

Peer Review File

Monitoring Long-term Cardiac Activity with Contactless Radio Frequency Signals



Open Access This file is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. In the cases where the authors are anonymous, such as is the case for the reports of anonymous peer reviewers, author attribution should be to 'Anonymous Referee' followed by a clear attribution to the source work. The images or other third party material in this file are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

Thank you for your detailed feedback on our manuscript "Monitoring Long-term Cardiac Activity with Contactless Radio Frequency Signals". We appreciate the thoughtful consideration you and your colleagues have given to our work. In what follows, we present detailed comments in response to the individual points raised by the associate editor and elaborate on how the manuscript has been revised.

Comment 1: *Specifically, other works claim that effective disentangling of breathing and cardiac signals is possible e.g 10.3390/s22208047 and 10.1109/INFOCOM48880.2022.9796905. Without a comprehensive discussion about these and similar works, along with convincing proof of novelty, the advances that your approach brings in are unclear.*

Response 1: We thank the editor for the comment. As suggested by the editor, we have included a more comprehensive discussion about prior work and revised the statements in the manuscript to demonstrate the novelty of our study. Specifically, while we acknowledge that other works have explored the disentanglement of breathing and cardiac signals, our work is fundamentally different from the existing approaches in several key aspects:

1. **For the first time, we disentangle heartbeat and respiration signals in far-field conditions without any model training.** Existing work by Sharma et al. [1] leverages a near-field radio-frequency (RF) sensor worn at the thorax to reduce the influence of abdomen motion caused by respiration. However, respiration can still lead to a certain level of thorax motion, which is significantly larger than heartbeat motion. Zhang [et.al](#) [2] propose to learn the mapping between RF signal and heartbeat waveform with deep neural networks. Nevertheless, their black-box machine learning model requires massive training data to generalize across different individuals. In contrast, we use a far-field RF sensor to directly disentangle vital signs from the body surface, which is fully contactless and considerably more challenging. Moreover, our system is a pure signal processing framework that is totally interpretable and generalizes well without any training data.

2. **For the first time, we leverage high-frequency ranges and heartbeat harmonics for precise heartbeat extraction.** Existing work by Sharma et al. [1] only extracts heartbeat signals from the heart rate frequency band (around 2 Hz). In contrast, we have discovered that heartbeat information predominates over other motions (e.g., respiration) in higher frequency ranges (around 15 Hz), allowing us to more effectively disentangle the heartbeat signal from the respiratory signal. Moreover, traditional approaches have considered heartbeat harmonics as detrimental and something that should be suppressed. However, we leverage the beat frequency effect to generate prominent heartbeat patterns by superimposing these heartbeat harmonics, which challenges the conventional views and extracts heart rate variability with **10x improvement in accuracy** compared to state-of-the-art methods.
3. **For the first time, we conduct extensive evaluations in both large-scale clinical settings and long-term daily scenarios.** Since existing work [1, 2] only conducted evaluation with fewer than 20 subjects, whether their system can support large-scale or long-term deployment is unclear. In contrast, we have extensively validated our system across a diverse patient population involving **7,090 participants** and a long-term monitoring spanning **60 days**, demonstrating its reliability and generalizability in a way that, to the best of our knowledge, has not been achieved by previous studies.

These key innovations set our work apart from the existing studies and enable us to achieve a novel and comprehensive system for contactless monitoring of long-term cardiac activity. We believe that these advancements, when presented in the full context of the existing literature, will make a significant contribution to the field of contactless cardiac monitoring.

Comment 2: *Moreover, per our journal's ethical guidelines, you should give more information about the clinical studies you carried out. In particular, you should confirm that all the patients gave informed consent for the recording and analyzing of their data, as well as indicate if this work has been carried out in the framework of a clinical trial. In this regard, it would be useful to have a protocol approved by an ethical commission, if available.*

Response 2: We thank the editor for the comment. We have included the **Ethics Statement and Informed Consent** in our manuscript. Specifically, the clinical studies were conducted after ethics review by the Ethics Review Board of The First Affiliated Hospital of the University of Science and Technology of China (Protocol 2023 KY Ethics Review No. 056). When registering, all participants of the clinical studies agreed that their data could be used for scientific research and

were made aware that they could revoke this consent at any time. Participants received no compensation for their participation.

[1] P. Sharma, Z. Zhang, T.B. Conroy, X. Hui, E.C. Kan, "Attention Detection by Heartbeat and Respiratory Features from Radio-Frequency Sensor." *Sensors*, 2022.

[2] S. Zhang, T. Zheng, Z. Chen and J. Luo, "Can We Obtain Fine-grained Heartbeat Waveform via Contact-free RF-sensing?," *IEEE Conference on Computer Communications*, 2022.

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

This work has demonstrated remarkable results of IBI estimation on a large scale study, and preliminary demonstration towards long-term overnight monitoring using contactless RF.

One of the novelty of proposed approach is using beat frequency effect for HR estimation involving higher order harmonics. Another value addition or highlight of the work is large-scale clinical evaluation.

- While this work is unique in only using higher order harmonics and beat frequency effect, I would recommend authors to acknowledge other works where higher order HR harmonics are leveraged.

Examples: 10.3390/s22010083 and 10.1109/TAES.2019.2917489

- The long-term claim of the authors are not clear. For reference, in abstract and on Pg 4, authors claim 60 days of monitoring. Pg 8 mentions 5 nights spanning 2 months. So is long-term monitoring performed over random 5 nights in 2 month duration? In that case, abstract and other instances should make that distinction clear.

- Also, 'long-term daily life' claim (Pg 4) cannot be asserted with demonstrated results, as long-term demonstration is done only in sleep state. The authors have shown performance with one relatively fixed orientation of sensor and subject. To make this claim, few important considerations need to be addressed including impact of multi-subject presence, sensor angle and distance change to user, subject clothes with multiple layers, and non-line of sight sensor.

- The performance comparison with mmHRV and V2iFi show >140 ms median RT-IBI error. As these papers claim median IBI error of 28 ms for mmHRV and 50 ms for V2iFi, can authors comment on why such a performance degradation is seen on their dataset? Further, author should clarify that they have used other algorithms on their dataset, instead of comparing with other papers' quoted results.

- Performance comparison can include a table, that summarizes details of other works (mmHRV, V2iFi) - including claimed results and #subjects, aims, multiuser capability.

- To support clinical data, further details regarding population set can be helpful. What kind of patients are involved, do they have any CVD, history of stroke or surgery, etc. Why only 30s data is collected in this scenario? Further, were any datasets rejected due to motion or other corruption issues, if so, what's that number? Was any sensor adjustment needed or performed based on subject age, height, weight or gender?

- In sleep data collection, data is divided into slices, what's the duration of each slice?

- I would recommend adding paragraph discussing problems faced during clinical study, if any, and limitations of the approach - both sensor and algorithm.

Reviewer #2 (Remarks to the Author):

Referee Report:

In this paper, Bin-Bin et al. introduce an FMCW radar-based sensing system designed for long-term monitoring of heart rate variability (HRV) and detecting heart abnormalities. They propose that their system has the potential to facilitate proactive self-assessment and enhance medical prevention through continuous, precise health monitoring. The authors utilized the VMD algorithm to extract heartbeat patterns. To demonstrate the system's efficacy in HRV monitoring and abnormality detection, they conducted an extensive evaluation, including a large-scale clinical study with 7,090 participants and a continuous monitoring study on one subject over 5 nights. However, the manuscript lacks sufficient novelty in sensing technology, lacks evidence demonstrating clinical-level usability, and struggles with presenting and visualizing its large dataset in a clear and coherent manner. Here are some suggestions the authors should consider enhancing the paper's content.

Major Comments:

- 1) This manuscript misses essential subject information, including statistics on gender, height, weight, effective recording time, health condition, and whether there is a ground truth comparison.
- 2) The authors used an FMCW radar system for heart function monitoring. Many published works have utilized FMCW radar for human vital signs monitoring ([1-3] and reference [23]). Additionally, using VMD for heart feature extraction from FMCW radar signals is a popular approach ([4,5]). What innovations do your proposed hardware and signal processing algorithms offer compared to previous works?

[1] Z. Ling, W. Zhou, Y. Ren, J. Wang and L. Guo, "Non-Contact Heart Rate Monitoring Based on Millimeter Wave Radar," in *IEEE Access*, vol. 10, pp. 74033-74044, 2022, doi: 10.1109/ACCESS.2022.3190355.

[2] J. Liu, Y. Li, C. Li, C. Gu and J. -F. Mao, "Accurate Measurement of Human Vital Signs With Linear FMCW Radars Under Proximity Stationary Clutters," in *IEEE Transactions on Biomedical Circuits and Systems*, vol. 15, no. 6, pp. 1393-1404, Dec. 2021, doi: 10.1109/TBCAS.2021.3123830.

[3] S. M. M. Islam, N. Motoyama, S. Pacheco and V. M. Lubecke, "Non-Contact Vital Signs Monitoring for Multiple Subjects Using a Millimeter Wave FMCW Automotive Radar," 2020 IEEE/MTT-S International Microwave Symposium (IMS), Los Angeles, CA, USA, 2020, pp. 783-786, doi: 10.1109/IMS30576.2020.9223838.

