In *Proceedings of the 28th Annual International Conference on Mobile Computing And Networking*. 268–281.

[13] Tianyue Zheng, Zhe Chen, Chao Cai, Jun Luo, and Xu Zhang. 2020. V2iFi: In-vehicle vital sign monitoring via compact RF sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 2 (2020), 1–27.

[14] Tianyue Zheng, Zhe Chen, Shujie Zhang, Chao Cai, and Jun Luo. 2021. MoRe-Fi: Motion-robust and fine-grained respiration monitoring via deep-learning UWB radar. In *Proceedings of the 19th ACM conference on embedded networked sensor systems*. 111–124.