[4] K. J. Wu and C. -L. Yang, "Heart Rate Extraction with VMD Algorithm in Non-Stationary Clutter Environment Based on FMCW Radar Systems," 2021 IEEE International Symposium on Radio-Frequency Integration Technology (RFIT), Hualien, Taiwan, 2021, pp. 1-3.

[5] Balasooriya, T., Nallabolu, P., Li, C. (2022). Application of Variational Mode Decomposition to FMCW Radar Interference Mitigation. In: Suryadevara, N.K., George, B., Jayasundera, K.P., Roy, J.K., Mukhopadhyay, S.C. (eds) *Sensing Technology. Lecture Notes in Electrical Engineering*, vol 886. Springer

- 3) In this manuscript, the authors emphasized the interference and signal morphology distortion caused by breathing. However, it is evident that there is a clear frequency band gap between breathing (fundamental tone and second harmonic) and heart rhythm in most healthy subjects. Could a high-pass filter be used to remove the low-frequency breathing signals? Additionally, did you employ any other signal pre-processing methods before applying VMD?

4) In this paper, the authors presented sum-to-product formulas for trigonometric functions, which are commonly known. They attempted to explain that the mixing of breathing signals and heartbeat harmonics exhibits unique features in the spectrum but did not provide sufficient evidence. This analysis appears unrelated to their heart rate variation study, as peak detection, which is more associated with the envelope of the IMF, was used. Please provide further clarification or consider removing this section.

5) In this paper, the heartbeat signal extraction highly relied on the VMD. How does each imf match with your raw signal frequency spectrum bands? Figure 5e has a strange layout and is unclear, and in supplementary figure 4, the authors only showed a selected band of spectrum which is a duplication of Figure 5d and 5e. Can you also show the whole spectrum, maybe also show the HHT spectrum of your imfs to further verify that? What is the HRV computation time of different signal epochs?

6) Several important technical details are missing. What are the sample digits and sampling rate of each radio channel? Are you transmitting both magnitude and phase signals or only the phase signal? What is your Wi-Fi data rate? Additionally, what is the time window for heart rate calculation, and what is the window overlap? A schematic of your radar system should also be included.

7) The authors proposed a radar system for long-term continuous HRV monitoring of subjects on a bed, emphasizing its non-contact nature. However, line-of-sight blockage and subject postures (supine or lateral) can affect the backscattered signal. The authors should provide more information on how their system performs under various postures and how these different postures impact heartbeat variation.

8) Furthermore, when the subject moved during the recording, how did you identify these movements, and did you still use these epochs to calculate heart rate? If moving epochs were excluded, what was the ratio of usable epochs (the windows used to calculate heart rate and HRV) in the long-term recording?

9) When comparing real-time IBI and post-processing IBI, the authors discard certain time segments to reduce 'error'. How can you determine if this 'missing' pattern is due to RF signal loss, heart halt, or another abnormality? How do you appropriately choose the threshold for this identification?

10) The authors included many algorithm comparison plots in the main figures. Are the algorithms you compared to widely used methods or just your previous works? To demonstrate your system's advantages or its potential clinical grade applications, comparing the IBI features and abnormality classification results with those from other clinical devices like ECG, PPG, or other radar methods such as Doppler radar or near-field RF sensors could make a stronger statement.

11) In the abstract, the authors emphasized "a long-term daily scenario spanning two months", which could be misleading. To clarify, the study involves recording individuals for 5 nights over two months.

12) In Supplementary Figure 2, the authors did not specify the definition of SNR, including whether it is based on amplitude or power spectrum density, or the bandwidth used for calculating the noise floor.

13) The authors provided examples of heartbeat patterns resulting from the 'beat frequency effect' without any accompanying labels or explanation in both the main manuscript and supplementary

materials. This oversight reflects negatively on the authors' attentiveness and responsibility to the reader.

Reviewer #2 (Remarks on code availability):

The authors didn't provide data or code in the submitted package.

Reviewer #3 (Remarks to the Author):

The noteworthy results are that, for the first time, the proposed system successfully overcomes the orders-larger interference from respiration motion in far-field conditions without any model training.

And the large-scale clinical study

proves that this system can achieve comparable performance to clinical-grade ECG-based systems.

Despite the noteworthy results, I have some comments and questions below:

1. In the introduction part, the authors stated that "we propose to leverage previously undiscovered higher frequency ranges where heartbeat information dominates over other motions." This seems to be the most important novelty of this paper. But even in ref [31] P. Sharma, Z. Zhang, T. B. Conroy, X. Hui, and E. C. Kan, "Attention detection by heartbeat and respiratory features from radio-frequency sensor," *Sensors*, vol. 22, no. 20, p. 8047, 2022. "To accurately extract the inter-beat interval (IBI) from the smooth NCS signal, a weaker, but sharper second harmonic heartbeat component of the heartbeat was used [32] and IBI was measured as the time for two cycles, as denoted in Figure 3b.", in this paper, heart rate was extracted from second harmonics to get more further away from respiration band. From this example, I wouldn't say higher frequency ranges are totally "undiscovered" and this method is used "for the first time".

2. The authors compared with other systems like mmHRV and V2Fi, I'm curious to see if a wider range of algorithms can be considered? For example, some learning based algorithms like CNN, especially considering the large datasets you have? A classical machine learning or deep learning algorithm may have a good result over large population and boost up the accuracy.

3. In Fig.1 d, both in clinical and daily life scenarios, the subjects are lying in bed. This has the least motion interference to the RF sensor, and easier to get higher signal quality. However, as the authors stated in the introduction, "our system shows immense potential in providing a comprehensive understanding of an individual's health conditions in their own homes", in real remote healthcare setting at home, the users will not always stay in bed. When users change to daily activities, like walking, talking, or exercising, will the accuracy degrades? I think that will be the important question to answer if the impact of this sensing system is more than algorithm improvement.

4. The authors mentioned that "Slices containing significant body motion, as indicated by pronounced RF

signal variations, were excluded from further analysis." Could you specify the percentage or total time of recording that was polluted by body motion? What percentage of time can you get accuracy read of HR?

5. In Fig.3.c, why the accuracy of ECG was not directly reported here?

Response to Review of Paper:

NCOMMS-24-30589A-Z

Monitoring Long-term Cardiac Activity with Contactless Radio Frequency Signals

We would like to thank all the reviewers for their thorough reviews and constructive comments. We have carefully revised the paper to clarify and address the reviewers' comments. In the document that follows, we describe changes made to the original version of the paper and address the comments of the reviewers.

I. RESPONSES TO REVIEWER 1

We would like to express sincere thanks to the reviewer 1 for the thorough and constructive comments. In what follows, we present detailed comments in response to the individual points raised by the reviewer and elaborate on how the manuscript has been revised.

Comment 1.1: *While this work is unique in only using higher order harmonics and beat frequency effect, I would recommend authors to acknowledge other works where higher order HR harmonics are leveraged. Examples: 10.3390/s22010083 [12] and 10.1109/TAES.2019.2917489 [8].*

Response 1.1: We thank the reviewer for the comment and suggestion. As suggested by the reviewer, we have revised the manuscript to acknowledge these works that utilize higher order heartbeat harmonics [8, 12]. For details, please refer to line 264 on page 11.

Comment 1.2: *The long-term claim of the authors are not clear. For reference, in abstract and on Pg 4, authors claim 60 days of monitoring. Pg 8 mentions 5 nights spanning 2 months. So is long-term monitoring performed over random 5 nights in 2 month duration? In that case, abstract and other instances should make that distinction clear.*

Response 1.2: We thank the reviewer for the comment. In the original manuscript, the assessment of long-term monitoring indeed involved randomly collecting 5 nights of data over a two-month period to conduct randomized follow-up experiments. As suggested by the reviewer, we have carefully revised the manuscript to clarify the descriptions related to long-term evaluation.

Additionally, to further assess the system's capability for long-term and continuous monitoring, we have supplemented an experiment with 21 consecutive nights of monitoring. The comparative performance of our system against two baseline methods, as evaluated using the aforementioned four metrics, is presented in Figure 1a-d. The results demonstrate that our system's performance is consistent across both the 21 consecutive nights of data and the data from 5 random nights over two months. Specifically, RF-HRV achieved a median IBI error of 39.5 ms, a median RMSSD error of 24.6 ms, a median SDRR error of 21.3 ms, and a median pNN50 error of 10.3%. Compared to mmHRV and V2iFi, our system significantly outperformed them across all evaluation metrics and achieved over 10x improvement in estimating RMSSD and SDRR. To

provide a clearer overview of how our system performs across the entire long-term dataset, we depict the comparison of CDF of the four metrics in Figure 1e-h. It is evident that our system consistently shows significantly better HRV estimation performance in long-term and continuous monitoring compared to existing systems.

We have included the additional experiments and results in the Supplementary Note 10 and Figure 11.

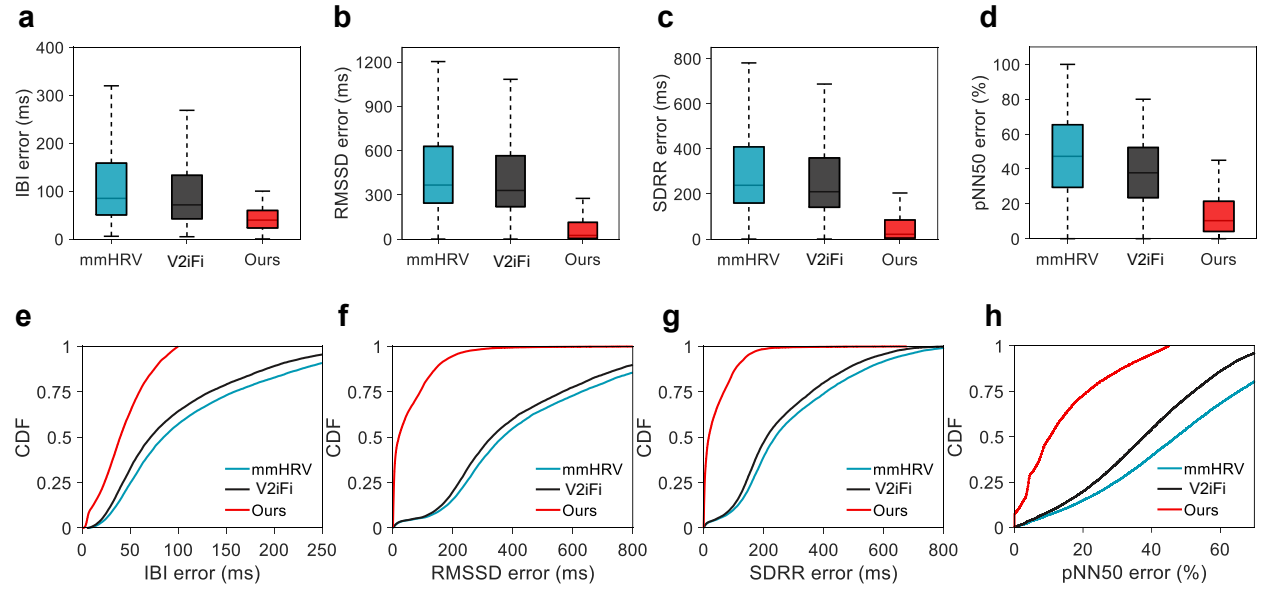


Fig. 1: Evaluation on 21 consecutive nights data. **a**, Comparison of real-time IBI error. RF-HRV achieves a 39.5 ms median RT-IBI error, while mmHRV and V2iFi achieve 84.9 ms and 71.3 ms median RT-IBI error, respectively. **b**, Comparison of RMSSD error. RF-HRV achieves a 24.6 ms median RMSSD error, while mmHRV and V2iFi achieve 366.1 ms and 329.5 ms median RMSSD error, respectively. **c**, Comparison of SDRR error. RF-HRV achieves a 21.3 ms median RT-IBI error, while mmHRV and V2iFi achieve 238.4 ms and 209.8 ms median SDRR error, respectively. **d**, Comparison of pNN50 error. RF-HRV achieves a 10.3% SDRR error, while mmHRV and V2iFi achieve 47.3% and 37.8% median SDRR error, respectively. **e,f,g,h**, The cumulative distribution function (CDF) plots of RT-IBI error, RMSSD error, SDRR error, and pNN50 error.

Comment 1.3: ‘Long-term daily life’ claim (Pg 4) cannot be asserted with demonstrated results, as long-term demonstration is done only in sleep state. The authors have shown

performance with one relatively fixed orientation of sensor and subject. To make this claim, few important considerations need to be addressed including impact of multi-subject presence, sensor angle and distance change to user, subject clothes with multiple layers, and non-line of sight sensor.

Response 1.3: We appreciate the reviewer’s insightful suggestion. As suggested by the reviewer, we have conducted additional experiments to investigate the impact of several critical factors on system performance, including variations in sensor angle, distance to the user, the multiple layers of clothing.

From Table I, Table II, Table III, we can see that as distance, angle, and the number of clothing layers increase, there is a slight decrease in system performance, mainly due to the reduction in the signal-to-noise ratio of the RF signal. At the same time, the experimental results also indicate that within a distance of 140 cm, 45 degrees, and about 3 layers of clothing, our system can achieve satisfactory HRV monitoring performance. This content has been added to line 247 on page 11 of the revised manuscript.

Furthermore, in this work, we focus on addressing the challenge of significant respiratory leakage that results in inaccurate HRV monitoring, supported by evaluations in single-person and line-of-sight scenarios. For scenarios involving multiple-user and non-line-of-sight, further algorithm design is needed to extend the application of our system. We leave these studies to future work and have elaborated on these points in the Discussion section.

TABLE I: Comparison of different distances from the subject.

		Metrics	RT-IBI Error (ms)	RMSSD Error (ms)	SDRR Error (ms)	pNN50 Error (%)
Distance (cm)	60	Mean	15.1	33.5	20.9	10.1
		Median	8.0	16.7	15.0	7.5
	100	Mean	17.5	56.2	25.0	12.5
		Median	19.8	56.9	24.5	12.4
	140	Mean	26.7	66.3	28.5	15.1
		Median	25.1	59.8	25.4	13.2

Comment 1.4: *The performance comparison with mmHRV and V2iFi show 140 ms median RT-IBI error. As these papers claim median IBI error of 28 ms for mmHRV and 50 ms for V2iFi, can authors comment on why such a performance degradation is seen on their dataset? Further,*

TABLE II: Impact of subject clothing with multiple layers.

Clothing with Multiple Layers	RT-IBI Error (ms)	RMSSD Error (ms)	SDRR Error (ms)	pNN50 Error (%)
1 layer	15.1	33.5	20.9	10.1
2 layers	17.3	38.6	24.1	12.5
3 layers	23.2	41.3	29.7	13.6

TABLE III: Impact of radar angle to subject.

Radar Angle (degrees)	RT-IBI Error (ms)	RMSSD Error (ms)	SDRR Error (ms)	pNN50 Error (%)
0	15.1	33.5	20.9	10.1
30	31.5	47.2	30.8	15.3
45	54.6	52.3	41.4	26.2
60	70.7	59.9	45.1	30.7

author should clarify that they have used other algorithms on their dataset, instead of comparing with other papers' quoted results.

Response 1.4: We thank the reviewer for the comment. As suggested by the reviewer, we have clarified that our reported results of V2iFi [13] and mmHRV [10] are obtained by re-implementing these algorithms and running them on our dataset. Moreover, the performance degradation of V2iFi and mmHRV on our dataset is primarily due to differences in evaluated datasets. Both V2iFi and mmHRV have limited and homogeneous test data, with the number of evaluated participants being 4 and 11, respectively. In contrast, we assessed the system performance with a large-scale outpatient dataset collected from a top-tier tertiary hospital, which consists of 6,222 participants with ages ranging from 3 to 94. The experimental results indicate that mmHRV and V2iFi struggle to maintain good performance on such a challenging dataset.

Comment 1.5: *Performance comparison can include a table, that summarizes details of other works (mmHRV, V2iFi) - including claimed results and subjects, aims, multiuser capability.*

Response 1.5: We thank the reviewer for the comment and suggestion. As suggested by the reviewer, we have included a table to show more details of mmHRV, V2iFi, and our method, as shown in Table IV. We have included the table in the Supplementary Note 9.

TABLE IV: Comparison of the proposed method with existing methods

Related Method	mmHRV	V2iFi	Our method
Number of Evaluated Subjects	11	4	6,222
Claimed Results (IBI error)	28 ms	50 ms	31.5 ms
Measured Results over 6,222 Subjects (IBI error)	168.3 ms	151.6 ms	31.5 ms
Frequency Band of Heartbeat Extraction	1 - 2 Hz	1 - 2 Hz	10 - 15 Hz
Heartbeat Pattern Extraction Technique	Signal Decomposition	Signal Decomposition	Signal Decomposition + Beat Frequency Effect
Outpatient Evaluation	No	No	Yes
Long-term Evaluation	No	No	Yes
Aims	HRV Monitoring	HRV Monitoring	HRV Monitoring & Abnormalities Identification
Multi-Subject Capability	Weak	Weak	Limited

TABLE V: Baseline demographic and characteristics. Simultaneous occurrences of different categories are present in ECG and laboratory findings.

Dataset	
7,090 subjects(eligible 6,222)	
Excluded from analysis	
Loosening of the electrode pads	868
Eligible subjects	6,222
Age (Years)	55±16(3 – 94)
Gender	
Male	2817 (45.3%)
Female	3405 (54.7%)
ECG and laboratory findings	
Sinus rhythm	5423 (87.2%)
Sinus bradycardia	428 (6.9%)
Sinus tachycardia	137 (2.2%)
Arrhythmia	245 (3.9%)
Left BBB	21 (0.3%)
Right BBB	125 (2.0%)
ST-segment changes	872 (14.0%)
T-wave changes	434 (7.0%)

Values are mean+-s.d. (minimum-maximum) or n(%).

Comment 1.6: *To support clinical data, further details regarding population set can be helpful. What kind of patients are involved, do they have any CVD, history of stroke or surgery, etc. Why only 30s data is collected in this scenario? Further, were any datasets rejected due to motion or other corruption issues, if so, what's that number? Was any sensor adjustment needed or performed based on subject age, height, weight or gender?*

Response 1.6: We thank the reviewer for the comment and have made the following clarification:

(1) We have provided more details of participants in Table V, including their age, sex, and the distribution of different CVDs.

(2) Collecting 30s data is a standard outpatient protocol.

(3) Data from 868 subjects were excluded due to loosening of the electrode pads when capturing the ground-truth ECG..

(4) During monitoring, we simply set the sensor at a certain height range (0.5m-1.4m) from the target, without the need to make special adjustments based on the subject's age, height, weight, or gender.

Comment 1.7: *In sleep data collection, data is divided into slices, what's the duration of each slice?*

Response 1.7: The duration of each slice is 30 seconds when dividing the sleep data.

Comment 1.8: *I would recommend adding paragraph discussing problems faced during clinical study, if any, and limitations of the approach - both sensor and algorithm.*

Response 1.8: We thank the reviewer for the suggestions. As suggested by the reviewer, we have included the problems faced during the outpatient study and the limitations of the approach in the revised manuscript as follows:

In this paper, we have validated the superior performance of our system in short-term monitoring (30s) across large-scale populations (6,222), as well as long-term monitoring (multiple nights) for a single subject. For a more comprehensive evaluation of the outpatient study, long-term validation on a large-scale population is necessary but yet to be implemented, which requires excessive resources recruiting more participants with additional funding support. This will be the direction in which we will continue our research in the future.

Moreover, while our RF-based system provides an important advancement over existing solutions, we also note that it has several limitations. First, we demonstrate the long-term monitoring capability of our system in single-person scenarios because our focus is to address the interference of respiratory leakage on heart rate estimation. However, in practice, two or more people may need to be monitored simultaneously. In such a case, the signals reflected by them may interfere with each other, requiring additional signal processing for multi-person sensing. Furthermore, most experiments conducted in this paper are predominantly in the scenario of an individual at rest in bed. While this activity presents hours of a day for people, the investigation of heart activity monitoring during sleep is of significant importance for the early identification and prevention of underlying health conditions. We also note that real-world applications often involve other short-term activities such as walking or exercising. Addressing these more dynamic scenarios, which are both significant and challenging, is crucial. We will dedicate more research efforts to advance the proposed HRV monitoring technique for covering a wider range of activities.

This above content has been included in revised manuscript on page 12, line 283.

II. RESPONSES TO REVIEWER 2

We would like to express sincere thanks to the reviewer 2 for the thorough and constructive comment. In what follows, we present detailed comments in response to the individual points raised by the reviewer and elaborate on how the manuscript has been revised.

Comment 2.1: *This manuscript misses essential subject information, including statistics on gender, height, weight, effective recording time, health condition, and whether there is a ground truth comparison.*

Response 2.1: We thank the reviewer for the comment and suggestion. As suggested by the reviewer, we have included statistics on gender, age, and health condition in Table VI. When gathering information about the experimental subjects, we did not collect their height and weight information. Moreover, we simultaneously collect radar data and ECG data with an effective recording time of 30 seconds, and use the ECG data as the ground truth of human heartbeats.

This table has been added in the revised manuscript on page 5.

Comment 2.2: *The authors used an FMCW radar system for heart function monitoring. Many published works have utilized FMCW radar for human vital signs monitoring ([5–7] and reference [23]). Additionally, using VMD for heart feature extraction from FMCW radar signals is a popular approach ([1, 11]). What innovations do your proposed hardware and signal processing algorithms offer compared to previous works?*

Response 2.2: We thank the reviewer for the comment and suggestion. While FMCW radar-based cardiac monitoring is gaining increasing attention in both academia and industry, a fundamental challenge, namely the substantial spectral leakage of the respiration motion into the heart rate frequency band, remains unresolved. Given that the amplitude of respiration significantly exceeds that of heartbeats, the interference of respiration harmonic can be observed in the heartbeat fundamental frequency band. Consequently, existing work cannot support long-term and accurate cardiac monitoring, especially for large-scale populations.

To address this issue, our study introduces two key innovations:

(i) We systematically analyze the impact of higher-order harmonics and observe a simple and fundamental yet unidentified phenomenon: respiration harmonics decay more rapidly than heartbeat harmonics. This is because the lower frequency of respiration compared to heartbeats results in higher-order respiratory harmonics that decay faster. Leveraging this discovery, we

TABLE VI: Baseline demographic and characteristics. Simultaneous occurrences of different categories are present in ECG and laboratory findings.

Dataset	
7,090 subjects(eligible 6,222)	
Excluded from analysis	
Loosening of the electrode pads	868
Eligible subjects	6,222
Age (Years)	55±16(3 – 94)
Gender	
Male	2817 (45.3%)
Female	3405 (54.7%)
ECG and laboratory findings	
Sinus rhythm	5423 (87.2%)
Sinus bradycardia	428 (6.9%)
Sinus tachycardia	137 (2.2%)
Arrhythmia	245 (3.9%)
Left BBB	21 (0.3%)
Right BBB	125 (2.0%)
ST-segment changes	872 (14.0%)
T-wave changes	434 (7.0%)

Values are mean±s.d. (minimum-maximum) or n(%).

develop the core insight that differs from existing works [1, 11], i.e., we extract more clean heartbeat information leveraging frequency ranges much higher than existing work, typically beyond 10-order heartbeat harmonics.

(ii) Based on the above insight, we successfully extract heartbeat signals from frequencies far from the heart rate frequency band leveraging the beat frequency effect. Specifically, extracting heartbeat from higher-frequency bands presents unique challenges, including the rapid attenuation of signal strength and the unknown order of heartbeat harmonics. To resolve this, we find that when two harmonics are superimposed, they generate a distinctive signal pattern with a frequency equal to the difference between the frequencies of the two harmonics. Therefore, we

generate strong heartbeat patterns that exhibit a frequency precisely equal to the heart rate by superimposing adjacent higher-order heartbeat harmonics.

In summary, while previous works have utilized FMCW radar for vital signs monitoring [5–7] and VMD for heart feature extraction [1, 11], our study is fundamentally different from them in both the approach to finding the heartbeat (in the fundamental frequency vs. in the higher-order harmonics) and the method of extracting the heartbeat (extracting the fundamental frequency vs. superimposing higher-order harmonics).

This content has been added to line 262 on page 11 of the revised manuscript.

Comment 2.3: *In this manuscript, the authors emphasized the interference and signal morphology distortion caused by breathing. However, it is evident that there is a clear frequency band gap between breathing (fundamental tone and second harmonic) and heart rhythm in most healthy subjects. Could a high-pass filter be used to remove the low-frequency breathing signals? Additionally, did you employ any other signal pre-processing methods before applying VMD?*

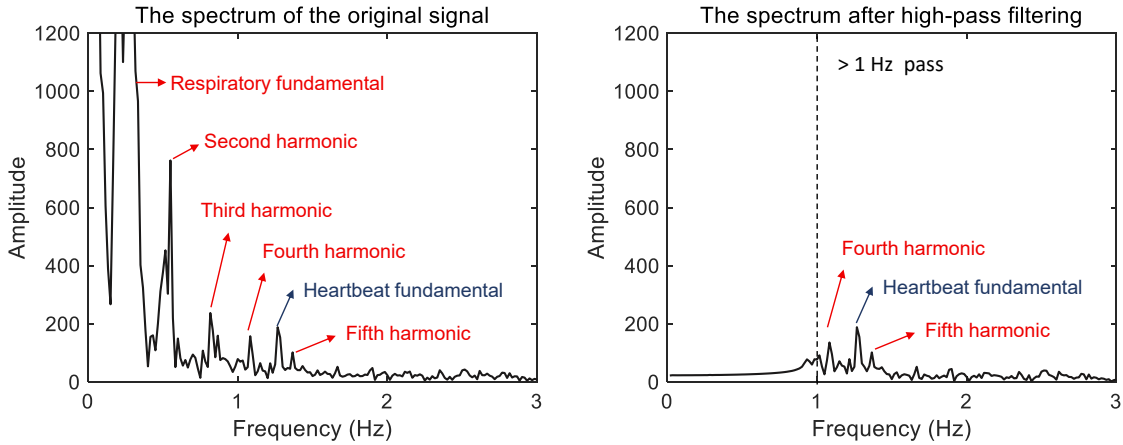


Fig. 2: Frequency spectrum of the mixed vital sign signal (left) and the spectrum after high-pass filtering (right).

Response 2.3: We thank the reviewer for the comment. While there is a certain gap between respiration (fundamental tone and second harmonic) and heartbeats, higher respiratory harmonics can be found in the vicinity of the heartbeat's fundamental frequency. This is because: (i) as breathing is a quasi-periodic signal, it encompasses not only the fundamental frequency and second-order harmonics but also higher-order harmonics; (ii) given that the amplitude of respiration significantly exceeds that of heartbeats, the higher-order respiratory harmonics, despite

their continuous attenuation, do not exhibit a substantial energy disparity within the fundamental frequency of heartbeats. Consequently, these higher-order breathing harmonics intermingle with the heartbeats, as illustrated in Figure 2(left). In this case, employing a high-pass filter—with a cutoff frequency set at 1 Hz—can effectively mitigate the interference from the fundamental frequency and the second harmonic of respiration on the heartbeat signal. However, it falls short in attenuating the impact of higher-order respiratory harmonics, as depicted in Fig. 2(right) (For more details, see in Supplementary Note 2 and Figure 1).

We did not employ any other signal pre-processing methods before applying VMD.

Comment 2.4: *In this paper, the authors presented sum-to-product formulas for trigonometric functions, which are commonly known. They attempted to explain that the mixing of breathing signals and heartbeat harmonics exhibits unique features in the spectrum but did not provide sufficient evidence. This analysis appears unrelated to their heart rate variation study, as peak detection, which is more associated with the envelope of the IMF, was used. Please provide further clarification or consider removing this section.*

Response 2.4: We appreciate the reviewer’s comment and apologize for any confusion caused by this section. We employed the sum-to-product formulas for trigonometric functions not to explain the mutual modulation phenomenon between breathing and heartbeat harmonics, but rather to elucidate the beat frequency effect present among the high-frequency heartbeat harmonics.

This effect describes how the higher-order harmonics of heartbeats exhibit a distinct beat pattern, characterized by periodic variations in amplitude due to the interference of two or more closely spaced frequencies. This understanding is crucial for accurately estimating heartbeats without needing to precisely determine the order of the higher-order harmonics.

Additionally, we have provided experimental evidence for the existence and significance of the beat frequency effect in the higher-order harmonics of heartbeats.

This content has been integrated into the updated Supplementary Information and is located on Page 8.

Comment 2.5: *In this paper, the heartbeat signal extraction highly relied on the VMD. How does each imf match with your raw signal frequency spectrum bands? Figure 5e has a strange layout and is unclear, and in supplementary figure 4, the authors only showed a selected band*

of spectrum which is a duplication of Figure 5d and 5e. Can you also show the whole spectrum, maybe also show the HHT spectrum of your imfs to further verify that? What is the HRV computation time of different signal epochs?

Response 2.5: We appreciate the reviewer’s comments and suggestions. As shown in Fig. 3, we have revised manuscript Figure 5e (Page 10) and Supplementary Figure 4 (Page 20) to clearly display each IMF’s frequency spectrum bands alongside the whole spectrum. Additionally, we have included the HHT spectrum of the five IMFs to further validate our findings, which aligns with the FFT spectrum. Moreover, we also recorded the computation time for Heart Rate Variability (HRV) with different signal durations, as shown in Table VII.

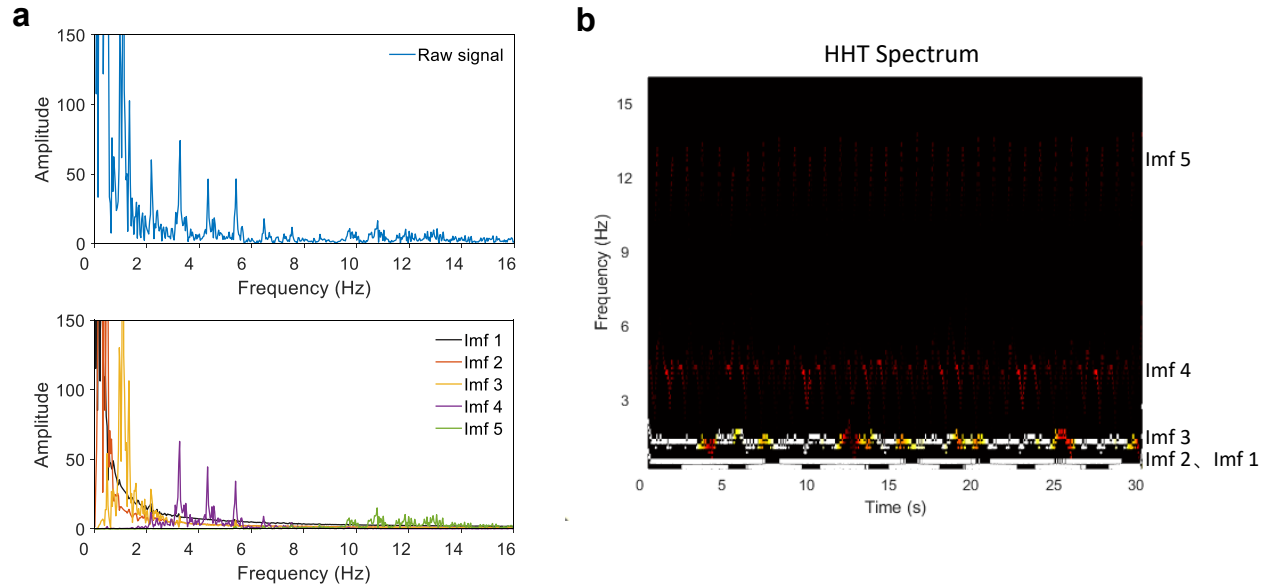


Fig. 3: (a) The raw signal frequency spectrum band and each IMF’s frequency spectrum bands. (b) The HHT spectrum of the five IMFs.

TABLE VII: The HRV computation time for different signal epochs.

Signal Duration (second)	30	60	90	120
Computation Time (second)	1.38	3.72	5.91	9.34

Comment 2.6: Several important technical details are missing. What are the sample digits and sampling rate of each radio channel? Are you transmitting both magnitude and phase signals or only the phase signal? What is your Wi-Fi data rate? Additionally, what is the time window

for heart rate calculation, and what is the window overlap? A schematic of your radar system should also be included.

TABLE VIII: Radar parameters

Parameters	Value
Start frequency	60 GHz
Rate of frequency change	65 MHz / μ s
Idle time	10 μ s
Chirp duration	60 μ s
Sampling frequency points	256
Sampling rate	5 MHz
Frame rate	100 Hz

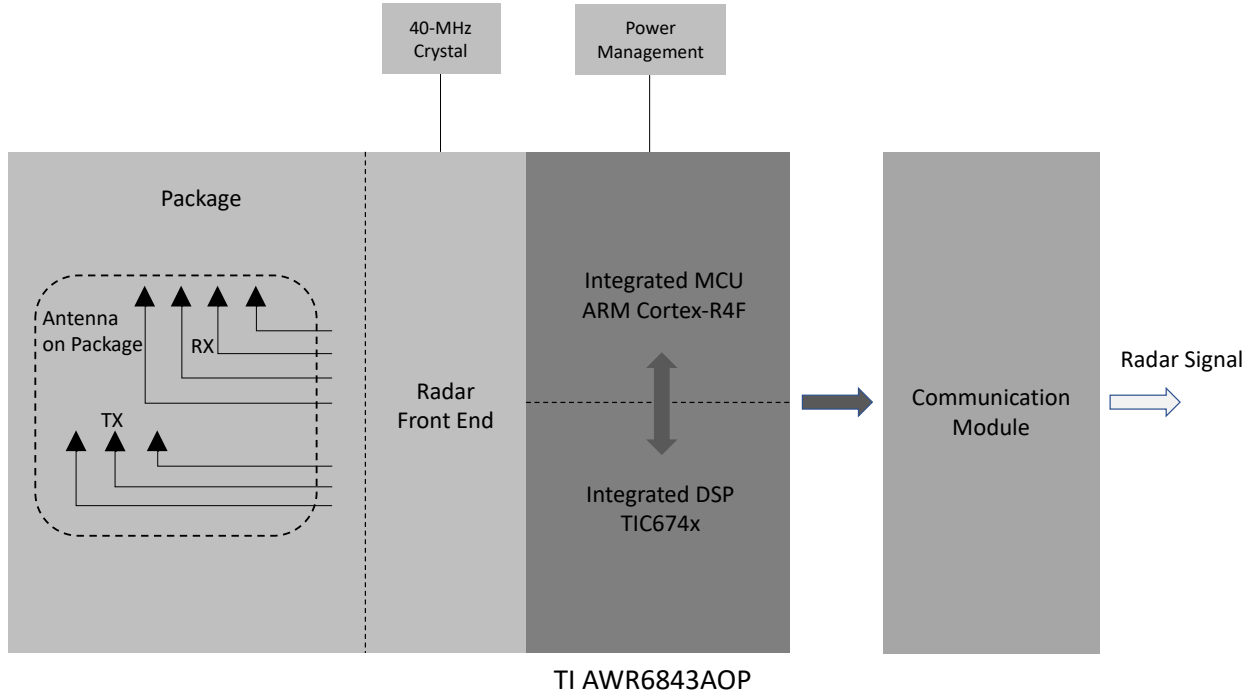


Fig. 4: The architecture of the radar system

Response 2.6: We appreciate the reviewer's insights and have included additional technical details in the revised manuscript. Specifically, the radio parameters are outlined in Table VIII. The transmitted signal includes both magnitude and phase. The Wi-Fi data rate is 54 Mbps,

and the time window for heart rate calculation is 30 seconds with no overlap. We have also incorporated a schematic of the radar system, as shown in Figure 4.

These contents have been included in Supplementary Table 9 and Figure 9.

Comment 2.7: *The authors proposed a radar system for long-term continuous HRV monitoring of subjects on a bed, emphasizing its non-contact nature. However, line-of-sight blockage and subject postures (supine or lateral) can affect the backscattered signal. The authors should provide more information on how their system performs under various postures and how these different postures impact heartbeat variation.*

Response 2.7: We thank the reviewer for the comment and suggestion. As suggested by the reviewer, we compared the system’s performance under supine (normal) and lateral positions, with the results shown in Table IX. When the subjects were lateral, the system’s performance showed a certain decline, mainly because the signal-to-noise ratio of the heartbeats obtained was lower in this position. Moreover, it is noteworthy that in our original manuscript, while all subjects maintain a normal posture in large-scale outpatient settings, the subjects’ postures are not restricted when performing overnight long-term monitoring.

We have added this result to Supplementary Information in Supplementary Note 8.

TABLE IX: Impact of the various postures on system’s performance.

Postures	RT-IBI Error (ms)	RMSSD Error (ms)	SDRR Error (ms)	pNN50 Error (%)
Supine	15.1	33.5	20.9	10.1
Lateral	23.5	86.2	44.8	21.6

Comment 2.8: *Furthermore, when the subject moved during the recording, how did you identify these movements, and did you still use these epochs to calculate heart rate? If moving epochs were excluded, what was the ratio of usable epochs (the windows used to calculate heart rate and HRV) in the long-term recording?*

Response 2.8: We thank the reviewer for the comment. During long-term monitoring, significant movements—such as turning over during sleep—cause drastic changes in the phase of the reflected signals, as illustrated in Fig. 5. We identify these movements by detecting these

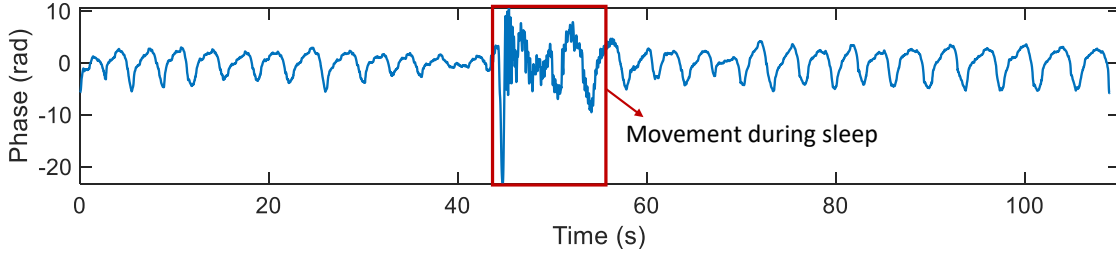


Fig. 5: Abrupt phase changes caused by significant movements.

abrupt phase changes and subsequently exclude those epochs from the analysis. As a result, the ratio of usable epochs in our long-term recordings is 80.8%.

The relevant information has been incorporated into the Supplementary Figure 10.

Comment 2.9: *When comparing real-time IBI and post-processing IBI, the authors discard certain time segments to reduce 'error'. How can you determine if this 'missing' pattern is due to RF signal loss, heart halt, or another abnormality? How do you appropriately choose the threshold for this identification?*

Response 2.9: We appreciate the reviewer's comment and apologize for any confusion caused by our previous explanations. To clarify, our work used real-time IBI instead of post-processing IBI for evaluation. More importantly, when calculating real-time IBI, we did not discard any segments, to provide a more objective and precise indication of the system's true performance (see how to calculate real-time IBI in Supplementary Figure 5). In contrast, the approach of discarding certain time segments was used in prior studies that employed post-processing IBI for evaluation. As the reviewer indicated, such evaluation may reduce errors and overestimate the system's capabilities. That is the reason we propose to evaluate the system with real-time IBI instead of post-processing IBI (see the comparison in Supplementary Figure 6).

Comment 2.10: *The authors included many algorithm comparison plots in the main figures. Are the algorithms you compared to widely used methods or just your previous works? To demonstrate your system's advantages or its potential clinical grade applications, comparing the IBI features and abnormality classification results with those from other clinical devices like ECG, PPG, or other radar methods such as Doppler radar or near-field RF sensors could make a stronger statement.*

Response 2.10: We thank the reviewer for the comment. The comparison algorithms, specifically mmHRV [10] and V2iFi [13], are state-of-the-art RF-based HRV sensing methods developed by research teams from the University of Maryland, College Park, and Nanyang Technological University.

To demonstrate our system’s potential clinical grade applications, we have simultaneously collected radar signals and ECG data. Then, we calculated the errors between the HRV estimated by our RF system and the HRV obtained from ECG data, using this as a basis for evaluating our system’s performance. We have revised the manuscript to make this evaluation approach more clear.

Moreover, as suggested by the reviewer, we have also employed PPG devices to sense the HRV of the subjects and compared the performance with our method. From Table X, we can see that our method has demonstrated superior performance on the RT-IBI metric, yet it falls short in comparison to PPG on RMSSD, SDRR, and pNN50. In summary, our method achieves, in a completely non-contact manner, comparable performance with PPG.

This section has been added to the revised manuscript, specifically on page 11, line 242.

Metrics	RT-IBI Error (ms)		RMSSD Error (ms)		SDRR Error (ms)		pNN50 Error (%)	
Methods	Ours	PPG	Ours	PPG	Ours	PPG	Ours	PPG
Mean	15.1	16.7	33.5	13.1	20.9	5.8	10.1	4.1
Median	8.0	13.0	6.7	10.0	5.0	4.1	1.5	4.2

TABLE X: Comparison of different devices.

Comment 2.11: *In the abstract, the authors emphasized “a long-term daily scenario spanning two months”, which could be misleading. To clarify, the study involves recording individuals for 5 nights over two months.*

Response 2.11: We appreciate the reviewer’s suggestion. In the original manuscript, the assessment of long-term monitoring indeed involved randomly collecting 5 nights of data over a two-month period to conduct randomized follow-up experiments. As suggested by the reviewer, we have carefully revised the manuscript to clarify the descriptions related to long-term evaluation.

Additionally, to further assess the system’s capability for long-term continuous monitoring, we

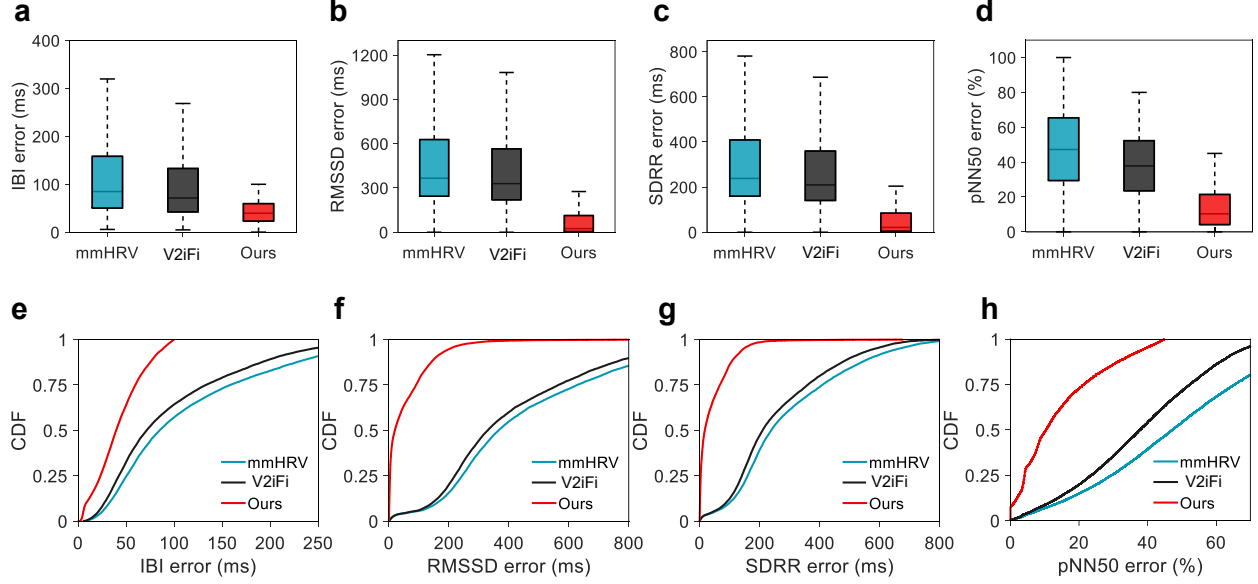


Fig. 6: Evaluation on 21 consecutive nights data. **a**, Comparison of real-time IBI error. RF-HRV achieves a 39.5 ms median RT-IBI error, while mmHRV and V2iFi achieve 84.9 ms and 71.3 ms median RT-IBI error, respectively. **b**, Comparison of RMSSD error. RF-HRV achieves a 24.6 ms median RMSSD error, while mmHRV and V2iFi achieve 366.1 ms and 329.5 ms median RMSSD error, respectively. **c**, Comparison of SDRR error. RF-HRV achieves a 21.3 ms median RT-IBI error, while mmHRV and V2iFi achieve 238.4 ms and 209.8 ms median SDRR error, respectively. **d**, Comparison of pNN50 error. RF-HRV achieves a 10.3% SDRR error, while mmHRV and V2iFi achieve 47.3% and 37.8% median SDRR error, respectively. **e,f,g,h**, The cumulative distribution function (CDF) plots of RT-IBI error, RMSSD error, SDRR error, and pNN50 error.

have also supplemented an experiment with 21 consecutive nights of monitoring. The comparative performance of our system against two baseline methods, as evaluated using the aforementioned four metrics, is presented in Figure 6a-d. The results demonstrate that our system’s performance is consistent across both the 21 consecutive nights of data and the data from 5 random nights over two months. Specifically, RF-HRV achieved a median IBI error of 39.5 ms, a median RMSSD error of 24.6ms, a median SDRR error of 21.3 ms, and a median pNN50 error of 10.3%. Compared to mmHRV and V2iFi, our system significantly outperformed them across all evaluation metrics and achieved over 10x improvement in estimating RMSSD and SDRR. To provide a clearer overview of how our system performs across the entire long-term dataset, we

depict the comparison of CDF of the four metrics in Figure 6e-h. It is evident that our system consistently shows significantly better HRV estimation performance in long-term and continuous monitoring compared to existing systems.

The content has been included in the Supplementary Note 10 and Figure 11.

Comment 2.12: *In Supplementary Figure 2, the authors did not specify the definition of SNR, including whether it is based on amplitude or power spectrum density, or the bandwidth used for calculating the noise floor.*

Response 2.12: We thank the reviewer for the comment. In Supplementary Figure 2, the signal-to-interference-plus-noise ratio (SINR) of heartbeat harmonics represents the ratio of the power spectral density of the heartbeat harmonics to the combined power spectral density of interference and noise. The bandwidth used for calculating the noise floor is 2 Hz.

We have added the explanation to the Supplementary Note 3.

Comment 2.13: *The authors provided examples of heartbeat patterns resulting from the ‘beat frequency effect’ without any accompanying labels or explanation in both the main manuscript and supplementary materials. This oversight reflects negatively on the authors’ attentiveness and responsibility to the reader.*

Response 2.13: We appreciate the reviewer’s feedback and apologize for any confusion caused. As suggested by the reviewer, we have included a detailed explanation of the examples of heartbeat patterns resulting from the beat frequency effect in the Supplementary Information (see Supplementary Note 4.2 and Figure 8). Additionally, we have included further clarification about the beat frequency effect in the revised manuscript, and made revisions throughout the article to enhance its readability.

Comment 2.14: *The authors didn’t provide data or code in the submitted package.*

Response 2.14: We thank the reviewer for the comment. We have uploaded the code and sample data to the officially recommended platform Code Ocean.

III. RESPONSES TO REVIEWER 3

We would like to express sincere thanks to the reviewer 3 for the thorough and constructive comment. In what follows, we present detailed comments in response to the individual points raised by the reviewer and elaborate on how the manuscript has been revised.

Comment 3.1: *In the introduction part, the authors stated that "we propose to leverage previously undiscovered higher frequency ranges where heartbeat information dominates over other motions." This seems to be the most important novelty of this paper. But even in ref [31] P. Sharma, Z. Zhang, T. B. Conroy, X. Hui, and E. C. Kan, "Attention detection by heartbeat and respiratory features from radio-frequency sensor," Sensors, vol. 22, no. 20, p. 8047, 2022. "To accurately extract the inter-beat interval (IBI) from the smooth NCS signal, a weaker, but sharper second harmonic heartbeat component of the heartbeat was used [32] and IBI was measured as the time for two cycles, as denoted in Figure 3b.", in this paper, heart rate was extracted from second harmonics to get more further away from respiration band. From this example, I wouldn't say higher frequency ranges are totally "undiscovered" and this method is used "for the first time".*

Response 3.1: We thank the reviewer for the comment and suggestion. While some studies reduce the interference of respiratory harmonics by utilizing the second or third harmonics of heartbeats, none of them have theoretically analyzed how respiratory harmonics might affect heartbeats. In contrast, we have found that due to the amplitude of respiration being much greater than that of heartbeats, even the fourth, fifth, or higher-order respiratory harmonics have energy similar to the fundamental frequency of heartbeat, as shown in Figure 2. Consequently, these higher-order respiratory harmonics can still intermingle with the fundamental frequency of heartbeats and the second harmonic of heartbeats.

To address this issue, we systematically analyze the impact of higher-order respiratory harmonics and find a fundamental phenomenon that prior work has never identified: respiration harmonics decay more rapidly than heartbeat harmonics. This is because the lower frequency of respiration compared to heartbeats results in higher-order respiratory harmonics that decay faster. Leveraging this discovery, we develop the core insight that differs from existing works, as we focus on extracting more clean heartbeat information leveraging frequency ranges much higher than existing work, typically beyond 10-order heartbeat harmonics.

Based on the above insight, we successfully extract heartbeat signals from frequencies far from the heart rate frequency band leveraging the beat frequency effect. Specifically, extracting heartbeat from higher-frequency bands presents unique challenges, including the rapid attenuation of signal strength and the unknown order of heartbeat harmonics. To resolve this, we find that when two harmonics are superimposed, they generate a distinctive signal pattern with a frequency equal to the difference between the frequencies of the two harmonics. Therefore, we generate strong heartbeat patterns that exhibit a frequency precisely equal to the heart rate by superimposing adjacent higher-order heartbeat harmonics.

In summary, while previous works have utilized 2-order/3-order heartbeat harmonics to estimate heart rate, our study is fundamentally different from them in both the approach to finding the heartbeat (in the 2-order/3-order heartbeat harmonics vs. beyond 10-order heartbeat harmonics) and the method of extracting the heartbeat (extracting a single low-order heartbeat harmonic vs. superimposing adjacent higher-order heartbeat harmonics).

The section has been added on page 11, line 262 of the revised manuscript.

Comment 3.2: *The authors compared with other systems like mmHRV and V2iFi, I'm curious to see if a wider range of algorithms can be considered? For example, some learning based algorithms like CNN, especially considering the large datasets you have? A classical machine learning or deep learning algorithm may have a good result over large population and boost up the accuracy.*

Response 3.2: We thank the reviewer for the comment and suggestion. As suggested by the reviewer, we have compared our algorithm with a classical learning-based method. Specifically, based on the characteristics of radar signals, we constructed a deep learning model (DL-HRV) using CNN, the structure of which is depicted in Figure 7. We employed the widely used UNet architecture to extract heartbeat patterns from the radar signals more effectively. We randomly selected 4,000 samples from the dataset to train the model and then conducted tests on the remaining 2,222 samples. The results, as shown in Table XI, indicate that the learning-based method has comparable yet slightly inferior performance compared to our method. This is because the generalization of learning-based methods heavily depends on the quantity, quality, and diversity of the training dataset. While we have collected a large-scale dataset for evaluation, it may still not be enough to train a robust neural network. Such results also validate another superiority of our method: it is interpretable and does not require any training data to achieve

generalization.

This experiment and results have been included in the Supplementary Note 11 and Figure 12.

Metrics	RT-IBI Error (ms)		RMSSD Error (ms)		SDRR Error (ms)		pNN50 Error (%)	
Methods	Ours	DL-HRV	Ours	DL-HRV	Ours	DL-HRV	Ours	DL-HRV
Mean	33.4	36.8	71.1	76.8	58.6	65.5	13.4	16.3
Median	28.5	31.2	55.2	56.9	31.3	35.7	11.8	14.9

TABLE XI: Comparison with deep learning model.

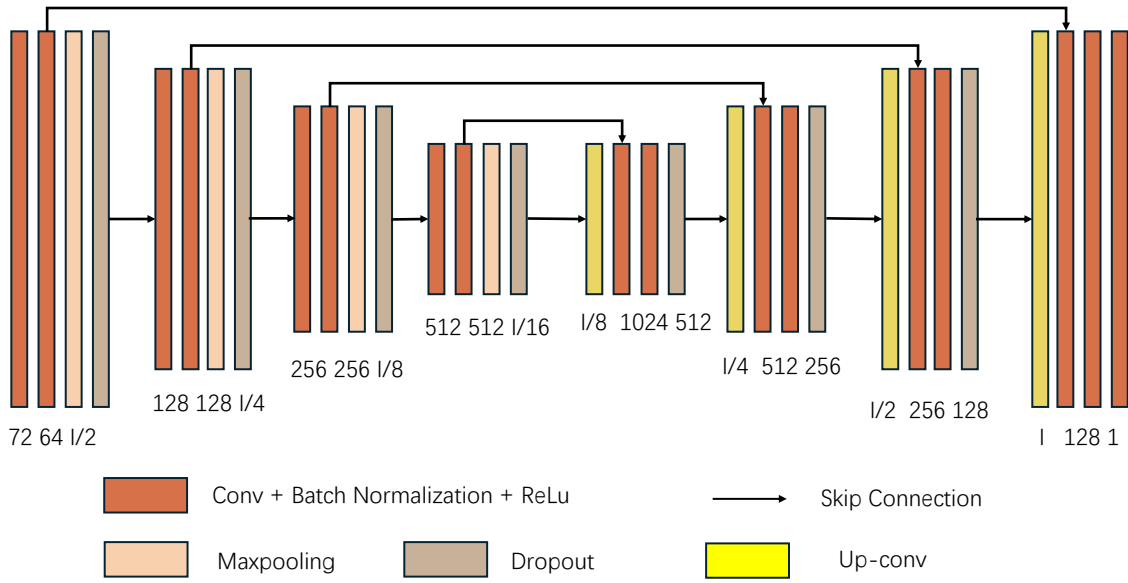


Fig. 7: The architecture of the deep learning model.

Comment 3.3: In Fig.1 d, both in clinical and daily life scenarios, the subjects are lying in bed. This has the least motion interference to the RF sensor, and easier to get higher signal quality. However, as the authors stated in the introduction, "our system shows immense potential in providing a comprehensive understanding of an individual's health conditions in their own homes", in real remote healthcare setting at home, the users will not always stay in bed. When users change to daily activities, like walking, talking, or exercising, will the accuracy degrades? I think that will be the important question to answer if the impact of this sensing system is more than algorithm improvement.

Response 3.3: We appreciate the reviewer's insightful comment regarding the challenges posed by body motion interference in RF vital sign monitoring.

The rationality of our experimental settings is that a variety of illnesses, including cardiovascular, cerebrovascular, and respiratory diseases like asthma [2, 9], are particularly prone to manifest during nighttime sleep. On average, adults dedicate roughly five to six hours or more of their day to slumber. Consequently, the investigation of heart activity monitoring during sleep is of significant importance for the early identification and prevention of underlying health conditions. However, accurately gauging cardiac activity, even with improved signal quality in the supine position at night, poses a significant challenge, especially when applied to extensive population datasets—a challenge adeptly tackled by the methodology presented in this paper.

Indeed, the extraction of heart activity signals while an individual is in motion is a significant and pioneering concern, with several research efforts striving to address it [3, 4]. Nonetheless, it continues to be an unresolved issue in the field, and will be one of the directions of our future research.

This section has been added to the revised manuscript, specifically on page 12, line 291.

Comment 3.4: *The authors mentioned that “Slices containing significant body motion, as indicated by pronounced RF signal variations, were excluded from further analysis.” Could you specify the percentage or total time of recording that was polluted by body motion? What percentage of time can you get accuracy read of HR?*

Response 3.4: We thank the reviewer for the comment. On average, in the entire night’s data, 19.2% of the data could not accurately estimate HRV due to body movement, while the remaining 80.8% was estimable for HRV.

This detail has been added in the revised manuscript, specifically on page 8, line 192.

Comment 3.5: *In Fig.3.c, why the accuracy of ECG was not directly reported here?*

Response 3.5: We thank the reviewer for the comment. The overall identification accuracy of the ECG device is 90.6%. We have included this information in Fig. 3 of the revised manuscript.

REFERENCES

- [1] Thilina Balasooriya, Prateek Nallabolu, and Changzhi Li. Application of variational mode decomposition to fmcw radar interference mitigation. In Nagender Kumar Suryadevara, Bobby George, Krishanthi P. Jayasundera, Joyanta Kumar Roy, and Subhas Chandra

- Mukhopadhyay, editors, *Sensing Technology*, pages 425–432, Cham, 2022. Springer International Publishing.
- [2] Peter J Barnes. Nocturnal asthma: mechanisms and treatment. *British Medical Journal (Clinical research ed.)*, 288(6428):1397, 1984.
 - [3] Zhe Chen, Tianyue Zheng, Chao Cai, and Jun Luo. Movi-fi: Motion-robust vital signs waveform recovery via deep interpreted rf sensing. In *Proceedings of the 27th annual international conference on mobile computing and networking*, pages 392–405, 2021.
 - [4] Jian Gong, Xinyu Zhang, Kaixin Lin, Ju Ren, Yaoxue Zhang, and Wenxun Qiu. Rf vital sign sensing under free body movement. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 5(3):1–22, 2021.
 - [5] Shekh M M Islam, Naoyuki Motoyama, Sergio Pacheco, and Victor M. Lubecke. Non-contact vital signs monitoring for multiple subjects using a millimeter wave fmcw automotive radar. In *2020 IEEE/MTT-S International Microwave Symposium (IMS)*, pages 783–786, 2020.
 - [6] Zongquan Ling, Weinan Zhou, Yuhao Ren, Jiping Wang, and Liquan Guo. Non-contact heart rate monitoring based on millimeter wave radar. *IEEE Access*, 10:74033–74044, 2022.
 - [7] Jingtao Liu, Yuchen Li, Changzhi Li, Changzhan Gu, and Jun-Fa Mao. Accurate measurement of human vital signs with linear fmcw radars under proximity stationary clutters. *IEEE Transactions on Biomedical Circuits and Systems*, 15(6):1393–1404, 2021.
 - [8] Yu Rong and Daniel W. Bliss. Remote sensing for vital information based on spectral-domain harmonic signatures. *IEEE Transactions on Aerospace and Electronic Systems*, 55(6):3454–3465, 2019.
 - [9] Francoise Roux, Carolyn D’Ambrosio, and Vahid Mohsenin. Sleep-related breathing disorders and cardiovascular disease. *The American journal of medicine*, 108(5):396–402, 2000.
 - [10] Fengyu Wang, Xiaolu Zeng, Chenshu Wu, Beibei Wang, and K.J. Ray Liu. mmhrv: Contactless heart rate variability monitoring using millimeter-wave radio. *IEEE Internet of Things Journal*, 8(22):16623–16636, 2021.
 - [11] Kuan Ju Wu and Chin-Lung Yang. Heart rate extraction with vmd algorithm in non-stationary clutter environment based on fmcw radar systems. In *2021 IEEE International Symposium on Radio-Frequency Integration Technology (RFIT)*, pages 1–3, 2021.
 - [12] Hongqiang Xu, Malikeh P. Ebrahim, Kareeb Hasan, Fatemeh Heydari, Paul Howley, and

- Mehmet Rasit Yuce. Accurate heart rate and respiration rate detection based on a higher-order harmonics peak selection method using radar non-contact sensors. *Sensors*, 22(1), 2022.
- [13] Tianyue Zheng, Zhe Chen, Chao Cai, Jun Luo, and Xu Zhang. V2ifi: in-vehicle vital sign monitoring via compact rf sensing. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 4(2), jun 2020.

REVIEWERS' COMMENTS

Reviewer #1 (Remarks to the Author):

The manuscript has improved substantially to demonstrate the uniqueness and improved performance of the algorithm with additional data and results. I appreciate the effort put into it and do not have any further requests or comments.

Reviewer #2 (Remarks to the Author):

In this paper, the authors present a novel RF-based system for long-term heart rate variability (HRV) monitoring with clinical-grade accuracy. It successfully addresses challenges like respiratory motion interference without requiring model training. The system's effectiveness is validated in large-scale outpatient and long-term daily life scenarios, demonstrating comparable performance to ECG-based methods. The approach offers a promising solution for non-invasive, precise, and long-term cardiac monitoring.

The authors have effectively addressed my comments, and the paper is now well-prepared for publication.

Reviewer #2 (Remarks on code availability):

The code provided can be successfully compiled without issues.

Reviewer #3 (Remarks to the Author):

Thank authors for answering my questions. I have follow-up questions here:

Could you indicate the frequency range you used in the RF sensor in the abstract clearly? I think this is important information for RF sensing. Apart from signal processing side, do you have other contributions in the hardware level? And are there thoughts on how to improve SNR and signal quality in hardware design level? For example, frequency division, adding transceiver channels, or other ideas?

Have you considered sensor fusion to improve overall accuracy? Like IMU? To improve motion artifact detection?

Response to Review of Paper:

NCOMMS-24-30589A-Z

Monitoring Long-term Cardiac Activity with Contactless Radio Frequency Signals

I. RESPONSES TO REVIEWER 3

We would like to express sincere thanks to the reviewer 3 for the thorough and constructive comment. In what follows, we present detailed comments in response to the individual points raised by the reviewer and elaborate on how the manuscript has been revised.

Comment 1: *Could you indicate the frequency range you used in the RF sensor in the abstract clearly? I think this is important information for RF sensing.*

Response 1: We thank the reviewer for the comment and suggestion. The frequency range of our RF sensor is 60 to 64 GHz. We have included this information in the revised abstract (page 1, line 17).

Comment 2: *Apart from signal processing side, do you have other contributions in the hardware level? And are there thoughts on how to improve SNR and signal quality in hardware design level? For example, frequency division, adding transceiver channels, or other ideas?*

Response 2: We appreciate the reviewer's comment. Our hardware design focuses on achieving a cost-effective (\$45.5) and portable (50g) solution. The device includes three main components: a commercial radar chip (TI AWR6843aop) for sensing, an ESPRESSIF ESP32 module for data transmission, and a SAMSUNG 256 Pro storage module for data retention.

While our primary focus has been on signal processing, we acknowledge the potential for hardware-level improvements to enhance SNR and signal quality. In our current design, we use three transmitters and four receivers to form multiple channels, which improves SNR by capturing more diverse signal paths. While adding additional antennas could further enhance SNR, it would increase hardware complexity and cost, presenting a trade-off.

In future iterations, we plan to explore additional techniques, such as frequency division pointed out by the reviewer, as well as advanced antenna designs tailored to focus on specific regions, which could further optimize performance for physiological sensing applications.

We have included this content in the Discussion section of the manuscript.

Comment 3: *Have you considered sensor fusion to improve overall accuracy? Like IMU? To improve motion artifact detection?*

Response 3: We thank the reviewer for this insightful comment. We have indeed explored the potential of integrating sensors like IMUs and depth cameras to enhance motion artifact detection. However, these approaches did not align with our design goals. Specifically, IMUs require direct attachment to the body, while depth cameras need clear line-of-sight conditions, which contradicts our aim of developing a non-contact, privacy-preserving cardiac monitoring system. Nevertheless, we agree with the reviewer that sensor fusion is a promising approach, and in future work, we plan to explore the integration of other wireless signals, such as ultrasound, to improve accuracy under motion artifacts.

We have included this content in the Discussion section of the manuscript.